# Microfinance Credit Prediction Using Machine Learning Techniques

**Agastya Nashier**

**Abstract: This research article will delve into the prediction of a defaulter in a microfinance credit industry using advanced machine learning techniques. The payment delinquency for a borrower is predicted using logistic regression, support vector classifier, random forest classifier, k-nearest neighbour and artificial neural network. The dataset is taken from an Indian microfinance company which contains an 18 feature set for this prediction derived from 4064 individuals living in different regions of India. The logistic regressor based classification technique performed slightly better than other techniques, reporting an accuracy of 87.94 %. The accuracy score of support vector classifier, random forest classifier and artificial neural network and K-nearest neighbours are 87.69 %, 87.82 %, 87.45 % and 87.69 % respectively. The models are further evaluated for their performance using confusion matrices, as well as receiver operating characteristic (ROC) curves and area under curve (AUC) values.**

**Nomenclature:**

ANN – Artificial Neural Network
MLP – Multi Layer Perceptron
DTI – Debt-to-income ratio
SVC – Support Vector Classifier
LR – Logistic Regression
KNN – K-Nearest Neighbor
ROC – Receiver Operating Characteristics
AUC – Area Under Curve
LR – Logistic Regression

**Keywords: Prediction, Classification, Feature Engineering, Machine Learning, Deep Learning, ANN, SVM, microfinance, random forest, logistic regression.**

## 1. Introduction

In the domain of microfinance it is difficult to predict the creditworthiness of an individual without previous credit data. In the past, many financial institutions have relied on the famous Five C's method of credit scoring which are namely, Character, Capacity, Capital, Collateral, and Conditions. Character refers to individual reputation, past track record in paying for debts and credit score (CIBIL score in Indian economy). Capacity is the borrower's income against debts which helps calculate an individual's debt-to-income (DTI) ratio. The DTI ratio helps in assessing a borrower's financial capability to repay a loan after covering existing expenses. Capital is the amount the borrower pays against a potential investment, such as a down payment for the purchase of an asset, which helps ensure the timely repayment of a loan at a given interest rate. Collaterals are assets that a borrower provides against a loan -- land, property, a bank savings deposit, etc. Finally, conditions, the internal and external financial conditions including the interest rate, market rates, etc. These conditions may or may not be under borrower's control.

From the research that has been going on for the past few decades there have emerged several indicators for making credit lending decisions [1]. Davis et. al., applied the machine learning technique for credit risk assessment and developed an integrated model -- a general computational model combined with an artificial neural network [2]. Although their dataset was small, their research proved that machine learning techniques could be efficiently applied in such scenarios. Another early work [3], developed an attribute selection metric to prevent non-monotonicity of the decision tree model without compromising the inductive accuracy. Galindo et. al., [4] developed a comparative analysis of the classification and regression decision tree (CART) along with an artificial neural network and KNN. They proved that the CART technique would be the best suited for risk prediction. Shi et. al., [5] used the datasets from Australian and German financial institutions and developed a novel-SVM and random forest technique which used the F1-score to infer the importance of features with given characteristics. Similarly [6, 22] SVM and RF techniques have been widely explored for credit risk prediction for financial institutions in the literature. Butaru et. al., [7] in their research showed that decision tree and random forest methods outperform the logistic regression technique, while Kruppa et. al., [8] compared RF, LR, KNN and bagged KNN. He found that RF is better than most algorithms for credit scoring.

In this research paper, we employ a similar machine-learning algorithm on Indian borrowers and estimate their performance in payment delinquency. This article is divided into 5 sections. We will first introduce the work done in the past and a problem statement. A theoretical explanation of different machine learning techniques that is used in this article is described in section 2. Section 3, covers the methodology that we adopted to approach the problem as well as feature selection, feature engineering, and parameter tuning and prediction metrics. Section 4 will discuss the results obtained from various models. Lastly, we will conclude this article with our main findings.

## 2. Machine Learning Techniques

Five machine learning classification techniques are used in this article namely logistic regression, k-nearest neighbour classifier, support vector classifier and random forest classifier, and artificial neural network or multi-layer perceptron. These techniques are explained in this section.

### 1. Logistic Regression

Logistic Regression is a technique used for binary classification of categorical features. This is the most common classification method based on the statistical approach of a discriminant analysis [3]. The technique comes with a regularization parameter C. Higher values of C lead to better fitting to the training set (i.e. less regularization). Low values of C places more emphasis on finding a coefficient vector closer to zero. Despite critical drawbacks, logistic regression is used due to its simplicity. It is also less computationally intense and time consuming.

### 2. K- Nearest Neighbor

K-Nearest neighbors technique (KNN) is one of the oldest, non-pragmatic machine learning techniques [10, 11] most prominently used for classification problems. The algorithm uses a large dataset for training with a label corresponding to each point in the dataset. The algorithm assumes that nearby points are of a similar nature, hence the term k-nearest neighbors. The classification algorithm assigns labels around the most observed k-nearest neighbors [12]. In this model, the similarity of two points depends upon the relative distance between the points [13]. Finally, the process is dependent on two critical independent processes, which form an adjacency matrix, which is constructed by estimating the edge's weights [14].

### 3. Support Vector Classifier

The main purpose of the support vector classifier technique (SVC) is to create a hyperplane that creates a division across a homogenous group of datapoints. The SVC separates the training set $(x_1, x_2, x_3, \dots x_n)$ which belong to the d-dimensional space into $y_i \in \{-1, 1\}$ which denotes different classes of the observation. The classes are separated by a hyperplane into a new feature space by a kernel function $K(x_i, x_j)$. The kernel function can be a linear function, radial basis function (RBF), polynomial function or a sigmoid function which is dependent on the problem set [16, 17, 18 ].

### 4. Random Forest Classifier

The main drawback of decision trees is that they tend to overfit the training data. Random forest solves this problem as it is basically a collection of decision trees forming a forest. Randomness in the random forest comes in as the trees are different from each thus preventing overfitting of the dataset and resulting in an improvement in overall accuracy [15]. It was proposed by Breiman in 2001 [19]. He showed prediction consistency using a simple version of random forest. The random forest classifier is a very powerful technique and does not require depth specification. It is often observed that little parameter tuning is required to get the highest efficiency in this model, making it fairly straightforward to implement.

### 5. Artificial Neural Network

The feed forward artificial neural network (ANN) is also known as multi-layer perceptron (MLP). The perceptron combines to form an ANN. Each input vector is fed into a hidden neural layer through a weighted matrix. The ANN comprises three sections -- an input layer, a hidden layer and an output layer. The neurons in the previous layer are connected to the neurons in the next layer. The training of the feed forward ANN goes as shown in figure 1.

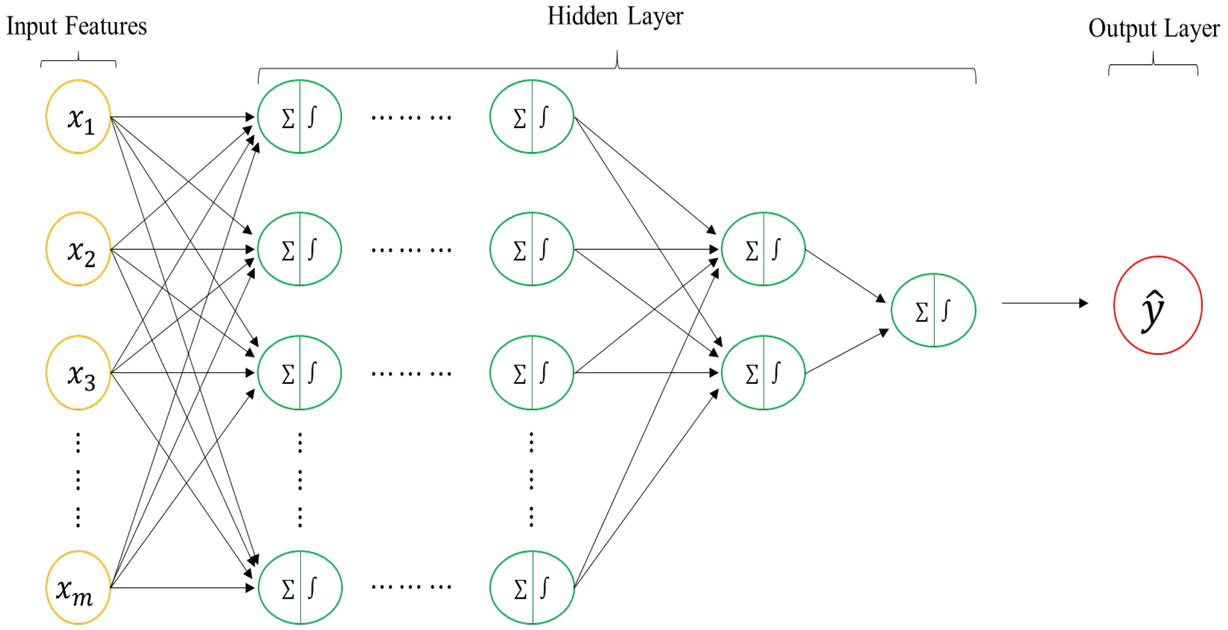

**Figure 1.** The feed forward artificial neural network.

**Table 1**. A detailed description of the dataset, features types and their respective values in the dataset.

| No. | Variable | Description | Values |
|---|---|---|---|
| **1** | Residence Status | Residence status of the borrower at the moment the loan is granted | Live with Parents |
| | | | Self |
| | | | No Value |
| **2** | Occupation Type | Evaluates whether the borrower is salaried or non-salaried | Salaried |
| | | | Self-Employed Non-Professional |
| | | | Self-Employed Professional |
| **3** | Industry | Borrower’s occupational industry | Aerospace, Agriculture, Apparels, Automobiles, Beverages, Capital Market, Chemicals, Construction, Communications, Education, Electronics and Equipment, Energy, Engineering, Financial Services, Food and Beverages, Furnishing, Gems & Jewellery, Healthcare, Industrial Equipment, Information & Technology, Media, Metals, Paper Products, Pharmaceuticals, Professional Services, Shipping, Speciality, Technology, Telecommunications, Textile, Transportation Logistics, Others. |
| **4** | Age | Borrower's age | value |
| **5** | Location | Borrower’s home address | Zipcode |
| **6** | Marital Status | Marital Status of the borrower | Single |
| | | | Married |
| **7** | Frequency | Frequency of the loan deposits | Monthly |
| | | | Structured |
| **8** | Gender | Gender of the individual borrower | Male |
| | | | Female |
| **9** | DSCR/DBR | Debt service coverage ratio / debt burden ratio | Values ranging between 0 and 1 |
| **10** | Purpose | Purpose for which the loan is taken | Acquisition of Assets |
| | | | Balance Transfer |
| | | | Business |
| | | | Family Celebrations |
| | | | Health and Wellness |
| | | | Home Construction |
| | | | New Vehicle Purchase |
| | | | Vacation |
| | | | Working Capital Requirement |
| | | | Multipurpose |
| | | | Others |
| **11** | Gross Credit | Amount borrowed by the borrower | values in rupees |
| **12** | Opening Principal Balance | First instalment amount paid by the borrower | value in rupees |
| **13** | Instalment amount | Amount left to be paid by the borrower | value in rupees |
| **14** | Principal Amount | Principal that the borrower has to pay as per the loan frequency | value in rupees |
| **15** | Interest Amount | Interest on the instalment amount that the borrower needs to pay. | value in rupees |
| **16** | Interest Rate | Interest rate on the loan amount | (in %) |
| **17** | CIBIL Score | Credit score of the borrower from the financial institution | Value of the score |
| **18** | Defaulter | Evaluates whether the borrower has defaulted on their loan | yes / No |

## 3. Experimental Methodology

Motivated by the preceding literature, we evaluated a wide range of machine learning algorithms (ANN, Logistic Regression, Support Vector, Random Forest and K-Nearest Neighbor) in our work on credit risk in micro-lending.

### 3.1. Dataset

The data is collected from an Indian firm for people located in the states of New Delhi, Karnataka, Maharashtra, Tamil Nadu, Telangana and Gujarat. The dataset includes people who are either working in a formal profession or running a business. Loan interest rate varies from 11% up to 28%. Loans are taken for a diverse range of reasons including healthcare expenditure, family celebrations, investment, business expansion, vacation, vehicle purchase, working capital requirement etc. Borrowers belong to industries such as, healthcare, capital goods, financial services, communication and technology, transportation logistics, energy, media, agriculture, apparel, automobile, aerospace, chemical, capital markets, construction, education, electronics and equipment, engineering, textiles, etc. The borrowed amount varies from a few hundred rupees to 31 million rupees. A detailed set of features and their respective possible values are mentioned in table 1.

For each borrower we have 5 features describing their personal character (residential, age, marital, gender, location), 4 features describing their earning and spending behaviour (Occupation, industry, DSCR/DBR, CIBIL), and 8 features describing their loan transactional behaviour. The borrower account is defined as a defaulter if they fail to keep-up with repayment instalments for more than 30 days as defined by their payment frequency (monthly, structured). In total, we have 18 variables defining a borrower's behaviour in order to make a fair prediction of payment delinquency after taking a loan.

### 3.2. Data Analysis

Data analysis is carried out to understand the distribution of the data and its impact on payment delinquency. Further, getting a general consensus would help us to identify the correct features that can be fed into the machine learning algorithm. From our analysis we can see that the dataset is imbalanced as only 21% were defaulters and 25% belong to the service industry. It is also observed the majority of defaulters come from high interest rate groups with interest rate varying between 19% to 28%. It appears that people taking loans with high interest rates have a high probability of defaulting. Also, it is observed that people with formal jobs have lower chances of payment delinquency when compared to people with no formal jobs.

### 3.3. Feature Selection

The original data had more features than the ones used in this article. Some features were of little or no importance such as account number. Few other variables such as applicant type (values: individual, non-individual), number of dependents, loan status, company's office branch either had NAN values or empty cells. Another factor that contributed to the selection of features was their usefulness in algorithms. Features such as loan status and branch name did not provide any information regarding the payment delinquency of the applicant. For example, the branch from which the loan was borrowed did not make much difference for the machine learning technique. Instead, the applicant's own location (zip code) was a more important feature for consideration. We chose customers' personal information, behavioural information and transactional information as a key feature for making our machine learning predictions.

### 3.4. Feature Engineering

For all the machine learning models (LR, SVC, RFC, KNN and ANN), post feature selection, feature engineering is carried out (see figure 2) to enable categorical inputs for the model to make the best possible predictions. As the dataset contains multiple integers and numerical values ranging from 0 to 1 million, it is necessary to scale the dataset. Before the dataset can be scaled the integers are encoded into 0's and 1's using a label encoder and one hot encoding method.

Label encoding generates a numerical value for each label in the column. The values can range from 0 to $n-1$ (where $n$ is the number of labels in the particular feature). In this case, to prevent any kind of hierarchy or order in the numeric values we applied label encoding on the feature set which had binary labels, as this will only generate 0's and 1's in response. Such encoding is done on features such as marital status, frequency and gender. For features having more than 2 labels the label is encoded using the 'one hot encoding' method. In this technique, each categorical value of a label is converted into a new column and 1 or 0 is assigned to the value of that column. This eliminates hierarchy issues in the feature set, although this expands the number of columns which depend on the number of labels in the feature.

For all the models fitted in this article, the dataset is split into an 80:20 ratio. 80% of the data belongs to the training set and 20% data to the test set. The testing set will be used to generate prediction accuracy or confidence estimate on the performance of the tuned model. It is ensured through stratified distribution technique that the training set and test set data contain equal number of classes (delinquent vs legitimate). The dataset is divided into training set and test set using 'train_test_split' method. Post splitting the dataset, the training set is fitted with the 'standard scaler' method which will scale the dataset into machine optimised values for easier training and better prediction.

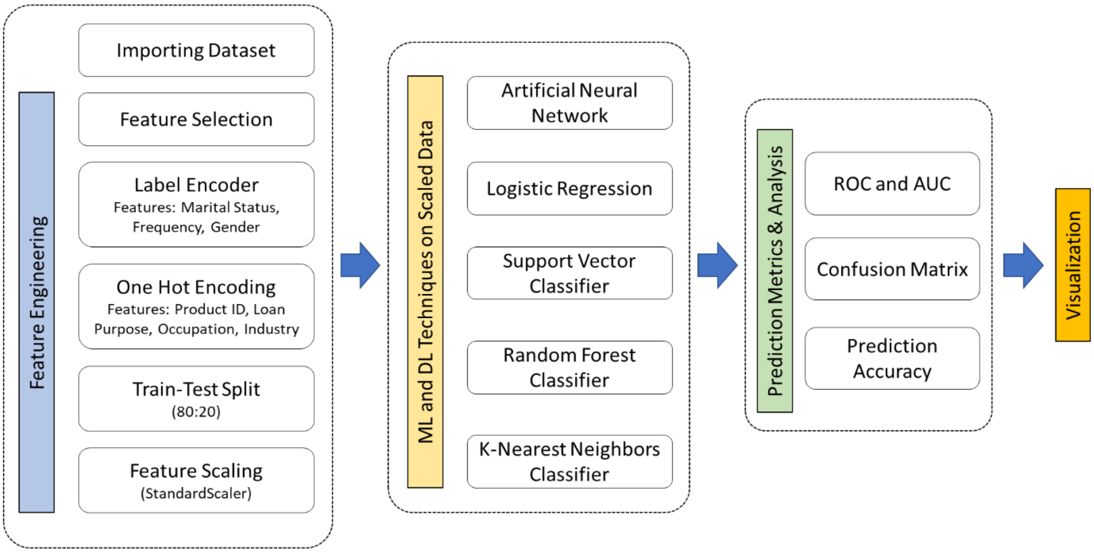
Feature Engineering
Importing Dataset
Feature Selection
Label Encoder
Features: Marital Status, Frequency, Gender
One Hot Encoding
Features: Product ID, Loan Purpose, Occupation, Industry
Train-Test Split
(80:20)
Feature Scaling
(StandardScaler)
ML and DL Techniques on Scaled Data
Artificial Neural Network
Logistic Regression
Support Vector Classifier
Random Forest Classifier
K-Nearest Neighbors Classifier
Prediction Metrics & Analysis
ROC and AUC
Confusion Matrix
Prediction Accuracy
Visualization

**Figure 2.** The process steps used to obtain the prediction metrics and its analysis.

### 3.5. Parameter Tuning

The ANN consists of 4 hidden layers along with one input layer and one output layer. A rectified linear unit is used as the activation function. Each hidden layer contains 20 nodes. An adaptive moment estimator is used to optimize the gradient descent. The batch size is 64. No accuracy improvement is observed after 150 epochs. It took approximately 7.5 sec to finish the compilation of the model. We also tried other activation functions, optimizers and loss functions to tune the model for best performance but the above-mentioned model parameters performed the best for the selected dataset.

The support vector classifier is imported from the sklearn library. The radial basis function was used as a kernel with C = 0.1, the value which gave us the best accuracy.

Logistic regressor classifier is imported from the linear_model method of sklearn library. The model was trained on the training set with C = 100 which gave us the best fitting model with the highest accuracy. It took significantly less time to compile this model.

Random forest classifier is an ensemble technique borrowed again from sklearn library. The trees that are built in the random forest are stored in the estimator attribute. n_estimator was set at 60 – the value which provided the best prediction.

KNN classifier is fairly easy to implement as well. This technique uses two main parameters, namely the number of neighbors and the metrics to measure the distance between the two points. In our case, we achieved good performance with 20 neighbors using Minkowski distance, a metric for real-valued vector space.

### 3.6. Prediction Metrics

Standard metrics are used to analyse the performance of the prediction classification models [20, 21, 22, 23, 24, 25]. Confusion matrix is a popular method in machine learning for performance measurement of a classifier. It is a matrix which compares the actual target values with the predicted target values. For a binary classification we have 2 x 2 matrix which give out 4 values depicting true positive (TP), false positive (FP), false negative (FN) and true negative (TN). The accuracy of a model is given by equation 1.

$$Accuracy = \frac{True\ Positive+True\ Negative}{True\ Positive+False\ Positive+True\ Negative+False\ Negative} \quad (1)$$

Additionally, Precision describes the number of correct predictions that were actually positive and the Recall informs us of the number of positive cases we were able to correctly predict with our model. Precision and recall are given by equation 2 and 3 respectively.

$$Precision = \frac{True\ Positive}{True\ Positive+False\ Positive} \quad (2)$$

$$Recall = \frac{True\ Positive}{True\ Positive+False\ Negative} \quad (3)$$

F1 Score, is the harmonic mean of the precision and the recall. F1 Score is highest when precision and recall are equal.

Another metric that we are using besides the confusion matrix is the receiver operating characteristic curve, popularly known as ROC curve and the area under the curve value which is popularly known as AUC value. The ROC curve is plotted on two parameters, true positive rate (TPR) vs false positive rate (FPR) where the values vary from 0 to 1. Ideally, the TPR should be 1 for all values of FPR for a classifier to be known as a perfect classifier. ROC curve analysis provides us with a tool to select the best possible model and discard suboptimal. Models having higher AUC values suggest highly accurate models with optimal fitting (i.e. no overfitting or underfitting).

## 4. Results And Discussion

**Table 2.** The predictive performance of the various model.

| | Artificial Neural Network | Support Vector Machine | Logistic Regression | Random Forest Classifier | K-Nearest Neighbor Classifier |
|---|---|---|---|---|---|
| True Positive (TP) | 659 | 672 | 664 | 662 | 662 |
| False Positive (FP) | 19 | 6 | 14 | 16 | 16 |
| False Negative (FN) | 83 | 94 | 84 | 83 | 83 |
| True Negative (TN) | 52 | 41 | 51 | 52 | 52 |
| Precision | 0.972 | 0.991 | 0.979 | 0.976 | 0.976 |
| Recall | 0.888 | 0.877 | 0.888 | 0.889 | 0.889 |
| F1 Score | 0.928 | 0.931 | 0.931 | 0.930 | 0.930 |
| AUC | 0.789 | 0.787 | 0.816 | 0.794 | 0.738 |
| Accuracy | 0.875 | 0.877 | 0.879 | 0.878 | 0.878 |
| Time (s) | 7.94 | 1.7 | 1.07 | 1.44 | 5.3 |

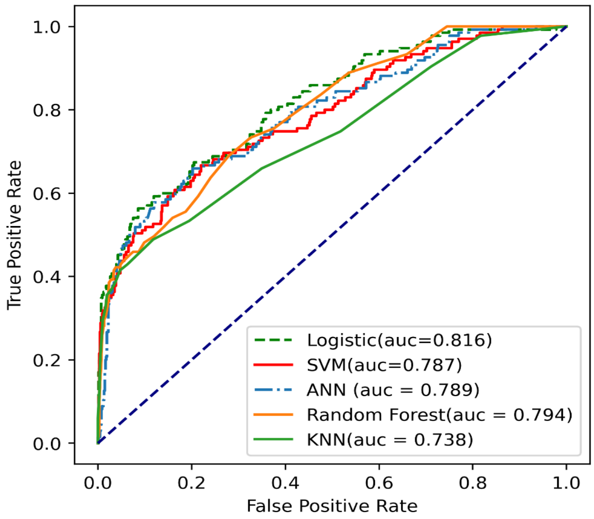

**Figure 3.** The ROC curve for LR, SVC, ANN, RFC and KNN along with their respective AUC values.

The model performance is listed in table 2. We can observe that SVC shows the lowest Type I error, which means that the model performs better for borrowers with low chances of payment delinquencies while for borrowers having high chances of being a defaulter, other models perform better. All models (except SVC) have almost equal prediction for type II errors. However, from the ROC curve (show in figure 3.) and AUC values we see that logistic regression performed the best: it was able to correctly predict whether the borrower will be a defaulter or not. From the ROC curve we can observe that the TPR increases exponentially for FPR upto 0.1. The real difference across models is observed between 0.1 and 0.8. This is where we see LR perform the best. While generally it is expected that a neural network would perform the best because of its efficient learning rate and robustness, this assumption did not hold true in our case.

Also, training the LR took the least time of 1.07 sec compared to other models. The ANN took the maximum time of 7.94 sec followed by the KNN of 5.3 sec, showing the complexity of the models and the number of features set. In our dataset we only trained on 4064 rows. The ANN might perform better with a larger dataset. Also, the prediction accuracy of 87.9% is highest with logistic regression. Other models achieved similar accuracy of above 87%.

## 5. Conclusions

In conclusion, in this article we developed a comprehensive list of models that were trained on a medium scale microfinance-based borrowers dataset. The dataset consisted of details belonging to 4064 individuals. Because of complex labels and multiple features, we have to rely heavily on feature engineering to create a sparse matrix which can be efficiently used as input data for our model. Post splitting the dataset into 80:20 ratio five different machine learning models are trained and model performance is estimated using the test data. Confusion matrices and ROC curves are used as performance metrics for the models. It is shown that logistic regression performed the best with prediction accuracy of 87.9%. ANN, SVC, RFC and KNN gave an accuracy of 87.45%, 87.70%, 87.82% and 87.82 % respectively.

Thus, our research suggests logistic regression as the best predictor for payment delinquencies -- although models by other techniques also show similar results. While on the surface, Logistic Regression may appear to be the simplest mode of analysis, it is important to note that it works robustly when applied to linear interactions. This is due to the fact that logistic regression is simply a special case of linear regression, albeit one that utilises binary response variables.

Then the accuracy of the logistic regression model must be understood along with the limitations of the model. Specifically, its inability to capture nonlinear and interactive effects of the features selected in this case. Further research on the subject is required to come to a conclusion regarding credit repayment delinquency and its ties to the selected features.

## References

1. Khandani, Amir E., Adlar J. Kim, and Andrew W. Lo. 2010. Consumer credit-risk models via machine-learning algorithms. Journal of Banking & Finance 34: 2767–87.
2. Davis R, Edelman D, Gammerman A (1992) Machine-learning algorithms for credit-card applications. IMA J Manag Math 4(1):43–51.
3. Ben-David, A. Monotonicity maintenance in information-theoretic machine learning algorithms. Mach Learn 19, 29–43 (1995). https://doi.org/10.1007/BF00994659
4. Galindo, J., Tamayo, P. Credit Risk Assessment Using Statistical and Machine Learning: Basic Methodology and Risk Modeling Applications. Computational Economics 15, 107–143 (2000). https://doi.org/10.1023/A:1008699112516
5. Shi, J., Zhang, Sy. & Qiu, Lm. Credit scoring by feature-weighted support vector machines. J. Zhejiang Univ. - Sci. C 14, 197–204 (2013). https://doi.org/10.1631/jzus.C1200205.
6. Dimitrios Niklis, Michael Doumpos, and Constantin Zopounidis. 2014. Combining market and accounting-based models for credit scoring using a classification scheme based on support vector machines. Appl. Math. Comput. 234, C (May 2014), 69–81. DOI: https://doi.org/10.1016/j.amc.2014.02.028
7. Butaru, F., Chen, Q., Clark, B., Das, S., Lo, A. W., & Siddique, A. (2016). Risk and risk management in the credit card industry. Journal of Banking & Finance, 72, 218–239.
8. Kruppa, J., Schwarz, A., Arminger, G., & Ziegler, A. (2013). Consumer credit risk: individual probability estimates using machine learning. Expert Systems with Applications, 40(13), 5125–5131.
9. Allison, P.D. (ed.): Logistic regression using the SAS system: theory and application. SAS Institute, Stanford (2012)
10. Fix, Evelyn. 1951. Discriminatory Analysis: Nonparametric Discrimination, Consistency Properties. San Antonio: USAF School of Aviation Medicine.
11. Fix, Evelyn, and Joseph L. Hodges, Jr. 1952. Discriminatory Analysis-Nonparametric Discrimination: Small Sample Performance. Technical report. Berkeley: University of California, Berkeley
12. Neath, Ronald, Matthew Johnson, Eva Baker, Barry McGaw, and Penelope Peterson. 2010. Discrimination and classification. In International Encyclopedia of Education, 3rd ed. Edited By Baker Eva, McGaw Barry and Penelope Peterson, London: Elsevier Ltd., vol. 1, pp. 135–41.
13. Weinberger, Kilian Q., and Lawrence K. Saul. 2009. Distance metric learning for large margin nearest neighbor classification. Journal of Machine Learning Research 10: 207–44.
14. Dornaika, Fadi, Alirezah Bosaghzadeh, Houssam Salmane, and Yassine Ruichek. 2017. Object categorization using adaptive graph-based semi-supervised learning. In Handbook of Neural Computation. Amsterdam: Elsevier, pp. 167–79.
15. Geurts, Pierre, Damien Ernst, and Louis Wehenkel. 2006. Extremely randomized trees. Machine Learning 63: 3–42.
16. Moula FE, Guotai C, Abedin MZ (2017) Credit default prediction modeling: an application of support vector machine. Risk Manag 19(2):158–187.
17. Pławiak P, Abdar M, Pławiak J, Makarenkov V, Acharya UR (2020) DGHNL: a new deep genetic hierarchical network of learners for prediction of credit scoring. Inf Sci 516:401–418.
18. Zhong H, Miao C, Shen Z, Feng Y (2014) Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings. Neurocomputing 128:285–295
19. Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001). https://doi.org/10.1023/A:1010933404324.
20. Feng, X., Xiao, Z., Zhong, B. et al. Dynamic weighted ensemble classification for credit scoring using Markov Chain. Appl Intell 49, 555–568 (2019). https://doi.org/10.1007/s10489-018-1253-8
21. Li W, Ding S, Chen Y, Wang H, Yang S (2019) Transfer learning-based default prediction model for consumer credit in China. J Supercomputer 75(2):862–884
22. Luo C (2020) A comprehensive decision support approach for credit scoring. Ind Manag Data Syst 120(2):280–290.
23. Morales EA, Ramos BM, Aguirre JA, Sanchez DM (2019) Credit risk analysis model in microfinance institutions in Peru through the use of Bayesian networks. In: 2019 Congreso Internacional de Innovación y Tendencias en Ingenieria (CONIITI), IEEE, pp 1–4.
24. Moula FE, Guotai C, Abedin MZ (2017) Credit default prediction modeling: an application of support vector machine. Risk Management 19(2):158–187.
25. Siami M, Gholamian MR, Basiri J (2013) An application of locally linear model tree algorithm with combination of feature selection in credit scoring. Int J Syst. Sci 45(10):2213–2222