



UNIVERSIDAD ESAN

FACULTAD DE INGENIERÍA

INGENIERÍA EN GESTIÓN AMBIENTAL

INGENIERÍA DE TECNOLOGÍAS DE INFORMACIÓN Y SISTEMAS

INGENIERÍA INDUSTRIAL Y COMERCIAL

**Técnicas de machine learning para la mejora del método de proyección de ventas de los
análisis de multirresiduos de plaguicidas en alimentos de Mérieux Nutrisciences**

Trabajo de Suficiencia Profesional presentado en satisfacción parcial de los requerimientos
para obtener el título profesional de Ingeniero en ingeniería en gestión ambiental, ingeniería
de tecnologías de información y sistemas e ingeniería industrial y comercial

AUTORES

Huisacayna Cutipa, Abigail Nelly

Jacinto Martell, Samuel Humberto

La Rosa Gadea, Marghore Susana

Machuca Abanto, Axl Boris

Torres Yupanqui, Rocio del Pilar Lesly

ASESOR

Fabian Arteaga, Junior John

ORCID N° 0000-0001-9804-7795

Noviembre, 2023

RESUMEN

En la actualidad es importante que toda empresa realice pronósticos de todo tipo, que ayuden a mitigar el impacto negativo y/o aprovechar los impactos positivos que los cambios generan. El propósito del presente trabajo es identificar la técnica del aprendizaje automático que mejore el método de proyección de ventas generadas por el análisis de multirresiduos de plaguicidas en alimentos de la empresa Mérieux Nutrisciences. Se emplea la metodología Cross-Industry Standard Process for Data Mining (CRISP-DM) para determinar el modelo predictivo óptimo para la empresa. Después de adquirir y adecuar la data, se aplica y analiza en las técnicas de regresión lineal, light gradient boosting machine (LightGBM), seasonal auto regressive integrated moving average (SARIMA) y long short-term memory (LSTM). Con la ejecución de los modelos establecidos, se concluyó que el uso de modelos predictivos permite a las empresas, tomar decisiones más acertadas y mejorar su gestión, además, se visualizó que el modelo LightGBM tiene una mayor precisión que los otros modelos con un 0.0152 de mean squared error (MSE). Se recomienda realizar el modelado con un mayor número de data para generar un pronóstico más preciso, contrastar con el laboratorio y realizar estudios adicionales para ajustar hiperparámetros propios del modelo.

Palabras claves: *aprendizaje automático, inteligencia artificial, crisp-dm, lightgbm.*

ABSTRACT

Nowadays, it is important for every company to make forecasts of all kinds, which help to mitigate the negative impact and/or take advantage of the positive impacts that changes generate. The purpose of this work is to identify the machine learning technique that improves the sales projection method generated by the analysis of multi-residues of pesticides in food of Mérieux Nutrisciences company. The Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology is used to determine the optimal predictive model for the company. After acquiring and adapting the data, it is applied and analyzed in linear regression techniques, light gradient boosting machine (LightGBM), seasonal auto regressive integrated moving average (SARIMA) and long short-term memory (LSTM). With the execution of the established models, it was concluded that the use of predictive models allows companies to make more accurate decisions and improve their management, in addition, it was visualized that the LightGBM model has a higher accuracy than the other models with a 0.0152 mean squared error (MSE). It is recommended to perform the modeling with a larger number of data to generate a more accurate forecast, to contrast with the laboratory and to perform additional studies to adjust the model's own hyperparameters.

Keywords: *machine learning, artificial intelligence, crisp-dm, lightgbm.*

RESULTADOS DE INFORME DE SIMILITUD

Trabajo de suficiencia profesional

ORIGINALITY REPORT

13%

SIMILARITY INDEX

12%

INTERNET SOURCES

3%

PUBLICATIONS

5%

STUDENT PAPERS

PRIMARY SOURCES

1

hdl.handle.net

Internet Source

2%

2

Submitted to Universidad ESAN -- Escuela de
Administración de Negocios para Graduados

Student Paper

2%

3

scielo.sld.cu

Internet Source

1%

4

repositorio.espe.edu.ec

Internet Source

1%

5

repositorio.esan.edu.pe

Internet Source

1%