



**UNIVERSIDAD PERUANA DE CIENCIAS APLICADAS**

**FACULTAD DE NEGOCIOS**

**PROGRAMA ACADÉMICO DE ADMINISTRACIÓN Y NEGOCIOS  
INTERNACIONALES**

**Índices globales de búsquedas *online* acerca del *Zingiber Officinale* como predictores  
de las exportaciones Peruanas de jengibre fresco (Partida arancelaria 091011) al  
resto del mundo 2013-2021: Un enfoque basado en redes neuronales**

**TESIS**

Para optar el título profesional de Licenciado en Negocios Internacionales

**AUTOR(ES)**

Perez Garcia, Carlos Andres (0000-0003-3359-7029)

Pimentel Vignes, Raphael (0000-0003-0357-3405)

**ASESOR**

Blas Rivera, Aldrudover (0000-0002-4661-9111)

**Lima, 26 de abril del 2023**

## DEDICATORIA

*A nuestros padres*

## RESUMEN

La presente investigación tuvo por objetivo determinar si los índices globales de búsquedas *online* acerca del jengibre (*Zingiber officinale*) eran predictores de las exportaciones mensuales peruanas de dicho producto en su versión no triturada ni pulverizada al resto del mundo durante el período 2013-2021, todo ello bajo en enfoque de las redes neuronales.

Los índices que representa búsquedas *online* acerca de un término específico son considerados una forma de *Big Data*. Es posible obtener los mismos con *Google Trends*. Los modelos que asocian a las búsquedas *online* con una eventual compra pueden ser asociados a la Teoría del comportamiento planificado. Sin embargo, los antecedentes revisados evidenciaron que ello era significativo para los bienes duraderos o de alto valor, mientras que para los alimentos los resultados no eran muy auspiciosos. No obstante, luego de analizar 108 series mensuales, los autores de esta investigación evidenciaron que las búsquedas *online* acerca del jengibre sí eran predictores de las exportaciones peruanas del mencionado producto.

Lo anterior se logró adicionando las distintas dimensiones de las búsquedas *online* a un modelo base que solo incluía rezagos de correspondientes a un año. Este último alcanzó una precisión de 66.15%, mientras que la versión que adicionó a las búsquedas en YouTube y de imágenes acerca del jengibre registró una precisión de 80.84%. Las búsquedas web ordinarias no contribuyeron a la precisión del modelo final. Adicionalmente, al emplear de forma complementaria la regresión lineal, se evidenciaron relaciones directas y significativas entre las búsquedas web y en YouTube, no ocurrió lo mismo con los índices asociados a la búsqueda de imágenes.

Términos clave: *Modelos predictivos, Big Data, Redes Neuronales, Jengibre.*

## ABSTRACT

This research aimed to determine whether global online search indices for ginger (*Zingiber officinale*) were predictors of Peruvian monthly exports of unground and unpulverised ginger to the rest of the world during the period 2013-2021, using a neural network approach.

Indexes representing online searches for a specific term are considered a form of Big Data. It is possible to obtain them with Google Trends. Models that associate online searches with an eventual purchase can be associated with the Theory of Planned Behaviour. However, the background research showed that this was significant for durable or high-value goods, while for food the results were not very auspicious. Nevertheless, after analysing 108 monthly series, the authors of this research found that online searches for ginger were indeed predictors of Peruvian ginger exports.

This was achieved by adding the different dimensions of online searches to a base model that included only one year's lags. The latter achieved an accuracy of 66.15%, while the version that added YouTube and image searches about ginger recorded an accuracy of 80.84%. Ordinary web searches did not contribute to the accuracy of the final model. In addition, when linear regression was used in a complementary way, direct and significant relationships between web and YouTube searches were evident, but the same was not true for the indices associated with image searches.

**Keywords:** *Predictive models, Big Data, Neural Networks, Ginger*

N°4604\_Índices globales de búsquedas online acerca del Zingiber Officinale como predictores de las exportaciones Peruanas de jengibre fresco (Partida arancelaria 091011) al resto del mundo 2013-2021:

INFORME DE ORIGINALIDAD

11%

INDICE DE SIMILITUD

11%

FUENTES DE INTERNET

4%

PUBLICACIONES

5%

TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	<a href="https://repositorio.usil.edu.pe">repositorio.usil.edu.pe</a> Fuente de Internet	2%
2	<a href="https://repositorioacademico.upc.edu.pe">repositorioacademico.upc.edu.pe</a> Fuente de Internet	1%
3	<a href="https://hdl.handle.net">hdl.handle.net</a> Fuente de Internet	1%
4	<a href="http://www.freshplaza.es">www.freshplaza.es</a> Fuente de Internet	1%
5	<a href="https://indicadores.bajacalifornia.gob.mx">indicadores.bajacalifornia.gob.mx</a> Fuente de Internet	1%
6	<a href="https://renati.sunedu.gob.pe">renati.sunedu.gob.pe</a> Fuente de Internet	1%
7	<a href="https://eduarea.wordpress.com">eduarea.wordpress.com</a> Fuente de Internet	1%
8	Submitted to Universidad San Ignacio de Loyola	<1%

---

Trabajo del estudiante

---

9	<a href="http://revistas.unheval.edu.pe">revistas.unheval.edu.pe</a> Fuente de Internet	<1 %
10	<a href="http://repositorio.usmp.edu.pe">repositorio.usmp.edu.pe</a> Fuente de Internet	<1 %
11	Submitted to imfice Trabajo del estudiante	<1 %
12	<a href="http://ideas.repec.org">ideas.repec.org</a> Fuente de Internet	<1 %
13	Jurado Piña Alan Rubén. "Desarrollo de un dispositivo tipo wearable orientado al entrenamiento HIIT utilizando aprendizaje supervisado", TESIUNAM, 2019 Publicación	<1 %
14	<a href="http://baixardoc.com">baixardoc.com</a> Fuente de Internet	<1 %
15	<a href="http://repositorio.puce.edu.ec">repositorio.puce.edu.ec</a> Fuente de Internet	<1 %
16	<a href="http://cybertesis.unmsm.edu.pe">cybertesis.unmsm.edu.pe</a> Fuente de Internet	<1 %
17	<a href="http://upc.aws.openrepository.com">upc.aws.openrepository.com</a> Fuente de Internet	<1 %
18	<a href="http://www.coursehero.com">www.coursehero.com</a> Fuente de Internet	<1 %

---

Submitted to Editorial Elearning S.L.

19	Trabajo del estudiante	<1 %
20	<a href="http://alicia.concytec.gob.pe">alicia.concytec.gob.pe</a> Fuente de Internet	<1 %
21	<a href="http://repositorio.une.edu.pe">repositorio.une.edu.pe</a> Fuente de Internet	<1 %
22	<a href="http://saber.ucv.ve">saber.ucv.ve</a> Fuente de Internet	<1 %
23	<a href="http://gestion.pe">gestion.pe</a> Fuente de Internet	<1 %
24	<a href="http://repositorio.ucv.edu.pe">repositorio.ucv.edu.pe</a> Fuente de Internet	<1 %
25	Submitted to Universidad Carlos III de Madrid Trabajo del estudiante	<1 %
26	Bano, Maria M.. "Intencion de Endeudamiento Personal con Tarjeta de credito en el Ecuador.", Pontificia Universidad Catolica del Peru - CENTRUM Catolica (Peru), 2020 Publicación	<1 %
27	<a href="http://aperta.ulakbim.gov.tr">aperta.ulakbim.gov.tr</a> Fuente de Internet	<1 %
28	<a href="http://es.usaswimofficials.com">es.usaswimofficials.com</a> Fuente de Internet	<1 %
29	<a href="http://worldwidescience.org">worldwidescience.org</a> Fuente de Internet	<1 %

30	Abhishek Singh, G. C. Mishra. "APPLICATION OF BOX-JENKINS METHOD AND ARTIFICIAL NEURAL NETWORK PROCEDURE FOR TIME SERIES FORECASTING OF PRICES", Statistics in Transition. New Series, 2015 Publicación	<1 %
31	dehesa.unex.es Fuente de Internet	<1 %
32	www.esdelatino.com Fuente de Internet	<1 %
33	www.tandfonline.com Fuente de Internet	<1 %
34	www.academia.edu Fuente de Internet	<1 %
35	Lehbib Mohamed-Salem Mohamed. "Integracion economica : el caso europeo y latinoamericano", TESIUNAM, 2005 Publicación	<1 %
36	mail.agraria.pe Fuente de Internet	<1 %
37	pirhua.udep.edu.pe Fuente de Internet	<1 %
38	www.slideshare.net Fuente de Internet	<1 %



39	Reyes Tello José Carlos. "El comercio electrónico en las empresas minoristas de México desde la teoría de recursos y capacidades", TESIUNAM, 2019 Publicación	<1 %
40	Submitted to University of Sheffield Trabajo del estudiante	<1 %
41	dokumen.pub Fuente de Internet	<1 %
42	patagoniataste.com.ar Fuente de Internet	<1 %
43	Submitted to Universidad Estatal de Milagro Trabajo del estudiante	<1 %
44	Submitted to Universidad Nacional del Centro del Peru Trabajo del estudiante	<1 %
45	issuu.com Fuente de Internet	<1 %
46	Submitted to Universidad Manuela Beltrán Virtual Trabajo del estudiante	<1 %
47	Vilca, Jesica Jannet Aleman. "Perfil de Lenguaje de Usuarios de Protesis Auditiva de 4 A 8 Anos del Colegio Fernando Wiese Eslava", Pontificia Universidad Catolica del Peru - CENTRUM Catolica (Peru), 2021	<1 %

Publicación

48	<a href="https://pubmed.ncbi.nlm.nih.gov">pubmed.ncbi.nlm.nih.gov</a> Fuente de Internet	<1 %
49	<a href="http://www.labome.org">www.labome.org</a> Fuente de Internet	<1 %
50	<a href="http://zagan.unizar.es">zagan.unizar.es</a> Fuente de Internet	<1 %
51	alvarez, Gloria   Jara, Giselle   Munarriz, Gianina   Luis Lara, Jose. "Planeamiento estrategico De La Industria De La chia En El Peru.", Pontificia Universidad Catolica del Peru - CENTRUM Catolica (Peru), 2020 Publicación	<1 %
52	<a href="http://dergipark.org.tr">dergipark.org.tr</a> Fuente de Internet	<1 %
53	<a href="http://www.timetoast.com">www.timetoast.com</a> Fuente de Internet	<1 %
54	Díaz Gonzalez Flores Pablo Claudio. "Análisis de comercio exterior de productos químicos en el contexto de economía abierta", TESIUNAM, 1996 Publicación	<1 %
55	Submitted to ISM International Academy Trabajo del estudiante	<1 %
56	<a href="http://cdn.bancentral.gov.do">cdn.bancentral.gov.do</a> Fuente de Internet	<1 %

---

Excluir citas	Apagado	Exclude assignment template	Activo
Excluir bibliografía	Activo	Excluir coincidencias	< 20 words

## Tabla De Contenidos

<b>INTRODUCCIÓN .....</b>	<b>1</b>
<b>CAPÍTULO I. MARCO TEÓRICO.....</b>	<b>4</b>
1.1. ANTECEDENTES.....	4
1.1.1. Antecedentes nacionales .....	4
1.1.2. Antecedentes internacionales .....	12
1.2. BASES TEÓRICAS.....	27
1.2.1. Teorías clásicas del Comercio Internacional vinculadas a la agricultura ....	27
1.2.2. Teorías modernas del Comercio Internacional vinculadas a la agricultura .	32
1.2.3. Aspectos teóricos vinculados a la predicción.....	38
1.2.4. Modelos predictivos.....	40
1.2.5. Evolución de la Industria 1.0 a la 4.0.....	42
1.2.6. Aspectos teóricos de la <i>Big Data</i> .....	46
1.2.7. Aspectos teóricos del <i>Machine Learning</i> .....	49
1.2.8. Las Redes Neuronales Artificiales y el Modelo Perceptrón Multicapa .....	51
1.2.9. Teoría de la Comportamiento Planificado .....	57
1.2.10. El Modelo AIDMA y el AISAS .....	61
1.2.11. El modelo basado en rezagos y la incorporación de la Big Data de Google Trends 64	
1.3. MARCO CONCEPTUAL.....	65
1.3.1. Beneficios de los modelos predictivos.....	66
1.3.2. Jengibre .....	67
1.3.3. El mercado mundial de las exportaciones de jengibre.....	69
1.3.4. Variable dependiente: Exportaciones peruanas de jengibre no triturado ni pulverizado .....	73
1.3.5. Variable independiente: Índices de búsquedas <i>online</i> obtenidos de <i>Google Trends</i> 78	
1.3.4. Ventajas y desventajas de trabajar con indicadores de <i>Google Trends</i> .....	79
1.3.5. Dimensiones de <i>Google Trends</i> relacionadas al consumo <i>online</i> .....	81
1.4. Viabilidad de la investigación .....	83
<b>CAPÍTULO II. PLAN DE INVESTIGACIÓN.....</b>	<b>85</b>

2.1. SITUACIÓN PROBLEMÁTICA .....	85
2.1.1. Problema General .....	88
2.1.2. Problemas específicos.....	88
2.2. OBJETIVOS .....	89
2.2.1. Objetivo general .....	89
2.2.2. Objetivos específicos .....	89
2.3. HIPÓTESIS .....	90
2.3.1. Hipótesis general .....	90
2.3.2. Hipótesis específicas.....	90
2.4. JUSTIFICACIÓN .....	91
2.4.1. Justificación teórica .....	91
2.4.2. Justificación metodológica.....	91
2.4.3. Justificación práctica .....	92
<b>CAPÍTULO III. METODOLOGÍA DE TRABAJO.....</b>	<b>94</b>
3.1. ENFOQUE DE LA INVESTIGACIÓN .....	94
3.2. OPERACIONALIZACIÓN DE VARIABLES.....	95
3.3. UNIVERSO O POBLACIÓN .....	97
3.4. MUESTRA.....	97
3.5. INSTRUMENTOS Y/O HERRAMIENTAS.....	98
3.6. VALIDEZ DE LOS DATOS OBTENIDOS DE GOOGLE TRENDS .....	99
3.7. RECOLECCIÓN DE DATOS .....	100
3.8. PLAN DE ANÁLISIS .....	102
3.9. LIMITACIONES DE LA METODOLOGÍA EMPLEADA.....	106
<b>CAPÍTULO IV. DESARROLLO Y APLICACIÓN .....</b>	<b>108</b>
4.1. RESULTADOS ASOCIADOS A LA HIPÓTESIS GENERAL.....	108
4.2. COMPROBACIÓN DE LA HIPÓTESIS GENERAL .....	117
4.3. DISCUSIÓN DE LA HIPÓTESIS GENERAL.....	118
4.4. RESULTADOS ASOCIADOS A LA HIPÓTESIS ESPECÍFICA 1 .....	119
4.5. COMPROBACIÓN DE LA HIPÓTESIS ESPECÍFICA 1.....	121
4.6. DISCUSIÓN DE LA HIPÓTESIS ESPECÍFICA 1 .....	121
4.7. RESULTADOS ASOCIADOS A LA HIPÓTESIS ESPECÍFICA 2 .....	123
4.8. COMPROBACIÓN DE LA HIPÓTESIS ESPECÍFICA 2.....	124

4.9. DISCUSIÓN DE LA HIPÓTESIS ESPECÍFICA 2 .....	125
4.10. RESULTADOS ASOCIADOS A LA HIPÓTESIS ESPECÍFICA 3 .....	126
4.11. COMPROBACIÓN DE LA HIPÓTESIS ESPECÍFICA 3 .....	127
4.12. DISCUSIÓN DE LA HIPÓTESIS ESPECÍFICA 3 .....	127
<b>CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES.....</b>	<b>129</b>
5.1. CONCLUSIONES .....	129
5.2. RECOMENDACIONES .....	130
<b>REFERENCIAS.....</b>	<b>133</b>

## ÍNDICE DE TABLAS

<b>Tabla 1</b> Matriz de antecedentes nacionales – Parte 1.....	10
Tabla 2 Matriz de antecedentes nacionales – Parte 2.....	11
<b>Tabla 3</b> Matriz de antecedentes internacionales- Parte 1.....	24
<b>Tabla 4</b> Matriz de antecedentes internacionales- Parte 2 .....	25
Tabla 5 Matriz de antecedentes internacionales- Parte 3 .....	26
Tabla 6: Principales exportadores mundiales de jengibre fresco 2012-2021.....	71
<b>Tabla 7</b> Vías de transporte empleados para las exportaciones de jengibre fresco de acuerdo al FOB 2014-2021 (en porcentaje).....	75
<b>Tabla 8</b> Principales destinos de las exportaciones peruanas de jengibre fresco 2012-2021 (en toneladas)....	76
<b>Tabla 9</b> Participación de los principales exportadores de jengibre fresco peruano 2014-2021.....	77
<b>Tabla 10</b> Operacionalización de variables.....	95
<b>Tabla 11</b> Lista de modelos que adicionan indicadores Google al modelo basado en rezagos .....	103
<b>Tabla 12</b> Comparativo de modelos empleados para la hipótesis general .....	108
<b>Tabla 13</b> Importancia de las variables predictoras en los modelos asociados a la hipótesis general.....	110
<b>Tabla 14</b> Parámetros de la red neuronal de mejor ajuste .....	112
<b>Tabla 15</b> Coeficientes de los modelos de regresión base y de mejor ajuste .....	116
<b>Tabla 16</b> Coeficientes de los modelos de regresión base y de mejor ajuste .....	117
<b>Tabla 17</b> Variables excluidas del modelo de regresión .....	117
<b>Tabla 18</b> Comparativo entre modelo base y modelo asociado a la hipótesis específica 1 .....	120
<b>Tabla 19</b> Importancia de las variables predictoras en los modelos asociados a la hipótesis específica 1 .....	120
<b>Tabla 20</b> Coeficientes del Modelo de regresión en base a rezagos y búsquedas web .....	120
<b>Tabla 21</b> Resumen del Modelo de regresión en base a rezagos y búsquedas web .....	121
<b>Tabla 22</b> Comparativo entre modelo base y el modelo asociado a la hipótesis específica 2 .....	123
<b>Tabla 23</b> Importancia de las variables predictoras en los modelos asociados a la hipótesis específica 2 ....	123
<b>Tabla 24</b> Coeficientes del modelo de regresión en base a rezagos y búsquedas en YouTube .....	124
<b>Tabla 25</b> Resumen del modelo de regresión en base a rezagos y búsquedas en YouTube .....	124
<b>Tabla 26</b> Comparativo entre modelo base y modelo asociado a la hipótesis específica 3 .....	126
<b>Tabla 27</b> Importancia de las variables predictoras en los modelos asociados a la hipótesis específica 3 .....	126
<b>Tabla 28</b> Coeficientes del modelo de regresión en base a rezagos y búsquedas en YouTube .....	127

## ÍNDICE DE FIGURAS

<b>Figura 1</b> Esquema de un perceptrón .....	52
<b>Figura 2</b> Arquitectura de una red neuronal.....	53
<b>Figura 3</b> Esquema de la Teoría de la Comportamiento Planificado .....	60
<b>Figura 4</b> Modelo AISAS .....	63
<b>Figura 5</b> Zingiber officinale .....	68
<b>Figura 6</b> Exportaciones de jengibre fresco a nivel global 2012-2021 (en millones de dólares).....	70
<b>Figura 7</b> Exportaciones peruanas de Jengibre del Perú al Resto del mundo 2012-2021 (en toneladas).....	74
<b>Figura 8</b> Variación porcentual de las exportaciones peruanas de Jengibre fresco del Perú al Resto del mundo 2012-2020 (en toneladas).....	75
<b>Figura 9</b> Market Share de las exportaciones peruana de jengibre con respecto a las globales 2012-2021 (en porcentaje) .....	76
<b>Figura 10</b> Flujograma de Operacionalización de variable de búsquedas online .....	96
<b>Figura 11</b> Interfase de uso de Google Trends.....	101
<b>Figura 12</b> Flujograma del SPSS Modeler.....	105
<b>Figura 13</b> Valores observados vs predichos para el modelo basado en rezagos .....	109
<b>Figura 14</b> Arquitectura de la red neuronal de mejor ajuste .....	110
<b>Figura 15</b> Función tangente hiperbólica pata una variable.....	111
<b>Figura 16</b> Valores observados vs predichos para el modelo de mejor ajuste que incluye datos de Google Trends .....	113
<b>Figura 17</b> Diagrama de dispersión para valores observados y predichos en los modelos base (izquierda) y el de mejor ajuste (derecha).....	114
<b>Figura 18</b> Valores predichos vs. observados para las exportaciones de jengibre (2013-2021).....	115

## ÍNDICE DE ANEXOS

Anexo 1 Datos de investigación .....	154
Anexo 2 Matriz de consistencia .....	157



## Lista de Abreviaturas

FMI: Fondo Monetario Internacional

OCDE: Organización para la Cooperación y el Desarrollo Económicos

OMC: Organización Mundial de Comercio

RNA: Red Neuronal Artificial

## Introducción

En las operaciones comerciales, una de las variables de mayor relevación es la estimación de la demanda. Una correcta estimación de la demanda contribuye a un mejor planeamiento financiero y de operaciones. Las operaciones de comercio exterior, desde luego, no son la excepción.

El estimar adecuadamente una demanda resulta crucial en productos que son perecibles. Por ello quienes tengan a su cargo el *forecasting* tienen como misión mejorar los pronósticos. Diversos investigadores sugieren incorporar nuevas variables. En recientes documentos de investigación se toma en cuenta indicadores *Big Data*.

Las búsquedas *online* son consideradas una forma de *Big Data*: se forman a una velocidad bastante alto, representan un considerable volumen de datos y se pueden dar en múltiples formas (desde una simple búsqueda en YouTube hasta una búsqueda realizada por voz en el buscador Google). Las búsquedas online representan el interés del consumidor, interés que puede terminar en una compra efectiva.

Bajo condiciones *ceteris paribus* la demanda del bien depende del precio. No obstante, en el mundo real la demanda tiene relación con el interés del individuo. Así, por ejemplo, antes de comprar un *Iphone* de cierto modelo, un potencial comprador puede buscar información en *Google* sobre dicho *smartphone* (y de otros productos similares, por supuesto), similar situación puede darse con autos, viajes turísticos y otros.

*Google Trends* es una aplicación diseñada por los ingenieros de *Google*. Consiste en un *dashboard* que permite obtener índices que representa a las búsquedas de un término cualquiera a lo largo de un intervalo de tiempo determinado y en cierto espacio geográfico

específico. Los datos del *dashboard* pueden ser accedidos desde cualquier computadora con acceso a internet.

El jengibre sin triturar y pulverizar (o fresco) es uno de los productos de agroexportación que mayor dinamismo ha presentado en los últimos años. Las causas todavía se discuten. Sin embargo, actualmente el Perú se ha convertido en uno de los principales exportadores de dicho insumo a nivel global. Cabe mencionar que las exportaciones han entrado en una nueva fase de expansión como resultado de la propagación del COVID-19.

Si es que las búsquedas online, medidas con *Google Trends*, permiten medir el interés de los potenciales compradores, cabe preguntarse entonces si las búsquedas de jengibre alrededor del mundo pueden contribuir a la predicción de las exportaciones de dicho bien. Ello puede ser demostrado empleando distintas técnicas estadístico-matemáticas. Quienes han empleado a los indicadores de *Google Trends* como elemento predictor han tomado modelos de referencia preexistentes y han logrado mejorar su precisión incorporando la data de *Google*.

En el primer capítulo se presentan los antecedentes nacionales e internacionales que han sido considerados como pertinentes. Asimismo, se definen y se sustentan los términos trabajados durante la investigación. Además, en un marco contextual, se detalla la situación actual del jengibre sin trituras ni pulverizar.

Luego, en el segundo capítulo se desarrolla el plan de investigación, en el que se plantean los problemas, los objetivos y las correspondientes hipótesis de trabajo. Se establecieron un conjunto de hipótesis a ser probadas. Asimismo, previo al análisis, en el tercer capítulo se desarrolla la metodología de la investigación y se definen las variables analizadas, así como la construcción del *dataset*. En el cuarto capítulo se muestran los resultados de las redes neuronales, opción de modelado empleada a lo largo de toda la tesis, y se comparan con los

dichos por los antecedentes. Finalmente, en el quinto capítulo se establecen las conclusiones y recomendaciones pertinentes tomando en consideración el potencial impacto de la *Big data* en la predicción de las exportaciones.

# CAPÍTULO I. MARCO TEÓRICO

## 1.1. Antecedentes

En las siguientes líneas, son descritos y citados todos las referencias internacionales y nacionales halladas con relación al tema a investigar. La documentación tiene vinculación con los índices de *Google Trends*, la predicción de series económico-financieras y de consumo de alimentos y las redes neuronales.

### 1.1.1. Antecedentes nacionales

Pilco y Sandoval (2018) orientaron su tesis llamada *Google Trends, inclusión de datos en tiempo real para la predicción de variables macroeconómicas en el Perú* hacia el análisis del uso de la información en tiempo real provista por Google Trends para predecir variables macroeconómicas como la inflación y la inversión. Para ello, utilizaron índices de Google (IG) y modelos ARMA y VAR. El estudio del tipo cuantitativo destacó la importancia de comprender mejor la actividad en línea de la población y cómo se puede aprovechar la enorme cantidad de datos que se ofrecen de forma gratuita en internet. Los datos analizados incluyeron series mensuales desde el 2008 hasta el 2014. Los resultados mostraron que los modelos que incluían los datos del popular motor de búsqueda fueron más efectivos en la predicción de variables que los modelos generales, y que la inclusión de estos datos permitió mejorar los estadísticos de predicción en mayor medida para la inversión que para la inflación. Estos hallazgos subrayan la importancia de utilizar este tipo de información como insumo para las predicciones económicas.

León y Villar (2018) realizaron la investigación *Estudio exploratorio del uso de indicadores de Big Data de Google como variable en el modelamiento del flujo mensual de visitantes al Complejo Arqueológico de Kuélap en el periodo 2011-2016* . El objetivo fue determinar si los datos obtenidos del aplicativo Google Trends tenían un aporte significativo en la

predicción del flujo de visitantes al Complejo Arqueológico de Kuélap, el cual incluyó a visitantes nacionales como de extranjeros. Para lograr esto, las investigadoras utilizaron la ecuación propuesta por Choi y Varian (2009) y analizaron datos mensuales de 2012 a 2016. Las búsquedas en Google, YouTube e Imágenes de Kuélap se emplearon como *inputs* y se incorporaron en un modelo autorregresivo para aumentar su precisión (incremento del R-cuadrado ajustado). El método estadístico utilizado fue la regresión lineal. Los resultados mostraron que sólo las búsquedas en Google contribuyeron a mejorar el ajuste del modelo, mientras que las búsquedas en YouTube y las de imágenes no tuvieron el mismo efecto. La relación entre las búsquedas y los arribos fue directa y significativa.

Sosa y Jeri (2019), en su investigación *Uso de indicadores Big Data para mejorar el nivel de ajuste de un modelo autorregresivo de arribos domésticos al Aeropuerto Internacional Jorge Chávez*, procuraron predecir la cantidad mensual de arribos domésticos al Aeropuerto Internacional Jorge Chávez. Las variables predictoras fueron índices de *Google Trends* relacionados que reflejaban las búsquedas en Google y YouTube acerca de pasajes aéreos a Lima. La muestra analizada abarcó series mensuales desde enero del 2013 hasta diciembre del 2017. Los investigadores concluyeron que la inclusión de las mencionadas variables electrónicas mejoró el nivel de ajuste de un modelo autorregresivo base en 8%. En la etapa cualitativa, los investigadores determinaron que los datos de *Google Trends* son ideales para medir patrones de productos que tienen una alta representatividad en entornos digitales.

Francia y Pacheco (2019), en su tesis *Variables determinantes y relacionadas en las exportaciones de jengibre fresco (Partida Arancelaria 0910.11.00.00) de la República del Perú al Reino de los Países Bajos, comprendido en los años 2012 – 2019*, analizaron los elementos condicionantes de las exportaciones de jengibre fresco (Partida Arancelaria. 0910.11.00.00) de la República del Perú al Reino de los Países Bajos, comprendido en los

años 2012 – 2019. En el estudio se ha identificaron tres variables clave para analizar la inestabilidad presente en el mencionado flujo comercial: el entorno macroeconómico, la competencia internacional y la gestión de inocuidad alimentaria. Tras realizar el análisis se concluyó que el entorno macroeconómico es un factor determinante para las exportaciones, mientras que la competencia internacional. El precio y el tipo de cambio fueron los predictores significativos, no ocurrió lo mismo con el ingreso pér-capita. En la etapa cualitativa, los resultados evidenciaron que la competencia internacional, en particular el indicador de características del producto, es una variable particularmente relevante. El modelo final utilizado ha logrado un R-cuadrado del 51,3%.

Alvarez (2020), en su trabajo de investigación *Determinación de la competitividad y los factores que afectan las exportaciones de jengibre peruano (2012 -2019)*, tuvo por objetivo principal fue identificar y analizar los factores que influyeron en las exportaciones de dicho producto con el fin de establecer si tuvieron un impacto importante que pudiera afectar su grado de competitividad. Mediante un modelo de regresión lineal múltiple, se determinó la conexión entre las variables independientes de las exportaciones de jengibre de la región sudamericana: el precio de exportación del producto, el volumen de producción y el tipo de cambio. Asimismo, el investigador analizó si las exportaciones de este producto para establecer si su crecimiento había sido o no consistente. En conclusión, los resultados adquiridos confirmaron la asociación entre la mayoría de los parámetros analizados; sin embargo, se estableció que la producción de jengibre no tuvo una relación significativa con la exportación de este producto agrícola a lo largo del periodo de estudio.

Lao (2020), en su tesis titulada *Reputación corporativa en un mercado extranjero y su relación con variables del social media - Caso del BCP Bolivia*, se propuso investigar si había una relación entre la reputación del BCP en Bolivia y ciertos aspectos del Social

Media. Para llevar a cabo este estudio, utilizó el índice MERCO, que proporciona datos anuales desde 2013 hasta 2018, y lo correlacionó con seis dimensiones que representaban al *Social Media*. En estas se incluyó el número de comentarios en YouTube en los videos subidos por la empresa cada año, el número de veces que el banco apareció en noticias en los portales de periódicos bolivianos, los índices de búsqueda relativos en Google y YouTube (medidos con Google Trends) sobre el negocio dentro del territorio boliviano, y el puntaje sentimental presente en los comentarios de YouTube y el contenido de los periódicos. Los resultados indicaron que las dos últimas variables, es decir, el puntaje sentimental de los comentarios de YouTube y el contenido de los periódicos, tuvieron una correlación significativa y directa con la reputación del BCP en Bolivia.

Ramirez y Vilela (2020), en su tesis *Modelado del valor de las acciones de Graña y Montero en la NYSE con indicadores del interés hacia el caso Odebrecht*, tuvieron por objetivo pronosticar el precio de apertura semanal de las acciones de Graña y Montero (GyM) en la *New York Stock Exchange* (Bolsa de Valores de Nueva York) empleando como *inputs* los niveles de interés hacia el caso Odebrecht, el mismo que involucraba a la mencionada constructora. El interés hacia el mencionado caso fue medido con *Google Trends* usando como término clave *Caso Odebrecht*. Los investigadores compararon una serie modelos incorporando datos de 151 semanas entre el 2016 y el 2019. Las redes neuronales fueron las opciones de modelado. Los investigadores concluyeron que tanto el interés por el Caso Odebrecht en Estados Unidos y el Perú incidieron en el valor de cada acción de GyM en la NYSE. El modelo final alcanzó una precisión 93.8% superior en 2 puntos porcentuales a la versión que solo incluyó rezagos de dos semanas.

Arevalo & Latour (2021), en su tesis *La participación de las empresas en ferias internacionales y su influencia positiva en la exportación de jengibre de la Región Junín*



*periodo 2016 – 2020*, se orientaron a determinar si la participación en ferias internacionales tuvo un efecto favorable en el aumento de las exportaciones de jengibre fresco de la región de Junín entre 2016 y 2020. Los investigadores evidenciaron que , durante este lapso de tiempo, las exportaciones de jengibre aumentaron enormemente y que Perú se había convertido en uno de los principales exportadores del mundo. Para cumplir con los objetivos especificados, se plantearon hipótesis para identificar los elementos relevantes en el incremento de las exportaciones de jengibre entre 2016 y 2020. Para validar las hipótesis planteadas, los investigadores emplearon una metodología mixta (cuantitativa y cualitativa) recurriendo dos instrumentos: entrevistas a ocho especialistas del sector agroexportador de jengibre y ferias internacionales, y encuestas a setenta representantes de empresas agroexportadoras de jengibre de la Región Junín. Una vez realizada la investigación, se determinó que las ferias internacionales tenían un efecto beneficioso en el crecimiento de las exportaciones de jengibre.

Cisneros et al. (2021), en el artículo *El jengibre y su relación con el comercio exterior en la economía peruana 2010 – 2019* , tuvieron por objetivo fue determinar el grado de incidencia de las exportaciones de jengibre en el comercio exterior de la economía peruana en el transcurso del 2010 al 2019. Para lograr este objetivo, se realizó un análisis del crecimiento de las exportaciones de jengibre, así como la tendencia de las mismas, y se estableció una correlación entre el comportamiento de las exportaciones de jengibre y el comportamiento del PBI total y del PBI Agrícola. Mediante el uso del razonamiento lógico e inductivo, el análisis descriptivo y correlacional, los investigadores examinaron el vínculo sustancial que existía entre los indicadores. Luego de realizar la investigación correspondiente, se pudo establecer una alta correlación entre las exportaciones de jengibre y las exportaciones totales ( $R. = 0,6794$ ); así como con las exportaciones no tradicionales ( $R. = 0,7296$ ); de igual manera, una alta relación con la expansión de las economía peruana, lo cual se evidencia en

la correlación con el Producto Interno Bruto Nacional ( $R. = 0,7971$ ); y el Producto Interno Bruto Agrícola ( $R. = 0,6885$ ).

Alvaro y Martínez (2022), en su tesis *La influencia del Covid-19 en el desempeño de las exportaciones peruanas de jengibre fresco al mercado de EE.UU en 2020*, tuvieron por objetivo explicar cómo la pandemia del COVID-19 afectó el rendimiento de las exportaciones peruanas de jengibre fresco al mercado de los Estados Unidos en el año 2020. Los investigadores utilizaron un enfoque cuantitativo con un diseño no experimental, correlacional-descriptivo y de nivel explicativo. Para recopilar los datos, se utilizó el análisis documental y se procesaron y analizaron los resultados utilizando la prueba t de Student. Los resultados indicaron una correlación positiva alta y significativa, respaldada por un valor p ( $0.001 < 0.05$ ), lo que permitió rechazar la hipótesis nula y aceptar la hipótesis de los investigadores de que el COVID-19 tuvo un impacto significativo en el rendimiento de las exportaciones peruanas de jengibre fresco al mercado estadounidense durante 2020. Los autores concluyeron que hubo un aumento significativo en el envío de jengibre al mercado estadounidense en 2020 en comparación con el año anterior.

En la siguiente tabla se muestra un consolidado de las principales investigaciones llevadas a cabo en el ámbito nacional:

**Tabla 1***Matriz de antecedentes nacionales – Parte 1*

<i>N<sup>o</sup></i>	Autor(es) y año	Carrera o ámbito	Título	Universidad	Tipo de estudio	Relación con la investigación
1	Pilco y Sandoval (2018)	Economía	<i>Google Trends, inclusión de datos en tiempo real para la predicción de variables macroeconómicas en el Perú</i>	Santo Toribio de Mogrovejo	Cuantitativo	Google Trends predice la inflación y la inversión privada.
2	León y Villar (2018)	Administración y Turismo	<i>Estudio exploratorio del uso de indicadores de Big Data de Google como variable en el modelamiento del flujo mensual de visitantes al Complejo Arqueológico de Kuélap en el periodo 2011-2016</i>	USIL	Mixto	Los arribos al atractivo se pudieron modelar en función a las búsquedas web, pero las búsquedas en YouTube y las de imágenes no aportaron de forma significativa.
3	Sosa y Jeri (2019)	Marketing	<i>Uso de indicadores Big Data para mejorar el nivel de ajuste de un modelo autorregresivo de arribos domésticos al Aeropuerto Internacional Jorge Chávez</i>	USIL	Cuantitativo	Las búsquedas en Google y YouTube contribuyen al pronóstico de arribos mensuales al Jorge Chávez.
4	Francia y Pacheco (2019)	Negocios Internacionales	<i>Variables determinantes y relacionadas en las exportaciones de jengibre fresco (Partida Arancelaria 0910.11.00.00) de la República del Perú al Reino de los Países Bajos, comprendido en los años 2012 – 2019</i>	UPC	Cuantitativo	El modelo base aplicado al jengibre logró una precisión de 0.513.
5	Alvarez (2020)	Negocios Internacionales	<i>Determinación de la competitividad y los factores que afectan las exportaciones de jengibre peruano (2012 - 2019)</i>	USIL	Cuantitativo	Modelo de referencia en cuanto a precisión

**Tabla 2***Matriz de antecedentes nacionales – Parte 2*

<i>N<sup>o</sup></i>	Autor(es) y año	Carrera o ámbito	Título	Universidad	Tipo de estudio	Relación con la investigación
6	Lao (2020)	Negocios Internacionales	<i>Reputación corporativa en un mercado extran"jero y su relación con variables del social media - Caso del BCP Bolivia</i>	USIL	Cuantitativo	Las búsquedas de YouTube y web no contriuyen a conocer la reputación corporativa de la empresa analizada, no ocurre lo mismo con el sentimiento reflejado en los comentarios.
7	Ramirez y Vilela (2020)	Negocios Internacionales	<i>Modelado del valor de las acciones de Graña y Montero en la NYSE con indicadores del interés hacia el caso Odebrecht</i>	USIL	Cuantittativo	La incorporación de datos de Google Trends incrementaron el <i>accuracy</i> del modelo basado en rezagos de 0.91 a 0.936, todo ello bajo el enfoque de las redes neuronales.
8	Arevalo y Latour (2021)	Negocios Internacionales	<i>La participación de las empresas en ferias internacionales y su influencia positiva en la exportación de jengibre de la Región Junín periodo 2016 – 2020</i>	UPC	Cuantittativo	Evidenciaron que la participación en ferias tenía relación con la apertura de mercado para el caso de los exportadores de Jengibre.
9	Cisneros et al. (2021)	Negocios Internacionales	<i>El jengibre y su relación con el comercio exterior en la economía peruana 2010 – 2019</i>	Hemilio Valdizán	Cuantittativo	El jengibre se correlaciona con la expansión del PBI.
10	Alvaro y Martinez (2022)	Negocios Internacionales	<i>La influencia del Covid-19 en el desempeño de las exportaciones peruanas de jengibre fresco al mercado de EE.UU en 2020</i>	Científica del Sur	Cuantittativo	El COVID-19 impulsó las exportaciones de jengibre

### 1.1.2. Antecedentes internacionales

Park (2015), en su investigación *Does the Rise of the Korean Wave Lead to Cosmetics Export?*, tuvo por objetivo determinar la relación entre la ola coreana (*korean wave*) y la exportación coreana de cosméticos. En lugar de emplear los datos convencionales de COMTRADE de la ONU, utilizados en investigaciones similares, este estudio incorporó el índice de consulta de Google Trends con la palabra clave "Korean drama" como variable proxy para aproximar el comercio influenciado por factores culturales. El investigador examinó si la popularidad de la ola coreana, representada por Google Trends, contribuyó al aumento explosivo de la exportación de cosméticos coreanos en los últimos años. Además, el estudio también investigó los posibles efectos de la ola coreana en las exportaciones, que podrían variar según los diferentes grupos comerciales clasificando a los países de importación en dos categorías: 74 países de todo el mundo y 9 países miembros de la Asociación de Naciones del Sureste Asiático (ASEAN). Como método estadístico de predicción, el investigador utilizó la regresión lineal. Los resultados revelaron que la popularidad de la ola coreana contribuyó significativamente a estimar las exportaciones de cosméticos a los países de la ASEAN, pero presentó una relación débil con las exportaciones de cosméticos a nivel mundial.

Abraham et al. (2017) , en el paper de conferencia titulado *Neural Network System to Forecast the Soybean Exportation on Brazilian Port of Santos*, tuvieron por objetivo construir, entrenar y simular una Red Neuronal Artificial (RNA) orientada al pronóstico de las exportaciones de soja de Brasil. Ello tomando en consideración que la agricultura desempeña un papel importante en la economía brasileña. El país fue uno de los principales productores de soja del mundo, con 114 millones de toneladas producidas durante la cosecha 2016-2017 ; Mato Grosso es el mayor productor de Brasil, con 30.500.000 toneladas, y el puerto de Santos tiene una gran demanda debido a su condición de mayor puerto de América

Latina. Sin embargo, la inadecuada infraestructura de las carreteras de transporte generó cuellos de botella en el envío de soja a través de los principales puertos. El investigador indicó que las (RNA) se utilizan globalmente en logística; por ello, propuso construir, entrenar y simular una RNA con el software MatLab para estimar la demanda de soja producida en Mato Grosso y exportada a través del puerto de Santos. La información fue recogida de los Sistemas de Información de Comercio Exterior (ALICE WEB). Este sistema era mantenido por el Ministerio de Industria, Comercio Exterior y Servicios y se basa en información recuperada del Sistema Integrado de Comercio Exterior (SISCOMEX). Dicho sistema proporcionó las estadísticas de las transacciones brasileñas de exportación e importación. El sistema generó una hoja de cálculo de Microsoft Excel, que simplemente se importó a una matriz de Matlab. En este estudio, el análisis de los datos de exportación de soja se presentó en toneladas y fue asociado con la producción del estado de Mato Grosso que salió por el puerto de Santos en el período 2001-2016. Los investigadores lograron anticipar un valor de 9,0 millones de toneladas para 2017, lo que representaba un aumento de alrededor del 26,5% en comparación con el volumen de 2016 de 7,1 millones de toneladas.

Boone et al. (2018), en el *journal article* titulado *Can Google Trends Improve your Sales Forecast?*, tuvieron por objetivo relacionar a las búsquedas *online* con las ventas de un minorista de alimentos. Este minorista se describió a sí mismo como un proveedor de alimentos especiales de primera calidad. Este proporcionó información sobre las ventas y los precios de una serie de productos (carnes especiales, especias, utensilios de cocina, vinos y quesos). Fueron analizados un total de 169 series semanales que incluyeron datos desde el 2013 hasta el 2015, las ventas de cada producto fueron correlacionadas con sus rezagos, precios y las búsquedas *online* asociadas a las marcas de cada producto (*Antipasto*, *Hors d'oeuvres*, entre otras), obtenidas de *Google Trends*. Luego de emplear un modelo de

regresión lineal, los investigadores evidenciaron que , al incorporar las búsquedas *online*, la precisión de los modelos que solo incluían rezagos y precios incrementaron su precisión en 0.02 y 0.06.

Woo y Owen (2018), en su *journal article* titulado *Forecasting private consumption with Google Trends data*, tuvieron por objetivo investigar la relación prevista entre los datos de *Google Trends* relacionados con el consumo y con las noticias y los cambios en el consumo privado en Estados Unidos. Los resultados indicaron que (1) los modelos aumentados con *Google Trends* proporcionaron información adicional sobre el consumo además de los indicadores de sentimiento del consumidor basados en encuestas, (2) los datos de *Google Trends* relacionados con el consumo proporcionaron información sobre las tendencias de investigación previas al consumo, (3) los datos de *Google Trends* relacionados con las noticias proporcionaron información sobre los cambios en el consumo de bienes duraderos, y (4) la combinación de datos relacionados con las noticias y el consumo mejora significativamente los modelos de previsión. Los investigadores demostraron que la utilización de estos datos mejora las estimaciones del crecimiento del consumo privado en comparación con las previsiones que no emplean los datos de *Google Trends* y con las previsiones que utilizan los datos de *Google Trends* pero que no tienen en cuenta la forma única en que informan las previsiones. Cabe destacar que los investigadores resaltaron que las búsquedas de noticias no inciden en el pronóstico de servicios o bienes no duraderos. Todo lo anterior se evidenció empleando modelos de regresión lineal.

Yu et al. (2019), en el *journal article* *Online big data-driven oil consumption forecasting with Google trends* , tuvieron por objetivo desarrollar un modelo de previsión del consumo de petróleo basado en la *big data* de *Google Trends*. Esta investigación constó de dos procesos principales, que son el análisis de las correlaciones y la mejora de las predicciones.

Para empezar, fue llevada a cabo una prueba de cointegración y un análisis de causalidad de Granger con el fin de comprobar estadísticamente el poder predictivo de las tendencias de Google en cuanto a tener una relación significativa con el consumo de petróleo. Estas pruebas se llevaron a cabo para determinar si existe o no una correlación entre las dos variables. En segundo lugar, los datos de *Google Trends* se incorporaron a las técnicas convencionales de previsión para anticipar los patrones de consumo de petróleo, así como los valores. El estudio empírico de la predicción del consumo mundial de petróleo demuestra que el trabajo de previsión en línea basado en *big data* con *Google Trends* supone una mejora significativa respecto a las técnicas convencionales que no utilizan *Google Trends*: 4% usando el modelo *Extreme Machine Learning* y 0.09% con el modelo de regresión lineal.

Akman et al. (2020), en su *journal article* titulado “*Visa trial of international trade: evidence from support vector machines and neural networks*”, fue orientado a emplear técnicas de aprendizaje automático: máquinas de soporte vectorial (SVM) y redes neuronales (NN), para examinar la influencia de la facilitación del visado en el comercio bilateral utilizando datos de exportación de Turquía del 2000 a 2014. Los investigadores incidieron en que el comercio internacional depende de la creación de redes, el compromiso y los encuentros cara a cara, lo que fomenta los viajes internacionales. Para potenciar el comercio bilateral, el turismo y las inversiones extranjeras directas, algunos países están aplicando políticas que fomentan la movilidad entrante. En la última década, Turquía ha sido una de las naciones que ha aplicado normas liberales en materia de visados para promover el comercio, lo que ha dado lugar a grandes ganancias en los volúmenes de exportación. La investigación desentrañó las variables con mayor impacto en el comercio y demostró que las políticas de visado tenían efectos considerables en las exportaciones de origen turco. El modelo final de mayor precisión, basado en redes neuronales y construido con *SPSS Modeler*, alcanzó una precisión de 79.21%, donde la presencia del visado fue el tercer



predictor en importancia dentro de un conjunto de 9 variables, dicha importancia fue estimada por el *SPSS Modeler*. En un esfuerzo por aumentar las exportaciones, los investigadores aconsejaron a los gobiernos que adopten normas de visado más permisivas.

Collison (2020) indicó en *Performance Assessment of Google Trend Index in Forecasting Car Sales in Argentina* que, en los últimos años, la literatura económica ha utilizado la tecnología para tratar de mejorar el ajuste de modelos predictivos. Con la *Big Data* y la gran cantidad de información relativa a los hábitos de consumo, muchos modelos de previsión de ventas han explotado la información *online*. Dado dicho contexto, el autor indicó que las búsquedas en *Google* pueden proporcionar información valiosa sobre las preferencias de los consumidores para un producto. El investigador tuvo como propósito evaluar la contribución de las búsquedas de Google en la previsión de ventas de automóviles Ford en Argentina. Para este objetivo, construyó un modelo que incorpora el Índice de Tendencias de Google para capturar la frecuencia de búsqueda en Internet, y comparamos este modelo con un modelo estándar sin dicho Índice como una variable explicativa. El investigador concluyó que la incorporación de los indicadores de Google contribuye a mejorar el R-cuadrado del modelo y a disminuir los residuos (diferencia entre los valores predichos y observados). Ello fue verificado comparando dos modelos base: uno basado en los rezagos correspondiente a un mes y año (en escala logarítmica) y otro que adicionó las búsquedas web. El primero alcanzó una precisión de 0.283; mientras que el segundo fue logró una precisión de 0.495.

Ma et al. (2020), en el *journal article* titulado *Media influences on consumption trends: Effects of the film Food, Inc. on organic food sales in the U.S.*, tuvieron por objetivo examinar empíricamente el impacto de *Food, Inc*, un influyente documental sobre la industria alimentaria estadounidense, en las ventas de alimentos orgánicos. Los investigadores emplearon el volumen de búsqueda en Google para medir la influencia de la

película. El conjunto de datos de IRI Marketing Science proporcionó información sobre la compra de alimentos orgánicos para tres categorías: yogur, mantequilla de cacahuete y café. Este conjunto de datos, que abarcó 11 años, ofreció información sobre las ventas semanales y las características de la mezcla de marketing de más de 1.700 supermercados que atendían al 47% de la población estadounidense. Los investigadores hipotetizaron que, si la exposición al documental *Food, Inc.* influyó en el consumo de alimentos orgánicos, deberían encontrar un aumento en el *market share* de los alimentos orgánicos en los mercados con un mayor número de búsquedas de la película en Google tras su estreno. En consecuencia, la técnica de identificación empírica dependió tanto de la variación longitudinal de la popularidad de la película a lo largo del tiempo como de la variación transversal asociada a los mercados. Los investigadores destacaron que estudios anteriores evidenciaron que, aunque los medios de comunicación pueden influir en el comportamiento de las personas a corto plazo, es poco probable que los efectos sean duraderos, ya que la atención de las personas cambia y los recuerdos se desvanecen. En general, *Food, Inc.* tuvo un impacto positivo en las cuotas de los alimentos orgánicos en dos de las tres categorías estudiadas: un aumento de una desviación estándar en las búsquedas de la película en Google corresponde a un incremento del 0,66% en la cuota de mercado del yogur orgánico (del 7,8% al 8,5%) y un aumento del 0,15% en la cuota de mercado de la mantequilla de cacahuete (del 2,7% al 2,7%). Sin embargo, los efectos disminuyeron drásticamente con el tiempo: los intervalos de duración del efecto para las categorías de yogur y mantequilla de cacahuete fueron de 4,7 meses y 5,5 meses, respectivamente, según los coeficientes de Google.

Mayasari et al. (2020) indicaron en *Pandemic on Food Security and Diet-Related Lifestyle Behaviors: An Analytical Study of Google Trends-Based Query Volumes* que el Coronavirus (SARS-CoV) -2 (COVID) -19 está teniendo efectos profundos en la economía mundial y el comercio de alimentos. Los investigadores resaltaron el hecho de que existen datos limitados

disponibles sobre cómo esta pandemia está afectando a nuestros comportamientos dietéticos y relacionados con el estilo de vida a nivel global. *Google Trends* se utilizó para obtener volúmenes de búsqueda relativos en todo el mundo que cubren un periodo de tiempo desde antes de la pandemia de COVID-19, más precisamente, desde el 1 de junio de 2019 a 27 de abril de 2020. Entre otros hallazgos, los investigadores determinaron que existe una correlación directa y significativa (0.350) entre el interés que suscita el COVID-19 y las búsquedas relacionadas al consumo de jengibre (*ginger*). Los investigadores sugieren que las búsquedas que involucran al jengibre tienen una relación con el consumo de dicho producto alrededor del mundo.

Chang y Meyerhoefer (2021), en su *journal article* titulado *COVID-19 and the Demand for Online Food Shopping Services: Empirical Evidence from Taiwan*, tuvieron como objetivo analizar el impacto de la epidemia del COVID-19 en la demanda de servicios de compra de alimentos en línea. Para ello utilizaron datos búsquedas acerca de Ubox, la mayor plataforma de comercio electrónico agroalimentario de Taiwán, junto con búsquedas en Google de palabras clave y artículos de prensa relacionados con el COVID-19. Fueron analizados series semanales desde enero del 2017 hasta junio del 2020, ello con el fin de comprender mejor los escenarios con la ausencia y presencia de la pandemia. Los investigadores encontraron una mayor correlación entre las ventas de Ubox y la frecuencia con la que se citaba a COVID-19 en los periódicos que entre las ventas de Ubox y las búsquedas de COVID-19 en Google. En concreto, un aumento del 1% en las búsquedas de palabras clave en Google se correlacionó con un aumento estadísticamente significativo del 0,8% en las ventas de Ubox; en comparación, un aumento del 1% en las menciones de COVID-19 en los periódicos está relacionado con un aumento del 3,8% en las ventas de Ubox, un incremento mucho mayor.

France et al. (2021), en su *journal article* titulado *Web Trends: A valuable tool for business research*, afirmaron que las estadísticas sobre empresas y la economía pueden extraerse de la información de las búsquedas *online*, que es un recurso muy importante. Los investigadores afirmaron que, en estudios anteriores, los economistas han utilizado los datos de *Google Trends* para hacer previsiones económicas. Los investigadores ampliaron el alcance ese tipo de trabajos proporcionando un modelo en donde el *Brand Value*, obtenido de *Interbrand.com*, se correlacionó con los volúmenes de búsquedas *online* realizadas desde Estados Unidos, obtenidos de *Google Trends*, para las 100 principales marcas de dicho país. Los investigadores evidenciaron que había una relación lineal entre ambas variables: el índice de búsquedas dependía del valor registrado en *Interbrand.com*. La data correspondió al año 2020. A modo de aplicación complementaria, los investigadores correlacionaron también las taquillas de películas exhibidas entre los años 2012 y 2013 (obtenidos de *boxofficemojo.com*) con las búsquedas *online* correspondientes a las 3 semanas anteriores al estreno de cada película. También fueron considerados otros predictores como el hecho de que la película tuviera una precuela, el número de cines en donde iba a ser exhibida la cinta, el *Budget* de la producción, entre otras. Con una precisión del modelo de 75.9%, los autores del artículo también concluyeron que las búsquedas *online* eran predictoras de las taquillas.

Başıyigit (2021) publicó su investigación *Can Google Trends improve the marble demand model: A case study of USA's marble demand from Turkey*. En la misma, el investigador tuvo por objetivo evidenciar el impacto de Google Trends en la demanda de mármol. Estados Unidos es el segundo mayor importador de minerales de Turquía, siendo el mármol el producto mineral que más se importa del país. Por ello, para naciones como Turquía, conocer la función de la demanda de mármol es una cuestión esencial a estudiar. El mármol se distingue de otros productos minerales por el hecho de que es utilizado por personas y familias. Las tendencias y la popularidad, al igual que con otros artículos de uso final, son

variables esenciales para impulsar la demanda de mármol en el mercado. La predilección por las casas unifamiliares en Estados Unidos, así como la personalización de estos hogares por parte de las familias individuales, contribuye a la creciente relevancia del mármol. Los estudios sobre la demanda en el ámbito de los productos minerales se han centrado hasta ahora únicamente en parámetros de la demanda como el PIB per cápita, que es un hecho relativamente reciente. Por otra parte, no han considerado el impacto de las tendencias evidenciadas en *Google*. Para determinar el impacto de las mismas en la demanda de mármol, se construyó un modelo base basado en los factores de paridad, PIB per cápita y precio. El modelo de base era estadísticamente significativo y tenía un alto poder explicativo ( $R^2 = 0.85$ ), lo que indicaba que tenía un buen potencial predictivo. El investigador decidió crear un segundo modelo que incorporara la variable de tendencias, que se recogió de *Google Trends*. La incorporación de las tendencias al modelo básico dio lugar a una mejora de los indicadores analíticos ( $R^2 = 0.90$ ).

Golovanova y Zubarev (2021), en el *journal article* titulado *Forecasting Aggregate Retail Sales with Google Trends*, tuvieron por objetivo evaluar la contribución de las consultas de búsqueda de diferentes grupos de productos al poder predictivo de los modelos de ventas minoristas de alimentos y productos que no son alimentos, así como de las ventas de grupos de productos específicos. La data analizada correspondió a series mensuales del 2015 al 2021, siendo el Banco de Rusia la fuente consultada para las ventas del sector Retail, mientras que las búsquedas web se obtuvieron de *Google Trends*. Fue empleado el método de regresión lineal. Adicionalmente fueron incorporadas variables macroeconómicas como el índice de volatilidad ruso, el precio del crudo Brent, el tipo de cambio efectivo real y el indicador de confianza de los consumidores. Los resultados evidenciaron que los índices de búsquedas web contribuyeron en la predicción de los modelos durante el intervalo de tiempo investigado. Esto es especialmente evidente en los modelos para las ventas minoristas no

alimentarias, donde el valor del coeficiente de determinación mejora en aproximadamente 0.2 en los mejores modelos; mientras que, para los alimentos, la mejora fue de 0.06. Los investigadores concluyen destacando que los resultados indicados anteriormente se han obtuvieron a partir de las búsquedas en Google, que sólo representa aproximadamente un tercio del mercado de los motores de búsqueda en Rusia. Sin embargo, *Yandex*, el líder del mercado en Rusia en el momento en que se redactó el informe, no ofrecía datos que cumplan los requisitos necesarios para la investigación.

Gopinath et al. (2021), en su investigación *International agricultural trade forecasting using machine learning*, recurrieron a la analítica de datos. En este estudio se emplearon datos secundarios para interpretar los patrones de comercio, concretamente utilizando el aprendizaje automático (o *Machine Learning*) supervisado, así como las redes neuronales. El estudio se centró en siete productos agrícolas principales que tienen un largo historial de comercio. Las técnicas de aprendizaje automático supervisado y las redes neuronales fueron ejecutadas con data que abarcó desde 1962 hasta el 2016 . Al ser comparados con los métodos tradicionales, que suelen consistir en evaluaciones subjetivas o en predicciones de series temporales, los resultados demostraron que los modelos de aprendizaje automático tienen una gran relevancia para predecir las tendencias comerciales a corto y largo plazo. En lo que respecta a los flujos comerciales agrícolas, los algoritmos de aprendizaje automático supervisados fueron capaces de cuantificar los factores económicos cruciales, pero los enfoques de redes neuronales ofrecen ajustes superiores a largo plazo.

Borup y Montes (2022), en su *journal article* titulado *In Search of a Job: Forecasting Employment Growth Using Google Trends*, tuvieron por objetivo demostrar que la actividad de búsqueda en Google sobre frases relevantes fue un robusto predictor que mostró el crecimiento del empleo en Estados Unidos entre 2004 y 2019, tanto a corto como a largo

plazo. Los investigadores utilizaron los datos de *Google Trends* para medir consultas de búsqueda semánticamente relevantes, ellos generaron un panel de 172 keywords basados en la búsqueda de empleo (“*indeed*”, “*part-time jobs*”, “*online jobs*”, etc.). El mejor modelo de Google Trends, que superó en precisión a modelos con una variedad de variables macroeconómicas y financieras, obtuvo un R2 fuera de muestra de entre el 29% y el 62% para horizontes que van de un mes a un año. Los autores señalaron que esta robusta predicción fue el resultado de la heterogeneidad de las frases de búsqueda y que se puede extender el modelo al crecimiento del empleo a nivel industrial y estatal utilizando la actividad de búsqueda específica de cada Estado.

Onomayo et al. (2022), en el *journal article* titulado *Factors influencing export performance of Ginger (Zinbiger officinale) in Nigeria*, se orientaron a evaluar los factores que afectaron al rendimiento de las exportaciones de jengibre (*Zinbiger officinale*) en Nigeria.. Las variables consideradas fueron la cantidad de exportación de jengibre (*Zinbiger officinale*), el tipo de cambio, el tipo de interés, la relación entre el precio del productor y el precio interno, y la relación entre el precio del productor y el precio de exportación. Las fuentes secundarias proporcionaron los datos desde 1995 hasta 2020. Los instrumentos estadísticos y econométricos utilizados para alcanzar los objetivos mencionados fueron el test de Dickey Fuller Aumentado (ADF), la prueba de Phillips-Perron (PP), la prueba de cointegración de Johansen y el modelo de corrección de errores vectorial (VECM). Los resultados de la prueba de raíz unitaria de Dickey-Fuller aumentado (ADF) indicaron que ninguna de las variables de prueba estaba nivelada. Todas las variables de prueba fueron estadísticamente significativas al nivel de probabilidad del 5% para la primera diferencia. La prueba de raíz unitaria de Phillip-Perron (PP) demostró que todas las variables de prueba eran estacionarias. La prueba de cointegración realizada por Johansen indicó la cointegración y las correlaciones a largo plazo entre las variables. Los resultados a corto plazo de las

exportaciones de jengibre en Nigeria estuvieron significativamente influenciados por el tipo de cambio (EXR) y el tipo de interés (INR), según el modelo de corrección vectorial de errores (VECM). El rendimiento de las exportaciones de jengibre (Zinbiger officinale) en Nigeria se fue significativamente influenciado a corto plazo por las relaciones entre el precio de producción y el precio interno (PP), el precio de producción y el precio de exportación (PX), y la cantidad de jengibre enviada (Y). Estas variables fueron estadísticamente significativas. Los coeficientes del modelo de corrección del error (MCE) fueron negativos, lo que evalúa la tasa de ajuste del equilibrio a largo plazo. El coeficientes de determinación múltiple (R<sup>2</sup>) registrado fude de 0.6710. Esto, de acuerdo a los investigadores, demostró la existencia de un modelo aceptable. El documento sugiere formulaciones e implementaciones de políticas que estabilicen el tipo de cambio, también sugirieron que el tipo de interés no debería superar un dígito.

A continuación se muestran las tablas que resumen a los antecedentes internacionales:



**Tabla 3***Matriz de antecedentes internacionales- Parte 1*

Número	Autores	Título	Journal o Conferencia	Fuente de datos	Aporte a la investigación
1	Park (2015)	<i>Does the Rise of the Korean Wave Lead to Cosmetics Export?</i>	Journal of Asian Finance, Economics and Business	UN COMTRADE	Relación entre las búsquedas online y las exportaciones
2	Abraham et al. (2017)	<i>Neural Network System to Forecast the Soybean Exportation on Brazilian Port of Santos</i>	IFIP Advances in Information and Communication Technology	ALICE WEB	Caso aplicado de las redes neuronales a las exportaciones
3	Boone et al. (2018)	<i>Can Google Trends Improve your Sales Forecast?</i>	Production and Operations Management	Google Trends y datos brindados por una empresa	
4	Woo y Owen (2018)	<i>Forecasting private consumption with Google Trends data</i>	Journal of Forecasting	Google Trends	Sugiere que los patrones de búsqueda se asocian al consumo pero no lo demuestra.
5	Yu et al. (2019)	<i>Online big data-driven oil consumption forecasting with Google trends</i>	International Journal of Forecasting	Google Trends	Google Trends como forma simplificada de <i>Big Data</i> .
6	Akman et al. (2020)	<i>Visa trial of international trade: evidence from support vector machines and neural networks</i>	Journal of Management Analytics	US Visa Waiver Program (VWP) y Banco de Turquía	Hace referencia al poder predictor de las variables

**Tabla 4***Matriz de antecedentes internacionales- Parte 2*

Número	Autores	Título	Journal o Conferencia	Fuente de datos	Aporte a la investigación
7	Collison (2020)	<i>Performance Assessment of Google Trend Index in Forecasting Car Sales in Argentina</i>	SSRN	Google Trends	Aplicación de la Big Data al comercio de autos.
8	Ma et al. (2020)	<i>Media influences on consumption trends: Effects of the film Food, Inc. on organic food sales in the U.S.,</i>	International Journal of Research in Marketing	Google Trends e IRI Marketing Science	Efecto de una película (cuantificada con Google) en el consumo de alimentos orgánicos.
9	Mayasari et al. (2020)	<i>Pandemic on Food Security and Diet-Related Lifestyle Behaviors: An Analytical Study of Google Trends-Based Query Volumes</i>	Nutrients	Google Trends	Existe la posibilidad de que las tendencias en el consumo se puedan medir con Google Trends
10	Chang y Meyerhoefer (2021)	<i>COVID-19 and the Demand for Online Food Shopping Services: Empirical Evidence from Taiwan</i>	American Journal of Agricultural Economics	Google Trends y UBox	Caso de aplicación al sector retail de alimentos.
11	France et al. (2021)	<i>Web Trends: A valuable tool for business research</i>	Journal of Business Research	Google Trends, Interbrand.com y boxofficemojo.com	Correlación entre variables empresariales y las búsquedas online.
12	Başyigit (2021)	<i>Can Google Trends improve the marble demand model: A case study of USA's marble demand from Turkey</i>	Resource Policy	Google Trends y Banco Mundial	Caso aplicado de las búsquedas online al comercio exterior.

**Tabla 5***Matriz de antecedentes internacionales- Parte 3*

Número	Autores	Título	Journal o Conferencia	Fuente de datos	Aporte a la investigación
13	Golovanova y Zubarev (2021)	<i>Forecasting Aggregate Retail Sales with Google Trends</i>	Bank of Russia	Google Trends y Banco de Rusia	Comparativo
14	Gopinath et al. (2021)	<i>International agricultural trade forecasting using machine learning</i>	Data & Policy	Gurevich & Herman (2018)	Sugiere incorporar al <i>machine learning</i> como método predictivo para hacer frente a una demanda cada vez más
15	Borup y Montes (2022)	<i>In Search of a Job: Forecasting Employment Growth Using Google Trends</i>	Journal of Business & Economic Statistics	Google Trends y U.S. Labour of Statistics	Uso combinado de <i>keywords</i> en Google Trends para mejora de pronósticos.
16	Onomayo et al. (2022)	<i>Factors influencing export performance of Ginger (Zinbiger officinale) in Nigeria</i>	International Journal of Agriculture, Environment and Food Sciences	Central Bank of Nigeria data base & Bureau of Statistics of Nigeria,	Modelo de referencia extranjero aplicado a la exportación de jengibre.

## 1.2. Bases Teóricas

### 1.2.1. Teorías clásicas del Comercio Internacional vinculadas a la agricultura

#### *1.2.1.1. Teoría del Mercantilismo*

En los anales de la historia económica, el término "mercantilista" se refiere a las primeras concepciones del comercio internacional; entre los siglos XVI y XVIII en Occidente, los argumentos y regulaciones en torno al comercio internacional fueron el principal motor del desarrollo de la idea (Allen, 1991; Verter, 2015). El modelo hacía hincapié en que los gobiernos debían desincentivar las importaciones mediante impuestos y cuotas y estimular las exportaciones mediante subvenciones y ayudas a la exportación, además de la recolección de metales preciosos, lo cual debería realizarse para maximizar la eficiencia económica (Verter, 2015).

Los defensores del mercantilismo favorecían el comercio de exportación, ya que se consideraba que mejoraba el bien (la riqueza) de una nación, y viceversa para el comercio de importación (Paul, 2008). La idea proponía que, para que una nación tuviera una balanza comercial positiva, era necesario estimular las exportaciones, la sustitución de importaciones y la creación de capital financiero (principalmente en forma de oro y plata) (Verter, 2015). El medio ordinario para aumentar nuestra riqueza y tesoro es el comercio exterior, en el que se debería observar siempre esta regla: vender más a los extranjeros anualmente de lo que consumimos de ellos en valor (Mun, 1964, como se cita en Verter, 2015). Según esta idea, existe una cantidad predeterminada y finita de riqueza en el mundo; en consecuencia, para que una nación aumente su riqueza, necesita robar directa o indirectamente algunos recursos de otra nación (Verter, 2015).

Desde luego, el modelo mercantilista tiene sus detractores, que creen que la idea proporciona "una falsa unidad a los acontecimientos heterogéneos", lo que, en cierta medida, ha ahogado

la prosperidad, especialmente en los países emergentes. Por ejemplo, el concepto de precio-especie-flujo de David Hume del siglo XVIII afirmaba que una balanza comercial favorable sería factible, pero sólo a corto plazo (Oxford Reference, s.f.). Sin embargo, esto sólo se aplica a la situación a corto plazo. El Chicago Booth Review (s.f.), subrayó que el sistema mercantilista no era más que una conspiración masiva de comerciantes y fabricantes en perjuicio de los consumidores; Smith también afirmó que la idea no ofrecía la posibilidad de que los clientes nacionales eligieran variaciones de artículos que se fabricaban en otras naciones, haciendo referencias a los artículos importados. Se pensaba que la teoría era un "juego de suma cero", también conocido como "juego de ganar y perder", que indica que cualquier beneficio obtenido por una nación puede traer una pérdida proporcional a la otra nación que participaba en la transacción; en otras palabras, la teoría era un "juego de ganar-perder" (Verter, 2015).

#### *1.2.1.2. Teoría de Comercio con Ventaja Absoluta*

El desarrollo de esta idea se atribuye a Adam Smith, considerado a menudo como el fundador de la economía moderna y el primero en defender el libre comercio (Oxford Reference, s.f.). La ventaja absoluta, según él, se refiere al proceso por el que un individuo o nación puede fabricar un determinado bien a un coste inferior al de otro individuo o nación. lo cual se caracteriza como una ventaja absoluta; por lo tanto, una nación que se dedique al comercio transfronterizo debe centrarse en fabricar el tipo de productos para los que tenga una clara ventaja competitiva sobre otras naciones (Britannica, s.f.). Según el razonamiento de Smith, lo que es sabiduría en la conducta de un solo hogar apenas puede ser insensatez en la de un país enorme: si una nación extranjera puede proporcionarnos una mercancía a un precio inferior al que podemos producir nosotros mismos, entonces nos conviene comprar esa

mercancía a esa nación y utilizar una parte de la producción de nuestra propia industria de manera que tengamos alguna ventaja competitiva.

Smith creía que el aumento de la especialización en la producción de bienes y servicios conduciría a un aumento de la cantidad de producción global; además, la eficiencia global aumenta en la utilización de los recursos disponibles cuando un país exporta una parte de los bienes que produce a un coste menor e importa los productos que produce su socio comercial a un coste menor que en su país; esto crea una situación en la que ambos países salen ganando (Leen, 2014; Verter, 2015). Smith también sostenía que era imposible que las naciones que participaban en un comercio se beneficiaran de tales interacciones de la forma en que las apoyaba el mercantilismo; esto se debe a que la exportación de un país es la importación de otro (Leen, 2014; Verter, 2015). El padre del liberalismo económico también sostiene que todas las naciones estarían mejor si adoptaran políticas de libre comercio y centraran sus economías en la producción de bienes al menor coste posible; ello sugiere que el comercio internacional es factible cuando una nación produce un determinado artículo con menor cantidad de mano de obra que la otra nación, y viceversa (Verter, 2015).

Smith sostenía que, en una época de mercados libres, aunque el empleo de un Estado pueda superar el consumo interno, esto motivaría a la nación a aumentar sus capacidades productivas. (Leen, 2014; Verter, 2015). Smith afirmó que este sería el caso incluso si el consumo interno era mayor que el empleo; como consecuencia directa de ello, el ritmo de acumulación de dinero y riqueza de la nación se aceleraría de forma espectacular (Verter, 2015). Smith creía que cada nación y cada individuo tienen algún tipo de superioridad sobre los demás. ¿Qué pasa si la nación goza de una importante ventaja competitiva en la producción de todo? ¿Seguirá el país produciendo todos los artículos tanto para el mercado interno como para el mercado internacional? Cuando una sola nación produzca todos los

artículos, ¿seguirá siendo factible que las dos naciones comercien entre sí y disfruten de los beneficios de hacerlo? Estos tópicos de discusión que se pueden solucionar tomando como base la teoría de Smith (Verter, 2015).

### *1.2.1.3. Teoría de la ventaja Competitiva*

La idea de la ventaja comparativa fue propuesta por David Ricardo como un intento de abordar varias preocupaciones que no fueron resueltas satisfactoriamente por la hipótesis de la ventaja absoluta (1817). Ricardo sostenía que las naciones se beneficiarían del comercio, aunque un país tuviera una ventaja absoluta sobre el otro en la producción de todos los artículos con los que comerciaban, incluso si ambos países comerciaban con la misma cantidad de productos (Verter, 2015). Ricardo subrayó la importancia de que la nación se especialice en la producción de aquellos artículos para los que tiene la mayor producción con el menor coste de oportunidad en relación con las otras naciones; mientras los costes de fabricación varíen entre las naciones que se dedican al comercio, se puede argumentar que cada nación posee una ventaja comparativa para un producto que registró la mayor eficiencia en la producción (Costinot & Donaldson, 2012; Verter, 2015).

La teoría de la ventaja comparativa se basa en una serie de supuestos, entre los que se encuentran los siguientes: sólo dos países participan en el comercio; sólo se comercian dos productos; existe una competencia perfecta; no hay efectos del comercio en la distribución de la renta entre los países que comercian juntos; el nivel tecnológico difiere entre los países; el comercio es necesario en parte debido a las diferencias en la productividad del trabajo entre los países; el trabajo es el único factor de producción; no hay restricciones al comercio; y existe un equilibrio de poder (Verter, 2015). Sin embargo, los detractores de esta idea subrayan que no explica las razones por las que las naciones que participan en el comercio tienen distintos niveles de productividad laboral y avances tecnológicos. La hipótesis de las

proporciones de los factores, que se basa en el modelo de la ventaja comparativa, es un intento de arrojar luz sobre las razones que explican la variación de la productividad nacional del trabajo y los avances tecnológicos (Bouare, 2009; Costinot & Donaldson, 2012).

#### *1.2.1.4. Heckscher – Ohlin Trade Theory*

Teoría de la ventaja comparativa basada en el hecho de que los países con mayor abundancia de capital tienden a exportar bienes intensivos en capital e importar bienes intensivos en mano de obra, mientras que los países con menor abundancia de capital tienden a exportar bienes intensivos en mano de obra y traer bienes más abundantes en capital. Esto se conoce como la teoría Heckscher-Ohlin en economía. Bertil Ohlin (1899-1979) creó la idea a partir de los trabajos realizados por su tutor, Eli Filip Heckscher (1879-1952). En 1977, Ohlin recibió el Premio Sveriges Riksbank de Ciencias Económicas en Memoria de Alfred Nobel por su trabajo sobre la teoría (Britannica, The Editors of Encyclopaedia, 2021).

El capital es abundante en ciertas naciones, y el trabajador medio tiene acceso a una variedad de máquinas y herramientas que le ayudan en su tarea.; los salarios en estas naciones tienden a ser altos, lo que hace que los productos que requieren mucha mano de obra, como los textiles, los artículos deportivos y la electrónica de consumo básica, sean más costosos de lo que serían en países donde la mano de obra es barata y los salarios son bajos (Lundahl, 2022; Verter, 2015). Los automóviles y los productos químicos, por ejemplo, tienden a ser más caros en las naciones con capital abundante y barato, ya que requieren mucho capital y sólo una pequeña cantidad de mano de obra. La capacidad de un país para crear productos intensivos en capital a bajo coste y luego exportarlos para pagar los bienes intensivos en mano de obra importados de otras naciones debería ser un hecho (Verter, 2015).



## 1.2.2. Teorías modernas del Comercio Internacional vinculadas a la agricultura

### *1.2.2.1. Country Similar Theory*

En 1961, el economista sueco llamado Staffan Burenstam Linder propuso esta hipótesis para analizar las pautas del comercio internacional; siguiendo la hipótesis de Leontief, Linder realizó una investigación empírica (Verter, 2015). Según sus conclusiones, más que disparidades en la oferta de factores de producción, como se espera según las proporciones de los factores resaltados por Heckscher – Ohlin, sus resultados implican que los productos suelen venderse en función de patrones de demanda comparables en las distintas naciones (McPherson et al., 2001; Verter, 2015).

Linder sugirió que las naciones que tienen deseos comparables de productos pueden construir muy fácilmente sectores vinculados entre sí; en consecuencia, entablarán relaciones comerciales entre sí, pero con productos distintos (Verter, 2015). Linder parte de la base de que los clientes que residen en naciones con niveles comparables de riqueza y desarrollo per cápita tienen probablemente preferencias comparables y pueden comprar proporcionalmente artículos de la misma calidad; en consecuencia, es probable que esas naciones realicen la misma cantidad y calidad de comercio y consumo de bienes y servicios (McPherson et al., 2001). Utilizando las metodologías de Linder, la mayoría de las investigaciones econométricas han descubierto una correlación positiva entre la proporción de la economía de un país que se dedica al comercio intraindustrial y el nivel medio de renta per cápita del país (Bergstrand, 1990; McPherson et al., 2001; Verter, 2015).

### *1.2.2.2. Teoría del ciclo de vida del producto internacional*

Raymond Vernon (1966) propuso el modelo como una solución al problema causado por la incapacidad del modelo Heckscher – Ohlin para predecir el comportamiento del mercado; él ofreció un modelo compuesto por cinco etapas del ciclo de vida del producto cuando se le

pidió que describiera la estructura del comercio internacional. Para empezar, la introducción de nuevos productos en el mercado incrementa el comercio, que suele producirse entre naciones o zonas comparables entre sí; en segundo lugar, la etapa de crecimiento es cuando se produce la competencia, el aumento de la intensidad del capital y el incremento de las exportaciones y la transferencia de tecnología de la nación que innovó a otros países, así como el aumento de la inversión extranjera; en tercer lugar, la maduración se traduce en una disminución de la cantidad de exportaciones procedentes del país que inventa.; la cuarta etapa es la de saturación, que se produce en un momento en el que las ventas o la distribución de los productos han alcanzado la posición máxima; la quinta y última etapa es la fase de caída al aumentar la intensidad de la producción en otros países. Se distingue por una concentración de la producción en las naciones emergentes, lo que hace que el país innovador se transforme en un importador neto de algunos artículos de los que había sido pionero en el mercado (Verter, 2015; Katsioloudes & Hadjidakis, 2017).

En las primeras fases del ciclo de vida de un producto, la mano de obra que se asocia a sus cualidades procede de los países que producen esas propiedades (Vernon, 1966, como se cita en Hadjidakis , 2017). Cuando un producto se hace conocido y popular en el mercado de un país, la fabricación de ese producto puede trasladarse a otro país. En pocas palabras, la tesis de Vernon se centraba en la idea de que las empresas fabricarían un bien en primer lugar para satisfacer la demanda local antes de exportarlo a otras naciones. Es posible que la empresa desplace tanto sus instalaciones de fabricación como sus puntos de venta a largo plazo. Siempre existe la posibilidad de que la ventaja comparativa de un determinado producto se desplace de una nación a otra (Verter, 2015; Katsioloudes & Hadjidakis, 2017).

### 1.2.2.3. La “New Trade Theory”

Paul Krugman (1979) y otros economistas fueron los principales defensores de la Nueva Teoría del Comercio (*NTT* por sus siglas en inglés); como señalaron estas personas, se requieren modelos de ventaja comparativa o de dotación de factores; sin embargo, estas ideas no lograron explicar eficazmente las pautas del comercio internacional (Ahmed, 2012; Verter, 2015). Krugman subrayó que los motores aparentes de las pautas comerciales eran los rendimientos crecientes de la escala y los efectos de red que tienen lugar en las empresas importantes (Verter, 2015). Estas empresas son denominadas por Krugman como *first movers* del mercado, y reciben los beneficios de ser los primeros, la mayoría de los cuales se basan en las ventajas de los costes (Jomo & Von Arnim, 2008; Verter, 2015). Krugman sostiene que, si existen enormes economías de escala y rendimientos crecientes de la especialización en un sector, es posible que la demanda de productos y servicios a escala mundial sólo pueda soportar una o un número menor de empresas; por lo tanto, a largo plazo, para que las empresas puedan incorporarse a un mercado y seguir siendo competitivas en una determinada nación, pueden necesitar subvenciones y otras formas de apoyo por parte del gobierno (Jomo & Von Arnim, 2008; Ahmed, 2012).

El modelo de Krugman tiene una serie de supuestos, uno de los cuales es que hay dos naciones que son iguales entre sí en cuanto a sus gustos y tecnologías; sólo hay un elemento de producción que no puede ser objeto de comercio, y es el trabajo; además, la dotación de cada nación es la misma; hay una cantidad importante de competencia en el mercado, y los artículos que ofrecen son de una amplia gama (es decir, diferenciación de productos); cada empresa tiene su propia selección de productos y opera como un monopolio; se puede argumentar que los países que fueron los primeros en entrar en este tipo de industrias tuvieron una ventaja conocida como ventaja de ser el primero, dado que los gustos de los

clientes son universales y consistentes en todas las naciones; y es posible que el precio disminuya con el paso del tiempo a medida que más empresas entran en el mercado (Verter, 2015; Katsioloudes & Hadjidakis, 2017).

Según la teoría de Krugman, la competencia intraindustrial tiene lugar siempre que hay un intercambio de múltiples tipos de los mismos productos diferenciados; el autor también menciona que los beneficios del comercio se derivan del aumento del número de tipos diferentes de productos que se ofrecen a los clientes: cuando hay una mayor producción de cada tipo, se produce el correspondiente aumento de la renta real, ya que los precios bajan al haber un mayor tamaño de mercado y más competencia (Verter, 2015; Katsioloudes & Hadjidakis, 2017).. Según Krugman, la ventaja comparativa no depende exclusivamente de las diferencias en la dotación de factores, sino que depende de las economías de escala y los efectos de red que se producen en las industrias que se consideran significativas (Jomo & Von Arnim, 2008; Ahmed, 2012).

#### *1.2.2.4. Teoría de la ventaja competitiva nacional de Porter*

Este modelo generalmente conocido como el “*Diamante de Porter*” fue acuñado por Michael Porter (1990). En este los patrones comerciales están determinados por la cantidad de empresas y las circunstancias económicas de las naciones que comercial, y el único criterio significativo de competitividad a nivel nacional es la productividad nacional (Verter, 2015). Además, Porter postuló cuatro claves para la ventaja competitiva de un Estado en contraste con las demás naciones: las condiciones de los factores; las circunstancias de la demanda; las industrias vinculadas y de apoyo; y la estrategia, la estructura y la rivalidad de las empresas (Alemu, 2010; Katsioloudes & Hadjidakis, 2017). Las condiciones de los factores son los determinantes básicos (es decir, los recursos naturales, la ubicación, la vegetación y las condiciones climáticas y la tierra fértil para la producción agrícola) y los

avanzados (es decir, la comunicación, los trabajadores cualificados, la desregulación de los mercados, la investigación y el desarrollo) disponibles en una economía determinada, las circunstancias de deseo son la cantidad de demanda de los clientes de los productos y servicios producidos en una determinada economía (Porter, 1990, como se cita en Verter, 2015) .Las industrias relacionadas y de apoyo están determinadas por la cantidad de inversiones en medios de producción avanzados y el desbordamiento de los sectores comparables que conducen a la competitividad local y global de las industrias (MBA Knowledge Base, s.f.).

En cuanto a la estrategia empresarial, la estructura y la rivalidad, Porter sostiene que son los factores de una nación que explican cómo se fundan, gestionan, estructuran y controlan las empresas, y que definen los aspectos de las contiendas nacionales (Verter, 2015). Porter afirma que los rivales locales y el deseo de obtener una ventaja competitiva dentro de un país pueden impulsar a las organizaciones o empresas con fundamento para alcanzar dicha ventaja competitiva en un ámbito internacional más amplio (MBA Knowledge Base, s.f.; Verter, 2015). Porter también hizo hincapié en que las naciones deberían exportar artículos de aquellas empresas en las que los cuatro pilares del diamante son favorables, y a la inversa, en aquella región que no son ventajosos, los países deberían importar (Verter, 2015). El Gobierno tiene un papel fundamental a la hora de garantizar que las empresas mantengan un alto nivel de producción, prestación de servicios y una sana competencia entre ellas (Grant, 1991). Es probable que las naciones se beneficien del comercio, aunque no varíen sus dotaciones de factores (Verter, 2015).

#### *1.2.2.5. La Nueva “New Trade Theory”*

La “*New Trade Theory*” tiene un defecto importante, ya que se basa en la premisa de una empresa representativa de Krugman que contrasta directamente con la situación real sobre

el terreno: en la mayoría de los casos, las empresas son más diversas que similares; esto significa que las empresas tienen niveles de productividad diferentes (Verter, 2015; Ranjan & Raychaudhuri, 2016). El análisis de Melitz (2003) sobre las consecuencias de la heterogeneidad empresarial para el comercio internacional está reconocido como un trabajo pionero en este ámbito. El trabajo realizado por Melitz fue también el responsable de sentar las bases de la “*New- New Trade Theory*” (Ranjan & Raychaudhuri, 2016). Aunque, por término medio, ninguna empresa de un determinado sector sea lo suficientemente productiva para exportar debido a la dispersión de sus productividades (Verter, 2015).

Melitz (2003) subraya que la heterogeneidad de las empresas es una fuente adicional de ventaja comparativa. La comparación de las productividades de las empresas es una fuente de ventaja comparativa. A pesar de ello, puede haber empresas que sean capaces de fabricar cantidades suficientes para ser exportadas. Esta constatación es esencial porque ofrece una explicación de por qué las naciones exportan o importan bienes en industrias en las que parecen tener una ventaja o desventaja comparativa (Verter, 2015; Ranjan & Raychaudhuri, 2016). Otro aspecto clave del modelo de Melitz (2003) es que el libre comercio no sólo conduce a reasignaciones de recursos dentro de los sectores, sino también entre ellos; los recursos se reasignan principalmente de las empresas con menor nivel de productividad a las empresas con mayor nivel de productividad (Verter, 2015; Ranjan & Raychaudhuri, 2016). Si esto no cambia, muchas empresas más pequeñas y exportadoras se verán obligadas a cerrar sus puertas, ya que no podrán competir con éxito con las empresas más grandes del mercado, al menos a corto plazo (Verter, 2015). Por otra parte, Melitz (2004) insiste mucho en el hecho de que las salvaguardias internas de las industrias nacientes obstaculizarán la productividad (Verter, 2015; Ranjan & Raychaudhuri, 2016).

### 1.2.3. Aspectos teóricos vinculados a la predicción

La información presente y anterior puede utilizarse para predecir el futuro (Fotios et al., 2020). Está muy extendida la idea de que los datos del pasado pueden utilizarse para encontrar patrones que sirvan para predecir valores futuros, especialmente en el caso de las series temporales; sin embargo, no se supone que los valores futuros puedan predecirse con exactitud (Ashoori & Mohammadi, 2011). Para una previsión de una única serie temporal en el futuro, existen varias posibilidades, como un valor esperado (también conocido como previsión puntual), un intervalo de predicción, un percentil y toda la distribución de la predicción (Fotios et al., 2020).

Un método de predictivo puede dar lugar a diferentes resultados (Miller, 2015). Como método asociado al *forecasting*, las series temporales pueden desempeñar sólo un pequeño papel en la predicción de eventos, como el fallo de un equipo (Leukel et al., 2021). Lo ideal es utilizar los procesos de previsión cuando se refieren a un problema que hay que abordar; una vez que se han captado las principales características de la situación, se puede formular una teoría; las conclusiones teóricas pueden entonces conducir a una mejor práctica (Fotios et al., 2020; Blanco, 2014).

Establecer predicciones puede ser descrito como un conjunto establecido de procesos que arrojan valores determinados para períodos futuros (Fotios et al., 2020 ; Kuhn & Johnson, 2013). Existen distintas opciones de modelado, pero no todos arrojan las mismas predicciones puntuales; además de proporcionar la capacidad de generar predicciones puntuales, un modelo también puede utilizarse para generar intervalos de predicción y distribuciones de predicción completas (Fotios et al., 2020). En los modelos estadísticos se hacen varias suposiciones sobre el proceso y las distribuciones de probabilidad que lo

acompañan. A pesar de que un modelo puede ser la base de un enfoque predictivo, no se garantiza que el modelo en sí mismo sea de naturaleza única (Kuhn & Johnson, 2013).

Hay que tener en cuenta la naturaleza de las variables y el papel que desempeñan en el proceso predictivo (Fotios et al., 2020). La previsión univariante es un tipo de modelo en el que se pronostica una serie temporal utilizando sólo los datos de esa serie temporal (Oskolkov, 2021). Por ejemplo, en la regresión de series temporales, se utilizan variables adicionales de la serie temporal para elaborar las predicciones (Frank et al., 2001). Las intervenciones son posibles tanto con la previsión univariante como con la multivariante (por ejemplo, promociones especiales, clima extremo) (Fotios et al., 2020).

Pueden utilizarse estructuras lineales o no lineales en las relaciones entre las variables y otras formas de entrada (por ejemplo, la penetración en el mercado de una nueva tecnología); y cuando no se puede acceder a una forma funcional explícita, se pueden utilizar enfoques como la simulación o las redes neuronales artificiales (Abhyankar & Singla, 2022). El desarrollo de estas conexiones puede beneficiarse en gran medida de las teorías extraídas de diversas disciplinas, como la economía, la epidemiología y la meteorología (Fotios et al., 2020). Además, la previsión multivariante puede referirse a la práctica de predecir numerosas variables simultáneamente (por ejemplo, los modelos econométricos) (Fuleky, 2020).

Los datos de las series temporales pueden observarse de diversas maneras, lo que podría restringir o dictar el enfoque de previsión utilizado; de hecho, cuando se utilizan enfoques predictivos, puede que no se disponga de ninguna observación histórica para el objeto de interés pertinente (por ejemplo, el tiempo que se tarda en completar la construcción de un nuevo aeropuerto) (Frank et al., 2001). Como los datos son tan complejos, es posible que sea necesario un nuevo enfoque. Cada minuto, cada hora, cada semana, cada mes y cada año son



posibilidades para la frecuencia de las observaciones; así, por ejemplo, la industria eléctrica necesita prever las cargas de la demanda a intervalos de una hora, así como la demanda a largo plazo para diez o más años adelante (Fotios et al., 2020).

Los datos pueden abarcar desde una única serie temporal significativa hasta miles de millones de series temporales (Brockwell & Davis, 2016). Hay numerosos factores que hay que tener en cuenta al realizar un análisis económico, y todos ellos interactúan entre sí. Es probable que las series temporales de las empresas tengan una jerarquía de series temporales, ya que es probable que sean significativas en muchos niveles distintos (como la unidad de tenencia de acciones, los componentes comunes o el contenedor de tamaño común). Las series temporales serán intermitentes si alguno o todos los valores son 0 y el número de formas de recoger los datos es casi infinito (Hung, 2019).

#### 1.2.4. Modelos predictivos

El proceso de desarrollar una herramienta o modelo matemático que genera una predicción precisa. La creciente presencia de modelos predictivos es un hecho que coincide con la cada vez mayor disponibilidad de data. Se pueden encontrar ejemplos de inteligencia artificial en todas partes: la máquina global de Google usa inteligencia artificial para interpretar crípticos y consultas humanas (Ahmed et al., 2017). Las compañías de tarjetas de crédito lo utilizan para rastrear el fraude. Netflix lo usa para recomendar películas a los suscriptores (Kuhn & Johnson, 2013). Y el sistema financiero lo usa para manejar miles de millones de operaciones (con solo la fusión ocasional). Ejemplos de los tipos de preguntas que uno quisiera predecir son: ¿cuántas copias venderá este libro?; ¿por cuánto se venderá mi casa en el mercado actual?; ¿tiene un paciente una enfermedad específica?; según las elecciones anteriores, ¿qué películas interesarán a este espectador?; ¿es un correo electrónico no deseado?; ¿responderá este paciente a esta terapia? (Kuhn & Johnson, 2013)

Las compañías de seguros, como otro ejemplo, deben predecir los riesgos de los posibles titulares de pólizas de automóviles, salud y vida. Esta información se utiliza para determinar si un individuo recibirá una póliza y, de ser así, con qué prima (Oldon & Wu, 2020). Al igual que las compañías de seguros, los gobiernos también buscan predecir riesgos, pero para propósito de proteger a sus ciudadanos (Kamel et al., 2010). Ejemplos recientes de modelos predictivos gubernamentales incluyen modelos biométricos para identificar sospechosos de terrorismo, modelos de detección de fraude y modelos de disturbios y agitación (Kamel et al., 2010). Los modelos predictivos ahora impregnan la existencia de los consumidores (Kuhn & Johnson, 2013).

Si bien los modelos predictivos pueden servir como guía hacia productos más satisfactorios, mejores tratamientos médicos e inversiones más rentables, también generan regularmente predicciones inexactas y proporcionar las respuestas incorrectas (Ashoori & Mohammadi, 2011). Por ejemplo, la mayoría de los usuarios ha recibido un correo electrónico importante debido a un modelo predictivo (filtro de correo electrónico-*antispam*) que identificó incorrectamente el mensaje como spam (Hofmann & Chisholm, 2015). Del mismo modo, los modelos predictivos (también conocidos como modelos de diagnóstico médico) diagnostican erróneamente enfermedades (Reece & Danforth, 2016). Algunos modelos orientados al *stock market* (también conocidos como algoritmos financieros), conducen a que las personas compren y vendan acciones por error y predican ganancias cuando, en realidad, encuentran pérdidas (Stibel, 2009).

Este último ejemplo de modelos predictivos que salieron mal afectó a muchos inversionistas en 2010 (Makridakis et al., 2010). Aquellos que siguen el mercado de valores probablemente estén familiarizados con la "caída repentina" del 6 de mayo de 2010, en la que el mercado perdió rápidamente más de 600 puntos, luego inmediatamente recuperó esos puntos.

Después de meses de investigación, la Comisión de Comercio de Futuros de Productos Básicos y la Comisión de Bolsa y Valores identificaron un modelo algorítmico erróneo como la causa del colapso (Comisión de Comercio de Futuros de Productos Básicos y Comisión de Bolsa y Valores de los EE. UU. 2010) (Kuhn & Johnson, 2013). Partiendo en parte del desplome repentino y otras fallas de los modelos predictivos, hay quienes plantean que los mismos fallan regularmente porque no tienen en cuenta variables complejas como el comportamiento humano. De hecho, las habilidades para predecir o tomar decisiones están limitadas por nuestro conocimiento presente y pasado y se ven afectadas por factores que no han sido considerados (Rosas-Romero & Medina-Ochoa, 2019). Estas realidades son límites de cualquier modelo, sin embargo, estas realidades no deberían impedir buscar mejorar procesos y construir mejores modelos (Liu, 2020).

Hay una serie de razones comunes por las que los modelos predictivos fallan. Las causas comunes incluyen (1) pre procesamiento inadecuado de los datos, (2) validación inadecuada del modelo, (3) extrapolación injustificada (por ejemplo, aplicación del modelo a datos que residen en un espacio que el modelo nunca ha visto), o, lo más importante, (4) ajuste excesivo (Kotu y Deshpande, 2014). Quienes tienen a su cargo el modelado predictivo a menudo solo exploran relativamente pocos modelos cuando buscan relaciones predictivas, esto generalmente se debe a la preferencia de los modeladores, el conocimiento o la experiencia en solo unos pocos modelos o la falta de software disponible que les permita explorar una amplia gama de técnicas (Kotu y Deshpande, 2014; Wendler & Gröttrup, 2016).

#### 1.2.5. Evolución de la Industria 1.0 a la 4.0.

Utilizando la energía del vapor y mecanizando la industria, la Primera Revolución Industrial comenzó en el siglo XVIII (Stearns, 2012). En el mismo tiempo que las ruecas ordinarias producían hilos, la versión mecanizada producía ocho veces más y ya se conocía la energía

del vapor. (Desoutter Industrial Tools, s.f.). Utilizar este último para usos industriales fue la innovación más importante para mejorar la producción humana de la época (Desoutter Industrial Tools, s.f.). En lugar de utilizar la fuerza para impulsar los telares, se podían utilizar máquinas de vapor; gracias a innovaciones como el barco de vapor y (unos 100 años después) la locomotora de vapor, las personas y las mercancías pudieron recorrer grandes distancias en menos tiempo (Desoutter Industrial Tools, s.f.; Stearns, 2012).

La segunda revolución industrial tuvo inicio en el siglo XIX, con el descubrimiento de la electricidad y la implementación de la producción en cadena. Henry Ford (1863-1947) adoptó el concepto de producción en masa de un matadero en Chicago, en el que los cerdos eran transportados en cintas y cada carnicero realizaba una tarea específica (Desoutter Industrial Tools, s.f.; Stearns, 2012). Ford aplicó estos principios en la fabricación de automóviles, lo que transformó radicalmente el proceso de producción. Antes, los automóviles eran ensamblados en una sola estación, mientras que actualmente, son fabricados en etapas sobre una cinta transportadora, lo que permitió un proceso más rápido y económico (Desoutter Industrial Tools, s.f.).

En los años setenta del siglo XX se inició la Tercera Revolución Industrial mediante la automatización parcial de los controles programables por memoria y los ordenadores. (Yin, Stecke, & Li, 2018) Desde la aparición de estas tecnologías, ahora es posible automatizar todo un proceso de producción sin necesidad de intervención humana; un ejemplo de ello son los robots que ejecutan secuencias preprogramadas (Desoutter Industrial Tools, s.f.).

Como resultado de la digitalización de la fabricación, la humanidad se encuentra en medio de una dramática revolución en la forma de fabricar productos (Yin et al., 2018). Este cambio es tan contundente que se ha denominado Industria 4.0 para reflejar la cuarta revolución industrial (Stearns, 2012). Desde la primera revolución industrial (la mecanización a través

del agua y la energía del vapor) hasta la producción en masa y las cadenas de montaje de la segunda revolución industrial utilizando la electricidad, la cuarta revolución industrial tomará lo que la tercera revolución industrial comenzó con la adopción de los ordenadores y la automatización y lo mejorará con sistemas inteligentes y autónomos impulsados por los datos y el aprendizaje automático (Desoutter Industrial Tools, s.f ; Yin et al., 2018).

La introducción de los ordenadores en la Industria 3.0 fue disruptiva debido a la introducción de una tecnología totalmente nueva (Desoutter Industrial Tools, s.f.). Con el desarrollo de la Industria 4.0, los ordenadores están interconectados y se comunican entre sí para tomar decisiones sin intervención humana (Yin et al., 2018). La Industria 4.0 y la fábrica inteligente son posibles gracias a la convergencia de los sistemas cibernéticos y el Internet de las cosas (Venery & Capasso, 2018). Las fábricas serán más eficientes, productivas gracias al apoyo de máquinas inteligentes, que toman mejores decisiones a medida que tienen acceso a más datos (Venery & Capasso, 2018; Desoutter Industrial Tools, s.f.). La verdadera fuerza de la Industria 4.0 deriva de la red de estas máquinas que están conectadas digitalmente entre sí y crean y comparten información (Desoutter Industrial Tools, s.f.).

A continuación, se presentan algunos ejemplos de aplicaciones potenciales dentro del contexto 4.0:

- Identificar oportunidades: La Industria 4.0 ofrece a los fabricantes la posibilidad de identificar oportunidades para optimizar sus operaciones de manera rápida y eficiente, ya que las máquinas conectadas recopilan grandes cantidades de datos que informan sobre el mantenimiento, el rendimiento y otros aspectos; además, el análisis de estos datos permite identificar patrones y perspectivas que serían difíciles de detectar para un ser humano en un tiempo razonable (Gilchrist, 2016; Marr, 2018). Por ejemplo, una mina de oro en África encontró un problema con los niveles de

oxígeno durante la lixiviación al analizar los datos de los sensores instalados en sus equipos (Marr, 2018).

- Optimizar la logística y las cadenas de suministro: Cuando se suministra nueva información, una cadena de suministro conectada puede adaptarse y acomodarse (Marr, 2018). Si un elemento meteorológico provoca que se retrase un envío, un sistema conectado puede adaptarse a esta realidad y cambiar las prioridades de fabricación (Marr, 2018; Grzybowska et al., 2020).
- Vehículos y equipos autónomos: Existen astilleros que utilizan grúas y camiones autónomos para optimizar las operaciones de recogida de contenedores en los barcos (Marr, 2018).
- Robots: La robótica, que antes sólo estaba al alcance de las grandes empresas con sus correspondientes presupuestos, es ahora más asequible y accesible para empresas de todos los tamaños (Javaid et al., 2021). Desde la recogida de productos en un almacén hasta su preparación para el envío, los robots autónomos pueden ayudar a los productores de forma rápida y segura. *Amazon.com* utiliza robots para transportar mercancías por sus almacenes, ahorrar gastos y aprovechar mejor el espacio (Marr, 2018).
- Fabricación aditiva (impresión 3D): Esta tecnología ha pasado de la creación de prototipos a la producción real durante los últimos diez años; el desarrollo de la fabricación aditiva aplicado a los metales ha creado numerosas oportunidades de producción (Marr, 2018).
- Internet de los objetos y la nube: El IOT (*Internet of Things*), que se define por los dispositivos conectados a la red, es un componente fundamental de la Industria 4.0 (Yin et al., 2018). Esto no solo ayuda a las operaciones internas, sino que, mediante el uso del entorno de la nube, donde se almacenan los datos, se pueden optimizar los

equipos y las operaciones aprovechando los conocimientos de otras personas que utilizan los mismos equipos (Marr, 2018).

La Industria 4.0 es todavía incipiente y puede que no se tenga una visión completa de su alcance hasta dentro de 30 años; sin embargo, las organizaciones que utilizan tecnologías de la Industria 4.0 reconocen su promesa (Marr, 2018). Estas mismas organizaciones también están luchando con la forma de capacitar a su personal actual para que asuma las nuevas obligaciones laborales relacionadas con la Industria 4.0, así como con la forma de contratar nuevo personal con las capacidades necesarias (Marr, 2018; Grzybowska et al., 2020).

#### 1.2.6. Aspectos teóricos de la *Big Data*

*Big data* es una palabra de moda que se ha disparado en popularidad en los últimos años. La *Big data* es la descripción de una gran cantidad de datos organizados o no organizados que se analizan para tomar una decisión o evaluación informada (Hung, 2019). Los datos se pueden tomar de una gran variedad de fuentes que incluyen historial de navegación, geolocalización, redes sociales, historiales de compra y registros médicos (Marconi & Lehman, 2014). *Big data* consiste en datos complejos que abrumarían el poder de procesamiento de los sistemas de bases de datos simples tradicionales. Hay tres principales características asociadas al *Big data*: volumen, variedad y velocidad (Hung, 2019).

- El Volumen es una característica que se utiliza para describir la gran cantidad de datos que se ha utilizado. Las cantidades de datos suelen oscilar entre gigabytes y yottabytes. Los macrodatos deberían poder manejar cualquier cantidad de datos incluso cuando se espera que crezca exponencialmente (Hung, 2019).
- La Variedad es una característica que se utiliza para describir los diferentes tipos de fuentes que se utilizan como parte de un sistema de análisis de *Big Data* (Boyd & Crawford, 2012). Hay múltiples formatos de almacenamiento de datos que utilizan

los dispositivos informáticos en todo el mundo. Hay datos estructurados como bases de datos, .csv, video, SMS y hojas de Excel. Los datos no estructurados pueden tener la forma de notas escritas a mano. Todos los datos de estas fuentes serían idealmente útiles para un sistema de análisis de Big Data (Blanco, 2014).

- La Velocidad es una característica que se utiliza para describir la rapidez con la que se generan los datos. También se utiliza para describir la velocidad a la que se procesan los datos generados. Con un clic de un botón, un minorista en línea puede ver rápidamente grandes datos sobre cierto cliente. La velocidad también es importante para garantizar que los datos estén actualizados y actualizado en tiempo real, lo que permite que el sistema funcione lo mejor posible (Hung, 2019). La velocidad es esencial, ya que la generación de datos en tiempo real ayuda a las organizaciones a acelerar procesos de operaciones; lo cual puede ahorrar a las organizaciones una gran cantidad de dinero (Marconi & Lehman, 2014).

No es sorprendente que no exista un umbral específico después del cual un conjunto de datos pueda considerarse "grande" (Boyd & Crawford, 2012; Hung, 2019). Muchos académicos han intentado describir las diferencias cualitativas que separan a *Big Data* de las fuentes de datos tradicionales (Oracle España, s.f.). Surgen dos temas: en primer lugar, los macrodatos se recopilan generalmente con fines distintos de la investigación académica y el modelado estadístico. En segundo lugar, generalmente requieren un procesamiento más allá de las capacidades del software estadístico estándar. Por mucho que ninguna de estas afirmaciones sea del todo correcta, hay algún valor en ambos (Hung, 2019).

Los datos administrativos se han utilizado para realizar previsiones mucho antes de la llegada de *Big Data*. Los datos de *Multiple Listing Services*, por ejemplo, han jugado un papel clave en la previsión de la demanda futura de vivienda durante décadas, mucho antes de que



alguien pensara en llamarla "grande"; datos recopilados con fines científicos de *Sloan Digital Sky Survey* constan de más de 175.000 espectros de galaxias, lo que ha llevado a los investigadores a desarrollar nuevas técnicas de gestión de datos (Hung, 2019). Sin embargo, a pesar de numerosas excepciones, muchos de los datos que ahora consideramos "grandes" se recopilan como parte de procesos comerciales o de gobierno regulares y, por lo tanto, como describiremos a continuación, presentan una serie de desafíos técnicos relacionados con la gestión y limpieza de datos (Anderson, 2019; Hung, 2019).

Por otro lado, que los macrodatos son "grandes" debido a los requisitos computacionales asociados con su análisis, es igualmente resonante, pero surge más de los orígenes de la disciplina en la informática que de sus usos en macroeconomía; la mayoría de las definiciones en este sentido se centran en si un conjunto de datos es "difícil de procesar y analizar en un tiempo razonable" o de forma más frívola cuando un conjunto de datos es tan grande que "no se puede utilizar STATA" (Hung, 2019). Y aunque estas definiciones presentan claramente un objetivo en movimiento, reflejan el hecho de que *Big Data* es un esfuerzo inherentemente interdisciplinario que a menudo requiere que los economistas colaboren con los científicos informáticos para estructurar los datos para el análisis (Anderson, 2019; Hung, 2019).

En lugar de centrarse en definiciones precisas, creemos que es más importante distinguir entre los tipos de *Big Data*, específicamente las dimensiones por las que un el conjunto de datos se puede considerar "grande". Se han hecho muchos intentos canónicos para definir estos parámetros, desde "volumen, velocidad y variedad", hasta su forma aumentada más reciente "volumen, velocidad, variedad y veracidad", hasta "alto, gordo y enorme" ( León & Villar, 2018; Hung, 2019). Pero lo más valioso para la predicción es la idea de que los datos de series de tiempo son grandes si son extensos en una o más de las siguientes dimensiones:

la cantidad de tiempo (días, trimestres, años) en que se recopilan los datos ( $T$ ), el número de muestras por unidad de tiempo que se realiza una observación ( $m$ ), 1 y el número de variables que se recogen a esta tasa ( $K$ ) para una matriz  $X$  de dimensiones ( $mT \times K$ ); por lo tanto, los datos de series de tiempo son Grandes si son altos ( $T$  enorme), anchos ( $K$  enorme), densos ( $m$  enorme) o cualquier combinación de ellos (Hung, 2019). El valor de este enfoque es que distingue entre los tipos de datos que son grandes porque se han recopilado durante mucho tiempo (población de EE. UU., por ejemplo), los que son grandes porque se recopilan muy frecuentemente (fluctuaciones de stock) y aquellos que contienen un número sustancial de variables (datos de imágenes de satélite) (Hung, 2019; Nooraeni, et al., 2022)

Desde una perspectiva de pronóstico, esta diferenciación proporciona un lenguaje común para determinar las fortalezas de un conjunto de datos para pronosticar la volatilidad frente a la tendencia tanto a corto como a largo plazo. Como señala Diebold (2016), los datos densos son en gran parte poco informativos para pronosticar tendencias a largo plazo (para las cuales se quieren datos altos), pero pueden ser bastante útiles para la estimación de volatilidad.

#### 1.2.7. Aspectos teóricos del *Machine Learning*

La programación que aprende de los datos se adapta a los cambios y mejora con la experiencia es el objetivo del *machine learning* (Blum, s.f.) . El aprendizaje automático ha crecido en importancia desde que se anticipa que los ordenadores se encargarán de cuestiones más complicadas y se entrelazarán más en nuestra vida cotidiana, además de ser una de las aspiraciones fundacionales de la informática (Burkov, 2019; Blum, s.f.). El proceso de creación de programas informáticos es similar al de dar instrucciones a un niño pequeño, lo cual es un proceso que requiere información muy literal, pero que también resulta ser miles de veces más rápido (Lee, 2019). A pesar de ello, muchas de las cuestiones

que actualmente se desean que resuelvan los ordenadores ya no son tareas que se pueda instruir directamente a un ordenador; la identificación de rostros en imágenes, la conducción autónoma en el desierto, la búsqueda de documentos relevantes en una base de datos (o la eliminación de los irrelevantes, como el correo electrónico basura), la búsqueda de patrones en grandes volúmenes de datos científicos y la optimización de los parámetros internos de los sistemas son ejemplos de estas aplicaciones (Burkov, 2019).

Resulta posible que las personas puedan ser expertos en reconocer individuos en imágenes, pero resulta más complejo instruir a un ordenador sobre cómo hacerlo por nuestra propia iniciativa. Las alternativas para superar estas dificultades parecen ser algoritmos que toman datos de entrenamiento etiquetados (por ejemplo, fotografías etiquetadas por quién aparece en ellas (Guarin & Dusseldorp, 2018), o mensajes de correo electrónico etiquetados por si son o no spam), luego los ordenadores aprenden reglas adecuadas a partir de los datos (Burkov, 2019). También se necesitan sistemas que se adapten a las situaciones cambiantes, que sean fáciles de usar al responder a las demandas de los usuarios individuales y que tengan la capacidad de mejorar el rendimiento (Blum, s.f.).

El aprendizaje de máquina o *Machine Learning*, también conocida como Teoría del Aprendizaje Computacional, trata de comprender los principios subyacentes del aprendizaje como proceso computacional mediante el uso de técnicas automatizadas (Blum, s.f.). Los investigadores de este campo intentan determinar a un nivel matemático preciso qué capacidades e información son fundamentalmente necesarias para que los ordenadores aprendan con éxito distintos tipos de tareas, así como los principios algorítmicos fundamentales que intervienen en el aprendizaje de los ordenadores a partir de los datos y la mejora de su rendimiento mediante la retroalimentación (Rebala et al., 2019). Los objetivos de esta teoría son ayudar a la construcción de mejores sistemas de aprendizaje automatizado

y comprender las preocupaciones básicas del propio proceso de aprendizaje (Burkov, 2019). El *Machine Learning* incorpora componentes tanto de la Teoría de la Computación como de la Teoría de la Estadística (Blum, s.f.).

#### 1.2.8. Las Redes Neuronales Artificiales y el Modelo Perceptrón Multicapa

Las redes neuronales artificiales (RNA) son un conjunto de algoritmos, modelados luego de tomar como inspiración al cerebro humano, que están diseñados para reconocer patrones (Cross et al., 1995). Interpretan los datos a través de una especie de percepción de la máquina, etiquetando o agrupando los datos de entrada en bruto. Los patrones que reconocen son numéricos, contenidos en vectores. Por ello, es necesario convertir los datos del mundo real, sean imágenes, sonido, texto o series de tiempo a patrones numéricos (Brenner, 2018). Las redes neuronales contribuyen a predecir y clasificar (Abhyankar & Singla, 2022). Pueden ser utilizado, también, en la tarea de agrupar datos que no cuenten con una etiqueta en congruencia con las similitudes entre las entradas y clasifican los datos cuando tienen un conjunto de los mismos etiquetados con los cuales se puede realizar un proceso de entrenamiento del *dataset* (Ciaburro & Venkateswaran, 2017). Las redes neuronales también están en la capacidad de extraer características que se alimentan de otros algoritmos para la agrupación y clasificación; en consecuencia, resulta posible pensar en las redes neuronales profundas como componentes de aplicaciones del *machine learning* de mayor complejidad que implica algoritmia para el aprendizaje de refuerzo, clasificación y regresión (Berzal, 2018).

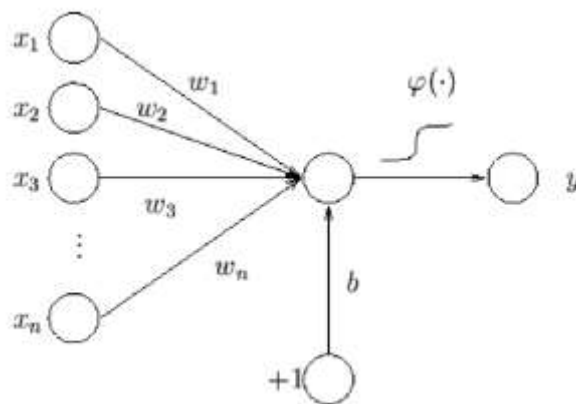
La unidad básica de la red neuronal es el perceptrón, el cual vendría a ser como una neurona del cerebro humano (Gibrán, 2020). La Oficina de Investigación Naval de Estados Unidos (*USONR* por sus siglas en inglés) anunció una idea sorprendente en julio de 1958 cuando se introdujeron tarjetas perforadas en un ordenador de 5 toneladas del tamaño de una

habitación. El ordenador aprendió a identificar entre las tarjetas etiquetadas a la izquierda y a la derecha tras 50 iteraciones de la tarea; según su inventor, Frank Rosenblatt, el "perceptrón" era "la primera máquina capaz de tener un pensamiento original" (Lefkowitz, 2019).

Los perceptrones se forman a partir de los elementos de entrada - *inputs* del modelo – de la forma  $x_1, x_2, \dots, x_n$  afectados por los pesos  $\omega_1, \omega_2, \dots, \omega_n$ . Estos elementos adicionan un *bias* o sesgo, elemento que cumple un rol similar de ajuste al que cumplen las constantes de las regresiones (Lefkowitz, 2019). Todos estos elementos se combinan en una función  $\varphi(\cdot)$  para dar paso al elemento de salida  $y$ , u *output* del perceptrón. En la siguiente figura se muestra la esquematización de un perceptrón:

**Figura 1**

*Esquema de un perceptrón*



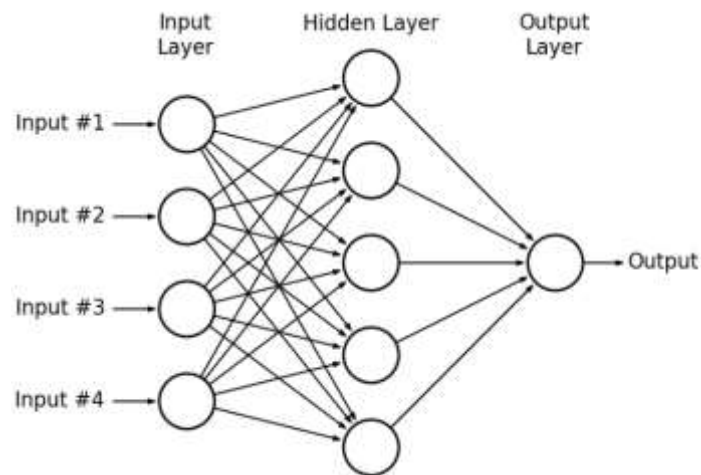
De "Predicting the success of nations in asian games using neural network"por Hematinezhad et al., 2011 ([https://www.researchgate.net/publication/320618036\\_Predicting\\_National\\_Team\\_Rank\\_in\\_Asian\\_Game\\_Using\\_Model\\_Tree](https://www.researchgate.net/publication/320618036_Predicting_National_Team_Rank_in_Asian_Game_Using_Model_Tree)).

El modelo perceptrón multicapa es una arquitectura de las redes neuronales. En ella se agrupan los *inputs* del modelo en la capa de entrada, mientras que el *output* se ubica en la

capa de salida. En la capa intermedia, también denominada como capa oculta, se organizan los perceptrones.

**Figura 2**

*Arquitectura de una red neuronal*



De “Assessment of Artificial Neural Network for bathymetry estimation using High Resolution Satellite imagery in Shallow Lakes: Case Study El Burullus Lake”, por Hassan et al., 2015 ([https://www.researchgate.net/publication/273768094\\_Assessment\\_of\\_Artificial\\_Neural\\_Network\\_for\\_bathymetry\\_estimation\\_using\\_High\\_Resolution\\_Satellite\\_imagery\\_in\\_Shallow\\_Lakes\\_Case\\_Study\\_El\\_Burullus\\_Lake](https://www.researchgate.net/publication/273768094_Assessment_of_Artificial_Neural_Network_for_bathymetry_estimation_using_High_Resolution_Satellite_imagery_in_Shallow_Lakes_Case_Study_El_Burullus_Lake)).

Las RNA existen desde los años 50; el hecho de que las computadoras hayan incrementado su capacidad de procesamiento ha provocado que las RNA recobren su popularidad en el ámbito académico (Van Gerven & Bohte, 2017). Entre las principales ventajas de usar una RNA se tiene:

El desarrollo de modelos de redes neuronales requiere menos entrenamiento estadístico que otros modelos: El desarrollo de RNA que funcionen puede ser realizado en un periodo de tiempo muy corto (es decir, de días a semanas) por aquellos que son nuevos en el ámbito, siempre que tengan acceso a un conjunto de datos aceptable y a un software de redes neuronales (Vu, 1996; Wendler & Gröttrup, 2016). Para desarrollar un modelo de predicción basado en una red neuronal, también puede emplearse cualquier conjunto de datos que pueda

examinarse mediante una regresión logística o lineal estándar (Purkait, 2019). En lo que respecta a las variables de entrada y salida, las redes neuronales pueden entrenarse tanto con variables continuas como categóricas, si bien pueden ser necesarias ciertas modificaciones de los datos en función del software que se utilice (Vu, 1996; IBM, s.f.; Wendler & Gröttrup, 2016).

Las interacciones no lineales complejas entre variables independientes y dependientes pueden ser detectadas por las redes neuronales de con precisión: La red modificará automáticamente los pesos de los enlaces en su estructura si existe un grado considerable de no linealidad entre las variables predictoras y los resultados relacionados en un conjunto de datos de entrenamiento (Vu, 1996). Un ejemplo de transformación no lineal son las variables predictoras de una red neuronal, que se transforman en la capa oculta (Wendler & Gröttrup, 2016).

Las redes neuronales tienen la capacidad de descubrir todas las interacciones imaginables entre las variables de predicción: La capa oculta de una red neuronal le da la capacidad de reconocer las correlaciones o interrelaciones entre todas las variables de entrada (Kavuncuoglu et al.,2018). Cuando existen interacciones críticas que no se recogen explícitamente en un modelo de regresión logística o lineal, se puede prever que un modelo de red neuronal prediga un determinado resultado mejor que un modelo de regresión (İcan & Buğra, 2017). A medida que aumenta el número de variables predictoras en cualquier modelo de regresión logística o lineal, el número de interacciones potenciales se incrementa simultáneamente, y los desarrolladores del modelo pueden encontrar pesado probar todas las interacciones bidireccionales concebibles, y mucho más todas las posibles interacciones de orden superior (Vu, 1996).

Se pueden utilizar varios métodos de entrenamiento alternativos para construir redes neuronales. Sin embargo, aunque el algoritmo de aprendizaje de retropropagación (o *BackPropagation*) es el usado en la presente tesis, hay una serie de otros métodos de entrenamiento de redes neuronales que pueden ser más eficaces, y también se han ideado variaciones del algoritmo de retropropagación; en contraste, las RNA tienen una serie de desventajas que deben ser tomadas en consideración por quienes las toman como opción de modelado (Hu et al., 2018).

Las redes neuronales son una "caja negra" (blackbox) y tienen una capacidad limitada para identificar explícitamente las relaciones causales putativas; los desarrolladores de modelos estadísticos suelen tener como objetivo principal la inferencia causal (Brenner, 2018). Aunque la presencia de una asociación estadística entre una variable predictora y un resultado no demuestra por sí sola la causalidad, los modelos de regresión logística son mejores que las redes neuronales para descubrir posibles relaciones causales (Vu, 1996; Khademi et al., 2015). En las regresiones logística y lineales, el creador del modelo es capaz de determinar qué variables son las que más predicen un resultado, basándose en el tamaño de los coeficientes y las correspondientes razones de probabilidad (Kostoglou et al., 1999; Rajabi Vandechali, et al., 2017). Además, mediante un proceso de selección de variables por pasos, el creador del modelo puede excluir una serie de variables independientes que no están relacionadas con un determinado resultado de interés (Montgomery et al., 2012).

Nuevamente, un modelo de red neuronal es una relativa "caja negra" en comparación con un modelo de regresión logística o lineal (Vu, 1996; Bocco et al., 2010). El desarrollador del modelo pone los datos de entrenamiento para la red y luego deja que se esfuerce por sí misma e identifique qué variables de entrada son las más significativas; no se puede



establecer con facilidad qué variables son las que más contribuyen a un resultado específico; un modelo de red neuronal puede incluir muchas variables predictoras insignificantes que el creador no logra comprender (Brenner, 2018). No existen criterios bien establecidos para interpretar los pesos que afectan a las variables predictoras (Tosun et al., 2016).

Determinar la estructura correcta de la red: No existe un método preciso para establecer la estructura de las redes neuronales artificiales, por lo que no hay forma de predecir la estructura de dichas redes (Vu, 1996). Se requiere experiencia, así como ensayo y error, para desarrollar una estructura de red adecuada (Brenner, 2018).

Se necesitan más recursos informáticos para el modelado de redes neuronales: La generación de modelos de redes neuronales es una operación computacionalmente compleja que requiere el uso de mucho más tiempo de computación (DeTienne & DeTienne, 2017). A inicios de los 90, en el caso de emplear hardware correspondiente ordenador personal estándar y usando el método de *backpropagation*, construir una red neuronal podía llevar desde unas horas hasta muchos días o incluso semanas antes de que una red alcance un estado de aprendizaje óptimo con la menor cantidad de errores (Vu, 1996; Gulli & Pal, 2017). Sin embargo, los modelos de regresión logística y lineales de conjuntos de datos razonablemente grandes pueden probarse en cuestión de segundos en ordenadores convencionales, y pueden evaluarse nuevos modelos a raíz de los resultados de las pruebas. Aunque se han creado variaciones más rápidas del método de retropropagación, todavía pueden necesitar una cantidad significativa de tiempo de entrenamiento (Brenner, 2018).

Las redes neuronales artificiales, que se introdujeron por primera vez en el mundo a mediados del siglo XX, avanzan hoy a un ritmo vertiginoso (Mijwil, 2018). En la investigación se han examinado las ventajas de las redes neuronales artificiales, así como las dificultades que pueden surgir cuando se utilizan. No hay que olvidar que los inconvenientes

de las RNA, que son un campo científico revalorado en parte por el crecimiento de la capacidad de procesamiento de los ordenadores, se están erradicando uno a uno, mientras que sus aspectos positivos se amplían cada día (Brenner, 2018). Esto sugiere que las redes neuronales artificiales se convertirán en un componente cada vez más importante de las vidas de las personas como resultado de este desarrollo (Mijwil, 2018).

#### 1.2.9. Teoría de la Comportamiento Planificado

Los consumidores de productos alimentarios de los países industrializados se enfrentan a un conjunto de opciones diversas en su vida cotidiana como resultado de la proliferación de opciones que tienen a su disposición (Ajzen, 2015). No sólo tienen que decidir entre una desconcertante selección de diferentes marcas de los mismos productos, sino que también tienen que tener en cuenta cuestiones más fundamentales en relación con sus decisiones de compra, incluso pueden verse involucrados factores culturales y étnicos (Fila & Smith, 2016). También las personas tienen que decidir la cantidad de carbohidratos y grasas que incluirán en su dieta; si comprarán productos cultivados de forma ecológica o convencional; si evitarán los pescados grandes con alto contenido en mercurio; si están dispuestos a aceptar los riesgos de los alimentos modificados genéticamente; si darán preferencia a los productos locales; etc. (Ajzen, 2015). No es de extrañar que este tipo de elecciones sean, en teoría, idénticas a las que se hacen en otros ámbitos (Ajzen, 2015; del Castillo et al., 2021).

La idea del comportamiento planificado ofrece una forma alternativa de abordar el problema de la comprensión de los procesos de toma de decisiones de los consumidores (Ajzen, 2015). Esta teoría no hace hincapié en una evaluación global o en la practicidad de un producto o servicio, sino en un tipo concreto de comportamiento del cliente que resulta interesante. En el contexto de la ingesta de alimentos, algunos ejemplos son "comprar platos vegetarianos cuando se sale a cenar" y "comprar yogures bajos en grasa" (Ajzen, 2015). Otros ejemplos

son "consumir alimentos modificados genéticamente en los próximos seis meses" o "consumir comida tipo *fastfood*" (Dunn et al., 2011). El objetivo de la teoría del comportamiento planificado, a menudo conocida como TPB, es ofrecer un marco global para comprender los factores que determinan dichas acciones. (Ajzen, 2015)

La teoría del comportamiento planificado como perspectiva transaccional del comportamiento, presentada por primera vez en 1985 (Ajzen, 1985), es actualmente uno de los modelos sociopsicológicos más conocidos para comprender y prever el comportamiento humano (Dunn, et al., 2011; Fila & Smith, 2016; del Castillo et al., 2021). En pocas palabras, según la TPB, la "intención" de realizar la conducta en cuestión se considera el antecedente inmediato de un comportamiento específico (Ajzen, 2015).

La literatura considera que los siguientes tres tipos de consideraciones o creencias contribuyen a la formación de esta intención.

- El primer tipo de creencias se conoce como "creencias conductuales", y se refiere a las consecuencias buenas o malas que se perciben al ejecutar la actividad, así como a los valores o evaluaciones subjetivas del individuo sobre estos resultados; la acumulación de ideas conductuales que se recuperan fácilmente de la memoria da lugar al desarrollo de una "actitud hacia la conducta" positiva o negativa (Ajzen, 2015; Rush, 2014).
- Un segundo tipo de factor a tener en cuenta es el nivel de motivación del individuo para cumplir con las expectativas y conductas de individuos o grupos de referencia clave, así como las expectativas y conductas percibidas de individuos o grupos de referencia importantes; estos factores se denominan creencias normativas, y las creencias normativas que son fácilmente recuperables en la memoria funcionan

conjuntamente para generar una sensación de presión social o una norma subjetiva en relación con la realización de la conducta (Ajzen, 2015).

- El tercer tipo de elemento a tener en cuenta se denomina creencias de control, y se centra en la forma en que una persona percibe la presencia de circunstancias que pueden tener un efecto sobre su capacidad para llevar a cabo una determinada conducta. Junto con el poder percibido de estos factores para facilitar o interferir en la realización de la conducta, las creencias de control fácilmente accesibles producen un cierto nivel de control conductual percibido (Ajzen, 2015; Rush, 2014)

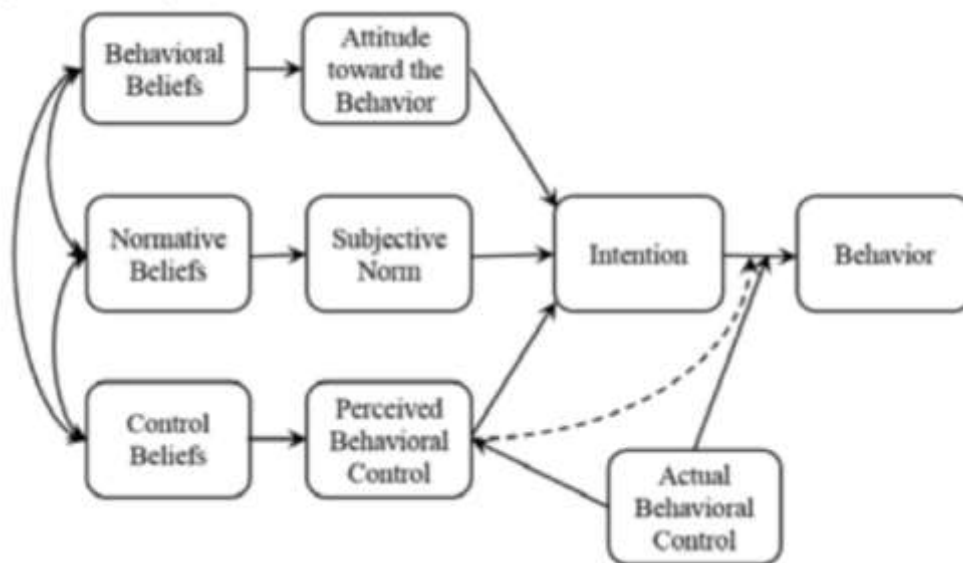
Por regla general, la probabilidad de que una persona forme una intención de realizar la conducta en cuestión aumenta en proporción a la positividad de la actitud hacia la conducta, así como a la norma subjetiva asociada a ella, y al grado en que el individuo percibe que tiene control sobre la situación (Ajzen, 2015; Seyal et al., 2017). En definitiva, se supone que las intenciones de las personas llevarán a la ejecución de la acción en la medida en que sean realmente capaces de hacerlo, o en la medida en que tengan realmente control sobre la conducta. Por lo tanto, es razonable prever que el control real de la conducta moderará la influencia que la intención ejerce sobre el comportamiento (Ajzen, 2015; Rush, 2014)

Sin embargo, en muchas aplicaciones de la teoría del comportamiento planificado, sería difícil o imposible identificar todos los elementos que influyen en el control real sobre la ejecución de la conducta que Ajzen (2015) cuestiona, lo cual se debe a su considerable número. Por ello, los investigadores suelen utilizar la medición del control conductual percibido como un indicador del control real, bajo la premisa de que las percepciones de control reflejan relativamente bien el control real. Esto se debe a que el control percibido está relacionado con el control real (Ajzen, 2015; Rush, 2014).

A continuación, se muestra se muestra el esquema del TPB planteado de Ajzen (2015):

**Figura 3**

*Esquema de la Teoría de la Comportamiento Planificado*



De “Consumer attitudes and behavior: The theory of planned behavior applied to food consumption decisions”, por Ajzen, 2015 (<https://doi.org/10.13128/REA-18003>).

Los bucles de retroalimentación que van de los comportamientos a las creencias no se incluyen en el diagrama (Ajzen, 2015). La realización de una conducta proporciona información sobre sus efectos, las reacciones de los demás, la facilidad o dificultad de realizar la actividad y la facilidad o dificultad de ejecutar la conducta (Ajzen, 2015; Seyal et al., 2017). Este conocimiento hará que, con toda seguridad, cambien algunas de las creencias conductuales, normativas y/o de control de la persona, lo que repercutirá en sus intenciones y su conducta en el futuro (Ajzen, 2015; Rush, 2014). Además de los factores que componen la propia teoría, la misma reconoce la importancia potencial de otras variables. Precisamente, otras variables a considerar incluyen las características demográficas (edad, sexo, raza, religión, educación, ingresos, etc.), los rasgos de personalidad, las actitudes generales y los valores vitales, la inteligencia, las emociones, etc. (Ajzen, 2015; Seyal et al., 2017). En la teoría del comportamiento planificado, estas variables se consideran factores de fondo; se prevé que sólo influyan en las intenciones y el comportamiento de forma

indirecta a través de los efectos que tienen en las creencias de comportamiento, normativas y de control.

Es esencial comprender que la teoría del comportamiento planificado no presupone un comportamiento estrictamente racional por parte del individuo que toma la decisión (Ajzen, 2015; Rush, 2014). Las creencias disponibles que proporcionan la base o la actitud hacia el comportamiento, la norma subjetiva y el control conductual percibido pueden estar mal informadas, reflejar sesgos inconscientes, tendencias paranoicas, ilusiones, motivos egoístas u otros procesos irracionales. (Ajzen, 2015; Rush, 2014). Estas creencias proporcionan la base del comportamiento, la norma subjetiva y el control conductual percibido (Ajzen, 2015).

Lo único que la teoría del comportamiento planificado asume como cierto es que las acciones e intenciones de las personas seguirán sus creencias de forma razonable y coherente, independientemente de cómo se hayan originado esas opiniones (Ajzen, 2015). Sólo en este sentido específico puede describirse un curso de acción como razonado o planificado (Ajzen, 2015; Seyal et al., 2017). Cuando una persona lleva a cabo una acción, no es razonable suponer que esté pasando por un proceso de análisis meticuloso de las creencias que subyacen a esa actividad.; cuanto más se repita una acción, más se convertirá en algo habitual y se llevará a cabo con menos atención consciente (Ajzen, 2015).

#### 1.2.10. El Modelo AIDMA y el AISAS

De acuerdo con Saputra & Kom (2017), fue Roland Hall, en Estados Unidos, quien originalmente hizo campaña por el modelo *AIDMA* en la década de 1920. Según Sugiyama & Andree (2011), Un cliente descubre un producto, un servicio o un marketing y decide adquirirlo, según el modelo *AIDMA*, que describe los siguientes pasos:

*Attention → Interest → Desire → Memory → Action*

El *eWoM* (*electronic Word of Mouth*) se ha ido constituyendo en uno de los medios que terminan llamando la atención del consumidor (Daugherty & Hoffman, 2013). En el caso de los alimentos, el interés de los consumidores se refleja en las distintas comunidades virtuales, ya sea como rumores o mediante contenidos generados por usuarios (Abdul Mumin & Abdulai, 2022). El interés de los consumidores puede reflejarse en el acto de realizar búsquedas acerca de un determinado producto o servicio (Moon et al., 2014). De ahí que para la estimación de ciertos productos o servicios se consideren a las búsquedas de internet como una variable significativa (Salisu et al., 2020).

Otro camino para conseguir la atención del consumidor es la publicidad. De hecho, conseguir la atención del consumidor y provocar un deseo por el producto son considerados los principales objetivos de la publicidad (Saputra & Kom, 2017). Los consumidores que veían un anuncio exitoso recordaban ese deseo e, idealmente, lo recordaban lo suficiente como para actuar en consecuencia la próxima vez que iban a la tienda: comprar ese producto o marca (Sama, 2019).

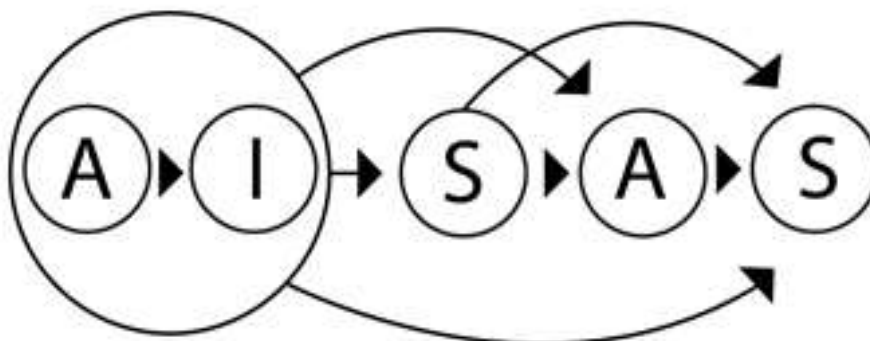
En la publicidad convencional de artículos relativamente básicos en la que el objetivo es persuadir a los consumidores para que elijan su marca entre una variedad de opciones, el *AIDMA* es una estrategia simple pero exitosa: implica que la información suministrada por la empresa a través de la publicidad es todo lo que el cliente necesita, y el objetivo es convencer al consumidor de que recuerde la marca y la promesa de marca en el momento de la compra (Saputra & Kom, 2017). El enfoque *AIDMA* puede ser apropiado para las empresas cuando los clientes tienen poca motivación para saber más sobre los artículos antes de hacer una compra (Sugiyama & Andree, 2011).

Sin embargo, en la era de Internet, en la que cualquiera puede acceder fácilmente a la información. Se ha observado un aumento significativo de lo que puede denominarse "contacto activo con la información", que ocurre cuando los consumidores se fijan en un producto, servicio o anuncio y, voluntariamente, profundizan y comparten con otros la información intrigante que han descubierto (Saputra & Kom, 2017). Además, buscar (es decir, adquirir) información y compartir la misma han surgido como elementos clave en la decisión de compra, también hay que considerar que el flujo de información de las empresas (los emisores) a los clientes (los receptores) (Sugiyama & Andree, 2011).

La compañía Dentsu propuso un nuevo modelo de comportamiento de consumo conocido como AISAS, que se basa en estos cambios en el entorno de la información (*Attention, Interest, Search, Action, Share*). Si comparamos este nuevo modelo de comportamiento de consumo, con el anterior modelo AIDMA, podemos ver que el proceso de transformación psicológica (AIDM) se ha reducido (Sugiyama & Andree, 2011), y Dentsu propone ahora un nuevo modelo de comportamiento de consumo que se basa en estos cambios en el entorno informativo (Saputra & Kom, 2017). Así, el modelo es el siguiente:

**Figura 4**

*Modelo AISAS*



De "The Dentsu Way: Secrets of Cross Switch Marketing from the World's Most Innovative Advertising Agency" por Sugiyama y Andree, 2011.



Fundamentalmente, el modelo *AIDMA* es lineal; refleja un proceso paso a paso que comienza con la "atención" y termina con la "acción". El modelo *AISAS*, en cambio, no pasa necesariamente por cada una de las cinco fases de Atención, Búsqueda y Acción. Es posible saltarse un paso o repetirlo. Un consumidor puede ver un anuncio de televisión de un producto similar al suyo e inmediatamente salir a comprarlo (Atención Interés Acción), o puede quedar tan prendado de la actriz que aparece en el anuncio de televisión que inmediatamente se sienta ante su ordenador y escribe sobre el anuncio en su blog (Atención Interés Acción). Cabe mencionar que la última fase *Share*, puede despertar la fase *Attention* en otras personas (Amornpashara et al.,2015).

#### 1.2.11. El modelo basado en rezagos y la incorporación de la Big Data de Google Trends

En un modelo basado en rezagos, los flujos comerciales mensuales pueden depender de los valores registrados hace un 12 mes o un año; también se conoce en econometría como modelo en base a rezagos (León & Villar, 2018). Una serie comercial puede modelarse como una función que de los rezagos correspondientes a hace un mes y hace un año.

$$x(t) = f(x(t - 12))$$

La demanda tiene relación con el interés del individuo por adquirir el mismo. Así, por ejemplo, antes de comprar un iPhone de cierto modelo, un potencial comprador puede buscar información en Google sobre dicho *smartphone* (y de otros productos similares, por supuesto), similar situación puede darse con autos, viajes turísticos y otros (León & Villar, 2018).

*Google Trends* es una aplicación diseñada por los ingenieros de Google. Consiste en un *dashboard* que permite obtener índices que representa a las búsquedas de un término cualquiera en un intervalo de tiempo determinado y a lo largo de un espacio geográfico

específico. Los datos del *dashboard* pueden ser accedidos desde cualquier computadora con acceso a internet. Los indicadores de Google Trends son una forma simplificada de *Big Data* (León & Villar, 2018).

Las exportaciones de bienes pueden ser modeladas (predichas) utilizando los índices de Google Trends, ya que los mismos representan el interés de los consumidores antes de una potencial compra; se espera que las relaciones entre las exportaciones y las búsquedas sean del tipo directa y lineales (Park, 2015). La idea central es que, al incorporar las búsquedas online a un modelo preestablecido, la precisión de este último se incrementa (Choi & Varian, 2009). Ello aplica también para el modelo autorregresivo. Si se adiciona la variable de búsquedas al modelo que incorpora rezagos de un año, la expresión quedaría de la siguiente manera.

$$x(t) = f(x(t - 12), Google\_Trends(t - 1))$$

¿Por qué debería emplearse un modelo basado en rezagos de un año y de búsquedas web? Ciertamente es posible acceder a las exportaciones semanales e incluso diarias. Sin embargo, ello requiere contar con una base de datos que se actualice con rapidez (como ADEX DATA TRADE o VERITRADE), mientras que el modelo propuesto puede ser implementado recurriendo a bases de datos de acceso libre como TRADEMAP, la cual trabaja con series mensuales Medrano & Rodríguez (2020) evidenciaron que una importante proporción de los exportadores de jengibre son reacios a utilizar sistemas de inteligencia comercial porque los ven como una inversión que solo las grandes empresas pueden permitirse. Un modelo que se pueda construir con bases de datos de libre acceso puede resultar adecuado para quienes tengan limitaciones financieras en cuanto a la adquisición de herramientas de *Business Intelligence* (Anderson, 2019).

### 1.3. Marco Conceptual

### 1.3.1. Beneficios de los modelos predictivos

Muchos fabricantes y distribuidores utilizan métodos de previsión anticuados para desarrollar estrategias comerciales; las soluciones *Big Data* ofrecen a los fabricantes de productos la posibilidad de evolucionar de la previsión tradicional a la planificación basada en los entornos digitales (Hosni & Vulpiani, 2018). Todos estos datos, que actúan como insumos para la analítica avanzada, pueden utilizarse para descubrir tendencias y patrones para hacer más eficientes los programas de fabricación y asegurarse de que las marcas vendan lo máximo posible sin sobreesaturar el mercado (Hassani & Silva, 2015).

La falta de datos oportunos puede dar lugar a desajustes entre la oferta y la demanda. Los datos en tiempo real pueden utilizarse para aumentar la precisión de las previsiones y, a su vez, pueden mejorar los resultados empresariales. La *Big Data* contribuye a la construcción de índices oportunos. Esto se traduce en un aumento de las ventas gracias a la reducción de las roturas de stock, la optimización de los niveles de inventario y la reducción de los costes logísticos, ya que las existencias se entregan sólo cuando se necesitan (Hassani & Silva, 2015).

La previsión implica mucha recopilación de datos, organización de los mismos y coordinación. Las empresas suelen emplear a un equipo de planificadores de la demanda que se encargan de elaborar los pronósticos (Ito, Masuda, Naito, & Takeda, 2021). Pero para hacerlo bien, los planificadores de la demanda necesitan una aportación sustancial de los equipos de ventas y marketing. Además, no es raro que los procesos sean manuales y requieran mucha mano de obra, con lo que se pierde mucho tiempo. Afortunadamente, si se dispone de la tecnología adecuada y los datos necesarios, esto es mucho menos problemático (Hassani & Silva, 2015).

### 1.3.2. Jengibre

El jengibre (*Zingiber officinale*) es una planta herbácea perenne de la familia *Zingiberaceae* que se utiliza como especia, ingrediente aromático culinario y medicamento (Gupta & Sharma, 2014). El nombre genérico de la especia, *Zingiber*, se origina en el griego *zingiberis*, que procede de la palabra sánscrita para la especia, singabera; el nombre latino, *Zingiber*, se traduce como "formado como un cuerno", y se relaciona con las raíces de la planta, que se asemejan a los cuernos de los ciervos en apariencia (Sharma, 2017).

Las culturas antiguas con una larga historia de la medicina, como China e India, tienen una amplia gama de plantas en sus farmacopeas que se utilizan en todo el sudeste asiático; en África y Sudamérica ocurre lo mismo (Ravindran & Babu, 2004). Así, un gran número de personas de todo el mundo obtienen su medicina de las plantas medicinales y aromáticas. Incluso la medicina occidental está haciendo cambios para incluir a la fitoterapia (Fürst & Zündorf, 2015).

El jengibre fue una de las primeras especias orientales que se descubrieron en Europa, ya que los antiguos chinos e hindúes lo cultivaban desde hacía mucho tiempo. En la antigüedad se le atribuían propiedades medicinales (Srinivasan, 2017). Las mujeres embarazadas lo utilizaban a menudo para aliviar las náuseas matutinas y otros síntomas (Sharma, 2017). La especia tiene un sabor algo amargo y suele utilizarse para aromatizar panes, salsas, curry, dulces, encurtidos y cerveza de jengibre. Normalmente se seca y se pulveriza antes de utilizarlo (Sharma, 2017; Srinivasan, 2017).

El rizoma fresco del jengibre se utiliza en la preparación de platos. Es uno de los alimentos más alcalinizantes que existen. Su pH oscila entre 5,6 y 5,9, comparable al de los higos, el hinojo, los puerros, las chirivías y la lechuga romana, entre otros alimentos (Sharma, 2017).

La alcalinidad de los alimentos viene determinada por diversos factores, como el entorno de crecimiento y el método de elaboración.

**Figura 5**

*Zingiber officinale*



De “Medicinal properties of Zingiber officinale Roscoe-A review”, por Kumar y Sharma, 2014 ([https://www.researchgate.net/profile/Anand-Sharma-8/publication/280724449\\_Medicinal\\_properties\\_of\\_Zingiber\\_officinale\\_Roscoe\\_-\\_A\\_Review/links/60ca366b92851ca3aca712ef/Medicinal-properties-of-Zingiber-officinale-Roscoe-A-Review.pdf](https://www.researchgate.net/profile/Anand-Sharma-8/publication/280724449_Medicinal_properties_of_Zingiber_officinale_Roscoe_-_A_Review/links/60ca366b92851ca3aca712ef/Medicinal-properties-of-Zingiber-officinale-Roscoe-A-Review.pdf))

A diferencia de otras plantas, los rizomas de jengibre tienen una forma irregular y pueden ser ramificados o palmeados (Sharma, 2017). La gama de colores de esta especie va del amarillo oscuro al marrón claro, pasando por el marrón pálido (Srinivasan, 2017). Aproximadamente el 2 por ciento del aceite esencial de la especia es *zingibereno*, que es el principal constituyente, y *zingeron*, que es el principio picante, además de otros constituyentes, el aceite se extrae de los rizomas y se utiliza, entre otras cosas, en la industria culinaria y de la perfumería (Stirling, 2004).

El interés por la inclusión y el uso de numerosas hierbas tradicionales o extractos de plantas en la medicina convencional se ha incrementado desde el 2000 (Sharma, 2017). Muchas de estas hierbas y extractos de plantas se basan en lo que se ha utilizado tradicionalmente como parte de los sistemas de medicina tradicional, y hay una cantidad sustancial de pruebas anecdóticas que apoyan su uso y utilidad en este contexto (Gupta & Sharma, 2014). Uno de

los problemas es la escasez de datos clínicos de alta calidad procedentes de ensayos clínicos aleatorios, y otro es que, aunque muchos de estos compuestos se utilizan en mezclas complejas en la práctica, los experimentos de laboratorio tienden a concentrarse en una o unas pocas sustancias químicas específicas dentro de una mezcla, lo que resulta problemático (Sharma, 2017).

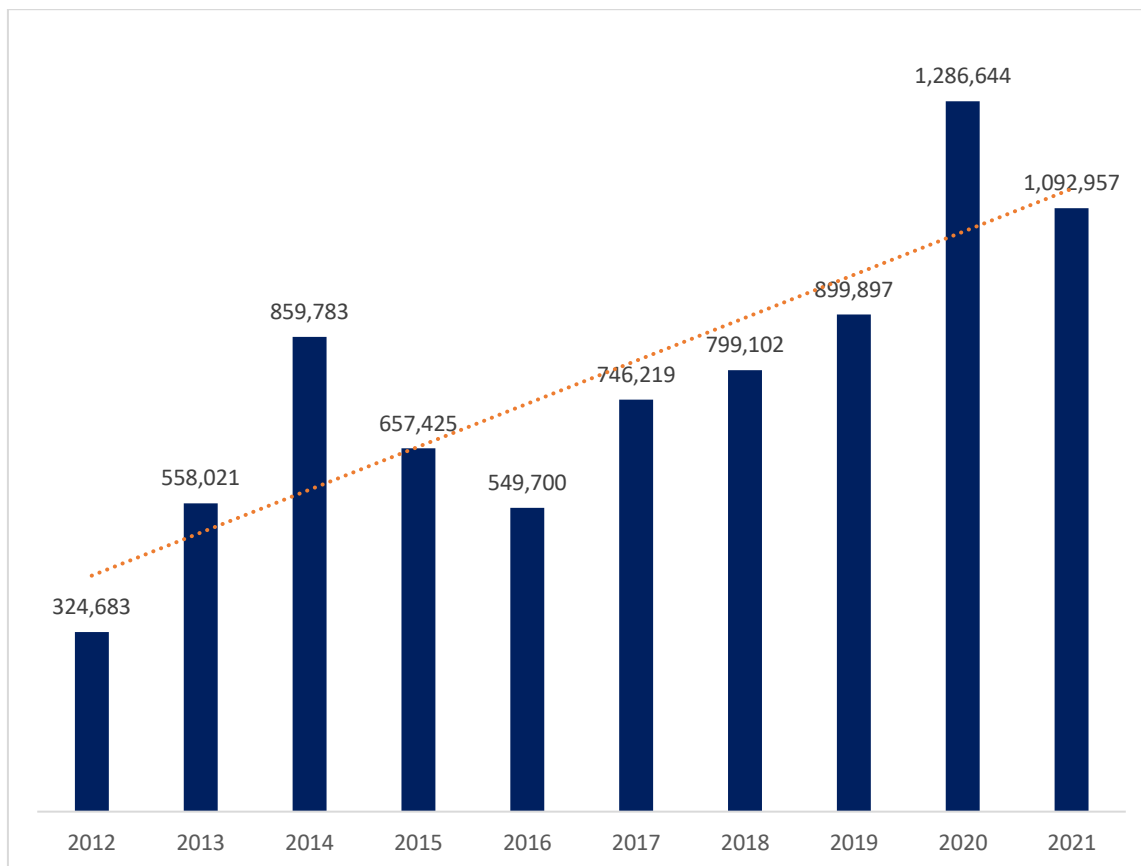
El jengibre incluye una variedad de sustancias químicas bioactivas diferentes, como los gingeroles, los shogaoles y los paradoles, y cuenta con muchas bioactividades, como capacidades antioxidantes, antiinflamatorias y antibacterianas (Sharma, 2017; Mao et al., 2019). Además, el jengibre puede utilizarse como ingrediente de alimentos funcionales o nutraceuticos, y también puede emplearse en el tratamiento y la prevención de diversas enfermedades, como el cáncer, las enfermedades cardiovasculares, la *diabetes mellitus*, la obesidad, las enfermedades neurodegenerativas, las náuseas, la emesis y los trastornos respiratorios (Mao et al., 2019). Se espera que en un futuro no muy lejano se consiga extraer e identificar otras sustancias químicas bioactivas presentes en el jengibre, y que se realicen investigaciones posteriores sobre sus actividades biológicas y los correspondientes mecanismos de acción. (Sharma, 2017; Mao et al., 2019). En particular, los ensayos clínicos sobre el jengibre y sus diferentes componentes bioactivos deben estar bien diseñados para demostrar que el jengibre es eficaz contra estas dolencias en personas reales (Mao et al., 2019).

### 1.3.3. El mercado mundial de las exportaciones de jengibre

Desde el 2012 al 2021, el valor de las exportaciones de jengibre ha presentado una tendencia creciente en cuanto a las exportaciones globales, ciertamente se han dado contracciones, pero existen externalidades que pueden condicionar las exportaciones, como el COVID - 19 (BBC, 2020).

**Figura 6**

*Exportaciones de jengibre fresco a nivel global 2012-2021 (en millones de dólares)*



*Nota.* Información al 3 de noviembre del 2022. Adaptado de “Exportaciones de jengibre fresco a nivel global 2012-2021”, por Trade Map, 2022a (<https://www.trademap.org/Index.aspx?lang=es>).

A continuación se muestra la tabla con los principales exportadores de jengibre fresco a nivel mundial:

**Tabla 6***Principales exportadores mundiales de jengibre fresco 2012-2021*

	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
China	431,531	363,792	249,534	408,243	521,253	440,888	476,775	523,619	497,941	443,726
Costa de Marfil	0	0	0	631	360	949	1,665	2,348	4,071	270,876
India	0	17,159	43,090	26,115	15,977	20,700	16,692	22,861	99,546	164,494
Tailandia	27,030	40,001	26,686	21,086	44,164	83,248	73,525	77,797	60,888	60,618
Países Bajos	31,257	29,346	34,480	33,298	42,362	43,999	48,311	50,679	57,921	58,874
Perú	1,904	3,431	12,291	10,662	14,330	26,409	20,773	22,801	49,657	54,178
Brasil	4,461	5,220	3,692	7,588	8,847	5,639	15,326	22,128	31,712	41,862
Burkina Faso	0	0	0	3	36	105	124	20	49,440	16,125

*Nota.* Información al 3 de noviembre del 2022. Adaptado de “Exportaciones de jengibre fresco a nivel global 2012-2021”, por Trade Map, 2022a (<https://www.trademap.org/Index.aspx?lang=es>).

En la actualidad parece haber un pequeño exceso de jengibre en el mercado mundial, lo que está provocando un descenso de los precios a pesar de una demanda generalmente estable (FreshPlaza.com, 2022). Tras la pandemia, las ventajas atribuidas al jengibre para la salud han contribuido a aumentar la popularidad del producto en el mercado norteamericano (Mayasari et al., 2020). Es posible que estos costes baratos no perduren, ya que Italia ha observado una subida de precios debido a la falta de suministro chino. Actualmente, Perú parece dominar los mercados europeo y norteamericano, en parte debido a la mencionada disminución de la oferta de China, el proveedor tradicional. Este año, la producción de jengibre del país oriental ha disminuido hasta un 30% en tierra y hasta un 50% en rendimiento (FreshPlaza.com, 2022).



En octubre del 2022, las regiones productoras de jengibre del norte de China han entrado en la temporada de cosecha; tras un largo periodo de caída de los precios, los precios del jengibre están aumentando (FreshPlaza.com, 2022). En primer lugar, el clima lluvioso durante el verano y, más recientemente, en octubre, ha dañado las raíces del jengibre y ha acelerado la cosecha, lo que ha provocado una disminución de las cantidades totales recogidas; en segundo lugar, el descenso mundial de la producción de jengibre está haciendo subir los precios (FreshPlaza.com, 2022). La superficie dedicada al cultivo de jengibre en China se redujo un 30% entre la temporada anterior y la actual, según una asociación agrícola china (Zicha, 2021; FreshPlaza.com ,2022). Los cultivadores informan de una pérdida de rendimiento de la cosecha de hasta el 50% como consecuencia del deterioro de las condiciones meteorológicas (FreshPlaza.com ,2022).

Con respecto al mercado estadounidense, actualmente, el suministro de jengibre procedente de Perú es estable y abundante. En el período 2022-2023 debería haber una menor oferta de jengibre orgánico en Europa en comparación con el año anterior debido a los nuevos requisitos para alimentos orgánicos en dicho continente, que también se aplican al jengibre de otros orígenes; por lo tanto, la oferta a otros destinos debería ser ligeramente mayor que el año pasado, según un expedidor citado por FreshPlaza.com (2022).

Brasil y China también exportan ya jengibre a Estados Unidos, aunque principalmente envían jengibre tradicional; el exceso de existencias es un síntoma de un problema mayor en el sector del jengibre; existen desafíos involucran a todos los actores de la cadena comercial del jengibre, desde el agricultor hasta las entidades gubernamentales, los exportadores y los importadores, entre otros, para tener una planificación correcta, responsable y sostenible de la producción del sector, evitando la sobreproducción y mejorando la calidad del jengibre (Zicha, 2021; FreshPlaza.com ,2022). También, para lograr esto, es importante mejorar la

infraestructura de las carreteras, los puertos y las unidades de transporte de tal manera que la logística sea más eficiente y menos costosa (FreshPlaza.com, 2022). Entonces Perú podrá aumentar su competitividad en las exportaciones (Gestión, 2017b; FreshPlaza.com, 2022).

Se prevé un aumento del consumo del kió en Estados Unidos en 2023 (FreshPlaza.com, 2022). Desde la pandemia, la gente es más consciente de las ventajas del jengibre para la salud; como resultado, numerosos artículos, incluyendo bebidas, galletas y aperitivos, utilizan el jengibre como ingrediente (Mayasari et al.,2020; FreshPlaza.com, 2022). Existe un descenso de los precios del jengibre debido a una ligera sobreoferta; esto, unido a los elevados costes logísticos, ha hecho bajar los precios en origen, ya que el cliente final no puede absorber los sobrecostes logísticos" y menos aún con la crisis mundial (FreshPlaza.com, 2022).

A medida que se acercan las fiestas navideñas, la mano de obra tiende a reducirse y se observa un pequeño descenso en la oferta de jengibre. La temporada del jengibre chino y brasileño termina en diciembre, y entonces hay una gran demanda de jengibre peruano; además, el invierno en el hemisferio norte aumenta el consumo de jengibre (Masih et al., 2019).

#### 1.3.4. Variable dependiente: Exportaciones peruanas de jengibre no triturado ni pulverizado

El jengibre o kion proveniente del Perú se destaca por su sabor excepcional, que lo distingue por completo del kion orgánico chino (FreshPlaza.com, 2022). De acuerdo con un artículo en Gestión (2017a), esto se debe en gran parte a que los agricultores peruanos prestan atención y cuidado especial al proceso de maduración del kion, mientras que en China se utilizan sistemas mecanizados para reducir costos y aumentar la eficiencia en la rotación de

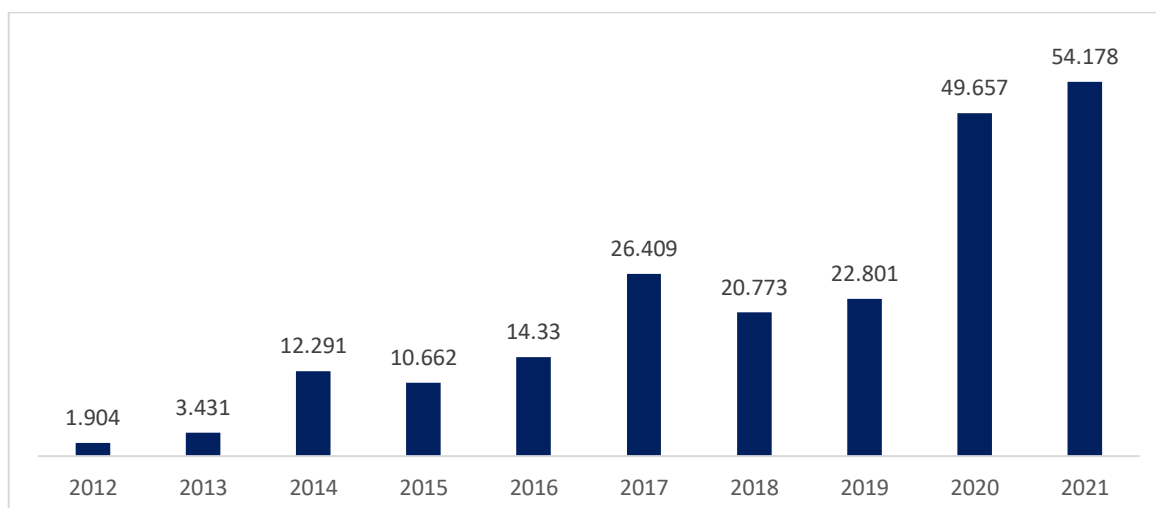
la tierra durante la siembra, cosecha y producción; como resultado, el sabor del jengibre peruano es mucho más concentrado y pronunciado que el chino.

El jengibre es utilizado para crear una variedad de productos con valor agregado, tales como bebidas, té, dulces, snacks, barras energéticas y cápsulas, entre otros; algunos ejemplos de empresas que utilizan jengibre en sus productos son *Simply Gum*, que incluye el sabor natural de jengibre en su goma de mascar y mentas; *Bearded Brothers*, que utiliza jengibre orgánico como ingrediente en su barra energética *Fabulous Ginger Peach*; y *Ginger People*, que ofrece una variedad de productos que contienen jengibre como ingrediente principal, incluyendo dulces, jugos y salsas (Gestión, 2017b)

De acuerdo a Trade Map (2021), el código del jengibre no triturado ni pulverizado fue creado en el 2012. Desde esa fecha, las exportaciones peruanas del producto han tenido una tendencia creciente tal como muestra la siguiente figura:

**Figura 7**

*Exportaciones peruanas de Jengibre del Perú al Resto del mundo 2012-2021 (en toneladas)*

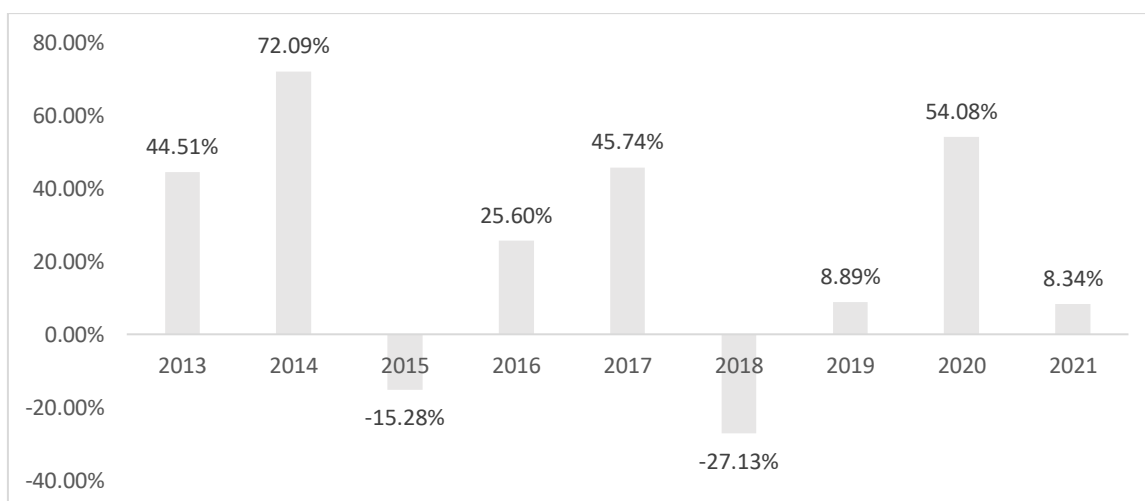


*Nota.* Información al 3 de noviembre del 2022. Adaptado de “Exportaciones peruanas de jengibre fresco 2012-2021”, por Trade Map, 2022b (<https://www.trademap.org/Index.aspx?lang=es>).

En la siguiente figura se muestra la variación porcentual de las exportaciones del jengibre, los mismos que se derivan de la tabla anterior.

**Figura 8**

*Variación porcentual de las exportaciones peruanas de Jengibre fresco del Perú al Resto del mundo 2012-2020 (en toneladas)*



*Nota.* Información al 3 de noviembre del 2022. Adaptado de “Exportaciones peruanas de jengibre fresco 2012-2021”, por Trade Map, 2022b (<https://www.trademap.org/Index.aspx?lang=es>).

En cuanto al modo de transporte de las exportaciones de jengibre, más del 96% se realiza por vía marítima, tal como se muestra en la siguiente tabla:

**Tabla 7**

*Vías de transporte empleados para las exportaciones de jengibre fresco de acuerdo al FOB 2014-2021 (en porcentaje)*

Marítimo	Aéreo	Carretero	Otras Vías	Aduana Postal
96.69%	2.09%	1.18%	0.03%	0.01%

*Nota.* Información al 3 de noviembre del 2022. Adaptado de “Reporte de las exportaciones peruanas de jengibre (2014-2021)”, por Veritrade, 2022 (<https://veritra.upc.elogim.com/es/mis-busquedas>).

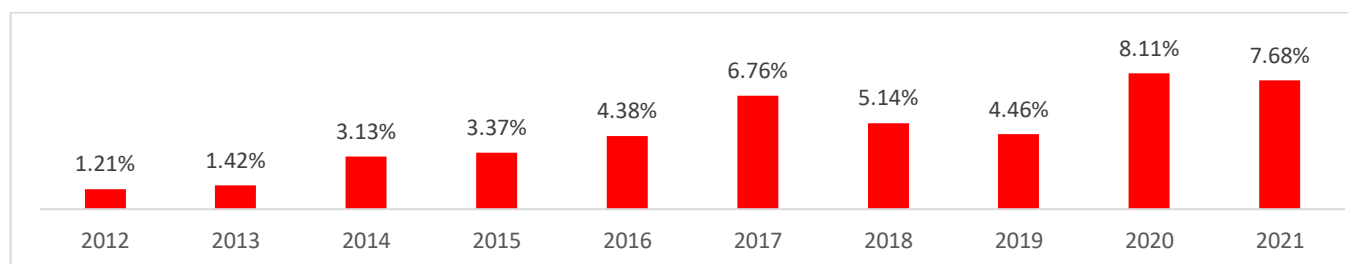
Del 2012 hasta el 2021, los principales países destino de las exportaciones peruanas de jengibre no triturado ni pulverizado han sido:

**Tabla 8***Principales destinos de las exportaciones peruanas de jengibre fresco 2012-2021 (en toneladas)*

	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Estados Unidos de América	953	1,843	5,668	4,518	5,672	9,302	8,851	8,437	18,589	21,345
Países Bajos	333	799	4,331	4,026	5,427	11,389	6,493	7,562	18,054	16,079
Canadá	24	61	185	334	458	990	1,067	1,216	2,374	2,553
Rusia,	0	0	39	15	4	115	0	915	1,562	2,502
España	6	244	419	292	172	338	546	1,603	2,487	2,451
Alemania	21	38	603	216	309	218	478	272	1,046	1,880
Chile	116	98	185	830	1,428	1,717	1,438	1,402	1,843	1,825
Reino Unido	0	6	36	0	0	40	57	47	688	1,558
Italia	0	0	44	0	152	364	319	267	490	714

*Nota.* Información al 3 de noviembre del 2022. Adaptado de “Reporte de las exportaciones peruanas de jengibre (2014-2021)”, por Veritrade, 2022 (<https://veritra.upc.elogim.com/es/mis-busquedas>).

En cuanto a la participación de mercado de las exportaciones peruanas como parte del mercado global se cuenta con la siguiente data:

**Figura 9***Market Share de las exportaciones peruana de jengibre con respecto a las globales 2012-2021 (en porcentaje)*

*Nota.* Información al 3 de noviembre del 2022. Adaptado de “Exportaciones de jengibre fresco a nivel global 2012-2021”, por Trade Map, 2022a (<https://www.trademap.org/Index.aspx?lang=es>).

Por otro lado, en cuanto a los principales exportadores de jengibre fresco se cuenta con los valores acumulados del 2014 al 2021.

**Tabla 9**

*Participación de los principales exportadores de jengibre fresco peruano 2014-2021*

Empresa	% de exportaciones
1 AGRONEGOCIOS LA GRAMA S.A.C.	14.25%
2 KION EXPORT S.A.C.	4.32%
3 AGROINDUSTRIAS JAS E.I.R.L.	3.83%
4 SOBIFRUIITS S.A.C.	3.59%
5 JCH ORGANIC SOCIEDAD ANÓNIMA CERRADA	3.43%
6 LA CAMPIÑA PERU S.A.C.	3.17%
7 ELISUR ORGANIC S.A.C.	3.13%
8 FRUITXCHANGE S.A.C - FC S.A.C	2.68%
9 INTERLOOM S.A.C.	2.48%
10 ASICA FARMS S.A.C.	2.44%
11 ORGANI-K S.A.C.	2.39%
12 RAINFOREST ORGANIC PERU S.A.C.	2.23%
13 NATIVA ORGANICS S.A.C.	2.23%
14 SADE TRADING SOCIEDAD ANONIMA CERRADA	2.11%
15 DOÑA DOROTEA S.A.	1.77%
16 PRAC AGRIBUSINESS FRESH S.A.C.	1.75%
17 EUROFRESH PERU S.A.C	1.71%
18 VALLE INKA PERU S.A.C.	1.65%
19 RTE AMAZON S.A.C.	1.57%
20 FAIRTRASA PERU S.A.	1.52%
21 JCH AGROEXPORTACIONES S.A.C.	1.47%
22 STERLING PERU S.A.C.	1.42%
23 AGRO DEVELOP PERU SOCIEDAD ANONIMA CERRADA	1.34%
24 HAPPY VEG S.A.C.	1.19%
25 TROPIC-X S.A.C.	1.18%
26 NATURAL GREEN S.A.C.	1.13%
27 INKA FRESH S.A.C.	1.05%
28 VANCARD PERU SOCIEDAD ANONIMA CERRADA	1.00%
29 AGRICOLA NUESTRA TIERRA S.A.	0.98%
30 ANAWI EXPORT S.A.C.	0.97%
31 HAMILLTON FARM PERU S.A.C.	0.93%
32 E & J GAVILAN EXPORT SOCIEDAD ANÓNIMA CE	0.92%
33 FOREST SUN SOCIEDAD ANONIMA CERRADA - FOREST SUN S.A.C.	0.90%
34 PRAC AGRIBUSINESS S.A.C.- PRACAB S.A.C.	0.88%
35 PRAC AGRIBUSINESS TRADING S.A.C.	0.77%
36 CORPORACION NATURAL GREEN S.A.C.	0.75%
37 AGROCOSTA PERU S.A.C.	0.75%
38 BEO S.A.C.	0.73%
39 EMPRESA AGRO EXPORT ICA S.A.C.	0.63%
40 PRODUCTOS DE LA SELVA DEL PERU S.A.C.	0.61%

*Nota.* Información al 3 de noviembre del 2022. Adaptado de “Reporte de las exportaciones peruanas de jengibre (2014-2021)”, por Veritrade, 2022 (<https://veritra.upc.elogim.com/es/mis-busquedas>).

### 1.3.5. Variable independiente: Índices de búsquedas *online* obtenidos de *Google Trends*

Internet es una red masiva de redes que conecta millones de ordenadores alrededor de todo el globo, formando una red en la que un ordenador cualquiera puede comunicarse con otro siempre que ambas estén conectadas a *Internet World Wide Web* -o simplemente Web- (Snell, 1995) .Con toda la información que existe, filtrarla manualmente llevaría una eternidad; por lo tanto, los motores de búsqueda fueron creados para hacer esta tarea por uno (Jun et al., 2018; Aaronson et al.,2022).

Un motor de búsqueda es un programa disponible a través de Internet que busca en los documentos y archivos las palabras clave que se le proporciona y devuelve los archivos que contienen esas palabras clave (Jun et al., 2018). Con el uso de algoritmos ingeniosamente escritos, la mayoría de los motores de búsqueda generalmente devuelven archivos y documentos según su importancia; el motor de búsqueda más popular y posiblemente el mejor disponible es Google. Durante el 2021, Google alcanzó un *market share* de 91.56% dentro de los buscadores web a nivel global (StatCounter, 2022).

El motor de búsqueda de *Google* es particularmente especial porque rastrea lo que las personas están buscando y ponen esta información a disposición del público a través de *Google* (Jun et al., 2018). *Google Trends* es un servicio web gratuito que muestra la frecuencia con la que se ingresa un término de búsqueda particular en relación directamente proporcional con el volumen de búsqueda total en diversas regiones del globo y en diferentes idiomas, todo ello en términos relativos (León & Villar, 2018). El acceso a dicha información podría ser útil para cualquier persona que quisiera hacer un estudio relacionado con el historial de búsqueda del público a través de Google (Varian, 2014).

*Google Trends* muestra en tiempo real los índices, valores numéricos que representan a los volúmenes de consultas que los usuarios ingresan en Google; la plataforma también da

acceso a lo que llama "datos en tiempo no real", que se refiere a datos históricos de 2004 hasta 36 horas antes de la actividad de búsqueda (de los que provienen los datos de *Google Trends*) (Jun et al., 2018). Los índices en tiempo no real pueden verse y descargarse en diferentes rangos de tiempo: última hora, últimas 4 horas, último día, últimos 7 días, últimos 30 días, últimos 90 días, últimos 12 meses, últimos 5 años, 2004 hasta el presente, e intervalo de tiempo personalizado (Jun et al., 2018). Sin embargo, la frecuencia de tiempo en el conjunto de datos varía según el rango de tiempo establecido por el usuario: datos por hora de la última hora hasta el último día; datos diarios de la semana pasada hasta 90 días; datos semanales del año pasado hasta 5 años; y datos mensuales para un rango de tiempo más allá de 5 años (Jun et al., 2018; León & Villar, 2018).

El motor de *Google Trends* tiene distintas funciones útiles para la investigación académica. Ofrece la función de autocompletado y depuración de contenido (León & Villar, 2018). Por ejemplo, si se inserta *Tita*, el aplicativo ofrece la posibilidad de elegir Titanio o *Titanic*, lo cual representa a la función de autocompletado; además, es posible depurar el contenido. Puede diferenciarse resultados para *Titanic*, la película, o *Titanic* como accidente propiamente dicho (Sosa & Jeri, 2019).

#### 1.3.6. Ventajas y desventajas de trabajar con indicadores de *Google Trends*

Entre las principales ventajas asociadas al uso de los índices de *Google Trends* en distintas investigaciones se tiene:

- Se puede acceder a *Google Trends* rápidamente y es completamente gratuita (León & Villar, 2018).
- Los datos pueden exportarse y utilizarse para comparar estudios (León & Villar, 2018).



- No hay una encuesta que recoja datos semanalmente en varios idiomas y lugares (Vosen & Schmidt, 2011). Dada su recopilación de datos a largo plazo y a nivel mundial, *Google Trends* es una fuente de información útil (Varian, 2014).
- Dado que la búsqueda de Google es el motor de búsqueda más popular, los datos de *Google Trends* tienen una alta representatividad (Varian, 2014).

En contraposición *Google Trends* posee una serie de desventajas:

- La aplicación sólo proporciona datos relativos; no hay forma de obtener datos absolutos del servicio. En consecuencia, hay que realizar una búsqueda comparable para obtener resultados que puedan ser comparados (León & Villar, 2018).
- Algunos de los fenómenos pueden tener un mayor impacto en aquellos que no tienen acceso a Internet, como las personas sin hogar (Scheffold, 2016).
- La búsqueda no parece tener un contexto claro. Es posible introducir una frase en la búsqueda en Google por una variedad de razones diferentes. El interés no es en absoluto el único factor a tener en cuenta (Blanco, 2014). Ello debido a la falta de información disponible sobre el individuo que buscó la palabra clave (Scheffold, 2016).
- Es necesario verificar la relevancia del material en su contexto. Hay una gran cantidad de palabras que pueden ser utilizadas para transmitir una variedad de significados (Scheffold, 2016; León & Villar, 2018; Aalborg et al., 2019). *Google Trends* ofrece la opción de buscar por categorías, lo que permite filtrar los resultados que no se desean ver en el interés de la búsqueda. Sin embargo, no está claro cómo se crean estas categorías ni qué algoritmo se utiliza para hacer coincidir las consultas en primer lugar (Scheffold, 2016).

- Es posible que ciertos resultados de búsqueda no se muestren porque se ha alcanzado la cuota de búsqueda. Sin embargo, debido al gran número de palabras que se buscan en total, un índice de búsqueda muy pequeño puede no ser intrascendente en términos absolutos a pesar de su reducido tamaño. *Google Trends*, en cambio, no proporciona esta información (Scheffold, 2016).
- Como las búsquedas en la web de Google se eligieron al azar, los valores difieren de una búsqueda a otra, lo que reduce la repetibilidad de los resultados. La solidez estadística de los resultados puede mejorarse repitiendo el proceso de búsqueda muchas veces. Sin embargo, dado que tanto la escala espacial como la temporal expresan magnitudes, esta deficiencia es tolerable en este contexto (Scheffold, 2016).

### 1.3.7. Dimensiones de *Google Trends* relacionadas al consumo *online*

#### 1.3.7.1.. *Busquedas web*

Representan a los búsquedas de internet en su estado más puro, es decir, aquella que es realizada desde el buscador *Google* ingresando un contenido textual (Varian, 2014). La cantidad de búsquedas representa el interés de los usuarios en un tópico específico, aunque no se puede determinar las motivaciones asociadas al hecho de buscar, pero sí se acepta que una búsqueda puede llevar a la eventual compra de un bien y servicio, sobretodo en aquellos bienes que poseen un alto valor. Se espera que los índices de *Google Trends* tengan una significancia relevante para los bienes duraderos y de alto valor (Woo & Owen, 2018). Existen otras herramientas que miden el tráfico web: *SEMRUSH*, por ejemplo, se orienta a *keywords*; mientras que *Baidu Index*, asociado al buscador *Baidu*, se orienta al mercado chino (Vaughan & Chen, 2014). Las búsquedas web propiamente dichas fueron usadas en la investigación de Choi & Varian (2009) como predictores de arribos de viajeros

y venta de autopartes. De hecho, dicha investigación fue la primera en incorporar a los índices de *Google Trends* como partes de un modelo econométrico (Jun et al.,2018).

En la actualidad, la cantidad de búsquedas web es un indicador muy tomado en cuenta por el marketing digital (León & Villar, 2018). La literatura pertinente ha relacionado a las búsquedas web con la demanda de cirugías plásticas, las ventas de *commodities* , las ventas de mascotas , y diversos productos y servicios (Başyiğit, 2021) .

#### *1.3.7.2.Búsquedas en YouTube*

Los equipos de marketing y ventas de las empresas han descubierto un nuevo canal para llegar a nuevos clientes en plataformas como *YouTube*, mientras que la mayoría de las empresas prefieren recurrir a los influencers para hacer llegar su mensaje (Zelenka & Hruška, 2018). En YouTube, la información que contiene un vídeo suele mayor a la que se puede encontrar en las experiencias de las personas contadas en redes sociales (Hiller & Kim, 2014). De ahí que hoy se esté impulsando disciplina como el *Video Analytics* orientado a detectar *insights* en el contenido del vídeo propiamente dicho (Gandzeichuk, 2020).

Otro punto a tener en cuenta es que hay otros creadores de contenido que, aunque no se consideren influencers, pueden, sin embargo, generar cierto interés por un determinado producto o servicio ; ni que decir tiene que los vídeos de YouTube están destinados a quienes tienen tiempo para verlos. De hecho, *YouTube* puede tener impacto en las actitudes de los consumidores hacia la comida orgánica (Vargas & Yamanaka, 2020). Cabe mencionar que los índices relaciones a la cantidad de búsquedas en YouTube han sido correlacionados con las taquillas de películas y las descargas de música (Hiller & Kim, 2014).

### 1.3.7.3. Búsquedas de imágenes

Las imágenes, al contrario de los vídeos, sirven como referencia rápida, permitiendo obtener una primera opinión de un producto o servicio prácticamente en seguida después de verlo (Oliveira & Casais, 2019). Aunque el uso de fotografías en la publicidad existe desde hace mucho tiempo, el éxito de las plataformas de medios sociales como Instagram puede atribuirse al hecho de que una parte importante de la población prefiere obtener información en formato de imagen (Ciriaco & Garayar, 2020).

Con la introducción de las cámaras digitales y las cámaras fotográficas, la capacidad de capturar y compartir fotografías y películas se ha hecho más común (Angulo et al., 2019). Por ejemplo, en el 2018, los comerciantes de *eBay* subieron aproximadamente 20 millones de fotos de productos al día, según las estadísticas de la empresa (Oliveira & Casais, 2019). Cabe mencionar que existen fotógrafos especializados para el rubro de alimentos (Folts et al., 2006).

En sitios como Amazon, Craigslist y eBay, las fotografías juegan un papel vital en el proceso de transacción de los negocios (Oliveira & Casais, 2019). Amazon estandariza las imágenes debido a que es un sitio de comercio por catálogo; en cambio, en eBay, debido a la gran variedad de productos, condiciones del producto y vendedores, hay grandes variaciones entre las imágenes en términos de cantidad y calidad. Amazon también estandariza las imágenes debido a que es un sitio de comercio por catálogo; una gran proporción de las fotografías del inventario son tomadas por vendedores individuales que no tienen habilidades fotográficas profesionales, lo que resulta en grandes variaciones en la calidad de la imagen (Di et al., 2014). Los consumidores pueden asociar el atractivo visual con la calidad de un producto o servicio (Ren et al., 2020).

## 1.4. Viabilidad de la investigación

La data necesaria para la investigación proviene de fuentes secundarias de libre acceso (Trade Map y *Google Trends*). Por tanto, la viabilidad es máxima desde una perspectiva de la disponibilidad de los datos.

# CAPÍTULO II. PLAN DE INVESTIGACIÓN

## 2.1. Situación Problemática

Los sectores agrícolas y alimentarios eficientes y con capacidad de respuesta son fundamentales para la vida humana (Gopinath et al., 2021). En los últimos años, diversas perturbaciones como catástrofes naturales, conflictos comerciales y pandemias han causado estragos en la industria agrícola mundial (Gopinath et al., 2021; The World Bank, 2022). La amplitud de las opciones que comienzan en los campos de cultivo y concluyen en el hogar se ha visto afectadas por esta extraordinaria incertidumbre comercial (Gopinath et al., 2021).

Las mencionadas perturbaciones han producido una cantidad de imprevisibilidad y complejidad sin precedentes durante las últimas décadas en el comercio agrícola (Gopinath et al., 2021; Nooraeni et al., 2022). Para complicar las cosas, está el carácter estático de la mayoría de los modelos comerciales, que suelen hacer análisis estáticos comparativos de los resultados comerciales resultantes de los cambios deterministas de la política comercial (Nooraeni et al., 2022).

¿Cómo hacer frente a escenarios en donde la demanda asociada a los flujos de comercio exterior es cada vez más imprevisible? Ciertamente hay escasez de asesoramiento sobre la modelización teórica de la incertidumbre comercial y sus implicaciones para las preferencias y el comportamiento de productores y consumidores (Gopinath et al., 2021). Para el caso de productos agrícolas bien se puede optar por el uso de modelo más robustos, como aquellos que se cimentan en el *machine learning* (Gopinath et al., 2021; Nooraeni et al., 2022). Las redes neuronales representan una alternativa atractiva para quienes estén dispuestos a optar por buscar alternativas a las relaciones lineales entre variables (Gopinath et al., 2021; Nooraeni et al., 2022).

Los modelos propuestos por Francia y Pacheco (2019) y Alvarez (2020) para explicar las exportaciones de jengibre fresco de origen peruano a Países Bajos y Estados Unidos respectivamente alcanzaron precisiones de 51.3% y 42.7%. Ciertamente ambas propuestas son valiosas para comprender que, en el ámbito académico peruano, todavía existen oportunidades de mejora para aproximarse a una mejor comprensión de los elementos que indican en el flujo de exportaciones del mencionado producto agrícola. Los mencionados antecedentes emplean variables macro que se usan con bastante frecuencia, tal es el caso del precio, el tipo de cambio el nivel de ingreso del mercado de destino, entre otros. Todo ello mediante relaciones lineales propias de un modelo de regresión múltiple. Cabe mencionar que Onomayo et al. (2022), para el contexto de las exportaciones de jengibre nigerianas, lograron un modelo cuya precisión logró un 69.10%.

En la última década, la creciente disponibilidad de grandes datos y los avances tecnológicos en materia de informática han supuesto un reto para las metodologías estadísticas y econométricas tradicionales a la hora de predecir patrones o relaciones económicas complicadas, especialmente en el sector financiero y comercial (Varian, 2014; Fuleky, 2020). Numerosas organizaciones internacionales y agencias gubernamentales pronostican flujos comerciales para orientar la elaboración de políticas a nivel nacional y mundial (Gopinath et al., 2021). Debido a que estas proyecciones se basan en una combinación de análisis basados en la opinión de los expertos, varias fuentes han señalado sus limitaciones, como la precisión de las previsiones, que es inferior al 35 % (Daugherty & Jiang, 2019), y la incapacidad de cuantificar la contribución de los factores económicos subyacentes (Gopinath et al., 2021). Son necesarios grandes volúmenes de datos para evaluar estos acontecimientos atípicos (The World Bank, 2022). No basta, entonces, contar con modelos más robustos, tal como se espera del *machine learning*: es necesario, también, contar con un volumen representativo de datos.

¿De dónde se puede obtener grandes datos que sean representativos para el comercio de alimentos? Mayasari et al. (2020) indicaron que *Google Trends* resulta de utilidad para conocer tendencias en los hábitos alimenticios de alimentos, el jengibre entre ellos. Dicho de otra forma, el investigador planteó que las búsquedas online pueden servir para conocer las tendencias de consumo de dicho producto. No obstante, no llega a evidenciar ello matemáticamente, simplemente describió como las búsquedas *online* acerca del jengibre se han incrementado como resultado de la pandemia. Por otro lado, Woo y Owen (2018) indicaron que si bien es cierto que el consumo privado puede modelarse en función a las búsquedas en internet (*Google Trends*), ello es solo significativo para bienes duraderos (autos, casas, etc.) o de alto valor (computadoras, joyas), pero que no tendrían un aporte significativo para bienes no duraderos (alimentos y similares); los investigadores demostraron ello luego de emplear las regresiones lineales y las búsquedas asociadas a las noticias de cada producto.

Una crítica que se puede realizar al trabajo de Woo y Owen (2018) radica en el hecho de que únicamente se haya abocado al emplear búsquedas relacionadas a contenido noticioso: una persona común y corriente puede realizar una búsqueda web no necesariamente impulsada por una noticia, sino por el mero hecho de querer encontrar un contenido más sencillo (una receta a base de jengibre, por ejemplo). También existen consumidores que visualizan videos en *YouTube* para informarse acerca de productos, servicios, procedimientos; mientras que otros se concentran en buscar imágenes referenciales de aquello que desean adquirir (Feng et al., 2019). Precisamente, las búsquedas en YouTube y acerca de imágenes, junto a las búsquedas comunes en Google, son medibles a escala global con *Google Trends* (León & Villar, 2018).



Por lo mencionado anteriormente, cabe preguntarse si las búsquedas web en Google, las búsquedas en YouTube y las búsquedas de imágenes acerca del mencionado producto alrededor del mundo contribuyen a la predicción del total de las exportaciones peruanas. Todo ello bajo el modelado basado en redes neuronales, un modelo del tipo *machine learning* sugerido por Gopinath et al. (2021).

La línea de investigación involucrada corresponde al ámbito de operaciones y logística, más precisamente a la subcategoría de pronósticos y modelos de demanda, ya que se emplea a las redes neuronales para pronosticar las exportaciones peruanas de jengibre al resto del mundo.

#### 2.1.1. Problema General

El problema de investigación que se desliga de todo lo mencionado anteriormente es:

PG: ¿Es correcto afirmar que los índices globales de búsquedas *online* acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales?

#### 2.1.2. Problemas específicos

Dadas las dimensiones de las búsquedas *online*, obtenidas de *Google Trends* y discutidas en el trabajo de León & Villar (2018), resulta posible plantear los siguientes problemas específicos.

PE1: ¿Es correcto afirmar que los índices globales de búsquedas web acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales?

PE2: ¿Es correcto afirmar que los índices globales de búsquedas en YouTube acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales?

PE3: ¿Es correcto afirmar que los índices globales de búsquedas de imágenes de jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales?

## 2.2. Objetivos

### 2.2.1. Objetivo general

El objetivo asociado a la pregunta de investigación puede formularse de la siguiente manera:

OG: Determinar si es correcto afirmar que índices globales de búsquedas *online* acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.

### 2.2.2. Objetivos específicos

Los objetivos asociados a las preguntas de investigación previamente indicados son los siguientes:

OE1: Determinar si es correcto afirmar que los índices globales de búsquedas web acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales

OE2: Determinar si es correcto afirmar que los índices globales de búsquedas en YouTube acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.

OE3: Determinar si es correcto afirmar que los índices globales de búsquedas de imágenes de jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.

### 2.3. Hipótesis

#### 2.3.1. Hipótesis general

HG: Los índices globales de búsquedas *online* acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.

#### 2.3.2. Hipótesis específicas

H1: Los índices globales de búsquedas web acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales

H2: Los índices globales de búsquedas en YouTube acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.

H3: Los índices globales de búsquedas de imágenes de jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.

## 2.4. Justificación

### 2.4.1. Justificación teórica

En los trabajos de investigación de negocios internacionales ubicables en los principales repositorios de universidades peruanas, no se ha ubicado trabajo alguno que use indicadores *Big Data* de *Google Trends* combinados con algún método *machine learning* aplicada al modelamiento de flujos de comercio exterior. Ciertamente Pilco y Sandoval (2018) emplearon las búsquedas *online* para pronosticar el empleo y la inversión, pero su investigación emplear métodos econométricos. Ello implica que la tesis tiene cierto grado de originalidad en el contexto peruano. Por otro lado, se pretende contrastar la idea de que los volúmenes de búsqueda no tienen un aporte significativo para modelar la demanda de un bien no duradero, tal como afirmaron Woo y Owen (2018).

### 2.4.2. Justificación metodológica

Por otro lado, el empleo de redes neuronales no es habitual en la tesis de negocios desarrolladas en el ámbito peruano (Ramirez & Vilela, 2020). Ello probablemente se deba a que su uso se había restringido a profesionales con formación en ciencias de la computación o ingeniería. Quienes puedan revisar la presente documentación podrán encontrar una visión introductoria al uso de las redes neuronales, método que generalmente no forma parte de los tópicos a tratarse en los cursos que forman parte de las carreras de negocios. Cabe mencionar que el empleo de métodos alternativos asociados al *Machine Learning* es una recomendación de organismos multilaterales (Gopinath et al., 2021). Finalmente, también se hace uso de las

regresiones lineales y, en consecuencia, resulta posible establecer un comparativo entre ambos métodos. Los *journal articles Predicting strength of recycled aggregate concrete using artificial neural network, adaptive neuro-fuzzy inference system and multiple linear regression* (Khademi et al., 2015); *Comparison of Regression and Neural Networks Models to Estimate Solar Radiation* (Bocco et al., 2010) y *Prediction of the antimicrobial activity of walnut (Juglans regia L.) kernel aqueous extracts using artificial neural network and multiple linear regression.* (Kavuncuoglu et al., 2018) son ejemplo de como se discuten ambas opciones de modelado desde distintos ámbitos de la ciencia.

#### 2.4.3. Justificación práctica

Finalmente, el modelo a desarrollar es *per se* un aporte práctico, ya que puede ser empleado por quienes tengan a su cargo la proyección de demanda del sector ante escenarios cada vez más incierto. Con escenarios con cada vez menor grado de certidumbre, resulta más necesario que nunca contar con técnicas alternativas para analizar, modelar y proyectar los flujos comerciales de comercio exterior (Gopinath et al., 2021).

Las decisiones de plantar, mantener el progreso de los cultivos, cosechar y comercializar a corto plazo, así como de invertir en activos agrícolas a medio plazo, se han visto afectadas por graves interrupciones de la oferta (por ejemplo, inundaciones o sequías), una incertidumbre significativa en la demanda (por ejemplo, las compras de soja de China) y un colapso repentino tanto de la oferta de insumos como de la demanda de productos, potenciadas por el COVID-19 (Gopinath et al., 2021).

Existen en el sector de exportación de jengibre ciertas deficiencias en el acceso de herramientas asociadas a la inteligencia comercial - *ADEX Data Trade*, por ejemplo - atribuidas a la percepción de que únicamente aquellas empresas con una participación significativa de mercado pueden acceder a las mismas (Medrano & Rodríguez, 2020). El

modelo propuesto puede construirse en base a fuentes de acceso libre como TRADE MAP y *Google Trends*.

# CAPÍTULO III. METODOLOGÍA DE TRABAJO

## 3.1. Enfoque de la investigación

### 3.1.1. Alcance de la investigación

Dado que es posible alcanzar los objetivos de la investigación empleando métodos numéricos como las redes neuronales, el enfoque a emplear es cuantitativo. Otro rasgo común de este tipo de investigaciones es el hecho de que conlleva una etapa para la recolección de datos que recibirán un tratamiento matemático (Hanneman et al., 2013).

## 3.2. Diseño de investigación

La investigación es no experimental debido a que tanto los datos de las exportaciones de jengibre como los indicadores de *Google Trends* se obtienen de fuentes externas sobre las cuales los investigadores no ejercen control alguno (Locharoenrat, 2017). Por otro lado, la investigación también es correlacional porque, como subproducto, se establecerán relaciones entre las distintas dimensiones de las búsquedas *online* y las exportaciones del producto. En contraste, cabe recalcar que las redes neuronales no establecen correlaciones propiamente dichas entre las variables de entrada y de salida del modelo (Ramirez & Vilela, 2020). Ello ocurre, por ejemplo, en los modelos de regresión lineales y logísticos (Khademi et al., 2015). La investigación trata de llenar las lagunas de la comprensión actual de un tema sobre el que se encuentran vacíos del conocimiento (Bairagi & Munot, 2019). La investigación alcanza, además, un nivel predictivo, pues se pretende pronosticar el valor de una variable (Hernandez & Mendoza, 2018), incluso si se trata una predicción de muy corto plazo. La investigación analiza las variables a través del tiempo por lo que puede ser considerada como longitudinal (León & Villar, 2018)

### 3.2. Operacionalización de variables

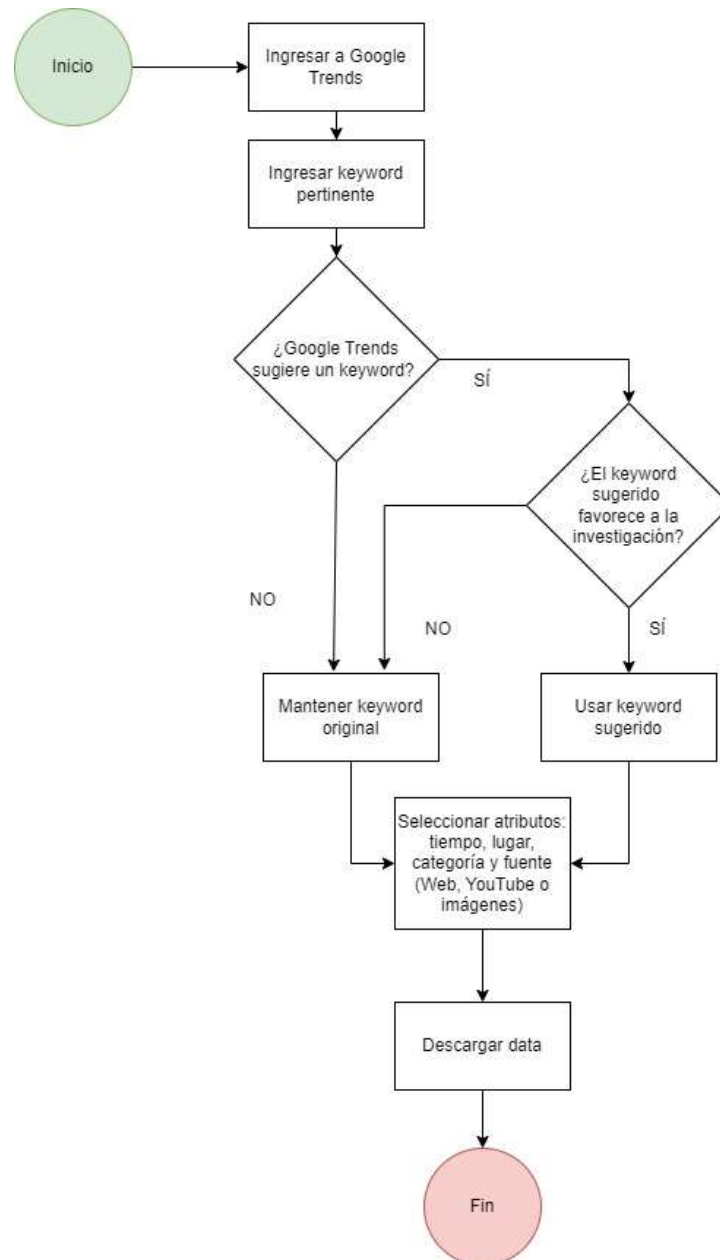
**Tabla 10** Operacionalización de variables

Variable	Dimensión	Etiqueta	Descripción	Operacionalización
Exportaciones de jengibre peruano al resto del mundo.	Peso neto de las exportaciones de jengibre	Expo	Es la variable dependiente del estudio. Representa al valor del peso neto de las exportaciones. Los rezagos correspondientes a un año cumplen el rol de variable independiente y tienen la etiqueta Expo-12	Se obtiene a partir de los indicadores de Adex Data Trade (2022)
Índices de Búsquedas globales online acerca del jengibre ( <i>Zingiber Officinale</i> )	Búsquedas web en el mundo acerca del jengibre durante el mes inmediato anterior a la exportación.	Web	Representa al índice mensual de las búsquedas web acerca del jengibre alrededor del mundo durante el mes inmediato anterior a la exportación. Es una variable cuantitativa e independiente.	Es obtenida a partir de <i>Google Trends</i> . Es necesario con ingresar el <i>keyword Zingiber officinale</i> en la aplicación y elegir la opción de búsquedas web en la categoría consumo alrededor del mundo (ver figura 10 y 11)
	Búsquedas en YouTube en el mundo acerca del jengibre durante el mes inmediato anterior a la exportación.	YouTube	Representa al índice mensual de las búsquedas en YouTube en el mundo acerca del jengibre durante el mes inmediato anterior a la exportación.	Es obtenida a partir de Google Trends. Es necesario con ingresar el <i>keyword Zingiber officinale</i> en la aplicación y elegir la opción de búsquedas en YouTube en la categoría consumo alrededor del mundo (ver figura 10)
	Búsquedas de imágenes en el mundo acerca del jengibre durante el mes inmediato anterior a la exportación.	Imágenes	Representa al índice mensual de las búsquedas de imágenes en el mundo acerca del jengibre durante el mes inmediato anterior a la exportación.	Es obtenida a partir de Google Trends. Es necesario con ingresar el <i>keyword Zingiber officinale</i> en la aplicación y elegir la opción de búsquedas de imágenes en la categoría consumo alrededor del mundo (ver figura 10)



Para el caso del jengibre se optó por usar el *keyword* de nombre científico debido a que fue sugerida por *Google Trends* y favorecía a la investigación al vencer la barrera idiomática. Se pudo también haber elegido como *keyword* a “kion”, pero dicho término es un regionalismo.

**Figura 10** *Flujograma de Operacionalización de variable de búsquedas online*



Adaptado de “Estudio exploratorio del uso de indicadores de Big Data de Google como variable en el modelamiento del flujo mensual de visitantes al Complejo Arqueológico de Kuélap en el periodo 2011-2016”, De León y Villar, 2018 (<http://repositorio.usil.edu.pe/handle/USIL/369>).

Si es que la variable dependiente fuera arribos al Jorge Chávez, el *keyword* adecuado sería Lima y la categoría a elegir sería “vuelos de avión” (Sosa & Jeri, 2019). Para el caso del jengibre se eligió a la categoría “compra”. Si es que la variable dependiente fuera arribos nacionales al complejo arqueológico de Kuélap, “Chachapoyas” sería un *keyword* adecuado, “viajes” la categoría pertinente y el territorio peruano la zona propicia (León & Villar, 2018). Si es que se quiere modelar la colocación de tarjeta de crédito del Banco de Comercio Português (BCP), puede optarse por ingresar el *keyword* “BCP”, Google Trends mostrará sus *keyword* sugeridos “BCP Perú”, “Banca Móvil BCP” y “Banco de Comercio Português”, naturalmente se elegirá la última sugerencia y la categoría pertinente sería “tarjetas de crédito” (Google Trends, 2022).

### 3.3. Universo o Población

En los modelos basados en series de tiempo, las unidades de análisis son cada una de las series que contribuyen a solucionar el problema de investigación (Golovanova & Zubarev, 2021). Para la presente investigación, una serie está compuesta por las exportaciones mensuales de jengibre fresco junto con sus respectivos rezagos. A dichos valores se deben adicionar las búsquedas *online* realizadas acerca del jengibre, ya sea como imagen, a través de *YouTube* o desde una búsqueda web ordinaria, todas ellas correspondientes a un mes previo a la exportación.

### 3.4. Muestra

En el *journal article* titulado *Comparison of correlation analysis techniques for irregularly sampled time series*, Rehfeld et al.(2011) parte de la premisa de que las series de tiempo también son también susceptibles de muestreo. Es necesario recordar que el modelo propuesto se basa en rezagos de un año (ver sección 1.2.11). En consecuencia, se analizan a las series desde enero del 2013. No resulta factible analizar meses anteriores; así, por

ejemplo, no se podría asociar a las exportaciones de jengibre realizadas desde territorio peruano en diciembre del 2012 con el valor correspondiente de diciembre del 2011: las exportaciones de jengibre fresco cuentan con una partida arancelaria propia desde el 2012 (Trade Map, 2022). Las series mensuales analizadas van desde enero del 2013 hasta diciembre del 2021. Ello con el fin de maximizar la cantidad de unidades muestrales para años completos. El número total de series mensuales analizadas fue 108 correspondiente a 9 años.

### 3.5. Instrumentos y/o Herramientas

En la tesis los investigadores no hicieron uso de instrumentos tales como encuestas o guía de pautas. El trabajo de investigación se cimenta en el uso de datos provenientes de fuentes secundarias, tal como ocurre en la totalidad de documentos citados en la sección de antecedentes internacionales (ver sección 1.1.2. del documento).

Sin embargo, la hoja de cálculo que consolida el *dataset* (conjunto de datos) es considerada una herramienta de investigación para el ámbito de las investigaciones operativas (Ferreira, 2015). Por otro lado, Pakhare et al. (2013), en su *journal article* titulado *Use of mobile phones as research instrument for data collection*, mencionaron que la recolección de datos es un componente integral de todo plan o programa de investigación; estos datos recogidos deben introducirse en una hoja de cálculo o en un programa estadístico para poder ser analizado; la recopilación de datos mediante teléfonos móviles, una forma de tecnología de captura de datos electrónicos en la que se combinan la recopilación y la introducción de datos, es una opción potencial, *Google Trends* funciona bajo el mismo principio: captura datos a partir de las consultas que realizan los internautas, ya sea desde un celular, una tablet, una laptop o una computadora de escritorio (Montoya, 2020; Huang et al., 2019).

¿Los datos de búsquedas online de *Google Trends* brindan indicadores que son equiparables a las encuestas? En “*Forecasting private consumption: survey-based indicators vs. Google trends*”, Vosen y Schmidt (2011) evidenciaron que el rendimiento predictivo de las búsquedas *online* se compara con el Índice de Sentimiento del Consumidor de la Universidad de Michigan y el Índice de Confianza del Consumidor del *Conference Board*, los dos indicadores basados en encuestas que se miden con mayor frecuencia en Estados Unidos.

### 3.6. Validez de los datos obtenidos de Google Trends

La validez de los datos de búsquedas obtenidos en *Google Trends* ha sido objeto de investigaciones de distintas disciplinas. En *Reliability of Google Trends: Analysis of the Limits and Potential of Web Inveillance During COVID-19 Pandemic and for Future Research*, Rovetta (2021) indicó que *Google Trends* ha demostrado ser fiables para la vigilancia del COVID-19 si es que se emplean búsquedas globales.

Por otro lado, en el *journal article* titulado *Is Google Trends a quality data source*, Cebrián y Domenech (2022) afirmaron que el valor de un atributo es legítimo si entra en un rango predeterminado de valores aceptables; por ejemplo, la edad de una persona sólo puede ser un número positivo; la validación de los atributos es un requisito esencial para garantizar la exactitud, pero no es suficiente.

En general, los datos de los informes de *Google Trends* son precisos (Cebrián & Domenech, 2022; Ahmed, et al., 2017). Sin embargo, en determinadas circunstancias, dicho valor devuelve un resultado no entero, concretamente " $<1$ "; esto indica que el volumen de búsqueda fue suficiente para ser incluido en el informe, pero menos de una centésima parte del período de tiempo más popular. *Google Trends* utiliza esta notación para evitar la confusión con el valor "0" (Cebrián & Domenech, 2022). Para la presente investigación, los datos que representan a búsquedas online no registraron valores de " $<1$ " (ver anexo 2). Esto

representa en requisito de validez para los datos (Varian, 2014). Si es que ello no hubiese ocurrido lo adecuado sería emplear una combinación de términos clave en lugar de un único *keyword* (Cebrián & Domenech, 2022).

### 3.7. Recolección de datos

Las exportaciones de jengibre no triturado ni pulverizado se obtuvieron del portal ADEX DATA TRADE (2020). Basta con retornar la partida arancelaria 091011.

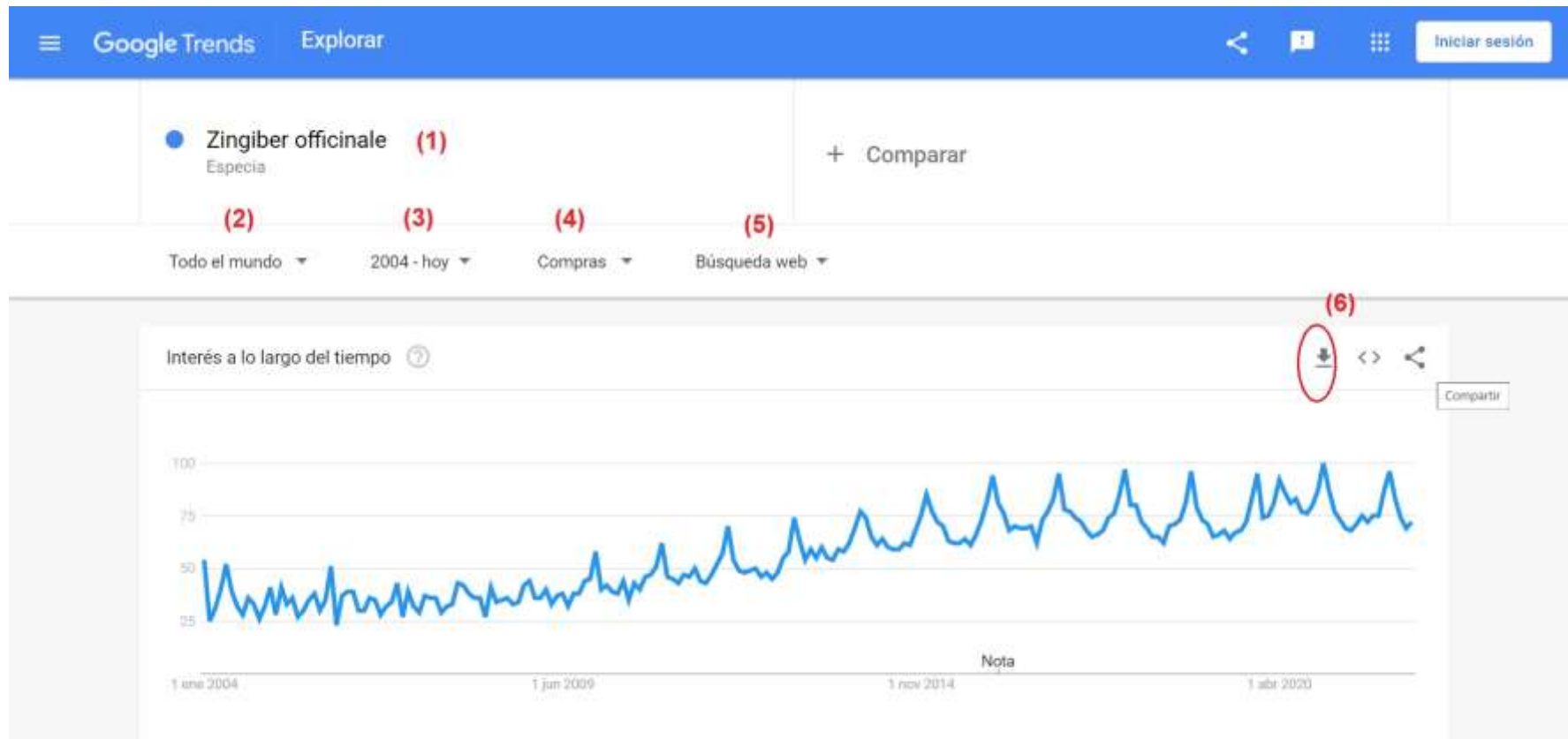
En el caso de los índices de *Google Trends* sí se requiere una secuencia de pasos:

- i. Ingresar a la aplicación *Google Trends*
- ii. Ingresar la palabra Jengibre en la caja donde se ubica la búsqueda y esperar a que el sistema reconozca al *keyword* como *Zingiber officinale*. Esto es una función de autocompletado.
- iii. Ingresar el lugar y fecha del que se desea obtener resultados: “Todo el mundo” y desde el 2004 para adelante.
- iv. En categoría se debe seleccionar *Compras*. Ello con el fin de de filtrar las búsquedas que son relacionas a ese ese fin.
- v. Seleccionar la dimensión a descargar. Estas son búsquedas web, búsquedas en YouTube y búsqueda de imágenes.
- vi. Descargar la data, la misma se exporta a un archivo CSV, compatible con Excel
- vii. Organizar los datos de tal manera de que al lado de cada dato de exportación aparezca las cantidades que se exportaron hace un mes y hace un año (rezagos), y los índices de Google Trends correspondiente al mes previo de las exportaciones.

En la siguiente figura se muestran los componentes del *dashboard* de Google Trends:

**Figura 11**

*Interfase de uso de Google Trends*



*Nota.* (1) representa a la zona de *keywords*, (2) área geográfica a abarcar, (3) tiempo a evaluar, (4) categoría de búsqueda, (5) dimensiones de *Google Trends*, (6) Botón de descarga. Adaptado de “Estudio exploratorio del uso de indicadores de Big Data de Google como variable en el modelamiento del flujo mensual de visitantes al Complejo Arqueológico de Kuélap en el periodo 2011-2016”. De León y Villar, 2018 (<http://repositorio.usil.edu.pe/handle/USIL/369>).

### 3.8. Plan de análisis

Como ya se ha mencionado, el análisis principal se orienta en mejorar el nivel de ajuste de un modelo autorregresivo base.

Modelo *Baseline* o en base un rezago incluye a  $X(t-12)$  en la predicción de  $X(t)$

A dicho opción de modelado se van a ir incorporando a los indicadores vinculados a *Google Trends*, tal como lo hicieron

- Modelo A: Incluye a los elementos del modelo AR-1 e incorpora búsquedas web y de YouTube
- Modelo B: Incluye a los elementos del modelo AR-1 e incorpora búsquedas web y de imágenes
- Modelo C: Incluye a los elementos del modelo AR-1 e incorpora búsquedas en YouTube y de imágenes
- Modelo D: Incluye a los elementos del modelo AR-1 e incorpora a todas las dimensiones de Google Trends.

En cuando a las hipótesis específicas, se requiere adicionar al modelo AR-1 los indicadores de *Google Trends*:

- Modelado E: asociado a la hipótesis específica 1 (AR-1 más búsquedas web)
- Modelado F: asociado a la hipótesis específica 2 (AR-1 más búsquedas en YouTube)
- Modelado G: asociado a la hipótesis específica 3 ((AR-1 más búsquedas de imágenes)

A continuación, se muestra una tabla con todos los modelos que han sido incluidos en la investigación:

**Tabla 11***Lista de modelos que adicionan indicadores Google al modelo basado en rezagos*

Modelo	Etiqueta alternativa	Rezago	Búsquedas web	Búsquedas en YouTube	Búsqueda de imágenes	Asociado a la hipótesis
A	RWY	Incluye	Incluye	Incluye	No incluye	General
B	RWI	Incluye	Incluye	No incluye	Incluye	General
C	RYI	Incluye	No incluye	Incluye	Incluye	General
D	RWYI	Incluye	Incluye	Incluye	Incluye	General
E	RW	Incluye	Incluye	No incluye	No incluye	Específica 1
F	RY	Incluye	No incluye	Incluye	No incluye	Específica 2
G	RI	Incluye	No incluye	No incluye	Incluye	Específica 3

En total se han generado 7 modelos, ello incorporando las distintas combinaciones posibles de índices de Google Trends al modelo basado en rezagos. La situación matemática se podría cuantificar de la siguiente manera:

$$\text{Modelos a emplear: } C_1^3 + C_2^3 + C_3^3$$

Donde  $C_k^n$  representa a las combinaciones posibles de k unidades a partir de n elementos.

Reemplazando los valores posibles se tiene

$$\text{Modelos a emplear: } \frac{3!}{(3-1)!1!} + \frac{3!}{(3-2)!2!} + \frac{3!}{(3-3)!3!}$$

$$\text{Modelos a emplear: } 3 + 3 + 1 = 7$$



### 3.6. Plan de análisis

El primer modelo que se construye es del tipo autorregresivo, es decir, uno que implica que las exportaciones solo dependen de sus rezagos y que todas las demás variables son independientes. A continuación, se crearán combinaciones de las variables de la web, YouTube y la búsqueda de imágenes para probar su eficacia. Es necesario llevar la cuenta de la cantidad de componentes de la capa oculta que tenemos y las funciones de activación que tenemos para cada modelo que se construye.

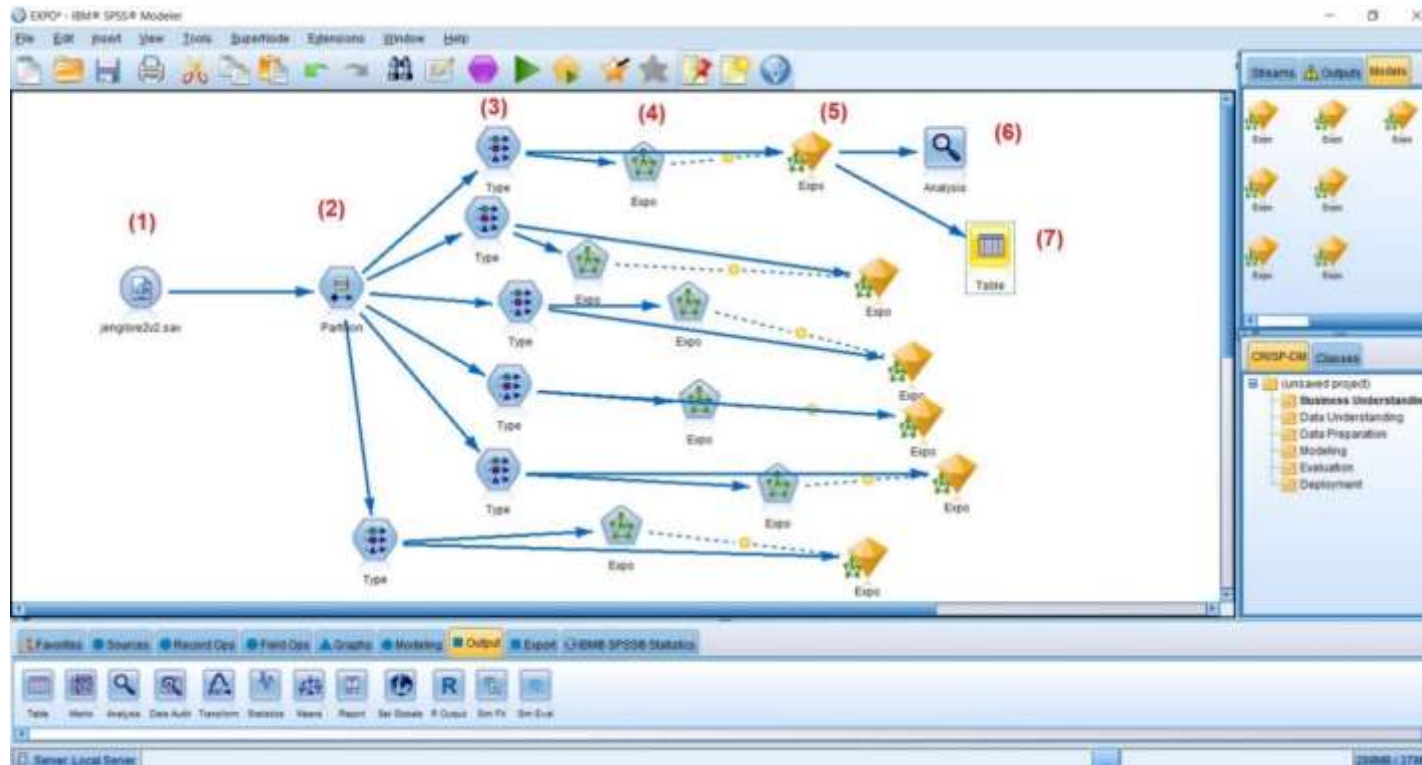
Para determinar el valor de verdad de cada una de las hipótesis de trabajo, es necesario comparar la precisión de los múltiples modelos vs. el modelo base de referencia que únicamente incorpora rezagos.

A continuación, el modelo que mejor se ajusta se compara con un modelo que sólo contiene los valores de las exportaciones anteriores, es decir, el modelo de referencia. En una etapa siguiente, se pueden ver las métricas relacionadas con las particiones de entrenamiento y de prueba. La primera es la que se utiliza para entrenar el modelo (es decir, la máquina aprende), y la segunda la que se utiliza para probar el entrenamiento (IBM, 2017). Un subconjunto de los datos se utiliza para el entrenamiento, mientras que el resto se utiliza para las pruebas. El consenso común es que una proporción de 70 por ciento de entrenamiento y 30 por ciento de pruebas es adecuada. Sin embargo, según Wendler & Gröttrup (2016), *SPSS Modeler* puede realizar modificaciones en esta proporción.

El software *SPSS Modeler* funciona como un flujo de órdenes lógicas ordenadas en cajas (IBM, 2017). En la siguiente figura se muestran los pasos necesarios para obtener las salidas que se muestran en la sección resultados:

Figura 12

Flujograma del SPSS Modeler



Adaptado de “Data Mining with SPSS Modeler, Theory, Exercises and Solutions”, por Wendler y Gröttrup , 2016.

Nota. (1) representa al origen de datos, (2) es donde se particiona la data , en (3) se indican las variables dependientes e independientes, (4) es la red neuronal propiamente dicha y (5) el resultado, mientras que (6) y (7) son las salidas del modelo .

La comparación de los modelos A, B, C y D con el modelo AR-1 pone a prueba la hipótesis general; mientras que las hipótesis particulares 1, 2 y 3 necesitan el modelado y el análisis de los modelos E, F y G son necesarios para las hipótesis específicas 1, 2 y 3.

### 3.9. Limitaciones de la metodología empleada

Siempre se debe tomar en consideración que los datos correspondientes a las búsquedas *online* son relativos y no absolutos (León & Villar, 2018; Sosa & Jeri, 2019). La aplicación *Google Trends* no indica, por ejemplo, si en enero del 2018 se produjeron 800,327 búsquedas acerca del jengibre en YouTube alrededor del mundo, simplemente se mostrará un valor que oscila desde 0 hasta 100, donde 100 representa al valor máximo de búsquedas que se dieron para el período analizado, estos índices, a diferencia de las exportaciones, son valores redondeados (León & Villar, 2018)

Cebrían y Domenech (2022) indicaron que los valores de Google Trends presentan un problema asociado a la exactitud de los datos puede ser considerados como especialmente crítico, ya que la falta de un valor preciso, ello podría constituir una importante fuente de sesgo. Además, cuando los datos se utilizan para estimar modelos econométricos, las estimaciones de los parámetros pueden verse afectadas, lo que, en última instancia, podría conducir a juicios económicos o políticos incorrectos (Cebrían & Domenech, 2022; Varian, 2014). Este es un sesgo que los investigadores asumen. A pesar de ello, los datos de búsquedas de *Google Trends* han sido empleado con éxito en investigaciones de negocios e ingeniería, tal como se muestra en los antecedentes nacionales e internacionales de la investigación.

Por otro lado, otra limitación principal radica en el propio uso de las redes neuronales: a diferencia de la regresión lineal, la red neuronal no permite establecer correlaciones directa o inversas que posean significancia (Brenner, 2018); sin embargo, los antecedentes

mencionados apuntan más al incremento de la precisión de un modelo debido a la incorporación de los índices de *Google Trends* como variables predictoras. A modo de complemento, los autores de la tesis también emplearon el modelo de referencia recurriendo al modelo de regresión lineal, debido a que es relativamente sencilla de interpretar.

## CAPÍTULO IV. DESARROLLO Y APLICACIÓN

### 4.1. Resultados asociados a la hipótesis general

En la siguiente tabla se muestran los resultados del modelo base, aquel enfocado en rezagos, versus los escenarios en donde se incorporaron las distintas dimensiones de las búsquedas online, medidas con *Google Trends*.

**Tabla 12**

*Comparativo de modelos empleados para la hipótesis general*

		Modelo Base (R)	Modelo A	Modelo B	Modelo C	Modelo D
Training	Sum of Squares Error	14.994	8.474	8.384	6.236	12.487
	Relative Error	.384	.253	.212	.183	.301
	Sum of Squares Error	2.913	7.911	7.841	3.254	.889
Testing	Relative Error	.214	.262	.480	.218	.206
	Cantidad de Perceptrones	2	2	4	4	2
1ra Función Activación	Tangente hiperbólica	Tangente hiperbólica	Tangente hiperbólica	Tangente hiperbólica	Tangente hiperbólica	Tangente hiperbólica
2 da Función Activación	Identidad	Identidad	Identidad	Identidad	Identidad	Identidad
Accuracy		66.15%	75.29%	70.99%	80.84%	71.54%

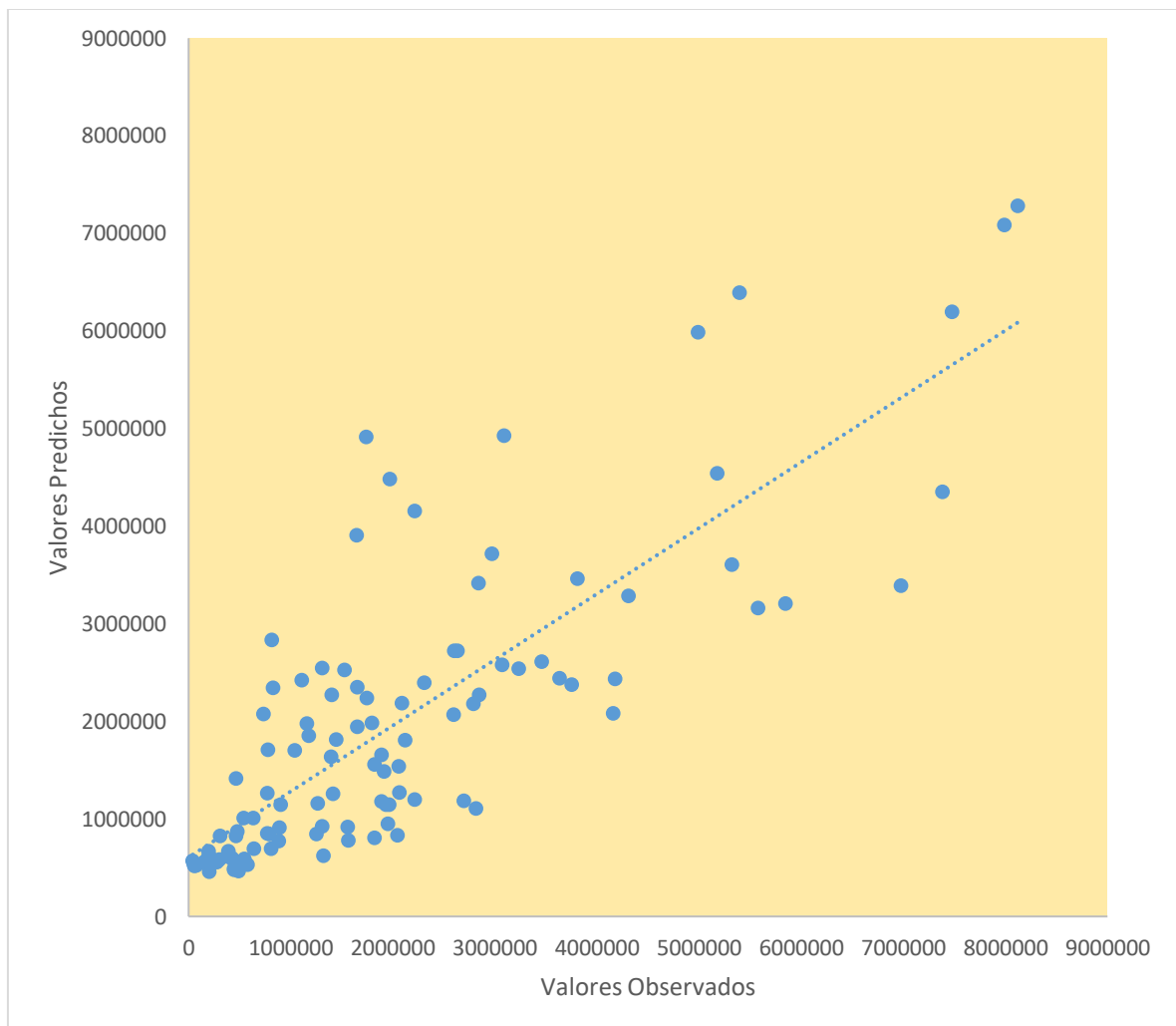
La opción de mejor ajuste es el modelo C, aquel que incorpora a las búsquedas YouTube y de imágenes (80.84%); incluso superando al modelo que incluye todos los predictores (modelo D). El modelo C presenta un incremento del *accuracy* de 14.79% sobre el modelo base basado en rezagos (Modelo A). El modelo de mejor ajuste, además, posee un menor nivel de error en comparación a sus pares. La función de activación son la tangente hiperbólica y la identidad.

Cabe mencionar, nuevamente, que el adicionar todas las dimensiones de las búsquedas *online* no es lo idóneo si lo que se busca es la máxima precisión posible. El modelo que incluyó a todos los predictores tuvo una precisión menor que aquel que aquel que combinó a las búsquedas en YouTube y de imágenes (71.54% < 80.84%).

En la siguiente figura se muestra el gráfico de dispersión correspondiente a los valores predichos vs. observados para el modelo basado en rezagos que sirve como *baseline*:

**Figura 13**

*Valores observados vs predichos para el modelo basado en rezagos*



En la siguiente tabla se muestra la importancia de las variables predictoras para cada modelo con el fin de comprender mejor el impacto de cada variable.

**Tabla 13**

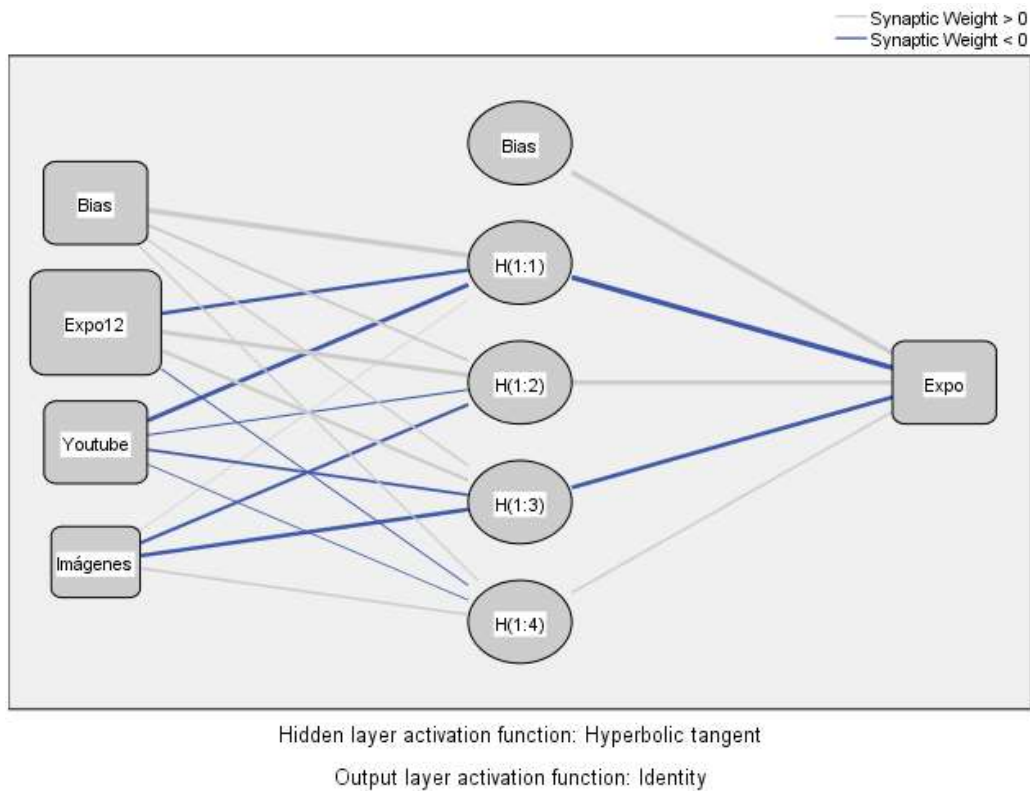
*Importancia de las variables predictoras en los modelos asociados a la hipótesis general*

Variable	Modelo Base	Modelo A	Modelo B	Modelo C	Modelo D
Rezago de 12 meses	100%	65%	77.9%	58.3%	47%
Búsquedas web	No aplicado	5.1%	9.8%	No aplicado	25%
Búsquedas en YouTube	No aplicado	29.8%	No aplicado	28.5%	15%
Búsquedas de imágenes	No aplicado	No aplicado	12.3%	13.2%	10%

De la tabla anterior se puede inferir que el rezago correspondiente a un año es la variable de mayor importancia. A ella le sigue las búsquedas en YouTube para el modelo de mejor ajuste. A continuación, se muestra la arquitectura de la red, en donde se aprecia las dos neuronas que forman parte de la capa oculta.

**Figura 14**

*Arquitectura de la red neuronal de mejor ajuste*



La arquitectura de la red neuronal que obtuvo mejor ajuste tiene como predictores a los rezagos correspondientes a un año de las exportaciones, y a las búsquedas en YouTube y de imágenes.

En una primera etapa, los elementos de entrada se combinaron empleando una función tangente hiperbólica, una función trigonométrica, para formar los perceptrones  $H(1,1)$  hasta el  $H(1,4)$ . Estos, junto al sesgo (o *Bias*) se combinan con la función identidad, estructura análoga a la regresión lineal múltiple, para dar paso al *output* del modelo: las exportaciones de jengibre no triturado ni pulverizado.

La tangente hiperbólica tiene como representación base la siguiente estructura para una sola variable:

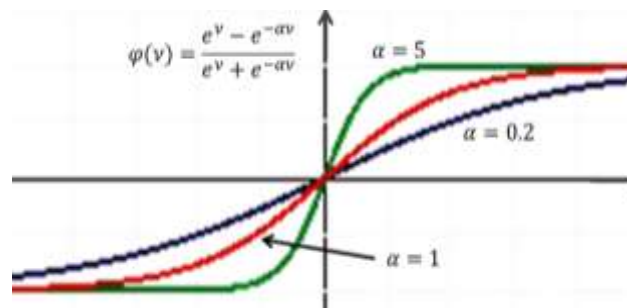
$$\tanh x = \frac{\operatorname{senhx}}{\operatorname{coshx}}$$

Dicha expresión puede ser expresada en función al logaritmo natural, tal como se expresa a continuación

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

**Figura 15**

*Función tangente hiperbólica para una variable*



De "Evaluación de dos técnicas de reconocimiento de patrones para su implementación en el Simulador de pilotaje automático de taller del STC Metro de la Cd. de México", por Arroyo-Fernández, 2013 (<http://repositorio.utm.mx:8080/jsui/handle/123456789/61>).



Esta función, aplicada a varias variables, es la primera etapa de la construcción de las redes neuronales. Con las variables de entrada del modelo (rezagos e indicadores de Google Trends) se forma la capa intermedia u oculta. Ya en la segunda etapa de la red se activa la función identidad. Desde luego, las redes neuronales poseen parámetros igual que los modelos de regresión, pero los mismos no poseen una interpretación propiamente dicha. Basta con ver la siguiente tabla:

**Tabla 14**

*Parámetros de la red neuronal de mejor ajuste*

Predictor		Predicted				Output Layer
		Hidden Layer 1				Expo
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	
Input Layer	(Bias)	2.552	.503	.350	.244	
	Expo12	-.704	1.023	.526	-.098	
	YouTube	-.851	-.139	-.461	-.079	
	Imágenes	.057	-.506	-.767	.416	
Hidden Layer 1	(Bias)					1.276
	H(1:1)					-1.769
	H(1:2)					1.115
	H(1:3)					-.778
	H(1:4)					.370

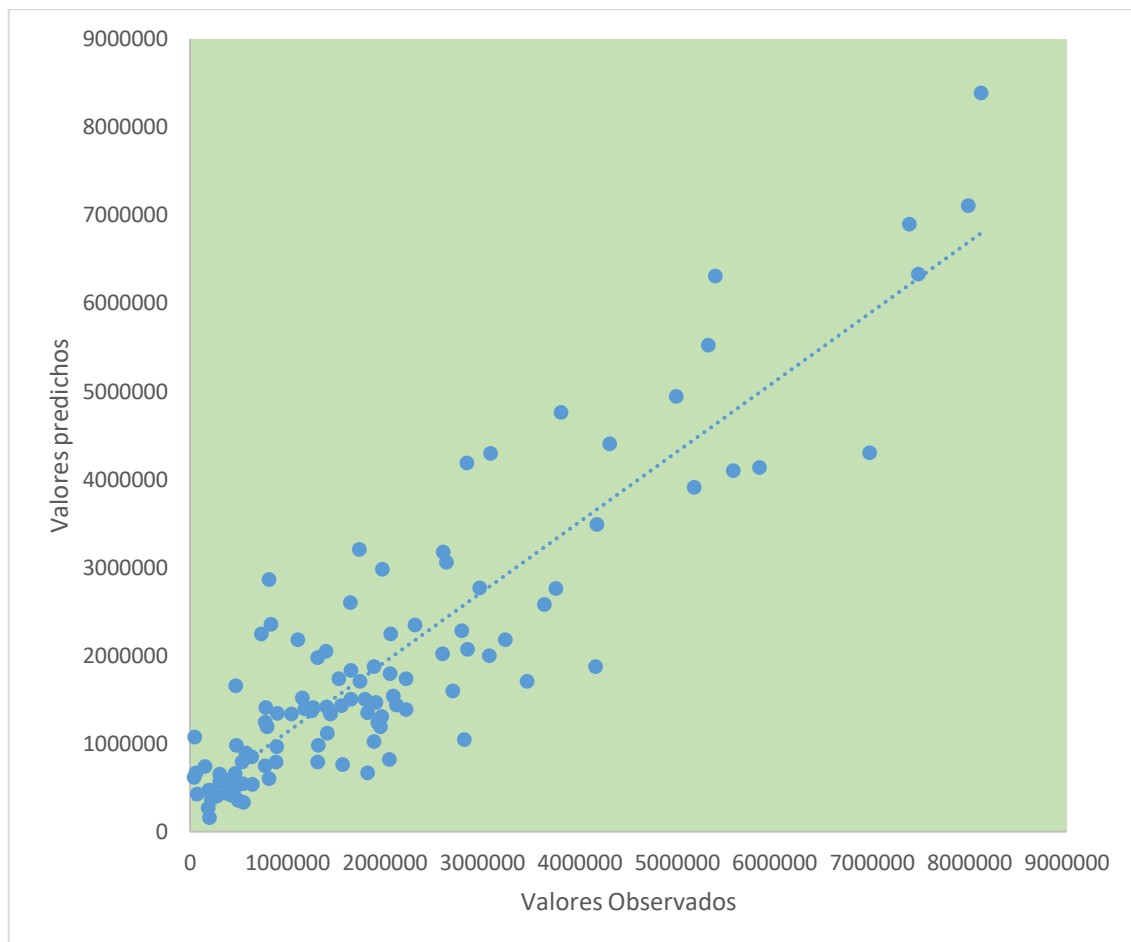
Los parámetros de la red son una combinación de coeficientes positivos y negativos a partir de los cuales no se puede inferir resultados en particular. Esta es una clara diferencia con los modelos de regresión (Khademi et al., 2015). Es por ello que se habla a menudo del problema de la interpretabilidad en las redes neuronales. La interpretación matemática incluye una combinación de la función trigonométrica tangente hiperbólica junto a una relación lineal, todo ello bajo un arreglo matricial. Al carecer de una interpretación correlacional, por lo general se omite la presentación de los parámetros de las redes neuronales en los *journal*

articles (Ramirez & Vilela, 2020). Ello no suele ocurrir con las regresiones lineales, donde los parámetros si poseen una interpretación propiamente dicha.

En la siguiente figura se muestra el diagrama de dispersión para los valores observados y predichos de la red neuronal de mejor ajuste:

**Figura 16**

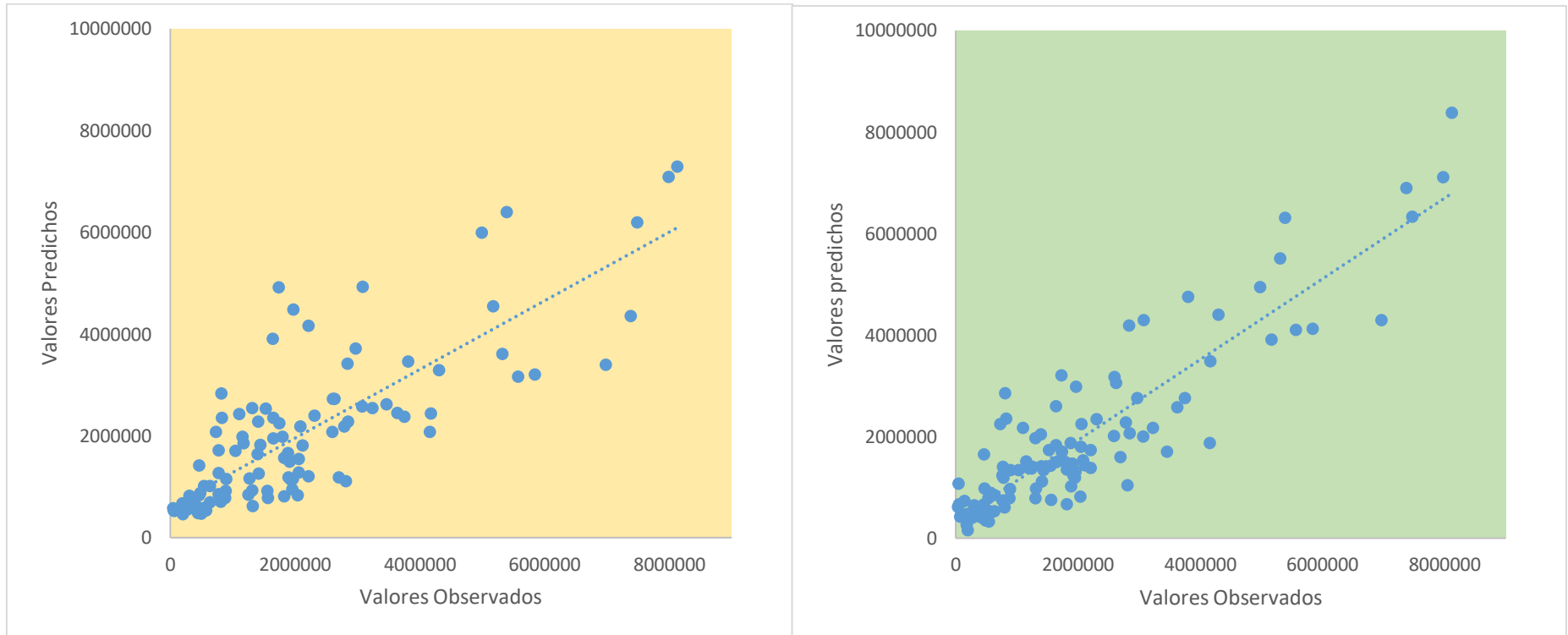
*Valores observados vs predichos para el modelo de mejor ajuste que incluye datos de Google Trends*



Gráficamente, el modelo de mejor ajuste posee un menor grado de dispersión que el modelo basado en rezagos (ver figura anterior). A continuación, se muestra un comparativa visual:

**Figura 17**

Diagrama de dispersión para valores observados y predichos en los modelos base (izquierda) y el de mejor ajuste (derecha)

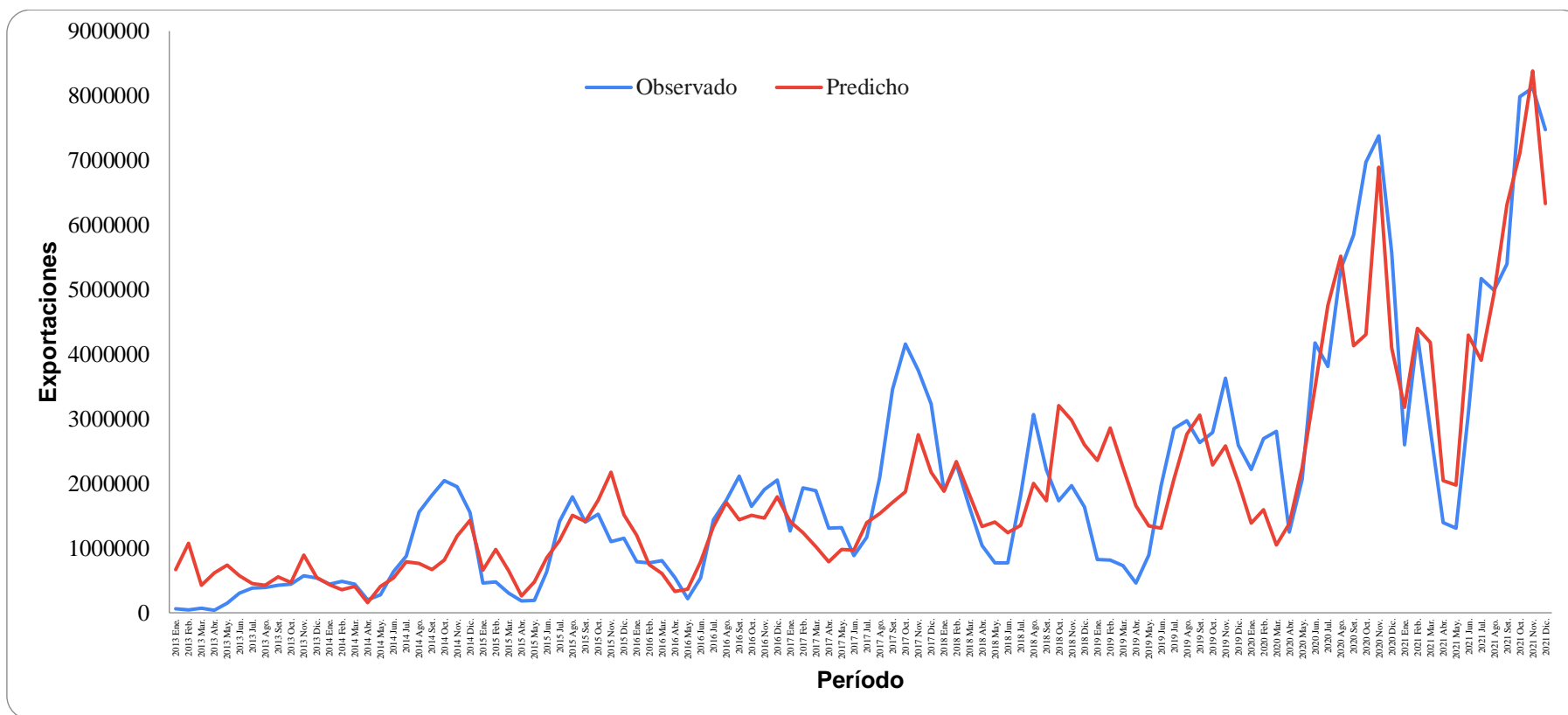


Nuevamente, el modelo de mayor ajuste, aquel que incluye a las búsquedas en YouTube e imágenes, posee un menor nivel de dispersión en cuanto a la relación de valores predichos y observados.

En la figura siguiente se muestra a los valores de las exportaciones peruanas de jengibre fresco versus los valores predichos por la red neuronal:

**Figura 18**

*Valores predichos vs. observados para las exportaciones de jengibre (2013-2021)*



A modo complementario se presentan los resultados en caso de que se optara por el modelo de regresión lineal, fue empleada la regresión por pasos, la misma que procura adicionar variables a medida que se produzca un incremento en el R-cuadrado ajustado (Wendler & Gröttrup, 2016):

**Tabla 15**

*Coefficientes de los modelos de regresión base y de mejor ajuste*

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	Estadísticas de colinealidad	
	B	Error estándar	Beta			Tolerancia	VIF
1 (Constante)	464670.477	146497.952		3.172	.002		
Expo-12	1.016	.069	.818	14.647	.000	1.000	1.000
2 (Constante)	-646074.426	266985.587		-2.420	.017		
Expo-12	.833	.074	.671	11.319	.000	.734	1.362
YouTube	25394.360	5288.886	.285	4.801	.000	.734	1.362

En el primer modelo se planteó que las exportaciones solo dependen de sus rezagos. La ecuación quedaría de la siguiente forma:

$$Expo(t) = 464670.477 + 1.016Expo(t - 12)$$

En el modelo 2, se evidencia que tanto las búsquedas en YouTube como los rezagos tienen un aporte significativo en la predicción de las exportaciones mensuales. La relación entre las variables es del tipo directa y puede ser expresada de la siguiente forma:

$$Expo(t) = -646074.426 + 0.833Expo(t - 12) + 25394.36YouTube(t - 1)$$

Pero es necesario contar con un indicador asociado a la calidad del modelo. El R-cuadrado representa al *accuracy* del modelo Wendler & Gröttrup (2016). En la siguiente tabla se muestra los valores de dicho indicador para los modelos de regresión lineal ya detallados

**Tabla 16***Coefficientes de los modelos de regresión base y de mejor ajuste*

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Durbin-Watson
1	.818 <sup>a</sup>	.669	.666	1062405.022739	
2	.854 <sup>b</sup>	.729	.724	966599.334425	.673

El modelo de mejor ajuste bajo el enfoque de regresiones alcanzó una precisión de 0.729, lo cual se ubica por debajo del modelo basado en redes neuronales.

Cabe destacar que las regresiones buscan el mejor ajuste empleado va agregando variables a medida que se incrementa el R-cuadrado ajustado.

**Tabla 17***Variables excluidas del modelo de regresión*

Modelo	En beta	t	Sig.	Correlación parcial	Estadísticas de colinealidad			
					Tolerancia	VIF	Tolerancia mínima	
1	Web	.118 <sup>b</sup>	2.064	.042	.197	.924	1.082	.924
	YouTube	.285 <sup>b</sup>	4.801	.000	.424	.734	1.362	.734
	Imágenes	.106 <sup>b</sup>	1.849	.067	.178	.934	1.071	.934
2	Web	-.017 <sup>c</sup>	-.276	.783	-.027	.679	1.472	.540
	Imágenes	.030 <sup>c</sup>	.541	.589	.053	.843	1.186	.663

Con el fin de alcanzar un mejor R-cuadrado ajustado, el SPSS excluyó a las búsquedas web y de imágenes del modelo final; ambas, además, superaron el nivel de significancia máxima establecido.

#### 4.2. Comprobación de la hipótesis general

“Los índices globales de búsquedas *online* acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las

redes neuronales.” VERDADERO. No obstante, ello no se cumple incorporando todas las dimensiones. El modelo de mejor ajuste bajo el enfoque de redes neuronales se obtiene adicionando las búsquedas en YouTube y las de imágenes. El modelo basado en redes neuronales supero también en precisión al de regresión lineal ( $0.8084 > 0.729$ ).

#### 4.3. Discusión de la hipótesis general

El modelo de mayor *accuracy* ha registrado una mayor precisión que las propuestas de Francia y Pacheco (2019) y Alvarez (2020) a Países Bajos y Estados Unidos respectivamente, estas registraron precisiones de 51.3% y 42.7% y son inferiores al 80.84% alcanzado con las redes neuronales. El modelo aplicado al contexto peruano también logro superar en precisión a la propuesta de Onomayo et al. (2022) aplicada a las exportaciones de jengibre realizadas desde Nigeria (67.10%). Los mencionados trabajos de investigación se cimentan en el uso de indicadores más tradicionales (tipo de cambio, precio FOB, nivel de ingreso, etc); mientras que para la presente tesis se optó por una variable del tipo *Big Data* que cuantifica las búsquedas *online* globales.

Lo sugerido por Gopinath et al. (2021) con respecto al uso del *machine learning*, entonces, posee sustento, de ello queda evidencia en la precisión del modelo final que superó al enfoque que empleó una regresión lineal ( $80.84\% > 72.9\%$ ). Luego, Collison (2020) y Yu et al. (2019), adicionando las búsquedas *online*, mejoraron las precisiones de sus modelos base que pronosticaron las ventas de autos en argentina y las de petróleo en 21% y 4% respectivamente, aunque el primero tenía una precisión inicial muy baja (28%). Por otra parte, León y Villar (2018) evidenciaron que las búsquedas de imágenes y en YouTube no aportaron significativamente al pronóstico de arribos a Kuélap. No obstante, no ocurrió lo mismo para el caso del jengibre peruano, aunque cabe destacar que las investigadoras usaron modelos de regresión en lugar de redes neuronales. Por otro lado, se ha podido establecer un ranking de importancia entre las variables predictoras, tal como lo realizó Akman et al.

(2020); a pesar de que la teoría de las redes neuronales describe el problema de interpretabilidad asociada a las mismas (Vu,1996; Bocco et al., 2010, Brenner, 2018), el hecho de probar distintas combinaciones y medir la importancia de los predictores contribuye a la comprensión del modelo final. Las búsquedas en YouTube juegan un rol particularmente relevante en las exportaciones de jengibre peruano fresco.

Mayasari et al. (2020) sugirieron que las búsquedas *online* acerca de productos alimenticios realizadas durante el COVID-19 (jengibre entre ellos) bien podrían describir cambios en los hábitos de consumo de la persona. Sin embargo, dichos investigadores se limitaron a describir los índices de búsquedas, obtenidos de *Google Trends*, acerca de una serie de alimentos antes y durante la pandemia. Los hallazgos para el caso del jengibre peruano y su *forecasting* con soporte en redes neuronales sugieren que en realidad la asociación entre búsquedas y consumo datan desde antes de la llegada del COVID-19. Woo y Owen (2018) indicaron que las búsquedas *online* acerca de alimentos no influenciaba en la demanda de los mismos; sin embargo, ellos solo emplearon la dimensión de contenido noticioso, la misma que no ha sido aplicada para la investigación del jengibre peruano. En este sentido, los resultados del modelo con soporte en redes neuronales evidencian que resulta más adecuado emplear otras dimensiones de las búsquedas *online* , tal como son las búsquedas en YouTube o las de imágenes. Otro aspecto a considerar es que para el caso de jengibre se filtraron las búsquedas a la categoría compra. Es posible que esto haya terminado impactando la calidad del modelo.

#### 4.4. Resultados asociados a la hipótesis específica 1

En la siguiente tabla se muestran los resultados del modelo base, aquel enfocado en rezagos, versus el escenario en donde únicamente se adiciona a las búsquedas web.



**Tabla 18***Comparativo entre modelo base y modelo asociado a la hipótesis específica 1*

		Modelo Base	Modelo E
Training	Sum of Squares Error	14.994	13.264
	Relative Error	0.384	.336
Testing	Sum of Squares Error	2.913	3.997
	Relative Error	0.214	.445
Cantidad de Perceptrones		2	2
1ra Función Activación		Tangente hiperbólica	Tangente hiperbólica
2 da Función Activación		Identidad	Identidad
Accuracy		66.15%	65.92%

El adicionar las búsquedas web a los rezagos no mejora la predicción de las exportaciones.

Por el contrario, el *accuracy* cae de 66.15 a 65.92%.

**Tabla 19***Importancia de las variables predictoras en los modelos asociados a la hipótesis específica 1*

Variable	Modelo Base	Modelo E
Expo-12	100%	84.4%
Web	No aplicado	15.6%

La dimensión búsquedas en la web es de menor relevancia que el rezago correspondiente a un año. Si es que optara por emplear el modelo de regresión lineal, los resultados a obtener serían los siguientes:

**Tabla 20***Coefficientes del Modelo de regresión en base a rezagos y búsquedas web*

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados			Estadísticas de colinealidad	
	B	Error estándar	Beta	t	Sig.	Tolerancia	VIF
(Constante)	-1001857.854	725091.922		-1.382	.170		
Expo-12	.975	.071	.786	13.730	.000	.924	1.082
Web	22473.126	10889.090	.118	2.064	.042	.924	1.082

La ecuación de regresión quedaría de la siguiente forma:

$$Expo(t) = -1001857.854 + 0.975Expo(t - 12) + 22473.126Web(t - 1)$$

La ecuación evidencia una relación directa entre las búsquedas web y las exportaciones de jengibre.

**Tabla 21**

*Resumen del Modelo de regresión en base a rezagos y búsquedas web*

R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Durbin-Watson
.826 <sup>a</sup>	.682	.676	1046438.553779	.553

El R-cuadrado del modelo es superior al que únicamente incluye a los rezagos (0.682 >0.669). Dicho de otra forma, las búsquedas *web* acerca del producto mencionado contribuyen a la predictibilidad de las exportaciones del mismo bajo el modelo de regresión, no ocurre lo mismo con las redes neuronales.

#### 4.5. Comprobación de la hipótesis específica 1

“Los índices globales de búsquedas web acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales” FALSO. No obstante, cabe mencionar que el enunciado sería verdadero si es que se basara en el modelo de regresión lineal.

#### 4.6. Discusión de la hipótesis específica 1

Park (2015) evidenció que las búsquedas web acerca del *Korean Drama* era un predictor de las exportaciones de cosmético surcoreanas, mientras que Başığit (2021) empleó las búsquedas web para pronosticar las exportaciones de mármol turco a Estados Unidos; ambos

productos son de valor relativamente alto, mientras que el jengibre es un bien fungible para los consumidores finales. En ambos casos fueron empleados regresiones. Si bien es cierto que para el caso del jengibre peruano sí se produce una mejora al adicionar las búsquedas web, dicha mejora es marginal (de 0.669 a 0.682) bajo el modelo de regresión lineal múltiple; mientras que para las redes neuronales se produce una disminución en la calidad del *accuracy*.

Ciertamente una mejora significativa tan pequeña coincide con los hallazgos de Golovanova y Zubarev (2021), quienes evidenciaron que al añadir las búsquedas web a los modelos pre-establecidos para el comercio de alimentos en el sector retail ruso se lograban una mejora de en la precisión de 0.02; mientras que para el caso del jengibre peruano la mejora fue de 0.013.

Por otro lado, debe considerarse que a diferencia del estudio de Boone et al. (2018), quienes emplearon *keywords* de marcas específicas y logró mejoras en el *accuracy* de entre 0.02 y 0.06, para el caso del jengibre peruano únicamente se empleó a un *keyword* generado por la función de autocompletado de *Google Trends*; la aplicación sugirió el nombre científico en lugar del nombre genérico ingresado por los investigadores (jengibre). Este hecho contrasta también con el estudio de Borup y Montes (2022), quienes emplearon a un conjunto de *keywords* asociados a la búsqueda de trabajo para modelar el nivel de empleo.

Si consideramos a las exportaciones nacionales de jengibre como una variable macroeconómica nacional, los resultados del modelo de regresión coinciden con los hallazgos de Pilco y Sandoval (2018), quienes fueron capaces de mejorar la precisión de modelos orientados a predecir el empleo y la inversión con búsquedas web en territorio peruano.

#### 4.7. Resultados asociados a la hipótesis específica 2

En la siguiente tabla se muestran los resultados del modelo base, aquel enfocado en rezagos, versus el escenario en donde únicamente se adiciona a las búsquedas en YouTube:

**Tabla 22**

*Comparativo entre modelo base y el modelo asociado a la hipótesis específica 2*

		Modelo Base	Modelo F
Training	Sum of Squares Error	14.994	12.477
	Relative Error	0.384	0.367
Testing	Sum of Squares Error	2.913	3.687
	Relative Error	0.214	.445
Cantidad de Perceptrones		2	2
1ra Función Activación		Tangente hiperbólica	Tangente hiperbólica
2 da Función Activación		Identidad	Identidad
Accuracy		66.15%	68.96%

Se observa una mejora en el *accuracy* del modelo de 2.81%. Esta situación es contraria al modelo que solo incorpora las búsquedas web a los rezagos, donde hubo una caída en la precisión. En cuanto a la importancia de las variables predictoras se tiene:

**Tabla 23**

*Importancia de las variables predictoras en los modelos asociados a la hipótesis específica 2*

Variable	Modelo Base	Modelo F
Expo-12	100%	59.3%
YouTube	No aplicado	40.7%

Nuevamente, para esta opción de modelado, los rezagos son más relevantes que los datos de origen electrónico (búsquedas en YouTube). Los rezagos correspondientes a un año constituyen el elemento predictor de mayor relevancia.

Si es que optara por emplear el modelo de regresión lineal, los resultados a obtener serían los siguientes:

**Tabla 24**

*Coefficientes del modelo de regresión en base a rezagos y búsquedas en YouTube*

	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	Estadísticas de colinealidad	
	B	Error estándar	Beta			Tolerancia	VIF
(Constante)	646074.426	266985.587		-2.42	0.017		
Expo-12	0.833	0.074	0.671	11.319	0	0.734	1.362
YouTube	25394.36	5288.886	0.285	4.801	0	0.734	1.362

$$Expo(t) = -646074.426 + 0.833Expo(t - 12) + 22473.126YouTube(t - 1)$$

Nuevamente aquí se evidencia una relación entre la variable predicha y los dos predictores (rezagos y búsquedas en YouTube).

**Tabla 25**

*Resumen del modelo de regresión en base a rezagos y búsquedas en YouTube*

R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Durbin-Watson
.854 <sup>a</sup>	.729	.724	966599.334425	.673

El R-cuadrado del modelo que adicionó a las búsquedas en YouTube es superior al que únicamente incluye a los rezagos (0.729 > 0.669). Dicho de otra forma, las búsquedas online acerca del producto mencionado contribuyen a la predictibilidad de las exportaciones del mismo bajo el modelo de regresión, ocurrió lo mismo con las redes neuronales.

#### 4.8. Comprobación de la hipótesis específica 2

“Los índices globales de búsquedas en YouTube acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni

pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.” VERDADERO. Ello ocurre tanto para ambas opciones de modelado.

#### 4.9. Discusión de la hipótesis específica 2

De acuerdo a Ma et. al. (2020), el interés suscitado por un documental en particular puede influir en el consumo de bienes orgánicos, pero solo de manera temporal y solo mientras el contenido despierte un interés significativo. En cambio, el interés hacia el jengibre, que pueden evidenciarse en plataformas como YouTube, es continuo; además, de acuerdo a Mayasari et al. (2020), es muy poco probable que el interés por dicho producto desaparezca en el corto plazo. De acuerdo a la *Theory of planned behavior* de Ajzen (1985, 2015), la probabilidad de que una persona forme una intención de realizar la conducta en cuestión aumenta en proporción a la positividad de la actitud hacia la conducta; con la llegada del COVID-19 se incrementó la actitud positiva hacia el consumo de jengibre (Mayasari et al., 2020); es probable que dicha actitud positiva pueda haberse visto potenciada tanto por los videos generados por los usuarios como por los comentarios acerca de los mismos.

Los resultados coinciden con los de Sosa y Jeri (2019), quienes evidenciaron que las búsquedas en YouTube eran predictores de los arribos al Jorge Chávez, la diferencia radica en que ellos se basaron únicamente en el modelo de regresión lineal, mientras que para el caso del jengibre peruano, la mejora en el *accuracy* se ha dado tanto en la regresión lineal como en las redes neuronales.

Al igual que como pasó con Golovanova y Zubarev (2021), una parte importante de la muestra de la prueba coincide con el periodo de una fuerte conmoción de la economía mundial producida por la pandemia, lo que aumenta la importancia de los resultados presentados para el caso del jengibre peruano. De acuerdo a France et al. (2021), los datos

de *Google Trends* son valiosos para los negocios porque permiten aproximarse a la demanda. Ello se ha podido verificar empíricamente para los datos de YouTube, pero los resultados son más valiosos en términos de precisión del modelo si es que se adicionan las búsquedas de imágenes.

#### 4.10. Resultados asociados a la hipótesis específica 3

En la siguiente tabla se muestran los resultados del modelo base, aquel enfocado en rezagos, versus el escenario en donde únicamente se adiciona de imágenes.

**Tabla 26**

*Comparativo entre modelo base y modelo asociado a la hipótesis específica 3*

		Modelo Base	Modelo G
Training	Sum of Squares Error	14.994	12.840
	Relative Error	0.384	0.352
Testing	Sum of Squares Error	2.913	6.030
	Relative Error	0.214	0.267
Cantidad de Perceptrones		2	2
1ra Función Activación		Tangente hiperbólica	Tangente hiperbólica
2 da Función Activación		Identidad	Identidad
Accuracy		66.15%	68.61%

Se observa que el *accuracy* del modelo se incrementa en 2.46 puntos porcentuales. Este resultado es similar al modelo que únicamente adicionó a las búsquedas en YouTube.

En cuanto a la importancia de los predictores se tiene:

**Tabla 27**

*Importancia de las variables predictoras en los modelos asociados a la hipótesis específica 3*

Variable	Modelo Base	Modelo G
Expo-12	27%	92.3%
Imágenes	No aplicado	7.7%

A diferencia del modelo que adicionó únicamente a las búsquedas en YouTube, la importancia relativa de las imágenes es mucho menor.

Si es que optara por emplear el modelo de regresión lineal, los resultados a obtener serían los siguientes:

**Tabla 28**

*Coefficientes del modelo de regresión en base a rezagos y búsquedas en YouTube*

	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	Estadísticas de colinealidad	
	B	Error estándar	Beta			Tolerancia	VIF
(Constante)	-759773.892	677803.136		-1.121	0.265		
Expo-12	0.98195289	0.071	0.791	13.836	0.000	0.934	1.071
Imágenes	18599.773	10058.215	0.106	1.849	0.067	0.934	1.071

$$Expo(t) = -759773.892 + 0.98195Expo(t - 12) + 18.599.773Imágenes(t - 1)$$

Nuevamente aquí se evidencia una relación entre la variable predicha y los dos predictores (rezagos y búsquedas de imágenes). Sin embargo, la relación no es significativa ( $p\text{-valor} = 0.067 > 0.05$ ), en consecuencia, se debe retirar a la variable asociada a imágenes del modelo.

#### 4.11. Comprobación de la hipótesis específica 3

“Los índices globales de búsquedas de imágenes de jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales”. VERDADERO. Ello bajo el modelo de redes neuronales.

#### 4.12. Discusión de la hipótesis específica 3

La búsqueda de imágenes es un predictor de las exportaciones de jengibre fresco. Ello ha sido evidenciado a el modelo basado en redes neuronales. No obstante, debe tomarse en consideración que una importante proporción de las imágenes que vemos en la web son



contenido generado por el usuario. En el modelo AISAS, la última “S” representa a *Share*, y esta acción llevada a cabo por un individuo puede llevar a la atención e interés (dos primeros componentes del AISAS) de otros. Así, por ejemplo, si un generador de contenido comparte imágenes en Instagram que involucre al jengibre (*Share*), ya sea desde una temática gastronómica o de salud (o ambas combinadas), puede terminar por generar reacciones en el consumidor (*Attention e Interest*), lo cual podría terminar en una potencial compra.

Debe tomarse en consideración que, al igual que Ramirez y Vilela (2020), la variable a predecir se ve afectada no solo por los rezagos, sino también por las búsquedas *online* previas. Ellos, usando las redes neuronales, evidenciaron que el interés hacia el caso Odebrechet en Estados Unidos y Perú terminaba impactado en el valor de las acciones de Graña y Montero en la NYSE. Para el caso del jengibre peruano se ha obtenido un resultado similar, aunque los rezagos fueron mensuales y no diarios; además se emplearon las búsquedas globales y no de países específicos.

Ciertamente el comportamiento de las exportaciones de jengibre peruano puede relacionarse significativamente con otras variables como la participación en ferias (Arevalo & Latour, 2021) o incluso con el PBI Agrícola peruano (Cisneros et al., 2021). No obstante, dichos elementos no puede ser considerados como *Big Data*, ni tampoco se actualizan en tiempo real.

De la misma manera que Abraham et al. (2017) logró construir una red neuronal para pronosticar las exportaciones de soja realizadas desde Brasil con el software MatLab y una hoja de cálculo que se construyó a partir de repositorios de datos; lo mismo se ha hecho para el caso del jengibre peruano, el SPSS Modeler ha resultado fundamental para construir todas las redes neuronales desarrolladas en la tesis.

# CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES

## 5.1. Conclusiones

Con los resultados obtenidos se ha alcanzado el objetivo general, se ha podido determinar que los índices globales de búsquedas *online* acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales. El mejor nivel de precisión ha sido alcanzado con la combinación de los rezagos correspondientes a un año y las búsquedas en YouTube y de imágenes. El *accuracy* del modelo final fue de 80.84%. Las búsquedas *online* a escala global, una forma de *Big Data*, contribuye a la predictibilidad de los mercados de productos perecibles.

En lo referido al primer objetivo específico, no es correcto afirmar que los índices globales de búsquedas web acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período indicado bajo el enfoque de las redes neuronales. El *accuracy* del modelo base, que incluye únicamente rezagos de un año, cae de 66.15% a 65.92% al adicionar las búsquedas web. No obstante, la relación entre predictor (búsquedas web) y *output* (exportaciones de jengibre) sí es del tipo directa y significativa si se usa el modelo de regresión lineal ( $\beta > 0$ , p-valor = 0.042)

En lo que respecta al segundo objetivo específico, es correcto afirmar que los índices globales de búsquedas en YouTube acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las

redes neuronales. El modelo *baseline* incrementó su precisión de 66.15% a 68.96% con la incorporación de los datos de YouTube. El incremento en la precisión también es verificable bajo el modelo de regresión lineal (de 66.9% a 72.9%) donde la relación entre ambas variables implicadas es significativa y directa ( $\beta > 0$ , p-valor =0.000). Cabe destacar que, en el modelo final basado en redes neuronales, las búsquedas en YouTube es el predictor de origen electrónico de mayor importancia.

En lo que respecta al tercer específico, es correcto afirmar que los índices globales de búsquedas de imágenes de jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período abarcado bajo el enfoque de las redes neuronales. Ello debido a que al *accuracy* del modelo base se incrementó de 66.15% a 68.81%) Por otro lado, las búsquedas de imágenes no contribuyeron a la predicción de las exportaciones de jengibre bajo el modelo de regresión lineal ( $\beta > 0$ , p-valor =0.067)

## 5.2. Recomendaciones

Las empresas pueden implementar el modelo de mayor precisión como parte de sus herramientas de prospectiva de corto plazo; incluso si no cuenta con herramientas que requieren el pago de licencias, bien podrían emplear soluciones gratuitas como *Orange 3*, herramienta gratuita con soporte en programación visual, opción que facilita la implementación de modelos (Bockermann, 2014). Los encargados del planeamiento comercial y de producción del jengibre fresco deben monitorear a las búsquedas *online*, no solo por el hecho de que contribuyan a la predicción de las exportaciones, sino también con el fin de identificar cambios repentinos en las tendencias de búsqueda. Y es que cualquier cambio repentino en el interés por el jengibre se reflejará primero en las búsquedas *online*.

En cuanto a las búsquedas web propiamente dichas, a pesar de no contribuir a la predictibilidad de las exportaciones, deben ser consideradas por la academia en futuras investigaciones, ya que el modelo puede ser replicado para ámbitos geográficos específicos (empleando búsquedas web en Estados Unidos acerca de *Ginger* o de *Zenzero* en Italia, por ejemplo). Otros investigadores, además, pueden emplear otras opciones de modelado como la máquina de soporte vectorial o alguna técnica econométrica propiamente dicha.

En cuanto a las búsquedas en YouTube, dada su importancia en la predictibilidad de las exportaciones de jengibre fresco, debe ser considerado como un indicador a tomar en cuenta en las publicaciones que realicen entidades estatales como PROMPERU. Ello no solo por su rol como predictor, sino también por la considerable cantidad de datos descriptivos que se pueden obtener desde *Google Trends*. Precisamente, Mayasari et al. (2020) cimentó su investigación en ello. Con *Google Trends* es posible medir en tiempo real el interés por el jengibre (evidenciado en búsquedas en YouTube) en cualquier ciudad del mundo en donde opere Google. Por otro lado, la academia también debe considerar que los comentarios en YouTube pueden ser estructurados (convertidos a patrón numérico) a través del *Text Mining*; *Google Trends* brinda información acerca de lo que los usuarios buscan en *YouTube*; sin embargo, no es posible determinar si dichas búsquedas se asocian a un interés positivo o negativo. Ello puede obtenerse analizando el contenido presente en la caja de comentarios. También resulta viable para la academia emplear otros *keywords* que potencialmente se asocien a la demanda de jengibre. El interés de las personas por el COVID-19, por ejemplo, es cuantificable también con *Google Trends*. El interés de los consumidores por la enfermedad pudo haber incidido en las exportaciones del jengibre no triturado ni pulverizado.

Las empresas exportadoras deben vigilar los índices asociados a la búsqueda de imágenes no solo porque tengan un rol predictor de las exportaciones si se combinan con las búsquedas en YouTube, sino también porque el contenido en formato de fotografía puede contener información relevante como los modos de uso. Por otro lado, la academia puede replicar el estudio para otros modelos las exportaciones de otros productos de origen peruano que tengan al elemento visual como un factor importante (las flores y determinadas prendas de ropa, por ejemplo).

## Referencias

- Aalborg, H., Molnár, P., & de Vries, J. (2019). What can explain the price, volatility and trading volume of Bitcoin? *Finance Research Letters*, 29, 255-265. <https://doi.org/10.1016/j.frl.2018.08.010>
- Aaronson, D., Brave, S., Butters, R., Fogarty, M., Sacks, D., & Seo, B. (2022). Forecasting unemployment insurance claims in realtime with Google Trends. *International Journal of Forecasting*, 38(2), 567-581. <https://doi.org/10.1016/j.ijforecast.2021.04.001>
- Abdul Mumin, Y., & Abdulai, A. (2022). Social networks, adoption of improved variety and household welfare: evidence from Ghana. *European Review of Agricultural Economics*, 49(1), 1-32. <https://doi.org/10.1093/erae/jbab007>
- Abhyankar, A., & Singla, H. (2022). Comparing predictive performance of general regression neural network (GRNN) and hedonic regression model for factors affecting housing prices in “Pune-India”. *International Journal of Housing Markets and Analysis*, 15(2), 451-477. <https://doi.org/10.1108/IJHMA-01-2021-0003>
- Abraham, E. d., Colossetti, A., Souza, A., & Tolo, R. (2017). Neural Network System to Forecast the Soybean Exportation on Brazilian Port of Santos. En H. Lödding, R. Riedel, K. Thoben, G. von Cieminski, & D. Kiritsis (Edits.), *Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing* (págs. 83-90). Springer, Cham. [https://doi.org/10.1007/978-3-319-66926-7\\_10](https://doi.org/10.1007/978-3-319-66926-7_10)
- ADEX DATA TRADE . (2022). *Reporte de Exportaciones*.
- Ahmed, F., Asif, R., Hina, S., & Muzammil, M. (2017). Financial Market Prediction using Google Trends. *International Journal of Advanced Computer Science and Applications*, 8(17). <https://doi.org/10.14569/IJACSA.2017.080752>
- Ahmed, G. (2012). Krugman Trade Theory and Developing Economies. *China-USA Business Review*, 11(12), 1557-1564.

- Ajzen. (2015). Consumer attitudes and behavior: The theory of planned behavior applied to food consumption decisions. *Italian Review of Agricultural Economics*, 70(2), 121-138. <https://doi.org/10.13128/REA-18003>
- Akman, E., Karaman, A., & Kuzey, C. (2020). Visa trial of international trade: evidence from support vector machines and neural networks. *Journal of Management Analytics*, 2020(2), 231-252. <https://doi.org/10.1080/23270012.2020.1731719>
- Alemu, H. (2010). *Analysis of the Competitiveness of Ethiopian Textile Sub-Sector Under Africa Growth and Opportunity Act (AGOA) Using Porter Diamond Model*. <https://doi.org/10.2139/ssrn.1833104>
- Allen, W. (1991). The World of Economics. En J. Eatwell, M. Milgate, & P. Newman (Edits.). London: The New Palgrave. Palgrave Macmillan.
- Alvarez, M. (2020). *Determinación de la competitividad y los factores que afectan las exportaciones de jengibre peruano (2012 -2019)*. [Trabajo de investigación para optar por el grado de bachiller, Universidad San Ignacio de Loyola]. Repositorio institucional de la USIL. <https://repositorio.usil.edu.pe/handle/usil/10663>
- Alvaro, L., & Martinez, D. (2022). *La influencia del Covid-19 en el desempeño de las exportaciones peruanas de jengibre fresco al mercado de EE.UU en 2020* [Tesis de licenciatura, Universidad Científica del Sur]. Repositorio institucional de la UCSUR. <https://hdl.handle.net/20.500.12805/2426>
- Amornpashara, N., Rompho, N., & Phadoongsitthi, M. (2015). A study of the relationship between using Instagram and purchase intention. *Journal for Global Business Advancement*, 8(3), 354-370. <https://doi.org/10.1504/JGBA.2015.071330>
- Anderson, C. (2019). Business Intelligence. En *Data Science in Practice. Studies in Big Data* (Vol. 46, págs. 97–118). Cham: Springer.
- Angulo, C., Espinoza, P., Manayay, A., & Valencia, M. (2019). *La migración a Instagram desde Facebook y Snapchat*. Universidad de Lima. Recuperado el 19 de noviembre de 2022, de [https://repositorio.ulima.edu.pe/bitstream/handle/20.500.12724/12855/La\\_migraci%C3%B3n\\_a\\_Instagram\\_desde\\_Facebook\\_y\\_Snapchat.pdf?sequence=1&isAllowed=y](https://repositorio.ulima.edu.pe/bitstream/handle/20.500.12724/12855/La_migraci%C3%B3n_a_Instagram_desde_Facebook_y_Snapchat.pdf?sequence=1&isAllowed=y)

- Arevalo, M., & Latour, A. (2021). *La participación de las empresas en ferias internacionales y su influencia positiva en la exportación de jengibre de la Región Junín periodo 2016 – 2020* [Tesis de licenciatura, Universidad Peruana de Ciencias Aplicadas]. Repositorio Académico UPC. <http://hdl.handle.net/10757/657832>
- Arroyo-Fernández, I. (2013). *Evaluación de dos técnicas de reconocimiento de patrones para su implementación en el Simulador de pilotaje automático de taller del STC Metro de la Cd. de México*. [Tesis de maestría, Universidad Tecnológica de Mixteca]. Repositorio institucional de la UTM. <http://repositorio.utm.mx:8080/jspui/handle/123456789/61>
- Ashoori, S., & Mohammadi, S. (2011). Compare failure prediction models based on feature selection technique: empirical case from Iran. *Procedia Computer Science*, 3, 568-573. <https://doi.org/10.1016/j.procs.2010.12.094>
- Bairagi, V., & Munot, M. (2019). *Research Methodology: A Practical And Scientific Approach*. Chapman and Hall/CRC.
- Baker, M., & Wurgler, J. (2007). Investor Sentiment in the Stock Market. *Journal of Economic Perspectives*, 21(2), 129-152.
- Başığit, M. (2021). Can Google Trends improve the marble demand model: A case study of USA's marble demand from Turkey. *Resource Policy*, 102073. <https://doi.org/10.1016/j.resourpol.2021.102073>
- Batarseh, F., Gopinath, M., Nalluru, G., & Beckman, J. (2019). Application of Machine Learning in Forecasting International Trade Trends. *AAAI FSS-19: Artificial Intelligence in Government and Public Secto*. Virginia. 3 de Marzo de 2022, de <https://arxiv.org/abs/1910.03112>
- BBC. (2020). Coronavirus en Perú | *El repentino boom del negocio del jengibre peruano durante la pandemia de covid-19*. BBC Mundo. Recuperado el 22 de junio del 2022, de <https://www.bbc.com/mundo/noticias-america-latina-53102394>
- Bergstrand, J. (1990). The Heckscher-Ohlin-Samuelson model, the Linder hypothesis and the determinants of bilateral intra-industry trade. *The Economic Journal*, 100(403), 1216-1229.



- Berzal, F. (2018). *Redes Neuronales & Deep Learning*. Edición independiente.
- Blanco, E. (2014). Herramientas de Big Data: ¿Podemos aprovechar Google Trends para pronosticar algunas variables macro relevantes?\*. *Asociación Argentina de Economía Política - XILX Reunión Anual*, (págs. 1-16). Buenos Aires. <https://doi.org/10.13140/RG.2.1.1514.2644>
- Blum, A. (s.f.). *Machine Learning Theory*. Carnegie Mellon University. Recuperado el 23 de julio del 2022, de <https://www.cs.cmu.edu/~avrim/Talks/mlt.pdf>
- Bocco, M., Willington, E., & Arias, M. (2010). Comparison of Regression and Neural Networks Models to Estimate Solar Radiation. *Chilean journal of agricultural research*, 70(3), 428-435. <https://doi.org/10.4067/S0718-58392010000300010>.
- Boone, T., Ganeshan, R., Hicks, R., & Sanders, N. (2018). Can Google trends improve your sales forecast? *Production and Operations Management*, 27(10), 1770-1774.
- Borup, D., & Montes, E. (2022). In Search of a Job: Forecasting Employment Growth Using Google Trends. *Journal of Business & Economic Statistics*, 40(1), 186-200. <https://doi.org/10.1080/07350015.2020.1791133>
- Bouare, O. (2009). An Evaluation of David Ricardo's Theory of Comparative Costs: Direct and Indirect Critiques. *Journal of Economic Development*, 34(1), 1-27.
- Boyd, D., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society: Provocations for a cultural, technological, and scholarly phenomenon*, 15(5), 662–679. <https://www.microsoft.com/en-us/research/publication/critical-questions-for-big-data-provocations-for-a-cultural/>
- Brenner, M. (23 de Junio de 2018). Artificial Neural Networks: What Every Marketer Should Know. *Marketing Insider Group*. Recuperado el 4 de octubre del 2021, de <https://marketinginsidergroup.com/content-marketing/artificial-neural-networks-every-marketer-know/>
- Britannica. (s.f.). *Absolute advantage*. Recuperado el 17 de enero del 2022, de [https://www.britannica.com/topic/absolute-advantage#:~:text=absolute%20advantage%2C%20economic%20concept%20that,more%20efficiently\)%20than%20another%20party.](https://www.britannica.com/topic/absolute-advantage#:~:text=absolute%20advantage%2C%20economic%20concept%20that,more%20efficiently)%20than%20another%20party.)

- Britannica, The Editors of Encyclopaedia. (13 de Diciembre de 2021). Heckscher-Ohlin theory. <https://www.britannica.com/topic/Heckscher-Ohlin-theory>
- Brockwell, P., & Davis, R. (2016). *Introduction to Time Series and Forecasting* (Tercera ed.). Springer.
- Burkov, A. (2019). *The Hundred Page - Machine Learning Book*. Andriy Burkov.
- Carbajal, A., & Ramos, J. (2020). *Factores determinantes en las exportaciones peruanas de mango a Países Bajos durante el periodo 2013 - 2018* [Tesis de licenciatura, Universidad Peruana de Ciencias Aplicadas]. Repositorio Institucional de la UPC. <http://hdl.handle.net/10757/651623>
- Cebrían, E., & Domenech, J. (2022). Is Google Trends a quality data source? *Applied Economics Letters*. <https://doi.org/10.1080/13504851.2021.2023088>
- Chang, H., & Meyerhoefer, C. (2021). COVID-19 and the Demand for Online Food Shopping Services: Empirical Evidence from Taiwan. *American Journal of Agricultural Economics*, 103(2), 448-465. <https://doi.org/10.1111/ajae.12170>
- Chicago Booth Review. (s.f.). *Adam Smith's Relevance for 1976*. Recuperado el 13 de octubre del 2021, de <https://www.chicagobooth.edu/review/adam-smiths-relevance-1976>
- Choi, H., & Varian, H. (2009). *Predicting the Present with Google Trends*. Recuperado el 2 de setiembre del 2021, de [https://static.googleusercontent.com/media/www.google.com/es//googleblogs/pdfs/google\\_predicting\\_the\\_present.pdf](https://static.googleusercontent.com/media/www.google.com/es//googleblogs/pdfs/google_predicting_the_present.pdf)
- Ciaburro, G., & Venkateswaran, B. (2017). *Neural Networks with R*. Packt Publishing.
- Ciriaco, S., & Garayar, D. (2020). *Modelamiento predictivo del engagement de fotografías en relación a sus atributos visuales en la red social Instagram para el caso de la Isla Taquile*. [Tesis de licenciatura, Universidad San Ignacio de Loyola]. Repositorio institucional de la USIL. <http://repositorio.usil.edu.pe/handle/USIL/10281>
- Cisneros, G., Lavado, C., & Estacio, R. (2021). El jengibre y su relación con el comercio exterior en la economía peruana 2010 - 2019. *Investigación Valdizana*, 15(3), 137-144. <https://doi.org/10.33554/riv.15.3.909>

- Collison, J. (2020). Performance Assessment of Google Trend Index in Forecasting Car Sales in Argentina A Case Study. *SSRN*. <https://doi.org/10.2139/ssrn.3512971>
- Costinot, A., & Donaldson, D. (2012). Ricardo's Theory of Comparative Advantage: Old Idea, New Evidence. *American Economic Review*, 102(3), 453-458.
- Cross, S., Harrison, R., & Kennedy, R. (1995). Introduction to neural networks. *The Lancet*, 346, 1075-1079.
- Daugherty, K., & Jiang, H. (2019). *Outlook for U.S. Agricultural Trade: August 2019*. Recuperado el 2 de abril del 2022, de <https://www.ers.usda.gov/publications/pub-details/?pubid=94836>
- Daugherty, T., & Hoffman, E. (2013). eWOM and the importance of capturing consumer attention within social media. *Journal of Marketing Communications*, 20(1), 82-102.
- de Castro, F., Andrade, A., dos Santos, J., Facó, J., & Goldman, S. (2020). Exploratory Study on the behavior of the Brazilian Financial Market using Google Trends. *International Journal of Advanced Engineering Research and Science*, 7(6). <http://journal-repository.com/index.php/ijaers/article/view/2066>
- De Long, J., Shleifer, A., Summers, L., & Waldmann, R. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy*, 98(4), 703-738.
- del Castillo, E., Díaz, R., & Gutiérrez, D. (2021). An Extended Model of the Theory of Planned Behaviour to Predict Local Wine Consumption Intention and Behaviour. *Foods*, 10(9), 2187. <https://doi.org/10.3390/foods10092187>
- Desoutter Industrial Tools. (s.f.). *Industrial Revolution - From Industry 1.0 to Industry 4.0*. News. Recuperado el 3 de mayo del 2022, de <https://www.desouttertools.com/industry-4-0/news/503/industrial-revolution-from-industry-1-0-to-industry-4-0>
- DeTienne, K., & DeTienne, D. (2017). Neural networks in strategic marketing: exploring the possibilities. *Journal of Strategic Marketing*, 25(4), 289-300. <https://doi.org/10.1080/0965254X.2015.1076881>
- Di, W., Sundaresan, N., Piramuthu, R., & Bhardwaj, A. (2014). Is a picture really worth a thousand words? - On the role of images in e-commerce. *Proceedings of the 7th ACM*

*international conference on Web search and data mining.*  
<https://doi.org/10.1145/2556195.2556226>

Dunn, K., Mohr, P., Wilspon, C., & Wittert, G. (2011). Determinants of fast-food consumption. An application of the Theory of Planned Behaviour. *Appetite*, 57(2), 349-357. <https://doi.org/10.1016/j.appet.2011.06.004>

Fernandes, N. (2020). *Economic Effects of Coronavirus Outbreak (COVID-19) on the World Economy*. SSRN Recuperado el 17 de mayo del 2021, de [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3557504](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3557504)

Ferreira, J. (2015). *La hoja de cálculo como herramienta para Investigación operativa* [Tesis de grado, Universidad Politécnica de Cartagena]. Repositorio digital de la UPCT. <https://repositorio.upct.es/bitstream/handle/10317/5096/tfg726.pdf?sequence=1>

Fila, S., & Smith, C. (2016). Applying the Theory of Planned Behavior to healthy eating behaviors in urban Native American youth. *International Journal of Behavioral Nutrition and Physical Activity*, 3(11). <https://doi.org/10.1186/1479-5868-3-11>

Folts, J., Lovell, R., & Zwhalen, F. (2006). *Handbook of Photography*. Pearson Learning.

Fotios, P., et al. (2020). *Forecasting: theory and practice*. *International Journal of Forecasting*. Recuperado el 4 de setiembre del 2022, de <https://www.sciencedirect.com/science/article/pii/S0169207021001758#:~:text=The%20theory%20of%20forecasting%20is,make%20predictions%20about%20the%20ofuture.>

France, S., Shi, Y., & Kazandjian, B. (2021). Web Trends: A valuable tool for business research. *Journal of Business Research*, 132, 666-679. <https://doi.org/10.1016/j.jbusres.2020.10.019>

Francia, P., & Pacheco, D. (2019). *Variables determinantes y relacionadas en las exportaciones de jengibre fresco (Partida Arancelaria 0910.11.00.00) de la República del Perú al Reino de los Países Bajos, comprendido en los años 2012 – 2019*. [Tesis de licenciatura, Universidad Peruana de Ciencias Aplicadas]. Repositorio Institucional de la UPC. <http://hdl.handle.net/10757/652587>

- Frank, R., Davey, N., & Hunt, S. (2001). Time Series Prediction and Neural Networks. *Journal of Intelligent and Robotic Systems*, 91–103. <https://doi.org/10.1023/A:1012074215150>
- FreshPlaza.com. (18 de Noviembre de 2022). *Global Market Overview Ginger*. Recuperado el 4 de diciembre del 2022, de <https://www.freshplaza.com/asia/article/9479045/global-market-overview-ginger/>
- Fuleky, P. (Ed.). (2020). *Macroeconomic Forecasting in the Era of Big Data*. Springer.
- Fürst, R., & Zündorf, I. (2015). Evidence-Based Phytotherapy in Europe: Where Do We Stand? *Planta Medica*, 81(12-13), 962-967. <https://doi.org/10.1055/s-0035-1545948>
- Gandzeichuk, I. (2020). *Video Analytics: Transforming Data Into An Asset*. Forbes. Recuperado el 13 de abril del 2022, de <https://www.forbes.com/sites/forbestechcouncil/2020/07/23/video-analytics-transforming-data-into-an-asset/?sh=7d1a301b58b5>
- Gestión. (2017a). *Kion peruano: conoce por qué lo prefieren los importadores de EE.UU.* Economía. Recuperado el 8 de junio del 2022, de <https://gestion.pe/economia/kion-peruano-conoce-prefieren-importadores-ee-uu-135228-noticia/#:~:text=La%20comercializaci%C3%B3n%20del%20kion%20o,atribuye%2C%20inform%C3%B3la%20Ocex%20del>
- Gestión. (2017b). *EE.UU. paga más por el jengibre peruano que por el importado desde China y Brasil.* Economía. Recuperado el 24 de agosto del 2022, de <https://gestion.pe/economia/ee-uu-paga-jengibre-peruano-importado-china-brasil-274336-noticia/>
- Gibrán, C. (2020). *Single-Perceptron inteligencia artificial*. NAPS. Recuperado el 2 de abril de 2022, de <https://naps.com.mx/blog/introduccion-a-la-inteligencia-artificial/single-perceptron/>
- Gilchrist, A. (2016). *Industry 4.0: The Industrial Internet of Things*. Apress.
- Golovanova, E., & Zubarev, A. (2021). Forecasting Aggregate Retail Sales with Google Trends. *Bank of Russia*, 80(4), 50-73. <https://doi.org/10.31477/rjmf.202104.50>

- Gopinath, M., Batarseh, F. A., Beckman, J., Kulkarni, A., & Jeong, S. (2021). International agricultural trade forecasting using machine learning. *Data & Policy*, 3, e1. [https://www.researchgate.net/publication/348680402\\_International\\_agricultural\\_trade\\_forecasting\\_using\\_machine\\_learning](https://www.researchgate.net/publication/348680402_International_agricultural_trade_forecasting_using_machine_learning)
- Grant, R. (1991). Porter's competitive advantage of nations: An assessment. *Strategic Management Journal*, 12(7), 535–548.
- Grzybowska, K., Awasthi, A., & Sawhney, R. (2020). *Sustainable Logistics and Production in Industry 4.0: New Opportunities and Challenges (EcoProduction)*. Springer.
- Guarin, D., & Dusseldorp, J. (2018). A Machine Learning Approach for Automated Facial Measurements in Facial Palsy. *JAMA Facial Plastic Surgery*, 20(4). <https://doi.org/10.1001/jamafacial.2018.0030>
- Gulli, A., & Pal, S. (2017). *Deep Learning with Keras: Implementing deep learning models and neural networks with the power of Python*. Packt Publishing.
- Gupta, S., & Sharma, A. (2014). Medicinal properties of Zingiber officinale - A Review. *Journal of Pharmacy and Biological Sciences*, 9(5), 124-129.
- Gupta, S., & Starr, M. (2014). *Production and Operations Management Systems*. CRC Press.
- Hagan, M., Demuth, H., Beale, M., & De Jesús, O. (2002). *Neural Network Design* (Segunda ed.).
- Hamulka, K., Jeruszka-Bielak, M., Górnicka, M., Drywień, M., & Zielinska-Pukos, M. (2020). Dietary Supplements during COVID-19 Outbreak. Results of Google Trends Analysis Supported by PLifeCOVID-19 Online Studies. *13*(1), 54. <https://doi.org/10.3390/nu13010054>.
- Hanneman, R., Kposowa, A., & Riddle, M. (2013). *Research Methods for the Social Sciences: Basic Statistics for Social Research*. Wiley.
- Hassan, M., Negm, A., Zahran, M., & Saavedra, L. (2015). Assessment of Artificial Neural Network for bathymetry estimation using High Resolution Satellite imagery in Shallow Lakes: Case Study El Burullus Lake. *Eighteenth International Water Technology Conference, IWTC18*. Sharm ElSheikh. [https://www.researchgate.net/publication/273768094\\_Assessment\\_of\\_Artificial\\_Ne](https://www.researchgate.net/publication/273768094_Assessment_of_Artificial_Ne)

ural\_Network\_for\_bathymetry\_estimation\_using\_High\_Resolution\_Satellite\_imagery\_in\_Shallow\_Lakes\_Case\_Study\_El\_Burullus\_Lake

- Hassani, H., & Silva, E. (2015). Forecasting with Big Data: A Review. *Annals of Data Science*, 5-9. <https://link.springer.com/article/10.1007/s40745-015-0029-9>
- Hematinezhad, M., Ramezaniyan, M., Gholizadeh, M., Shafiee, S., & Ghazi Zahedi, A. (2011). Predicting the success of nations in asian games using neural network. *Pamukkale Journal of Sport Sciences*, 2(3), 22-36.
- Hernandez, R., & Mendoza, C. (2018). *Metodología de la investigación*. Mc Graw Hill.
- Hiller, R., & Kim, J. (2014). Online Music, Sales Displacement, and Internet Search: Evidence from YouTube. *Center of Analysis of Property Rights an Innovation*, 13(2), 1-30. <https://www.key4biz.it/files/000263/00026398.pdf>
- Hofmann, M., & Chisholm, A. (2015). *Text Mining and Visualization: Case Studies Using Open-Source Tools*. Chapman and Hall/CRC .
- Hosni, H., & Vulpiani, A. (2018). Forecasting in Light of Big Data. *Philosophy & Technology*, 31, 557–569. <https://doi.org/10.1007/s13347-017-0265-3>
- Hu, H., Tang, L., & Wang, H. (2018). Trends, Predicting the direction of stock markets using optimized neural networks with Google. *Neurocomputing*, 188-195. <https://doi.org/10.1016/j.neucom.2018.01.038>
- Huang, M., Rojas, S., & Convery, P. (2019). Forecasting stock market movements using Google Trend searches. *Empirical Economics*. <https://doi.org/10.1007/s00181-019-01725-1>
- Hung, P. (2019). *Big Data: Applications and Uses Cases*. Springer.
- Ibarra, D. (2016). Crítica a la Teoría Clásica del Comercio Internacional, un enfoque de equilibrio general entre país grande y país pequeño Critique to the Classic Teory of the International Trade, an approach of General Balance between big country and small country. *Economía Informa*, 397, 61-79. <https://doi.org/10.1016/j.ecin.2016.03.004>
- IBM. (2017). *SPSS Modeler*. Products. Recuperado el 8 de marzo de 2022, de <https://www.ibm.com/pe-es/products/spss-modeler>

- IBM. (s.f.). *IBM SPSS Neural Networks 21*. Documentation. Recuperado el 8 de enero de 2022, de [ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/21.0/es/client/Manuals/IBM\\_SPSS\\_Neural\\_Network.pdf](ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/21.0/es/client/Manuals/IBM_SPSS_Neural_Network.pdf)
- İcan, Ö., & Buğra, T. (2017). Stock Market Prediction Performance of Neural Networks: A Literature Review. *International Journal of Economics and Finance*, 9(11), 100-104.
- Ito, T., Masuda, M., Naito, A., & Takeda, F. (2021). Application of Google Trends-based sentiment index in exchange rate prediction. *Journal of Forecasting*, 40(7), 1154-1178. <https://doi.org/10.1002/for.2762>
- Javaid, M., Haleem, A., Singh, R., & Suman, R. (2021). Substantial capabilities of robotics in enhancing industry 4.0 implementation. *Cognitive Robotics*, 1, 58-75. <https://www.sciencedirect.com/science/article/pii/S2667241321000057>
- Jomo, K., & Von Arnim, R. (2008). Trade Theory Status Quo despite Krugman. *Economic and Political Weekly*, 43(49), 29–31.
- Jun, S., Yoo, H., & Choi, S. (2018). Ten years of research change using Google Trends: From the perspective of big data utilizations and applications. *Technological Forecasting and Social Change*, 130, 69-87. <https://doi.org/10.1016/j.techfore.2017.11.009>
- Kamel, M., Sanfilippo, A., Corley, C., & Wheeler, S. (2010). Social Web mining and exploitation for serious applications: Technosocial Predictive Analytics and related technologies for public health, environmental and national security surveillance. *Computer Methods and Programs in Biomedicine*, 100(1), 16-23. <https://doi.org/10.1016/j.cmpb.2010.02.007>
- Katsioloudes, M., & Hadjidakis, S. (2017). *International Business*. Routledge.
- Kavuncuoglu, H., Kavuncuoglu, E., Karatas, S., Benli, B., Sagdic, O., & Yalcin, H. (2018). Prediction of the antimicrobial activity of walnut (*Juglans regia* L.) kernel aqueous extracts using artificial neural network and multiple linear regression. *Journal of microbiological methods*, 148, 78-86. <https://doi.org/10.1016/j.mimet.2018.04.003>



- Khademi, F., Jamal, S., Deshpande, N., & Londhe, S. (2015). Predicting strength of recycled aggregate concrete using artificial neural network, adaptive neuro-fuzzy inference system and multiple linear regression. *International Journal of Sustainable Built Environment*, 5(2), 355-369. <https://core.ac.uk/download/pdf/82746981.pdf>
- Kostoglou, G., Anagnostou, J., Ftikos, C., & Marinos, J. (1999). Development of a multiple linear regression model for predicting the 28-day compressive strength of Portland pozzolan cement. 32(98-102).
- Kotu, V., & Deshpande, B. (2014). *Predictive Analytics and Data Mining: Concepts and Practice with RapidMiner*. Morgan Kaufmann.
- Krugman, P. (1979). Increasing returns, monopolistic competition and international trade. *Journal of International Economics*, 9, 469-479.
- Kumar, S., & Sharma, A. (2014). Medicinal properties of Zingiber officinale Roscoe-A review. *J. Pharm. Biol. Sci*, 9, 124-129
- Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. New York: Springer. <https://doi.org/10.1007/978-1-4614-6849-3>
- Lao, C. (2020). *Reputación corporativa en un mercado extranjero y su relación con variables del social media - Caso del BCP Bolivia* [Tesis de licenciatura, Universidad San Ignacio de Loyola]. Repositorio institucional de la USIL. <https://repositorio.usil.edu.pe/handle/usil/10017>
- Lee, W. (2019). *Python Machine Learning*. Wiley.
- Leen, A. (2014). *Adam Smith's Policy of International Trade: Trade policy in a mercantilist world*. ResearchGate . Recuperado el 4 de julio de 2022, de [https://www.researchgate.net/publication/348190040\\_Adam\\_Smith's\\_Policy\\_of\\_International\\_Trade\\_Trade\\_policy\\_in\\_a\\_mercantilist\\_world](https://www.researchgate.net/publication/348190040_Adam_Smith's_Policy_of_International_Trade_Trade_policy_in_a_mercantilist_world)
- Lefkowitz, M. (2019). *Professor's perceptron paved the way for AI – 60 years too soon*. Cornell Chronicle. Recuperado el 15 de diciembre de 2021, de <https://news.cornell.edu/stories/2019/09/professors-perceptron-paved-way-ai-60-years-too-soon>

- León, P., & Villar, O. (2018). *Estudio exploratorio del uso de indicadores de Big Data de Google como variable en el modelamiento del flujo mensual de visitantes al Complejo Arqueológico de Kuélap en el periodo 2011-2016* [Tesis de licenciatura, Universidad San Ignacio de Loyola]. Repositorio institucional de la USIL. <http://repositorio.usil.edu.pe/handle/USIL/3694>
- Leukel, J., González, J., & Riekert, M. (2021). Adoption of machine learning technology for failure prediction in industrial maintenance: A systematic review. *Journal of Manufacturing Systems*, 61, 87-96. <https://doi.org/10.1016/j.jmsy.2021.08.012>
- Liu, Y., Peng, G., Hu, L., Dong, J., & Zhang, Q. (2019). Using Google Trends and Baidu Index to analyze the impacts of disaster events on company stock prices. *Industrial Management & Data Systems*, 120(2), 350-365. <https://doi.org/10.1108/IMDS-03-2019-0190>
- Locharoenrat, K. (2017). *Research Methodologies for Beginners*. Pan Stanford.
- Lundahl, M. (2022). The Inspiration for the Heckscher–Ohlin Theorem. En *Twelve Figures in Swedish Economics*. Cham: Palgrave Macmillan. [https://doi.org/10.1007/978-3-030-94327-1\\_2](https://doi.org/10.1007/978-3-030-94327-1_2)
- Ma, J., Seenivasan, S., & Yan, B. (2020). Media influences on consumption trends: Effects of the film Food, Inc. on organic food sales in the U.S. *International Journal of Research in Marketing*, 37(2), 320-335.
- Makridakis, S., Hogarth, R., & Gaba, A. (2010). Why Forecasts Fail. What to Do Instead. *MIT Sloan Management Review*. <https://sloanreview.mit.edu/article/why-forecasts-fail-what-to-do-instead/>
- Malhotra, N. (2008). *Investigación de mercados* (Quinta ed.). Pearson - Prentice Hall.
- Mao, Q., Xu, X., Cao, S., Gan, R., Corke, H., Beta, T., & Li, H. (2019). Bioactive Compounds and Bioactivities of Ginger (*Zingiber officinale* Roscoe). *Foods*, 8(6), 185. <https://doi.org/10.3390/foods8060185>
- Marconi, K., & Lehman, H. (2014). *Big Data and health Analytics*. CRC Press.
- Marr, B. (2018). *What is Industry 4.0? Here's A Super Easy Explanation For Anyone*. Forbes. Recuperado el 13 de junio de 2022, de

<https://www.forbes.com/sites/bernardmarr/2018/09/02/what-is-industry-4-0-heres-a-super-easy-explanation-for-anyone/?sh=17b9959c9788>

- Masih, J., Verbeke, W., Deutsch, J., Sharma, A. S., Rajkumar, R., & Matharu, P. (2019). Big data study for gluten-free foods in India and USA using online reviews and social media. *Agricultural Sciences*, *10*(3), 302-320.
- Mayasari, N., Ho, D., Lundy, D., Skalny, A., Tinkov, A., Teng, I., . . . Chang, J. (2020). Pandemic on Food Security and Diet-Related Lifestyle Behaviors: An Analytical Study of Google Trends-Based Query Volumes. *Nutrients*, 3103.
- MBA Knowledge Base. (s.f.). *National Competitive Advantage Theory of International Trade – Porters Diamond Model*. International Finance. Recuperado el 13 de abril de 2022, de <https://www.mbaknol.com/international-finance/national-competitive-advantage-theory-of-international-trade/>
- McPherson, M., Redfearn, M., & Tieslau, M. (2001). International trade and developing countries: an empirical investigation of the Linder hypothesis. *Applied Economics*, *33*(5), 649-657. <https://doi.org/10.1080/00036840122575>
- Medrano, L., & Rodríguez, P. (2020). *Estrategias competitivas empleadas por las empresas exportadoras de jengibre de la región Junín relacionadas con el proceso de internacionalización hacia el mercado de Estados Unidos (Periodo 2014-2018)*. [Tesis de licenciatura, Universidad Peruana de Ciencias Aplicadas]. Repositorio académico UPC. <https://repositorioacademico.upc.edu.pe/handle/10757/652464>
- Melitz, M. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, *71*(6), 1695-1725.
- Mijwil, M. (2018). *Artificial Neural Networks Advantages and Disadvantages*. LinkedIn. Recuperado el 21 de mayo de 2022, de <https://www.linkedin.com/pulse/artificial-neural-networks-advantages-disadvantages-maad-m-mijwel/>
- Miller, T. (2015). *Marketing Data Science: Modeling Techniques in Predictive Analytics with R and Python*. PH Professional Business.
- Montgomery, D., Peck, E., & Vining, G. (2012). *Introduction to Linear Regression* (Quinta ed.). Wiley.

- Montoya, N. (2020). Using Google data to understand governments' approval in Latin America. *OECD Development Centre Working Papers*, 343. <https://doi.org/10.1787/89ed5e8f-en>.
- Moon, S., Bae, S., & Kim, S. (2014). Predicting the near-weekend ticket sales using web-based external factors and box-office data. In *2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)* (Vol. 2, pp. 312-318). IEEE.
- Mun, T. (1964). *England's treasure by foreign trade*. New York: Macmillan and Co.
- Nooraeni, R., Nickelson, J., Rahmadian, E., & Yudho, N. (2022). New recommendation to predict export value using big data and machine learning technique. *Statistical Journal of the IAOS*, 38(1), 277-290.
- Oldon, D., & Wu, D. (2020). *Predictive Data Mining Models* (Segunda ed.). Springer.
- Oliveira, B., & Casais, B. (2019). The importance of user-generated photos in restaurant selection. *Journal of Hospitality and Tourism Technology*, 10(1), 2-14.
- Onomayo, A. (2022). Factors influencing export performance of Ginger (*Zingiber officinale*) in Nigeria. *International Journal of Agriculture, Environment and Food Sciences*, 6(3), 370-377.
- Oracle España. (s.f.). *¿Qué es big data?* Big Data. Recuperado el 3 de junio de 2022, de <https://www.oracle.com/es/big-data/what-is-big-data.html>
- Oskolkov, N. (2021). *Univariate vs. Multivariate Prediction*. Toward Data Science. Recuperado el 25 de noviembre de 2022, de [towardsdatascience.com/univariate-vs-multivariate-prediction-c1a6fb3e009](https://towardsdatascience.com/univariate-vs-multivariate-prediction-c1a6fb3e009)
- Oxford Reference. (s.f.). *Overview Mercantilism*. <https://www.oxfordreference.com/view/10.1093/oi/authority.20110803100150843#:~:text=In%20a%20celebrated%20essay%20of,way%20by%20government%20interference%20with>
- Pakhare, A., Bali, S., & Kalra, G. (2013). Use of mobile phones as research instrument for data collection. *Indian Journal of Community Health*, 25(2), 95-98.

- Park, Y. (2015). Does the Rise of the Korean Wave Lead to Cosmetics Export? *Journal of Asian Finance, Economics and Business*, 2(4), 13-20. <https://www.koreascience.or.kr/article/JAKO201516863230395.page>
- Paul, J (2008) *International Business* Prentice-Hall of India
- Peterson, S., & Fabozzi, F. (2010). *The Basics of Finance: An Introduction to Financial Markets, Business Finance, and Portfolio Management*. Wiley.
- Petropoulos, A. (2021). Employing Google Trends and Deep Learning in Forecasting Financial Market Turbulence. *Journal of Behavioral Finance*. <https://doi.org/10.1080/15427560.2021.1913160>
- Pilco, J., & Sandoval, K. (2018). *Google Trends, inclusión de datos en tiempo real para la predicción de variables macroeconómicas en el Perú* [Tesis de licenciatura , Universidad Católica Santo Toribio de Mogrovejo]. Repositorio de tesis USAT. <http://hdl.handle.net/20.500.12423/1518>
- Porter, M. (1990). *The competitive advantage of nations*. Free Press.
- Pudaruth, S. (2014). Predicting the Price of Used Cars using Machine Learning Techniques. *International Journal of Information & Computation Technology*, 4(7), 753-764.
- Puente, A., & Silva, A. (2020). *Factores que contribuyen al incremento de la exportación de jengibre fresco producido en la región Junín para el periodo 2012-2018* [Tesis de licenciatura, Universidad Peruana de Ciencias Aplicadas]. Repositorio académico UPC. <https://repositorioacademico.upc.edu.pe/handle/10757/651575>
- Purkait, N. (2019). *Hands-On Neural Networks with Keras: Design and create neural networks using deep learning and artificial intelligence principles*. Packt Publishing.
- Rajabi Vandechali, M., Abbaspour-Fard, M., & Rohani, A. (2017). Multiple linear regression analysis of some important performance parameters of a conventional diesel engine in different working conditions. *Modares Mechanical Engineering*, 17(5), 363-373.
- Ramirez, H., & Vilela, M. (2020). *Modelado del valor de las acciones de Graña y Montero en la NYSE con indicadores del interés hacia el caso Odebrecht*. [Tesis de

licenciatura, Universidad San Ignacio de Loyola]. Repositorio institucional de la USIL. <https://repositorio.usil.edu.pe/handle/usil/10497>

- Ranjan, P., & Raychaudhuri, J. (2016). The “New-New” Trade Theory: A Review of the Literature. En M. Roy, & S. Sinha Roy (Edits.), *International Trade and International Finance* (págs. 3-21). Springer. [https://doi.org/10.1007/978-81-322-2797-7\\_1](https://doi.org/10.1007/978-81-322-2797-7_1)
- Ravindran, P., & Babu, N. (2004). *Ginger: : the genus Zingiber*. CRC press.
- Rebala, G., Ravi, A., & Churiwala, S. (2019). An Introduction to Machine Learning. <https://doi.org/10.1007/978-3-030-15729-6>
- Reece, A., & Danforth, C. (2016). Instagram photos reveal predictive markers of depression. *EPJ Data Science*, 6, 15-21.
- Ren, M., Vu, H., Li, G., & Law, R. (2020). Large-scale comparative analyses of hotel photo content posted by managers and customers to review platforms based on deep learning: implications for hospitality marketers. *Journal of Hospitality Marketing & Management*. <https://doi.org/10.1080/19368623.2020.1765226>
- Ricardo, D. (1817). On the Principles of Political Economy and Taxation. (P. Sraffa, Ed.)
- Rosas-Romero, R., & Medina-Ochoa, J. (2019). Learning Financial Time Series for Prediction of the Stock Exchange Market. *EPiC Series in Computing*, 58, 418-427.
- Rovetta, A. (2021). Reliability of Google Trends: Analysis of the Limits and Potential of Web Infoveillance During COVID-19 Pandemic and for Future Research. *Frontiers in research metrics and analytics*, 6(1), 670226. <https://doi.org/10.3389/frma.2021.670226>
- Rush, V. (2014). *Planned Behavior: Theory, Applications and Perspectives*. Nova Science Publishers.
- Salisu, A., Ogbonna, A., & Adewuyi, A. (2020). Google trends and the predictability of precious metals. *Resources Policy*, 65(101542). <https://doi.org/10.1016/j.resourpol.2019.101542>

- Sama, R. (2019). Impact of Media Advertisements on Consumer Behaviour. *Journal of Creative Communications*, 14(1), 54-68.  
<https://doi.org/10.1177/0973258618822624>
- Sánchez, H., & Reyes, C. (2015). *Metodología y Diseños en la Investigación Científica* (Quinta ed.). Business Suport.
- Saputra, D., & Kom, M. (2017). *AIDMA Model & AISAS model in Digital Marketing Strategy*.<https://bbs.binus.ac.id/ibm/2017/06/aidma-model-aisas-model-in-digital-marketing-strategy/>
- Scheffold, M. (2016). Advantages and disadvantages of Google Trends. <https://scales-course.cen.uni-hamburg.de/main-script/1-the-scales-concept/14-third-axes/141-google-trends-search-interest-example-for-a-third-axis/1411-advantages-and-disadvantages-of-google-trends/>
- Seyal, A., Noah, M., & Rahman, A. (2017). *Theory of Planned Behavior: New Research*. Nova Science Publishers.
- Sharma, Y. (2017). Ginger (Zingiber officinale)-An elixir of life a review.
- Snell, N. (1995). *Internet ¿Qué hay que saber?* Madrid: Sams Publishing.
- Sosa, Y., & Jeri, J. (2019). *Uso de indicadores Big Data para mejorar el nivel de ajuste de un modelo autorregresivo de arribos domésticos al Aeropuerto Internacional Jorge Chávez* [Tesis de licenciatura, Universidad San Ignacio de Loyola].Repositorio institucional de la USIL.  
[http://repositorio.usil.edu.pe/bitstream/USIL/8765/1/2019\\_Jeri-Jong.pdf](http://repositorio.usil.edu.pe/bitstream/USIL/8765/1/2019_Jeri-Jong.pdf)
- Srinivasan, K. (2017). Ginger rhizomes (Zingiber officinale): A spice with multiple health beneficial potentials. *PharmaNutrition*, 5(1), 18-28.  
<https://doi.org/10.1016/j.phanu.2017.01.001>
- StatCounter. (2022). *Search Engine Market Share Worldwide*. Search Engine Market Share Worldwide. Recuperado el 2 de abril del 2022, de <https://gs.statcounter.com/search-engine-market-share>
- Stearns, P. (2012). *The Industrial Revolution in World History* (Cuarta ed.). Routledge.

- Stibel, J. (2009). *Why We Can't Predict Financial Markets*. <https://hbr.org/2009/01/why-we-cant-predict-financial>
- Stirling, A. (2004). The causes of poor establishment of ginger (*Zingiber officinale*) in Queensland, Australia. *Australasian Plant Pathology volume, 33*, 203–210. <https://doi.org/10.1071/AP04003>
- Sugiyama, K., & Andree, T. (2011). *The Dentsu Way: Secrets of Cross Switch Marketing from the World's Most Innovative Advertising Agency*. McGraw Hill.
- Teng, S., Khong, K., Sharif, S. P., & Ahmed, A. (2020). YouTube Video Comments on Healthy Eating: Descriptive and Predictive Analysis. *JMIR public health and surveillance, 6*(4), e19618.
- The Economist. (4 de Marzo de 2020). *Tracking the economic impact of covid-19 in real time*. United States. Recuperado el 14 de Noviembre del 2022, de <https://www.economist.com/united-states/2020/03/14/tracking-the-economic-impact-of-covid-19-in-real-time>
- The World Bank. (2022). *Russian Invasion to Shrink Ukraine Economy by 45 Percent this Year*. Press Release. Recuperado el 3 de diciembre del 2022, de <https://www.worldbank.org/en/news/press-release/2022/04/10/russian-invasion-to-shrink-ukraine-economy-by-45-percent-this-year>
- Torrecillas, M., & Jareño, F. (2013). Inflation News Impact on Stock Market: A Review. *Pensée, 75*(11), 414-419.
- Tosun, E., Aydin, K., & Bilgili, M. (2016). Comparison of linear regression and artificial neural network model of a diesel engine fueled with biodiesel-alcohol mixtures. *Alexandria Engineering Journal, 55*(4), 3081-3089.
- Trade Map. (2022a). *Exportaciones de jengibre fresco a nivel global 2012-2021*, Recuperado el 3 de diciembre del 2022, de <https://www.trademap.org/Index.aspx?lang=es>
- Trade Map. (2022b). *Exportaciones peruanas de jengibre fresco 2012-2021*. Recuperado el 3 de diciembre del 2022, de <https://www.trademap.org/Index.aspx?lang=es>



- Vallejos, J., & McKinnon, S. (2013). Logistic regression and neural network classification of seismic records. *International Journal of Rock Mechanics and Mining Sciences*, 62, 86-95. <https://doi.org/10.1016/j.ijrmms.2013.04.005>
- Van Gerven, M., & Bohte, S. (2017). Editorial: Artificial Neural Networks as Models of Neural Information Processing. *Frontiers in Computational Neuroscience*, 19. <https://doi.org/10.3389/fncom.2017.00114>
- Vargas, X., & Yamanaka, T. (2020). Food Communication and its Related Sentiment in Local and Organic Food Videos on YouTube. *Journal of Medical Internet Research*, 22(8), :e16761. <https://doi.org/10.2196/16761>
- Varian, H. (2014). Big Data: New Tricks for Econometrics. *The Journal of Economic Perspectives*, 28(2), 3-27.
- Vaughan, L., & Chen, Y. (2014). Data Mining From Web Search Queries: A Comparison of Google Trends and Baidu Index. *Journal of the Association for Information Science and Technology*, 66(1). <https://doi.org/10.1002/asi.23201>
- Veritrade (2022). *Reporte de las exportaciones peruanas de jengibre (2014-2021)*. Recuperado el 3 de diciembre del 2022 de <https://veritra.upc.elogim.com/es/mis-busquedas>
- Venery, G., & Capasso, A. (2018). *Hands-On Industrial Internet of Things: Create a powerful Industrial IoT infrastructure using Industry 4.0*. Packt Publishing.
- Vernon, R. (1966). International investment and international trade in the product life cycle. *Quarterly Journal of Economics*, 80(2), 190-207.
- Verter, N. (2015). The Application of International Trade Theories to Agriculture. *Mediterranean Journal of Social Sciences*, 6(6), 209-219. <https://doi.org/10.5901/mjss.2015.v6n6s4p209>
- Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: survey-based indicators vs. Google trends. *Journal of Forecasting*, 30(6), 565-578. <https://doi.org/10.1002/for.1213>

- Vu, J. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, 49(11), 1225-1231. [https://doi.org/10.1016/s0895-4356\(96\)00002-9](https://doi.org/10.1016/s0895-4356(96)00002-9)
- Wendler, T., & Gröttrup, S. (2016). *Data Mining with SPSS Modeler, Theory, Exercises and Solutions*. Springer.
- Woo, J., & Owen, A. (2018). Forecasting private consumption with Google Trends data. *Journal of Forecasting*, 38(2), 81-91. <https://doi.org/10.1002/for.2559>
- Yin, Y., Stecke, K., & Li, D. (2018). The evolution of production systems from Industry 2.0 through Industry 4.0. *International Journal of Production Research*, 56(1-2), 848-861. <https://doi.org/10.1080/00207543.2017.1403664>
- Yu, L., Zhao, Y., Tang, L., & Yang, Z. (2019). Online big data-driven oil consumption forecasting with Google trends. *International Journal of Forecasting*, 35(1), 213-223. <https://doi.org/10.1016/j.ijforecast.2017.11.005>
- Zelenka, J., & Hruška, J. (2018). Ways and effectiveness of social media utilization by airlines. *Tourism*, 66(2), 227-238.
- Zicha, J. (2021). *US ginger prices fall amid plentiful supply in China*. Mintec. Recuperado el 8 de junio del 2022, de <https://www.mintecglobal.com/top-stories/us-ginger-prices-fall-amid-plentiful-supply-in-china>

**Anexo 1** Datos de investigación

Año	Mes	Expo	Expo-12	Web	YouTube	Imágenes
2013	Ene.	58583.20	107380.00	69	50	70
2013	Feb.	47183.20	119617.80	57	33	86
2013	Mar.	73299.20	115329.05	50	23	58
2013	Abr.	39343.10	163562.20	52	35	74
2013	May.	149735.72	146356.00	52	39	75
2013	Jun.	300392.40	181254.20	55	42	70
2013	Jul.	385859.80	275831.20	53	37	62
2013	Ago.	394822.31	205264.40	51	45	63
2013	Set.	422739.77	190360.69	55	38	54
2013	Oct.	446673.95	146752.20	54	28	57
2013	Nov.	573731.20	122869.10	57	40	78
2013	Dic.	538980.00	129032.80	65	40	71
2014	Ene.	441996.48	58583.20	71	62	60
2014	Feb.	489670.33	47183.20	71	60	57
2014	Mar.	440482.88	73299.20	60	57	53
2014	Abr.	198364.60	39343.10	58	29	64
2014	May.	274201.00	149735.72	59	49	57
2014	Jun.	635661.87	300392.40	56	36	59
2014	Jul.	879908.82	385859.80	56	45	51
2014	Ago.	1561369.57	394822.31	55	62	51
2014	Set.	1820917.60	422739.77	58	35	71
2014	Oct.	2043096.73	446673.95	58	44	52
2014	Nov.	1951034.96	573731.20	64	45	78
2014	Dic.	1554706.60	538980.00	73	74	76
2015	Ene.	460666.54	441996.48	80	44	64
2015	Feb.	473339.40	489670.33	73	46	73
2015	Mar.	305000.80	440482.88	67	30	61
2015	Abr.	185425.50	198364.60	67	21	67
2015	May.	191313.18	274201.00	59	29	61
2015	Jun.	631251.10	635661.87	58	44	62
2015	Jul.	1409613.55	879908.82	59	22	60
2015	Ago.	1791966.72	1561369.57	60	49	60
2015	Set.	1401514.82	1820917.60	59	21	50
2015	Oct.	1527172.43	2043096.73	63	26	66
2015	Nov.	1104374.31	1951034.96	69	41	76
2015	Dic.	1153562.73	1554706.60	77	32	64

2016	Ene.	786475.86	460666.54	89	37	85
2016	Feb.	769565.91	473339.40	76	40	57
2016	Mar.	806520.28	305000.80	73	41	69
2016	Abr.	542739.80	185425.50	66	31	63
2016	May.	219814.40	191313.18	66	35	67
2016	Jun.	533962.58	631251.10	65	25	64
2016	Jul.	1442300.47	1409613.55	64	40	53
2016	Ago.	1743543.49	1791966.72	64	17	74
2016	Set.	2117163.81	1401514.82	59	30	65
2016	Oct.	1649136.25	1527172.43	68	19	63
2016	Nov.	1909067.89	1104374.31	73	41	75
2016	Dic.	2054337.20	1153562.73	77	56	79
2017	Ene.	1263650.49	786475.86	90	56	74
2017	Feb.	1928779.03	769565.91	73	39	78
2017	Mar.	1888558.09	806520.28	73	33	73
2017	Abr.	1308688.70	542739.80	69	45	58
2017	May.	1315297.68	219814.40	66	59	72
2017	Jun.	887504.98	533962.58	64	26	53
2017	Jul.	1174465.36	1442300.47	61	35	57
2017	Ago.	2086103.08	1743543.49	64	46	59
2017	Set.	3458142.35	2117163.81	64	60	58
2017	Oct.	4158485.26	1649136.25	68	54	68
2017	Nov.	3752479.12	1909067.89	72	68	79
2017	Dic.	3232887.20	2054337.20	80	62	67
2018	Ene.	1884857.09	1263650.49	92	79	86
2018	Feb.	2306894.54	1928779.03	76	61	71
2018	Mar.	1647108.40	1888558.09	73	59	63
2018	Abr.	1036927.21	1308688.70	68	53	56
2018	May.	772964.17	1315297.68	64	56	59
2018	Jun.	770000.72	887504.98	61	51	68
2018	Jul.	1820086.03	1174465.36	60	50	63
2018	Ago.	3068554.45	2086103.08	59	53	67
2018	Set.	2213442.49	3458142.35	66	55	53
2018	Oct.	1737180.67	4158485.26	67	58	77
2018	Nov.	1971054.87	3752479.12	69	60	74
2018	Dic.	1642120.43	3232887.20	78	66	67
2019	Ene.	824977.95	1884857.09	89	77	90
2019	Feb.	811061.27	2306894.54	75	67	73

2019	Mar.	731363.50	1647108.40	69	61	73
2019	Abr.	463379.09	1036927.21	66	56	75
2019	May.	897006.80	772964.17	61	49	76
2019	Jun.	1963367.44	770000.72	62	52	73
2019	Jul.	2846187.60	1820086.03	64	49	72
2019	Ago.	2971026.85	3068554.45	61	61	72
2019	Set.	2629721.48	2213442.49	63	86	63
2019	Oct.	2786206.25	1737180.67	63	78	63
2019	Nov.	3633426.04	1971054.87	69	62	86
2019	Dic.	2590344.86	1642120.43	77	81	92
2020	Ene.	2214269.72	824977.95	89	81	97
2020	Feb.	2696280.33	811061.27	70	68	78
2020	Mar.	2811238.27	731363.50	71	68	51
2020	Abr.	1247633.89	463379.09	76	76	72
2020	May.	2060234.72	897006.80	86	96	72
2020	Jun.	4174431.27	1963367.44	80	88	76
2020	Jul.	3808510.52	2846187.60	77	95	64
2020	Ago.	5316362.43	2971026.85	77	92	79
2020	Set.	5843012.97	2629721.48	72	86	86
2020	Oct.	6975669.84	2786206.25	71	90	67
2020	Nov.	7380772.85	3633426.04	75	93	82
2020	Dic.	5574594.30	2590344.86	83	90	67
2021	Ene.	2599799.14	2214269.72	100	100	100
2021	Feb.	4306763.41	2696280.33	82	88	72
2021	Mar.	2838486.89	2811238.27	72	83	74
2021	Abr.	1392338.80	1247633.89	68	80	64
2021	May.	1306090.92	2060234.72	65	70	60
2021	Jun.	3084688.42	4174431.27	65	81	62
2021	Jul.	5173254.41	3808510.52	67	80	65
2021	Ago.	4990350.33	5316362.43	70	71	73
2021	Set.	5392352.56	5843012.97	67	75	75
2021	Oct.	7985313.36	6975669.84	70	74	66
2021	Nov.	8120928.11	7380772.85	71	72	88
2021	Dic.	7477039.67	5574594.30	81	72	84

## Anexo 2 Matriz de consistencia

Problema	Objetivos	Hipótesis	Variable	Dimensiones	Metodología
PG: ¿Es correcto afirmar que los índices globales de búsquedas online acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales?	OG: Determinar si es correcto afirmar que índices globales de búsquedas online acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales	HG: Los índices globales de búsquedas online acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.	Y: Exportaciones peruanas de jengibre no triturado ni pulverizado al resto del mundo.	Unidimensional (expresado en peso neto)	Cuantitativa No experimental Correlacional Predictiva
PE1: ¿Es correcto afirmar que los índices globales de búsquedas web acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales?	OE1: Determinar si es correcto afirmar que los índices globales de búsquedas web acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.	H1: Los índices globales de búsquedas web acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales	Búsquedas globales online acerca del jengibre (ZINGIBER OFFICINALE)	X1: Búsquedas web	<b>Muestra</b> 108 series mensuales de enero del 2013 al 2021
PE2: ¿Es correcto afirmar que los índices globales de búsquedas en YouTube acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales?	OE2: Determinar si es correcto afirmar que los índices globales de búsquedas en YouTube acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.	H2: Los índices globales de búsquedas en YouTube acerca del jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.		X2: Búsquedas en YouTube	Plan de análisis Redes neuronales Regresión Lineal
PE3: ¿Es correcto afirmar que los índices globales de búsquedas de imágenes de jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales?	OE3: Determinar si es correcto afirmar que los índices globales de búsquedas de imágenes de jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.	H3: Los índices globales de búsquedas de imágenes de jengibre sí llegaron a constituirse como predictores de las exportaciones mensuales peruanas de la versión no triturada ni pulverizada de dicho producto al resto del mundo durante el período 2013-2021 bajo el enfoque de las redes neuronales.		X3: Búsquedas de imágenes	

Elaboración propia