Carolina University

Enhancing Credit Risk Evaluation in Automotive Finance: A Machine Learning Approach

by

Omoshola Owolabi Simon

Submitted to the Master of Data Science Program
April 2024

This document is provided in partial fulfillment of the requirements for the completion of a Master of Science in Data Science at Carolina University.

Approved by
Pereira Denzil.
Director, Data Governance, Volvo Financial Services

Murchtricia K. Jones, Ph.D.
Assistant Professor of Data Science

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

ACKNOWLEDGEMENT

I extend my profound gratitude to Dr. Murchtricia K. Jones, Ph.D., and all members of the supervisory committee, including Dr. Praharshin Senadeera, Ph.D., for their invaluable guidance, critical insight, and constant encouragement throughout the duration of my program.

I am immensely grateful to Volvo Financial Services for their generous approval to use their resources for the facilitation of this research work, which was needed to complete this project. Special gratitude goes to Pereira Denzil, who is the Director of Data Governance at Volvo Financial Service, for his unrelenting follow-up on the project. This appreciation is extended to all members of his team, too.

Special thanks to individuals who willingly shared their time, experience, and support for the smooth completion of this program and research. Paul Oghenekome and Ifeanyi Anyakora. Their contributions have provided invaluable support to this research.

My sincere appreciation also goes to my colleagues at Carolina University, particularly Prince Uche and Nathaniel Adeniken, whose discussions enriched this work significantly. Their perspective and critiques were invaluable to this research project.

Finally, I would like to express my heartfelt gratitude to my wife, Mrs. Toluwani Owolabi, as well as my friends for their understanding and support throughout this journey. Their belief in my abilities has been a source of inspiration, motivation, and resilience throughout my research journey.

Enhancing Credit Risk Evaluation in Automotive Finance: A Machine Learning Approach

Omoshola Owolabi
Department of Data Science
Carolina University
Winston Salem - North Carolina,
United States
owolabio@carolinau.edu

***Abstract* - Credit risk assessment is important for auto finance companies to reduce financial losses, defaults, and repossessions while promoting market stability. Traditional credit scoring models that rely on limited data have struggled to accurately predict the probability of default. This capstone project explores and focuses on the use of advanced machine learning techniques to improve credit risk assessment by leveraging large datasets and algorithmic complexity. This capstone project focuses on developing a machine learning model to predict consumer default in the automotive finance industry using Volvo Financial Services credit data. The project aims to identify important determinants of credit risk and evaluate how machine learning systems work well to predict the probability of defaults. The study begins with a comprehensive exploratory data analysis, which includes various financial indicators such as net worth, working capital, and revenue. This analysis helps in understanding the distribution of the target variable and identifying outliers that could affect the risk assessment mode This analysis helps in understanding the distribution of the target variable and identifying outliers that could affect the risk assessment model. The project follows a methodological approach that includes data preprocessing to address missing values and outliers and feature engineering to enhance the predictive power of financial key indicators in credit risk. The research work aims to improve the accuracy and efficiency of credit risk assessment by adopting machine learning models that can incorporate real-time data and capture non-linear relationships. This would help automotive finance companies predict and manage credit risk more effectively, thereby drastically reducing defaults and repossessions. Overall, this abstract highlights the process and expected outcomes of the study, with emphasis on the potential of machine learning to transform risk assessment practices in automotive finance.**

Keywords - Automotive, Finance, Credit risk assessment, Machine Learning, Predictive Modelling, Default Probability, Financial Indicators, Creditworthiness.

I. INTRODUCTION

A. Background

In the realm of automotive finance, evaluating credit risk has become crucial for ensuring reduced financial losses and repossessions, as well as strengthening market stability and minimizing default probabilities. As the industry continues to innovate new ways of increasing market share and minimizing risk, leveraging machine-learning approaches to assess potential credit risks has gained significant attention. This is due to the potential that the data science field has to enhance the efficiency and accuracy of credit risk assessment. As machine learning (ML) becomes increasingly representative and influential in finance, especially in the automotive industry, finance professionals are finding it capable of discovering subtle relationships, capturing various non-linear properties, and making dependable predictions for the management of risks.[1]

B. Problem

Volvo Financial Services (VFS), as well as other auto finance companies, in deciding to optimally minimize credit risk in this ever-changing market, traditional credit scoring systems have been found to have limitations in accurately predicting the probability of default because these methods rely on limited and static datasets. To address this issue, VFS is looking into the exploration of the use of machine learning to improve its credit risk assessment process, and also aims to improve the accuracy and efficiency of credit risk assessment in the automotive industry. By incorporating real-time processing into the credit risk evaluation model, auto finance companies at large can adequately predict and react swiftly to changing market situations and make informed decisions regarding auto loan approvals.

C. Objective

This capstone project aims to design and develop a machine learning model capable of predicting the probability of default of customers of auto companies. Historical financial data from VFS will be analyzed to identify key predictors of credit risk. These predictors will then be used in developing the machine learning model for the probability of default prediction. Various machine learning algorithms will be adopted with different evaluation metrics, and the most suitable one will be selected based on accuracy and performance. The models will also be tested on unseen data to ensure reliability and accuracy.

II. EXPLORATORY DATA ANALYSIS

Exploratory analysis of the data (EDA) was conducted to gain insight into the datasets and understand the relationship between variables. The EDA involved examining variable distributions, identifying patterns, and detecting outliers or anomalies. The dataset used for this analysis is 12,701 records and 25 features, with a mix of both categorical and numerical variables. The dataset contains various columns related to customer information and financial indicators such as the total net worth, working capital, total exposure, revenue, depreciation, EBIT, depreciation, net profit, assets and liabilities, credit rating score, and payment discipline score.

A. U*nivariate Analysis*

In the univariate analysis, the distribution of the target variable was analyzed to understand the central tendency and the range of the target variable, providing insight into overall credit risk assessment. The distribution of customer information and financial indicators are also examined to identify outliers and patterns that could influence the robustness of the risk assessment model. The distribution of the target variable shows the skewed distribution of the target.

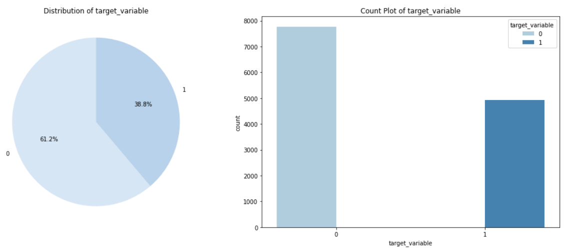

Fig. 1. Pie and Bar graph showing the distribution of the target variable

Fig. 1. above shows two graphs: a pie chart and a bar graph. The pie chart represents the distribution of the target variable, divided into the binary classes of the target variables '0' and '1'. Class '0' makes up 61.2% of the total distribution, while Class '1' accounts for approximately 40%.

The bar graph visually represents this distribution, the bar for class '0' being much taller, indicating a higher count, and the bar for class '1' being shorter, reflecting a lower count. This suggests that the target variable relates to creditworthiness, where category '0' represents low-risk assessment, while category '1' relates to high-risk or non-creditworthy cases.

This information can be used in the credit risk evaluation for automotive finance to determine credit approvals for different classes.

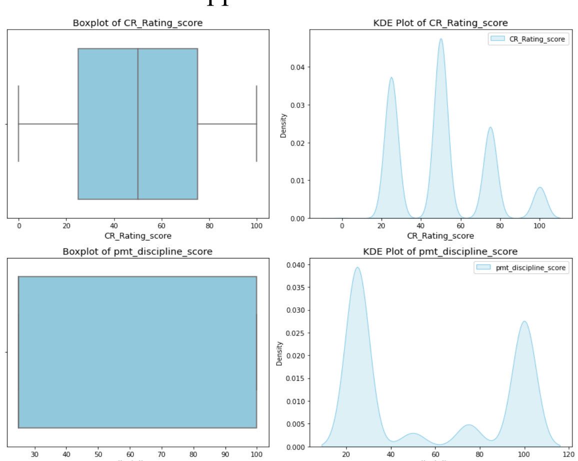

Fig. 2. Boxplot and the distribution of credit rating score and payment discipline score.

In Fig. 2. above, the boxplot on the left represents the distribution of credit rating, showing that most data points fall within the interquartile range (IQR), and there are no visible outliers. The IQR suggests that a significant number of individuals have scores between approximately 40 and 60. On the right, the kernel density estimation (KDE) plot of credit rating exhibits multiple peaks, indicating a multimodal distribution with prominent peaks around scores of 20, 50, and 80.

These peaks suggest distinct clusters of individuals with these scores. Moving to the payment discipline, the boxplot in the second row reveals a concentrated distribution with almost all values concentrated around a specific score. The KDE plot on the right confirms this observation, as it has a single prominent peak around a score of approximately 70. These findings provide important insights into the distribution and concentration of credit risk scores in automotive finance, which should be considered.

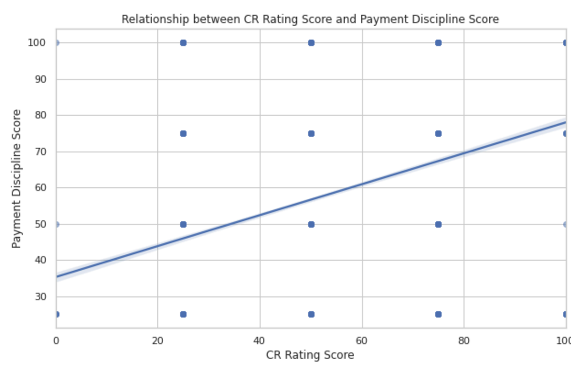

Fig.3. Relationship between the credit rating score and payment discipline score

Fig. 3. shows data points representing individual cases or observations. The credit rating and payment discipline have a positive correlation, meaning that as the credit rating score increases, so does the payment discipline score. This is supported by the upward trajectory of the blue trend line. A higher credit rating score suggests better creditworthiness or

financial stability, while a higher payment discipline score indicates a more reliable payment history. Essentially, individuals or entities with higher CR Rating Scores are more likely to have better payment discipline. This information is important for assessing credit risk in automotive finance.

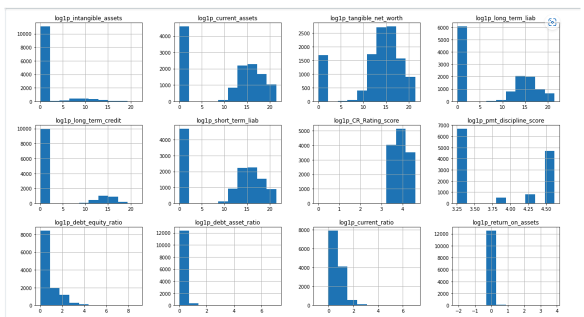

Fig. 4. Histograms for key financial indicators

The Fig. 4. above presents a series of histograms that visualize key financial indicators for credit risk assessment; these histograms provide insights into the distribution of the data, the frequency, and trends capable of aiding decision-making. Some metrics exhibit skewed distribution, indicating concentration at specific values.

We analyze three key metrics, ranging from Debt Asset Ratio, Current Ratio, and Return on Assets (ROA). The Debt Asset Ratio measures the level of debt in proportion to the total assets while reeling a company's financial risks. The Current Ratio assesses a company's liquidity by comparing current assets to current liabilities. ROA quantifies the company's profitability relative to its assets. Our analysis of these metrics shows the following statistical summaries. Debt Asset Ratio has a mean of approximately 0.34 with a wide standard deviation and a range from -8.36 to 1205.0. The Current Ratio has a mean of approximately 1.38, with a broad range from negative values to 11.28.0. ROA has a positive mean of about 0.03 with a standard deviation indicating varying efficiency in asset utilization. Negative ROA values, such as -21.39, indicate potential operational challenges. Overall, these metrics provide valuable information about a customer's financial situation and performance.

B. *Bivariate Analysis*

In the bivariate analysis, the relationship between customer information and financial indicators was explored to identify potential trends or correlations. This analysis aimed to determine if the simplicity or complexity of financial ratios and indicators had a significant impact on the probability of default.

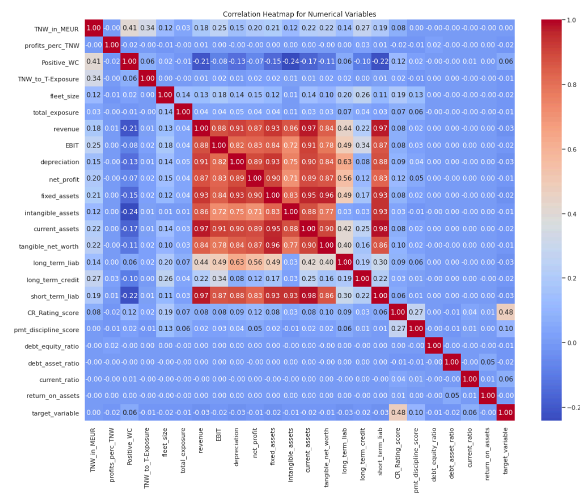

Fig.5. Correlation heatmap between key features

In Fig. 5. the heatmap represents the correlation coefficient between two variables. The color scale goes from dark red for positive correlation to dark blue for negative correlation, with white showing no correlation. The x-axis and y-axis labels represent different financial metrics. Notable correlations include a strong positive correlation between profits and positive working capital, indicating that an increase in profits leads to an increase in positive working capital. Tangible net worth and long-term liability have a strong negative relationship, implying that a decrease in tangible net worth leads to an increase in long-term liabilities. Depreciation and net profit, on the other hand, have almost no relationship. Understanding these correlations is important for credit risk assessment, as positive correlations suggest favorable financial health, negative correlations indicate potential risk and variables with no correlation are independent of each other.

III. LITERATURE REVIEW

A. *Credit Risk in Automotive Finance*

Machine learning is a popular choice for accessing credit risk in the automotive industry. Khashman (2011) explored neural networks as one of the machine learning models for credit risk evaluation in the automotive industry, highlighting its ability to capture complex patterns and non-linear relationships in data. However, the author cautioned about their lack of interpretability and potential for overfitting. The author also compared the neural networks with traditional statistical methods, noting their applicability, advantages, and potential limitations. Butaru et al. (2016) focused on credit risk management in the credit card industry, with emphasis on the need to integrate data from multiple sources and utilize advanced analytics, such as machine learning techniques, for fraud detection and credit scoring. They highlighted the importance of robust data structure and advanced modeling techniques to efficiently and effectively manage credit risk in a rapidly changing industry.[2][3]

B. *Techniques for Credit Risk Assessment*

The credit risk assessment in the automotive finance industry has attracted attention for its potential to reduce financial risk and losses, repossessions, and defaults while promoting market stability. machine learning techniques have been utilized to enhance the accuracy and efficiency of credit risk evaluation. Machine learning has proven effective in capturing subtle relationships and non-linear properties, resulting in reliable risk prediction.[4] Deep learning techniques, including recurrent neural networks and long-short-term memory networks, have also been explored for credit risk assessment. These models have demonstrated their ability to identify complex patterns and improve credit risk management. Additionally, a dynamic ensemble classification approach has been proposed for credit risk, which combines multiple ML models using soft probability. This approach has been shown to enhance the accuracy and stability of credit risk assessment, particularly in the presence of concept drifts and class imbalance. In general, leveraging machine learning techniques and methodologies can greatly improve credit risk assessment in the automotive finance industry[5]

C. *Traditional Models and Machine Learning*

Khandani et al. (2010) found that machine learning models, such as decision trees, neural networks, and support vector machines, were more effective in predicting consumer credit risk than traditional statistical models. ML models were able to capture non-linear patterns and interactions between variables, resulting in superior predictive performance. Traditional credit scoring systems, which rely on limited and static datasets, have limitations in accurately predicting default probabilities. To address this, several studies have explored the application of marching learning algorithms for credit risk assessment.[6] Lessmann et al. (2015) conducted a comprehensive review and benchmarking of various classification algorithms, including a machine learning model, for credit scoring. They evaluated these algorithms on multiple datasets and provided recommendations for selecting appropriate models based on data characteristics and business objectives. Their study emphasized the importance of considering data quality, class imbalance, and model interpretability when choosing ML models for credit risk assessment.[7]

D. Advanced Machine Learning Approach

Luo et al. (2017) found that deep learning models are able to capture complex patterns and temporal dependencies, which are important for credit risk assessment. However, interpreting the predictions of these models can be challenging.[8] Similarly, Xia et al. (2017) explored the use of cost-sensitive boosted tree models for loan evaluation in peer-to-peer lending. They addressed the problem of imbalanced credit risk data by using cost-sensitive learning models, aiming to improve the assessment of minority classes and reduce financial losses. While these studies are not specific to the automotive industry, they contribute to the knowledge of cost-sensitive approaches and handling imbalanced data for credit risk assessment.[9]

E. Automotive Credit Decision

Artificial intelligence technology is transforming the automotive industry by streamlining credit decision-making. Evaluating borrower creditworthiness and determining loan terms is critical for lenders and consumers. Automating this process reduces time and errors associated with traditional manual underwriting, ensuring safer and more efficient transactions.[10]

F. Credit Scoring System

Credit scoring is a crucial tool used by lenders to evaluate lending to borrowers. It involves analyzing factors like credit history, income stability, and financial obligations to assign a numerical score. This helps to determine the borrower's likelihood of repaying the loan. Credit scoring is important for various reasons such as. Firstly, it aids in risk assessment, with higher scores indicating lower risks. It also provides a standardized approach to evaluating credit applications, reducing biases. The automated nature of credit scoring systems enables quick and efficient processing of auto loan applications, which is valuable in consumer markets. Factors like credit history and debt-to-income ratio. However, there are challenges and limitations, including data quality, potential bias, security and privacy concerns.[11]

G. Algorithmic Bias in Credit Risk Assessment

Machine learning models have advantages but also face challenges, including algorithmic bias. [12] addressed bias in machine learning models for credit risk assessment, which is important in automotive finance. Biased models can result in unfair credit decisions, risking regulations and the reputation of financial institutions. This literature advocates for accuracy and fairness in machine learning algorithms to avoid perpetuating societal biases.

H. Integration of ML in Automotive Finance Systems

Fayyaz, Rasouli, and Amiri (2020) proposed a novel approach for credit risk prediction in automotive finance systems. Their network-aware, data-driven model enhances accuracy and provides a holistic view of risk, considering the interconnectedness of financial actors within the industry. It offers valuable insights for institutions looking to integrate machine learning into their existing systems.[4]

IV. METHODOLOGY

We will conduct a comparative study of four machine learning models (Logistic Regression, SVM, Naive Bayes, and Neural Networks) to enhance credit risk evaluation in automotive finance. The dataset, collected from reputable financial institutions, includes customer attributes and credit risk indicators. The raw data will undergo preprocessing, addressing missing values, outliers, and quality issues. Categorical variables will be encoded, while the numerical variables will be standardized and normalized as required by assumptions of each machine learning model to be adopted. The dataset will be split into training and testing sets for robust model evaluation. We will carry out hyper-parameterization and cross-validation to ensure the effectiveness of the model. We will extract meaningful features and improve model predictability. Relevant features will be selected using correlation analysis, and domain expertise was consulted for the engineering of key financial indicators and ratios, for the improvement of the model performance and interpretability.

A. *Data Preprocessing*

Data clearing involves dropping irrelevant columns, such as the data metadata, and filtering records with missing or irrelevant data. Data transformation includes renaming columns for clarity and correcting data types for accurate analysis. Missing values are handled by imputing numeric columns with mean values from non-missing data to maintain the statistical properties of the dataset.

B. *Feature Engineering*

Feature engineering is critical in enhancing the predictive power of the model. It involves aggregating financial variables like revenue, EBIT, and net profit over different periods. Non-numerical features such as credit rating and payment discipline are also converted into numerical scores. Engineering financial ratios provide insight into the financial health of entities, including debt-to-equity and asset ratios, as well as liquidity ratios like the current ratio. The target variable for the model is defined based on financial ratio, specifically, the current ratio threshold, credit score, and payment discipline. This methodical approach ensures that the dataset is robust and ready for subsequent analytical phases in predicting credit risk in automotive finance.

C. *Machine Learning Models*

1) *Logistic Regression - Theoretical Understanding*

The theoretical basis of logistic regression, a statistical technique widely used in machine learning for binary classification, is important to understand its application in various fields, including credit risk analysis in automotive finance.

Logistic regression is basically a regression model where the dependent variable is categorical. It estimates the probability that a given input point belongs to a class, usually through the use of a logistic function. This function creates an S-shaped curve that can literally take any number value and map to a value from 0 to 1, but never exactly within those limits. This property makes it suitable for binary classification problems, such as predicting the probability of credit risk analysis.

Mathematically, logistic regression applies a logistic function to a linear combination of input characteristics. The coefficients of this linear combination are known from the training data, usually using maximum likelihood estimation. This strategy requires the refinement of the simulations to improve the observed data under the model. Optimization is usually done using conventional methods such as gradient descent.

Logistic function

The sigmoid function is defined as

$$\sigma(z) = \frac{1}{1+e^{-z}} \tag{13}$$

where (e) is the base of the natural logarithm and (z) is the linear combination of the input features (x), as given:

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + + \beta_n u x_n$$

. (13)

Probability Calculation

Logistic regression models probability as follows:

$$P(Y = 1|X = x) = \sigma(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + +\beta_n x_n);$$

. (14)

where $(P(Y = 1|X = x))$ is the probability that output (Y) is class 1 given the input vector (X).

Model Training

Model parameters are estimated using maximum likelihood estimation. The probability function is:

$$L(\beta) \prod_{i=1}^{n} P(y_i|x_i;\beta)^{y_i}(1 - P(y_i|x_i;\beta))^{(1-y_i)}$$

. (15)

Log-Likelihood

The log-likelihood is used for optimization:

$$\ell(\beta) = \sum_{me=1}^{n} [y_i \log(P(y_i|x_i;\beta)) + (1 - y_i)\log(1 - P(y_i|x_i;))]$$

. (13)(15)

This model is adopted due to the ability to treat two outcomes of logistic regression—either predetermined or unpredicted—as an appropriate choice for predicting credit risk. It predicts the probability of an instance belonging to a class by fitting a logistic function to the input features. It aims to find optimal coefficients through techniques like gradient descent or maximum likelihood estimation. Logistic regression assumes a linear relationship between input features and the target variables's log-odds. In our case, the model was used to predict the probability of credit default using financial metrics. The model was trained with Spark's LogisticRegression: features were normalized and scaled for better prediction. It achieved an accuracy of 89.7% on training data and 89.1% on test data, which demonstrated an effective prediction of credit risk.

2) *Support Vector Machines - Theoretical Understanding*

Support Vector Machine (SVM) is a versatile algorithm for both regression and classification tasks. It creates hyperplanes in high-dimensional spaces to maximize the margin between classes. By using kernel functions, SVM can handle nonlinear boundaries by mapping features to higher-dimensional space where classes are linearly separable. In SVM, the margin is the distance between the closest data points of each class and the decision plane. These points are known as support vectors. The optimal hyperplane is the one that maximizes the margin. A hyperplane is an n-dimensional space defined by an equation.

Hyperplane Definitions

$$\mathbf{w} \cdot \mathbf{x} - b = 0 \quad (16)$$

where $(\mathbf{w})$ is the weight vector, $(\mathbf{x})$ represents the input features, and (b) is the bias term. [16]

SVM Classification

In binary classification, SVM aims to find the hyperplane that best separates the two classes. The decision function is:

$$f(x) = \text{sign}(\mathbf{w} \cdot \mathbf{x} - b) \quad (17)$$

This function returns $+1$ or -1 depending on which side of the hyperplane the data point lies. [17]

SVM was applied to the problem of credit risk evaluation in automotive finance as It highlights the model's accuracy in classifying customers into high or low-credit-risk categories based on various financial indicators. This project demonstrates the SVM's utility in handling complex, high-dimensional datasets and its effectiveness in binary classification tasks. Spark's LinearSVC class was used to train an SVM model for classifying customers into high or low credit risk. The model achieved

87.9% accuracy on the training set while achieving 867% accuracy on the test set, confirming its efficacy in binary credit risk classification.

3) Theoretical Understanding - Naive Bayesian

Naive Bayes is a widely used probabilistic classifier model, which is based on Bayes' theorem. It comes with the assumption that features are conditionally independent. The algorithm calculates the probability of each class and selects the one with the highest probability. Baye's Theorem is mathematically expressed as

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)} \quad (18)(19)$$

This theorem updates the probability estimate for a hypothesis as more evidence becomes available.[6][7] Bayesian analysis involves updating beliefs about unknown parameters through the prior distribution of $(P(H))$ and refining it to the posterior distribution $(P(H|E))$, combining prior beliefs with the likelihood of observed data.[5] The likelihood function $(P(E|H))$ represents the probability of observing the data given certain parameter values and is a key component in updating the posterior. Bayesian inference is a process of continuously updating the belief about parameters as new data is observed, refining the posterior distribution over time.[20][21]

In the context of credit risk evaluation in automotive finance, Bayesian analysis provides a framework for incorporating prior economic conditions and updating these beliefs with customer data. This approach can improve the accuracy and robustness of risk assessment. The Gaussian model of Naive Bayes was employed for analysis, taking into account the continuous nature of financial variables. It achieved 77.5% accuracy on the training data and 77.9% on the test data. This indicates potential limitations compared to alternative models for this dataset.

4) Neural Networks

Neural networks, a cornerstone of modern machine learning, provide powerful tools for modeling complex relationships and patterns in data. They are particularly suited for tasks such as credit risk analysis, where traditional statistical methods may struggle to capture nonlinear interactions between variables. Neural networks consist of a network of neurons, each designed to perform a specific computation. The basic configuration includes an input layer, one or more stored layers, and an output layer. Each neuron in the layer is connected to many others in the next layer, and notably, its weight is variable as the network learns.

Activation Functions

Neurons use activation functions to convert their weighted input into an output signal. Sigmoid is a common function that maps onto input to a value from 0 to 1, making it optimal for binary classification tasks. ReLU (Rectified Linear Unit), which is used to add nonlinearity to the model, is provided it learns more complex applications. In forward propagation, input data is passed through the network from the input layer to the output layer. At each neuron, the following calculation occurs:

$$a^{(l+1)} = \sigma(W^{(l)}a^{(l)} + b^{(l)}) \quad (22)$$

where $(a^{(l)})$ is the activation from the previous layer, $(W^{(l)})$ is the weight matrix for layer (l), $(b^{(l)})$ is the bias, and (σ) is the activation function. While backpropagation is crucial for neural network learning as it calculates the gradient of the loss function, adjusting weights to minimize errors and improve predictions.[22]

V. RESULT AND DISCUSSION

A. Evaluation Metrics for Credit Risk Models

In credit risk assessment, classification metrics are vital for evaluating a predictive model's ability to differentiate between clients who will default on a loan and those who will not. Precision helps to minimize the risk of loss by avoiding unnecessary loan denial. Recall is important to identify actual defaulters and minimize the risk of bad debt. The F1-score balances the cost of false positives and false negatives. This research work adopts precision, recall, accuracy, and f1-score for the evaluation metrics for the classification models adopted.

Classification report helps to understand the models's efficacy and fine-tune them according to specific application needs. It is essential for understanding model performance, especially credit risk by considering the implications of different types of errors. Empirical evaluation of machine learning models analyzes their performance across all the metrics. Logistic regression, SVM, Naive Bayes, and Neural network models are compared and evaluated on the Databriks platform using PySpark.

B. Interpretation of Evaluation Metrics and Key Insights

The table below shows the comparison between the machine learning models, such as the logistic regression, support vector machine, naive bayes, and neural networks. It summarizes their performance metrics, aiding in easy comparison for selecting the most suitable model for different predictive analytics tasks.

The results are presented in the table below:

TABLE I. CONSOLIDATED PERFORMANCE METRICS

Model	Precision (Positive Class)	Recall (Positive Class)	F1-Score (Positive Class)	Training Accuracy	Test Accuracy
Logistic Regression	0.74	0.59	0.66	76.04%	74.26%
Support Vector Machine (SVM)	0.78	0.91	0.84	86.43%	85.33%
Naive Bayes	0.4	0.98	0.57	45.02%	42.69%
Neural Networks	0.67	0.76	0.71	75.88%	76.25%

Some of the key insights from Table I above are that the SVM has the highest accuracy and recall, making it suitable for critical applications. Naive Bayes has a high recall but with low precision, suggesting numerous false positives. Neural Networks have consistent accuracy and balance between precision and recall, suitable for complex patterns. Logistic regression has decent precision but the lowest recall, potentially missing more positive cases. However, its simplicity and speed make it advantageous for less critical applications where interpretability and quick decision-making are needed.

C. Error Metrics and Key Insights

The table below shows the tabular comparison of the error metrics for each of the models expressed in percentages. These metrics include True Positives(TP), True Negatives(TN), False Positives(FP), and False Negatives(FN), providing a detailed view of each of the model's performance in handling classification tasks. These metrics were generated from the computation of the confusion matrix for each of the models.

TABLE II. ERROR METRICS IN PERCENTAGES

Model	True Positive (%)	True Negative (%)	False Positive (%)	False Negative (%)
Logistic Regression	87%	77%	23%	13%
Support Vector Machine (SVM)	91%	83%	17%	9%
Naive Bayes	98%	8%	92%	2%
Neural Networks	76%	76%	24%	24%

The above error metrics in Table II provide information on different measures:

- True Positive (TP) - the correct identification of actual positives.
- True Negative (TN) - the correct identification of actual negatives.
- False Positive (FP) - the incorrect identification of negatives as positives.
- False Negative (FN) - the incorrect identification of positives as negatives.

A higher TP and TN percentage indicates better model performance, while lower FP and FN percentages indicate fewer incorrect positive predictions and missed positive cases, respectively.

From the table, the Support Vector Machine (SVM) is highly accurate in classifying positive and negative cases, with the highest TP and TN percentages. Naive Bayes, on the other hand, tends to classify most inputs as positive, leading to a low TN percentage. Neural Networks show consistent performance but have room for improvement in reducing false positives (FP) and false negatives (FN). Logistic Regression exhibits a good balance but is less effective in identifying positive cases compared to SVM. These metrics are crucial for understanding the strengths and weaknesses of each model in terms of error handling, aiding in model selection based on specific needs regarding error sensitivity.

The objective is to find a balance between minimizing false positives and false negatives while correctly predicting the true classifications. The impact of reducing one type of error depends on the model's performance and the project's specific credit risk requirements. In credit risk analysis, it is important to minimize the Type I error where eligible customers are denied credit, affecting business opportunities and customer satisfaction. Similarly, Type II errors should be minimized to prevent high-risk individuals from being categorized as low-risk, which can lead to financial losses for the company.

D. *Interpretation of Model Performance and Precision*

Table I above shows the performance of different models on various evaluation metrics, alongside their interpretability and computational efficiency; in contrast, Table II shows the classification of instances across different classes.

Logistic Regression - This model demonstrates a relatively decent and moderate precision of 0.74 and recall of 0.59, resulting in an F1-score of 0.66. This indicates that it is moderately effective in identifying positive cases, which is important in applications with serious consequences for false negatives. The training and test accuracies were 76.04% and 74.26%, respectively, demonstrating stable performance on unseen data. Logistic Regression is favored for its simplicity and interpretability, making it suitable for initial risk assessments that require regulatory compliance and stakeholder interpretability. However, its relatively lower recall rate may result in a higher rate of missed defaults and potential financial losses.

Support Vector Machine (SVM)- The SVM demonstrates strong performance in credit risk analysis with an accuracy of 86.43% and a test accuracy of 85.33%. It's precision for the positive class reached 0.78, and the recall was 0.91, culminating in an F1-score of 0.84. These figures indicate a strong ability to generalize well from training to test data, making it particularly suitable for datasets with complex non-linear relationships. It is effective in identifying positive cases but with many false positives. This sensitivity is useful in preliminary screening for rare events. SVM, with high precision and recall, is effective in complex risk assessment scenarios, such as distinguishing defaulting and non-defaulting loans. It performs well with nonlinear relationships and high-dimensional data in automotive finance, including borrower's credit history, loan terms, and economic factors.

Naive Bayes - Naive Bayes showed a remarkable recall rate of 0.98 but a low precision of 0.40 for the positive class, resulting in an F1-score of 0.57. This indicates that the model is highly effective at detecting positive cases but at the expense of a high number of false positives. This characteristic could prove useful in preliminary screening situations where missing rare events is more problematic than dealing with false alarms, such as disease outbreaks. However, in the financial domain, Naive Bayes' low precision and high false positive rate could result in many customers being mistakenly identified as high-risk, leading to missed opportunities and customer dissatisfaction. The automotive industry, where financing approval affects sales, needs to carefully manage the balance between sensitivity and specificity.

Neural Networks - Configured with two hidden layers, Neural Networks showed balanced performance with a precision of 0.67 and a recall of 0.76 for positive class. Training and test accuracies were closely matched at 75.88% and 76.25%, respectively. This model's ability to capture intricate patterns without extensive feature engineering makes it valuable, especially in complex predictive tasks where traditional models struggle with linearity assumptions. Neural Networks are well-suited for capturing complex patterns in large datasets typical of the automotive industry, which include variables from macroeconomic indicators to individual credit histories. Despite being computationally intensive, they are highly effective for dynamic and nuanced risk assessments in modern credit environments.

Practical Implications in the Automotive Industry

The varied performance across these models highlights the importance of selecting the right model for different applications. SVM is ideal for scenarios where failing to detect an event could have severe consequences. Logistic Regression provides a good trade-off between speed and accuracy. Naive Bayes is best used in preliminary tests, and Neural Networks offer a flexible architecture. The analysis of precision and recall across models reveals their strengths and weaknesses. SVM is effective in tasks requiring accurate identification of true positives, while logistic regression offers quicker computations and scalability. Decision-makers should consider accuracy, computational efficiency, and model interpretability when choosing a model. In the automotive industry, SVM is crucial for revenue generation and customer retention. Naive Bayes and SVM are vital for minimizing credit losses. Neural Networks offer a compromise between computational complexity and performance.

VI. CONCLUSION

The integration of machine learning techniques in automotive finance marks a significant improvement compared to traditional credit scoring methods. This project has demonstrated that machine learning can greatly enhance the accuracy and efficiency of credit risk assessments. Different models, like Logistic Regression, Support Vector Machines, Naive Bayes, and Neural Networks, each showed unique strengths in handling complex and dynamic credit risk data. While some models excelled in accuracy, others offered benefits in computational efficiency or non-linear data handling.

Applying these models to real-world data from Volvo Financial Services (VFS) has provided valuable insights into default predictors, enabling more informed decision-making. Incorporating these predictive models into risk assessment frameworks allows automotive finance companies to decrease loan defaults, minimize financial losses, and better manage risks in a rapidly changing economic environment. This highlights the transformative potential of machine learning in financial risk assessment, suggesting a shift towards data-driven, analytical approaches in the automotive finance industry.

Further research is encouraged to refine these models, explore new algorithms, and expand their application in various sectors of financial services. Machine learning improves credit risk assessment in auto finance by enabling precise decision-making. Updating data evaluation and models is essential to adapt to economic shifts and changing consumer behaviors.

Although implementing these models offers benefits such as improved prediction accuracy, operational efficiency, and reduced financial risks, there are also shortcomings to address. Algorithmic bias and interpreting complex model outputs need attention to ensure fairness and transparency in credit evaluations. Future research areas include integrating alternative data sources like key macroeconomic and microeconomic indicators to enhance predictive accuracy. Additionally, exploring advanced machine learning techniques such as deep learning and ensemble methods can improve model performance with non-linear relationships and large datasets. Developing strategies to mitigate bias and real-time analytics capabilities for dynamic risk assessment are also important for equitable credit decisions across different demographics.

REFERENCES

[1] Fayyaz, M. R., Rasouli, M., & Amiri, B. (2020). A data-driven and network-aware approach for credit risk prediction in supply chain finance. In Ind. Manag. Data Syst. (Vols. 121, pp.785-808). Ind. Manag. Data Syst. https://doi.org/10.1108/imds-01=2020-0052

[2] F. Butaru, Q. Chen, B. Clark, S. Das, A. Lo, and A. Siddique, “Risk and Risk Management in the Credit Card Industry”, Financial Crises eJournal. Financial Crises eJournal, Jun. 01, 2015. [Online]. Available: https://doi.org/10.2139/ssrn.2618746

[3] A. Khashman, “Credit risk evaluation using neural networks: Emotional versus conventional models”, Appl. Soft Comput., vol. 11. Appl. Soft Comput., pp. 5477–5484, Dec. 01, 2011. [Online]. Available: https://doi.org/10.1016/j.asoc.2011.05.011

[4] M. R. Fayyaz, M. Rasouli, and B. Amiri, “A data-driven and network-aware approach for credit risk prediction in supply chain finance”, Ind. Manag. Data Syst., vol. 121. Ind. Manag. Data Syst., pp. 785–808, Jul. 02, 2020. [Online]. Available: https://doi.org/10.1108/imds-01-2020-0052

[5] X. Feng, Z. Xiao, B. Zhong, J. Qiu, and Y. Dong, “Dynamic ensemble classification for credit scoring using soft probability”, Appl. Soft Comput., vol. 65. Appl. Soft Comput.,

pp. 139–151, Apr. 01, 2018. [Online]. Available: https://doi.org/10.1016/j.asoc.2018.01.021

[6] A. Khandani, A. Kim, and A. Lo, "Consumer Credit Risk Models Via Machine-Learning Algorithms", American Finance Association Meetings (AFA). American Finance Association Meetings (AFA), Mar. 11, 2010. [Online]. Available: https://doi.org/10.2139/ssrn.1568864

[7] S. Lessmann, B. Baesens, H. Seow, and L. C. Thomas, "Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research", Eur. J. Oper. Res., vol. 247. Eur. J. Oper. Res., pp. 124–136, Nov. 16, 2015. [Online]. Available: https://doi.org/10.1016/j.ejor.2015.05.030

[8] C. Luo, D. Wu, and D. Wu, "A deep learning approach for credit scoring using credit default swaps", Eng. Appl. Artif. Intell., vol. 65. Eng. Appl. Artif. Intell., pp. 465–470, Oct. 01, 2017. [Online]. Available: https://doi.org/10.1016/j.engappai.2016.12.002

[9] Y. Xia, C. Liu, and N. Liu, "Cost-sensitive boosted tree for loan evaluation in peer-to-peer lending", Electron. Commer. Res. Appl., vol. 24. Electron. Commer. Res. Appl., pp. 30–49, Jul. 01, 2017. [Online]. Available: https://doi.org/10.1016/j.elerap.2017.06.004

[10] H. Huo, D. Luo, and Z. Yan, "Automaker's credits strategy considering fuel consumption and endurance capacity constraints under dual-credit policy in China", vol. 10. Jan. 12, 2023. [Online]. Available: https://doi.org/10.3389/fenrg.2022.963900

[11] Y. Hayashi, "Emerging Trends in Deep Learning for Credit Scoring: A Review", Electronics. Electronics, Oct. 03, 2022. [Online]. Available: https://doi.org/10.3390/electronics11193181

[12] Y. Wei, "Application of Machine Learning and Artificial Intelligence in Credit Risk Assessment", 2023 2nd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS). 2023 2nd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS), pp. 150–156, Jul. 01, 2023. [Online]. Available: https://doi.org/10.1109/AIARS59518.2023.00037

[13] D. W. Hosmer, . Jr., S. Lemeshow, and R. X. Sturdivant, Applied Logistic Regression. John Wiley & Sons, 2013. [Online]. Available: http://books.google.com/books?id=64JYAwAAQBAJ&dq=Applied+Logistic+Regression.+John+Wiley+%26+Sons&hl=&source=gbs_api

[14] T. Hastie, R. Tibshirani, and J. H. Friedman, The Elements of Statistical Learning. 2009. [Online]. Available: http://books.google.com/books?id=eBSgoAEACAAJ&dq=The+Elements+of+Statistical+Learning.+Springer&hl=&source=gbs_api

[15] C. M. Bishop, Pattern Recognition and Machine Learning. Springer Verlag, 2006. [Online]. Available: http://books.google.com/books?id=kTNoQgAACAAJ&dq=Pattern+Recognition+and+Machine+Learning.+Springer&hl=&source=gbs_api

[16] B. Boser, I. M. Guyon, and V. Vapnik, "A training algorithm for optimal margin classifiers". pp. 144–152, Jul. 01, 1992. [Online]. Available: https://doi.org/10.1145/130385.130401

[17]N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, 2000. [Online]. Available: http://books.google.com/books?id=_PXJn_cxv0AC&dq=An+Introduction+to+Support+Vector+Machines+and+Other+Kernel-based+Learning+Methods&hl=&source=gbs_api

[18] T. Bayes, R. Price, and J. Canton, An Essay Towards Solving a Problem in the Doctrine of Chances. 1763. [Online]. Available: http://books.google.com/books?id=Xi2wpwAACAAJ&dq=An+Essay+towards+solving+a+Problem+in+the+Doctrine+of+Chances.&hl=&source=gbs_api

[19] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, Bayesian Data Analysis, Third Edition. CRC Press, 2013. [Online]. Available: http://books.google.com/books?id=ZXL6AQAAQBAJ&dq=Bayesian+Data+Analysis.+CRC+Press&hl=&source=gbs_api

[20] J. M. Bernardo and A. F. M. Smith, Bayesian Theory. John Wiley & Sons, 2009. [Online]. Available: https://play.google.com/store/books/details?id=11nSgIcd7xQC&source=gbs_api

[21] C. M. Bishop, Pattern Recognition and Machine Learning. 2023. [Online]. Available: http://books.google.com/books?id=0BEs0AEACAAJ&dq=Pattern+Recognition+and+Machine+Learning&hl=&source=gbs_api

[22] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016. [Online]. Available: http://books.google.com/books?id=Np9SDQAAQBAJ&dq=Courville,+A.+(2016).+Deep+Learning&hl=&source=gbs_api

APPENDIX

DATA DICTIONARY

TABLE III. DATA DICTIONARY FOR THE DEFINITION OF VARIABLES

Column Name	Description	Data Type	Nullable
customer_name	The name of the customer could be an individual or a company.	String	TRUE
TNW_in_MEUR	Total Net Worth (TNW) in Million Euros, representing the financial value of the customer's net worth.	Double	TRUE
profits_perc_TNW	Profits as a percentage of Total Net Worth, indicating the proportion of profits relative to the total net worth.	Double	TRUE
Positive_WC	Positive Working Capital, representing the excess of current assets over current liabilities, indicates liquidity.	Double	TRUE
TNW_to_T-Exposure	The ratio of Total Net Worth to Total Exposure indicates the financial leverage of the customer.	Double	TRUE
fleet_size	Size of the fleet owned by the customer, such as the number of vehicles or assets.	Double	TRUE
total_exposure	Total financial exposure of the customer, representing the total financial risk.	Double	TRUE
revenue	Total income generated by the customer.	Double	TRUE
EBIT	Earnings Before Interest and Taxes, indicating operating profit before deducting interest and taxes.	Double	TRUE
depreciation	Depreciation expenses represent the decrease in the value of assets over time.	Double	TRUE
net_profit	Net Profit, representing the profit after deducting all expenses from revenue.	Double	TRUE
fixed_assets	Value of fixed assets, representing long-term tangible assets owned by the customer.	Double	TRUE
intangible_assets	Value of intangible assets, such as patents or goodwill.	Double	TRUE
current_assets	Value of current assets expected to be converted into cash within a year.	Double	TRUE
tangible_net_worth	Tangible Net Worth, representing the net worth excluding intangible assets.	Double	TRUE
long_term_liab	Long-term liabilities, representing debts or obligations due beyond one year.	Double	TRUE
long_term_credit	Long-term credit represents credit obtained for long-term purposes.	Double	TRUE
short_term_liab	Short-term liabilities, representing debts or obligations due within one year.	Double	TRUE
CR_Rating_score	Credit Rating Score, indicating the creditworthiness of the customer.	Integer	TRUE
pmt_discipline_score	Payment Discipline Score, indicating the payment behavior of the customer.	Integer	TRUE
debt_equity_ratio	Debt to Equity Ratio, representing the proportion of debt to equity in the customer's capital structure.	Double	TRUE
debt_asset_ratio	Debt to Asset Ratio, representing the proportion of debt to total assets.	Double	TRUE
current_ratio	The current ratio, representing the ratio of current assets to current liabilities, indicates short-term liquidity.	Double	TRUE
return_on_assets	Return on Assets, representing the profitability of the customer's assets.	Double	TRUE
target_variable	The Target Variable used for predictive modeling or analysis could be categorical or numerical.	Integer	TRUE