**LITERATURE REVIEW**

**AJAY MASADI**

**Background/Motivation**

This study investigates how machine learning (ML) can enhance the loan approval process within the banking industry. The problem of loan approvals presents considerable obstacles, leading to financial instability and a rise in non-performing assets. Current systems, which frequently depend on rule-based methods or simple statistical techniques, struggle to manage complex, multi-dimensional datasets efficiently. The study focus on significant deficiency in earlier research, particularly the absence of sophisticated ML methods utilized for predicting loan approvals. Numerous current studies have failed to use ensemble techniques that can more effectively capture complex patterns in data and enhance prediction accuracy. Filling this gap is vital, since enhancing the precision and dependability of loan prediction models directly affects banks. Through enhancing resource distribution and reducing risks, these models aid in improving decision-making processes. This research emphasizes showcasing the use of AdaBoosting, a well-known ensemble learning technique, in attaining enhanced predictive results.

**Methods Used**

The authors utilized several machine learning techniques, such as Logistic Regression, Decision Trees, Random Forest, and AdaBoosting. Preprocessing techniques were applied to ready the data for analysis. This involved managing absent values, encoding categorical variables, and normalizing numerical characteristics to guarantee consistency among features. AdaBoosting was chosen for its capacity to improve weak classifiers by merging them into a strong model. This renders it especially efficient for datasets that exhibit class imbalances, as it mainly targets on the most difficult instances throughout training. This study's main innovation is the combination of a well-organized preprocessing pipeline with AdaBoosting. Moreover, the complete evaluation process, which utilized metrics like accuracy, precision, recall, and F1-score, guaranteed a strongest assessment of model performance.

**Significance of the Work**

The results demonstrated that AdaBoosting significantly outperformed the other machine learning methods evaluated in the study, achieving an impressive accuracy rate of 99.99%. Its primary focus is on the method's potential to transform financial predictive modeling. The outcomes demonstrate the advantages of ensemble learning techniques over traditional single-model approaches. The study makes a compelling case for financial institutions to apply these strategies in order to enhance decision-making and reduce loan default rates. The study provides a valuable reference for upcoming investigations into machine learning applications in the banking industry by demonstrating the effectiveness of AdaBoosting.

**Connection to Other Work**

The study expands on previous work that used machine learning tools to evaluate financial risk. In contrast to earlier research that mostly concentrated on statistical methods or simple algorithms, this study emphasizes the enhanced effectiveness of ensemble approaches, especially AdaBoosting.This study shows that, in contrast to traditional research that typically employs individual classifiers, merging multiple models can yield a more readily predictive tool. A further level of rigor is added by the meticulous attention to preprocessing and evaluation in comparison to many previous investigations. Along with citing significant studies in ensemble learning, the authors highlight significant publications on the principles and applications of AdaBoosting. These citations provide background information for the study's methodology and demonstrate the study's significance in many sectors.

**Relevance to Capstone Project**

The methods and results of this research closely match the aims of my forthcoming capstone project, which centers on utilizing ML techniques to forecast loans. The implementation of AdaBoosting, specifically, offers a useful and efficient approach that can be applied to the project.

Various components from the paper, such as the preprocessing methods, ensemble learning strategy, and assessment metrics, can be incorporated into my future project work. In my project, I aim to investigate other ensemble techniques, including Gradient Boosting Machines (GBM) or XGBoost, to evaluate performance results.

**References**

Haque, F. M. A., & Hassan, Md. M. (2024). Bank Loan Prediction Using Machine Learning Techniques. *American Journal of Industrial and Business Management*. Retrieved from <https://www.scirp.org/journal/paperinformation?paperid=138439>