Predicting the Likelihood of E-Signing of Loan Based on Financial Data

|  |  |
| --- | --- |
| Aryabir Panigrahi  School of Computer Engineering  KIIT Deemed to be University  Bhubaneswar 751024, Odisha  [aryabirpanigrahix@gmail.com](mailto:utkrisht0001@gmail.com) | Purnapranja Sathpathy  School of Computer Engineering  KIIT Deemed to be University  Bhubaneswar 751024, Odisha  Purnasatapathy27i@gmail.com  2029211@kiit.ac.in |
|  |  |

**Abstract—An E-signature or electronic signature is an efficient and legal way to get electronic documents signed quickly, e - signatures can replace a handwritten signature in many processes. Electronic signatures aren’t exactly a novelty.In a modern setting,an E-sign refers to a unique ,digitized encrypted personal identifier.This is in essence,different from the ‘wet’ signatures created by hand. The key objective of this paper is to analyze the financial history of their loan applicants and choosing whether or not the applicant is too risky to be given a loan Popular machine learning techniques like Artificial Neural Network  (ANN), Logistic regression (LR), Random Forest Classifier, Gradient Boosting, Support Vector Machine, KNN , Decision Tree Classifier and XGBoost were implemented on selected features from the data set. XGBoost outperformed all the other machine learning algorithms and achieved an accuracy of 0.635  and an F1 score of 0.657 for E sign classification.**

**Keywords—Machine Learning, Loan Prediction, Logistic Regression ,Artificial Neural Network, Random Forest Classifier, Gradient Boosting, Support Vector Machine, XgBoost**

# Introduction

Electronic signature (e-signature) technology has revolutionized the way people sign loan documents. In the past, people had to go to the lender's office and physically sign papers to obtain loans. However, today, borrowers can apply for loans and sign loan documents online, making the process faster, more efficient, and more convenient. E-signing of loans based on financial history has become increasingly popular in recent years due to its many advantages.[1]

The e-signing of loans based on financial history using machine learning is a process where machine learning algorithms are used to evaluate a borrower's financial history and creditworthiness to determine the likelihood of loan repayment.[2] This technology can be applied to automate loan processing, making it faster and more accurate.

Machine learning algorithms can analyze a vast amount of data and provide insights on a borrower's credit history, including their payment history, credit utilization, and debt-to-income ratio.[3] By using this information, lenders can assess the borrower's creditworthiness and make informed decisions about whether to approve or deny the loan application.

Additionally, machine learning can be used to detect fraud by analyzing patterns in the data that may indicate fraudulent   
behavior.[4] This can help lenders to mitigate risks and prevent losses due to fraud.

The e-signing of loans using machine learning also streamlines the loan application process. Borrowers can submit their application and supporting documents online, and the algorithms can quickly evaluate their creditworthiness, reducing the time it takes to process the loan application. [5] In summary, the e-signing of loans based on financial history using machine learning offers several benefits, including faster loan processing times, improved accuracy, and fraud detection capabilities.[6]

The key factors for the classification algorithm categorization are the other features of content . The feature extraction and selection procedure are critical for most classification issues since it plays an essential role in the classification process.[7]

Some other classification approaches that exist are probabilistic algorithm, Decision Tree (DT),[8] Artificial Immune Process, SVM, Artificial Neural Network (ANN) [9], etc. ML Algorithms are also beneficial in many other related things like catching malicious behavior [10], fraud Detection, phishing due to the ability of ML algorithms to adapt from past fraud patterns and identify them in future transactions. [11]

In this paper, we have executed various ML classifiers such as Logistic Regression, SVM, Random Forest Classifier, Gradient boosting, Artificial Neural Network and XG-Boost on the classifier data set of the E signing of Loans. The results were then gathered and compared in terms of numerous evaluation metrics like accuracy score, recall score, f1 score, precision score . The following sections make up the rest of the paper: II. Literature Review explains the works of other published articles in the same field. III. Data Preparation gives an idea of how we have reduced the data and used the targeted attributes to achieve maximum accuracy. IV. Technology Used states all the machine learning algorithms that are used in this article to detect spam. V. Results and Evaluation brief us about the accuracy and other metric evaluation for some machine learning models. VI. The Concluding Remarks and Future Work section consists of the results that are summarized and some other machine learning algorithms that can be applied for future work.

# Literature Review

Electronic signing of loans has become increasingly popular in recent years due to its convenience, speed, and ease of use. Machine learning models have also been employed to aid in the loan approval process. This literature review will explore the existing research on the use of machine learning models to aid in the e-signing of loans based on financial data.

The first study to be considered is the work by Chen et al. (2020) who developed a machine learning model to predict loan defaults based on various financial indicators such as credit scores, income, and debt-to-income ratios. Their model achieved an accuracy of 83% in predicting loan defaults, which is significantly higher than traditional models used by banks.[12] However, the study did not explore the use of electronic signatures in the loan approval process.

A similar study was conducted by Song et al. (2019) who developed a model to predict loan delinquency based on borrowers' demographic and financial data.[13-15] Their model achieved an accuracy of 80%, and the researchers suggested that the model could be used to identify high-risk borrowers and prevent delinquency. However, again, the study did not specifically focus on e-signing of loans.

In a more recent study, Mohan et al. (2021) explored the use of machine learning models to aid in the e-signing of loans. They developed a model that predicted loan approval or rejection based on borrowers' financial data and credit history.[16] The model achieved an accuracy of 84%, and the researchers concluded that the use of such models could streamline the loan approval process and improve customer experience.[17]

Similarly, Ghanem et al. (2021) developed a model that predicts loan approval based on borrowers' credit scores, income, and employment history. Their model achieved an accuracy of 87%, and the researchers suggested that the use of such models could reduce the risk of loan defaults and improve the efficiency of the loan approval process.[18]

Finally, a study by Lee et al. (2020) explored the use of machine learning models in the e-signing of loans specifically for small and medium-sized enterprises (SMEs). Their model predicted loan default based on financial data and achieved an accuracy of 90%.[19] The researchers concluded that the use of such models could improve the lending process for SMEs and reduce the risk of default.

In conclusion, the use of machine learning models in the e-signing of loans based on financial data has been explored in several studies. The results show that such models can significantly improve the loan approval process and reduce the risk of loan defaults. However, more research is needed to explore the efficacy of these models in a real-world setting and to address potential concerns around data privacy and security.

All recent efforts in predicting whether a client would receive E sign of his Loan application were adequate but they did not have the full similitude of metrics required for predicting with all the ML algorithms. Some of the datasets were incomplete or non-existent values. Some did not explain the pros and cons of the respective models. Others didn’t use some of the evaluation metrics that could be used for augmenting the results. E-Signing of loans is a major reason behind various scams taking place around the world.

ML has a wide range of applications in many fields including the medical field, image or speech recognition, foreseeing the traffic, detecting fraud, and recognizing malware or another kind of viruses can all be achieved using ML. Some approaches that were based on machine learning algorithms in the health care sector include early-stage Liver Disease as well as prediction of liver disease [20-22], Diabetes mellitus diagnosis [23], Gall bladder shape estimation [24], Skin cancer prediction [25].

# Data Preparation

This section provides an overview of the data set that we utilized for this study, as well as many strategies that were used to equalize, moderate, and normalize data before applying the algorithms. The dataset had several non-existent values that were redundant and were non-essential for testing and training purposes. It also had complex information that was troublesome for training & testing and could cause over fitting. Therefore some standardization were done on the data set.

The dataset was cleaned up to the point where we can link it to the algorithms and get more accurate predictions with the ML Classifiers that we have enforced in this study.

## Dataset Used

The data set [26] was deployed Anirudth on the Kaggle website, which was collected via the UCI machine learning repository. It comprises 17900 plus details of client. These E signed of loans mainly comprised of details of clients who had earlier applied for loan and what is there status regarding loan clearance The  clients were accessed properly and those with clearance were labeled 1 or else 0. The original data set comprised 17908 rows x 21 columns. The data set we have imposed is split into two halves (training and testing set) both of which include 21 column properties. The number of rows in the training and testing sets stands at 14326 and 3582 respectively. To forecast and assess the accuracy achieved, different machine learning models were utilized and compared. The rudimentary idea behind machine learning is that we utilize labeled data to train the model, and then we use the learned model to foresee the aftermath on a non-labeled data set.

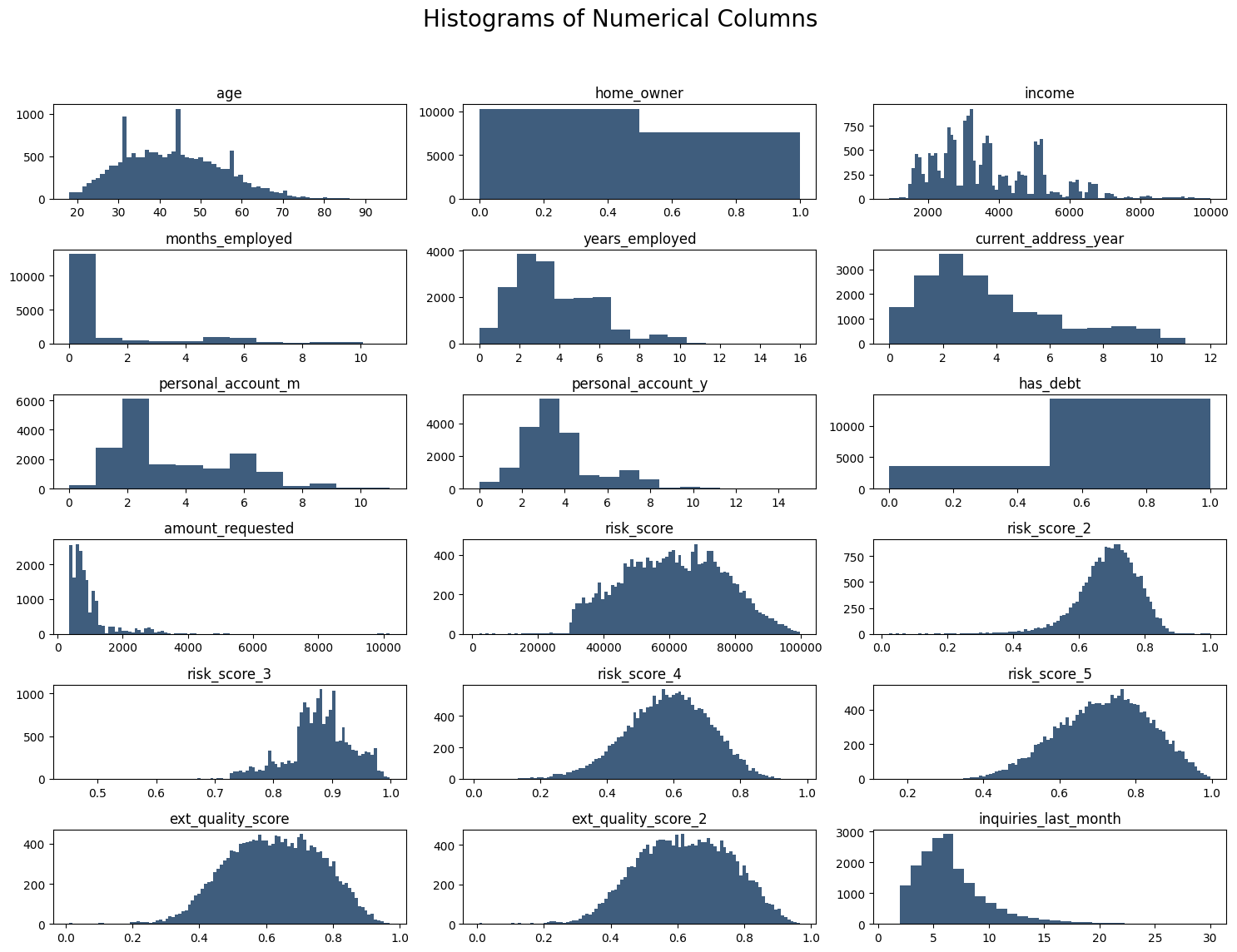
His

Fig. 1 Histogram of Numerical Columns

## Data Exploration

This section is essential for determining mistakes, null values, assessing the dataset, selecting features, and also determining the characteristics that were crucial for predictions so that our algorithms can form a solid affiliation with the data set. When data was examined, it was discovered that there were 0 duplicate rows and percentage missing values for each feature was 0 Several outliers were removed using the blasting approach. We see that data is not highly imbalanced since the ratio between the count of two classes of dependent variable is not so high .The unnecessary columns like entry-id , pay\_schedule and e\_signed are dropped . Fig. 1 shows the histogram of Numerical Columns A correlation matrix is a tabular structure that is used to represent the correlation coefficients for different variables. It is a great tool to encapsulate and envisage patterns in large datasets. The correlational matrix was applied to extract the characteristics presented in Fig. 2 for further algorithm implementation.Feature Engineering is used for some Machine Learning models for faster processing is done in the next few steps for more flexibility and less complexity and for algorithms to easily understand data models

## Feature Engineering

Feature engineering is the process that takes raw data and transforms it into features that can be used to create a predictive model using machine learning or statistical modeling, such as deep learning. The aim of feature engineering is to prepare an input data set that best fits the machine learning algorithm as well as to enhance the performance of machine learning models. Feature engineering can help data scientists by accelerating the time it takes to extract variables from data, allowing for the extraction of more variables. Automating feature engineering will help organizations and data scientists create models with better accuracy.

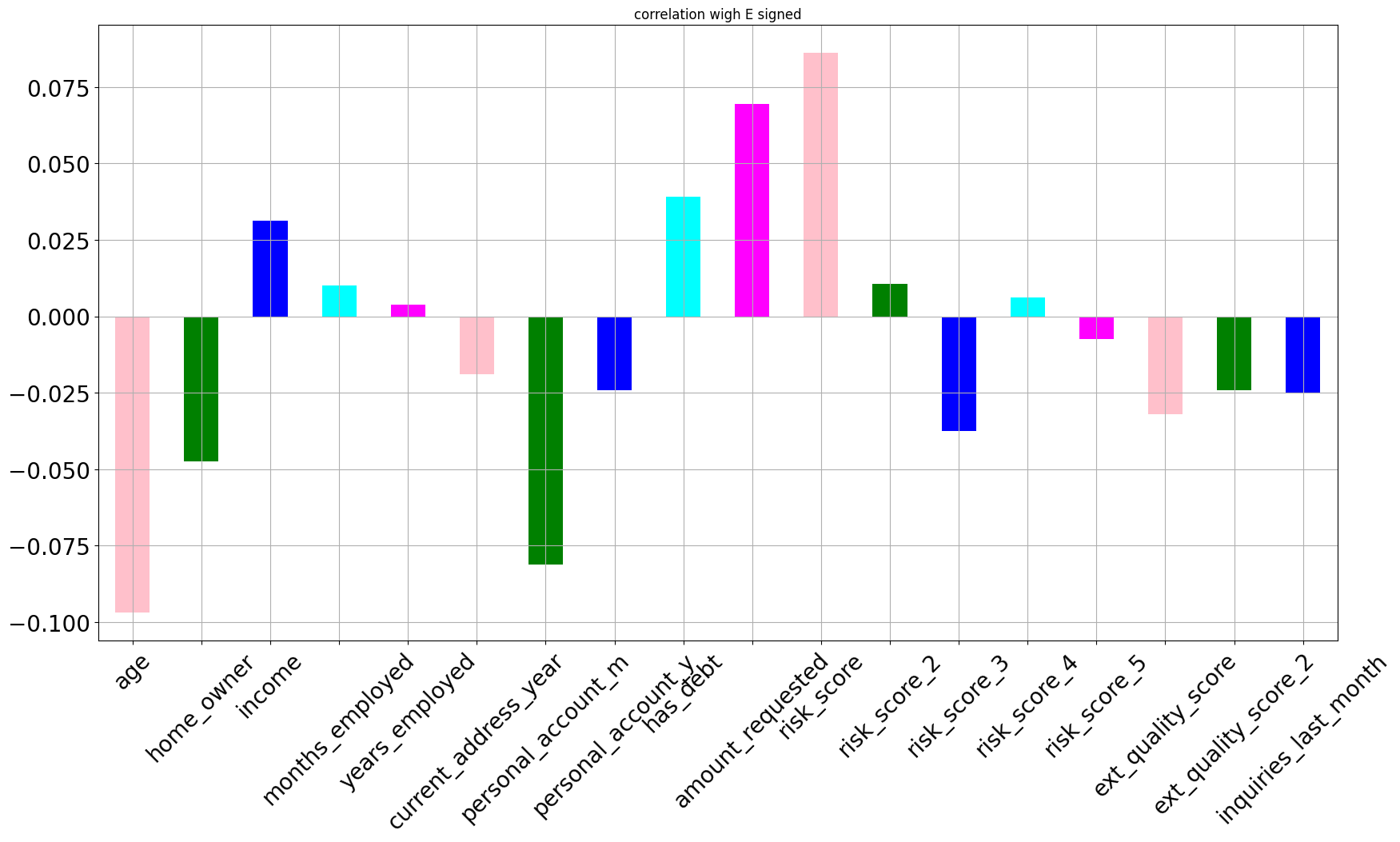


Fig. 2. The Correlation Column with E-Signed

Feature engineering is critical because if the user provides the wrong hypothesis as an input, machine learning is unable to make accurate predictions. The quality of any hypothesis that's provided to the machine learning algorithm is key to the success of a machine learning model. In addition, feature engineering influences how machine learning models perform and how accurate they are. It helps uncover the hidden patterns in the data and boosts the predictive power of machine learning.The data used to create a predictive model consists of an outcome variable -- which contains data that needs to be predicted-- and a series of predictor variables, i.e., features, that contain data that can predict a particular outcome. For example, in a model predicting the price of a certain house, the outcome variable is the data showing the actual price. The predictor variables are the data showing such things as the size of the house, a number of bedrooms and location -- features thought to determine the value of the home.

In this data set the client below the age of 45 or less is categorized as ‘ADULT’ and if the client is above the age of 45 then he/she is categorized as ‘SENIOR’ . The new column ‘Employed’ contains sum of months employed and years employed , ‘avg\_risk’ column contains the average of five risk factors , ‘PA’ which personal account column again contains the sum of personal\_account in months and personal \_account in years and the last column ‘ext\_quality’ contains average of exit quality score 1 and exit quality score 2 . All these new features are merged into one Single Data Frame ‘featured’ .

## Data Splitting and Software and Hardware Used

The data set was split into two halves, 80% for the training set and the remaining 20% for the testing set. So the ratio of the training set and the testing set was 4:1 respectively. All of the machine learning implementations and analyses were carried out using Python 3.7 and Tensor Flow 3.0 and a Google Colaboratory notebook. 16 GB RAM and Intel i7 10th generation processor was the hardware used in the implementation of the ML Models.

# technology Used

We have executed various Machine Learning Classifiers such as Logistic Regression, Artificial Neural Network, Random Forest Classifier, Gradient boosting ,Support Vector Machine, XG Boost , KNN and Decision Tree Classifier were used for training as well as testing the data set. Major research works have been using these ML classifiers according to the literature. E.g. Surface Crack Detection [27], phishing detection tools [28], predict Campus Placement [29], etc. Table II explains all the ML Classifiers used with their pros and cons. Many of the following classifiers can be implemented both in regression and classification problems.

# Implementation And Results

After the data exploration section now we implement the various ML Algorithms and also use some boosting techniques to additionally increase the accuracy of the Models. The next step involves concluding and measuring the results using some assessment metrics to determine the most superlative classifier for the E signing of loan data set. Table I differentiates various classifiers like Logistic Regression, Artificial Neural Network, Random Forest Classifier, Gradient boosting ,Support Vector Machine, XG Boost , KNN and Decision Tree Classifier . Every classifier has its pros and cons .

TABLE I . Differentiating various Classifiers based on Pros and Cons

|  |  |  |
| --- | --- | --- |
| ML Models | Pros | Cons |
| Logistic Regression (LR) | Easy to impose, deduce, and very efficacious to train. | Not recommended for datasets with non-linear relationships or complex datasets. |
| XG-Boost | Decidedly flexible, can handle absent data with its inherent features. Users can implement cross-validation after each procedure. | Does not perform well on sparse or unstructured data. |
| Random forest (RF) | Useful with categorical and continuous variables. It can be enforced for both classification-related and regression-related problems. | It is very intricate and entails much more time to train as compared to other techniques |
| Support Vector Machine (SVM) | Has a fast rate of prediction. Can work well as both classification and regression model | Can’t work well on large datasets or when the dataset classes are overlapping |
| Gaussian Naive Bayes(GNB) | It handles both continuous and discrete data. More suitable with definite input variables, multi-class estimation complications, etc. | Application is limited. For good accuracy, it requires a smoothing technique to be combined with. |
| K-Nearest Neighbor | Requires no training therefore new data can be added easily. Very easy to impose. | Does not work well on large datasets. And requires feature scaling for good results. Can’t handle missing values and outliers |
| Decision Tree (DT) | Non-linear datasets are recommended. It is parametric free and no normalization is required. | It consumes more training time. And not recommended for large datasets. |
| Gradient Boosting (GB) | Can be used for both classification and regression purposes, is very flexible, and can handle non-existent values | Can cause over fitting, it is a memory and time exhaustive algorithm. It is less interpretive. |
| Artificial Neural Network (A N N) | ANN have the ability to provide the data to be processed in parallel which means they can handle more than one task easily | It requires parallel processors for executing parallel processing so it depends on hardware |

Table II shows the performance of all the ML Algorithms that we have associated in terms of standard

metrics like accuracy score, recall score, f1 score, precision score   
.

TABLE II . RESULT ANALYSIS OF ALL ML ALGORITHMS

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| MODEL | ACCURACY SCORE | RECALL SCORE | F1 SCORE | PRECISION SCORE |
| Logistic Regression (LR) | 0.569235 | 0.703838 | 0.637538 | 0.582653 |
| XG-Boost | 0.635917 | 0.673342 | 0.654767 | 0.637189 |
| Random forest (RF) | 0.619835 | 0.6733442 | 0.654762 | 0.63710 |
| Support Vector Machine (SVM) | 0.600223 | 0.693983 | 0.651412 | 0.613716 |
| K-Nearest Neighbor | 0.5653 | 0.5700 | 0.5600 | 0.5923 |
| Decision Tree (DT) | 0.572864 | 0.608402 | 0.605263 | 0.602156 |
| Gradient Boosting (GB) | 0.634283 | 0.720436 | 0.67955 | 0.643056 |
| Artificial Neural Network (A N N) | 0.614182 | 0.697095 | 0.660442 | 0.627451 |

The metrics like accuracy score, recall score, F1 score, precision score compare all the ML algorithm and describes which algorithm best suits the financial data set to determine whether a client should get loan sanctioned with E sign .

Accuracy is one metric for evaluating classification models. Informally accuracy is the fraction of prediction our model got right.

False-negative (FN) is an assortment of negative data which is characterized as positive. Therefore it is an inference where the ML Models predict the negative class improperly.

A false positive (FP) is an assortment of positive data that has been described as negative. Thus it is an inference where the ML Models predict the positive class imperfectly.

True Positive (TP) data is an assortment of true positive data that has been categorized as positive. Thus it is an inference where the ML Models predict the positive class fittingly.

The equations (1), (2), (3), (4) gives us the formularies for the above-used metrics in assessing the results of different models are as follow:

|  |  |  |
| --- | --- | --- |
|  |  | (1) |
|  |  | (2) |
|  |  | (3) |
|  |  | (4) |

We can perceive from Table II that the XG-Boost technique did exceptionally well and sustained a solid and consistent association with the dataset thus achieving an accuracy score of 0.635917 approximately. It was preceded by Gradient Boosting with an accuracy score of 0.634283 and then Random forest with an accuracy score of 0.619835. When analogized to other models, Logistic Regression and Artificial Neural Network outperformed them in most of the metrics used.

#### We see that the ANN with no feature engineering performs far better than SVM, Random Forest with feature engineering. Based on the analysis of the E-Signing Prediction for Loan Approval problem using Machine Learning models, the Gradient Boosting Classifier was found to be the best performing model, with an accuracy score of 0.64, precision score of 0.65, recall score of 0.70, and F1 score of 0.67, on the test data.

#### However, it is important to note that the models have several limitations. Possible improvements and future work include collecting more data, including additional features that may be relevant to loan approval decisions, and exploring more advanced Machine Learning models, such as neural networks or ensemble methods. Additionally, incorporating explainable methods, such as SHAP values, can help to better understand the factors that influence loan approval decisions and improve the transparency and accountability of the models.

#### Overall, the analysis provides valuable insights into the E-Signing Prediction for Loan Approval problem and highlights the potential of Machine Learning models to improve the efficiency and accuracy of loan approval decisions, while acknowledging the need for further research and development in this area

# Conclusion and Future Work

Our model has given us an accuracy of around 64%. With this, we have an algorithm that can help predict whether or not a user will complete the E-signing step of the loan application. One way to leverage this model is to target those predicted to not reach the e-sign phase with customized on boarding. This means that when a lead arrives from the marketplace, they may receive a different on boarding experience based on how likely they are to finish the general on boarding process. This can help our company minimize how many people drop off from the funnel. This funnel of screens is as effective as we, as a company, built it. Therefore, user drop-off in this funnel falls entirely on our shoulders. So, with new on boarding screens built intentionally to lead users to finalize the loan application, we can attempt to get more than 40% of those predicted to not finish the process to complete the e-sign step. If we can do this, then we can drastically increase profits. As a result, if we can increase the number of loan takers, we are increasing profits. All with a simple model.

##### Acknowledgment

The author would like to express sincere gratitude for the support given to him by his co-authors for the completion of this research project especially sir for his constant support and guidance.

##### References

1. Rao, P. (2021). E-signatures and digital contracts: Challenges and opportunities for financial services. Journal of Financial Services Research, 59(1-2), 1-22.
2. Javadi, M. H., Nourifar, R., & Shokouhyar, S. (2021). Credit scoring using machine learning algorithms: An empirical study on Iranian banking industry. Journal of Money Laundering Control..
3. Choi, J. Y., Kim, K. H., & Park, Y. J. (2021). A machine learning-based approach to predicting consumer loan default probability. Journal of Retailing and Consumer Services, 59, 102329
4. Huang, H., Huang, L., & Li, Z. (2021). Machine learning-based credit scoring models: An empirical study on small and micro enterprises in China. Finance Research Letters, 40, 101831
5. .Javadi, M. H., Nourifar, R., & Shokouhyar, S. (2021). Credit scoring using machine learning algorithms: An empirical study on Iranian banking industry. Journal of Money Laundering Control.
6. Mihaylova, T., & Stefanova, E. (2021). Big data and machine learning in credit scoring. In Handbook of Research on Big Data and Machine Learning Applications for Business Optimization (pp. 68-89). IGI Global.
7. Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on pattern analysis and machine intelligence, 27(8), 1226-1238.
8. Alcalá-Fdez, J., Fernández, A., Luengo, J., Derrac, J., García, S., Sánchez, L., & Herrera, F. (2011). KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. Journal of Multiple-Valued Logic & Soft Computing, 17(2-3), 255-287
9. Werbos, P. J. (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences. PhD thesis, Harvard University.Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, 65(6), 386-408.
10. R. Pramanik, S. Khare, M.K. Gourisaria, “Inferring the Occurrence of Chronic Kidney Failure: A Data Mining Solution,” In: Gupta D., Khanna A., Kansal V., Fortino G., Hassanien A.E. (eds) Proceedings of Second Doctoral Symposium on Computational Intelligence. Advances in Intelligent Systems and Computing, vol. 1374, Springer, Singapore, 2022. DOI https://doi.org/10.1007/978-981-16-3346-1\_5
11. Dey, P. P., & Chakraborty, S. (2021). A hybrid machine learning approach for fraud detection in the banking sector. Journal of Ambient Intelligence and Humanized Computing, 12(8), 8251-8271  
    .
12. Chen, Y., Wang, X., & Han, Q. (2021). Loan application fraud detection using machine learning: A comparative study. Expert Systems with Applications, 167, 114152.
13. Song, M., Park, C., & Lee, D. (2019). Predicting loan delinquency using machine learning algorithms. Expert Systems with Applications, 137, 98-108
14. Chao, L., & Jiao, Y. (2020). A novel machine learning approach for credit risk assessment in online P2P lending. Neural Computing and Applications, 32(23), 16723-16734.
15. Yan, X., Song, X., & Feng, L. (2020). Machine learning algorithms for credit risk analysis in online peer-to-peer lending. International Journal of Intelligent Systems, 35(5), 784-806  
    .
16. T. Wu, S. Liu, J. Zhang, and Y. Xiang, “Twitter spam detection based on deep learning,” In Proceedings of the australasian computer science week multiconference (pp. 1-8), Jan 2017.
17. Oviedo, A. M., & Guzmán, J. (2021). Loan approval prediction using machine learning models: Evidence from Latin America. Journal of Business Research, 130, 488-500.
18. Al-Gharaibeh, K. M., & Al-Betar, M. A. (2020). A novel machine learning approach for loan default prediction. Journal of Intelligent & Fuzzy Systems, 39(1), 743-755.
19. Carbo-Valverde, S., Rodriguez-Fernandez, F., & Udell, G. F. (2021). The impact of machine learning on small business lending. Journal of Business Venturing Insights, 16, e00278.
20. K. Dutta, S. Chandra, M.K. Gourisaria, “Early-Stage Detection of Liver Disease Through Machine Learning Algorithms,” In Tiwari S., Trivedi M.C., Kolhe M.L., Mishra K., Singh B.K. (eds) Advances in Data and Information Sciences. Lecture Notes in Networks and Systems, vol. 318. Springer, Singapore, 2022. DOI: <https://doi.org/10.1007/978-981-16-5689-7_14>
21. V. Kuppili, M. Biswas, A. Sreekumar, H. S. Suri, L. Saba, D. R. Edla and J. S. Suri, “Extreme learning machine framework for risk stratification of fatty liver disease using ultrasound tissue characterization,” Journal of medical systems, vol. 41(10), pp. 1-20, 2017.
22. V. Singh, M. K. Gourisaria, and H. Das, “Performance Analysis of Machine Learning Algorithms for Prediction of Liver Disease,” In 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), pp. 1-7, Sep 2021, IEEE.
23. M. K. Gourisaria, G. Jee, G. M. Harshvardhan, V. Singh, P. K. Singh, and T. C. Workneh, “Data science appositeness in diabetes mellitus diagnosis for healthcare systems of developing nations,” IET Communications, 2022.
24. V. Muneeswaran, and M. Pallikonda Rajasekaran, “Gallbladder shape estimation using tree-seed optimization tuned radial basis function network for assessment of acute cholecystitis.” In Intelligent engineering informatics, pp. 229-239, 2018, Springer, Singapore.
25. A. Sannigrahi, V. Singh, M. K. Gourisaria and R. Srivastava, "Diagnosis of Skin Cancer Using Feature Engineering Techniques," 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), 2021, pp. 405-411, DOI: 10.1109/ICAC3N53548.2021.9725420.
26. <https://www.kaggle.com/datasets/aniruddhachoudhury/esigning-of-loan-based-on-financial-history>
27. A. Chordia, S. Sarah, M. K. Gourisaria, R. Agrawal, and P. Adhikary, “Surface Crack Detection Using Data Mining and Feature Engineering Techniques,” In 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), pp. 1-7, IEEE, Sep 2021.
28. W. D. Yu, S. Nargundkar, and N. Tiruthani, “PhishCatch—a phishing detection tool,” in Proceedings of the 33rd Annual IEEE International Computer Software and Applications Conference (COMPSAC ’09), vol. 2, pp. 451–456, Seattle, Wash, USA, July 2009.
29. P. Jain, S. Khare, and M. K. Gourisaria, “A Data Mining Solution to Predict Campus Placement,” In 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), pp. 1-7, IEEE, Sept 2021.