
ISSN: 2321-2152
IJMECE
International Journal of modern
electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

ISSN2321-2152

www.ijmece .com

Vol 12, Issue.2 April 2024

ENHANCING BANKING

Mr.M.Arokya Muthu[1]K.Sourabhya Reddy[2]R.Rishika Reddy[3] K.Sai Kiran[4]Y.Vara Laxmi[5]
Assistant Professor/CSE(DS) TKR College of Engineering and Technology Telangana, India Arokyamuthu@tkrcet.com
IV Final Year of CSE (DS) TKR College of Engineering and Technology Telangana, India Ksourabhyareddy@gmail.com
IV Final Year of CSE (DS) TKR College of Engineering and Technology Telangana, India Rishikareddy3999@gmail.com
IV Final Year of CSE (DS) TKR College of Engineering and Technology Telangana, India Ksaikiran129@gmail.com
IV Final Year of CSE (DS) TKR College of Engineering and Technology Telangana, India Yasamlakshmi2712@gmail.com

ABSTRACT

This paper emphasizes the crucial role of customer retention in the contemporary banking sector, recognizing the substantial expenses associated with acquiring new customers.startingwith the pressing issue of credit card fraud detection, this project employs machine learning algorithms to proactively identify fraudulent activities in real-time.Moving on to loan prediction, this project utilizes historical data and customer profiles to develop precise credit assessment models.By accurately evaluating, the bank can make informed loan approval, thereby minimizing risks and ensuring responsible lending practices.Nowadays, about 1.5 million customers switch banks each year, with numbers going up. They might switch for reasons like better services, lower fees, or convenient locations. To handle this, we use prediction models to guess who might leave in the future

Keywords:

Predicting churn, utilizing classification and clustering methods. Employing correlation attribute ranking and machine learning algorithms.Evaluating performance through ROC curve analysis in credit card fraud detection, emphasizing the use of deep learning algorithms.

1. INTRODUCTION

In the contemporary data-centric banking sector, customer retention and process optimization take center stage. This overview underscores the impact of data-driven approaches through three noteworthy studies. The initial study introduces a churn prediction model, enabling decision-makers with targeted strategies for customer identification and retention. The second study focuses on enhancing efficiency in loan approval processes by leveraging machine learning algorithms, benefiting both clients and institutions. Lastly, the third study employs deep learning algorithms for robust fraud detection, reinforcing security measures and minimizing financial losses. Collectively, these studies exemplify how advanced analytics and machine learning contribute to efficiency, innovation, and heightened customer satisfaction within the banking sector.

2. LITERATURE SURVEY

The literature on personalized recommendations, notably in e-commerce, push notifications, and movie suggestions, has found success with Collaborative Filtering (CF) methods. Implementing CF has notably improved user experience across various industries. The rise of mobile internet technology has facilitated the collection, analysis, and utilization of extensive travel scenario data, leading to a surge in research on context-sensitive mobile recommendations.

Context-aware recommendations, a growing research field, provide personalized services based on users' contextual information. Initially focusing on context prefiltering, researchers proposed strategies to align contextual information with recommendation systems. For example, multiple dimensions like date, time, weather, and companionship were used to suggest restaurants to mobile users, illustrating the broad applications of context-aware recommendations.

Recent research has shifted towards context modeling, user preference characterization, and recommendation enhancement. By incorporating scenario information into preference modeling and utilizing contextual similarity calculation, matrix decomposition, and nearest neighbor clustering techniques, researchers aim to refine recommendation systems for improved user satisfaction and engagement.

Simultaneously, industries are increasingly adopting data-driven decision-making, driven by the growth of big data and analytics capabilities. In the airline industry, big data analytics holds promise for optimizing aviation operations, enhancing customer service, and revenue management. Airfare price prediction, a complex task, has advanced with machine learning algorithms such as Linear Regression, Support Vector Machines, and Random Forests. Recent studies explore more sophisticated models like Deep Regressor Stacking to improve airfare price prediction accuracy, showcasing the continuous evolution of predictive modeling in the airline industry.

However, challenges persist in data availability and organization, particularly concerning airline ticket data. While datasets like T100 and DB1A/1B exist, the limited association between prices and specific flight information poses obstacles to comprehensive analysis. Researchers often rely on web crawling or private data sources, making replication and comparison challenging.

3. PROBLEM DEFINATION

In the contemporary interconnected business environment, banking organizations grapple with various challenges in effectively utilizing predictive analytics. Key focus areas encompass customer churn, loan approval processes, and fraud detection. While traditional methods like decision trees and logistic regression are prevalent, theyencounter difficulties in handling intricate scenarios and evolving data dynamics. Although feature selection techniques enhance model performance, the integration of diverse data sources remains a persistent challenge. While supervised machine learning techniques prove effective in fraud detection, the advent of technologies like deep learning opens up new avenues. The development of an integrated predictive analytics framework is paramount, capitalizing on the synergies between banking and fraud detection domains. Through the amalgamation of methodologies such as hybrid modeling, feature selection, and deep learning, organizations can drive data-driven decision-making, optimize operations, and effectively mitigate risks.

3.1 Limitationsofexistingsystem

While data-driven approaches hold promise, they encounter limitations. Churn prediction models may face challenges due to the dynamic nature of customer behavior and evolving factors influencing churn. Relying on historical data for model training may constrain adaptability to changing market conditions. In the realm of loan approval prediction, issues arise from data quality and potential bias in algorithmic decision-making, particularly regarding sensitive attributes like race or gender. Deploying machine learning for fraud detection requires careful consideration of false positive rates to avoid increased operational costs and customer inconvenience. Implementing these solutions demands substantial investment in infrastructure, talent, and ongoing maintenance, presenting financial and organizational challenges for some institutions. Despite these limitations, proactive management and continuous refinement can help mitigate risks and unlock the transformative potential of data-driven approaches in the banking sector

.

3.2 Proposed system

Upon reviewing various research papers, the decision is to employ widely-used algorithms suitable for the dataset. The proposed integrated predictive analytics system serves as a comprehensive solution to address challenges in customer churn prediction, loan approval automation, and credit card fraud detection. Through advanced machine learning techniques and meticulous data preprocessing, the system ensures data quality and relevance in all three domains. Data preprocessing involves noise removal, handling missing values, and feature selection using methods like Information Gain and Correlation Attributes Ranking Filter, retaining only the most predictive attributes for further analysis.

In the customer churn prediction domain, the system utilizes classification algorithms such as Random Forest, Decision Stump, and Logistic Regression to categorize customers into churn and non-churn categories. Analyzing transactional behavior patterns, the system identifies potential churners and suggests personalized retention strategies for enhanced customer loyalty. For loan approval automation, the system assesses eligibility using machine learning algorithms like Logistic Regression, Decision Trees, and Random Forest. By scrutinizing features like loan amount and dependents, the system streamlines the approval process and enhances decision-making in banking institutions.

In credit card fraud detection, the system employs advanced techniques like Convolutional Neural Networks (CNN) and Support Vector Machines (SVM) to identify fraudulent activities. Analyzing transactional data and customer behavior patterns allows the system to detect suspicious transactions, minimizing financial losses for credit card companies. The integrated predictive analytics system furnishes organizations with actionable insights and recommendations, optimizing operations, improving customer satisfaction, and effectively mitigating risks in customer churn prediction, loan approval automation, and credit card fraud detection domains.

IJMECE
International Journal of modern electronics and communication engineering

ISSN2321-2152

www.ijmece .com

Vol 12, Issue.2 April 2024

4. FIGURES

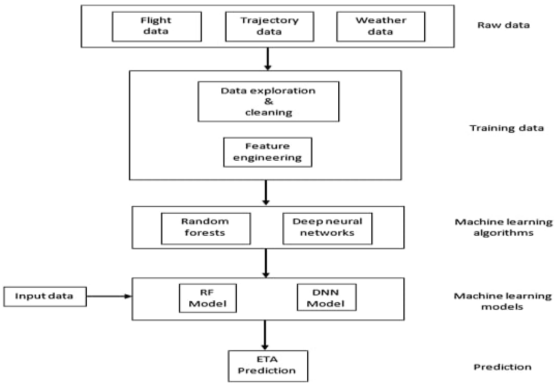

Figure 1: SYSTEM ARCHITECTURE

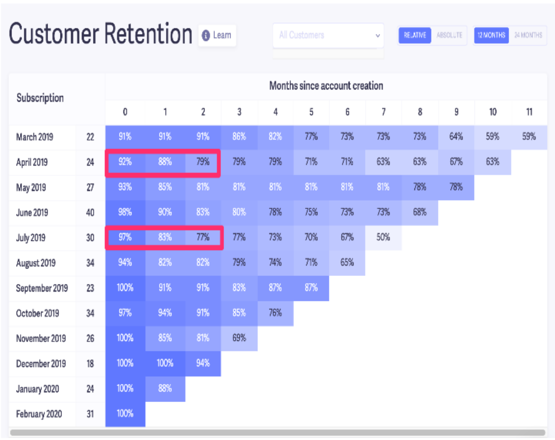

Figure 2 Analysis on Different Attributes for Churn Analysis

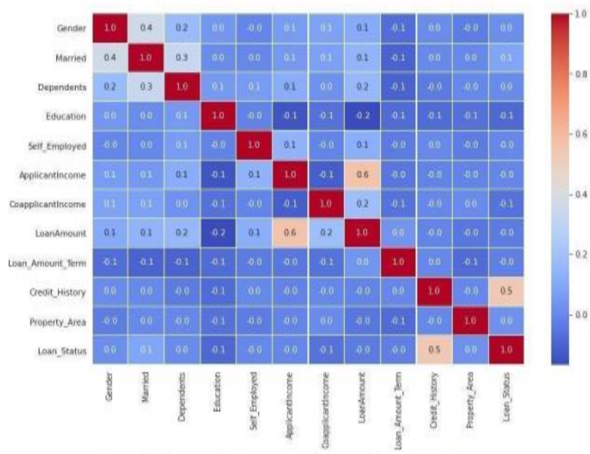

Figure 2: The correlation among all the attributes in the dataset (https://www.kaggle.com/code/vinodkumargr/loan-prediction)

Fig. 3 Co- Relation Map

.

5. MODULES

Customer Churn Prediction:

The proposed model aims to categorize customers into churn and non-churn categories through data preprocessing and clustering techniques.

It utilizes recommender systems to personalize retention strategies based on customer behavior and preferences.

The model employs similarity-based approaches to target similar customers and offers loyalty programs or special packages to enhance retention.

ISSN2321-2152

www.ijmece .com

Vol 12, Issue.2 April 2024

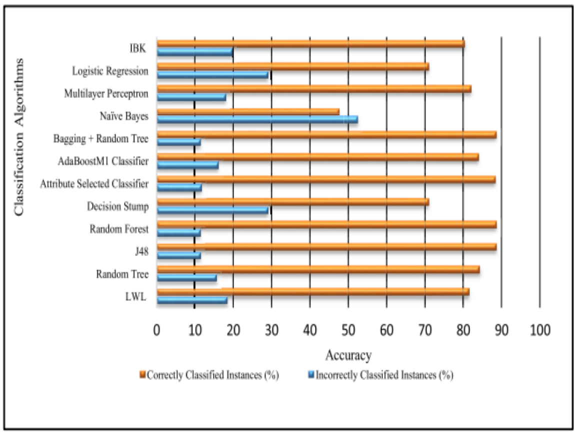

FIGURE 2. Accuracy performance of classification algorithms on the churn-balance dataset.

Fig 5.1 Accuracy of Model's

Loan Approval Prediction:

The model utilizes machine learning algorithms such as Linear Regression, Logistic Regression, Support Vector Machine, and Naïve Bayes for loan approval prediction.

Features like income, credit score, and loan amount are analysed to streamline the approval process and improve decision-making in banking institutions.

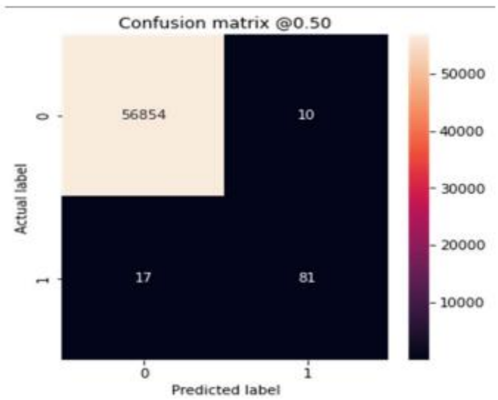

Fig 5.2 Confusion Matrix

Baseline Model	Epoch 35		Epoch 14	
Metrics	Training Accuracy (%)	Validation Accuracy (%)	Training Accuracy (%)	Validation Accuracy (%)
Precision	93	42	91	89
Recall	90	85	80	68
AUC	98	97	94	95
PRC	56	22	84	80
Accuracy	98	96	99	99

Fig 5.3 Accuracy of Model's

Credit Card Fraud Detection:

The model leverages machine learning methods like Decision Trees, K-nearest Neighbours, and Random Forest for fraud detection.

It analyses transactional data and customer behaviour patterns to identify suspicious activities and minimize financial losses for credit card companies.

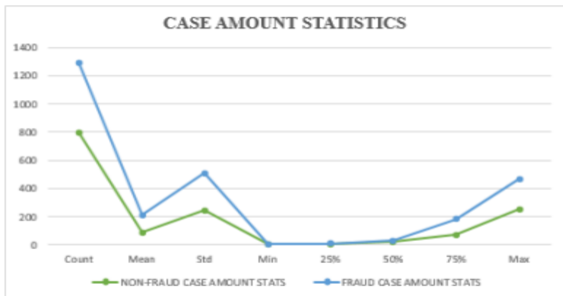

FIGURE 7. The case count statistics for fraud and non-fraud transactions.

Model Name	Existing Accuracy (%)	New Accuracy (%)
CNN	93.00 94.6	96.34
BL	83	99.72
RF	97.55 92.3	99.92
SVM	97.43	99.93
KNN	93.27 91.11	99.91
DT	97.08 66.5	99.93
LR	97.18 67.8	99.91

Utilized to improve prediction accuracy by averaging results from multiple decision trees.

By integrating these models and techniques, organizations can enhance their operational efficiency, improve customer satisfaction, and mitigate risks effectively across customer churn prediction, loan approval automation, and credit card fraud detection domains.

6. Deployment

Flask and Streamlit are both popular frameworks for web development. Flask, a lightweight Python web framework, offers flexibility and control for building versatile web applications, including RESTful APIs and dynamic web pages. It's favored for its simplicity and scalability, suitable for both simple projects and complex applications. On the other hand, Streamlit is a Python library designed for data science and machine learning practitioners. It simplifies the creation of interactive web applications by abstracting away the complexities of web development, making it ideal for rapidly prototyping and sharing data-centric applications. While Flask provides general-purpose flexibility, Streamlit excels in streamlining the development of data-centric web applications.

7. ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Mr.M.Arokia Muthu sir for their invaluable guidance, mentorship, and support throughout this machine learning project. Their expertise, insightful feedback, and encouragement have been instrumental in shaping the direction and outcomes of this endeavor.

I am deeply thankful to K.Sourabhya for their collaboration, dedication, and contributions to various aspects of the project. Their collective efforts and teamwork have enriched the project's development and fostered a conducive environment for innovation and learning.

I extend my appreciation to [TKR College of Engineering & Technology] for providing the necessary resources, facilities, and opportunities that facilitated the execution of this project. Their continued support has been pivotal in our pursuit of excellence and achievement of project milestones

8. REFERENCES

[1]. S. Babu, D. N. Ananthanarayanan, and V. Ramesh, ‘‘A survey on factors impacting churn in telecommunication using datamininig techniques,’’ Int. J. Eng. Res. Technol., vol. 3, no. 3, pp. 1745–1748, Mar. 2014.

[2]. C. Geppert, ‘‘Customer churn management: Retaining high-margin customers with customer relationship management techniques,’’ KPMG &Associates Yarhands Dissou Arthur/Kwaku Ahenkrah/David Asamoah, 2002.

[3]. W. Verbeke, D. Martens, C. Mues, and B. Baesens, ‘‘Building comprehensible customer churn prediction models with advanced rule induction techniques,’’ Expert Syst. Appl., vol. 38, no. 3, pp. 2354–2364, Mar. 2011.

[4]. Y. Huang, B. Huang, and M.-T. Kechadi, ‘‘A rule-based method for customer churn prediction in telecommunication services,’’ in Proc. Pacific– Asia Conf. Knowl. Discovery Data Mining. Berlin, Germany: Springer, 2011, pp. 411–422.

[5]. Awodele, O., Alimi, S., Ogunyolu, O., Solanke, O., Iyawe, S., &Adegbie, F. (2022, November).

[6]. Cascade of Deep Neural Network And Support Vector Machine for Credit Risk Prediction. In 2022 5th Information Technology for Education and Development (ITED) (pp. 1-8). IEEE.

[7]. In 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART) (pp. 423-426). IEEE. https://jupyter.org/ [Last accessed on 15-06-2023]

[8]. Y. Abakarim, M. Lahby, and A. Attioui, ‘‘An efficient real time model for credit card fraud detection based on deep learning,’’ in Proc. 12th Int. Conf. Intell. Systems: Theories Appl., Oct. 2018, pp. 1–7, doi: 10.1145/3289402.3289530.

[9]. H. Abdi and L. J. Williams, ‘‘Principal component analysis,’’ Wiley Interdiscipl. Rev.,

[10]. V. Arora, R. S. Leekha, K. Lee, and A. Kataria, ‘‘Facilitating user authorization from imbalanced data logs of credit cards using artificial intelligence,’’ Mobile Inf. Syst., vol. 2020, pp. 1–13, Oct. 2020, doi: 10.1155/2020/8885269.

[11]. J. Błaszczyński, A. T. de Almeida Filho, A. Matuszyk, M. Szelg¸, and R. Słowiński, ‘‘Auto loan fraud detection using dominance-based rough set approach versus machine learning methods,’’ Expert Syst. Appl., vol. 163, Jan. 2021, Art. no. 113740, doi: 10.1016/j.eswa.2020.113740.

[12]. B. Branco, P. Abreu, A. S. Gomes, M. S. C. Almeida, J. T. Ascensão, and P. Bizarro, ‘‘Interleaved sequence RNNs for fraud detection,’’ in Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2020, pp. 3101–3109, doi: 10.1145/3394486.3403361