

RESUMEN

Este trabajo tiene como objetivo encontrar uno o varios modelos analíticos que determinen la probabilidad de compra de un seguro de vida individual, mediante el uso y aplicación de metodologías de inteligencia de negocios y minería de datos, que definen los pasos y permiten estructurar el análisis de información histórica de clientes corporativos para determinar la probabilidad de compra de un producto de vida individual en base a las características de los clientes que ya han comprado un seguro permitiendo, mejorando la efectividad de las ventas. Para este estudio se han utilizado metodologías y herramientas que permiten determinar los siguientes pasos que son: definir las tareas, recolectar y analizar datos, elegir y configurar el modelo o los modelos, formatear datos, evaluar resultados y entregar resultado o el reporte a los tomadores de decisión, se han evaluaron y probaron cuatro modelos analíticos: Árboles de decisión, Clasificación Bayesiana, Redes neuronales y Random Forest los mismos que fueron entrenados y posteriormente comparados con datos históricos para determinar su grado de confianza. Se debe destacar que los datos utilizados para este análisis son datos demográficos como sexo, edad, ciudad de residencia, estado civil, número de hijos entre otros, también se analizan aspectos socio económicos como ingresos, si es bancarizado o si tiene tarjeta de crédito, toda información que se pueda obtener del cliente es importante para descubrir el perfil del cliente, todos estas actividades son el aporte de este trabajo, que busca definir un marco de trabajo a ser reutilizados en las futuras campañas de venta o en otros análisis de información.

Palabras clave: seguros de vida, minería de datos, plataformas de aprendizaje de máquinas, modelos predictivos, ciencia de datos.

Abstract.

This work aims to find one or more analytical models that determine the probability of purchasing individual life insurance, through the use and application of business intelligence and data mining methodologies, which define the steps and allow structuring the analysis of Historical information from corporate clients to determine the probability of purchasing an individual life product based on the characteristics of clients who have already purchased insurance, thus improving sales effectiveness. For this study, methodologies and tools have been used to determine the following steps, which are: define the tasks, collect and analyze data, choose and configure the model or models, format data, evaluate results and deliver the result or report to policy makers. decision, four analytical models have been evaluated and tested: Decision Trees, Bayesian Classification, Neural Networks and Random Forest, the same ones that were trained and later compared with historical data to determine their degree of confidence. It should be noted that the data used for this analysis are demographic data such as sex, age, city of residence, marital status, number of children, among others, socio-economic aspects such as income, whether you have a bank account or have a credit card, are also analyzed. All information that can be obtained from the client is important to discover the client's profile, all these activities are the contribution of this work, which seeks to define a framework to be reused in future sales campaigns or in other information analysis.

Keywords: life insurance, data mining, machine learning platforms, predictive models, data science.