

**Predicting
Credit Card Default
Using Supervised Machine
Learning Approaches**

Wahalathantrige Chamod Harsha Bagya Perera

**Department of Computing
Informatics Institute of Technology, Sri Lanka
Collaboration with
Robert Gordon University, Scotland**

2023

Research Summery

Abstract

Based on current economic context the customers obtained loans, advances and credit card repayment probability is very less when considering the past days. May be several factor or factors affect for that result but it is not good behaviour pattern for the financial institutes like banks. So banks need to take serious analyse to identify those Non-Performing or the default customers in earlier. In the financial sector credit cards represents the main segments of the advances and loans in banking sector. Banks basically issuing the credit cards examine the customer's credit card application in initial step. But somehow their pattern of repayment may be not good after few time after issued the credit card.

The increasing prevalence of credit card usage has heightened the need for effective tools to predict credit card default. The ability to accurately identify customers at risk of default can empower financial institutions to make informed decisions, mitigate risks, and implement proactive measures to minimize financial losses. This thesis focus on explore and find various machine learning models to find out and predict credit card default based on the past 6-month behaviour of credit card customers.

The research begins with a comprehensive review of related literature on credit card default prediction, highlighting the significance of the problem and the existing approaches employed in the field. A thorough analysis of a real-world credit card dataset is conducted, focusing on key variables representing the past 6-month behaviours of customers. The dataset is pre-processed, scaling numerical features, encoding categorical variables, and handling missing values.

Numerous ML models are trained and evaluated on the pre-processed dataset, Logistic Regression, Support Vector Machine (SVM), Decision Trees, Random Forest Algorithm, Naïve Bayes Algorithm, K – Nearest Neighbour and Artificial Neural Networks (ANN).The evaluation is performed getting exact performance matrixes such as F1 score, recall, precision, and accuracy. The models are compared based on their predictive capabilities and the ability to identify credit card default accurately.

The outcomes in comparative analyzation highlight the 100% right accurate model for credit card default prediction. The chosen model demonstrates superior performance in terms of accuracy and

provides valuable insights into the influential factors contributing to default risk. Visualization techniques, including confusion matrices and feature importance plots, are utilized to enhance the interpretation and storytelling of the analysis.

The findings of this thesis commit to the onset knowledge that already on credit card default prediction and give efficient implications to financial institutions. In this recommended model will contribute as a valuable method for assessing credit risk, enabling proactive interventions and tailored strategies to prevent default. The thesis gives with an explanation on the limitations and for future research gives permanent paths, emphasizing the ongoing importance of accurate credit card default prediction in the dynamic financial landscape.

Keywords

Credit card default, machine learning, predictive modelling, comparative analysis, risk assessment.