

Sistema inteligente para la segmentación de clientes con el fin de potencializar y optimizar la conversión de la campaña depósito financiero.

Sebastián Herrera¹, Edinson Neira¹ y Sergio Rairán.¹

¹ Pontificia Universidad Javeriana, Bogotá, Colombia

Resumen. las campañas de marketing directas son una son una técnica frecuentemente por entidades financieras. En este artículo veremos el problema en el que una institución bancaria portuguesa, se propone por medio de una campaña de marketing basada en llamadas telefónicas, predecir si un usuario se suscribirá a un depósito a plazos. De cada usuario se tiene atributos como el trabajo actual, el nivel de educación, el incumplimiento financiero, etc. Actualmente, la institución realiza las llamadas sin clasificar ni priorizar clientes, como consecuencia se tiene un tiempo elevado en la labor, costos elevados e incomodidad en clientes. Se evaluará por medio de cuatro métodos diferentes a los clientes y se harán comparaciones entre los métodos usando tres tipos de métricas diferentes para concluir cual es la mejor manera de implementar la campaña teniendo una mejor colocación de depósitos y así una reducción en los costos de ejecución de la campaña.

Palabras Clave: Marketing Bancario Directo, Redes Neuronales Artificiales, Sistemas Inteligentes.

1 Motivación y Contextualización del Problema

Una institución bancaria portuguesa, propone a través de una campaña de marketing directo basada en llamadas telefónicas, predecir si un usuario se suscribirá a un depósito financiero a plazos. El banco ha recopilado información sobre las características y atributos de los usuarios objetivo. Ejemplo de las características recopiladas son el trabajo actual, el nivel de educación, el incumplimiento financiero, etc. Inicialmente, la institución realiza las llamadas sin clasificar ni priorizar clientes, como consecuencia se tiene un tiempo elevado en la labor e incomodidad en clientes que no cumplen con el perfil. Entendiendo que los resultados de las campañas de marketing directas, en mercados competitivos, depende de una clasificación detallada de los posibles clientes buscando predecir el interés sobre el producto ofrecido, en este proyecto abordaremos el anterior problema.

1.1 Targeting direct marketing campaigns by neural networks. Journal of Marketing Management

Este estudio muestra cómo el uso de Redes Neuronales Artificiales (ANN) puede mejorar la efectividad de las campañas de marketing de correo directo gracias a una mejor predicción de la tasa de respuesta de los sujetos incluidos en la población objetivo de acuerdo con los factores que se cree que tienen un impacto en su intención de compra. Los resultados muestran la efectividad de las Redes Neuronales Artificiales, en comparación con el análisis de regresión múltiple y el análisis de regresión logística, para identificar relaciones complejas entre los datos y, en particular, para perfilar clientes y prospectos y anticipar su comportamiento.

1.2 Bank direct marketing based on neural network and C5.0 Models

Este documento presenta aplicaciones de modelos recientes e importantes de minería de datos; Red neuronal de perceptrón multicapa (MLPNN) y el modelo de árbol de decisiones de Ross Quinlan (C5.0). El objetivo es examinar el rendimiento de los modelos MLPNN y C5.0 en datos reales de suscripción de depósitos bancarios. El objetivo es aumentar la efectividad de la campaña mediante la identificación de las principales características que inciden en el éxito (el depósito suscrito por el cliente) en base a MLPNN y C5.0. Los resultados experimentales demuestran, con mayor precisión, el éxito de estos modelos en la predicción de la mejor campaña de contacto con los clientes para la suscripción de depósitos. Los rendimientos se miden mediante tres medidas estadísticas; precisión, sensibilidad y especificidad de la clasificación

1.3 Evaluating marketing campaigns of banking using neural networks. In Proceedings of the World Congress on Engineering

Este artículo propone redes neuronales para evaluar el marketing bancario. En este artículo se utiliza una red neuronal de propagación retroalimentada con funciones de transferencia tan-sigmoideas para predecir si el cliente suscribe el depósito. El conjunto de datos se obtiene del repositorio de aprendizaje automático de UCI. Los resultados de aplicar la metodología de redes neuronales propuesta para predecir la suscripción en función de los parámetros de las llamadas telefónicas seleccionadas muestran la capacidad de la red para aprender los patrones correspondientes a las suscripciones del cliente depósito. El porcentaje clasificado correctamente en la muestra de simulación por la red neuronal propuesta es del 90 por ciento.

1.4 Variable selection in clustering for marketing segmentation using genetic algorithms

En este documento, se utilizan algoritmos genéticos (GA) para la selección de variables y para determinar el número de grupos. Se utiliza un caso real de un conjunto de datos bancarios para ilustrar la aplicación de la segmentación de marketing. Los resul-

tados muestran que la selección de variables a través de GA puede encontrar efectivamente la solución óptima global, y la precisión del modelo clasificado aumenta drásticamente después de la agrupación.

1.5 Fuzzy based decision making for promotional marketing campaigns

Este artículo presenta una predicción de la selección de clientes para campañas de marketing promocional incorporando un enfoque difuso. Se propone un modelo borroso para la selección de clientes a los que se debe apuntar para esquemas de suscripción de depósitos. Los criterios de selección se formulan sobre la base del saldo del préstamo del cliente, la edad del cliente y los ingresos anuales. El enfoque de lógica difusa se utiliza para la toma de decisiones para apuntar a clientes específicos para la suscripción.

2 Descripción de la Tarea.

Usando el set de datos públicos de una institución financiera de Portugal que se encuentra alojado en Center of Machine Learning and Intelligent System (UCI)¹, se desarrollaran 4 diferentes modelos de IA que buscan mejorar la colocación de una campaña de depósito financiero.

2.1 Visión general

Contrastar 4 modelos de IA aplicados a la base de datos del banco financiero de Portugal, donde por medio de las métricas de especificidad, sensibilidad, curva ROC y precisión, donde cada una buscara superar un valor ingenuo definido a continuación:

- Especificidad
- Sensibilidad
- ROC
- Precisión

Las técnicas serán trabajadas en COLAB, como lenguaje principal Python 3.7, sin embargo, serán usados otras herramientas como R, Matlab, entre otras para robustecer el tratamiento de la información, mapeado de los datos y el cálculo de métricas.

El proyecto tendrá una duración de tres meses donde al final de este periodo se tendrá un análisis comparativo entre los mejores modelos encontrados para cada una de las técnicas y se concluirá sobre cuál de ellas es la mejor opción para cumplir el objetivo planteado.

1 Repositorio de datos: <https://archive.ics.uci.edu/ml/datasets/Bank+Marketing>

2.2 Restricciones y alcances

Se trabajará únicamente sobre 4 técnicas de IA, estas serán: Redes neuronales, Lógica difusa, algoritmos genéticos y aprendizaje de máquina. El objetivo de estas técnicas será predecir que clientes que pueden adquirir un depósito financiero CDT.

3 Análisis de Potencialidades para el Uso de Herramientas de Inteligencia Artificial

La técnica de enfoque más frecuente en la literatura son las redes neuronales artificiales. El artículo original en el que se describe el problema propone una red neuronal con 79% de precisión [6]. Guido et al. [1] abordan el problema utilizando técnicas de aprendizaje de máquina como regresión múltiple y regresión logística, y también se desarrolla usando redes neuronales, la que se concluye como una mejor técnica para este problema. Elsalamony et al. [2] utilizan las metodologías redes neuronales de perceptrón multicapa y modelo de árbol de decisiones de Ross Quinlan, y concluyen que ambos son modelos precisos. Al-Shayea et al. [3] proponen las redes de propagación retroalimentadas y las redes neuronales artificiales, concluyendo que estas últimas tienen mejor rendimiento en este tipo de problemas. Liu et al [4] utilizan el algoritmo genético para clasificar las características que influyen sobre la variable respuesta del problema. Finalmente, Khan et al. [5] abordan el problema utilizando lógica difusa, convirtiendo cada variable en categórica y clasificando los clientes con mayores posibilidades de aceptación de la campaña.

4 Análisis de Viabilidad de Validación Experimental

Debido a que el problema a resolver es de clasificación, en la que para cada usuario se conoce la variable respuesta, es decir, la decisión de inscribirse al depósito financiero o no, se seleccionará un conjunto de datos para entrenar el modelo y otro conjunto de datos para la validación experimental. Analizaremos, técnicas de balanceo de datos como el submuestreo, el sobremuestreo y el algoritmo Smote-Tomek para el mejoramiento de los resultados en los modelos. Finalmente, realizaremos validaciones cruzadas de K iteraciones para comprobar los resultados del modelo con conjuntos de datos distintos.

5 Referencias

[1] Guido, G., Prete, M. I., Miraglia, S., & De Mare, I. (2011). Targeting direct marketing campaigns by neural networks. *Journal of Marketing Management*, 27(9-10), 992-1006.

- [2] Elsalamony, H. A., & Elsayad, A. M. (2013). Bank direct marketing based on neural network and C5. 0 Models. *International Journal of Engineering and Advanced Technology (IJEAT)*, 2(6).
- [3] Al-Shayea, Q. K. (2013). Evaluating marketing campaigns of banking using neural networks. In *Proceedings of the World Congress on Engineering (Vol. 2)*.
- [4] Liu, H. H., & Ong, C. S. (2008). Variable selection in clustering for marketing segmentation using genetic algorithms. *Expert Systems with Applications*, 34(1), 502-510.
- [5] Khan, N., & Khan, F. (2013). Fuzzy based decision making for promotional marketing campaigns. *International Journal of Fuzzy Logic Systems*, 3(1), 64-77.
- [6] Moro, S., Cortez, P., & Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 62, 22-31.