Bulletin of Business and Economics, 13(3), 30-35
https://bbejournal.com
https://doi.org/10.61506/01.00425

Striking a Balance: Evaluating Credit Risk with Traditional and Machine Learning Models

Faraz Ahmed[1], Kehkashan Nizam[2], Zubair Sajid[3], Sunain Qamar[4], Ahsan[5]

Abstract
This research assesses machine learning models' validity, clarity, and equity, compared to classical models and especially logistic regression in credit risk evaluation. In the traditional model of data management, efficiency and the accuracy of information are challenges; an issue of machine learning models is model selection and multicollinearity. The study intends to help financial institutions establish the best strategy for their needs. Furthermore, it delves into the effect of heterogeneous data sources on the credit risk model using machine learning. The research analyses the implications of using machine learning in assessing credit risk. Interestingly, focusing on peer-to-peer lending platforms, the research aims to deal with the need for more attention to combining machine learning and traditional models in the literature. The deductive method is the application of inferential analyses, the Traditional model is logistic regression, and the Machine Learning model is a neural network (CNN model) based on secondary data from the Kaggle peer-to-peer lending dataset. With likely findings expected to comprise prediction of the probability of default and better availability of loans, risk analysis leads to formulated lending decisions managing a financial portfolio.
Keywords: Credit Risk, Credit Scoring, Default, Traditional Model, Machine Learning Model

1. Introduction
The credit risk with traditional models may face several complications in the context of inefficient management of data, reducing the risk of acquiring inappropriate information, etc. Machine learning models may also encounter several credit-related risks, including the importance of selecting specific models, addressing Multicollinearity problems, etc. By considering these aspects, the study aims to evaluate the reliability, clarity, and fairness of machine learning models compared to traditional models for credit risk. Financial institutions may find this comparison helpful in determining which strategy best meets their unique requirements. Using machine learning techniques, the study may look into the effects of adding additional data sources to credit risk models. In this study, the core focus of the researcher is to ensure investigation, which may result in credit rating techniques that are more comprehensive and accurate.

This study also sheds light on the consequences of employing machine learning to credit risk evaluation in light of changing legal requirements and moral issues. This research also helps consumers and programmers create ethical and acceptable models. SME lending or microfinance can efficiently employ machine learning to a particular credit risk category (Mhlanga, 2021). This might entail investigating possible advantages and difficulties and customizing models to the particulars of the selected segment. The report also attempts to project how machine learning and associated technologies may boost the management of credit risks in the future. However, this study also highlights models' efficacy in enhancing financial implications within the economic sector.

In this study, the researcher has developed a topic based on the research gap so that it can be fulfilled proficiently. Limited studies are available on integrating machine learning and traditional models for credit risk (Noriega, Rivera, and Herrera, 2023). The appropriate methodology for both models also has a gap in the literature (Razali et al., 2021). However, considering these implications, the researcher has focused on the available peer-to-peer lending platform so that in-depth insight can be acquired proficiently. To acquire significant information in evaluating credit risk with traditional learning and model learning, the researcher has adopted a deductive approach in which the researcher relies on the quality and representativeness of the data to eliminate the risk of credit. In addition to this, the researcher also approached financial sectors where this dataset is used with both or single models. This comprehensive implication provides a comparison of both models. The data collection method for this study relied on secondary sources, from which the researcher approached the online peer-to-peer lending data set available at Kaggle. This site allowed the researcher to utilize in-depth insight regarding machine learning and traditional models with credit risk (Koskimäki, 2021). This site also ensures the recent data is not too old so that current complications and techniques are highlighted proficiently. Loan terms, financial history, demographics, and performance indicators underpinned the data collection method. These factors are also recognized as inducing credit risk and causing complications for the financial sector.

The researcher adopted regression and machine learning techniques for the data collection technique to acquire critical analysis of the models. These approaches allow the researcher to evaluate the capabilities of both models in the context of credit risk prediction. In addition, the regression model offers a detailed analysis of the relationship among variables, and the machine learning model ensures predictive accuracy (Kigo, Omondi, and Omolo, 2023). For the underlined topic, the researcher addresses the risk of credit about linear and non-linear relationships among the dataset. These tools are essential for providing the credit risk with traditional and machine learning models within the financial sector.

After analyzing these data, I found that the advantages of statistical and intelligent (like machine learning) approaches include the ability to identify intricate linkages in credit data. The findings may reveal default likelihood forecasts may be more reliable, reducing risks for lenders and facilitating better loan availability for worthy applicants. Moreover, credit data frequently has unequal classifications, with much fewer defaults than non-defaults. The regression method may highlight the replication approaches that can address this problem by enabling models to gain knowledge more efficiently from the underrepresented class and producing more difficult and equitable risk evaluations.

The researcher may also acquire results related to intelligent methods, which are prominent for their adaptability and capacity to take in new information. Apart from this, the machine learning models shed light on the accuracy of recognizing credit risk

[1] PhD Scholar, Department of Business Administration, IQRA University, Karachi, Pakistan, faraz.57964@iqra.edu.pk
[2] PhD Scholar, Department of Business Administration, IQRA University, Karachi, Pakistan, kehkashan.60003@iqra.edu.pk
[3] Lecturer, Department of Computer Science, IQRA University, Karachi, Pakistan, zubair.sajid@iqra.edu.pk
[4] MBA, Department of Management Sciences, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology University, Karachi, Pakistan, sunain.qamar@gmail.com
[5] Lecturer, Department of Commerce, Benazir Bhutto Shaheed University, Lyari, Karachi, Pakistan, ahsan@bbsul.edu.pk

(Alonso and Carbó, 2020) in the context of improving the generalization capacity of models and increasing their dependability across various borrower segments and economic situations, as statistical approaches frequently yield comprehensible outcomes. However, banks may recognize practical advantages from this strategy, and while maintaining responsible credit availability, more precise credit risk assessments can provide better loan pricing, enhanced portfolio administration, and more knowledgeable lending decisions. In conclusion, this study emphasized the relationship between traditional models and machine learning models for credit risk, in which the researcher identified a gap in the literature. Based on the underlined gap, the researcher developed a conceptual model along with appropriate methodology so that in-depth insight could be accomplished.

2. Literature Review

An inclusive related work on the proposed topic of credit risk with traditional and machine learning models explores a growing interest in integrating modified advanced techniques to improve the accuracy, authenticity, and efficiency of credit risk assessment. Various studies have investigated the strengths and limitations of traditional models compared to machine learning algorithms. Noriega, Rivera, and Herrera (2023) shed light on the integration of machine learning and conventional models for credit risk assessment, and Razali et al. (2021) contributed by highlighting the limited methodologies available for both traditional and machine learning models to ensure the reliability of credit risk assessment models. The focus has been turned toward Kaggle lending platforms as a valuable context for studying credit risk. The extraordinary characteristics of the Kaggle dataset, which leverages data, provide a ground for exploring the integration of traditional and machine learning. This proposed article aims to fill the existing void in the literature and contribute to a more improved and complete understanding of credit risk assessment in the developing era of financial technology. This study focuses on credit risk assessment using data mining techniques such as logistic regression, specifically when dealing with unbalanced datasets.

The survey by Dumitrescu et al. (2021) approached Machine Learning for Credit Scoring, using Logistic Regression with a Non-Linear decision tree for improvement. Research work typically involves exploring methods to improve the accuracy and authenticity of credit risk assessment models, mainly under conditions where the data is imbalanced, which means that there might be significantly more instances of one class (example, non-default) than the another (example, default) (Alonso and Carbo, 2022). Machine learning techniques involve the logistic regression model technique, which is famous for its ability to handle complex data patterns, making it more versatile and suitable for solving tasks related to credit risk assessment.

Furthermore, Kaggle lending platforms collect various data about borrowers, including financial history, credit scores, employment details, etc. Kaggle dataset, like in the study of Naik (2021), for credit risk assessment is a usual and ordinary exciting application. Researchers often build models by using this data to predict the probability of a borrower defaulting on a loan. Generally, defaults are less frequent than successful repayments, so the data set arrives unbalanced. Also, Mutembete and Mathias B's (2022) study used a Machine Learning model for Credit Risk Assessment.

According to the study by Quaranta, Calefato, and Lanubile (2021), Kaggle datasets are used to evaluate the performance of logistic regression models in predicting credit risk. The study equated the efficiency of logistic regression with other algorithms and offered insights into the strengths and limitations of each approach. On the Kaggle platform, logistic regression models have been a famous selection because of their simplicity and ease of implementation; for instance, Wang et al. (2020) expanded on the traditional logistic regression approach by merging feature engineering and ensemble methods with the help of Kaggle datasets, the study shows the comprehend predictive accuracy compared to standalone logistic regression models, demonstrating the potential efficiency for evaluating credit risk assessment through a combination of traditional and advanced techniques.

Kaggle datasets provide diverse scenarios and variables (Bojer and Meldgaard, 2020). It is crucial to acknowledge that the universality of models trained on these datasets to real-world credit risk scenarios demands careful considerations, as the literature emphasizes the significance of validation on external datasets or real-world data to make a model more accurate and to ensure the robustness and proficiency of the model developed on Kaggle datasets. It is also essential to note that the selection of the regression model depends on the features of the data and the particular requirements of the credit risk assessment tasks. A combination of traditional and machine learning models is often used to obtain a more accurate prediction (Hussin Adam Khatir and Bee,2022).

3. Methodology

The research methodology used in this study seeks to analyze credit and evaluate it more embracingly through the use of traditional models, logistic regression models, machine learning models, and neural networks (CNN model). The methodology includes the line of research in the inductive approach concerning the reliability, accuracy, and fairness of these models of work for fairness in credit risk assessment (Breeden, 2021). The empirical analysis mainly utilizes secondary information from Kaggle's open-source peer-to-peer lending dataset, selected for its dynamic and detailed content about machine learning and traditional credit risk models. This approach is deductive because it allows us to check existing theories and concepts as they relate to credit risk processes. High-quality and representative data are used in the study here to help minimize biases, thereby generating a high value of authentic and valid results. Financial sectors that apply either model or both are approached to get relevant data to make a comprehensive comparison.

The independent variables are Borrower's current delinquency accounts, Trades opened in the past 24 months, Borrower's self-reported annual income, Co-borrowers combined self-reported annual income, Loan application type (individual or joint), The average current balance of all accounts, The ratio of total current balance to high credit/credit limit for bankcard accounts, Charge-offs within the last 12 months,30+ days past-due incidences in the past 2 years, The past-due amount owed for delinquent accounts, Borrower's debt-to-income ratio, Co-borrowers debt-to-income ratio, Employment length in years, Home ownership status, Credit inquiries in the past 12 months, Inquiries in the past 6 months (excluding auto and mortgage),Monthly payment amount, Loan interest rate, Number of mortgage accounts, Months since Borrower's last delinquency, Months since most recent 90-day or worse rating, Accounts ever 120 or more days past due, Satisfactory bankcard accounts, Number of bankcard accounts, Number of instalment accounts, Number of open revolving accounts, Number of revolving accounts, Number of revolving trades with balance >0, Satisfactory accounts, Accounts currently 120 days past due (updated in past 2 months), Accounts currently 30 days past due (updated in past 2 months), Accounts 90 or more days past due in last 24 months, Accounts opened in the past 12 months, Total credit revolving balance, Charge-offs within previous 12 months for secondary applicant, Collections within

last 12 months for secondary applicant (excluding medical), Months since most recent 90-day or worse rating for secondary applicant.
With the information appearing on the dataset from Kaggle, loan terms, financial history, demographics, and performance indicators were involved. These are, in fact, essential aspects that affect the level of credit risk and significance for the finance industry. The online peer-to-peer lending dataset is most relevant because it represents real-life situations; thus, the findings can be easily implemented in solving the current complications and techniques in the industry (Ariza-Garzon et al., 2020). The conceptual framework of the study consists of independent variables and dependent variables. The independent variable (refer to Annexure I) and the dependent variable are credit risk. Both logistic regression models and machine learning techniques are implemented for a comprehensive assessment. Logistic Regression analysis facilitates a nuanced look at relationships among the variables and affects prediction accuracy (Dumitrescu et al., 2021).
Predictions of results liable to be made are better default likelihood forecasts, which mean less risk for lenders and allow better loan access for creditworthy candidates. The analysis predicts that the machine learning models will show flexibility and correctness in credit risk discernment; thus, the generalization capacity will be increased across different borrower groups and scenarios (Alonso Robisco and Carbó Martínez, 2022). The study also covers the possible advantages for financial organizations, in particular, improved pricing on loans, tools for solid portfolio management, and better-informed decisions by lenders. Although the selected data set is beneficial, one should understand its limits. Some of these could be sample representativeness, biases from the online peer-to-peer lending platform, and the inherent implications of secondary data. In sum, the outlined research methodology allows one to get a comprehensive analysis of credit risk assessment with classical and machine learning methods. The study applies a deductive approach, using a broad database, to provide essential comments to the study of credit risk assessment in the financial sector.

4. Data Analysis

The findings highlighted the reliable performance of logistic regression (Fullerton and Anderson,2021) in specific scenarios, mainly when interpretability and simplicity are crucial; the findings are obtained with the help of the Regression model, as it plays a fundamental role in providing a tool in statistics, and machine learning for evaluating the relationship between variables Especially when it comes to credit risk assessment, logistic regression is widely employed to predict the probability of a customer or Borrower defaulting on a loan.
Logistic regression is commonly used in credit risk modeling because of its simplicity and interpretability (Amaro, 2020). It is well suited for binary classification problems where the results are either 0 or 1 (default or non-default). The model further estimates the likelihood of an event occurring and is particularly effective when the relationship between the independent variables and the log odds of the dependent variable is approximately linear.
Similar to traditional methods, data preprocessing in machine learning also involves cleaning data to ensure that it is suitable for machine learning algorithms. With the help of feature engineering, capturing non-linear relationships is easy. Furthermore, sci-kit-learn in Python is used to train a logistic regression model, ensuring more flexibility and scalability than traditional methods.
It can be seen by the graphical illustrations that the data set is taken from Kaggle, and the visualizations are of Terms for which the loan is taken, Loan status, and the top 10 purposes for getting loans by implementing a logistic regression model to predict if the loan will be charged off or fully paid on time. Further, the evaluation occurred to modify the performance of the model with the help of the accuracy score and confusion matrix.

```
Total Loans:  396030
Total Number of Fully paid Loans:  318357
Defaulters / Not paid Fully:  77673

Percentage of Defaulters:  20.0 %
Percentage of Non-Defaulters:  80.0 %
```

Figure 1: Loan Status

The above representation shows the total number of Loans, which is 396,030. It also shows the total number of loans that were fully paid and the number of Defaulters, those loans that were not fully paid. Then, for better understanding, we have calculated the percentage as well: for those who have paid the loan fully and for those who have not paid the loan and are Defaulters.

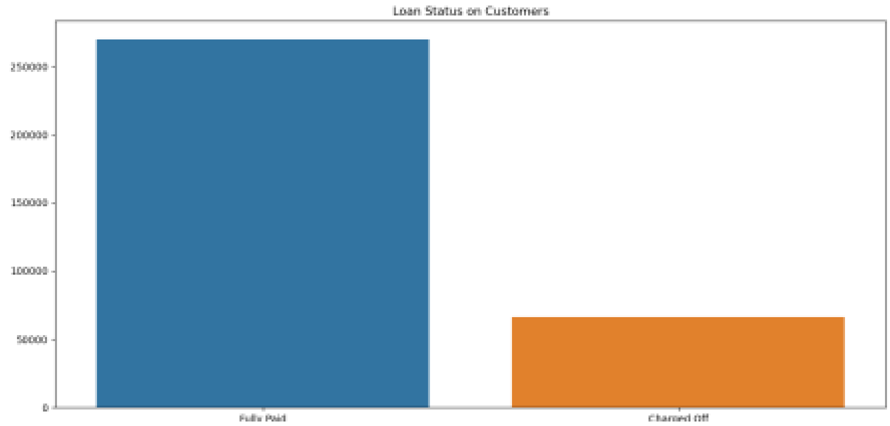

Figure 2: Loan Status of Customers

Figure 2 showcases the results of the customers' loan status. The blue bar indicates the value of customers who have fully paid the loan, while the orange indicates the value of customers who have charged off the loan.

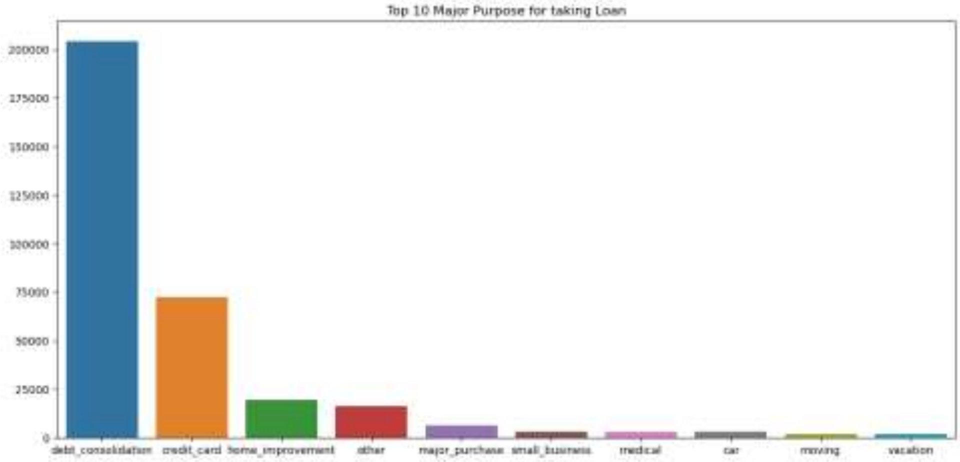

Figure 3: Top 10 Major Purposes for taking loans

Figure 3 sheds light on the 10 significant purposes for customers taking loans, including the bars. The most common purpose is debt consolidation, moving towards the orange bar, which indicates a credit card purpose, home improvement, other significant purchases, small businesses, medical, car, moving, and vacation. The dataset and the results show that the critical purpose is debt consolidation.

The first traditional model, the logistic regression model, is implemented.

Figure 4: Implementing Logistic Regression

The findings demonstrate that the model has achieved 80% accuracy, a positive sign. It suggests that the chosen features have a significant impact on predicting the target variables, and the positive and negative coefficients indicate the positive and negative relationship concerning the relative target variables, P- values associated with each coefficient, a low p-value (typically < 0.5) indicate that the coefficient is statistically significant, and this helps in understanding whether the observed relationship is likely to be genuine or just because of random chance.

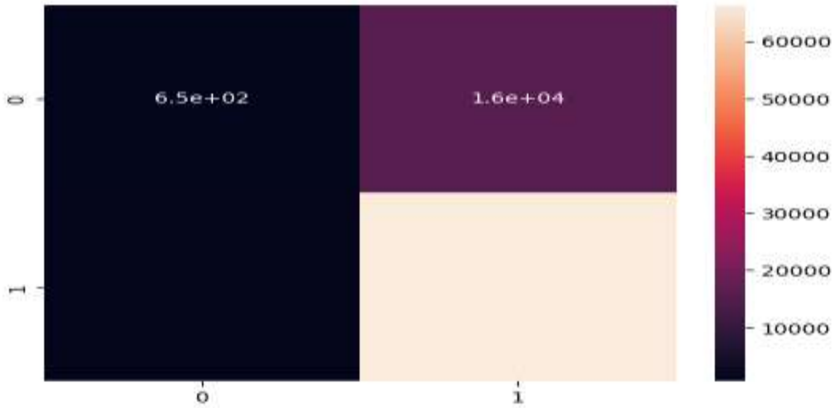

Figure 5: Confusion Matrix

Figure 5 depicts a confusion matrix, a commonly used evaluation metric for classification models; logistic regression can be used for binary classification tasks, and in those cases, a confusion matrix is applied. The matrix derived from the confusion matrix includes the accuracy, precision, sensitivity or actual positive rate, actual negative rate, and F1 score, which indicates the harmonic mean of precision and recall (Heydarian, Doyle, and Samavi, 2022). Adjusting the threshold can impact the confusion matrix and associated metrics.

```
[ ] # Split the data into training and testing sets
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

    # Standardize the data
    scaler = StandardScaler()
    X_train = scaler.fit_transform(X_train)
    X_test = scaler.transform(X_test)


    # Reshape data for Conv1D
    X_train = X_train.reshape(X_train.shape[0], X_train.shape[1], 1)
    X_test = X_test.reshape(X_test.shape[0], X_test.shape[1], 1)
```

```
    # Define your CNN model
    cnn_model = Sequential([
        Conv1D(32, kernel_size=3, activation='relu', input_shape=(X_train.shape[1], 1)),
        MaxPooling1D(pool_size=2),
        Flatten(),
        Dense(128, activation='relu'),
        Dense(1, activation='sigmoid')
    ])

    # Compile the model
    cnn_model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
```

Figure 6: Implementing CNN

Then, a machine learning model, the neural network (CNN model), is implemented, as shown in Figure 6.

Figure 7: Training and Validation Accuracy

This graph shows the training and validation score, where training accuracy has increased, and validation accuracy initially increased and then decreased.

```
2100/2100 [==============================] - 11s 5ms/step
Test Accuracy: 80.51
8397/8397 [==============================] - 41s 5ms/step
Training Accuracy: 80.6
```

Figure 8: Accuracy Scores

The findings demonstrate that the model has achieved 80.51% test accuracy, which is a positive sign, and 80.6% Training Accuracy. This suggests that the chosen features have a significant impact. Then, both models, the traditional model (Logistic Regression) and the Machine Learning Model (CNN), are compared to show which model better predicts default.

```
Accuracy of Traditional Model:  80.0 %

Accuracy of CNN Model:  80.6 %
```

Figure 9: Accuracy Scores Between Models (LR & CNN)

The accuracy results are around 80% in both the traditional and machine learning models. This is a nice accuracy figure. The CNN model performs slightly more accurately (80.6%) than the conventional model. Even this little difference is vital to the implementation and accuracy in various cases.

5. Conclusion

In conclusion, the study embarked on credit risk assessment models, comparing well-established ratios, logistic regression-based approaches, and Machine Learning-based techniques. The lack of extra data sources using the machine learning approaches served as a reminder for further consideration of this direction of using machine learning approaches. In this research gap closure, the study highlighted the need for more literature to integrate machine learning and traditional models for credit risk. The choice of the deductive research approach, the data from the online peer-to-peer lending dataset from Kaggle was collected, allowed practical analysis of the models. The research outcomes predicted included better forecasts of default likelihoods to diminish uncertainty risks for the lenders in allocating resources toward more economically viable applicants and improving loan availability for credit-worthy borrowers.

The research projected that machine learning models would demonstrate resilience and performance, contributing positively to such insights through better loan decisions and portfolio management. The importance of this study is found in the fact that it gives possible directions to financial institutions by recognizing relevant evaluation approaches to credit risk. At the end of this research, the connection between traditional and machine learning models in credit risk assessment can be shown as quite complicated. The results change how credit risk assessment is done in the financial market, leading to better lending and improved risk management. The ongoing monitoring and validation of machine learning models in the real world, a research collaboration between academia and industry practitioners, and continued focus on changing law and ethics will be necessary to ferment credit risk assessment.

References

Alonso Robisco, A. and Carbó Martínez, J.M. (2022). Measuring the model risk-adjusted performance of machine learning algorithms in credit default prediction. *Financial Innovation*, 8(1).

Alonso, A. and Carbó, J.M. (2020). *Machine Learning in Credit Risk: Measuring the Dilemma Between Prediction and Supervisory Cost*.

Amaro, M.M. (2020). *Credit scoring: comparison of non-parametric techniques against logistic regression*.

Ariza-Garzon, M.J., Arroyo, J., Caparrini, A. and Segovia-Vargas, M.-J. (2020). Explainability of a Machine Learning Granting Scoring Model in Peer-to-Peer Lending. *IEEE Access*, 8, pp.64873–64890.

Bojer, C.S. and Meldgaard, J.P. (2020). Kaggle Forecasting competitions: an Overlooked Learning Opportunity. *International Journal of Forecasting*, 37(2).

Breeden, J. (2021). *A Survey of Machine Learning in Credit Risk*.

Doumpos, M., Christos Lemonakis, Dimitrios Niklis and Constantin Zopounidis (2019). *Analytical Techniques in the Assessment of Credit Risk. EURO advanced tutorials on operational research*.

Dumitrescu, E., Hué, S., Hurlin, C. and Tokpavi, S. (2021). Machine Learning for Credit Scoring: Improving Logistic Regression with Non-Linear Decision-Tree Effects. *European Journal of Operational Research*, 297(3).

Fullerton, A.S. and Anderson, K.F. (2021). Ordered regression models: A tutorial. *Prevention Science*, pp.1-13.

Heydarian, M., Doyle, T.E. and Samavi, R. (2022). MLCM: Multi-Label Confusion Matrix. *IEEE Access*, 10, pp.19083–19095.

Hussin Adam Khatir, A.A. and Bee, M. (2022). Machine learning models and data-balancing techniques for credit scoring: What is the best combination?. *Risks*, *10*(9), p.169.

Kigo, S.N., Omondi, E.O. and Omolo, B.O. (2023). Assessing predictive performance of supervised machine learning algorithms for a diamond pricing model. *Scientific Reports*, *13*(1), p.17315.

Koskimäki, M. (2021). Default prediction in peer-to-peer lending and country comparison. *lutpub.lut.fi*.

Mhlanga, D. (2021). Financial inclusion in emerging economies: Applying machine learning and artificial intelligence in credit risk assessment. *International journal of economic studies*, *9*(3), p.39.

Naik, K.S. (2021). Predicting Credit Risk for Unsecured Lending: A Machine Learning Approach. *arXiv preprint arXiv:2110.02206*.

Naili, M. and Lahrichi, Y. (2020). The determinants of banks' credit risk: Review of the literature and future research agenda. *International Journal of Finance & Economics*, 27(1). Alonso Robisco, A. and Carbó Martínez, J.M. (2022). Measuring the model risk-adjusted performance of machine learning algorithms in credit default prediction. *Financial Innovation*, 8(1).

Noriega, J.P., Rivera, L.A. and Herrera, J.A. (2023). Machine Learning for Credit Risk Prediction: A Systematic Literature Review. *Data*, *8*(11), p.169.

Quaranta, L., Calefato, F. and Lanubile, F. (2021), May. Kgtorrent: A dataset of python jupyter notebooks from Kaggle. In *2021 IEEE/ACM 18th International Conference on Mining Software Repositories (MSR)* (pp. 550-554). IEEE.

Razali, N.A.M., Shamsaimon, N., Ishak, K.K., Ramli, S., Amran, M.F.M. and Sukardi, S. (2021). Gap, techniques, and evaluation: traffic flow prediction using machine learning and deep learning. *Journal of Big Data*, *8*(1), pp.1-25.

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.