



## Inteligencia artificial aplicada al método Backward Seismic Analysis

### Artificial Intelligence Applied to the Backward Seismic Analysis Method

### Inteligência artificial aplicada ao método de Análise Sísmica Retroativa

Patricia-Andrea Möller-Acuña<sup>1</sup>

Patricio-Andrés Pineda-Nalli<sup>2</sup>

**Recibido:** febrero de 2022

**Aceptado:** junio de 2022

**Para citar este artículo:** Möller-Acuña, P. A., y Pineda-Nalli, P. A. (2022). Inteligencia artificial aplicada al método Backward Seismic Analysis. *Revista Científica*, 45(3), 369-377. <https://doi.org/10.14483/23448350.18556>

#### Resumen

Este trabajo presenta aplicaciones del método Backward Seismic Analysis (BSA) para estanques de acero de acuerdo con una base de datos que recopila más de 382 estanques de acero en operación durante terremotos de subducción: el ocurrido en Valdivia en 1960, en Chile Central en 1985, en Tocopilla en 2007, el último de gran magnitud registrado en el Maule en 2010, en Alaska en 1964 y otros de Estados Unidos entre 1933 y 1995 (subductivos y corticales). Se ha registrado que gran parte de los estanques sin sistemas de anclajes han fallado durante grandes terremotos. Estos han sido diseñados principalmente con los códigos API 650-E, AWWA-D100 y NZSEE, los cuales proponen procedimientos equivalentes para estimar las demandas sísmicas, pero con métodos de diseño distintos. Durante diferentes conferencias se evaluaron las causas que originaron las fallas, concluyendo que los estanques estaban diseñados principalmente con el estándar API 650-E y no disponían de sistemas de anclajes. Además,

los códigos de diseño más utilizados no consideran en la actualidad los aspectos relevantes que condicionan la respuesta sísmica de los estanques de acero. Este trabajo desarrolla un modelo de predicción basado en la información histórica ya descrita, capaz de predecir de manera eficiente si un estanque presentará fallas durante algún terremoto. Se evaluaron diversos algoritmos, encontrando que el método Random Forest exhibe los mejores resultados. Los resultados obtenidos en la predicción de fallas de estanques alcanzan más del 90 % de eficiencia en la mayoría de los escenarios evaluados.

**Palabras clave:** anclado; aprendizaje automático; Backward Seismic Analysis; estanque; Random Forest; subducción.

#### Abstract

This work presents applications of the Backward Seismic Analysis (BSA) method for steel storage tanks using a data base of more than 382 steel storage tanks in operation during large subductive

1. Universidad Autónoma de Chile (Talca, Chile). Contacto: [patricia.moller@uautonoma.cl](mailto:patricia.moller@uautonoma.cl)  
2. PPN Ingeniería y consultoría (Talca, Chile).\

earthquakes: Valdivia 1960, Central Chile 1985, Tocopilla 2007, El Maule 2010, Alaska 1964, and others in the United States between 1933 and 1995 (subductive and cortical). It has been recorded that most of the steel storage tanks without anchor systems have failed during large earthquakes. These have been designed with the standards API 650-E, AWWA-D100, and NZSEE, which propose similar procedures for estimating seismic forces, but with different design methods. During different conferences, the causes of the failures were evaluated, concluding that the tanks were designed mainly with the API 650-E code and were unanchored. Moreover, the design codes employed do not consider relevant aspects that condition the seismic response of steel storage tanks. This work develops a prediction model based on the historical information already described, which is capable of efficiently predicting if a steel storage tank will suffer any failures during an earthquake. Various algorithms were evaluated, finding that the Random Forest method exhibits the best results. The results obtained in the prediction of steel storage tank failures reach more than 90% efficiency in most of the evaluated scenarios.

**Keywords:** anchored; Backward Seismic Analysis; machine learning; Random Forest; subduction; tank.

### Resumo

Este trabalho apresenta aplicações do método Backward Seismic Analysis (BSA) para tanques de aço de acordo com um banco de dados que compila mais de 382 tanques de aço em operação durante terremotos de subducção: o que ocorreu em Valdivia em 1960, no Chile Central em 1985, em Tocopilla em 2007, o último de grande magnitude registrado no Maule em 2010, no Alasca em 1964 e outros nos Estados Unidos entre 1933 e 1995 (subdutivo e crustal). Muitas das lagoas sem sistemas de ancoragem foram registradas como falhando durante grandes terremotos. Estes foram projetados principalmente com os códigos API 650-E, AWWA-D100 e NZSEE, que propõem procedimentos equivalentes para estimar demandas sísmicas, mas com métodos de projeto diferentes. Durante diferentes conferências, foram avaliadas as causas que originaram as falhas, concluindo que os tanques foram projetados principalmente com a norma API 650-E e não possuíam sistemas de

ancoragem. Além disso, os códigos de projeto mais utilizados atualmente não consideram os aspectos relevantes que condicionam a resposta sísmica dos tanques de aço. Este trabalho desenvolve um modelo de previsão baseado nas informações históricas já descritas, capaz de prever com eficiência se uma lagoa apresentará falhas durante um terremoto. Vários algoritmos foram avaliados, verificando-se que o método Random Forest apresenta os melhores resultados. Os resultados obtidos na previsão de falhas de lagoas atingem mais de 90% de eficiência na maioria dos cenários avaliados.

**Palavras-chaves:** ancorado; Análise Sísmica Regressiva; aprendizado de máquina; Floresta Aleatória; lago; subducção.

### Introducción

Dentro de las ramas de la inteligencia artificial se encuentra el aprendizaje automático, que es un proceso a través del cual un computador es capaz de aprender de una experiencia para aplicar lo aprendido sobre una tarea de aprendizaje, la cual es finalmente medida por la performance, que indica que tan bien aprendió el computador, específicamente el modelo con el que se está trabajando, según [Mitchell \(1990\)](#). Es así como de forma más específica se generan algoritmos que le permiten al computador generalizar comportamientos a partir de la experiencia proporcionada con el objetivo de repetir esta generalización en nuevos ejemplos.

Actualmente el diseño sísmico de estanques de acero se basa en las disposiciones de los códigos de diseño más utilizados en el mundo, tales como: API 650 en su apéndice E (API, 2013), AWWA-D100 (AWWA, 2011) y NZSEE (NZSEE, 2009), a pesar de esto, gran parte de los estanques de acero no anclados presentaron daños durante grandes terremotos y habían sido diseñados con el estándar API 650-E. De acuerdo con esto, se requiere actualizar y modificar los métodos de diseño para el cálculo de los esfuerzos admisibles del manto, dado que proponen métodos equivalentes para el cálculo de la demanda sísmica, pero con diferentes métodos de diseño. Actualmente en

Chile se utiliza la norma NCh2369 y el anteproyecto de norma en preparación por el ICHA (Instituto Chileno del Acero), este último se basa en el método Backward Seismic Analysis (BSA) (Pineda y Saragoni, 2003; [Pineda, Saragoni y Arze, 2011](#); Pineda y [Saragoni, 2017](#); [Pineda, 2019](#); [Pineda y Saragoni, 2019](#); [Pineda, 2020](#); ICHA, 2021; [Pineda, 2021](#)). Con el método BSA se pueden calcular los coeficientes sísmicos para el diseño de los estanques de acero, considerando la respuesta sísmica para una cantidad superior a 380 estanques de acero en operación durante los terremotos de Valdivia en 1960, Chile Central en 1985, Tocopilla en 2007, El Maule en 2010, además de Alaska en 1964 y de Estados Unidos entre los años 1933 y 1995.

En este artículo se presenta un sistema con un modelo de predicción de daño de estanque que utiliza información histórica de carácter internacional. Este modelo es capaz de predecir de manera eficiente y con hasta 93 % de eficiencia esta predicción, además se entrega un análisis de importancia de cada uno de los atributos considerados para realizar los análisis de predicción.

## Metodología

Por medio de inteligencia artificial se validará el método BSA que propone un procedimiento para determinar la respuesta sísmica en estanques de acero, considerando las propiedades geométricas de los estanques en operación durante los terremotos: geometría y espesores de planchas, alturas de llenado, características de los suelos, normas de diseño aplicadas, registros sísmicos instrumentales, efectos de la directividad sísmica, daños observados, con presencia de pandeo del manto y colapsos. El método BSA requiere disponer de información de la respuesta sísmica de estanques de acero, esto para clasificarlos según sus dimensiones y condiciones de esbeltez, para rangos seguros y daños menores, considerables y reparables hasta valores con alto riesgo de colapso. El método requiere de información de los estanques:

dimensiones y espesores del manto, características de las fundaciones y tipos de anclajes, suelo de fundación, tipo de líquido almacenado, niveles de llenado al momento del terremoto, efectos de la sismicidad y aceleraciones máximas del suelo de las zonas de operación del estanque, daños observados luego de terremotos, características del techo: tipo (cónico, flotante, domo), estructuración, peso, criterios de diseño y códigos utilizados, planos de fabricación y As Built. El método BSA se resume en la aplicación de una fórmula para el cálculo del coeficiente sísmico que representa el comportamiento estructural de los estanques durante grandes terremotos, por lo que su uso propone una predicción de la respuesta sísmica en terremotos futuros usando la ecuación (1) de Coeficiente sísmico para estanques de acero - Método BSA.

$$C_{ps} = \frac{\pi \cdot t_s \cdot (D_e)^2 \cdot F_c \cdot V_s}{M_t \cdot M_s} \quad (1)$$

La clasificación supervisada es una de las tareas más frecuentes llevadas a cabo por IA. Puede ser a través de modelos desarrollados en función de la estadística (regresión logística, análisis discriminante) o bien por IA (redes neuronales, inducción de reglas, árboles de decisión, redes bayesianas), capaz de realizar las tareas propias de la clasificación ([Möller, Ahumada y Reyes, 2017](#); [Japkowicz, 2000](#)), obteniendo de esta forma un modelo de clasificación que indicará si los estanques presentarán o no daño, con esta información se puede construir la matriz de confusión que permite visualizar el desempeño del modelo.

## Caracterización de los datos

Se cuenta con 382 casos de estanques, de los cuales 72 presentaron daños y 308 tuvieron buen comportamiento sísmico durante los terremotos considerados para los estudios.

1.- Pre-procesamiento de datos (*data preprocessing*). La limpieza de datos es la tarea fundamental de la inteligencia artificial ([Deshmukh y Wangikar, 2015](#)). Los datos de los estanques se analizaron para eliminar, corregir y estandarizar aquellos que presentaban anomalías en función del resto mediante algoritmos en lenguaje Python (CRAN, s.f.a)

2.- Diseño de atributos (*feature desing*). Se entien- de por atributo al conjunto de mediciones de los estanques durante un proceso sísmico, tales como:

Mag = magnitud del sismo

G = peso específico del líquido en el tanque

Hm(m) = altura del manto tanque

HLL(m) = altura del líquido en el tanque durante el sismo

D/H= relación de esbeltez del tanque

Cps = coeficiente sísmico obtenido del método BSA

ts = espesor de las planchas del manto del primer anillo inferior

De = diámetro del estanque

Fc = tensión admisible de compresión del manto del estanque

Vs = esfuerzo de corte sísmico

Mt = masa total del estanque

Ms = momento volcane sísmico

Fy = tensión de fluencia del acero

E = módulo de elasticidad de acero

Anclaje = el tanque se encuentra o no anclado

Daño = el tanque presento daño o no durante el sismo

3.- Selección de atributos (*feature selection*). La selección de atributos permite reconocer las características principales de un conjunto de datos que están relacionadas con la variable de respuesta ([Hernández, 2007](#)). ¿Cuáles son las características que se relacionan con la presencia o no de daño?

- **Chi-square test.** Una prueba de chi-square se usa en estadísticas para probar la independencia de dos eventos. Con las variables se puede obtener el recuento observado O y el recuento esperado E

([McHugh, 2013](#)). Chi-square mide como el recuento E y el recuento O se desvían entre sí (ecuación (2)).

$$X^2 = \sum \frac{(O_i - E_i)^2}{E_i} \quad (2)$$

Esta prueba se realizó a través de un algoritmo implementado en Python, utilizando la librería Scikit-Learn ([Pedregosa et al., 2011](#)).

- **F-test.** Es una prueba estadística que se utiliza para comparar entre modelos y verificar si la diferencia es significativa entre ellos. F-test hace un modelo de prueba de hipótesis X y Y.

Para evaluar se comparan errores de mínimos cuadrados en ambos modelos y se verifica si la diferencia en los errores entre el modelo X y Y es significativa o se introdujo de forma azarosa. Para realizar este test se utilizó Anova a través de un algoritmo programado en Python utilizando la librería Scikit-Learn ([Pedregosa et al., 2011](#)).

4.- Balance de clases. Para realizar un balance de clases se utilizó la librería Scikit-Learn de Python ([Pedregosa et al., 2011](#)), específicamente la función *resample* (en una proporción de 70 - 30) que integra muestras sintéticas de la clase minoritaria utilizando diversos algoritmos que siguen la tendencia del grupo minoritario.

5.- Machine Learning. Dado el set de referencia anteriormente descrito, se evaluaron diversos métodos de clasificación como Naive Bayes ([Dagnino, 2014](#)), Random-Forest ([Rish, 2001](#); CRAN, s.f.b), árboles de decisión ([Möller, Ahumada y Reyes, 2017](#)), SVM ([Breiman, 2001](#)), regresión logística (CRAN, s.f.c), Gradient Boosting ([Wang et al., 2019](#)), todos los cuales fueron implementados en el lenguaje de programación de Python, específicamente utilizando la librería Scikit-Learn([Pedregosa et al., 2011](#)).

6.- Entrenamiento y testeo (*training and testing*). Se realizaron cinco segmentaciones del set de datos y se entrenaron, parametrizaron y testearon los

modelos de forma simultánea con el fin de encontrar el modelo óptimo para realizar la tarea designada a través de un algoritmo implementado en Python con el uso de la librería Scikit-Learn.

A partir de los resultados obtenidos, se extrajeron los siguientes valores:

- VP son los verdaderos positivos: número de estanques que no falló y el modelo indico que no falló.
- FP son los falsos positivos: número de estanques que falló y el modelo indico que no falló.
- FN son los falsos negativos: número de estanques que no falló y el modelo indico que falló.
- VN son los verdaderos negativos: número de estanques que falló y el modelo indico que falló.
- EG eficiencia global: corresponde a tasa de éxito general de la predicción.

De estos se derivan las medidas de performance definidas para este problema:

Matrices de confusión: permite visualizar el desempeño de un algoritmo, la cual se puede ver en la [Tabla 1](#).

Realidad	Predicción	
	TN	FP
	FN	TP

**Tabla 1.** Matriz de confusión binaria

Curva ROC: muestra el rendimiento de un modelo de clasificación en todos los umbrales de clasificación. Esta curva representa dos parámetros: tasa de verdaderos positivos (TP) y tasa de falsos positivos (FP)

Precisión: mide la calidad del modelo.

$$\text{Precisión} = TP / (TP + FP)$$

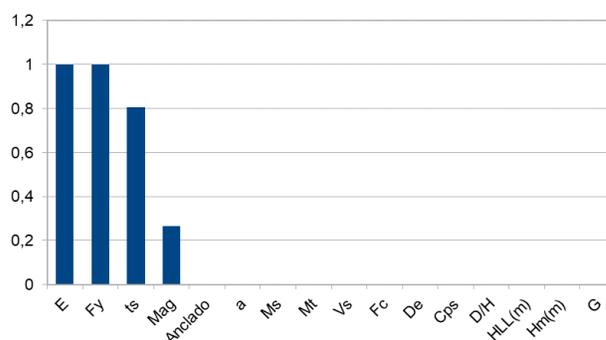
*Recall*: la métrica de exhaustividad informa sobre la cantidad de ejemplos que el modelo es capaz de identificar.

$$\text{Recall} = TP / (TP + FN)$$

## Resultados

### Chi-square test

Se utilizó esta prueba para evaluar la dependencia del daño de los estanques con los atributos mencionados anteriormente. Se considera que existe una dependencia cuando el valor de la prueba es cercano a 0, por lo tanto, todos los atributos que muestran un valor superior son variables independientes que no presentan relación con el daño del tanque ([Figura 1](#)).



**Figura 1.** Test Chi-square

### F-test (Anova)

Se utilizó esta prueba para medir la importancia de los atributos dependientes y estimar cuáles son los más relevantes para predecir el posible daño en un tanque. Todos los atributos presentan un grado de relación con el resultado analizado (con o sin daño). Para realizar el modelo de clasificación, se seleccionaron los seis atributos que muestran los mayores puntajes en F-test ([Figura 2](#)).

\* Anclado = 25,6

\* G = 14,08

\* Fc = 12,37

\* a = 10,23

\* Mt = 7,81

\* Vs = 7,13



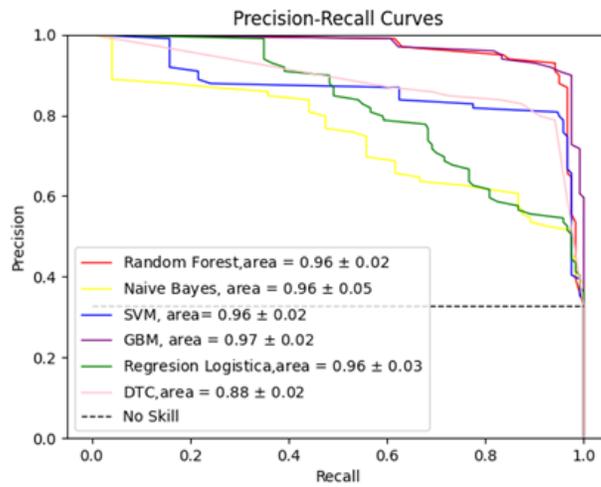


Figura 4. Curvas de precisión

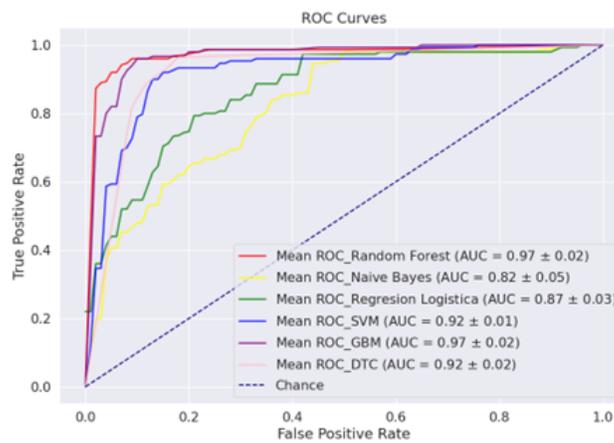


Figura 5. Curva ROC de atributos seleccionados por F-test

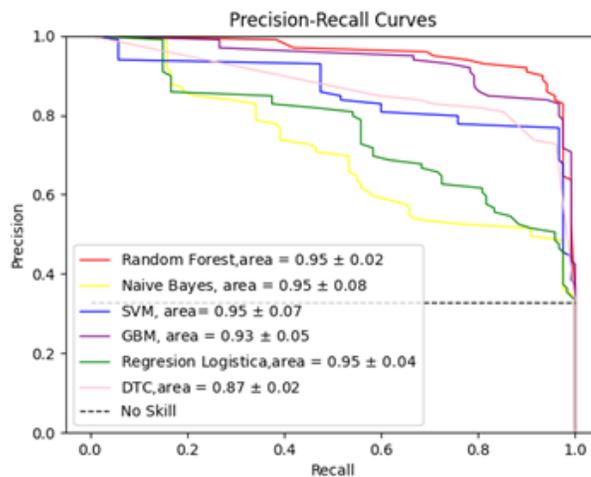


Figura 6. Gráfico de relación de precisión vs. recall

automático. Puede concluirse que el análisis de datos y el uso de herramientas de IA permiten analizar cuáles son las características y factores más importantes para que un estanque presente o no daño. El anclaje del estanque cumple un papel clave, ya que se presenta como la característica más relevante para que un estanque no presente daño; sin embargo, solo controlar el anclaje del estanque no asegura que se evite un daño total del mismo, por lo que se deben considerar otras variables para evitar este daño, por ejemplo, la magnitud del sismo.

Diversos métodos de aprendizaje automático fueron evaluados para esta tarea, encontrando que RF y GBM son los mejores modelos para realizar el proceso de clasificación de este problema (presenta o no daño) en los distintos escenarios y condiciones evaluadas para registros. Esto indica que estos algoritmos se adaptan mejor a las condiciones planteadas en la problemática propuesta en este trabajo, asociada a “predecir de forma temprana la ocurrencia de daño de un estanque utilizando información histórica”. El modelo desarrollado alcanza una eficiencia del 93 % de eficiencia. Utilizando el modelo de clasificación GBM se podría estimar si un tanque va o no a presentar daño durante un sismo, en función de su anclaje: G, a, Fs, Mt, Vs. Este trabajo por tanto entrega resultados concluyentes en cuanto a que es si posible determinar los principales atributos que cumplen un papel clave en el modelo recién mencionado.

Finalmente, se determinan los resultados con el bagaje teórico que antecede y sustenta al trabajo desarrollado. Supone una muestra de las reflexiones del autor en relación con el nuevo conocimiento generado en diálogo con el conocimiento ya existente.

## Agradecimientos

Agradecimiento al Grupo de Investigación Ingeniería Informática y Computación de la Universidad Autónoma de Chile, Facultad de Ingeniería.

## Referencias

- American Petroleum Institute (API). (2013). *API Standard 650: Welded Tanks for Oil Storage*
- American Water Works Association (AWWA) (2011). *AWWA D100-11: Welded Carbon Steel Tanks for Water Storage*
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32. <https://doi.org/10.1023/A:1010933404324>
- CRAN. (s.f.a). *The Comprehensive R Archive Network*. <https://cran.r-project.org>
- CRAN. (s.f.b). *randomForest*. <https://cran.r-project.org/web/packages/randomForest/randomForest.pdf>
- CRAN. (s.f.c). *RWeka*. <https://cran.r-project.org/web/packages/RWeka/RWeka.pdf>
- Dagnino, S. (2014). Análisis de varianza. *Revista Chilena de Anestesia*, 43(4), 306-310
- Deshmukh, R., Wangikar, V. (2015). Data cleaning: Current approaches and issues. En *IEEE International Conference on Knowledge Engineering*
- Hernández, J. (2007). *Introducción a la minería de datos*. Alhambra
- Instituto Chileno del Acero (ICHA). (2021). Anteproyecto Norma Técnica Chilena. Diseño sísmico de estanques atmosféricos de acero apoyados en el suelo. Chile
- Japkowicz, N. (2000). Learning from imbalanced data sets: A comparison of various strategies. *AAAI Workshop* (pp. 10-15)
- McHugh, M. L. (2013). The Chi-square test of independence Lessons in biostatistics. *Biochemia Medica*, 23(2), 143-149. <https://doi.org/10.11613/bm.2013.018>
- Mitchell, T. (1990). *An Artificial Intelligence Approach*
- Möller-Acuña, P., Ahumada-García, R., Reyes-Suárez, J. A. (2017). Machine learning for prediction of frost episodes in the Maule region of Chile. En *Ubiquitous Computing and Ambient Intelligence* (pp. 715-720). [https://doi.org/10.1007/978-3-319-67585-5\\_69](https://doi.org/10.1007/978-3-319-67585-5_69)
- New Zealand Society for Earthquake Engineering (NZSEE). (2009). *Seismic Design of Storage Tanks*.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Pasos, A., Cournapeau, D., Brucher, M., Perrot, M.,

- Duchesnay, E., (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830
- Pineda, P. (2019). *Análisis sísmico Backward de estanques atmosféricos de acero*. (Tesis de Magíster). Universidad de Chile
- Pineda, P. (2020). Tank design recommendations for seismic codes on critical industrial facilities. En *17th World Conference on Earthquake Engineering*
- Pineda, P. (2021). Proposal of seismic coefficient and estimate for horizontal sliding for steel tanks by Backward Seismic Analysis (BSA) method. En *10th International Conference on the Behaviour of Steel Structures in Seismic Areas*. [https://doi.org/10.1007/978-3-031-03811-2\\_120](https://doi.org/10.1007/978-3-031-03811-2_120)
- Pineda, P., Saragoni, G. R. (2003). Backward seismic analysis of steel tanks. En *8th International Conference on Behaviour of Steel Structures in Seismic Areas*
- Pineda, P., Saragoni, G. R. (2017). Analysis of steel tanks in Chile subduction earthquakes. En *16th World Conference on Earthquake Engineering*
- Pineda, P., Saragoni, G. R. (2019). Diseño sísmico de estanques de acero basado en el método análisis sísmico Backward. En *XII Congreso Chileno de Sismología en Ingeniería Sísmica*
- Pineda P., Saragoni, G. R., Arze, E. (2011). Performance of steel tanks in Chile 2010 and 1985 earthquakes. En *7th International Conference on Behaviour of Steel Structures in Seismic Areas* (pp. 337-342)
- Rish, I. (2001). An empirical study of the naive Bayes classifier. En *JCAI Workshop on Empirical Methods in AI*
- Wang, Q. Q., Yu S. C., Qi X., Hu Y. H., Zheng W. J., Shi J. X., Yao H. Y. (2019). Overview of logistic regression model analysis and application. *Chinese Journal of Preventive Medicine*, 53(9), 955-960.

