ISSN 1816-6075 (Print), 1818-0523 (Online)
Journal of System and Management Sciences
Vol. 14 (2024) No. 3, pp. 115-126
DOI:10.33168/JSMS.2024.0307

Applying Machine Learning Algorithms to Predict Liquidity Risks

Tisa Maria Antony, Sathish Kumar B

School of Commerce Finance and Accountancy,
CHRIST (Deemed to be University), Bengaluru, India
tisamariaantony@gmail.com

Abstract. Liquidity risk is a significant financial threat, and its mismanagement can lead to substantial financial losses. This study examines the application of machine learning techniques like KNN, SVM, decision tree, RF and XGBoost for predicting liquidity risk in Indian banks. Financial data from 2013-2022 for 31 commercial banks is analyzed. The models use financial ratios as predictors and liquidity risk is proxied by liquid assets to total assets and loan to deposit ratio. Despite limitations in generalizability due to the small sample, the results exhibit the potential of algorithms like KNN and XGBoost for accurate liquidity risk forecasting. The study's findings revealed that the model that used liquid asset to total assets to proxy liquidity risk gave the best results with the KNN, and gave MAE and MSE scores of 0.129 and 0.027, respectively. When loan to deposit was used to proxy liquidity risk, DT was the best-performing algorithm, with an MAE and RMSE score of 0.191 and 0.231, respectively. It was also found that MLP did not perform well when compared to other selected models. The practical implications include developing liquidity early warning systems for Indian banks using these techniques.

Keywords: machine learning, liquidity risk, feature selection, testing

1. Introduction

Banks are exposed to several financial risks, including liquidity risk, credit risk, operational risk, and market risk. In comparison to other risks, liquidity risk might be far more significant (Adalsteinsson, 2014). Failure to meet deposit withdrawal requests when required, either due to the inability to efficiently sell liquid assets or buy bank liabilities, could harm the bank's reputation of effective liquidity risk management. The liquidity risk is often the focal point of every financial crisis. The financial crisis 2008 that caused multiple bank failures can be blamed on liquidity problems. The recent crisis at Signature Bank and Silicon Valley Bank can also be attributed to liquidity issues.

In recent years, Indian banks have also been exposed to credit, liquidity and market risk as an outcome of various economic and non-economic events. The Indian government has introduced the bad bank to ease the burden of non-performing assets in Indian commercial banks (Khurana, 2021 &Grover, 2021). The previous decade has witnessed scams and failures associated with major financial institutions, including Lakshmi Vilas Bank, IL&FS, Diwan Housing Finance Corporation (DHFL), Punjab and Maharashtra Cooperative Bank (PMC), and Yes Bank, thereby alarming the Indian Banking system. Therefore it is essential to constantly evaluate and manage the levels of financial risk, including credit, liquidity, and market risks.

All this calls for a need to build an accurate model that precisely predicts the liquidity risks. Though several traditional models are developed over time to predict liquidity risk, there need to be more rigid forecasting models that can precisely forecast the liquidity risk of commercial banks. To address this issue, this study has explored artificial intelligence (AI), a major technological advancement that has become highly important in addressing various problems related to risk assessment and prediction problems. Building models, in particular statistical models that can predict and forecast, is an essential aspect of machine learning. Machine learning (ML) algorithms can be used to make predictions from underlying data. Using past data ML, algorithms can forecast values relating to events that have not yet occurred. This paper uses various ML algorithms to predict the liquidity risk of Indian commercial banks, using the key determinants of liquidity risk. To examine if fewer features could enhance the outcome, the study has further developed models using feature selection wherein only variables with minor correlation are considered, and the remaining variables are screened out. Finally, the study further compares the predictive capabilities of different ML algorithms to identify the best-performing algorithm.

The remainder of the paper is structured as follows. Section 2 details the related literature and discusses the research gap. Section 3 formulates the process of the proposed methodology in detail. The results are presented and further discussed in Section 4. Finally, section 5 concludes the study and provides the limitations and scope for further research.

2. Review of literature

Risk management using machine learning is a well-studied area in business finance literature. However, most studies focus only on works relating to credit risk prediction, risks concerning bankruptcy, financial fraud and financial failure. Khalid et al.(2022) applied ML techniques such as random forest (RF), decision tree (DT), naïve bayes (NB) and k-nearest neighbour (KNN) to assess corporate risk, and findings indicated that RF performs better than other models. Similarly, Mousa et al. (2022) also indicated that RF outperforms other models, such as LDA and QDA, in assessing financial risk. Predicting financial risk using K means, Expectation-maximization (EM), COBWEB, Repeated bisection approach, Graph-partitioning algorithm, and Density-based method reveals that the repeated bisection approach was found to perform the best in comparison to other models (Kou et al.,2014). Studies that have tried to predict bankruptcy using models like SVM and NN with dropout and autoencoder have revealed that NN with added layers with dropout has the highest accuracy (N. Wang, 2017). Credit risk using traditional ML models and hybrid models based on MLP revealed that hybrid models based on MLP perform better (Chi et al., 2019). An evaluation and comparison of algorithms like SVM, NB, ANN, KNN, RF, logistic regression, and bagging in detecting financial accounting fraud indicated that RF without feature selection -oversampling model performs better than other models (Hamal & Senvar, 2021). Studies have also compared the performance of DT, NN, and Bayesian belief networks (BBN) for detecting fraudulent financial statements, and the results showed that the BBN have higher accuracy than DT and NN (Kirkos et al., 2007). Studies have also employed DT, ANN, NB, SVM, and BBN for building predictive models using intentional and unintentional financial restatements, indicating that ANN performs better than other algorithms based on accuracy and area under the receiver operating characteristic curve (Dutta et al., 2017). Studies employed ANN for predicting fraudulent financial reporting in small market capitalization companies. The findings indicated that ANN showed a higher prediction result (94.87%) than traditional statistics, linear regression and other techniques (Omar et al., 2017). Ravisankar et al. (2011) used multilayer feed-forward NN, SVMs, genetic programming, group method of data handling, logistic regression, and probabilistic neural network (PNN) to identify financial statement fraud. The models were assessed with and without feature selection. Results revealed that PNN performs better than other techniques without feature selection, and with feature selection, both genetic programming models and PNN performs better than other models. Yao et al. (2018) also developed an optimized financial fraud detection model that combined feature selection and ML classification. Results indicated that RF outperformed the other methods, such as SVM, DT, ANN, and linear regression. As to feature selection methods, XGBoost was found to perform the best. Studies have also analyzed the performance of CART, CHAID, BBN, SVM, and ANN in constructing fraudulent financial statement detection models and findings indicated that CHAID – CART performs better than other models under study (Chen, 2016).

Comparatively, fewer studies have been published on liquidity risk prediction. However, liquidity risk should be periodically monitored, considering the severe repercussions of a liquidity shortage in the securities and financial markets during a potential financial crisis (Saleemi, 2014). Studies have employed ML models for measuring, managing and predicting liquidity risk. Popular deep-learning models like ANN and Bayesian networks have also been used in liquidity risk management (Leo et al., 2019). Studies have also investigated whether ML techniques can develop liquidity risk models for providing insights during stress-testing scenarios and have compared ML algorithms to a traditional statistical model for benchmarking. Findings indicate that XGBoost performs better than other methods for this classification problem (Guerra et al., 2022). Studies have also revealed that ANN and BN implementations could distinguish the most critical risk factors and assess the risk by a functional approximation and a distributional estimation (Tavana et al., 2018). Studies have also used ML algorithms to develop an early warning system to identify banks likely to have liquidity crises. The study compared the predictive performance of LASSO, RF and XGBoost, and their combination revealed that combined models achieve an extremely low percentage of false negatives. Results also indicate that ML algorithms perform better when compared to traditional logistic regression (Drudi & Nobili, 2021). ANN and CART models are also popular in their capability to forecast the financial failure of unsuccessful businesses (Aksoy & Boztosun, 2020).

As has already been mentioned, there is a lack of studies in the area of LRM using ML algorithms, specifically in the Indian banking sector. To compensate for this, the study tries to predict the liquidity risk of Indian commercial banks. The current study supports and expands the existing literature in several ways. By outlining the numerous methods that can be utilized for precise risk prediction, the study broadens the usage of artificial intelligence technologies in the Indian banking industry. The study further develops a machine learning framework that may assist academic practitioners and industrialists in employing the best-fit AI technique with the highest degree of precision and the least error level in the prediction of liquidity risk.

3. Methodology

The study aims to identify the best ML model for liquidity risk prediction. The data has been analyzed using Python 3.10.7. The data was initially processed to develop financial ratios to estimate liquidity risks (detailed in Sec. 3.1). The input dataset is split into training and testing datasets. The training dataset contains 80% of the data, while the testing dataset uses the remaining 20%. Six commonly used ML algorithms, including KNN, SVR, DT, RF, XGBoost and MLP, are applied for prediction and are then evaluated using key evaluation metrics, MAE, MSE and RMSE (outlined in Sec. 3.2). Models are developed with and without feature selection. When the feature selection model is applied, only the most significant variables are reassigned to prediction models (detailed in Sec. 3.3). Figure 1 summarizes the methodology of this study.

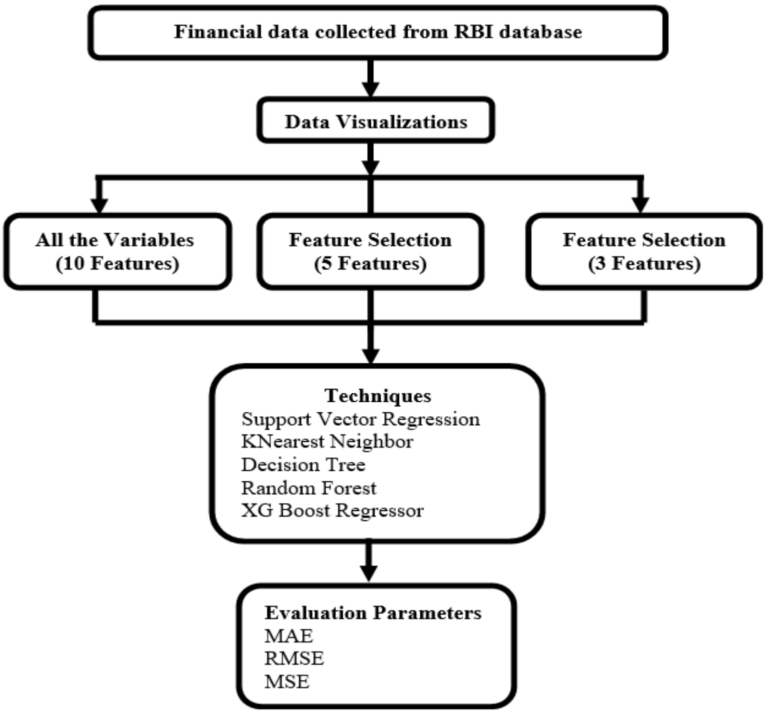

Fig.1: Methodology process overview.

3.1. Data and Variables

The study has retrieved financial data from the DBIE published by the Reserve Bank of India. The period covered in the study is ten years, ranging from 2013 to 2022. There are 33 Indian commercial banks. In this, two banks that lacked sufficient information to construct all proxy measures were excluded, and finally, 31 banks were used in the study.

For predicting liquidity risk and developing various machine learning algorithms, the dependent variable used in the study is liquidity risk, which is proxied by LTA and LDT. LTA was measured as liquid asset by total assets, and LDT was measured as loan to deposit ratio (Liao et al., 2009;Godlewski, 2006). The independent variables in this study include return on assets (ROA), which was measured as net profit by total assets (Masood & Ashraf, 2012), age of the bank (Misman and Bhatti, 2020) and bank size (BSZE) was measured as the natural logarithm of total assets (N. Gupta & Mahakud, 2020). Operating efficiency (OEI) was calculated as total operating expense divided by operating income (Al-Qudah et al., 2022). Capitalization (CAP) was measured as total equity to total assets (Salike& Ao, 2018), and bank diversification (BDF) is calculated as noninterest income divided by total income

(Louzis et al., 2010). Bank ownership indicates whether the bank is owned by the government or private parties (J. Gupta & Kashiramka, 2020), and whether the bank has undergone any merger or acquisition in a given year is proxied by dummy variables (Battaglia& Mazzuca, 2014).

3.2. Machine learning models and evaluation metrics

The study initially used the KNN algorithm, an essential technique for handling regression problems. KNN algorithm calculates the distances between each observed data value and the new data value with the unknown target. SVR is supervised learning to make predictions about discrete values. DT regression is a tree-structured model whereby internal nodes indicate the features of a dataset, branches indicate the decision rules, and the individual leaf nodes indicate the result. RF is a popular ML algorithm that is used in regression problems. It is an ensemble method that reduces over-fitting by averaging the outcome, making it superior to a single DT. Boosting algorithms, XGBoost, have also been employed in the study. XGBoost is a DT-based ensemble ML algorithm that uses a gradient-boosting framework. Multi-layer perceptron (MLP) is a supplement of feed-forward NN. It consists of three layers: the input layer, the output layer and the hidden layer.

In statistics, the MAE is a metric that measures how close forecasts or predictions are to the eventual outcomes.

$$\mathrm{MAE} = \frac{|(y_i - y_p)|}{n} \quad (1)$$

The Mean Squared Error assesses how closely a regression line matches a set of data points and is a risk function corresponding to the anticipated amount of the squared error loss.

$$\mathrm{MSE} = \frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2 \quad (2)$$

Root mean square error is a popular metric for assessing the quality of predictions, as it illustrates the Euclidean distance between measured true values and forecasts.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2} \quad (3)$$

3.3. Data transformation and Feature selection

The data is initially transformed, wherein data is cleaned, and missing values are dealt with. Further, ML models are applied after considering all the variables under study. Thereafter, a correlation-based feature–selection model is applied, wherein the correlation is applied to determine the most significant independent variables to be applied in the prediction models. In the feature selection model, the most relevant variables (variables with the least correlation) are reassigned to the prediction models. To examine if fewer features could enhance the outcome, the study selected 5 features and 3 features with minor correlation and screened out the remaining variables. Thus, six models are developed with and without feature selection and consider LTA and LDT as proxies of liquidity risk. The variables used to build each model, with and without feature selection methods, are detailed in Figure 2.

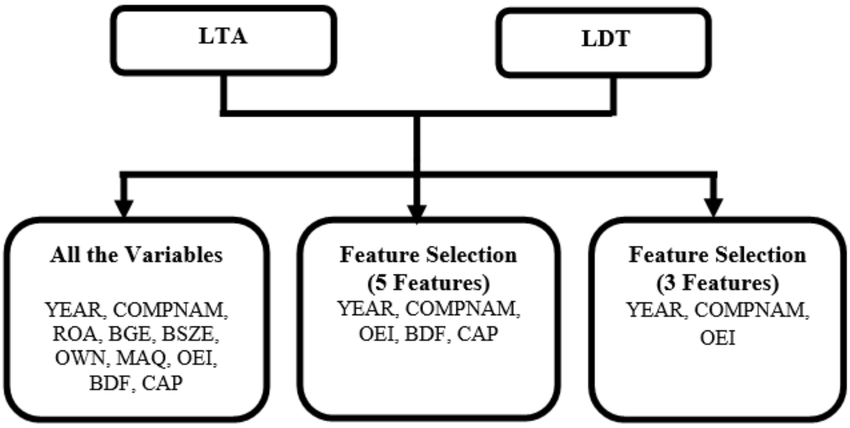

Fig.2: Variables used to build each model with and without feature selection methods.

4. Results

The study initially trained the dataset with the selected ML models without applying feature selection to identify the best model for predicting the banks' liquidity risk. Figure 3 shows the accuracy of ML models for predicting liquidity risk (proxied by LTA) using all the variables (i.e., without feature selection). Results reveal that KNN performs the best as compared to other models. XGBoost, RF and DT follow KNN. The deep learning algorithm MLP was found to be the worst performing algorithm.

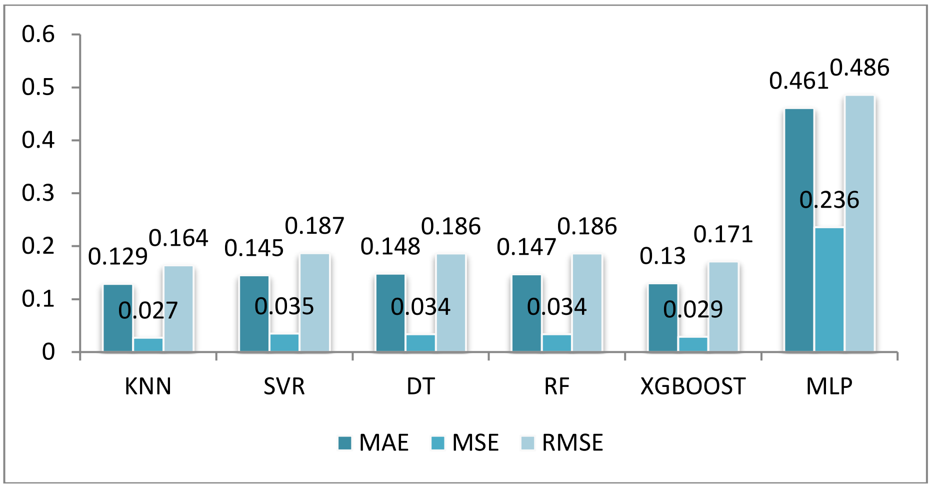

Fig.3: Accuracy of ML models for predicting liquidity risk (LTA) using all the features.

After applying the feature selection model, the 5 variables with the least correlation are reassigned to the prediction models. As shown in Figure 4, results reveal that KNN performs the best compared to other models. XGBoost, RF and DT follow KNN. The deep learning algorithm MLP was found to be the worst-performing algorithm. However, results reveal that MLP performs better using the feature selection model compared to the model that did not use feature selection.

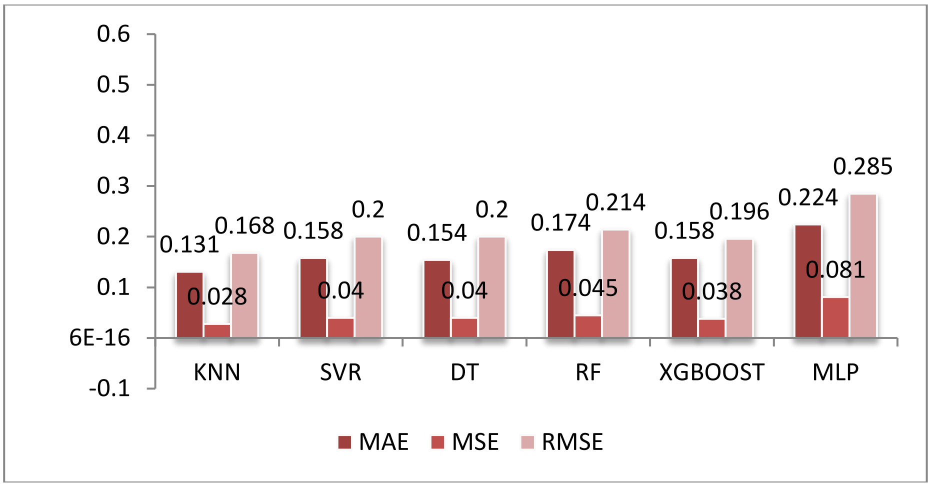

Fig.4: Accuracy of ML models for predicting liquidity risk (LTA) using all 5 features.

Figure 5 shows the accuracy of ML models for predicting liquidity risk (LTA) using 3 features. Results reveal that XGBoost performs the best compared to other models and is in line with the studies of Guerra et al. (2022). RF, KNN, and DT also perform well compared to other models. The deep learning algorithm MLP was found to be the worst-performing algorithm. Nevertheless, as found previously, MLP performs significantly better using feature selection than without feature selection.

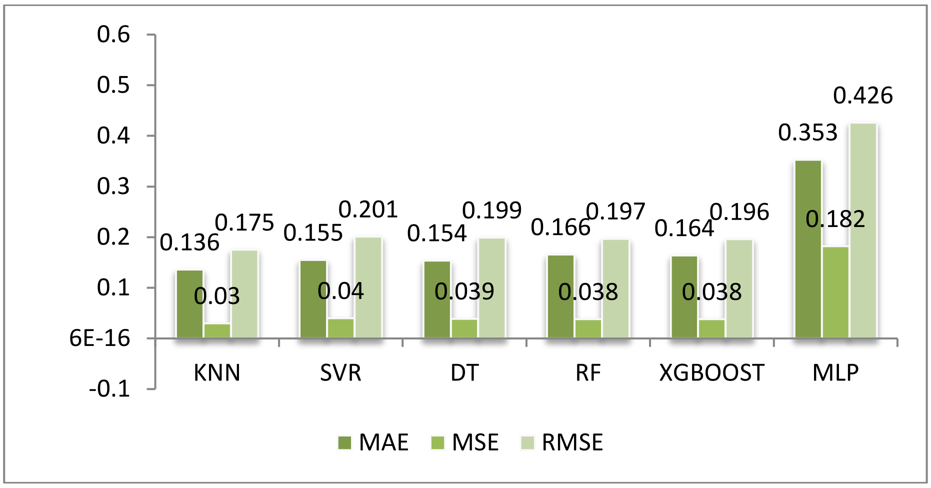

Fig.5: Accuracy of ML models for predicting liquidity risk (LTA) using 3 features.

Figure 6 compares the accuracy of selected ML models for predicting liquidity risk using LDT as a measure. Performance metrics of the model, which considered all the selected variables (without feature selection), reveal that DT performs best compared to other models. KNN and XGBoost were also found to perform well. The deep learning algorithm MLP was found to be the worst-performing algorithm based on MAE, MSE and RMSE scores.

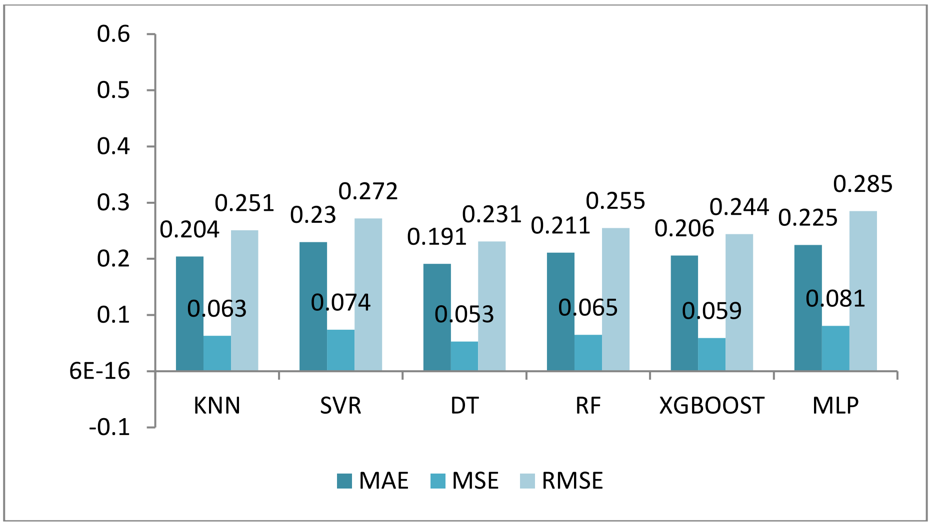

Fig.6: Accuracy of ML models for predicting liquidity risk (LDT) using all the features.

Figure 7 shows the accuracy of ML models for predicting liquidity risk using LDT using 7 features, with minimal correlation. Results reveal that XGBoost is the best-performing algorithm. KNN, RF, MLP and DT also performed moderately well. Results also show that SVR was found to be the worst-performing algorithm.

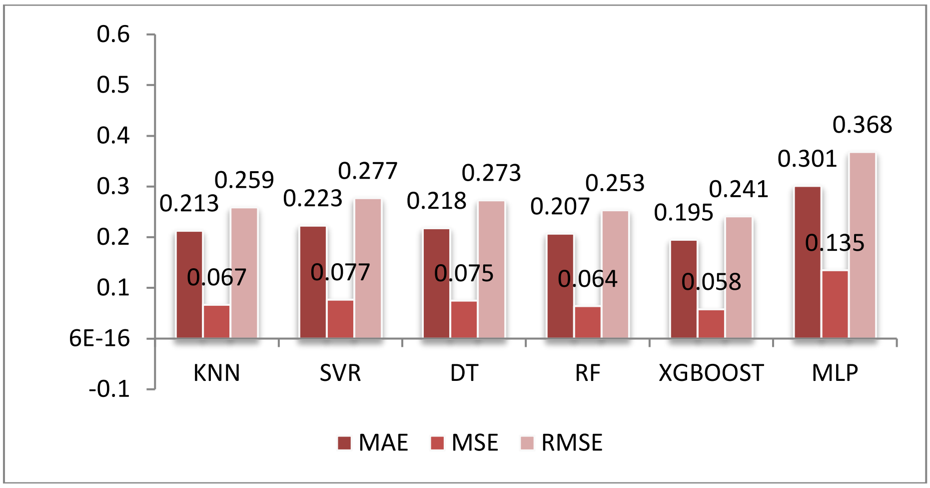

Fig.7: Accuracy of ML models for predicting liquidity risk (LDT) using 7 features.

Figure 8 shows the accuracy of ML models for predicting liquidity risk, proxied by LDT using 3 most uncorrelated features. Results reveal that KNN is the best-performing algorithm based on MAE, MSE and RMSE. Based on MAE, MSE and RMSE scores, MLP is the worst-performing algorithm. RF performed well based on MAE, MSE and RMSE scores.

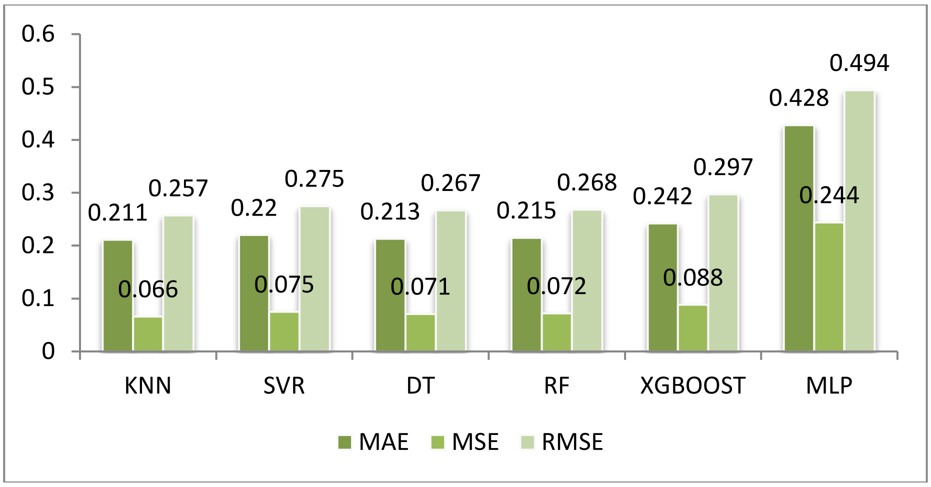

Fig.8: Accuracy of ML models for predicting liquidity risk (LDT) using 3 features

5. Discussions

Machine learning has been a major technological advancement that has grabbed much attention due to its immense potential. To obtain a reliable system for assessing a bank's liquidity risk, the study has compared the predictive capabilities of different machine learning technique such as linear regression, KNN, SVR, DT, RF, and XGBoost. An overall comparison of the accuracy of various ML models for predicting liquidity risk revealed that XG Boost, with feature selection (3 features), provides substantial competitive gains across all machine learning methods. The aforesaid model used LTA to proxy liquidity risk. This finding is in line with the results of Guerra et al., (2022). Without feature selection, MLP was found to be the worst-performing algorithm, using both LTA and LDT as proxies of liquidity risk and is contrary to the findings of Chi et al.(2019). After performing feature selection, MLP performed worst when LTA was used to proxy liquidity risk, whereas when LDT was used to proxy liquidity risk, SVR performed worst. Even so, the result suggests that machine learning models can be used as one promising tool for liquidity risk assessment. Even though the sample has been restricted to Indian commercial banks, future studies can also incorporate information from other countries thereby extending the scope of this research. Further, more machine learning algorithms and hybrid models can be used for predicting liquidity risk.

6. Conclusions and Implications

This study analyzed various ML techniques for predicting liquidity risk in Indian banks using financial data from 2013-2022. The study's findings revealed that the model that used LTA to proxy liquidity risk gave the best results with the XGBoost algorithm with feature selection. When LDT was used to proxy liquidity risk, DT was found to be the best-performing algorithm. RF and KNN also performed well with and without feature selection. Collectively, for both models that used LTA and LDT as target variables, it was found that MLP did not perform well compared to other selected models. The results exhibit potential for algorithms like KNN and XGBoost to enable accurate liquidity risk forecasting. However, limitations exist regarding sample size and generalizability. Further research should validate findings across larger, more diverse banking samples. Further, more machine learning algorithms and hybrid models can be used for predicting liquidity risk. The results can be further validated using accuracy, F1 score, and confusion matrix evaluation metrics.

Overall, this paper aims to contribute to the growing body of knowledge pertaining to the use of AI and focuses explicitly on predicting liquidity risk with the help of machine learning algorithms. The results offer insights for developing liquidity early warning systems using machine learning models tailored to the Indian banking context. Moreover, regulatory bodies, central banks, and policymakers can benefit from this study to make the periodic assessment of a bank's liquidity risk.

References

Adalsteinsson, G. (2014). The Liquidity Risk Management Guide: From Policy to Pitfalls. John Wiley & Sons Ltd. https://doi.org/10.1002/9781118858035

Aksoy, B., & Boztosun, D. (2020). Comparison of Machine Learning Methods in Prediction of Financial Failure of Businesses in The Manufacturing Industry: Evidence from Borsa İstanbul. Anadolu Üniversitesi Sosyal Bilimler Dergisi, 20(4), 237–268. https://doi.org/10.18037/ausbd.845792

Al-Qudah, A. A., Hamdan, A., Al-Okaily, M., & Alhaddad, L. (2022). The impact of green lending on credit risk: evidence from UAE's banks. Environmental Science and Pollution Research. https://doi.org/10.1007/s11356-021-18224-5

Battaglia, F., & Mazzuca, M. (2014). Securitization and Italian banks' risk during the crisis. The Journal of Risk Finance, 15(4), 458–478. https://doi.org/10.1108/JRF-07-2014-0097

Chen, S. (2016). Detection of fraudulent financial statements using the hybrid data mining approach. Springer Plus, 5(89). https://doi.org/https://doi.org/10.1186/s40064-016-1707-6

Chi, G., Uddin, M. S., Abedin, M. Z., & Yuan, K. (2019). Hybrid model for credit risk prediction: An application of neural network approaches. International Journal on Artificial Intelligence Tools, 28(5), 1–33. https://doi.org/10.1142/S0218213019500179

Drudi, M. L., & Nobili, S. (2021). A liquidity risk early warning indicator for Italian Banks: A machine learning approach.

Dutta, I., Dutta, S., & Raahemi, B. (2017). Detecting financial restatements using data mining techniques. Expert Systems with Applications, 90(30), 374–393. https://doi.org/https://doi.org/10.1016/j.eswa.2017.08.030

Godlewski, C. J. (2006). Regulatory and Institutional Determinants of Credit Risk Taking and a Bank's Default in Emerging Market Economies: A Two-Step Approach. Journal of Emerging Market Finance, 5(2), 183–206. https://doi.org/10.1177/097265270600500204

Grover, P. (2021). Bad bank: A good alternative mechanism for resolution of the stressed assets. International Journal of Advanced Research in Commerce, Management & Social Science, 4(1), 213–218.

Guerra, P., Castelli, M., & Côrte-Real, N. (2022). Machine learning for liquidity risk modelling: A supervisory perspective. Economic Analysis and Policy, 74, 175–187. https://doi.org/https://doi.org/10.1016/j.eap.2022.02.001

Gupta, J., & Kashiramka, S. (2020). Financial stability of banks in India: Does liquidity creation matter? Pacific Basin Finance Journal, 64(September), 101439. https://doi.org/10.1016/j.pacfin.2020.101439

Gupta, N., & Mahakud, J. (2020). Ownership, bank size, capitalization and bank performance: Evidence from India. Cogent Economics and Finance, 8(1). https://doi.org/10.1080/23322039.2020.1808282

Hamal, S., & Senvar, Ö. (2021). Comparing performances and effectiveness of machine learning classifiers in detecting financial accounting fraud for Turkish SMEs. International Journal of Computational Intelligence Systems, 14(1), 769–782. https://doi.org/https://doi.org/10.2991/ijcis.d.210203.007

Khalid, S., Khan, M. A., Mazliham, M. S., Alam, M. M., Aman, N., Taj, M. T., Zaka, R., & Jehangir, M. (2022). Predicting Risk through Artificial Intelligence Based on Machine Learning Algorithms: A Case of Pakistani Nonfinancial Firms. Complexity, 2022. https://doi.org/10.1155/2022/6858916

Khurana, K. S. (2021). The Goodness within the Bad Bank in India. RGNUL Financial and Mercantile Law Review, 8(2).

Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data Mining techniques for the detection of fraudulent financial statements. Expert Systems with Applications, 32(4), 995–1003. https://doi.org/https://doi.org/10.1016/j.eswa.2006.02.016

Kou, G., Peng, Y., & Wang, G. (2014). Evaluation of clustering algorithms for financial risk analysis using MCDM methods. Information Sciences, 275, 1–12. https://doi.org/doi.org/10.1016/j.ins.2014.02.137

Leo, M., Sharma, S., & Maddulety, K. (2019). Machine learning in banking risk management: A literature review. Risks, 7(1). https://doi.org/10.3390/risks7010029

Liao, H. H., Chen, T. K., & Lu, C. W. (2009). Bank credit risk and structural credit models: Agency and information asymmetry perspectives. Journal of Banking and Finance, 33(8), 1520–1530. https://doi.org/10.1016/j.jbankfin.2009.02.016

Louzis, D. P., Vouldis, A. T., & Metaxas, V. L. (2010). Macroeconomic and Bank-Specific Determinants of Non-Performing Loans in Greece: A Comparative Study of Mortgage, Business and Consumer Loan Portfolios. In Working Paper (Issue September). https://doi.org/10.2139/ssrn.1703026

Masood, O., & Ashraf, M. (2012). Bank-specific and macroeconomic profitability determinants of Islamic banks: The case of different countries. Qualitative Research in Financial Markets, 4(2/3), 255–268. https://doi.org/10.1108/17554171211252565

Misman, F. N., & Bhatti, M. I. (2020). The Determinants of Credit Risk: An Evidence from ASEAN and GCC Islamic Banks. Journal of Risk and Financial Management, 13(5), 89. https://doi.org/10.3390/jrfm13050089

Mousa, G. A., Elamir, E. A. H., & Hussainey, K. (2022). Using machine learning methods to predict financial performance: Does disclosure tone matter? International Journal of Disclosure and Governance, 19(1), 93–112. https://doi.org/https:// doi. org/ 10. 1057/ s41310- 021- 00129-x

Omar, N., Johari, Z. 'Amirah, & Smith, M. (2017). Predicting fraudulent financial reporting using artificial neural network. Journal of Financial Crime, 362–387(24), 2. https://doi.org/https://doi.org/10.1108/JFC-11-2015-0061

Ravisankar, P., Ravi, V., Rao, G. R., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. Decision Support Systems, 50(2), 491–500. https://doi.org/https://doi.org/10.1016/j.dss.2010.11.006

Saleemi, J. (2014). An empirical analysis of cost-based market liquidity measures for U.S & Norwegian banks. Universitetet i Nordland, Bodø, Norway.

Salike, N., & Ao, B. (2018). Determinants of bank's profitability: role of poor asset quality in Asia. China Finance Review International, 8(2), 216–231. https://doi.org/https://doi.org/10.1108/CFRI-10-2016-0118

Tavana, M., Abtahi, A. R., Di Caprio, D., & Poortarigh, M. (2018). An Artificial Neural Network and Bayesian Network model for liquidity risk assessment in banking. Neurocomputing, 275, 2525–2554. https://doi.org/10.1016/j.neucom.2017.11.034

Wang, N. (2017). Bankruptcy Prediction Using Machine Learning. Journal of Mathematical Finance, 7(4), 908–918. https://doi.org/https:// doi. org/ 10. 4236/ jmf. 2017. 74049

Yao, J., Zhang, J., & Wang, L. (2018). A Financial Statement Fraud Detection Model Based on Hybrid

Data Mining Methods. International Conference on Artificial Intelligence and Big Data, 57–61. https://doi.org/10.1109/ICAIBD.2018.8396167