# Predicting Customer Default Payments with Key Features Using Data Mining Techniques


## Abstract

Credit card default prediction poses a significant challenge for the financial industry, affecting lenders and consumers alike. This study aims to tackle this challenge by utilizing a dataset from Taiwanese credit card users to explore the predictive power of various machine learning models. Employing Recursive Feature Elimination with Cross-Validation, we identified key predictors of default and assessed the performance of seven different models. Our findings reveal that feature selection can substantially improve model accuracy, with Random Forest models showing the most significant improvement. This research provides valuable insights into effective predictors of credit card default, offering a solid foundation for financial institutions to enhance their risk management protocols. Despite encountering data quality issues, the methodologies applied herein demonstrate the robust potential of machine learning in financial risk analysis.



***Keywords***: Machine Learning, Recursive Feature Elimination, Cross-Validation, Random Forest, Data Quality


## Introduction

Credit card default is a pervasive issue in the financial industry. The inability of individuals to meet their credit card payment obligations not only impacts their financial well-being but also poses substantial risks to lenders in terms of revenue loss and increased credit risk. The Federal Reserve Bank of New York's Quarterly Report on Household Debt and Credit for Q3 2023 revealed a total household debt increase of $228 billion (1.3%), reaching $17.29 trillion, with 3% of outstanding debt in some stage of delinquency by the end of September. Credit card delinquencies were particularly high among borrowers aged 30-39. Despite this, less than 1% of aggregate student debt was 90+ days delinquent due to a policy delaying the reporting of missed payments until Q4 2024 [27]. These statistics underscore the critical need for effective prediction models to mitigate these risks.

Understanding the factors contributing to default behavior and developing effective prediction models is paramount in risk management and decision-making processes within the financial sector. Accurately predicting customer default payments enables financial institutions to implement proactive risk mitigation strategies, such as adjusting credit limits, offering financial counseling, or initiating collection procedures. For borrowers, the ability to predict default behavior can lead to more informed financial decisions, potentially averting financial distress and improving overall financial health.

Our research aims to address this pressing issue by leveraging advanced data mining techniques to develop predictive models for credit card default payments. By analyzing a comprehensive dataset of credit card transactions in Taiwan [31], which contains information on customer default payments, we aim to identify key predictors of default behavior and evaluate the performance of various predictive modeling approaches. Through this study, we strive to provide valuable insights and actionable recommendations for financial institutions, policymakers, and individuals alike, ultimately contributing to more robust risk management practices and fostering financial stability in the credit card industry, specifically in Taiwan.

## Background

Extensive research has been conducted on predicting credit card defaults. Many researchers have utilized datasets from the same sources [32], [28] as those used in our study, applying various data mining techniques.

Yeh and Lien [32] proposed six data mining techniques such as K-nearest neighbor (KNN), Logistic Regression (LR), discriminant analysis (DA), Naïve Bayes (NB), artificial neural networks (ANNs), and

classification trees (CTs) and used the same dataset as our project. ANNs showed the highest accuracy, especially in validation data.

Venkatesh and Gracia [28] also used the same dataset as our project. They assessed algorithms like BayesNet, Meta-Stacking, Naïve Bayes, Random Forest, SMO, and ZeroR. Feature selection methods, such as Correlation Feature Subset and Information Gain, significantly improved prediction accuracy, with Random Forest being the most effective.

Ma, Sha, Wang, Yu, Yang, and Niu [18] used LightGBM and XGBoost algorithms to predict customer defaults based on real-life peer-to-peer (P2P) transactions from the Lending Club. LightGBM outperformed XGBoost with an error rate of 19.9% and an accuracy of 80.1%.

Kvamme, Sellereite, Aas, and Sjursen [17] implemented Convolutional Neural Networks (CNN) to predict mortgage defaults using time series data related to customer transactions in current accounts, savings accounts, and credit cards. CNN showed promising results with an AUC of 0.918, which improved to 0.926 when combined with a Random Forest classifier.

Koutanaei, Sajedi, and Khanbabaei [15] proposed a hybrid credit scoring model, testing four feature selection algorithms and ensemble learning classifiers. Principal Component Analysis (PCA) was the best feature selection method, and ANN-AdaBoost was the best classification model.

Kruppa, Schwarz, Arminger, and Ziegler [16] used machine learning methods to estimate the probability of default rather than binary classification. They compared nonparametric regression-based approaches (random forests (RF), k-nearest neighbors (KNN), and bagged k-nearest neighbors (bNN)) with the parametric standard method of logistic regression. Random forests outperformed the other methods regarding AUC scores on the test data.

Bequé and Lessmann [3] introduced a type of neural network called Extreme Learning Machine (ELM). They compared its performance with other methods such as artificial neural networks, decision trees, support vector machines, and regularized logistic regression. They argue that this new approach combines significant prediction performance with noticeable computational efficiency.

Harris [10] conducted a study on predicting credit risk using a support vector machine algorithm applied to two definitions of default: a broader rule for up to 90 days overdue payments and a narrower definition for customers with more than 90 days late payments. He claims that the model used for the broader definition has higher accuracy than the narrower one and is a reliable and accurate method for predicting credit unworthiness compared to the traditional judgment approach.

Khandani, Kim, and Lo [13] proposed combining standard credit scoring features, such as the debt-to-income ratio, with more detailed characteristics like consumer banking transactions as input for the model, arguing that the latter significantly increases predictive power. In contrast, Khashman [14] introduced a novel approach to predicting credit risk for application scoring using an emotional neural network. This model incorporates anxiety and confidence during the learning process, simulating decisions made by a human expert. Khashman [14] concluded that the emotional neural network outperformed conventional neural networks in terms of speed, simplicity, accuracy, and minimum error.

# Methodology

### Dataset

The UCI Machine Learning Repository hosts the dataset for our project on customer default payments in Taiwan [31]. It has 30,000 instances with 24 features, as described in Table 1. The dataset contains integer and binary types without any missing values.

| Feature ID | Feature Name | Feature Description |
|---|---|---|
| X1 | LIMIT_BAL | Amount of the given credit (NT dollar) |
| X2 | SEX | Gender (1 = male; 2 = female) |
| X3 | EDUCATION | Education (1 = graduate school; 2 = university; 3 =high school; 4 = others) |
| X4 | MARRIAGE | Marital status (1 = married; 2 = single; 3 = others) |
| X5 | AGE | Age (year) |
| X6–X11 | PAY_0, PAY_2 to PAY_6 | History of past monthly payment records (from April to September 2005):<br>X6 = the repayment status in September...<br>X11 = the repayment status in April 2005,<br>The measurement scale for the repayment status is -1 = pay duly; 1= payment delay for one month; 2 = payment delay for two months; ...; 8 = payment delay for eight months; 9 = payment delay for nine months and above |
| X12–X17 | BILL_AMT1 to BILL_AMT6 | Amount of bill statement (NT dollar)<br>X12 =amount of bill statement in September 2005...<br>X17 = amount of bill statement in April 2005 |
| X18–X23 | PAY_AMT1 to PAY_AMT6 | Amount of previous payment (NT dollar)<br>X18 =amount paid in September 2005…<br>X23 = amount paid in April, 2005 |
| X24 | default payment next month | Default (1) and Non-Default (0) |

**Table 1:** Dataset Description

**Data Analysis**

In our project, we used Google Colab and Python, primarily with the Scikit-learn (SkLearn) library. Google Colab is a free cloud service that supports Python and offers free GPU access [21]. SkLearn is a powerful Python module for machine learning, providing tools for model fitting, data preprocessing, model selection, and evaluation [8]. Together, they offer a robust environment for developing and testing machine learning models.

Figure 1 below illustrates the steps taken in our analysis, inspired by these papers [1], [5], [19], [22], [29]. First, we imported the datasets into Google Colab for data inspection, exploration, and preprocessing. We omitted the first row of the dataset to simplify further analysis.

Data inspection helps us understand the nature and quality of the data. By exploring the dataset, we gain initial insights by viewing the top 5 and last 5 rows. Essential information includes data types, missing values per column, total rows per column, data duplication, and a statistical summary. Our dataset has 30,000 rows and 25 features, with no missing values or duplications.

Data visualization and exploration are critical steps in the data analysis process because they allow us better to understand the underlying patterns and trends within the dataset. Importantly, they make complex information more accessible and interpretable, uncovering insights that might not be immediately apparent from raw data alone.

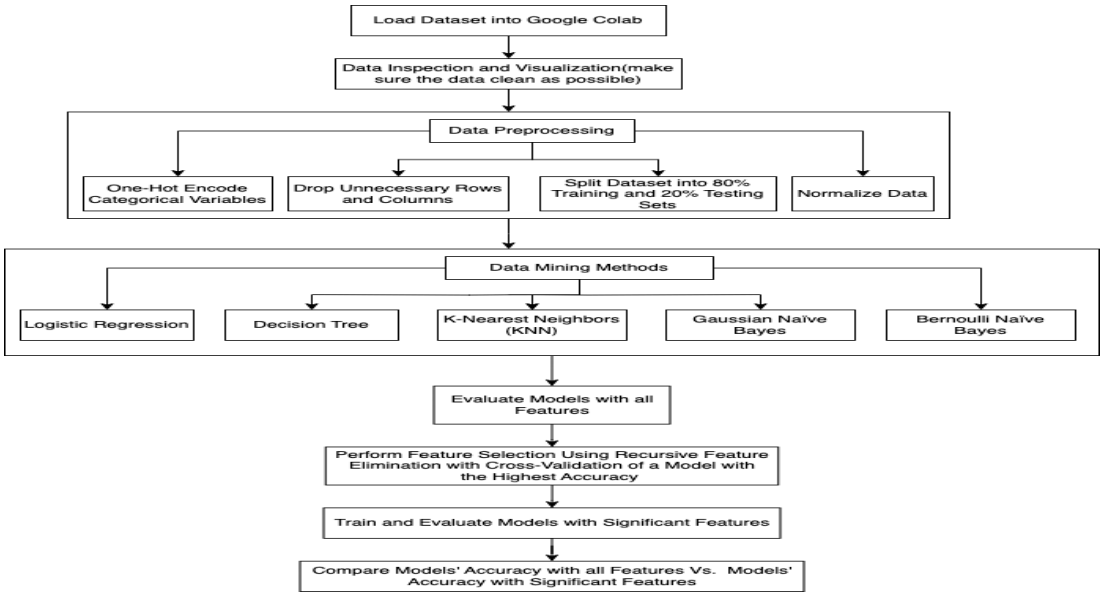
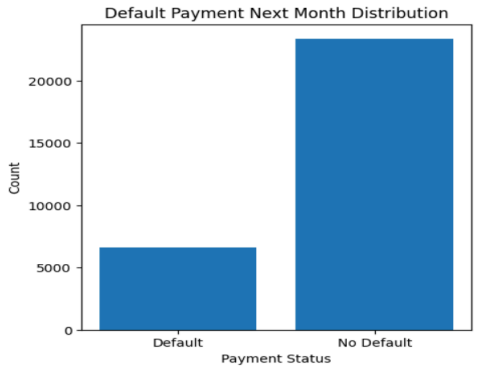



**Figure 1:** Analysis Steps

Data visualization helps us understand the distribution and anomalies within our dataset. The outcome variable is "default payment next month," and Figure 2 highlights a class imbalance with 6,636 defaults and 23,364 non-defaults.



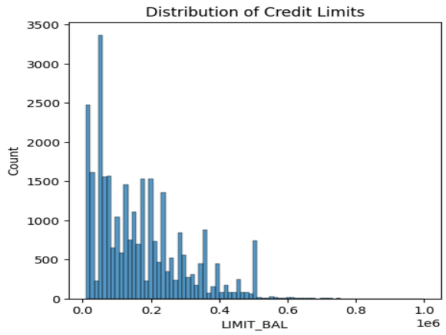
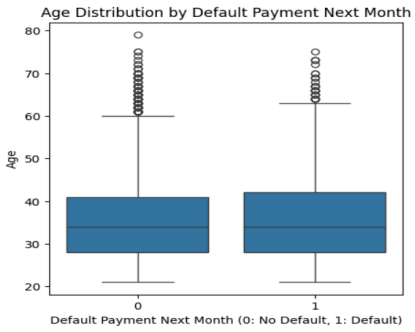




**Figure 2:** Default Payment Next Month Distribution

**Figure 3:** Distribution of Credit Limits

Figure 3 presents a histogram of the credit limits within our dataset, indicating a right-skewed distribution with the most frequent credit limits on the lower end of the spectrum. This suggests a large number of accounts with lower credit limits and fewer accounts with higher credit limits.

Figure 4 displays a box plot of the age distribution for non-defaulters (0) and defaulters (1). Non-defaulters have a median age in the mid-30s, an IQR from the late 20s to late 40s, and outliers up to the early 60s. Defaulters also have a median age in the mid-30s, an IQR from the late 20s to early 50s, and outliers up to around 80.



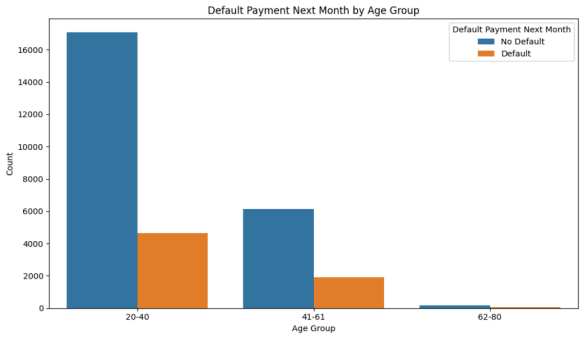




**Figure 4:** Age Distribution by Default Payment Next Month

**Figure 5:** Default Payment Next Month by Age Group

Figure 5 illustrates the count of individuals by age group who either defaulted (orange) or did not default (blue) on their credit card payment the next month. The 20-40 age group has the highest count, with non-defaulters significantly outnumbering defaulters. In the 41-61 age group, both counts decrease, but non-defaulters still predominate. The 62-80 age group has the lowest counts, indicating fewer credit card users and defaults. Overall, younger individuals (20-40) form the majority of users and defaulters, but non-defaulters surpass defaulters in all age groups.

The next major step is data preprocessing, which is crucial in the data analysis process, especially when using machine learning algorithms. This involves several steps. First, we one-hot encoded categorical variables like "SEX", "EDUCATION", "MARRIAGE", "PAY_0", "PAY_2", "PAY_3", "PAY_4", "PAY_5", and "PAY_6" to make them suitable for ML algorithms [20]. Next, we dropped unnecessary rows and columns, removing those that didn't align with our variables or contained negative values. We then split the dataset into training (80%) and testing (20%) sets using a random state of 42, allowing us to evaluate the model's performance on unseen data. Finally, we normalized the data to adjust the scale of numeric

features, ensuring that all features are treated equally by the algorithm, which enhances model performance and convergence [11].

Data mining methods are powerful techniques that enhance organizational decision-making through in-depth data analysis, serving both descriptive and predictive purposes via machine learning algorithms [30]. These techniques are crucial for fraud detection, user behavior analysis, identifying bottlenecks, and detecting security breaches [30].

We employed seven machine learning models in our project: Logistic Regression, Decision Tree, Random Forest, K-Nearest Neighbour (KNN), Support Vector Machine (SVM), Gaussian Naïve Bayes, and Bernoulli Naïve Bayes. Logistic Regression models the probability of an event's occurrence as a linear combination of explanatory variables, but it can struggle with non-linear interactions [25]. Decision Trees are effective for classification and regression tasks, visualizing decisions as a tree diagram [6]. K-Nearest Neighbour (KNN) predicts the label of a new point based on the closest training samples and is simple to implement for both classification and regression [12]. Random Forest is a meta-estimator that fits multiple decision tree classifiers on various sub-samples and uses averaging to improve accuracy and control overfitting [23]. Support Vector Machine (SVM) finds the best boundary that divides data into classes by identifying the hyperplane with the maximum margin, working well with both linear and non-linear data [26]. Naïve Bayes classifiers use Bayes' theorem with strong independence assumptions. Gaussian Naïve Bayes assumes continuous features follow a normal distribution [7], while Bernoulli Naïve Bayes is used for binary features [4].

After preprocessing, we applied all features to the seven machine learning models to find the highest accuracy. Using Recursive Feature Elimination with Cross-Validation (RFECV), we identified the most significant features by iteratively discarding the least important ones [24]. Cross-validation further assessed predictive performance [24].

With significant features identified, we removed insignificant ones, trained models with the refined feature set, and compared their accuracy to models using all features.

# Result

Below, we compare the accuracy rate of all features after data preprocessing to that of the significant features identified using Recursive Feature Elimination with Cross-Validation (RFECV) with a Logistic Regression model. We identified 11 key features, as shown in Figure 6 and Table 2.

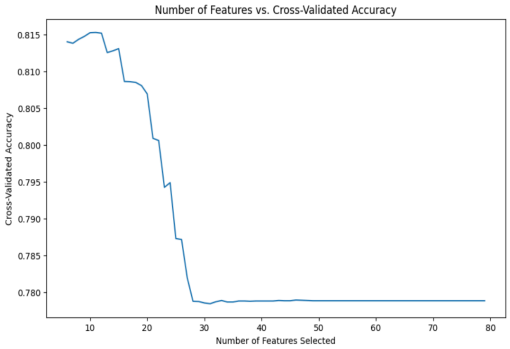



**Figure 6:** Number of Key Features

| Significant Feature Name | Feature Description |
|---|---|
| EDUCATION_1 | Graduate school |
| EDUCATION_2 | University |
| EDUCATION_3 | High school |
| PAY_0_-1 | Pay duly in September 2005 |
| PAY_0_2 | Payment delay for two months in September 2005 |
| PAY_0_3 | Payment delay for three months in September 2005 |
| PAY_2_-1 | Pay duly in August 2005 |
| PAY_3_2 | Payment delay for two months in July 2005 |
| PAY_5_2 | Payment delay for two months in May 2005 |
| PAY_6_2 | Payment delay for two months in April 2005 |

**Table 2:** Significant Features Description

Based on Table 3, Logistic Regression achieved the highest accuracy with all features at 81.87% and ranked third with significant features at 81.91%. The K-Nearest Neighbor (KNN) algorithm provided the second-highest accuracy with all features (81.76%) and with significant features (81.98%). The Random Forest algorithm ranked third with all features (80.69%) and achieved the highest accuracy with significant features (82.08%). Gaussian Naïve Bayes ranked third with all features (80.69%) and sixth with significant features (79.72%). Bernoulli Naïve Bayes was fifth with all features (79.19%) and fifth with significant features (81.33%). The Decision Tree had lower accuracy with all features (73.16%) but improved to fourth with significant features (81.33%). SVM had the lowest accuracy with both all features

(43.53%) and significant features (46.63%).

| Model | Accuracy with All Features (%) | Accuracy with Significant Features(%) |
|---|---|---|
| Logistic Regression | 81.87 | 81.91 |
| K-Nearest Neighbour (KNN) | 81.76 | 81.98 |
| Random Forest | 80.69 | 82.08 |
| Gaussian Naïve Bayes | 80.69 | 79.72 |
| Bernoulli Naïve Bayes | 79.19 | 81.33 |
| Decision Tree | 73.16 | 81.33 |
| Support Vector Machine (SVM) | 43.53 | 46.63 |

**Table 3:** Model Accuracy Comparison (All Features vs. Significant Features)

# Discussion

The adoption of significant features has notably enhanced the accuracy rates across almost all machine learning models utilized in our study. Specifically, the Random Forest model exhibited a remarkable improvement in accuracy, from 80.69% to 82.08%, when the number of leaf nodes was optimized from 24 to 16. The K-Nearest Neighbor (KNN) algorithm's accuracy improved to 81.98% with an optimized K value of 14, down from 20. Logistic Regression and the Decision Tree also saw improvements in their accuracy rates to 81.91% and 81.33%, respectively. Conversely, Gaussian Naïve Bayes experienced a slight decrease in accuracy, highlighting a potential area for further investigation. Bernoulli Naïve Bayes and the Support Vector Machine (SVM), however, showed improvements in their accuracy rates, demonstrating the benefit of feature selection across diverse algorithms.

Some limitations of this study relate to the quality of the data. In particular, it appears that missing values were filled with zeros. Some rows for the Amount of bill statement (NT dollar) and Amount of previous payment (NT dollar) contain negative values, and many columns on which we performed one-hot encoding contain extra values that do not align with the given data description. We treated negative values as mistakes, so we dropped the rows with negative values and also dropped columns of -2 and 0 for past monthly payment records.

# Conclusion

Our study leverages advanced data mining techniques to enhance the prediction models for credit card default payments. By applying Recursive Feature Elimination with Cross-Validation (RFECV), we successfully identified significant predictors of the default behavior, thereby improving the accuracy of our machine-learning models. This research not only contributes to the financial sector's risk management strategies but also offers a methodology for efficiently selecting features in predictive modeling. While the study is subject to certain data quality limitations, its findings underscore the importance of thorough data preparation and the potential for machine learning to refine risk assessment practices.

Future studies can explore the application of advanced machine learning algorithms, such as Heterogeneous Ensemble methods. Heterogeneous Ensemble methods, by integrating various machine learning algorithms, can leverage the strengths of each model and mitigate their weaknesses, potentially leading to more robust and accurate predictions [2], [9]. This approach can be particularly beneficial in further enhancing the financial sector's ability to manage credit risk effectively.

# Bibliography

[1] Aleksandrova, Y. 2021. “Comparing performance of machine learning algorithms for default risk prediction in peer to peer lending,” *TEM Journal*, pp. 133–143 (doi: 10.18421/tem101-16).

[2] Alshdaifat, E., Al-hassan, M., and Aloqaily, A. 2021. “Effective heterogeneous ensemble classification: An alternative approach for selecting base classifiers,” *ICT Express* (7:3), pp. 342–349 (doi: 10.1016/j.icte.2020.11.005).

[3] Bequé, A., and Lessmann, S. 2017. “Extreme Learning Machines for credit scoring: An empirical evaluation,” *Expert Systems with Applications* (86), pp. 42–53 (doi: 10.1016/j.eswa.2017.05.050).

[4] “Bernoulli naive Bayes.,” 2023a. *GeeksforGeeks*, GeeksforGeeks, October 25 (available at https://www.geeksforgeeks.org/bernoulli-naive-bayes/; retrieved July 8, 2024).

[5] Burkart, N., and Huber, M. F. 2021. “A survey on the explainability of supervised machine learning,” *Journal of Artificial Intelligence Research* (70), pp. 245–317 (doi: 10.1613/jair.1.12228).

[6] “Decision tree.,” 2024a. *GeeksforGeeks*, GeeksforGeeks, May 17 (available at https://www.geeksforgeeks.org/decision-tree/; retrieved July 8, 2024).

[7] “Gaussian naive Bayes.,” 2023b. *GeeksforGeeks*, GeeksforGeeks, November 13 (available at https://www.geeksforgeeks.org/gaussian-naive-bayes/; retrieved July 8, 2024).

[8] “Getting started.,” (n.d.). *scikit*, scikit-learn.org (available at https://scikit-learn.org/stable/getting_started.html; retrieved July 8, 2024 a).

[9] Gunakala, A., and Shahid, A. H. 2023. “A comparative study on performance of basic and ensemble classifiers with various datasets,” *Applied Computer Science* (19:1), pp. 107–132 (doi: 10.35784/acs-2023-08).

[10] Harris, T. 2013. “Quantitative credit risk assessment using support vector machines: Broad versus narrow default definitions,” *Expert Systems with Applications* (40:11), pp. 4404–4413 (doi: 10.1016/j.eswa.2013.01.044).

[11] Jaiswal, S. 2024. “What is normalization in machine learning? A comprehensive guide to data rescaling,” *DataCamp*, DataCamp, January 4 (available at https://www.datacamp.com/tutorial/normalization-in-machine-learning; retrieved July 8, 2024).

[12] “K-Nearest Neighbor(KNN) algorithm.,” 2024b. *GeeksforGeeks*, July 5 (available at https://www.geeksforgeeks.org/k-nearest-neighbours/; retrieved July 8, 2024).

[13] Khandani, A. E., Kim, A. J., and Lo, A. W. 2010. “Consumer credit-risk models via machine-learning algorithms,” *Journal of Banking & Finance* (34:11), pp. 2767–2787 (doi: 10.1016/j.jbankfin.2010.06.001).

[14] Khashman, A. 2011. “Credit risk evaluation using neural networks: Emotional versus conventional models,” *Applied Soft Computing* (11:8), pp. 5477–5484 (doi: 10.1016/j.asoc.2011.05.011).

[15] Koutanaei, F. N., Sajedi, H., and Khanbabaei, M. 2015. “A hybrid data mining model of feature selection algorithms and Ensemble Learning Classifiers for credit scoring,” *Journal of Retailing and Consumer Services* (27), pp. 11–23 (doi: 10.1016/j.jretconser.2015.07.003).

[16] Kruppa, J., Schwarz, A., Arminger, G., and Ziegler, A. 2013. “Consumer credit risk: Individual probability estimates using machine learning,” *Expert Systems with Applications* (40:13), pp. 5125–5131 (doi: 10.1016/j.eswa.2013.03.019).

[17] Kvamme, H., Sellereite, N., Aas, K., and Sjursen, S. 2018. “Predicting mortgage default using convolutional neural networks,” *Expert Systems with Applications* (102), pp. 207–217 (doi: 10.1016/j.eswa.2018.02.029).

[18] Ma, X., Sha, J., Wang, D., Yu, Y., Yang, Q., and Niu, X. 2018. “Study on a prediction of P2P network loan default based on the machine learning lightgbm and xgboost algorithms according to different high dimensional data cleaning,” *Electronic Commerce Research and Applications* (31), pp. 24–39 (doi: 10.1016/j.elerap.2018.08.002).

[19] Munkhdalai, L., Munkhdalai, T., Namsrai, O.-E., Lee, J. Y., and Ryu, K. H. 2019. “An empirical comparison of machine-learning methods on bank client credit assessments,” *Sustainability* (11:3), p. 699 (doi: 10.3390/su11030699).

[20] “One hot encoding in machine learning.,” 2024c. *GeeksforGeeks*, March 21 (available at https://www.geeksforgeeks.org/ml-one-hot-encoding/; retrieved July 8, 2024).

[21] Priya, B. C. 2022. “Google colab tutorial for data scientists,” *DataCamp*, DataCamp, February 4 (available at https://www.datacamp.com/tutorial/tutorial-google-colab-for-data-scientists?irclickid=SVLQthXe6xyPW5ZyPjWjHwiqUkHUI325pyh-Q00&irgwc=1&utm_medium=affiliate&utm_source=impact&utm_campaign=000000_1-2003851_2-mix_3-all_4-na_5-na_6-na_7-mp_8-affl-ip_9-na_10-bau_11-Bing%2BRebates%2Bby%2BMicrosoft&utm_content=BANNER&utm_term=EdgeBingFlow; retrieved July 8, 2024).

[22] Rahmani, R., Parola, M., and Cimino, M. 2024. “A machine learning workflow to address credit default prediction,” *Proceedings of the 26th International Conference on Enterprise Information Systems* (doi: 10.5220/0012640200003690).

[23] “Randomforestclassifier.,” (n.d.). *scikit* (available at https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html; retrieved July 8, 2024 b).

[24] “Recursive feature elimination with cross-validation in Scikit learn.,” 2023c. *GeeksforGeeks*, GeeksforGeeks, January 23 (available at https://www.geeksforgeeks.org/recursive-feature-elimination-with-cross-validation-in-scikit-learn/; retrieved July 8, 2024).

[25] Sperandei, S. 2014. “Understanding logistic regression analysis,” *Biochemia Medica*, pp. 12–18 (doi: 10.11613/bm.2014.003).

[26] “Support Vector Machine (SVM) algorithm.,” 2024d. *GeeksforGeeks*, July 4 (available at https://www.geeksforgeeks.org/support-vector-machine-algorithm/; retrieved July 8, 2024).

[27] “Total household debt reaches $17.29 trillion in Q3 2023; driven by mortgage, credit card, and student loan balances.,” (n.d.). *Total Household Debt Reaches $17.29 Trillion in Q3 2023; Driven by Mortgage, Credit Card, and Student Loan Balances - FEDERAL RESERVE BANK of NEW YORK* (available at https://www.newyorkfed.org/newsevents/news/research/2023/20231107; retrieved July 9, 2024).

[28] Venkatesh, A., and Gracia, S. 2016. “Prediction of credit-card defaulters: A Comparative Study on performance of classifiers,” *International Journal of Computer Applications* (145:7), pp. 36–41 (doi: 10.5120/ijca2016910702).

[29] Wang, Y., Zhang, Y., Lu, Y., and Yu, X. 2020. “A comparative assessment of Credit Risk Model based on machine learning ——a case study of Bank Loan Data,” *Procedia Computer Science* (174), pp. 141–149 (doi: 10.1016/j.procs.2020.06.069).

[30] “What is data mining?,” 2024. *IBM*, IBM, June 28 (available at https://www.ibm.com/topics/data-mining; retrieved July 8, 2024).

[31] Yeh, I.-C. 2016. “Default of credit card clients,” *UCI Machine Learning Repository*, UCI Machine Learning Repository, January 25 (available at https://archive.ics.uci.edu/dataset/350/default+of+credit+card+clients; retrieved July 8, 2024).

[32] Yeh, I.-C., and Lien, C. 2009. “The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients,” *Expert Systems with Applications* (36:2), pp. 2473–2480 (doi: 10.1016/j.eswa.2007.12.020).