LIFE INSURANCE RISK PREDICTION SYSTEM

Harindi Wanigaratne

A dissertation submitted in partial fulfilment of the requirement for Bachelor of Science
(Honours) degree in Computer Science

School of Computing

Informatics Institute of Technology, Sri Lanka
in collaboration with
University of Westminster, UK

2023

ABSTRACT

Life insurance is an essential financial tool that gives policyholders and their beneficiaries financial stability. Insurance businesses must accurately forecast risk levels in order to set policy premiums and keep long-term risk management procedures. This research paper provides a thorough analysis of the use of the CatBoost Machine Learning model to forecast risk levels in life insurance.

A overview of the literature is included at the start of the paper, highlighting methods and models currently used in risk prediction, including logistic regression, decision trees, and neural networks. The explanation for choosing the CatBoost model, a machine learning technique based on gradient boosting renowned for its great performance, handling of categorical data, and resilience to overfitting, is then covered in detail.

The data collecting, preprocessing, feature engineering, and model creation steps of the research approach are then followed by a thorough quantitative evaluation of the CatBoost model's performance using several evaluation criteria. The paper also evaluates the most practical product's usability and user interface, as well as the model's shortcomings, contributions to the field, and potential improvements.

The results of this study show that the CatBoost model performs better in risk level prediction when compared to conventional methods, making it an attractive tool for the life insurance sector. To further increase the effect of this study, future research should look at increasing the feature collection, enhancing model interpretability, and investigating cross-industry applications. In order to increase the model's applicability and relevance in the life insurance industry, the study finishes with a discussion of the successes in achieving the research's goals and objectives.

Keywords: Life Insurance, Risk Level prediction, CatBoost, Machine Learning, Risk Management