Modelos Predictivos para el Éxito de Nuevos Productos en la Industria Embotelladora

Daniela Jiménez Téllez, Lautaro Gabriel Coteja, Andrés Villarreal González, Hibran H. Tapia Fernández

Instructores: Dr. Edgar González Fernández, Dr. Mauricio González Soto

TC3007 - Agosto-Diciembre 2024

Inteligencia Artificial Avanzada para la Ciencia de Datos II, Escuela de Ingeniería y Ciencias, Tecnológico de Monterrey, Monterrey, México

Introducción

Arca Continental es una de las embotelladoras más grandes de América Latina y un actor clave en el mercado global. En su búsqueda por maximizar el éxito de los productos nuevos, este proyecto analizó datos de clientes y productos para desarrollar modelos predictivos que identificaran patrones de consumo y segmentaciones clave. El análisis utilizó técnicas avanzadas de Machine Learning e Inteligencia Artificial para predecir el éxito de productos de lanzamiento, ofreciendo a la empresa una herramienta estratégica para minimizar riesgos y optimizar recursos. **Este enfoque** no solo **mejora la precisión en la toma de decisiones**, sino que **también fortalece la relación con los clientes al entender mejor sus necesidades y preferencias.**

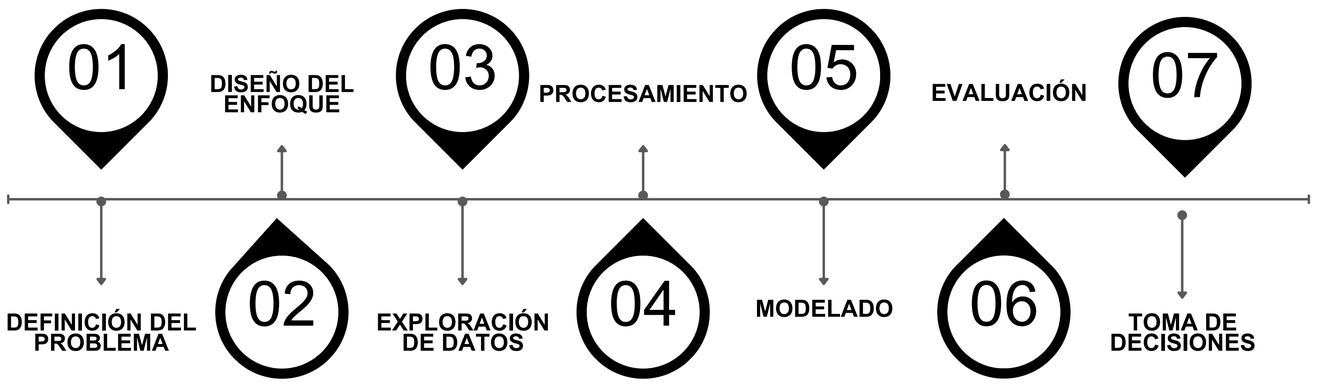
Objetivo

Aplicar modelos de Machine Learning para identificar clientes con alta probabilidad de éxito en la compra de productos de lanzamiento y evaluar el impacto de factores demográficos e infraestructurales en el desempeño de los productos.

Metodología

El análisis del éxito de productos en la red de tiendas de Arca Continental se realizó integrando datos de ventas, productos y clientes. Tras una limpieza y preprocesamiento de datos, se codificaron variables categóricas, se estandarizaron las numéricas y se generaron nuevas columnas derivadas para enriquecer la información y facilitar el análisis. Se entrenaron tres modelos: Regresión Logística, Random Forest y Red Neuronal, aplicando técnicas de optimización de hiperparámetros en cada uno para maximizar su rendimiento y asegurar predicciones más precisas.

Diagrama de Flujo



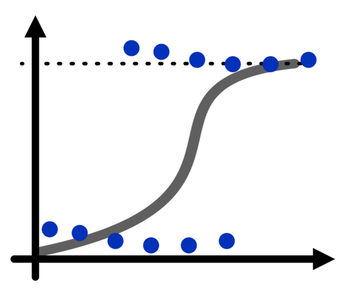
Resultados

Los modelos desarrollados fueron efectivos para predecir el éxito de productos de lanzamiento, ofreciendo diferentes perspectivas. La regresión logística permitió identificar patrones generales y tendencias iniciales, mientras que las redes neuronales captaron relaciones más complejas entre los datos. Sin embargo, **Random Forest destacó por su precisión y capacidad para manejar datos complejos**, permitiendo identificar clientes con alta confiabilidad de comprar productos de lanzamiento. Estos resultados refuerzan el valor de los modelos de Machine Learning como herramientas estratégicas para optimizar decisiones, minimizar riesgos y mejorar la implementación de nuevos productos en la industria embotelladora.

Conclusión

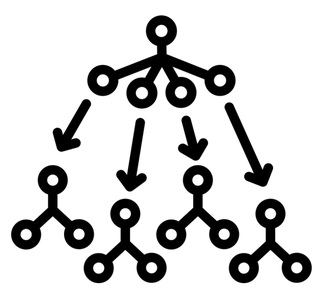
Este proyecto resalta el impacto positivo de la Inteligencia Artificial en el análisis de productos y el comportamiento de los clientes. A través de modelos los modelos aquí expuestos, **se alcanzó una alta precisión en la predicción del éxito de productos de lanzamiento y en la identificación de patrones clave en las preferencias de los consumidores.** Este enfoque optimiza la toma de decisiones estratégicas, reduce riesgos, y también fortalece la relación con los clientes al comprender mejor sus necesidades. Los resultados obtenidos consolidan el papel de estas herramientas para potenciar la competitividad en la industria embotelladora.

Regresión Logística



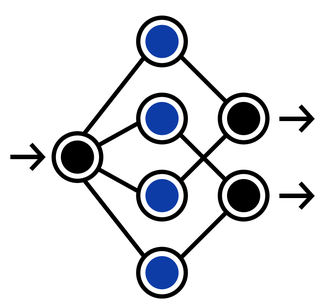
La regresión logística predice probabilidades para eventos binarios, como determinar si un producto será exitoso o no, modelando la relación entre variables independientes y una dependiente categórica.

Bosques Aleatorios



Bosque Aleatorio, o Random Forest, combina múltiples árboles de decisión para clasificar si un producto será exitoso o no, reduciendo errores y mejorando la precisión del modelo.

Redes Neuronales



Las redes neuronales son modelos inspirados en el cerebro humano, capaces de aprender patrones complejos mediante capas de neuronas interconectadas, clasificando si un producto será exitoso o no en el mercado.

Referencias

Arca Continental. (s.f.). Nuestra compañía. https://www.arcacontal.com/nuestra-compa%C3%B1%C3%ADa.aspx

Fernandes, A. A. T., Figueiredo Filho, D. B., Rocha, E. C. da, & Nascimento, W. da S. (2020). Read this paper if you want to learn logistic regression. Revista de Sociologia e Política, 28(74), e006. https://www.scielo.br/j/rsocp/a/RWjPthhKDYbFQYydbDr3MgH/?lang=en

Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., Czogalik, Ł., Dudek, P., Magiera, M., Lis, A., Paszkiewicz, I., Nawrat, Z., Cebula, M., & Gruszczyńska, K. (2023). What is machine learning, artificial neural networks and deep learning?—Examples of practical applications in medicine. Diagnostics, 13(15), 2582. https://www.mdpi.com/2075-4418/13/15/2582

Schonlau, M., & Zou, R. Y. (2020). The random forest algorithm for statistical learning. The Stata Journal, 20(1), 3–29. https://journals.sagepub.com/doi/10.1177/1536867X20909688