PREDICTORS OF LIFE INSURANCE POLCY LAPSATION

Mallika Arachchige Subhashi Jayasumana

A dissertation submitted in partial fulfilment of the requirement for Master of Science degree in Business Analytics

Department of Computing

Informatics Institute of Technology, Sri Lanka

in collaboration with

Robert Gordon University, Aberdeen, United Kingdom

2022

Abstract

Human life is the most valuable asset, and life insurance can be considered as the most vital sort of insurance since it protects a person's family financially in the event of unforeseeable dangers or harm. Individuals are given safety and security by life insurance, and it also encourages saving. However, Life insurance penetration and density is still low in Sri Lanka compared to most of its regional peers mainly due to lack of awareness on importance of life insurance. Despite the low penetration levels, the higher rate of life insurance lapsation has become a major challenge for life insurance companies. Policy lapse is a crucial factor in life insurance since it has an impact on future pricing and the insurer's ability to remain solvent. Accurate prediction of lapses will help insurance companies to implement personalized strategies.

The study focuses on how life insurance policy lapsation determine by twelve selected independent variables (Policy type, Term of the policy, Premium frequency, Policy selling method, Sum assured, Customer location, Annualized premium, Gender of the customer, Occupation of the customer, Annual income of the customer, Age at the purchase of the policy and No of remaining years to pay). The study is based on 65,535 customer data collected from ABC Insurance company, a life insurance company operating in Sri Lanka. The major contribution for the study is comparison of five machine learning classification models (Logistic regression, Decision tree, Random forest, K-Nearest neighbor, and Naïve Bayes classification) for life insurance lapse prediction. The model accuracy was observed by comparing Precision, Recall F1 Score and Accuracy of each model.

The findings of the study suggest that Term of the Policy, Premium frequency, Annualized premium, Gender of the customer, Occupation of the customer, Annual income of the customer and Age at the purchase of the policy significantly affect the life insurance policy lapsation. The analysis of classification models recommends that use of Decision tree and Random forest technique when predicting life insurance policy lapsation.