



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

UNIVERSIDAD DE SALAMANCA

Departamento de Informática y Automática
Facultad de Ciencias

TESIS DOCTORAL

Arquitectura Fog Computing para entornos FinTech

Autor:

Dña. Elena Hernández Nieves

Directores:

Dr. D. Juan Manuel Corchado Rodríguez

Dra. Dña. Sara Rodríguez González

Octubre de 2021



VNiVERSIDAD
D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL

UNIVERSITY OF SALAMANCA

Department of Computer Science and Automation
Faculty of Sciences

DOCTORAL THESIS

**Fog Computing Architecture for FinTech
environments**

Author:

Dña. Elena Hernández Nieves

Advisors:

Dr. D. Juan Manuel Corchado Rodríguez

Dr. Dña. Sara Rodríguez González

October, 2021



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

UNIVERSIDAD DE SALAMANCA

Departamento de Informática y Automática

Facultad de Ciencias

Arquitectura Fog Computing para entornos FinTech

Autor:

Dña. Elena Hernández Nieves

Directores:

Dr. D. Juan Manuel Corchado Rodríguez

Dra. Dña. Sara Rodríguez González

TRIBUNAL

Presidente:

Vocal:

Secretario:

Suplentes:

FECHA DE LECTURA:

CALIFICACIÓN:

Autorización del Director

En Salamanca, a 16 de septiembre de 2021.

HACEMOS CONSTAR:

Que como directores de la Tesis Doctoral de Elena Hernández Nieves, con D.N.I.: 70889719Q, autorizamos a presentar la tesis doctoral Arquitectura Fog Computing para entornos FinTech, mediante la modalidad de compendio de artículos al disponer de los siguientes artículos publicados:

1. Hernández-Nieves, E., Hernández, G., Gil-González, A. B., Rodríguez-González, S., & Corchado, J. M. (2020). Fog computing architecture for personalized recommendation of banking products. *Expert Systems with Applications*, 140, 112900.
2. Hernández-Nieves, E., Hernández, G., Gil-González, A. B., Rodríguez-González, S., & Corchado, J. M. (2021). CasE: A CasE-Based Reasoning Application to recommend banking products. *Engineering Applications of Artificial Intelligence*, <https://doi.org/10.1016/j.engappai.2021.104327>.
3. Hernández-Nieves, E., Parra-Domínguez, J., Chamoso, P., Rodríguez-González, S., & Corchado, J. M. (2021). A Data Mining and Analysis Platform for Investment Recommendations. *Electronics*, 10(7), 859.

Los directores:

Dr. Juan Manuel Corchado Rodríguez

Dra. Sara Rodríguez González

“No todos los que vagan estan perdidos”

Frase de “El Hobbit” de J.R.R. Tolkien (1892-1973)

Resumen

El sector financiero ha sido uno de los primeros en aceptar y utilizar los nuevos avances de las TIC (Tecnologías de la Información y la Comunicación). Hay que tener en cuenta que, con la aparición de nuevos competidores en un entorno económico que cambia rápidamente, junto con unos requisitos normativos exigentes y costosos, los proveedores de servicios bancarios y financieros de hoy en día se enfrentan a una transformación sustancial: las finanzas digitales.

Los modelos de negocio tradicionales, basados en la tienda física y que no hacen uso de las TIC para llevar a cabo las campañas de marketing o ventas, están en desventaja ante una clientela más informada tecnológicamente. En este contexto, la mayoría de las partes interesadas en el sector intentan ofrecer productos y servicios financieros personalizados, inteligentes y versátiles, al tiempo que persiguen nuevos niveles de orientación hacia el cliente.

La tesis doctoral que se presenta viene motivada por los nuevos retos del sector bancario que han surgido a raíz de la incorporación de las TIC, en los que se busca satisfacer las necesidades de los clientes a través de los servicios que ofrecen, es por ello que esta tesis propone una solución tecnológica que consiste en el diseño de una arquitectura global que integra tecnologías *Fog Computing* (FC), OVA (Organizaciones Virtuales de Agentes) y sistemas CBR (*Case-Based Reasoning*) que permita su implementación en entornos bancarios.

La arquitectura propuesta en esta investigación se denomina FOBA (*Fog Oriented Banking Architecture*) y propone mejorar aspectos de los servicios de atención al cliente de una entidad bancaria, especialmente, conseguir una mayor seguridad, mayor transparencia y agilidad de los procesos, así como reducir los costes de gestión de la entidad. La arquitectura presentada incluye nodos de niebla donde los datos son procesados por agentes inteligentes ligeros que permiten la implementación de sistemas de recomendación.

En este trabajo se presenta la descripción conceptual y técnica de la arquitectura propuesta, así como el componente experimental del estudio, que consistió en el desarrollo de dos herramientas de recomendación: la primera herramienta diseñada y desarrollada fue la creación del CBR (*Case-Based Reasoning*) denominado CEBRA (*CasE-Based Reasoning Application*), como resultado de la investigación, se presenta el sistema de recomendación, incluyendo, el algoritmo de recomendación y una interfaz

REST (*REpresentational State Transfer*) para su uso. Las recomendaciones que realizan se basan en el perfil del usuario, en valoraciones anteriores y/o en conocimientos adicionales como la información contextual del usuario. La segunda herramienta diseñada y desarrollada es un sistema recomendador del que se obtienen señales de compra/venta a partir de los resultados de los análisis técnicos y de las previsiones realizadas para empresas que operan en el mercado continuo español. Tiene un diseño modular para facilitar la escalabilidad del modelo y la mejora de las funcionalidades.

Abstract

The financial sector has been one of the first to embrace and use new ICT (Information and Communication Technology) developments. It should be noted that with the emergence of new competitors in a rapidly changing economic environment, coupled with demanding and costly regulatory requirements, today's banking and financial service providers are facing a substantial transformation: digital finance.

Traditional, physical store-based business models, which do not make use of ICT for marketing or sales campaigns, are at a disadvantage in the face of a more technologically informed customer base. In this context, most industry stakeholders are seeking to offer personalised, smart and versatile financial products and services, while pursuing new levels of customer orientation.

The doctoral thesis presented here is motivated by the new challenges that have arisen in the banking sector as a result of the incorporation of ICT. Its aim is to find a solution for the optimal digital transformation of banks and of the services they offer, so as to fully satisfy customer needs. Hence, this thesis proposes a global architecture, to be implemented in banking environments, which integrates FC (Fog Computing) technologies, VAO (Virtual Agent Organisations) and CBR (Case-Based Reasoning) systems.

The architecture proposed in this research is called FOBA (Fog Oriented Banking Architecture) and it offers improvements in relation to customer services in banks, in particular, achieving greater security, greater transparency and agility of processes, as well as reducing bank management costs. The presented architecture includes fog nodes where data is processed by intelligent agents that enable the implementation of recommender systems.

This research presents the conceptual and technical description of the proposed architecture, as well as the experimental part of the study, which consisted of the development of two recommendation tools: the first of the designed and developed tools was a CBR (Case-Based Reasoning) system called CEBRA (CasE-Based Reasoning Application), as a result of this research is provided a Recommender System, including a recommendation algorithm and a REST (REpresentational State Transfer) interface for its use. The recommendations are based on the user's profile, previous ratings and/or additional knowledge such as the user's contextual information. The second of the designed and developed tools is a recommender system from which buy/sell signals

are obtained on the basis of the results of technical analyses and forecasts made for companies operating on the Spanish continuous market. It has a modular design to facilitate the scalability of the model and the improvement of its functionalities.

Agradecimientos

Me gustaría comenzar agradeciendo a mis directores: Juan Manuel Corchado Rodríguez y Sara Rodríguez González. A Juan Manuel, agradecerle el haberme ofrecido la oportunidad que a día de hoy me ha llevado a finalizar esta etapa en el camino en forma de tesis doctoral, sin él no podría haberlo logrado. A Sara, agradecerle el indicarme el camino a seguir, su guía y apoyo a lo largo de estos cuatro años. Gracias por dejarme ser un pequeña parte de vuestro trabajo y gracias a ambos por vuestra confianza y apoyo.

A mis padres, gracias por creer en mi, gracias por tener paciencia infinita. Todo lo que he logrado es por y para vosotros.

Rubén, gracias por tu apoyo, paciencia y motivación para ir siempre a por más. Hemos crecido juntos. Por muchos años más.

Carla, hija, eres mi vida. Gracias por existir. Que estés con nosotros es un motivo más para seguir adelante.

A mi hermana, eres mi nudo de ocho, siempre me sostienes. Gracias por estar ahí incondicionalmente.

Gracias a Ana y a Guillermo, gracias por las correcciones, los consejos, los ánimos y todo el esfuerzo que habéis realizado durante estos años de trabajo en equipo.

Gracias a mis amigas, Inés, Nerea y Raquel, gracias por todos estos años acompañándome, aconsejándome y ayudándome. Cerca o lejos, gracias por vuestro apoyo.

Agradecer también a todo el equipo BISITE, el tiempo, los cafés, las risas y los momentos que siempre permanecerán en mi recuerdo.

Finalmente, al Programa de Becas Predoctorales de la Junta de Castilla y León por permitirme completar una etapa más de mi formación.

Índice general

Resumen	i
Abstract	iii
1 Introducción	1
1.1 Planteamiento del Problema	3
1.2 Motivación e hipótesis	5
1.3 Metodología de Investigación	7
1.4 Estructura de la memoria	9
2 Estado del Arte	11
2.1 Fog Computing	15
2.2 Sistemas multi-agentes. Organizaciones Virtuales	16
2.2.1 Tecnología financiera y Organizaciones Virtuales	19
2.3 Conclusiones	20
3 Contribuciones	23
3.1 Introducción	25
3.2 Arquitectura propuesta	26
3.2.1 Descripción técnica: Arquitectura Fog Computing que incorpora Organizaciones Virtuales de Agentes	27
3.3 CEBRA: modelo de razonamiento basado en casos para la banca comercial	30
3.3.1 Diseño de CEBRA	34
3.3.2 Experimentación y resultados	44
3.4 Una plataforma de análisis y extracción de datos para recomendaciones de inversión	46
3.4.1 Modelo propuesto	48
3.4.2 Experimentación y resultados	54
3.5 Conclusiones	57
4 Evidencias y Resultados	59
4.1 Publicaciones	61
4.1.1 Publicaciones en revistas científicas internacionales	61
4.1.2 Publicaciones en congresos internacionales y workshops	62

4.2	Proyectos	63
5	Conclusiones	65
5.1	Conclusiones	67
5.2	Trabajo futuro	71
	Bibliografía	73

Siglas y acrónimos

ANN	A rtificial N eural N etwork
API	A pplication P rogramming I nterface
AR	A ction R esearch
CBR	C ase B ased R easoning
CEBRA	C as E B ased R easoning A pplication
EMA	E xponential M oving A verage
FC	F og C omputing
FinTech	F inancial T echnology
FOBA	F og O riented B anking A rchitecture
IA	I nteligencia A rtificial
IoT	I nternet o f T hings
kNN	k -Nearest Neighbours
LP	L argo P lazo
MAS	M ulti A gent S ystems
ML	M achine L earning
MLP	M ulti L ayer P erceptron
MVC	M odelo V ista C ontrolador
OHLC	O pen - H igh - L ow - C lose
OV	O rganización V irtual

OVC	O rganización V irtual de C liente
OVRD	O rganización V irtual de R ecuperación de D atos
OVTD	O rganización V irtual de T oma de D ecisiones
REST	R Epresentational S tate T ransfer
RSI	R elative S trength I ndex
SI	S istema I nteligente
SR	S istema de R ecomendación
STOCH	S tochastic O scillator
SVM	S upport V ector M achines
TIC	T ecnologías de la I nformación y la C omunicación
ULTOSC	U LTimate O SCillator
WAN	W ide A rea N etwork
WILLR	W ILLams % R
WSN	W ireless S ensor N etwork

Capítulo 1

Introducción



**VNiVERSIDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

Introducción

1.1 Planteamiento del Problema

Algunos sectores de la industria financiera ofrecen productos y servicios basados en la web que no pueden obtenerse de un banco o un proveedor similar. Esto da lugar a un nuevo entorno competitivo, los productos de los nuevos actores van desde soluciones digitales de pago y servicios de información, ahorro y captación de depósitos, hasta servicios modernos como la banca online, el asesoramiento multicanal y la negociación de valores, sin olvidar las soluciones de financiación [Dapp, 2014].

El sector financiero es consciente de la necesidad de aplicar la tecnología para mejorar su actividad; esto se refleja en la acuñación del término FinTech o Tecnología Financiera, utilizado para denominar el uso de programas informáticos y otras tecnologías para apoyar o habilitar los servicios bancarios y financieros. En general, la tendencia actual de los desarrollos FinTech incluye los créditos online, el análisis de riesgos y el tratamiento de grandes volúmenes de datos. Sin embargo, hemos detectado una falta de soluciones tecnológicas en el ámbito de la banca comercial y de los servicios financieros de los gestores personales que contribuyan a mejorar los servicios a los clientes en las instituciones financieras. Ante este nuevo panorama de soluciones que requieren la captura y tratamiento de datos masivos, proponemos el uso de desarrollos de *Fog computing*.

La computación de niebla o *Fog Computing* (FC) surge como un nuevo paradigma computacional que extiende los servicios de la computación en la nube al borde de la red [Networking Index, 2016]. A través de la agrupación de los recursos locales, la computación, la comunicación, el control y el almacenamiento están más cerca de los usuarios finales [Datta et al., 2015] en lugar de estar íntegramente en la nube. Esta

característica mejora la baja latencia, la movilidad, el ancho de banda de la red, la seguridad y la privacidad.

Según los autores en [Hajibaba and Gorgin, 2014] los atributos que caracterizan al FC son:

1. Proximidad de los datos a los usuarios finales: la información ya no se aloja en grandes centros de datos alejados del cliente, sino que se sitúa más cerca del usuario final, mejorando la latencia y el acceso.
2. Organización jerárquica. La arquitectura está organizada en múltiples niveles para soportar una baja latencia y escalabilidad.
3. Localización en el borde de la red, conocimiento de la ubicación, baja latencia: los datos generados por los dispositivos y sensores se adquieren de los colectores *Fog*. Esto proporciona la localización y proporciona una baja latencia y conciencia del contexto. [Bonomi et al., 2012].
4. Amplia distribución geográfica: la distribución geográfica contribuye a un análisis más rápido de los datos y a un mejor soporte de los servicios basados en la localización.
5. Amplia red de sensores y gran número de nodos: el FC utiliza una red de sensores a gran escala para supervisar el entorno.
6. Soporte para la movilidad: los gestores del *host* pueden controlar el acceso de los usuarios sin tener que desplazarse por toda la red de área amplia (WAN), lo que mejora el rendimiento de los usuarios, la calidad del servicio, la seguridad y la privacidad.
7. Interacciones en tiempo real: los servidores pertenecen a la misma red que los usuarios finales, por lo que el FC incluye aplicaciones interactivas en lugar de procesamiento por lotes. El análisis de datos en tiempo real es más frecuente en la informática de niebla.
8. Predominio del acceso inalámbrico: cumple con la visión del Internet de las Cosas (IoT) teniendo en cuenta que se considera un objetivo importante que los dispositivos que participan en las tareas computacionales utilizando los recursos de la red soporten la comunicación inalámbrica. [Madsen et al., 2013].

9. Heterogeneidad: los nodos *Fog* son altamente dinámicos y heterogéneos para soportar la baja latencia y los requisitos de las aplicaciones. [Hong et al., 2013].
10. Optimización dinámica por usuario: a diferencia de la nube, que está separada por *WANs*, un servidor *Fog* tiene la ventaja de conocer localmente las condiciones de la red para el usuario final, ya que pertenecen a la misma red. Por lo tanto, el servidor *Fog* puede tener conocimiento de cada usuario lo que puede ayudar a elegir los mejores parámetros y personalizar la optimización [Zhu et al., 2013].

Debido al panorama tecnológico descrito, esta tesis doctoral aborda los desafíos de la Tecnología Financiera y sus entornos heterogéneos al proponer una Arquitectura Inteligente *Fog Computing* que integre sistemas predictivos dirigidos a la prestación de servicios personalizados al cliente en entornos bancarios. La memoria que se presenta esta basada en el modelo de compendio de artículos y se estructura en capítulos. El primero es una Introducción en la que se explica la motivación, hipótesis y objetivos de la investigación. Posteriormente el Capítulo 2 presenta un estado del arte de las tecnologías aplicadas. El Capítulo 3 profundiza en el planteamiento del problema y describe la arquitectura propuesta como solución y las herramientas predictivas diseñadas para la recomendación de productos bancarios y gestión de activos.

En el Capítulo ?? se presentan las publicaciones originales que forman parte de esta memoria por compendio de artículos. Cada publicación se acompaña de su correspondiente introducción y las conclusiones de los resultados obtenidos en cada una. Los resultados alcanzados mediante el desarrollo del estudio y que se validan a través de las publicaciones realizadas en: revistas, conferencias, workshops y capítulos de libros; así como los proyectos en los que se ha participado en calidad de estudiante del programa de doctorado se presentan en el Capítulo 4. Finalmente, las conclusiones y trabajo futuro de investigación que se han generado a raíz de esta tesis doctoral se detallan en el Capítulo 5.

1.2 Motivación e hipótesis

La banca actual, altamente competitiva, impulsada en parte por el rápido crecimiento de los nuevos paradigmas informáticos, junto con el FinTech, está empujando al sector a

buscar formas de seguir mejorando las relaciones con los clientes. Los procesos analíticos en entornos *Cloud* pueden aprovechar grandes volúmenes de datos para llevar a cabo un procesamiento computacional que incluya técnicas de aprendizaje automático para mejorar la fiabilidad, la configuración automática y el rendimiento [Mendhurwar and Mishra, 2018], sin embargo, el FC es una convergencia tecnológica que responde a los nuevos requerimientos que han surgido debido a la ubicuidad de los dispositivos, a las demandas de una gestión más ágil de las redes y servicios, y a los nuevos requerimientos relacionados con la privacidad de los datos [Vaquero and Rodero-Merino, 2014].

A la hora de diseñar la propuesta de arquitectura FC para un entorno bancario se han tenido en cuenta los siguientes aspectos:

1. Un servicio de atención al cliente más cercano. Partimos de la base de que los servicios de atención al cliente y los centros de datos deben trasladarse a entornos más cercanos al cliente (a la capa *Fog*). De esta forma se espera una mejora de la experiencia del usuario mediante interfaces adaptativas y la personalización de los servicios.
2. Aspectos de seguridad y transparencia. Son cruciales en este sector para la validación de las identidades de las transacciones, las alertas de fraude, la atención personalizada al cliente, etc.
3. Necesidad de agilidad en los procesos. Pagos, tasas, solicitud de créditos, cambio de divisas, criptomonedas, etc.
4. Reducción de los costes de gestión de la entidad bancaria ya que al implementar servicios seguros el cliente puede contratar productos financieros sin necesidad de la atención del personal de la oficina bancaria.

Por tanto, la hipótesis de partida es que es posible mejorar aspectos relacionados con los servicios de atención al cliente de una entidad bancaria, como lograr una mayor transparencia y agilidad de los procesos, al diseñar una arquitectura flexible, eficiente y personalizable basada en FC a partir de un modelo de agentes que facilite una gestión distribuida y dinámica, capaz de recopilar y analizar los datos recogidos por los dispositivos móviles, aplicaciones y equipos locales de la entidad, para recomendar productos bancarios.

Teniendo en cuenta las características del FC comentadas en la subsección 1.1, se propone una arquitectura flexible, eficiente, segura y personalizable denominada FOBA (*Fog Oriented Banking Architecture*), basada en FC, para la recogida y análisis de los datos recopilados por los dispositivos móviles, aplicaciones y equipos locales de la entidad bancaria. La propuesta consiste además en integrar sistemas predictivos para la recomendación de productos bancarios y gestión de activos. La recopilación y tratamiento de los datos se realiza a través de nodos *Fog* que incorporan Agentes Inteligentes para su procesamiento en tiempo real con características del entorno. La motivación detrás de esta investigación y el objetivo principal, es mejorar aspectos de los servicios de atención al cliente, especialmente, lograr una mayor seguridad, mayor transparencia y agilidad de los procesos, así como reducir los costes de gestión de la entidad.

Para alcanzar el objetivo principal, es necesario definir un listado de objetivos específicos, que se describen en el siguiente orden:

- **(OB1)** Identificar los requerimientos existentes en los entornos bancarios, en el ámbito de la tecnología financiera, y que se pueden fortalecer a través de soluciones basadas en *Fog Computing*.
- **(OB2)** Revisar las arquitecturas de referencia existentes en las que el *Fog Computing* es el eje principal.
- **(OB3)** Diseñar una arquitectura *Fog Computing* modular y escalonada, basada en agentes para dotar al sistema de inteligencia, orientada a entornos bancarios.
- **(OB4)** Conocer el estado de la técnica en cuanto a sistemas inteligentes de recomendación en el ámbito de la tecnología financiera.
- **(OB5)** Proporcionar sistemas de recomendación integrables en una arquitectura *Fog Computing* para recomendar productos bancarios y de gestión de activos de una manera eficiente, buscando una reducción de costes para la entidad bancaria.

1.3 Metodología de Investigación

La metodología empleada ha sido la *Action-Research* (acción-investigación), de aquí en adelante AR. El término fue acuñado por Kurt Lewin en 1952 [Herrerias, 2004] y

surgió como respuesta a la necesidad de una investigación que no se centrara sólo en la producción sino que hubiera un proceso cíclico de exploración, actuación y valoración de resultados, es una metodología orientada hacia el cambio. Actualmente se define como un proceso metodológico para realizar las actividades de mejora y mantener aquello en lo que se ha mejorado. El pilar principal sobre el que se apoya esta metodología es la posibilidad de sacar el máximo partido al conocimiento que reside en los casos prácticos, de esta manera el conocimiento se vuelve incremental a la vez que se van solucionando los problemas que surgen a medida que se va investigando [Coghlan, 2019].

De entre todas las características que posee esta metodología, la presente tesis doctoral se beneficia de que:

1. se construye desde y para la práctica;
2. pretende mejorar y comprender a la vez;
3. implica análisis crítico de las situaciones;
4. parte de un ciclo de planificación, acción, observación y reflexión;
5. exige más rigor en la investigación y más justificaciones teóricas [Coughlan and Coghlan, 2002].

Como se menciona en [Tanco and Camarero, 2013] la generación de conocimiento debe ser prioritaria en la investigación y este conocimiento debe traducirse en proposiciones empíricamente contrastables y que puedan generar una teoría. Así, esta metodología, al constituir una reflexión en la práctica y desde la práctica, permite que la investigación fluya de los datos a la teoría. Los proyectos de AR son situaciones específicas y no pretenden generar conocimiento universal, pero, al mismo tiempo, deben tener algunas implicaciones más allá de la acción o el conocimiento del proyecto concreto.

A continuación se exponen las fases genéricas que conlleva la metodología AR:

1. Planteamiento: formulación de un problema concreto por solucionar, una hipótesis y unos indicadores. Recogido en el punto 1.2 Motivación e hipótesis.
2. Revisión global: se forma un grupo para actuar, reflexionar y aprender juntos, de esta manera se mezcla el conocimiento local y el científico y se hace posible

el identificar áreas de mejora. Esta fase se correspondería con todo el trabajo recogido en el Capítulo 2 del documento ya que durante el desarrollo de esta tesis doctoral se ha realizado la revisión bibliográfica sobre arquitecturas FC, CBR y OVA. Un ejemplo de la revisión global sería el artículo de Breidbach *et al.* donde se plantea una agenda para guiar futuras investigaciones de servicios orientados a la transformación digital de los sistemas financieros a través del FinTech [Breidbach et al., 2019].

3. Revisión específica: análisis del problema y de las alternativas de solución. Se elabora una guía de revisión, mapas de procesos e indicadores.
4. Planificar y aplicar mejoras: se elabora un plan que incorpora las tareas que hay que realizar, las personas implicadas y las fechas de cumplimiento de cada tarea.
5. Revisión del plan de mejoras.

Como tareas concretas dentro de la tesis doctoral se han identificado las siguientes:

1. la revisión bibliográfica sobre arquitecturas FC, sistemas de recomendación basados en casos y sistemas multi-agente;
2. la estructuración de objetivos, fases y tareas para el programa de investigación;
3. el tratamiento de datos y aplicación de algoritmos;
4. el diseño de casos de estudio para validar las técnicas seleccionadas;
5. el análisis de los resultados obtenidos con los modelos desarrollados.

Entre las ventajas que presenta esta metodología [Benbasat et al., 1987] se encuentran las soluciones que aporta a las organizaciones y la contribución científica que realiza acerca de un tema, es decir la generación de conocimiento. Suponiendo además, para el investigador, la obtención de conocimiento en primicia.

1.4 Estructura de la memoria

Con la finalidad de comprobar la hipótesis establecida para esta investigación y cumplir con los objetivos definidos, se estructura la presente memoria a través de seis capítulos.

El primero de ellos es introductorio. El capítulo 2 presenta el estado del arte del Fog Computing y se revisan los conceptos de sistemas multiagentes y organizaciones virtuales, y cómo se han incorporado al sector financiero. En esta sección se muestran las investigaciones publicadas hasta la fecha, así como sus principales aportaciones. A modo de resumen, se presentan varias tablas que recopilan las aportaciones que se han considerado en los trabajos de investigación.

Las contribuciones de esta tesis doctoral se detallan en el capítulo 3, en el cual, se incluye una descripción conceptual y técnica de la arquitectura propuesta, denominada *Fog Oriented Banking Architecture* (FOBA) y los dos sistemas de recomendación diseñados y desarrollados para ser alojados en el arquitectura. El sistema de recomendación basado en casos denominado *CasE-Based Reasoning Application* (CEBRA) y la plataforma de recomendación de inversiones que aplica tanto algoritmos de machine learning como indicadores de análisis técnico e indicadores *Momentum* para proporcionar al usuario el algoritmo con la mejor precisión.

En el capítulo ?? se presentan las publicaciones originales que forman parte de esta memoria por compendio de artículos. Cada publicación se acompaña de su correspondiente introducción y las conclusiones de los resultados obtenidos en cada una.

Los resultados alcanzados mediante el desarrollo del estudio y que se validan a través de las publicaciones realizadas en: revistas, conferencias, workshops y capítulos de libros; así como los proyectos en los que se ha participado durante el periodo en calidad de estudiante del programa de doctorado se presentan en el capítulo 4.

El capítulo 5 presenta las principales conclusiones que han surgido del trabajo de investigación desarrollado, las aportaciones más relevantes y se definen las líneas para el desarrollo de trabajo futuro. Para finalizar, se incluye el listado de todas las fuentes bibliográficas que se han citado en esta memoria con el propósito de respaldar las afirmaciones y conceptos que se presentan.

Capítulo 2

Estado del Arte



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

Estado del Arte

La tecnología financiera está experimentando un rápido crecimiento con el concepto de FinTech, entendido como el conjunto de operaciones y actividades que realizan las organizaciones utilizando las nuevas tecnologías aplicadas a los servicios financieros que ofrecen. Sin embargo, este concepto no sólo abarca la idea de tecnología aplicada a los servicios, el FinTech aborda otros retos del sector financiero como son: la seguridad, las amenazas, los ciberataques, el Internet de las Cosas (IoT), la Inteligencia Artificial y el conjunto de tecnologías relacionadas con las redes móviles, la computación en la nube, el análisis de grandes volúmenes de datos, que forman parte del día a día de los usuarios del sector [Mehrban et al., 2020, Xie, 2019]. El interés cada vez mayor por el término FinTech obliga a organizaciones, profesionales e investigadores a seguir avanzando en este campo, aportando conocimientos precisos y actualizados en este ámbito [Cao, 2020]. En este sentido, en este capítulo se presenta una revisión de las tendencias y principales líneas de investigación en el ámbito de la tecnologías financiera en los últimos años.

La disrupción de la banca online ha provocado retos para las organizaciones y los investigadores, entre los más importantes se encuentran: la recopilación, el procesamiento y el análisis de grandes volúmenes de datos generados por los clientes y los servicios, la ciberseguridad para garantizar la integridad y protección de los datos, los equipos y los clientes, y la evolución del comercio electrónico [Mohan, 2016].

En la última década, la aplicación de las Tecnologías de la Información y la Comunicación (TIC) al sector financiero, han supuesto una transformación en el acceso a los datos y su disponibilidad, impactando directamente en la evolución de los servicios ofrecidos por las organizaciones [Breidbach and Maglio, 2016]. La investigación sobre el impacto de las TIC en el sector financiero no es reciente, sin embargo, varios autores cuestionan las aportaciones de los trabajos publicados al no abordar las tecnologías disruptivas y

cómo dan lugar a los nuevos servicios FinTech y sus consecuentes procesos de gestión [Breibach et al., 2019, Brust et al., 2017, Lusch and Nambisan, 2015].

Es importante destacar los aportes de Roumani *et al.* (2016) quienes desarrollaron un estudio para determinar cómo se asocian los registros financieros de las organizaciones con sus debilidades de ciberseguridad [Roumani et al., 2016]; en [Chen and Weiss, 2014] los autores utilizaron agentes inteligentes para monitorear y predecir los riesgos financieros en diversas operaciones de negocios. Otro grupo de autores ha trabajado en cómo la seguridad afecta a la operación empresarial en su conjunto, analizando la fiabilidad del asesoramiento financiero [Gai et al., 2018, Nussbaumer et al., 2012]. Yin y Gai han centrado su investigación en las técnicas de análisis de datos aplicadas a las FinTech [Yin and Gai, 2015]; en [Gai et al., 2018, Yin and Gai, 2015], los autores investigan el impacto y la aceptación de la computación en la nube en el sector financiero; en [Shi et al., 2017] los autores continúan la investigación de Wu *et al.* para obtener información sobre los volúmenes de datos generados por las operaciones financieras incorporando instrumentos cuantitativos [Wu et al., 2013]. Sagraves y Connors, destacan la importancia de ofrecer nuevos servicios personalizados, como sistemas de recomendación o asesoramiento bancario, independientemente del servicio financiero solicitado por el cliente [Sagraves and Connors, 2017]. Toda esta literatura ha motivado la aparición de numerosas líneas de investigación que adaptan diferentes enfoques orientados a las tecnologías financieras:

- La transformación de los servicios financieros existentes con nuevas tecnologías, como: criptomonedas y Blockchain [Breibach et al., 2019, Fosso Wamba et al., 2020];
- La aparición de nuevos servicios financieros que surgen a través de la aplicación/incorporación de técnicas de Inteligencia Artificial en FinTech. [Mogaji et al., 2020]. Las técnicas de IA han motivado el nacimiento de nuevos servicios financieros inteligentes como: asesoramiento, préstamos, vigilancia y atención al cliente [Xie, 2019].
- Identificar los requisitos del mercado y los procesos adecuados para la gestión de estos servicios, que garanticen no sólo su eficacia, sino también la seguridad de las organizaciones y sus usuarios [Hommel and Bican, 2020, Zetsche et al., 2017].

Hay que destacar cómo los avances tecnológicos, entre los que se encuentran la digitalización, el *Big Data*, el Internet de las cosas o la Inteligencia Artificial, están transformando el sector bancario [Llewellyn, 2018]. Al igual que otras organizaciones se transforman gracias a los datos sobre el comportamiento de los consumidores, los bancos deben adaptarse a clientes menos fieles y más activos que exigen acceso a información en tiempo real [Stulz, 2019].

Es en este contexto donde comienza el desarrollo de esta tesis doctoral. Una vez planteadas las líneas de investigación que contemplan varios enfoques orientados al FinTech, partimos de la segunda línea donde se pone de manifiesto que las técnicas de IA han motivado el nacimiento de nuevos servicios financieros inteligentes como: asesoramiento, préstamos, vigilancia y atención al cliente [Xie, 2019]. Es por eso que las siguientes secciones estarán dedicadas a las tecnologías adoptadas para dar forma a las aplicaciones de recomendación de productos bancarios y gestión de activos que se han desarrollado a lo largo de esta tesis doctoral incluyendo también el estado del arte relacionado con la arquitectura informacional escogida para dar cabida a ambas aplicaciones, que en este caso ha sido el *Fog Computing*.

2.1 Fog Computing

Se han revisado diferentes soluciones de FC desarrolladas para campos de aplicación específicos. Los principales ejes de desarrollo en las aplicaciones de FC corresponden al ámbito sanitario y a las Smart Cities, así como a otros campos de aplicación.

1. Soluciones FC para aplicaciones sanitarias: El uso de dispositivos IoT para la monitorización continua de las constantes vitales de los pacientes hace necesario que las aplicaciones sanitarias envíen información en tiempo real. Sin embargo, debido a su sensibilidad a la latencia, las arquitecturas Fog se han convertido en las preferidas para el desarrollo de estas aplicaciones. Los artículos que hemos encontrado dentro de esta categoría van desde aplicaciones de atención domiciliaria a personas mayores hasta las aplicaciones de modelado y la simulación para mejorar el flujo del servicio en entornos sanitarios. Las arquitecturas FC aplicadas en el campo de la salud y el bienestar del paciente, que han sido revisadas pueden verse en la Tabla 3.1 aún así hay un aumento en el número de arquitecturas dirigidas

a este campo debido a que el FC permite personalizar mejor la aplicación para su implementación ya que es capaz de filtrar, procesar y analizar la información recogida por dispositivos específicos de manera global y en tiempo real, algo que es muy útil en el caso de las aplicaciones de salud.

2. Soluciones FC para ciudades inteligentes: Los sensores se instalan en los hogares no sólo con fines sanitarios, sino también para lograr una mayor eficiencia energética. Ha habido innumerables propuestas que utilizan los datos de los sensores para fomentar un comportamiento más eficiente desde el punto de vista energético en los usuarios. Las redes inteligentes son uno de los campos en los que el FC tiene mucho que aportar. En esta sección revisaremos las arquitecturas que hemos encontrado en este campo.
3. Soluciones propuestas por las grandes empresas comerciales: El FC se está convirtiendo en un paradigma destacado en las soluciones propuestas por las grandes empresas comerciales. Aparte del ámbito académico, también es posible encontrar plataformas comerciales de IoT de grandes empresas como Huawei, Dell o Intel [Ai et al., 2018].

En este apartado se recopilan y estudian las soluciones encontradas en diversos artículos, describiendo su alcance, el enfoque que siguen y su principal aportación en formato de tabla (Tabla 2.1).

2.2 Sistemas multi-agentes. Organizaciones Virtuales

Hoy en día, la irrupción del Internet de las Cosas (IoT) ha llevado a los investigadores a centrarse en los avances relacionados con la sensórica, los sistemas embebidos, las comunicaciones inalámbricas, etc. En [Savaglio et al., 2020] los autores presentan cómo se pueden simular directamente los sistemas basados en agentes para estudiar tanto los patrones individuales emergentes como los fenómenos colectivos. Los autores consideran que la simulación basada en agentes es un enfoque natural que resulta útil para comprender y gestionar la dinámica global de sistemas complejos, como las redes de distribución y suministro, las ciencias sociales, etc. También subrayan su importancia a la hora de evaluar sistemas distribuidos que presentan un comportamiento discreto, no

TAB. 2.1: Resumen del estado del arte en el ámbito *Fog Computing*.

Soluciones	Contribución	Conclusiones
Ámbito sanitario [Al-khafajiy et al., 2019, Al Ridhawi et al., 2017, Fratu et al., 2015, Kyriazakos et al., 2016, Monteiro et al., 2016, Oueida et al., 2018, Stantchev et al., 2015]	Monitoreo permanente Sensibles a la latencia Notificaciones en tiempo real	Almacenamiento a L.P. Mejor personalización Procesamiento en tiempo real
Ciudades inteligentes [Aloqaily et al., 2019, Brzoza-Woch et al., 2016, Tang et al., 2015, Yan and Su, 2016]	Uso eficiente de energía Gestión de la ciudad	Mejor almacenamiento Control de infraestructuras
Empresas	Evaluación de riesgos Interconexión productos Paquetes de IoT y Edge computing	Sistemas automatizados Soluciones de datos, enchufes inteligentes Sensores y software a medida, Software de gestión y supervisión a gran escala

lineal y adaptativo. En [González-Briones et al., 2020] los autores señalan la complejidad intrínseca de modelar y diseñar arquitecturas distribuidas, la relación entre los agentes del sistema y su comunicación con el resto de elementos del sistema. En su investigación se propone una arquitectura de Organización Virtual (OV) de agentes que permite adaptarse a comportamientos y preferencias muy variables de los usuarios. Es capaz de manejar *Wireless Sensor Networks* (WSN), *Artificial Neural Networks* (ANN) y sistemas *Case Based Reasoning* (CBR) y así obtener las preferencias de los usuarios y predecir su comportamiento. Otro ejemplo de cómo se ha aplicado la investigación con OV al campo del IoT se puede ver en [Shokri Gazafroudi et al., 2019] donde se propuso una arquitectura de OV para el comercio de energía entre los agentes de la red de distribución. Otros autores centran su investigación en la portabilidad, las limitaciones de recursos, la seguridad y la escalabilidad de las plataformas de procesamiento de agentes. En [Bosse, 2016] proponen el despliegue de sistemas multi-agente (MAS, por sus siglas en inglés) en redes heterogéneas fuertes, incluyendo Internet. Su propuesta es desplegar MAS, implementando los agentes directamente en JavaScript. En [Alsboui et al., 2020] los autores introducen una arquitectura de agentes móviles que gestiona recursos y permite la implementación de aplicaciones IoT. El enfoque de agentes propuesto es eficiente y facilita la interacción local, así como la recogida y agregación de datos.

Una de las ventajas expresadas por los investigadores en el campo de IoT sobre los sistemas multi-agentes es la auto-organización y la auto-adaptación como se muestra en la investigación realizada en [do Nascimento and de Lucena, 2017] donde se propuso un marco basado en agentes IoT para crear aplicaciones (controladores autónomos para el tráfico de automóviles, maquinaria, alumbrado público, electrodomésticos inteligentes y hogares inteligentes).

A modo de resumen, la Tabla 2.2 resume las aportaciones consideradas en esta investigación.

TAB. 2.2: Estado del arte de las OV aplicadas al IoT

Referencias	Aplicaciones
[Savaglio et al., 2020]	Los autores presentan cómo se pueden simular directamente los agentes para estudiar patrones, etc.
[González-Briones et al., 2020]	Los autores proponen una arquitectura de agentes que se adapta al comportamiento y las preferencias del usuario.
[Shokri Gazafroudi et al., 2019]	Los autores proponen una arquitectura de agentes para el comercio de energía.
[Bosse, 2016]	La propuesta de los autores es desplegar un sistema multi-agente, implementando los agentes directamente en JavaScript.
[Kwan et al., 2016]	La propuesta de los autores es la agentificación de los objetos (para crear objetos virtuales conectados).
[do Nascimento and de Lucena, 2017]	Los autores propusieron un marco basado en agentes de IoT.
[Alsbouï et al., 2020]	Los autores proponen una arquitectura de agentes móviles para el IoT.

Como comentario final a la revisión de esta sección, decir que en la revisión del estado del arte presentada por [Pico-Valencia and Holgado-Terriza, 2018], se destaca que tanto los modelos experimentales como los no experimentales publicados, que proporcionan un proceso de integración entre el IoT y los agentes, muestran el creciente interés de los investigadores por la creación de redes cognitivas de objetos IoT. Como por ejemplo, en la investigación presentada en [Kwan et al., 2016] los autores proponen una nueva forma de utilizar el IoT. La propuesta consiste en objetos conectados a través de un enfoque multi-agente. Los autores explican que a través de la agentificación, los objetos conectados pueden ser utilizados para crear objetos virtuales conectados que se

utilizarían de la misma manera que los objetos físicos en la red de un sistema de objetos conectados. Como conclusión de su investigación, los autores en [Pico-Valencia and Holgado-Terriza, 2018] muestran que el modelado de objetos como agentes sirve para la gestión de las redes de objetos IoT a través del sistema multi-agente. Es posible abordar el componente cognitivo, la capacidad social y generar una red colaborativa, capaz de adaptarse y tener autonomía.

2.2.1 Tecnología financiera y Organizaciones Virtuales

En esta subsección se presentan algunas investigaciones que han utilizado Sistemas Multi-Agente (MAS) u Organizaciones Virtuales de Agentes para proponer, diseñar o crear servicios útiles en el ámbito de la Tecnología Financiera. Además, se identifican las principales líneas de investigación y aportaciones al área de FinTech. Estudios como el de Sagraves y Connors [Sagraves and Connors, 2017] que abordan los retos del sector bancario para satisfacer las necesidades de los clientes a través de los servicios que ofrecen, son una de las principales motivaciones para el desarrollo de esta tesis doctoral.

La gran aceptación de las Organizaciones Virtuales de agentes ha permitido su aplicación en diversos campos, sin embargo, existe poca investigación sobre la implementación de Organizaciones Virtuales en el campo del FinTech [Hernández-Nieves et al., 2020]. En [Kim, 2003], los autores compararon el rendimiento de los algoritmos de aprendizaje automático con los sistemas de razonamiento basados en casos (CBR) en la previsión financiera. Kim afirmó que el rendimiento de los algoritmos que utilizó (SVM) era superior al del sistema CBR. Posteriormente, otros autores objetaron sus resultados, indicando que contienen un sesgo de búsqueda de datos [Henrique et al., 2019]. En la investigación realizada en [Said et al., 2018] se han empleado MAS en un entorno financiero artificial para reflejar la complejidad del sistema del mercado financiero. Los autores investigan las interacciones entre los agentes y su entorno, basándose en una investigación previa desarrollada por quienes utilizaron MAS con el objetivo de explicar estas interacciones. En [Qi and Xiao, 2018], los autores destacan cómo la Inteligencia Artificial permite a las organizaciones financieras diseñar e implementar sistemas inteligentes con nuevas ofertas de productos. En [Bunnell et al., 2020] se propone una arquitectura para un sistema de recomendación de objetivos financieros basado en agentes llamado FinPathlight. Los resultados teóricos obtenidos en esta

investigación proporcionan una guía para el diseño de sistemas de recomendación financiera para FinTech basados en agentes y sus ontologías.

A continuación, la Tabla 2.3 resume la investigación considerada en esta subsección en el área del FinTech y las Organizaciones Virtuales de Agentes.

TAB. 2.3: Organizaciones Virtuales en el campo del FinTech.

Referencias	Enfoques
[Kim, 2003]	Los autores compararon el rendimiento de los algoritmos de aprendizaje automático con los sistemas de razonamiento basados en casos en la previsión financiera.
[Henrique et al., 2019]	Otros autores objetaron los resultados presentados en [Kim, 2003], indicando que contienen un sesgo de búsqueda de datos.
[Said et al., 2018]	Sistemas multi-agentes (MAS) en un entorno financiero artificial para reflejar la complejidad del sistema del mercado financiero.
[Qi and Xiao, 2018]	La Inteligencia Artificial permite a las organizaciones financieras diseñar e implementar sistemas inteligentes con nuevas ofertas de productos.
[Bunnell et al., 2020]	Se propone una arquitectura para un sistema de recomendación de objetivos financieros basado en agentes denominado FinPathlight.

2.3 Conclusiones

Los sistemas de Inteligencia Artificial (IA) que utilizan agentes, sistemas multi-agentes (MAS) y Organizaciones Virtuales (OV) aumentan su capacidad para resolver nuevos tipos de problemas, mejorando la eficiencia de los sistemas anteriores [Jennings and Wooldridge, 1998]. Sin embargo, en las últimas décadas, la investigación en MAS ha introducido el concepto de organización [Ponomarev and Voronkov, 2017], cuya aparición dio lugar al desarrollo de OV que coordinan, negocian y desarrollan acciones sociales más complejas para el intercambio de información y servicios.

Desde una perspectiva social, las OV son sistemas abiertos formados por entidades heterogéneas agrupadas para colaborar entre sí pero con independencia en sus formas, funciones y comportamiento [Ferber et al., 2003]. La incorporación de agentes, organizaciones virtuales o sistemas multi-agente, requiere de mecanismos

de razonamiento con capacidad de aprendizaje y adaptación, como los sistemas de razonamiento basados en casos [Luis Reboredo et al., 2014]. En la actualidad, los sistemas CBR (*Case-based reasoning*) se siguen aplicando en diversos campos porque utilizan situaciones del pasado y recomiendan nuevas soluciones gracias a su capacidad de adaptación y comunicación con varias instancias y otras bases de casos [Bergenti et al., 2006].

Para concluir la revisión del estado del arte cabe destacar que las organizaciones financieras se caracterizan por ser un sector puntero en la implantación de nuevas tecnologías para el desarrollo de modelos de negocio competitivos y eficientes. Sin embargo, en el nuevo escenario FinTech, las organizaciones se enfrentan a procesos de transformación en los que necesitan obtener información en tiempo real para ofrecer nuevos productos y soluciones que satisfagan las necesidades de sus clientes [Gomber et al., 2017, Milian et al., 2019]. Aunque la disrupción de la IA en las finanzas ha incrementado el número de proyectos de investigación orientados a la aplicación de modelos, arquitecturas y técnicas, algunos de ellos como los MAS, basados en un único sistema, trabajando de forma centralizada, no satisfacen los requerimientos actuales de los entornos FinTech [Milian et al., 2019]. Esto se debe a que los datos provienen de fuentes heterogéneas y deben ser recogidos de forma distribuida, por lo que se requiere que las OVA trabajen de forma colaborativa [Hugoson, 2007, Ponomarev and Voronkov, 2017] y que se basen en arquitecturas *Fog Computing*, siendo esta afirmación la línea de base de la tesis doctoral que se presenta.

Capítulo 3

Contribuciones



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

Contribuciones

3.1 Introducción

El concepto de *Fog Computing* ha sido descrito en detalle en los Capítulos 1 y 2, así como los escenarios de aplicación. También se han estudiado las tecnologías escogidas para compartir la inteligencia de negocio y mejorar la comunicación intrabancaria, y las tecnologías escogidas para realizar mejoras en cuanto a la tasa de aceptación de productos bancarios ofrecidos por la sucursal, tanto a clientes ya asentados como a clientes potenciales.

Esta tesis doctoral aborda los desafíos del FinTech y sus entornos heterogéneos al proponer una Arquitectura Inteligente FC basada en agentes que integre sistemas predictivos para recomendar productos bancarios y gestión de activos financieros.

El diseño basado en agentes dota al sistema de capacidad de adaptación y de adquirir conocimientos así como de tomar decisiones adecuadas. Como se explica en [Chamoso et al., 2018] las OV ofrecen al sistema la posibilidad de desarrollar un núcleo de software flexible, con gran independencia y modularidad en la aplicación de metodologías de recomendación para proporcionar la mejor solución.

En este capítulo se presenta la arquitectura propuesta, que ha sido diseñada de manera escalable para futuros requerimientos, su descripción técnica y las principales aportaciones del trabajo desarrollado. También se incorporan como secciones dentro de este capítulo las herramientas de recomendación diseñadas y desarrolladas para un entorno bancario, más concretamente se presenta el sistema CBR CEBRA: *CasE-Based Reasoning Application* y una plataforma para realizar recomendaciones de inversión bursátil a través de la aplicación de algoritmos de *Machine Learning*

(ML) (*Random Forest Regressor*, *Gradient Boosting Regressor*, *SVM-LinearSVR*, *MLPRegressor*, *kNNNeighbors Regressor*). Por último, en la sección de conclusiones se describen las contribuciones científicas que se han generado.

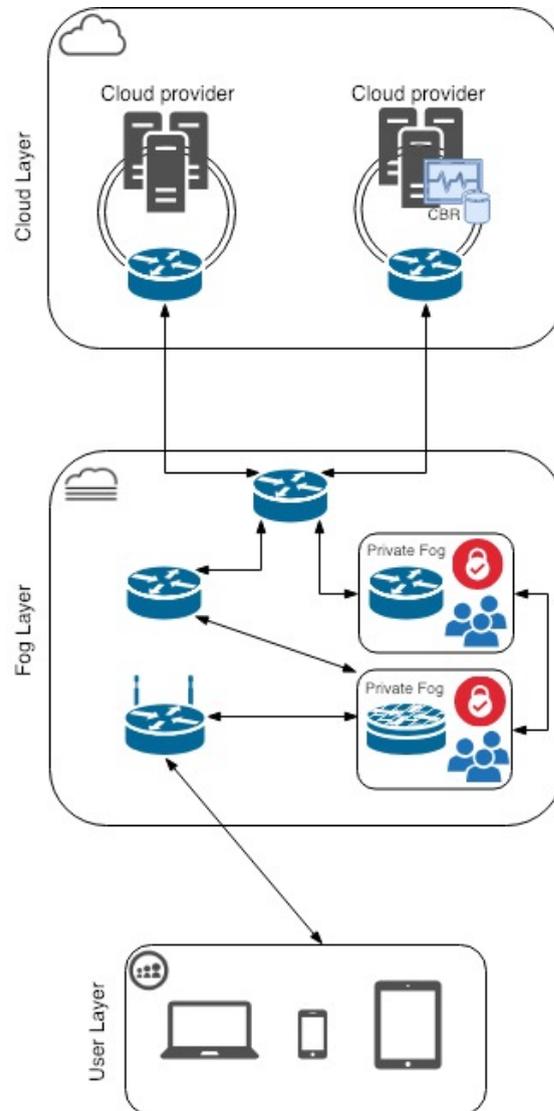
3.2 Arquitectura propuesta

Esta sección presenta: *Fog Oriented Banking Architecture* (FOBA) [Hernández-Nieves et al., 2020]. La Figura 3.1 resume la arquitectura, las tecnologías clave, los retos y las cuestiones abiertas de esta propuesta. La capa *Fog* comprende los nodos FC que pueden crearse utilizando dispositivos con poca capacidad de cálculo y conectividad de red o almacenamiento. Estos nodos FC son fácilmente desplegados y pueden interactuar con los usuarios en el borde de la red y recoger datos. La propuesta incluye nodos en los que se pueden procesar los datos con agentes inteligentes ligeros que permiten implementar sistemas de recomendación contextual. Estos nodos colaboran con los usuarios en el borde para lograr ciertas tareas de gestión, así como de almacenamiento de comunicaciones.

Se propone un sistema inteligente (SI) para ejecutar algoritmos de sistema de recomendación (SR) y para ayudar a filtrar los productos bancarios entre el gran volumen de productos disponibles. Este motor de recomendación analiza la información adquirida a través de los sensores y la información disponible en el banco, para proporcionar sugerencias al usuario a través de un agente personal mediante una API (Application Programming Interface) que utiliza métodos de filtrado colaborativo. Además se propone la verificación de las tasas de éxito de la recomendación para el almacenamiento/eliminación de casos del CBR.

Esta propuesta facilita la compartición, a través del CBR en la Nube y en las capas *Fog* de cada sucursal de la entidad, de toda la inteligencia de negocio, de esta forma se pretende mejorar el servicio de atención al cliente por parte de la entidad al ofrecer un servicio más rápido y personalizado. También permite establecer, al mismo tiempo, recomendaciones a través de datos de contexto en los nodos FC de la sucursal física.

FIG. 3.1: FOBA Architecture diagram.



3.2.1 Descripción técnica: Arquitectura Fog Computing que incorpora Organizaciones Virtuales de Agentes

En esta sección se describen los aspectos relacionados con la arquitectura FOBA y la comunicación de datos entre las sucursales del banco, así como el uso de esta información para conseguir un sistema de apoyo a la toma de decisiones para la banca comercial. CEBRA, *CasE-Based Reasoning Application*, es un agente asesor de banca minorista que opera en una plataforma *Fog*, apoyando el proceso de decisión de la banca comercial mediante una combinación de modelos de decisión locales y globales y datos locales.

En un entorno complejo es difícil determinar cuándo y cómo recomendar productos bancarios de forma eficaz. Para abordar este problema desde un punto de vista

innovador, este enfoque tiene en cuenta las sociedades humanas como inspiración. En las sociedades humanas se pueden encontrar estructuras organizativas que se crean y evolucionan mediante comportamientos emergentes o deliberativos complejos. La tecnología de agentes puede imitar a las sociedades humanas mediante la constitución de organizaciones virtuales dinámicas de agentes. Estos sistemas son capaces de tomar decisiones de forma autónoma y flexible, cooperando con otros sistemas dentro de una organización [García-Fornes et al., 2011, Oyenan et al., 2009, Rodríguez et al., 2011]. Diferentes estudios han aportado distintas perspectivas sobre cómo deben estructurarse las organizaciones para adaptarse fácil y eficazmente a los cambios de su entorno; adaptando los antiguos roles a las nuevas circunstancias o creando otros nuevos [Artikis, 2009, Carrascosa et al., 2009]. Las Organizaciones Virtuales (OV) basadas en agentes permiten la descripción de composiciones estructurales y comportamientos funcionales, así como la inclusión de regulaciones normativas para el control del comportamiento de los agentes, para la entrada/salida dinámica de componentes y para la formación dinámica de grupos de agentes [Echeverry et al., 2012]. Un análisis de las posibilidades y beneficios derivados de la implantación de sociedades artificiales muestra que las organizaciones virtuales son una tecnología adecuada para el funcionamiento complejo y altamente dinámico de la banca minorista. La Figura 3.2 muestra las organizaciones virtuales propuestas, adaptadas a la arquitectura FOBA de 3 niveles.

1. La capa de usuario corresponde a todos los dispositivos inteligentes que generan datos. Los datos se procesarán en el mismo terminal y se enviarán a los nodos *Fog* situados en la capa *Fog*.
2. La capa *Fog* incluye OV de agentes como método de coordinación entre las distintas sucursales de una misma entidad bancaria, siendo por tanto un método de comunicación intrabancaria. La figura 3.2 muestra tres secciones, correspondiéndose cada una de ellas con una OVA relacionada con una zona geográfica concreta, pudiendo ser ciudades u otras divisiones territoriales en función de su extensión geográfica. De este modo manera, OV1 se podría corresponder con la zona de Madrid, OV2 con Zaragoza,... De esta manera se pretende compartir toda la inteligencia empresarial. Dentro de esta misma capa *Fog*, se encuentra el sistema CBR CEBRA.

3. La capa *Cloud* corresponde a los servidores generales de la entidad. También se propone la inclusión del sistema CBR CEBRA para esta capa con conjuntos de datos más grandes, dada la mayor capacidad de computación disponible.

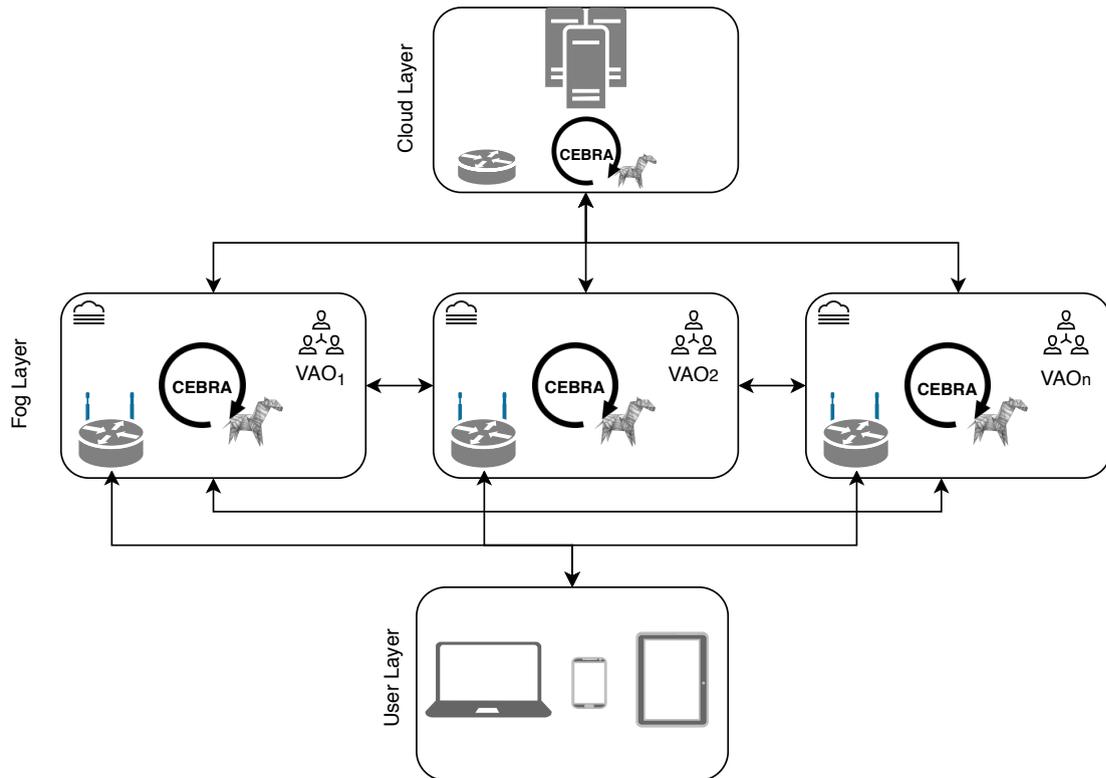


FIG. 3.2: Arquitectura FC que aloja a CEBRA. Cada OV corresponde a una red *Fog* definida en una determinada localización geográfica (por ejemplo, una zona de la ciudad), donde se aloja CEBRA. CEBRA también se encuentra en la capa *Cloud* destinada a trabajar con conjuntos de datos más grandes

Estas OV utilizarán sus capacidades de comunicación y coordinación para compartir los resultados obtenidos por CEBRA. Esto significa que si suponiendo por ejemplo un caso simulado en el que una OV de Zaragoza detecta que un cliente del sistema de Valladolid presenta características similares a un caso de éxito almacenado en el sistema CEBRA de Zaragoza, se comunicará con el sistema de Valladolid para compartir el caso de éxito, mejorando así el sistema de recomendación. Se considera que para gestionar el sistema los agentes son:

1. OV de Recuperación de datos (OVRD): Los agentes se encargan de comunicar y obtener los datos recogidos por los nodos *Fog* situados en la capa *Fog*. Estos agentes están conectados a los sensores distribuidos a través de una capa intermedia o *middleware*.

2. OV de Cliente (OVC): Esta organización contiene CEBRA. Este OVC recupera y gestiona los casos relevantes. En primer lugar, el sistema simula las soluciones aportadas por CEBRA, de modo que sólo se eligen los casos de éxito con mayor índice de aceptación. Esta organización cuenta con un agente de datos del cliente, de modo que el operador del banco puede ver los casos de éxito y puede interactuar con ellos. En esta organización, los agentes conocen las preferencias de los usuarios.
3. OV de Toma de decisiones (OVTD): Esta organización toma sus decisiones a partir de la información recibida de OVRD y OVC. Hay un agente para enviar casos de éxito de una sucursal a otra si se detecta un cliente similar a los alojados en su base de datos y hay otro agente para almacenar los casos enviados por otra sucursal para su propia base de datos. Estas acciones se pueden llevar a cabo mediante la comunicación entre los agentes y los nodos *Fog*.

Lo que se desarrolla en esta subsección presenta como eje principal la incorporación de CEBRA a una OV de agentes para dotar al sistema con la posibilidad de desarrollar un software central más flexible, con gran independencia y modularidad en la aplicación de metodologías de recomendación y de esta manera conseguir proporcionar la solución que mejor se adapte a cada cliente. Así, en función de la retroalimentación de CEBRA, el agente encargado de gestionarlo mejora su eficiencia, rendimiento y capacidad de adaptación al entorno cambiante a medida que se detectan nuevas adquisiciones de productos bancarios.

3.3 CEBRA: modelo de razonamiento basado en casos para la banca comercial

El proceso de la banca minorista implica la recomendación de productos a los clientes; un servicio que es importante en el desarrollo del sector FinTech y que, por tanto, está en constante proceso de crecimiento. La tasa de aceptación de los productos bancarios sería mayor entre los clientes si las recomendaciones estuvieran orientadas a sus necesidades reales. Hoy en día, cualquier empresa con una gran base de clientes y contactos, necesita herramientas especializadas para gestionar los datos y cruzar la información de diferentes bases de datos de forma eficiente. En el caso de los bancos e instituciones financieras esto se vuelve aún más crucial debido a la naturaleza crítica de la información que se

maneja. Estas herramientas comerciales no sólo sirven para fidelizar a los clientes y ofrecer un mejor servicio a los ya existentes, sino que también mejoran la captación de clientes al incidir en el proceso de venta de la organización bancaria o financiera. En este contexto, existen dos tipos diferentes de tecnologías que apoyan la estrategia de relación con el cliente [Al-Weshah, 2017, Jarrar and Neely, 2002]:

1. CRM (*Customer Relationship Management*). Se trata de una herramienta totalmente interna, que proporciona un servicio más rápido y personalizado a los clientes. Hay que tener en cuenta que las soluciones CRM adaptadas al sector financiero suelen ofrecer una serie de garantías en cuanto a la cobertura legal debido a que cuentan con datos muy sensibles y privados de los clientes.
2. Inteligencia de mercado. Proporciona herramientas para la captación, el almacenamiento, el tratamiento, el acceso, la organización y el análisis de los datos de los clientes. La adquisición de nuevos sistemas, como los de modelización predictiva, supone una inversión considerable. Estos sistemas construyen modelos de comportamiento para predecir las tasas de respuesta, las oportunidades de venta cruzada, los potenciales de fraude y los candidatos a crédito.

En la investigación realizada por [Melnychenko et al., 2020], hay tres etapas en la evolución de la banca digital. La primera etapa consiste en la introducción de cajeros automáticos y centros de llamadas en las actividades bancarias, lo que mejoró el servicio al cliente. En la segunda etapa, se aplicaron las tecnologías *Cloud*, las redes sociales, la analítica y el acceso móvil, lo que permitió a las instituciones bancarias personalizar sus servicios bancarios. Y, por último, la tercera etapa implica el desarrollo de la banca digital mediante el uso de la inteligencia artificial, el blockchain, las interfaces de programación y la robotización de los procesos de negocio individuales. Los autores señalan una encuesta de 2019 que muestra que el 76% de los bancos de América del Norte y Europa están tratando de maximizar el uso de soluciones de tecnología financiera para los servicios de pago y lo hacen para retener a los clientes en ciertos segmentos, mientras que el 28% de las empresas encuestadas indican que utilizan proveedores no bancarios para los pagos. La encuesta revela que casi la mitad de las empresas utilizan o están interesadas en utilizar nuevos servicios de pago, centrándose en los pagos en tiempo real (55%), las cámaras de compensación automatizadas (44%) y la implantación de servicios de blockchain (35%).

En este contexto se propone la aplicación de técnicas de razonamiento basadas en casos. Los sistemas de recomendación están diseñados para ofrecer a los usuarios productos personalizados, impulsados por mecanismos inteligentes automatizados que les permiten aprender de las experiencias de usuarios anteriores [Naumov et al., 2019, Nikzad-Khasmakhi et al., 2019]. Entre la amplia gama de tecnologías de ML, este trabajo hace uso del Razonamiento Basado en Casos (CBR) como paradigma de aprendizaje y razonamiento a través de la experiencia. El CBR utiliza un razonamiento automatizado que permite resolver nuevos problemas mediante la adaptación y personalización de soluciones anteriores [Chen and Burrell, 2001]. Utiliza su memoria para resolver nuevos casos, distinguiéndose de otros mecanismos de resolución de problemas por no partir de cero [Jubair et al., 2018]. El CBR se basa en diferentes tipos de métricas de similitud y recomienda los elementos que cumplen los requisitos especificados, y el concepto de crítica es compatible con él [Sridevi et al., 2016].

La motivación de esta investigación consiste en avanzar en el estudio anterior donde se proponía una plataforma de *Fog Computing* para la recomendación de productos bancarios [Hernández-Nieves et al., 2020]. El caso de uso en el trabajo referenciado comenzó con la aplicación del filtrado colaborativo. El filtrado colaborativo es un proceso en el que las valoraciones de los productos se calculan o estiman utilizando la opinión de diferentes personas. Se trata de usuarios con preferencias similares, o productos con valoraciones similares. Por lo tanto, estas recomendaciones se basan en las valoraciones de otros usuarios de esos productos o de otros productos similares. En [Hernández-Nieves et al., 2020] se presentó un caso de uso del algoritmo kNN (*K-Nearest Neighbours*). Este algoritmo predice la valoración de un usuario sobre un producto, teniendo en cuenta las valoraciones realizadas por los usuarios (vecinos) que han realizado valoraciones similares para los mismos productos. En este artículo, se incluye esta técnica en el proceso de reutilización de CEBRA como contrapeso a las recomendaciones obtenidas por el CBR, para que la recomendación final sea la más precisa.

Para la recomendación de productos bancarios, se decidió desarrollar un CBR en lugar de un sistema de recomendación basado en el filtrado colaborativo, porque se pretende ir más allá de un método de clasificación por distancia, como el kNN. Un CBR es capaz de interactuar con el entorno, evaluando sus decisiones en el mundo real. De este modo, el sistema juzga lo buenas que son las soluciones que propone y, en el futuro, evita los

errores que ha cometido antes. Aun así, se entiende que la calidad del CBR dependerá principalmente de 5 factores:

1. De la capacidad de comprender nuevas situaciones a partir de las anteriores.
2. De la experiencia inicial del sistema.
3. De su capacidad de adaptación.
4. De su capacidad de evaluación.
5. De su capacidad para incorporar nuevas experiencias a la base del caso

Esta plataforma requiere un mecanismo de razonamiento y toma de decisiones a nivel local y global. Por ello, en esta investigación se propone CEBRA (*CasE-Based Reasoning Application*). Durante la búsqueda de *frameworks* encontramos Colibri y myCBR como plataformas generales para el desarrollo del Razonamiento Basado en Casos; en [Roth-Berghofer et al., 2012] se describen y comparan ambos. Los autores explicaron que la principal característica de Colibri es que reutiliza sistemas CBR previamente definidos y proporciona un catálogo de sistemas ya desarrollados. myCBR, por su parte, ofrece una estación de trabajo para el desarrollo de modelos de conocimiento para sistemas CBR, como la estructura de casos y las medidas de similitud. Como Colibrí y myCBR son difíciles de integrar en una sola herramienta al no ser librerías integrables, se decidió contribuir al estado del arte proponiendo y desarrollando un nuevo marco integrador que pueda complementar las necesidades de segmentos especializados, incrementando con técnicas de aplicación de contexto en este segmento, como es el bancario en nuestro caso, incluyendo la documentación requerida y las instrucciones para modificar las medidas de similitud en los casos que lo requieran. El *framework* desarrollado durante esta investigación, como Api/REST, ha sido escrito en un código legible y ordenado para que sea posible cualquier tipo de modificación y actualización y por tanto pueda ser reutilizado en otros proyectos [Hernandez-Nieves, 2020]. Para probar el rendimiento de CEBRA, se ha generado un conjunto de datos con 60.000 casos. Una vez creada la base de casos, CEBRA permite definir un perfil para la recomendación, seleccionando el género, la edad, el estado civil, el tipo de trabajo, etc. dando como resultado una lista ordenada de productos bancarios con los mayores índices de aceptación en los casos más similares.

3.3.1 Diseño de CEBRA

Existen diferentes especializaciones o variedades de razonamiento basado en casos, especialmente en lo que respecta a la representación, la indexación o los mecanismos de razonamiento aplicados a los casos [Aamodt and Plaza, 1994, Richter and Weber, 2016]. Normalmente, un sistema CBR se caracteriza por el concepto de caso. Un caso es una pieza de conocimiento contextualizada que representa una experiencia que proporciona una lección fundamental para alcanzar los objetivos del razonador [Kolodner and Leake, 1996], y debe contener un cierto nivel de información y una cierta complejidad en su estructura interna. Este tipo de sistemas son capaces de adaptarse a diferentes entornos o contextos, posibilitando la generalización y permitiendo un cierto grado de independencia del entorno [de Mantaras, 1999]. En esta sección, se presenta al CBR como a un sistema de recomendación adecuado, ya que las aplicaciones comerciales han demostrado un gran éxito con el uso del mismo, sobre todo debido a la ventaja de que las nuevas recomendaciones (es decir, las soluciones) se pueden derivar de las antiguas recomendaciones más fácilmente [Skjold and Øynes, 2017].

El modelo propuesto por Riesbeck and Schank [Riesbeck and Schank, 2013] como uno de sus dos componentes fundamentales, está formado por cuatro procesos secuenciales: recuperar, reutilizar, revisar y retener [De Mantaras et al., 2005] al igual que CEBRA:

1. Recuperación. Es la primera etapa que realiza el sistema CBR. Es aquí donde se realiza la recuperación de casos. Se realizan dos funciones diferentes: el acceso a los casos almacenados y el establecimiento de la similitud entre los casos. Es necesario establecer el algoritmo de acceso a los casos almacenados y las técnicas que permiten determinar la similitud entre los casos.
2. Reutilización. En esta etapa se reciben los casos más similares de la etapa anterior. El objetivo es modificar y combinar o decidir cuál es el más óptimo y reutilizarlo.
3. Revisar. Verificación de la adecuación del caso propuesto en la etapa anterior. Se requiere un sistema de conocimiento experto o un experto humano. El resultado será un nuevo caso si la solución ha sido satisfactoria o la reparación de la solución si ha sido incorrecta.

4. Retener. Consiste en aprender de las nuevas experiencias. Se almacena el caso actual y la solución aplicada para resolverlo. Se asignan eficiencias al caso. A veces puede ser necesario reorganizar la base de casos.

El otro componente fundamental es la base de casos o memoria de casos. De esta base se extraen las soluciones anteriores y se almacena lo aprendido. La base de casos se encarga de mantener la representación y organización de los casos. Debe tener en cuenta la estructura de los casos y debe tratar de facilitar, en la medida de lo posible, cada una de las operaciones del ciclo de vida del CBR.

A continuación se presentan las diferentes etapas de CEBRA. En particular, se incorporan técnicas para mejorar los resultados en las etapas de recuperación y reutilización y que se presentan más adelante. El diagrama del ciclo de vida propuesto se muestra en la Figura 3.3. Como se puede observar, en la etapa de recuperación se incluyen dos siglas: Sim_demographic representa los casos similares relacionados con datos demográficos e indicadores socioeconómicos almacenados en la base de datos; y Sim_SN representa los casos similares relacionados con opiniones y aficiones extraídos de las redes sociales (Twitter).

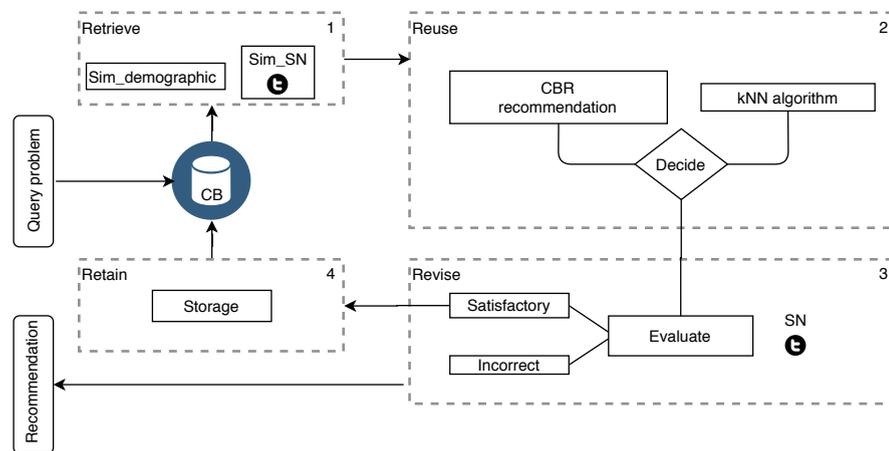


FIG. 3.3: Ciclo de vida de CEBRA

Para construir un sistema CBR hay que utilizar un marco formal. Para ello, se ha tomado como fundamento los trabajos de [Corchado and Laza, 2003, Corchado et al., 2004] donde se establece un formalismo analítico, proporcionándose una notación para los componentes aptitudinales:

1. Conjunto de bases de casos (β). Una base de casos $B \in \beta$ es un conjunto finito de casos que está indexado. Una base de casos se define como una tupla:

$(\{c_1, c_2, \dots, c_n\}, \iota)$. Siendo $\{c_1, c_2, \dots, c_n\}$ los casos que conforman la base de casos e ι el conjunto finito de características que permite indexar los casos.

2. Un caso (c) está representado por una secuencia de estados ambientales:
 $c = \{\text{start_state}, \{\text{action} \times [\text{intermediate_state}]\}^+, \text{end_state}\}$. Cada estado está representado por un conjunto de atributos que definen el entorno en el que se encuentra el sistema CBR. Los estados se dividen en tres grupos:
 - (1) Conjunto de estados iniciales (ini_state) representando la descripción del problema a resolver.
 - (2) Conjunto de estados intermedios (intern_state) que describen los diferentes estados por los que pasa el entorno antes de alcanzar el estado final.
 - (3) Conjunto de estados finales (final_state) que representan la descripción del entorno una vez alcanzados los objetivos iniciales.

Además de estados, un caso contiene acciones, que representan el conjunto de acciones aplicadas a cada uno de los estados. Se definen mediante un nombre y un conjunto de argumentos.

3. Un conjunto finito de atributos (k) es un conjunto de propiedades que permiten describir un estado.
4. Un conjunto de índices (I) es un conjunto de características de ι , con ι incluida en k .
5. Un conjunto de funciones de similitud (A) permite determinar el grado de similitud entre un problema a resolver y un caso.

La investigación comenzó con una introducción de los atributos y su importancia, basándonos en la metodología descrita por Richter, M. M., *et al.* en [Richter and Weber, 2016] para definir las fases para desarrollar el sistema CBR. Las fases a tener en cuenta son:

1. Identificar un escenario adecuado de alto nivel.
2. Introducir las características y tipos que serán aplicables a los productos bancarios para la clase elegida.

3. Definir la importancia de los atributos para la clase de cliente a ponderar.
4. Definir medidas de similitud a nivel local.
5. Definir las normas de finalización y adaptación según sea necesario.
6. Construir una base de casos.

3.3.1.1 Etapa de recuperación de CEBRA

Para comenzar el caso de uso que da cabida a la aplicación CEBRA, se comenzó con la construcción de una base de casos, estando formado cada caso por 41 atributos. Los atributos del caso considerados por el CBR, que se muestran más adelante en este ejemplo, se han obtenido de una fuente que no podemos revelar por razones de confidencialidad. El diccionario consta de 606 tablas organizadas en 30 categorías, cada una de las cuales se compone de otro subconjunto. Dentro de estos subconjuntos de tablas, el número de sub-tablas varía. Por ejemplo, la tabla número 1 "*Activos*" se compone de 47 sub-tablas, mientras que la tabla número 2 "*Auxiliar*" se compone de 254 sub-tablas. La selección de los atributos del caso dentro del diccionario puede representarse de esta manera:

$$\begin{aligned} \text{ATR} = \{ & T_1(\text{Sub42}) + T_2(\text{Sub250}) + T_6(\text{Sub8}, \text{Sub12}) \\ & + T_11(\text{Sub6}) + T_14(\text{Sub5}) + T_25(\text{Sub10}) \} \end{aligned} \quad (3.1)$$

A modo de resumen se presenta La tabla 3.1, donde se pueden observar los atributos considerados para la construcción del caso, organizados en tres columnas: la primera columna corresponde a las categorías que han sido seleccionadas dentro del diccionario de 606 tablas mencionado anteriormente, la segunda columna muestra el subconjunto elegido dentro de la categoría anterior, y finalmente, la tercera columna muestra los atributos que conforman un caso en CEBRA. Además de los atributos de caso considerados, se incluyen los intereses expresados por los usuarios en sus perfiles de redes sociales en la etapa de recuperación (Tabla 3.2). Estos intereses se extraerán aplicando clasificadores al contenido textual. Aunque se extraen en la primera fase del ciclo CBR, los intereses extraídos no se utilizarán en esta fase, sino que se proporcionarán

a los revisores expertos para completar el caso. Este razonamiento se explicará con más detalle en la fase de revisión.

TAB. 3.1: Atributos del caso de uso de CEBRA en el diccionario de la base de datos

Categoría	Subconjuntos	Atributos
Activos	Persona	Edad, Género, Estado Civil, Ciudad, Hijos, Trabajo, Tipo de contrato, Nº de casas en propiedad, Nº de coches en propiedad, Nº de vehículos de otro tipo en propiedad
Auxiliar	Producto	Solicitó una hipoteca, Hipoteca concedida, Devolución anticipada de la hipoteca, Primera vez que realiza una devolución anticipada de la hipoteca, Primera vez que solicita una hipoteca, Solicitó un préstamo, Préstamo concedido, Primera vez que solicita un Préstamo, Solicitó un depósito, Primera vez que solicita un depósito
Contabilidad	Tipo de cuenta Balance mensual	Cuenta corriente, Cuenta nómina, Cuenta de ahorro Saldo positivo a final de mes
Fondos	Contratos	Solicita un fondo de inversión, Solicita por primera vez un fondo de inversión
Incumplimiento	Balance de situación	¿Ha habido alguna vez un impago?, Impagos habituales, Impagos ocasionales
Seguros	Producto	Solicita un seguro de hogar, Solicita por primera vez un seguro de hogar, Solicita un seguro de accidentes, Solicita por primera vez un seguro de accidentes, Solicita un seguro de vida, Solicita por primera vez un seguro de vida, Solicita un seguro de salud, Solicita por primera vez un seguro de salud, Solicita un seguro de coche/moto, Solicita por primera vez un seguro de coche/moto

TAB. 3.2: Ejemplos de intereses de clientes que pueden ser extraídos por CEBRA de *Twitter* en la etapa de recuperación.

Categorías	Atributos
Entretenimiento	Películas, TV, Radio, Festivales de música y conciertos, Teatro y musicales
Tiempo libre	Viajes, Deportes
Automóviles	Coches, Motos, F1, Grandes premios, Dakar, Moto GP
Mercado mundial	Empresas, Marcas, Productos
Comida y bebida	Restaurantes, ¿Cómo cocinar?, Guía MICHELIN, <i>Haute cuisine</i>
Gaming	Nuevos juegos, Videoconsolas
Salud	Comida sana, Dieta, Vegano, Yoga, <i>Wearable</i>
Estilo y moda	Ropa, Marcas de ropa, Tendencias, Revistas de moda
Interiorismo y hogar	Muebles, Reformas de la casa, Jardinería, Tendencias de decoración

3.3.1.1.1 Cálculo de las funciones de similitud

Una vez cubiertos todos los atributos del caso, se ha considerado también el problema de la consulta (Table 3.3). Se ha generado un perfil aleatorio respondiendo a los atributos del caso (k). A continuación se presenta una comparación entre el problema de la

consulta (ini_stateQ) y varias descripciones de los problemas a resolver (ini_state1), (ini_state2), (ini_state3), (ini_state4), (ini_state5).

TAB. 3.3: Atributos del caso en el problema de la consulta

ini_stateQ	k	ini_state1	ini_state2	ini_state3	ini_state4	ini_state5
34	Edad	45	68	25	38	74
F	Género	M	M	M	F	M
C	Estado Civil	C	C	S	C	V
Sa	Ciudad	Sa	Sa	Sa	Sa	Sa
1	Hijos	2	0	0	2	3
S	Trabajo	S	J	S	S	J
T	Tipo de contrato	F	P	T	C. F.	P
1	Nº de casas en propiedad	1	1	0	0	1
0	Nº de coches en propiedad	1	1	0	1	0
0	Nº de otro tipo de vehículos en propiedad	0	0	1	0	0
S	Cuenta corriente	S	S	S	S	S
S	Cuenta nómina	S	S	S	S	S
N	¿ha habido alguna vez un incumplimiento?	N	S	S	N	S
N	Incumplimientos habituales	N	N	N	N	N
N	Incumplimientos ocasionales	N	S	S	N	S
H	Saldo positivo a final de mes (habitual, medio, bajo)	H	B	M	H	H
N	Cuenta de ahorros	N	S	N	N	S
S	Tarjeta de débito	S	S	S	S	S
S	Tarjeta de crédito	S	N	N	S	S
S	Solicitó una hipoteca	N	N	N	S	N
S	Hipoteca concedida	n/a	n/a	n/a	S	n/a
N	Devolución anticipada de la hipoteca	n/a	n/a	n/a	N	n/a
n/a	Es la primera vez que se hace una devolución anticipada de la hipoteca	n/a	n/a	n/a	n/a	n/a
S	Es la primera vez que se solicita una hipoteca	n/a	n/a	n/a	S	n/a
N	Solicitó un préstamo	S	N	S	N	S
n/a	Préstamo concedido	S	n/a	S	n/a	S
n/a	Es la primera vez que solicita un préstamo	S	n/a	S	n/a	N
N	Solicitó un depósito	N	N	N	N	S
n/a	Primera vez que se solicita un depósito	n/a	n/a	n/a	n/a	S
N	Solicitó un fondo de inversión	N	S	N	N	S
n/a	Es la primera vez que solicita un fondo de inversión	n/a	N	n/a	n/a	N
S	Solicitó un seguro de hogar	S	N	N	S	N
S	Es la primera vez que solicita un seguro de hogar	S	n/a	n/a	S	n/a
N	Solicitó un seguro de accidentes	N	N	N	N	N
n/a	Es la primera vez que se solicita un seguro de accidentes	n/a	n/a	n/a	n/a	n/a
S	Solicitó un seguro de vida	S	N	N	S	N
S	Es la primera vez que solicita un seguro de vida	S	n/a	n/a	S	n/a
N	Solicitó un seguro de salud	N	N	N	S	N
n/a	Es la primera vez que solicita un seguro de salud	n/a	n/a	n/a	S	n/a
N	Ha solicitado un seguro de coche/motocicleta	N	N	S	N	N
n/a	Es la primera vez que solicita un seguro de coche/moto	n/a	n/a	S	n/a	n/a

Género: F (Femenino), M (Masculino)
Estado Civil: C (Casado), S (Soltero), V (Viudo)
Ciudad: Sa (Salamanca, Spain)
Tipo de contrato: F. (Funcionario), C. F. (Contrato Fijo), P (Pensionista), T (Temporal)
Saldo positivo de la cuenta a final de mes: H (Habitual), M (Medio), B (Bajo)

Cada atributo del caso requiere su propia función de similitud. Como regla general, se da un valor de similitud de 1 cuando los valores de dos atributos son iguales; y un valor de similitud de 0 cuando los valores no son iguales. En este sentido, para dos valores que son diferentes entre sí pero que pueden considerarse medianamente similares, se puede utilizar un valor de 0,5 [Richter and Weber, 2016]. En este apartado se describen los valores de similitud que son más difíciles de definir, es decir, que difieren de 0 y 1. Para definir las funciones de similitud para cada uno de los atributos, en el caso de los números cuantitativos como la edad, se construye una forma analítica para cada uno de ellos:

- (i) La edad. Se establece una relación lineal con el valor absoluto de la diferencia de

edad, siendo 0 la similitud en los casos en los que la edad es de 40 años. En los casos en los que la diferencia es mayor, la similitud sigue siendo cero.

Formalmente:

$$\text{sim}_{\text{edad}}(x, y) = \max \left\{ 1 - \frac{|x - y|}{40}, 0 \right\} \quad (3.2)$$

El comportamiento de esta función se muestra en la figura 3. Obsérvese que, aunque podrían definirse modelos más complicados para captar con mayor precisión las diferencias entre las edades, se ha optado por el modelo uniparamétrico que se presenta aquí por simplicidad. En futuras revisiones cabría la posibilidad de modelar una función de similitud más complicada, siempre que esté respaldada por datos experimentales.

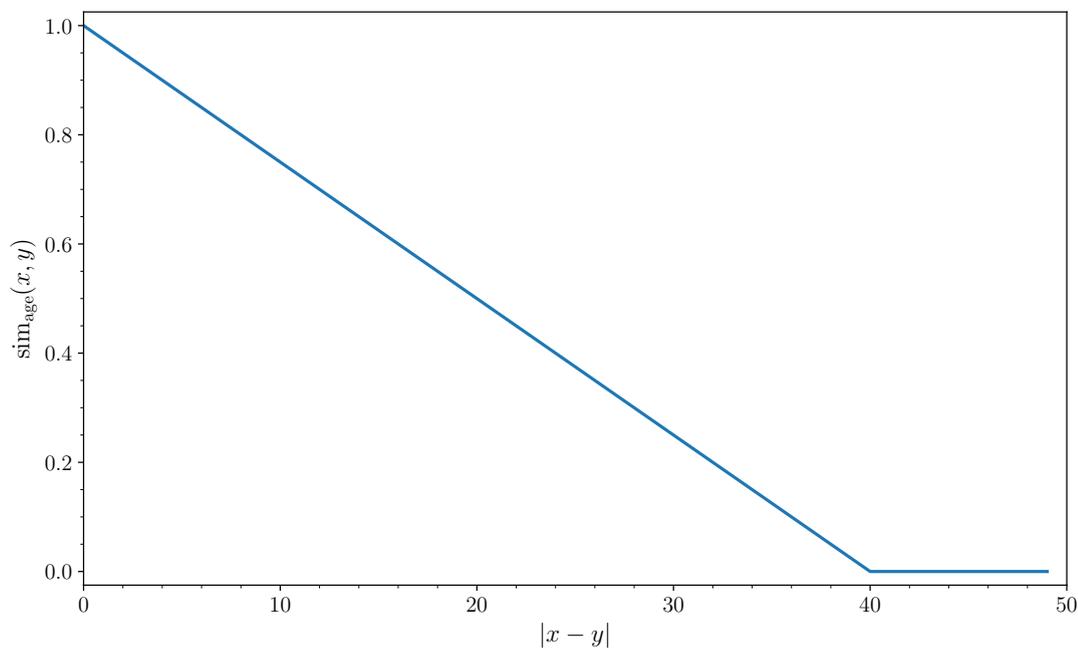


FIG. 3.4: Función de similitud de edad

- (ii) Hijos. La relación de similitud se establece de forma que para un valor de 0 hijos contra 0 hijos se da un valor de 1, para un valor de 1 contra 0 hijos se da un valor de 0,5; para un valor de dos hijos contra un valor de 0 se da un valor de 0,13; etc.

$$\text{sim}_{\text{hijos}}(x, y) = \left(\frac{1}{2} \right)^{|x-y|}. \quad (3.3)$$

La motivación de esta expresión es modelar un rápido decaimiento geométrico con las diferencias en los valores del atributo. Expresado como una matriz sería

$$\text{sim}_{\text{hijos}}(x, y) = \begin{pmatrix} 1 & 0.5 & 0.25 & 0.13 & \dots \\ 0.5 & 1 & 0.5 & 0.25 & \dots \\ 0.25 & 0.5 & 1 & 0.5 & \dots \\ 0.13 & 0.25 & 0.5 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (3.4)$$

Para los valores nominales o discretos, se define la matriz de similitud simétrica:

- (i) Trabajo. Se considera una similitud de 1 para la comparación con los estudiantes; a una persona sin trabajo frente a un estudiante se le da un valor de 0,9; el valor 0 se da a una persona que tiene trabajo frente a un estudiante. También se da un valor de 0,9 a una persona jubilada frente a una persona con trabajo, [Richter and Weber, 2016].

$$\text{sim}_{\text{trabajo}}(x, y) = \begin{pmatrix} 1 & 0.9 & 0 & 0 \\ 0.9 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0.9 \\ 0 & 0 & 0.9 & 1 \end{pmatrix} \quad (3.5)$$

- (ii) Tipo de contrato. En este caso, se ha establecido un orden de prioridad: los contratos de funcionariado se consideran el tipo más importante, seguido del contrato permanente, la pensión y el temporal. Por tanto, la comparación entre el tipo de contrato 1 y el 2 daría un valor de 0,66; el tipo de contrato 1 comparado con el 3 daría un valor de 0,33 y el tipo de contrato 1 comparado con el 4 daría un valor de 0.

Formalmente:

$$\text{sim}_n = 1 - \frac{|x - y|}{n - 1} \quad (3.6)$$

Donde x e y son el orden en la clasificación y n el número de categorías contenidas.

Cuando se expresa como una matriz:

$$\text{sim}_{\text{contrato}}(x, y) = \begin{pmatrix} 1 & 0.66 & 0.33 & 0 \\ 0.66 & 1 & 0.66 & 0.33 \\ 0.33 & 0.66 & 1 & 0.66 \\ 0 & 0.33 & 0.66 & 1 \end{pmatrix} \quad (3.7)$$

- (iii) Para trabajar con atributos en los que puede no estar definido un valor relevante, ampliamos su rango de valores para incluir una opción especial: n/a (no aplicable). Al trabajar con otros valores para medir la similitud entre atributos se aplicará la regla general:

$$\text{sim}(n/a, x) = n/a \quad (3.8)$$

Por último, al tomar la media ponderada de los valores que da lugar a la similitud total, no se tienen en cuenta los componentes individuales con valor n/a, es decir,

$$\text{sim}(x, y) = \frac{\sum_{i \in I} \text{sim}_i(x, y) w_i}{\sum_{i \in I} w_i}, \quad (3.9)$$

donde I es el conjunto de atributos con valores no nulos.

Así, cada atributo será ponderado, ya que ninguno de los atributos del caso (k) es igualmente relevante. Además, el problema de escribir la importancia de los atributos del caso también se denota mediante un valor numérico, como se muestra en la Tabla 3.4.

Como resumen de este subapartado, podría decirse que el trabajo propuesto por (Richter *et al.* 2016) se utiliza como punto de partida para establecer las funciones de similitud. Los pesos asignados a las mismas surgen de la intuición y pueden ser perfeccionados en un escenario real. Lo ideal sería utilizar un conjunto de datos reales de recomendaciones existentes, para poder ajustarlas formalmente.

El proceso de recuperación formal se resume en Alg. 1.

TAB. 3.4: Ponderación e importancia de los atributos del caso dentro de cada problema definido.

k	ini_state1	ini_state2	ini_state3	ini_state4	ini_state5	Importancia
Edad	0,72	0,15	0,77	0,99	0	3
Género	0	0	0	1	0	3
Estado Civil	1	1	0	1	0,5	3
Ciudad	1	1	1	1	1	4
Hijos	0,5	0,5	0	0,5	0,25	2
Trabajo	1	0,9	1	1	0,9	7
Tipo de contrato	0	0,66	1	0,33	0,66	7
Nº de casa en propiedad	1	1	0	0	1	4
Nº de coches en propiedad	0	0	1	0	1	4
Nº de otros vehículos en propiedad	1	1	0	1	1	4
Cuenta corriente	1	1	1	1	1	1
Cuenta nómina	1	1	1	1	1	1
¿Ha habido alguna vez un impago?	1	0	0	1	0	3
Incumplimientos habituales	1	1	0	1	1	5
Incumplimientos ocasionales	1	0	0	1	0	4
Saldo positivo a final de mes (habitual, medio, bajo)	1	0	0,5	1	1	8
Cuenta de ahorros	1	0	1	1	0	3
Tarjeta de débito	1	1	1	1	1	2
Tarjeta de crédito	1	0	0	1	1	2
Solicitó una hipoteca	0	0	0	1	0	9
Hipoteca concedida	n/a	n/a	n/a	1	n/a	3
Devolución anticipada de la hipoteca	n/a	n/a	n/a	1	n/a	3
Primera vez que se hace una devolución anticipada de la hipoteca	n/a	n/a	n/a	n/a	n/a	3
Es la primera vez que se solicita una hipoteca	n/a	n/a	n/a	1	n/a	3
Solicitó un préstamo	0	1	0	1	0	9
Préstamo concedido	n/a	n/a	n/a	n/a	n/a	3
Es la primera vez que solicita un préstamo	n/a	n/a	n/a	n/a	n/a	3
Solicitó un depósito	1	1	1	1	0	9
Primera vez que se solicita un depósito	n/a	n/a	n/a	n/a	n/a	3
Solicitó un fondo de inversión	1	0	1	1	0	9
Es la primera vez que solicita un fondo de inversión	n/a	n/a	n/a	n/a	n/a	3
Solicitó un seguro de hogar	1	0	0	1	0	9
Es la primera vez que solicita un seguro de hogar	1	0	0	1	0	3
Solicitó un seguro de accidentes	1	1	1	1	1	9
Es la primera vez que se solicita un seguro de accidentes	n/a	n/a	n/a	n/a	n/a	3
Solicitó un seguro de vida	1	0	0	1	0	9
Es la primera vez que solicita un seguro de vida	1	n/a	n/a	1	n/a	3
Solicitó un seguro de salud	1	1	1	0	1	9
Es la primera vez que solicita un seguro de salud	n/a	n/a	n/a	n/a	n/a	3
Ha solicitado un seguro de coche/motocicleta	1	1	0	1	1	9
Es la primera vez que solicita un seguro de coche/moto	n/a	n/a	n/a	n/a	n/a	3
Similitud global [Eq. (3.9)]	0.784	0.528	0.463	0.862	0.486	

Algorithm 1 Recuperación de k usuarios más parecidos a c_0

```

function RECUPERAR( $c_0, k$ )
   $l$  List() ▷ Lista de similitudes
  for  $c \in CB$  do ▷ Para cada caso de la base de casos
    Append( $l, \text{sim}(c_0, c)$ ) ▷ Calcular la similitud (Eq. 3.9)
  end for
   $l$  ArgSort( $l$ ) ▷ Recuperar los índices con los valores en orden decreciente
   $C \leftarrow \text{List}()$  ▷ Lista de casos de salida
  for  $i \in \{1, \dots, k\}$  do
    Append( $C, CB[i]$ ) ▷ Busca el  $i$ -ésimo caso más similar
  end for
  return  $C$ 
end function

```

3.3.1.2 Etapa de reutilización de CEBRA. Rendimiento de la clasificación*K-Nearest Neighbor*

En esta fase, el CBR hace una recomendación basada en los productos más comprados por los vecinos más cercanos en función de su similitud. Para mejorar la capacidad

de recomendación del sistema, pueden añadirse adicionalmente otras técnicas de recomendación, como el filtrado colaborativo [Hernández-Nieves et al., 2020].

3.3.1.3 Fase de revisión de CEBRA

Como se ha explicado anteriormente, el experto dispondrá de los intereses que los usuarios han compartido públicamente en sus redes sociales, de forma que se pueda comprobar la idoneidad de la propuesta. La recomendación podrá refinarse comprobando perfiles similares con preferencias e intereses parecidos que hayan comprado productos que el sujeto de estudio no haya contratado. En este ejemplo, como se muestra en la Tabla 3.4), se observa que el problema de la Consulta muestra una similitud general con el problema 4 (*ini_state4*) de 0.862. Observamos que la única diferencia es que el sujeto del problema 4 no tiene seguro médico, por lo que la recomendación dada en nuestro problema de consulta (*ini_stateQ*) definido en la Tabla 3.3, sería adquirirlo. Teniendo en cuenta los intereses extraídos de la red social del usuario, podría establecerse también que el sujeto es una mujer, que le gusta el deporte, y que es aficionada a la escalada, por lo que la recomendación podría estar orientada a un seguro de accidentes además de contratar un seguro de salud.

3.3.2 Experimentación y resultados

CEBRA se ha implementado como una API/REST, como se muestra en la Figura 3.5, para Swagger UI [Varanasi and Belida, 2015]. Se ha generado un conjunto de datos sintéticos con 60000 casos para probar el funcionamiento de CEBRA y se han utilizado herramientas de simulación para la construcción de la base de casos. CEBRA es una aplicación desarrollada para que las entidades bancarias puedan recomendar productos de forma eficaz y aumentar sus ventas. El código fuente está disponible en un repositorio público de GitHub [Hernandez-Nieves, 2020]. El objetivo era construir una base de casos que contuviera atributos realmente utilizados por los bancos, la parte que realmente pertenece a un banco es el diccionario del que extrajimos los atributos, por lo que no se han utilizado datos reales para construir la base de casos. Hay que tener en cuenta que la investigación se centra en la arquitectura y el diseño.

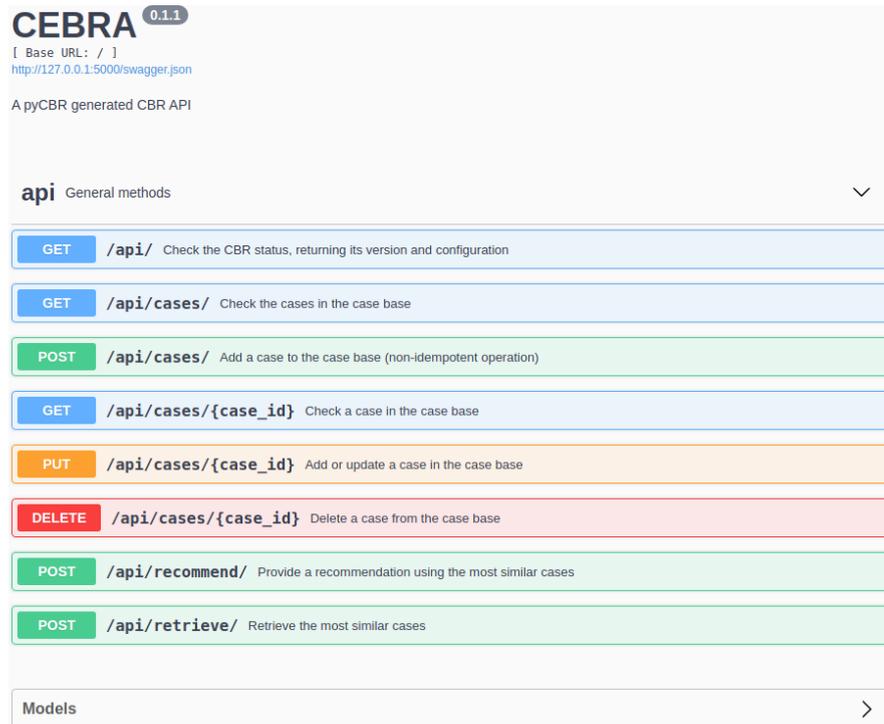


FIG. 3.5: CEBRA como una API/REST, visualizada mediante Swagger UI

La Api permite recuperar los usuarios más parecidos a aquel al que se le quiere hacer la recomendación. Por lo tanto, este *endpoint* *Api/retrieve* corresponde a la etapa de recuperación. Además, el *endpoint* *api/recommend* proporciona una recomendación al usuario. Esta recomendación podría completarse con un sistema externo a CEBRA, como se muestra en la etapa de reutilización de la Figura 3.3. La fase de revisión se llevaría a cabo de forma externa al sistema, donde un equipo de expertos utilizaría toda la información sobre los casos recuperados, además de la medición de las similitudes con los casos en los que se ha basado la recomendación. Por último, cuando se ha hecho una recomendación positiva, los casos en CEBRA pueden actualizarse con el *endpoint* */api/cases/{case_id}*, que correspondería a la etapa de retención. La API se completa con el *endpoint* */api/cases*, que permite consultar la base de casos de forma global, y */api/*, que permite comprobar la configuración de CEBRA, incluyendo aspectos como las medidas de similitud que incorpora.

Cuando se desarrolla software es necesario definir requisitos y verificarlos. Los requisitos de usuario, normalmente definidos mediante lenguaje natural, tablas y diagramas, son descritos para que puedan ser entendidos por los usuarios. Los requisitos pueden tener varios orígenes, como el dominio del problema (requisitos del dominio). Los requisitos de dominio son requisitos de usuario que describen las características y necesidades del

dominio (comunes a todas las organizaciones de ese sector). Los problemas que pueden surgir con este tipo de requisitos son principalmente la inteligibilidad y los equívocos. Los de comprensión porque los requisitos utilizan el lenguaje y el vocabulario habitual en el dominio de la aplicación y no son correctamente entendidos por los ingenieros de software que van a desarrollarla y, por supuesto, porque en muchas ocasiones se ignoran los requisitos de dominio por ser perfectamente conocidos por los expertos en el área. El sistema CBR CEBRA incluye varios requisitos de dominio:

1. El usuario (la entidad bancaria) debe tener en cuenta el Reglamento General de Protección de Datos (UE 2016/279) relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (este texto incluye la corrección de errores publicada en el DOUE de 23 de mayo de 2018).
2. Debe tenerse en cuenta el Reglamento (UE) 2018/1725 por el que se establecen las normas aplicables al tratamiento de datos personales por las instituciones, órganos y organismos de la Unión.
3. También hay que tener en cuenta, si procede, la normativa específica de cada país fuera de la Unión Europea.

3.4 Una plataforma de análisis y extracción de datos para recomendaciones de inversión

El análisis de datos es un proceso de inspección, limpieza, transformación, clasificación y modelización de datos con el fin de encontrar información útil, llegar a conclusiones y tomar decisiones adecuadas. En estadística, el análisis de datos se divide en análisis descriptivo, análisis exploratorio y análisis predictivo.

La analítica predictiva se define como la rama de la analítica utilizada para hacer predicciones sobre eventos futuros a los que se enfrenta, por ejemplo, una organización. Para ello, utilizará diversos métodos como la minería de datos, la minería de textos, la inteligencia artificial, la estadística o el modelado de datos, entre otros. Además, la analítica predictiva gestiona las tecnologías de la información, los métodos de análisis y

la modelización de los procesos de negocio con el fin de anticiparse a los acontecimientos futuros que puedan suceder a la organización en cuestión.

En esta investigación, la atención se centra en el análisis predictivo con un enfoque específico hacia el análisis bursátil. Se parte de la base de que una predicción bursátil se considera acertada si consigue los mejores resultados utilizando la mínima entrada de datos y el modelo bursátil menos complejo [Atsalakis and Valavanis, 2009]. Dentro del campo de la Inteligencia Artificial, la aparición del ML y el aumento de las prestaciones informáticas han permitido desarrollar nuevos servicios sobre la base de los productos financieros tradicionales, proporcionando instrumentos económico-financieros que aportan mayor versatilidad y rapidez. Como señala la investigación realizada en [Patel et al., 2015], predecir el valor de una acción es difícil por la incertidumbre que conlleva el gran número de determinantes potenciales. Los autores proponen un método que incluye tanto el análisis fundamental como el técnico combinado con algoritmos de ML, para tratar de mejorar su eficacia.

A diferencia de otras investigaciones que se centran en un único modelo para la recomendación de inversiones, como las Redes Neuronales Artificiales o los árboles de decisión optimizados, en esta investigación se aplican una serie de algoritmos (*Random Forest Regressor*, *Gradient Boosting Regressor*, *SVM-LinearSVR*, *MLP Regressor*, *kNNNeighbors Regressor*) junto con el análisis técnico, combinando Indicadores *Momentum* y Medias Móviles. El sistema de recomendación propuesto eliminará la subjetividad del proceso después de evaluar y validar los algoritmos y proporcionará al usuario el algoritmo con la mejor precisión. Sin embargo, la principal ventaja del sistema investigado consiste en la posibilidad de consultar todo el proceso que ha realizado el sistema (análisis, predicción y recomendaciones de inversión).

En esta sección se describe el desarrollo de un sistema de recomendación para obtener señales de compra/venta a partir de los resultados del análisis técnico y de las previsiones realizadas para las empresas que operan en el mercado continuo español.

Tiene un diseño modular para facilitar la escalabilidad del modelo y la mejora de las funcionalidades. Los módulos de los que consta son: análisis y minería de datos, el sistema de previsión, el módulo de análisis técnico, el sistema de recomendación y la plataforma de visualización. Se presenta la especificación de cada módulo, así como las dependencias y la comunicación entre ellos. Además, la propuesta incluye una

plataforma de visualización para la interacción de alto nivel entre el usuario y el sistema de recomendación. Esta plataforma presenta las conclusiones abstraídas de los valores resultantes.

3.4.1 Modelo propuesto

Tiene un diseño modular para facilitar la escalabilidad del modelo y la mejora de las funcionalidades. Los módulos de los que consta son: análisis y minería de datos, el sistema de previsión, el módulo de análisis técnico, el sistema de recomendación y la plataforma de visualización. La propuesta incluye una plataforma de visualización para la interacción de alto nivel entre el usuario y el sistema de recomendación. Esta plataforma presenta las conclusiones abstraídas de los valores resultantes.

1. Análisis y extracción de datos. El módulo de extracción de datos funciona de la siguiente manera: primeramente, se recibe una petición del usuario a través de un API *endpoint*. En segundo lugar, si el sistema ha recibido esa petición, incluirá en la cabecera el nombre de la empresa y el rango de fechas (si se han solicitado los datos históricos) o sólo extraerá el nombre de la empresa (si no se han solicitado los datos históricos). Una vez extraídos los datos históricos de una acción, se puede proceder al análisis de los datos.
2. El sistema de previsión. Para predecir el comportamiento futuro de una acción se aplican los algoritmos de regresión de ML [Arora et al., 2019, Khaidem et al., 2016, Pimprikar et al., 2017, Soni, 2011]. El objetivo es determinar el precio de cierre de la bolsa, para ello se ha definido el conjunto de valores de apertura como variables de entrada y el conjunto de valores de cierre como variables de salida, es decir, los valores de cierre son la variable objetivo del algoritmo. Dada la naturaleza del problema, es necesario aplicar algoritmos de regresión. Esto se debe a que cuando se trabaja con datos continuos, los algoritmos de regresión pueden indicar el patrón en un conjunto de datos dado. En consecuencia, estos algoritmos se aplican en los casos en los que se quiere modelar la relación entre una variable dependiente escalar o variable objetivo Y y una o más variables de entrada X . La siguiente lista describe los algoritmos utilizados por el sistema para predecir el último valor de cierre (desconocido) basado en datos históricos del mercado, a partir del último valor de apertura (conocido):

- (a) *Random Forest Regressor*: estos algoritmos son un método de ML para la clasificación, la regresión y otras tareas. Consiste en un metaestímulo que ajusta una serie de árboles de decisión de clasificación en varias submuestras del conjunto de datos y utiliza los medios para mejorar la precisión productiva y el control del ajuste.
 - (b) *Gradient Boosting Regressor*: técnica de ML que construye el modelo de forma escénica, al igual que los métodos basados en el refuerzo. Generaliza los modelos permitiendo la optimización de una función de pérdida arbitraria y diferenciable.
 - (c) *SVM-LinearSVR*: modelo de ML que analiza los datos para la clasificación y el análisis de regresión. El algoritmo SVM se basa en la construcción de un modelo de asignación de nuevos ejemplos a una u otra categoría, por lo que es un clasificador lineal binario no probabilístico. En SVR se intenta ajustar el error dentro de un determinado umbral.
 - (d) *MLP Regressor*: es un tipo de red neuronal artificial de retroalimentación. Utiliza una técnica de aprendizaje supervisado denominada retropropagación para la construcción de la red. Además, sus múltiples capas y su activación no lineal distinguen a MLP de un perceptrón lineal. También permite distinguir datos que no son linealmente separables.
 - (e) *KNNeighbors Regressor*: método no paramétrico utilizado para la clasificación y la regresión.
3. Análisis técnico. A partir de los datos históricos de las empresas del mercado continuo español, se realiza un análisis técnico del mercado, en este caso combinando Indicadores de *Momentum* y Medias Móviles. Esto se realiza para varias ventanas temporales previamente definidas para cada uno de los diferentes factores a calcular en función del estándar del tamaño de las ventanas temporales. El análisis técnico es un análisis utilizado para sopesar y evaluar las inversiones. Identifica oportunidades para adquirir o vender acciones en función de las tendencias del mercado. A diferencia del análisis fundamental, que intenta determinar el precio exacto de una acción, el análisis técnico se centra en la detección de tendencias o patrones en el comportamiento del mercado para la identificación de señales de compra o venta de activos, junto con diversas

representaciones gráficas que ayudan a evaluar la seguridad o el riesgo de una acción [Edwards et al., 2018].

Este tipo de análisis puede utilizarse en cualquier producto financiero siempre que se disponga de datos históricos. Es necesario incluir tanto los precios de las acciones como el volumen. El análisis técnico se emplea muy a menudo cuando se requiere un análisis a corto plazo, por lo que puede ayudar a abordar adecuadamente el problema planteado en esta investigación, donde se predice el valor de cierre de una acción en un día. En el análisis se tienen en cuenta los siguientes indicadores [Dash and Dash, 2016]:

- (a) *Relative Strength Index* (RSI): se trata de un Indicador de *Momentum* (estos indicadores reflejan la diferencia entre el precio de cierre actual y el precio de cierre de los N días anteriores), que mide el impacto de los cambios frecuentes en el precio de una acción, identificando las señales de sobrecompra o sobreventa. La representación del RSI se muestra en un oscilador, es decir, una línea cuyo valor oscila entre dos extremos, que en este caso está entre 0 y 100.

$$\text{RSI}_{\text{primer paso}} = 100 - \left[\frac{100}{1 + \frac{\text{Ganancia media}}{\text{Pérdida media}}} \right] \quad (3.10)$$

- (b) *Stochastic Oscillator* (STOCH): se trata de un Indicador de *Momentum* que compara el precio de cierre de una acción en un día determinado con el rango de valores de cierre de esa acción durante un determinado periodo de tiempo, definido por la ventana temporal. Se puede ajustar la precisión del STOCH mediante la realización de ajustes en la ventana temporal o mediante el cálculo de la media móvil del resultado del indicador. Al igual que el RSI, identifica las señales de sobrecompra o sobreventa de las acciones dentro de un rango de 0 a 100 valores posibles.

$$\%K = 100 - \left(\frac{C - L14}{H14 - L14} \right) \times 100 \quad (3.11)$$

Donde C es el precio de cierre más reciente, $L14$ es el precio más bajo negociado de las 14 sesiones anteriores, $H14$ es el precio más alto negociado durante el mismo período de 14 días y $\%K$ es el valor actual del indicador estocástico.

- (c) *Ultimate Oscillator* (ULTOSC): se trata de un Indicador de *Momentum* usado para medir la evolución de una acción a lo largo de una serie de marcos temporales utilizando una media ponderada de 3 ventanas o marcos temporales diferentes. Por lo tanto, adquiere una menor volatilidad e identifica menos señales de compra-venta que otros osciladores que sólo dependen de un único marco temporal. Cuando las líneas generadas por el ULTOSC divergen de los valores de cierre de una acción, se identifican señales de compra y venta para la misma.

$$UO = \left[\frac{(A_7 \times 4) + (A_{14} \times 2) + A_{28}}{4 + 2 + 1} \right] \times 100 \quad (3.12)$$

Donde UO es el Ultimate Oscillator y A es la media. El cálculo de la media sigue las siguientes fórmulas.

$$A_7 = \left[\frac{\sum_{p=1}^7 BP}{\sum_{p=1}^7 TR} \right] \quad (3.13)$$

$$A_{14} = \left[\frac{\sum_{p=1}^{14} BP}{\sum_{p=1}^{14} TR} \right] \quad (3.14)$$

$$A_{28} = \left[\frac{\sum_{p=1}^{28} BP}{\sum_{p=1}^{28} TR} \right] \quad (3.15)$$

Donde BP es la presión de compra y PC es el cierre previo

$$BP = Close - Min(Bajo, PC) \quad (3.16)$$

Donde TR es el Rango Verdadero

$$TR = Max(Alto, Cierre previo) - Min(Bajo, Cierre previo) \quad (3.17)$$

- (d) *Williams %R* (WILLR): También conocido como *Williams Percent Range*, es un Indicador de *Momentum* que fluctúa entre -100 y 0. Sirve para medir e identificar los niveles de sobrecompra o sobreventa de acciones. Es muy similar al STOCH en su uso y se utiliza con el mismo propósito. Su función es la comparación del valor de cierre de una acción con el rango entre los

valores máximos y mínimos dentro de un marco temporal determinado

$$\text{Williams\%K} = \frac{\text{Max Alto} - \text{Close}}{\text{Max Alto} - \text{Min Bajo}} \quad (3.18)$$

Donde *Max Alto* es el precio más alto en el periodo de revisión, que suele ser de 14 días, *Cierre* es el precio de cierre más reciente y *Min Bajo* es el precio más bajo en el periodo de revisión, que suele ser de 14 días.

Las medias móviles también se utilizan en el Análisis Técnico, ya que también representan el *Momentum* o cambio de valor en un marco temporal N . Ayudan a comprender la tendencia del mercado y permiten identificar las señales de compra y venta a partir de los datos históricos de una acción en un marco temporal N previamente mencionado. En esta investigación, se ha aplicado la media móvil simple (SMA) y la media móvil exponencial (EMA) para marcos temporales de 5, 10, 20, 50, 100 y 200 días, por lo que habrá indicadores en diferentes períodos.

- (a) *Simple Moving Average* (SMA): Es una media móvil aritmética. Se calcula sumando los valores de cierre recientes de una acción para una ventana de tamaño N y dividiendo esa suma por el tamaño de la ventana. Así, cuando el tamaño de la ventana temporal N es bajo, responde rápidamente a los cambios en el valor de la acción; si el tamaño de la ventana N es alto, responde más lentamente.

$$\text{SMA} = \frac{A_1 + A_2 + \dots + A_n}{n} \quad (3.19)$$

Donde A_n es el precio de un activo en el periodo n y n es el número de períodos totales.

- (b) *Exponential Moving Average* (EMA): También se denomina *Exponentially Weighted Moving Average* ya que pondera las observaciones recientes, es decir, los precios de cierre de una acción más cercanos al actual. Puede decirse que las EMA responden mejor que las SMA a los cambios recientes en el precio de una acción.

$$\text{EMA}_{\text{Hoy}} = \left(\text{Valor}_{\text{Hoy}} \times \left(\frac{\text{Smoothing}}{1 + \text{Dias}} \right) \right) + \left(\text{EMA}_{\text{Ayer}} \times \left(\frac{\text{Smoothing}}{1 + \text{Dias}} \right) \right) \quad (3.20)$$

Donde *EMA* es la media móvil exponencial. El *Smoothing factor* se calcula como sigue:

$$\text{Smoothing} = \frac{2}{n + 1} \quad (3.21)$$

Donde n representa el número de períodos que utiliza la EMA.

Dado que tanto las predicciones algorítmicas como los resultados de los cálculos del factor técnico y de la media móvil dan como resultado el próximo valor de cierre de una acción, la recomendación se basa en la identificación de señales de compra y venta a partir de la comparación del valor previsto con el valor que tiene la acción en el momento actual.

4. El sistema de recomendación. A partir de los resultados obtenidos de los sistemas de previsión y análisis técnico, se procede a la especificación del diseño del Sistema de Recomendación, en el que se ponderan los resultados obtenidos para identificar las señales de compra/venta con el fin de poder realizar una recomendación. El diseño del paquete propone la creación de un sistema neutro que, a partir del análisis de las señales de compra/venta, determine la acción a realizar para/con una acción. Además del cálculo de las medias móviles y de los ratios de análisis técnico, se incluye un análisis mediante algoritmos de regresión. Los algoritmos de regresión se utilizan cuando hay que hacer una predicción sobre un conjunto de datos continuos [Montgomery et al., 2021]. Este es el caso de los datos de las series temporales históricas de una acción. La salida del algoritmo es una cantidad que puede medirse de forma flexible, en función de las entradas que se pasen al algoritmo. Los algoritmos de clasificación se limitarían a un conjunto de etiquetas. La regresión lineal puede definirse como un enfoque para modelizar la relación entre una variable escalar dependiente y , y una o más variables explicativas denominadas x . Matemáticamente, se expresa en la forma presentada en la ecuación 3.22.

$$y_i = \beta_0 + \beta_1 x_1 \quad (3.22)$$

donde la variable a predecir es y_i , la constante es β_0 , la pendiente β_1 y la variable de entrada x_i .

En el escenario actual, dado que se dispone del conjunto de datos históricos de una acción, la variable explicativa x da los valores de apertura del mercado, y la variable objetivo y da los valores de cierre del mercado. Así, la entrada del modelo

es x y la salida esperada y , donde y es la variable dependiente y x la variable independiente, de modo que el valor de apertura del mercado condiciona el valor de cierre del mercado.

3.4.2 Experimentación y resultados

La plataforma de visualización recoge las conclusiones de la investigación realizada en el resto de los módulos y proporciona dos opciones de visualización a nivel de usuario: una visión general como resultado del análisis exploratorio de los datos y el resultado del sistema de recomendación subyacente.

De este modo, se detallan las diferentes fases o herramientas utilizadas para el desarrollo de la arquitectura de la plataforma, a partir de los resultados del estudio del resto de módulos. Cabe destacar que el desarrollo de una plataforma de visualización sólo pretende acercar los resultados al usuario, sin ser la parte central del sistema propuesto.

El patrón de diseño utilizado, denominado MVC (Modelo-Vista-Controlador), se centra en la división del proyecto web según las funcionalidades de cada una de sus partes. El objetivo de la plataforma no es sólo ser usable e intuitiva, sino también permitir que cualquier usuario, sea o no experto en bolsa, pueda extraer sus propias conclusiones de los datos y evaluar la información analizada por el sistema.

La funcionalidad cubre la extracción y la visualización básica de los datos. El sistema recupera el perfil de la empresa y los datos históricos de los últimos 5 años de la acción. A partir de esos datos, elabora una serie de representaciones:

1. Series temporales: ofrece una representación gráfica de los datos históricos recuperados, donde los ejes X e Y representan el valor de la acción en euros, y la fecha en la que la acción alcanzó ese valor, respectivamente.
2. Gráfico de velas: esta representación muestra los valores de apertura y cierre para cada fecha y la diferencia entre los valores máximos y mínimos para la misma fecha.
3. Tabla de datos: representa los valores disponibles. Se denominan OHLC (*Open-High-Low-Close*).

La funcionalidad de "Visión General y Recomendación" es la misma que la comprobación de entrada del usuario, en el sentido de que también extrae el perfil de la empresa y los datos históricos. Sin embargo, esta funcionalidad también incluye factores técnicos y medias móviles con la consiguiente recomendación de compra/venta. Los gráficos generados comparan los diferentes algoritmos que el sistema ha aplicado para hacer la predicción. Esto permite al usuario identificar aquellos que han tenido una mayor precisión 3.6. Además, la plataforma presenta las conclusiones extraídas de los valores resultantes. Muestra la recomendación de compra/venta en base a esos valores. Por tanto el proceso de predicción y recomendación realizado por el sistema es transparente para el usuario.

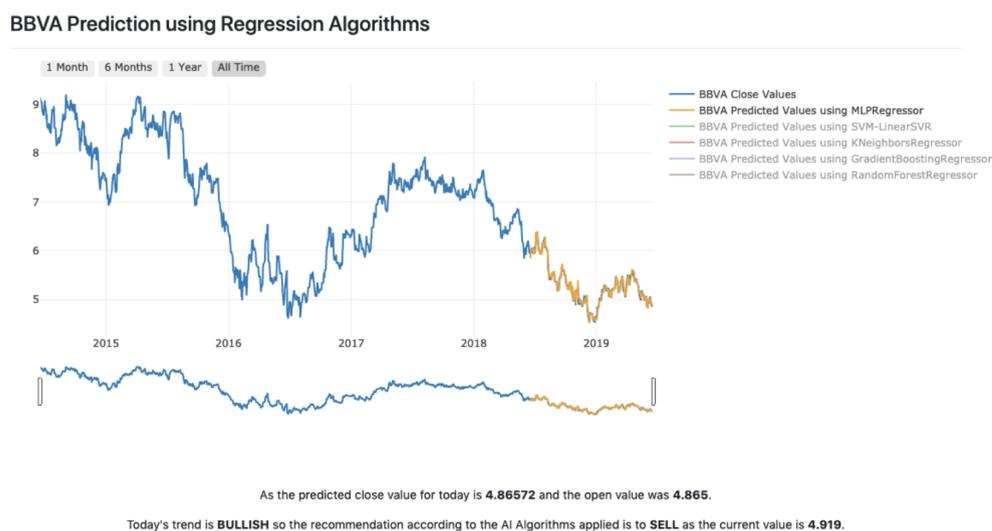


FIG. 3.6: Predicción mediante algoritmos de regresión

Una vez mostrada la justificación de los algoritmos de regresión utilizados por el sistema, el algoritmo que presenta una mejor precisión es el que muestra sus resultados por defecto (Figure 3.7). Aun así, la plataforma da la opción de mostrar diferentes ventanas temporales y visualizar los resultados de todos los algoritmos. Por último, la plataforma muestra un apartado en el que indica las conclusiones extraídas del estudio de los valores resultantes de la predicción, por lo que muestra la recomendación de compra/venta en base a estos valores.

Como conclusión se puede decir que la plataforma muestra las predicciones, que posteriormente combinará con los resultados del análisis técnico financiero, calculando los factores técnicos denominados Indicadores de *Momentum*, que indican la tendencia

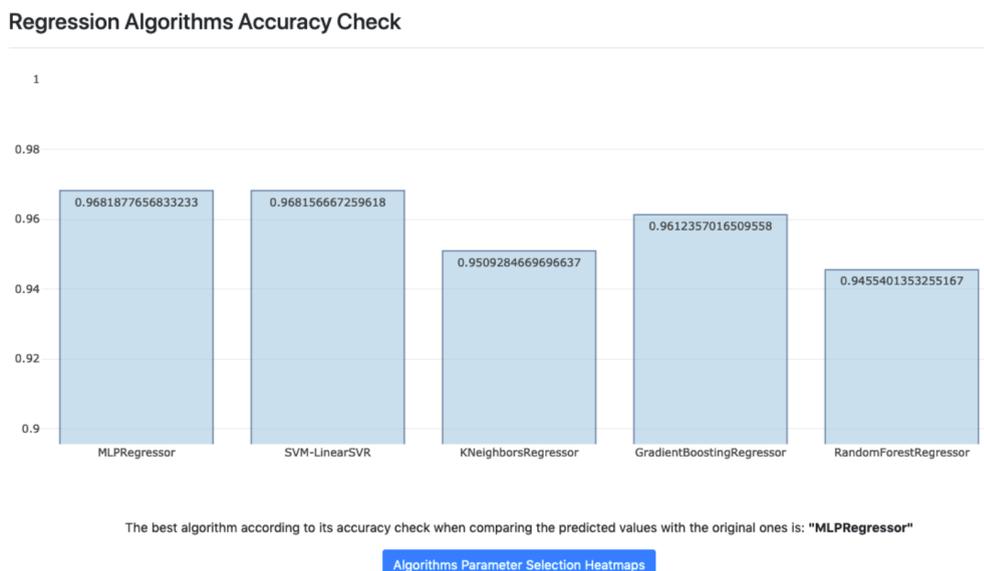


FIG. 3.7: Comprobación de la exactitud de los algoritmos de regresión

del mercado basándose en cálculos que toman diferentes ventanas de tiempo (Figure 3.8).

• Technical Factors

Tech. Factor	Result	Signal
RSI - 14 days	43.63047	SELL
STOCH - 9,6 days	38.29643	SELL
ULTOSC - 7,14,28 days	53.80424	BUY
WILLR - 14 days	-55.2381	SELL

[Back](#)

FIG. 3.8: Recomendaciones de compra/venta

Por tanto el sistema es capaz de realizar las recomendaciones y de soportar toda la carga computacional que conlleva realizarla. Además el usuario conoce el proceso de predicción y recomendación que realiza el sistema, conociendo lo ocurrido en cada una de las etapas del proceso. Esto supone un aumento de la confianza en la predicción al no haber sido alterada en beneficio de terceras personas.

3.5 Conclusiones

Las conclusiones obtenidas del análisis del estado del arte presentado en el Capítulo 2, así como la Sección 1.2 del Capítulo 1, permitieron el diseño de las herramientas pensadas para ser alojadas en la arquitectura FOBA. A continuación se describen las contribuciones científicas que han generado:

- El problema para gestionar los datos generados por el Internet de las Cosas, así como el coste de desplegar arquitecturas comerciales ha dado origen a la solución propuesta: una arquitectura global, escalable y modular basada en FC.
- Las capas de FOBA integran OV's de Agentes para compartir toda la inteligencia de negocio.
- FOBA ha sido diseñada para alojar sistemas de recomendación que mejoren la tasa de aceptación de productos bancarios.
- CEBRA ha sido creado para ser incorporado a una arquitectura Fog Computing que apoya el proceso de decisión de la banca comercial utilizando una combinación de modelos de decisión locales y globales y datos locales. La principal diferencia entre CEBRA y otros CBR es la incorporación de varias técnicas dentro de su ciclo de vida. Su principal aportación es que crea un perfil de usuario más completo mediante la extracción de información de las redes sociales. En la primera fase del ciclo, además de recoger los datos, se extraen los intereses de los clientes de sus redes sociales aplicando clasificadores al contenido textual. Posteriormente, esta información se entrega a revisores expertos para completar el caso. En la fase de reutilización, se propone incorporar kNN para mejorar la capacidad de recomendación.
- A la hora de hacer una recomendación, la API de CEBRA permite recuperar los usuarios que son altamente similares a aquel para el que se está haciendo la recomendación. Esta API permite dar una recomendación al usuario, actualizar los casos y comprobar la base de casos de forma global. En conclusión, la aplicación desarrollada es capaz de recomendar productos y se proporciona a la comunidad académica el material necesario para su uso y adaptación. También se

proporciona el conjunto de datos generado, para que la comunidad investigadora pueda utilizarlo.

- Adicionalmente al diseño e implementación de CEBRA, se ha diseñado y desarrollado una plataforma que permite llegar a conclusiones sobre el comportamiento futuro del mercado. La investigación realizada proporciona una primera aproximación al análisis de datos y al uso combinado de algoritmos y técnicas de ML junto con el análisis tradicional del mercado.
- Tras identificar las señales de compra y venta de acciones, se ha podido crear un sistema que recomienda al usuario comprar, mantener o vender una acción en un momento determinado del día, según la predicción obtenida por los algoritmos de regresión.

Capítulo 4

Evidencias y Resultados



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

Evidencias y Resultados

En este Capítulo se presentan el conjunto de publicaciones en revistas científicas internacionales, que han permitido el desarrollo de este trabajo y los resultados obtenidos. La Sección 4.1 detalla las publicaciones en: revistas científicas de alto impacto, revistas internacionales, conferencias y workshops internacionales. La sección 4.2 menciona los proyectos en los que se ha participado y que han sustentado la investigación desarrollada.

4.1 Publicaciones

4.1.1 Publicaciones en revistas científicas internacionales

1. Hernández-Nieves, E., Hernández, G., Gil-González, A. B., Rodríguez-González, S., & Corchado, J. M. (2020). Fog computing architecture for personalized recommendation of banking products. *Expert Systems with Applications*, 140, 112900. Factor de Impacto: 6,954 - Q1 (2020).
2. Hernández-Nieves, E., Hernández, G., Gil-González, A. B., Rodríguez-González, S., & Corchado, J. M. (2021). CEBRA: A CasE-Based Reasoning Application to recommend banking products. *Engineering Applications of Artificial Intelligence*, 104, 104327. Factor de Impacto: 6,212 - Q1 (2020).
3. Hernández-Nieves, E., Parra-Domínguez, J., Chamoso, P., Rodríguez-González, S., & Corchado, J. M. (2021). A Data Mining and Analysis Platform for Investment Recommendations. *Electronics*, 10(7), 859. Factor de Impacto: 2,397 - Q3 (2020).

4.1.2 Publicaciones en congresos internacionales y workshops

1. Hernández, E., Sittón, I., Rodríguez, S., Gil, A. B., & García, R. J. (2018, June). An investment recommender multi-agent system in financial technology. In *The 13th International Conference on Soft Computing Models in Industrial and Environmental Applications* (pp. 3-10). Springer, Cham.
2. Nieves, E. H. (2018, June). Stock Recommendation Platform Based on the Environment. INSIDER. In *International Symposium on Distributed Computing and Artificial Intelligence* (pp. 395-399). Springer, Cham.
3. Hernández, E., Sittón, I., Rodríguez, S., Gil, A. B., & García, R. J. (2018, June). An investment recommender multi-agent system in financial technology. In *The 13th International Conference on Soft Computing Models in Industrial and Environmental Applications* (pp. 3-10). Springer, Cham.
4. Hernández, E., González, A., Pérez, B., de Luis Reboledo, A., & Rodríguez, S. (2018, June). Virtual organization for fintech management. In *International Symposium on Distributed Computing and Artificial Intelligence* (pp. 201-210). Springer, Cham.
5. Hernández, E., Öztürk, M., Sittón, I., & Rodríguez, S. (2019, June). Data Protection on Fintech Platforms. In *International Conference on Practical Applications of Agents and Multi-Agent Systems* (pp. 223-233). Springer, Cham.
6. Hernández-Nieves, E., del Canto, Á. B., Chamoso-Santos, P., de la Prieta-Pintado, F., & Corchado-Rodríguez, J. M. (2020, June). A Machine Learning Platform for Stock Investment Recommendation Systems. In *International Symposium on Distributed Computing and Artificial Intelligence* (pp. 303-313). Springer, Cham.
7. Nieves, E. H. (2020, June). New Approach to Recommend Banking Products Through a Hybrid Recommender System. In *International Symposium on Ambient Intelligence* (pp. 262-266). Springer, Cham.
8. Nieves, E. H. (2020, June). Virtual Agent Societies to Provide Solutions to an Investment Problem. In *International Symposium on Distributed Computing and Artificial Intelligence* (pp. 241-246). Springer, Cham.

9. Nieves, E. H. (2020, June). Data Management Applied to Service Provision in Banking Environments. In International Congress on Blockchain and Applications (pp. 165-170). Springer, Cham.
10. Hernández-Nieves, E., García-Coria, J. A., Rodríguez-González, S., Gil-González, A. B. (2021, April) Distributed decision blockchain-secured support system to enhance stock market investment process (in press). In Sustainable Smart Cities and Territories International Conference. Springer, Cham.

4.2 Proyectos

- Nombre del proyecto: ROBIN: Robo-advisor intelligent.
Entidad/es financiadora/s: UD Ibérica.
Ciudad entidad financiadora: Madrid, Comunidad de Madrid, España.
Nombre del programa: Proyectos de I+D de Cooperación Nacional
Fecha de inicio-fin: 07/2018 - 01/2021.
- Nombre del proyecto: Emprendimiento inclusivo: Oportunidad y negocio en el ámbito transfronterizo.
Entidad/es financiadora/s: Fondo Europeo de Desarrollo Regional.
Nombre del programa: Interreg España-Portugal.
Fecha de inicio-fin: 10/2015 - 12/2019.
- Nombre del proyecto: BeEMP: Inteligencia social para la dinamización de la empleabilidad.
Entidad/es financiadora/s: Ministerio de Economía y Competitividad.
Ciudad entidad financiadora: Madrid, Comunidad de Madrid, España.
Nombre del programa: Retos colaboración 2016.
Fecha de inicio-fin: 03/2016 - 12/2017.

Capítulo 5

Conclusiones



**VNiVERSIDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

Conclusiones

En esta Tesis Doctoral se ha presentado una Arquitectura FC para entornos FinTech. Las publicaciones que conforman esta memoria por compendio de artículos presentan la arquitectura propuesta con el nombre de *Fog Oriented Banking Architecture* (FOBA, por sus siglas en inglés).

FOBA es una arquitectura por niveles diseñada para integrar sistemas predictivos en el proceso de recomendación de productos bancarios y de gestión de activos. Como complemento se han diseñado y desarrollado dos herramientas de recomendación, primeramente se ha implementado un CBR al que se ha denominado CEBRA (*CasE Based Reasoning Application*) y que está orientado hacia la recomendación de productos bancarios, y para finalizar, se ha implementado una plataforma de gestión de activos que tras aplicar algoritmos de regresión emite decisiones de compra/venta de acciones.

Este Capítulo presenta en la Sección 5.1, las principales conclusiones que se han obtenido durante la realización de la Tesis Doctoral. Finalmente, las líneas futuras de investigación que han surgido durante el proceso de desarrollo de esta memoria se presentan en la Sección 5.2.

5.1 Conclusiones

El FC surge como un nuevo paradigma computacional que extiende los servicios de la computación en la nube al borde de la red [Networking Index, 2016]. A través de la agrupación de recursos locales la computación, la comunicación, el control y el almacenamiento están más cerca de los usuarios finales en lugar de estar completamente en la nube [Datta et al., 2015]. Esta característica mejora la baja latencia, la movilidad, el ancho de banda de la red, la seguridad y la privacidad.

Podemos decir que el FC es una convergencia tecnológica que responde a los nuevos requerimientos que han surgido debido a la ubicuidad de los dispositivos y a las demandas de una gestión más ágil de las redes y los servicios, y de la privacidad de los datos [Vaquero and Rodero-Merino, 2014].

Para esta Tesis Doctoral se estableció como hipótesis que mediante el diseño de una arquitectura flexible, eficiente y personalizable basada en FC a partir de un modelo de agentes que facilite una gestión distribuida y dinámica, capaz de recopilar y analizar los datos recogidos por los dispositivos móviles, aplicaciones y equipos locales de la entidad, para recomendar productos bancarios, se lograrían mejorar aspectos de los servicios de atención al cliente, especialmente, lograr mayor transparencia y agilidad de los procesos, así como reducir los costes de gestión de la entidad.

El planteamiento de esta hipótesis permitió definir el objetivo principal y los objetivos específicos, mediante los cuales se ha demostrado la veracidad de la hipótesis de partida. En este contexto, el objetivo principal establecido consiste en la propuesta de una arquitectura modular y escalonada, capaz de gestionar soluciones complejas dirigidas a la recomendación de productos bancarios y gestión de activos.

La metodología Investigación - Acción fue el marco metodológico en el cual se enmarcó el desarrollo de esta tesis doctoral. Se definieron cinco fases que permitieron alcanzar los objetivos definidos. La revisión del estado del arte realizada da soporte a esta investigación, al permitir identificar que, aunque la disrupción de la IA en las finanzas ha incrementado el número de proyectos de investigación orientados a la aplicación de modelos, arquitecturas y técnicas, algunos de ellos como los Sistemas Multi-Agente (MAS), basados en un único sistema o agente trabajando de forma centralizada, no satisfacen los requerimientos actuales de los entornos FinTech [Milian et al., 2019]. Esto se debe a que los datos provienen de fuentes heterogéneas y deben ser recogidos de forma distribuida, por lo que se requiere que las OV's trabajen de forma colaborativa [Hugoson, 2007, Ponomarev and Voronkov, 2017] y que se basen en arquitecturas FC, siendo esta afirmación la línea de base de la tesis doctoral que se presenta. En consecuencia, se propone una arquitectura global capaz de abordar todas las necesidades y requerimientos importantes asociados a los entornos FinTech heterogéneos.

Para la mejora de aspectos de los servicios de atención al cliente, transparencia y agilidad de los procesos, y reducción de los costes de gestión de la entidad, planteados

en la hipótesis, se desarrollaron dos herramientas de recomendación, permitiendo su evaluación y validación. Las conclusiones obtenidas permiten afirmar que se ha cumplido con los objetivos definidos (OB1, OB2, OB3, OB4, OB5) y comprobar la veracidad de la hipótesis, las mismas se detallan en el siguiente orden:

- **Objetivos OB1, OB2 y OB3**

En referencia a los objetivos: OB1, OB2 y OB3, en la primera publicación de este compendio de artículos se logró [Hernández-Nieves et al., 2020]:

- Identificar los requerimientos existentes en los entornos bancarios, en el ámbito de la tecnología financiera, y que se pueden fortalecer a través de soluciones basadas en *Fog Computing* (FC).
- Revisar las arquitecturas de referencia existentes en las que el FC es el eje principal.
- Diseñar una arquitectura FC modular y escalonada, basada en agentes para dotar al sistema de inteligencia, orientada a entornos bancarios.

- **Objetivo OB4**

Para el cumplimiento del Objetivo 4 (OB4) se revisaron los trabajos realizados en el campo de los sistemas inteligentes de recomendación, consiguiéndose los siguientes hitos:

- Comprender que los sistemas de razonamiento basados en casos (CBR, por sus siglas en inglés) son sistemas capaces de adaptarse a diferentes entornos o contextos, posibilitando la generalización y permitiendo un cierto grado de independencia del entorno [de Mantaras, 1999]. En el segundo artículo de la tesis [Hernández-Nieves et al., 2021a] se adoptó como sistema de recomendación ya que las aplicaciones comerciales han demostrado un gran éxito con el uso de CBR, sobre todo por la ventaja de que se pueden derivar nuevas recomendaciones (es decir, soluciones) a partir de las antiguas recomendaciones con mayor facilidad [Skjold and Øynes, 2017].
- Comprender que más allá de las investigaciones que proponen las ANN como método de predicción bursátil, hay investigaciones que destacan la necesidad de contar con datos bursátiles históricos tras revisar diversas técnicas de

aprendizaje automático para la predicción bursátil [Yoo et al., 2005]. Durante la revisión se destacó que la predicción de los valores bursátiles es un reto debido a la falta de certeza [Patel et al., 2015]. Como conclusión a los trabajos revisados, ya que los autores atribuyen la falta de certeza a la imprevisibilidad de un entorno cambiante y proporcionan un enfoque mixto que utiliza tanto algoritmos de aprendizaje automático como análisis fundamentales y técnicos, en el tercer artículo de esta tesis doctoral [Hernández-Nieves et al., 2021b] se optó por este método.

• Objetivo OB5

Se estableció el objetivo de proporcionar sistemas de recomendación integrables en una arquitectura *Fog Computing* para recomendar productos bancarios y de gestión de activos de una manera eficiente, buscando una reducción de costes para la entidad bancaria. El diseño y desarrollo del CBR CEBRA y de la plataforma de recomendación de compra/venta de acciones, que se corresponden con el artículo segundo y tercero, respectivamente, de esta tesis doctoral [Hernández-Nieves et al., 2021a,b], cumplen con el objetivo. Se han conseguido los siguientes hitos:

- La primera herramienta desarrollada como una API/REST permite dar una recomendación al usuario, actualizar los casos y consultar la base de casos de forma global. A la hora de hacer una recomendación, la API permite recuperar los usuarios que son muy parecidos a aquel para el que se hace la recomendación. En conclusión, la aplicación desarrollada es capaz de recomendar productos y se pone a disposición de la comunidad académica con el material necesario para su uso y adaptación.
- La segunda herramienta desarrollada proporciona una primera aproximación al análisis de datos y al uso combinado de algoritmos y técnicas de ML, con el análisis tradicional del mercado. Su uso permite a la plataforma propuesta llegar a conclusiones sobre el comportamiento futuro del mercado. Así, se puede concluir que cuando los algoritmos de ML se entrenan con una cantidad de datos suficientemente grande, es posible predecir con éxito el valor de cierre sobre la base del valor de apertura actual del mercado. Así, tras identificar las señales de compra y venta, se ha podido crear un sistema que recomienda al

usuario comprar, mantener o vender una acción en un momento determinado del día, según la predicción obtenida por los algoritmos de regresión.

En base a las conclusiones obtenidas y descritas en este apartado, se demuestra que existe aún un amplio margen de mejoras para los desafíos asociados a los entornos FinTech.

5.2 Trabajo futuro

Durante el desarrollo de la tesis y una vez alcanzados los objetivos inicialmente propuestos, han surgido nuevas líneas de investigación que generan interés como trabajo futuro. En este apartado se presentan algunas de ellas:

- Despliegue de la arquitectura FOBA: la investigación futura se centrará en la realización de una prueba piloto en la que se desplegaría la arquitectura FC incorporando la OV de Agentes en la capa *Fog*. Teniendo en cuenta que no se desplegará en una oficina bancaria real, el objetivo es estudiar el funcionamiento de la red, probar las velocidades de procesamiento, la carga computacional soportada, la escalabilidad, etc. Se considerará el despliegue de sensores de reconocimiento facial que alerten de la entrada de clientes al agente encargado de esa tarea y el posterior acceso a la base de datos para iniciar la recomendación. De esta forma, el piloto cubrirá el ciclo completo de la investigación propuesta en esta tesis doctoral.
- Organización Virtual de Agentes: diseño de un mecanismo para los agentes que les permita ajustar sus pesos de similitud en el sistema CBR CEBRA según su funcionamiento. Por lo tanto, en futuras investigaciones crearemos un entorno virtual donde los agentes puedan realizar esta tarea.
- Plataforma de recomendación de compra/venta de acciones: adicionalmente, se propone realizar un estudio de los algoritmos aplicados a otros mercados, ya que el sistema propuesto está orientado a un mercado muy específico; el mercado continuo español. Será necesario realizar un estudio para determinar los mejores algoritmos para los mercados bursátiles de cada uno de los países a incorporar. Además, se considera la posibilidad de seguir investigando sobre las identificaciones de eventos que pueden utilizarse para elegir mejor la operación realizada (compra/venta) y las características sociales de las diferentes comunidades.

Bibliografía

- Aamodt, A. and Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI communications*, 7(1):39–59.
- Ai, Y., Peng, M., and Zhang, K. (2018). Edge computing technologies for internet of things: a primer. *Digital Communications and Networks*, 4(2):77–86.
- Al-khafajiy, M., Baker, T., Chalmers, C., Asim, M., Kolivand, H., Fahim, M., and Waraich, A. (2019). Remote health monitoring of elderly through wearable sensors. *Multimedia Tools and Applications*, pages 1–26.
- Al Ridhawi, I., Mostafa, N., Kotb, Y., Aloqaily, M., and Abualhaol, I. (2017). Data caching and selection in 5g networks using f2f communication. In *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1–6. IEEE.
- Al-Weshah, G. A. (2017). Marketing intelligence and customer relationships: empirical evidence from jordanian banks. *Journal of Marketing Analytics*, 5(3):141–152.
- Aloqaily, M., Al Ridhawi, I., Salameh, H. B., and Jararweh, Y. (2019). Data and service management in densely crowded environments: Challenges, opportunities, and recent developments. *IEEE Communications Magazine*, 57(4):81–87.
- Alsoubi, T., Qin, Y., Hill, R., and Al-Aqrabi, H. (2020). Enabling distributed intelligence for the internet of things with iota and mobile agents. *Computing*, pages 1–19.
- Arora, N. et al. (2019). Financial analysis: stock market prediction using deep learning algorithms. In *Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM)*, Amity University Rajasthan, Jaipur-India.

- Artikis, A. (2009). Dynamic protocols for open agent systems. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 97–104.
- Atsalakis, G. S. and Valavanis, K. P. (2009). Surveying stock market forecasting techniques—part ii: Soft computing methods. *Expert systems with applications*, 36(3):5932–5941.
- Benbasat, I., Goldstein, D. K., and Mead, M. (1987). The case research strategy in studies of information systems. *MIS quarterly*, pages 369–386.
- Bergenti, F., Gleizes, M.-P., and Zambonelli, F. (2006). *Methodologies and software engineering for agent systems: the agent-oriented software engineering handbook*, volume 11. Springer Science & Business Media.
- Bonomi, F., Milito, R., Zhu, J., and Addepalli, S. (2012). Fog computing and its role in the internet of things. In *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*, pages 13–16. ACM.
- Bosse, S. (2016). Mobile multi-agent systems for the internet-of-things and clouds using the javascript agent machine platform and machine learning as a service. In *2016 IEEE 4th international conference on future internet of things and cloud (FiCloud)*, pages 244–253. IEEE.
- Breidbach, C. F., Keating, B. W., and Lim, C. (2019). Fintech: research directions to explore the digital transformation of financial service systems. *Journal of Service Theory and Practice*.
- Breidbach, C. F. and Maglio, P. P. (2016). Technology-enabled value co-creation: An empirical analysis of actors, resources, and practices. *Industrial Marketing Management*, 56:73–85.
- Brust, L., Breidbach, C. F., Antons, D., and Salge, T.-O. (2017). Service-dominant logic and information systems research: A review and analysis using topic modeling. In *ICIS*.
- Brzoza-Woch, R., Konieczny, M., Nawrocki, P., Szydło, T., and Zielinski, K. (2016). Embedded systems in the application of fog computing—levee monitoring use case.

- In *2016 11th IEEE Symposium on Industrial Embedded Systems (SIES)*, pages 1–6. IEEE.
- Bunnell, L., Osei-Bryson, K.-M., and Yoon, V. Y. (2020). Finpathlight: Framework for an multiagent recommender system designed to increase consumer financial capability. *Decision Support Systems*, page 113306.
- Cao, L. (2020). Ai in fintech: A research agenda. *arXiv preprint arXiv:2007.12681*.
- Carrascosa, C., Giret, A., Julian, V., Rebollo, M., Argente, E., and Botti, V. (2009). Service oriented mas: an open architecture. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 1291–1292.
- Chamoso, P., Rivas, A., Rodríguez, S., and Bajo, J. (2018). Relationship recommender system in a business and employment-oriented social network. *Information sciences*, 433:204–220.
- Chen, D. and Burrell, P. (2001). Case-based reasoning system and artificial neural networks: A review. *Neural Computing & Applications*, 10(3):264–276.
- Chen, S. and Weiss, G. (2014). An intelligent agent for bilateral negotiation with unknown opponents in continuous-time domains. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, 9(3):1–24.
- Coghlan, D. (2019). *Doing action research in your own organization*. Sage.
- Corchado, J. M. and Laza, R. (2003). Constructing deliberative agents with case-based reasoning technology. *International Journal of Intelligent Systems*, 18(12):1227–1241.
- Corchado, J. M., Pavón, J., Corchado, E. S., and Castillo, L. F. (2004). Development of cbr-bdi agents: a tourist guide application. In *European Conference on Case-based Reasoning*, pages 547–559. Springer.
- Coughlan, P. and Coghlan, D. (2002). Action research for operations management. *International journal of operations & production management*.
- Dapp, T. (2014). Fintech—the digital (r) evolution in the financial sector. deutsche bank research.

- Dash, R. and Dash, P. K. (2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. *The Journal of Finance and Data Science*, 2(1):42–57.
- Datta, S. K., Bonnet, C., and Haerri, J. (2015). Fog computing architecture to enable consumer centric internet of things services. In *2015 International Symposium on Consumer Electronics (ISCE)*, pages 1–2. IEEE.
- de Mantaras, R. L. (1999). Case-based reasoning. In *Advanced Course on Artificial Intelligence*, pages 127–145. Springer.
- De Mantaras, R. L., McSherry, D., Bridge, D., Leake, D., Smyth, B., Craw, S., Faltings, B., Maher, M. L., T COX, M., Forbus, K., et al. (2005). Retrieval, reuse, revision and retention in case-based reasoning. *The Knowledge Engineering Review*, 20(3):215–240.
- do Nascimento, N. M. and de Lucena, C. J. P. (2017). Fiot: An agent-based framework for self-adaptive and self-organizing applications based on the internet of things. *Information Sciences*, 378:161–176.
- Echeverry, G. A. I., Mejía, M. H., Ossa, L. F. C., Morales, A., and Méndez, N. D. D. (2012). Network management using multi-agents system. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 1(3):49–54.
- Edwards, R. D., Magee, J., and Bassetti, W. C. (2018). *Technical analysis of stock trends*. CRC press.
- Ferber, J., Gutknecht, O., and Michel, F. (2003). From agents to organizations: an organizational view of multi-agent systems. In *International workshop on agent-oriented software engineering*, pages 214–230. Springer.
- Fosso Wamba, S., Kala Kamdjoug, J. R., Epie Bawack, R., and Keogh, J. G. (2020). Bitcoin, blockchain and fintech: a systematic review and case studies in the supply chain. *Production Planning & Control*, 31(2-3):115–142.
- Fratu, O., Pena, C., Craciunescu, R., and Halunga, S. (2015). Fog computing system for monitoring mild dementia and copd patients-romanian case study. In *2015 12th International Conference on Telecommunication in Modern Satellite, Cable and Broadcasting Services (TELSIKS)*, pages 123–128. IEEE.

- Gai, K., Qiu, M., and Sun, X. (2018). A survey on fintech. *Journal of Network and Computer Applications*, 103:262–273.
- Garcia-Fornes, A., Hübner, J. F., Omicini, A., Rodriguez-Aguilar, J. A., and Botti, V. (2011). Infrastructures and tools for multiagent systems for the new generation of distributed systems. *Engineering Applications of Artificial Intelligence*, 24(7):1095–1097.
- Gomber, P., Koch, J.-A., and Siering, M. (2017). Digital finance and fintech: current research and future research directions. *Journal of Business Economics*, 87(5):537–580.
- González-Briones, A., Prieto, J., De La Prieta, F., Demazeau, Y., and Corchado, J. M. (2020). Virtual agent organizations for user behaviour pattern extraction in energy optimization processes: A new perspective. *Neurocomputing*.
- Hajibaba, M. and Gorgin, S. (2014). A review on modern distributed computing paradigms: Cloud computing, jungle computing and fog computing. *Journal of computing and information technology*, 22(2):69–84.
- Henrique, B. M., Sobreiro, V. A., and Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications*, 124:226–251.
- Hernandez-Nieves, E. (2020). CEBRA.
- Hernández-Nieves, E., Hernández, G., Gil-González, A.-B., Rodríguez-González, S., and Corchado, J. M. (2020). Fog computing architecture for personalized recommendation of banking products. *Expert Systems with Applications*, 140:112900.
- Hernández-Nieves, E., Hernández, G., Gil-González, A. B., Rodríguez-González, S., and Corchado, J. M. (2021a). Cebra: A case-based reasoning application to recommend banking products. *Engineering Applications of Artificial Intelligence*, 104:104327.
- Hernández-Nieves, E., Parra-Domínguez, J., Chamoso, P., Rodríguez-González, S., and Corchado, J. M. (2021b). A data mining and analysis platform for investment recommendations. *Electronics*, 10(7):859.
- Herreras, E. B. (2004). La docencia a través de la investigación-acción. *Revista iberoamericana de educación*, 35(1):1–9.

- Hommel, K. and Bican, P. M. (2020). Digital entrepreneurship in finance: Fintechs and funding decision criteria. *Sustainability*, 12(19):8035.
- Hong, K., Lillethun, D., Ramachandran, U., Ottenwalder, B., and Koldehofe, B. (2013). Mobile fog: A programming model for large-scale applications on the internet of things. In *Proceedings of the second ACM SIGCOMM workshop on Mobile cloud computing*, pages 15–20. ACM.
- Hugoson, M.-Å. (2007). Centralized versus decentralized information systems. In *IFIP Conference on History of Nordic Computing*, pages 106–115. Springer.
- Jarrar, Y. F. and Neely, A. (2002). Cross-selling in the financial sector: customer profitability is key. *Journal of Targeting, Measurement and Analysis for Marketing*, 10(3):282–296.
- Jennings, N. R. and Wooldridge, M. (1998). Applications of intelligent agents. In *Agent technology*, pages 3–28. Springer.
- Jubair, M. A., Mostafa, S. A., Mustapha, A., and Hafit, H. (2018). A survey of multi-agent systems and case-based reasoning integration. In *2018 International Symposium on Agent, Multi-Agent Systems and Robotics (ISAMSR)*, pages 1–6. IEEE.
- Khaidem, L., Saha, S., and Dey, S. R. (2016). Predicting the direction of stock market prices using random forest. *arXiv preprint arXiv:1605.00003*.
- Kim, K.-j. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, 55(1-2):307–319.
- Kolodner, J. L. and Leake, D. (1996). A tutorial introduction to case-based reasoning.
- Kwan, J., Gangat, Y., Payet, D., and Courdier, R. (2016). An agentified use of the internet of things. In *2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pages 311–316. IEEE.
- Kyriazakos, S., Mihaylov, M., Anggorojati, B., Mihovska, A., Craciunescu, R., Fratu, O., and Prasad, R. (2016). eWALL: an intelligent caring home environment offering personalized context-aware applications based on advanced sensing. *Wireless Personal Communications*, 87(3):1093–1111.

- Llewellyn, D. T. (2018). Financial technology, regulation, and the transformation of banking. In *The European Money and Finance Forum (SUERF) conference: Financial Disintermediation and the Future of the Banking Sector 30th October, 2018 Madrid*.
- Luis Reboredo, A. d. et al. (2014). *Modelo de asignación de roles y distribución de tareas en organizaciones virtuales*.
- Lusch, R. F. and Nambisan, S. (2015). Service innovation: A service-dominant logic perspective. *MIS quarterly*, 39(1):155–176.
- Madsen, H., Burtschy, B., Albeanu, G., and Popentiu-Vladicescu, F. (2013). Reliability in the utility computing era: Towards reliable fog computing. In *2013 20th International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 43–46. IEEE.
- Mehrban, S., Nadeem, M. W., Hussain, M., Ahmed, M. M., Hakeem, O., Saqib, S., Kiah, M. M., Abbas, F., Hassan, M., and Khan, M. A. (2020). Towards secure fintech: A survey, taxonomy, and open research challenges. *IEEE Access*, 8:23391–23406.
- Melnychenko, S., Volosovych, S., and Baraniuk, Y. (2020). Dominant ideas of financial technologies in digital banking. *Baltic Journal of Economic Studies*, 6(1):92–99.
- Mendhurwar, S. and Mishra, R. (2018). Emerging synergies between internet of things and social technologies.
- Milian, E. Z., Spinola, M. d. M., and de Carvalho, M. M. (2019). Fintechs: A literature review and research agenda. *Electronic Commerce Research and Applications*, 34:100833.
- Mogaji, E., Soetan, T. O., and Kieu, T. A. (2020). The implications of artificial intelligence on the digital marketing of financial services to vulnerable customers. *Australasian Marketing Journal (AMJ)*.
- Mohan, D. (2016). How banks and fintech startups are partnering for faster innovation. *Journal of Digital Banking*, 1(1):13–21.
- Monteiro, A., Dubey, H., Mahler, L., Yang, Q., and Mankodiya, K. (2016). Fit: A fog computing device for speech tele-treatments. In *2016 IEEE International Conference on Smart Computing (SMARTCOMP)*, pages 1–3. IEEE.

- Montgomery, D. C., Peck, E. A., and Vining, G. G. (2021). *Introduction to linear regression analysis*. John Wiley & Sons.
- Naumov, M., Mudigere, D., Shi, H.-J. M., Huang, J., Sundaraman, N., Park, J., Wang, X., Gupta, U., Wu, C.-J., Azzolini, A. G., et al. (2019). Deep learning recommendation model for personalization and recommendation systems. *arXiv preprint arXiv:1906.00091*.
- Networking Index, C. V. (2016). Forecast and methodology, 2016-2021, white paper. *San Jose, CA, USA*, 1:0–4.
- Nikzad-Khasmakhi, N., Balafar, M., and Feizi-Derakhshi, M. R. (2019). The state-of-the-art in expert recommendation systems. *Engineering Applications of Artificial Intelligence*, 82:126–147.
- Nussbaumer, P., Matter, I., and Schwabe, G. (2012). “enforced” vs. “casual” transparency—findings from it-supported financial advisory encounters. *ACM Transactions on Management Information Systems (TMIS)*, 3(2):1–19.
- Oueida, S., Kotb, Y., Aloqaily, M., Jararweh, Y., and Baker, T. (2018). An edge computing based smart healthcare framework for resource management. *Sensors*, 18(12):4307.
- Oyenan, W. H., DeLoach, S. A., and Singh, G. (2009). Exploiting reusable organizations to reduce complexity in multiagent system design. In *International Workshop on Agent-Oriented Software Engineering*, pages 3–17. Springer.
- Patel, J., Shah, S., Thakkar, P., and Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert systems with applications*, 42(1):259–268.
- Pico-Valencia, P. and Holgado-Terriza, J. A. (2018). Agentification of the internet of things: A systematic literature review. *International Journal of Distributed Sensor Networks*, 14(10):1550147718805945.
- Pimprikar, R., Ramachadran, S., and Senthilkumar, K. (2017). Use of machine learning algorithms and twitter sentiment analysis for stock market prediction. *Int J Pure Appl Math*, 115(6):521–526.

- Ponomarev, S. and Voronkov, A. (2017). Multi-agent systems and decentralized artificial superintelligence. *arXiv preprint arXiv:1702.08529*.
- Qi, Y. and Xiao, J. (2018). Fintech: Ai powers financial services to improve people's lives. *Communications of the ACM*, 61(11):65–69.
- Richter, M. M. and Weber, R. O. (2016). *Case-based reasoning*. Springer.
- Riesbeck, C. K. and Schank, R. C. (2013). *Inside case-based reasoning*. Psychology Press.
- Rodriguez, S., Julián, V., Bajo, J., Carrascosa, C., Botti, V., and Corchado, J. M. (2011). Agent-based virtual organization architecture. *Engineering Applications of Artificial Intelligence*, 24(5):895–910.
- Roth-Berghofer, T., Recio Garcia, J. A., Sauer, C., Bach, K., Althoff, K.-D., Diaz-Agudo, B., and Gonzales Calero, P. A. (2012). Building case-based reasoning applications with mycbr and colibri studio.
- Roumani, Y., Nwankpa, J. K., and Roumani, Y. F. (2016). Examining the relationship between firm's financial records and security vulnerabilities. *International Journal of Information Management*, 36(6):987–994.
- Sagraves, A. and Connors, G. (2017). Capturing the value of data in banking. *Applied Marketing Analytics*, 2(4):304–311.
- Said, Y. B., Kanzari, D., and Bezzine, M. (2018). A behavioral and rational investor modeling to explain subprime crisis: Multi agent systems simulation in artificial financial markets. In *Financial Decision Aid Using Multiple Criteria*, pages 131–147. Springer.
- Savaglio, C., Ganzha, M., Paprzycki, M., Bădică, C., Ivanović, M., and Fortino, G. (2020). Agent-based internet of things: State-of-the-art and research challenges. *Future Generation Computer Systems*, 102:1038–1053.
- Shi, X., Zhang, P., and Khan, S. U. (2017). Quantitative data analysis in finance. In *Handbook of Big Data Technologies*, pages 719–753. Springer.
- Shokri Gazafroudi, A., Prieto, J., and Corchado, J. M. (2019). Virtual organization structure for agent-based local electricity trading. *Energies*, 12(8):1521.

- Skjold, K. and Øynes, M. S. (2017). Case-based reasoning and computational creativity in a recipe recommender system. Master's thesis, NTNU.
- Soni, S. (2011). Applications of anns in stock market prediction: a survey. *International Journal of Computer Science & Engineering Technology*, 2(3):71–83.
- Sridevi, M., Rao, R. R., and Rao, M. V. (2016). A survey on recommender system. *International Journal of Computer Science and Information Security*, 14(5):265.
- Stantchev, V., Barnawi, A., Ghulam, S., Schubert, J., and Tamm, G. (2015). Smart items, fog and cloud computing as enablers of servitization in healthcare. *Sensors & Transducers*, 185(2):121.
- Stulz, R. M. (2019). Fintech, bigtech, and the future of banks. *Journal of Applied Corporate Finance*, 31(4):86–97.
- Tanco, J. A. A. and Camarero, L. A. (2013). Investigación en acción: Cómo impulsar la contribución de la universidad en la competitividad de las organizaciones. *Harvard Deusto Business Research*, 2(2):89–101.
- Tang, B., Chen, Z., Hefferman, G., Wei, T., He, H., and Yang, Q. (2015). A hierarchical distributed fog computing architecture for big data analysis in smart cities. In *Proceedings of the ASE BigData & SocialInformatics 2015*, page 28. ACM.
- Vaquero, L. M. and Rodero-Merino, L. (2014). Finding your way in the fog: Towards a comprehensive definition of fog computing. *ACM SIGCOMM Computer Communication Review*, 44(5):27–32.
- Varanasi, B. and Belida, S. (2015). Documenting rest services. In *Spring REST*, pages 91–104. Springer.
- Wu, X., Zhu, X., Wu, G.-Q., and Ding, W. (2013). Data mining with big data. *IEEE transactions on knowledge and data engineering*, 26(1):97–107.
- Xie, M. (2019). Development of artificial intelligence and effects on financial system. In *Journal of Physics: Conference Series*, volume 1187, page 032084. IOP Publishing.
- Yan, Y. and Su, W. (2016). A fog computing solution for advanced metering infrastructure. In *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, pages 1–4. IEEE.

- Yin, H. and Gai, K. (2015). An empirical study on preprocessing high-dimensional class-imbalanced data for classification. In *2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems*, pages 1314–1319. IEEE.
- Yoo, P. D., Kim, M. H., and Jan, T. (2005). Machine learning techniques and use of event information for stock market prediction: A survey and evaluation. In *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06)*, volume 2, pages 835–841. IEEE.
- Zetsche, D. A., Buckley, R. P., Arner, D. W., and Barberis, J. N. (2017). From fintech to techfin: The regulatory challenges of data-driven finance. *NYUJL & Bus.*, 14:393.
- Zhu, J., Chan, D. S., Prabhu, M. S., Natarajan, P., Hu, H., and Bonomi, F. (2013). Improving web sites performance using edge servers in fog computing architecture. In *2013 IEEE Seventh International Symposium on Service-Oriented System Engineering*, pages 320–323. IEEE.