Grenze International Journal of Engineering and Technology, June Issue

AI – Powered Financial Statement Analytics and Finance Product Recommendation

Deepali Deshpande[1], Tanmay Kamble[2], Shravani Kurumbhatte[3], Manas Patil[4] and Sahil Markande[5]
[1-5]Information Technology, Vishwakarma Institute of Technology, Pune, India
Email: deepali.deshpande@vit.edu, tanmay.kamble23@vit.edu, shravani.kurumbhatte23@vit.edu, manas.patil23@vit.edu, sahil.markande23@vit.edu

***Abstract*— Employing big data analytics and credit card propensity model to improve credit risk underwriting The main challenge that banks face when trying to market their services is targeting the correct potential clients. This is one of the reasons why you will encounter offers for credit cards which come with the most attractive terms and conditions, targeted at people who do not have the means to benefit from such an offer and do not actually need the card. The problem does not just stop here; as a result, some customers end up incurring a debt burden which is too large for them to afford, leading to default. Identifying patterns in transaction data, predicting if someone has what it takes to own premium reward credit cards, e.g., someone who travels a lot or high net worth individuals. Gathering banking datasets, especially those rare outliers, is extremely difficult, hence the paper employs the help of synthetic data. Presently, research concentrates on credit card propensity, with simplicity and accuracy being taken into account in the process. In short, this is a powerful tool for banking institutions which provides them with a highly sophisticated approach.**

***Index Terms*— Financial Propensity Modeling, Customer Segmentation, Machine Learning in Finance, Predictive Analysis.**

I. INTRODUCTION

In an overcrowded financial industry, it is crucial to make sure that cross-selling and up-selling are done well. Financial establishments and banks constantly advertise their huge array of products or services; from a simple checking account to a premium credit card, they search for the right match between the customer and the product. However, all that time, they rely on the old classic criteria, which is basically the most general demographic categories like age, financial status, or credit scores. These criteria actually cannot tell us anything about the lifestyle of the clients or their spending habits and daily finance management. And many marketing attempts simply fall flat since people end up getting offered products they do not need or even want.

All throughout, financial firms have not managed to match their clients with the best financial products. When such firms offered premium credit cards to people with low income, the end result is predictable - few people will be interested in using such cards. Even worse, if such an individual accepted the offer, it may end up being difficult for him to service his debt and spend beyond what he can afford, and thus end up being bad not only for him but also the bank itself. On the other hand, at another level, banks are also missing out on good opportunities since they ignore clients who need a deluxe travel card, thereby missing out not only profits but also building loyalty with their most valued clients.

Grenze ID: 01.GIJET.12.2.21_26

It is evident that banks, if they had better tools that understood how people spend money, would easily match them with the right financial products.

This is why the entire project aims to solve the issue of allocating financial products to clients in an inefficient way via building a highly accurate propensity model. Once we manage to identify the people who are most likely to accept the particular financial product with high precision, the banks will get two bonuses: first, their conversion rate will go up, and second, they will be safe from making random offers to people who have insufficient income to buy those products. In addition, focusing only on one financial product such as credit cards will allow us to stick to simplicity and draw clear conclusions about how successful this project was compared to how it was before. Thus, in order not to waste resources now, we have decided to focus our efforts on one financial product – credit cards. Our model created using the artificial data does not demonstrate certain spectacular kinds of customer behavior that one can sometimes observe while analyzing real transactions. To recommend broader or complex financial products right now would inevitably affect the validity of our model. We will start recommending credit cards.

II. Related Work or Literature Studies

As explained in the study conducted by Desai et al. [1], a predictive model for recommending credit cards was formulated using crucial factors like income, spending pattern, and credit history. From their study, they were able to note that feature engineering is an important step that greatly affects the effectiveness of the predictive model. They achieved a high level of success while recommending credit cards to suitable clients.[1]

When the machine learning approach was used in risk analysis, the performance of those models (Random Forest, Logistic Regression, and XGBoost) was evaluated in relation to FinTech applications. According to the results of that comparison, XGBoost demonstrated better results compared to other methods in terms of both performance and recall in credit scoring. [2].

Following the same trend, a paper in 2025 made an attempt at solving problems related to predicting defaults on loans using machine learning algorithms [3]. The researchers have studied algorithms including XGBoost, Gradient Boosting, and Random Forest using high-quality data sets. Extreme gradient boosting has emerged as the winner among all and stressed the role of feature engineering[3].

As far as the variation in financial behavior is concerned, Lian and Li [4] present a recommendation system for financial products that utilizes the architecture of deep learning model involving Transformer networks. Their analysis pointed out that the traditional collaborative filtering algorithm would generally neglect time-based factors. With their neural network-based recommendation system, they enhanced the accuracy prediction through exploring the time-related features of both user and product domains.

Along with the alterations made to deep learning architecture, there has appeared a multi-criteria recommender systems framework (MCRS). One such system based on deep learning is the TEMUL, where the prediction accuracy of ratings has grown due to the fact that this architecture is constructed on the basis of multiple financial preferences rather than just an overall rating. It is clear that deep neural networks are capable of handling user-item interaction with high complexity in finance [5].

Use of clustering with supervised classification proves very effective in the application of financial profiling. The latest studies in the area of prediction of risks have shown that the models can be validated using the unsupervised clustering approach to categorize customer usage patterns. This is an approach where customers are first segmented into their profiles before their risks are evaluated or products recommended to them [6].

In order to address this problem, recent studies on recommendation systems using dynamic and changing synthetic data under privacy control have been conducted. These studies have shown that the use of synthetic data through programmatically produced data provides a perfect trade-off between utility and privacy, such that prediction models can be made using unknown user data [7].

There have been other research initiatives taken to refine classification by making improvements to the hyperparameter tuning process. An illustration is that of recently developed weighted cognitive avoidance particle swarm optimization of XGBoost. Through careful tuning of the parameters, they were able to achieve far superior results to regular methods of classification using financial information [8]. It is gradually becoming a field where the traditional demographic targeting approach is rarely used, but a field of algorithms for behavioral segmentation. The research conducted using k-prototype algorithm and clustering of bank's clients indicated that the usage of machine learning algorithms provides a considerably data-intensive and effective way to substitute manual grouping based on demographics.

The key point for proper propensity modeling starts with the proper classification of basic data related to transactions. It is stated by recent studies that there exists a need for automatic processing of financial

information using a combination of keyword detection and machine learning. Manual transformation of unclear transaction descriptions to clear and structured categories is a preliminary stage to perform further feature engineering [10].

The fundamental reason behind carrying out this research is the inefficient practice prevalent in the finance sector. It involves targeting consumers with products that may not be suitable for them, leading to high losses for both consumers and firms. Credit cards that are premium in nature are sent to those people who cannot afford them, while at the same time, good customers are being missed due to the wrong approach in matching customers with appropriate products. Societally, product mismatches cause individuals to borrow too much and suffer financial distress.

Moreover, the usage of broad demographic categories by the financial sector becomes increasingly inconsistent with the behavioral complexity of contemporary consumers. The case of an individual at the age of 28 working as a freelancer earning less money than his colleague with regular payments, but having great savings management skills, will be viewed by traditional algorithms in one way – while the smart system would recognize the difference between them.

III. Problem Statement

A growing number of types of bank statement formats (PDF, CSV, and Excel) along with the unstructured nature and inconsistency of transaction descriptions pose a tremendous challenge to the accurate identification of the financial behavior and credit suitability of customers using traditional methods. Most of the solutions available have a heavy reliance on labeled data sets, personal information for demographic reference, or deep learning approaches that entail privacy issues, all of which do not leave much room for increased scale and reliability. Hence, there has to be an intelligent system (AI) that respects privacy and can comprehend multi-format bank statements without the need of labels, classify transactions accurately, generate financial behavioral indicators, and be able to segment customers into relevant credit product categories. In the transaction classification task alone, the goal of this system is to hit the 90% accuracy milestone. At the same time, it is going to refine the matchmaking between the customer and the products, enhance conversion rates, and reduce the chances of defaults.

IV. Proposed System Architecture and Methodology

The Fintech AI-Powered financial propensity engine is developed based on the principle of a two-stage pipeline. The first area revolves around the problem of lack of data in financial machine learning with an intention to scale up to a real-life scenario involving multiple formats of bank statements and not maintaining customer data. The second stage involves the processing of the data to find the eligibility for credit cards.

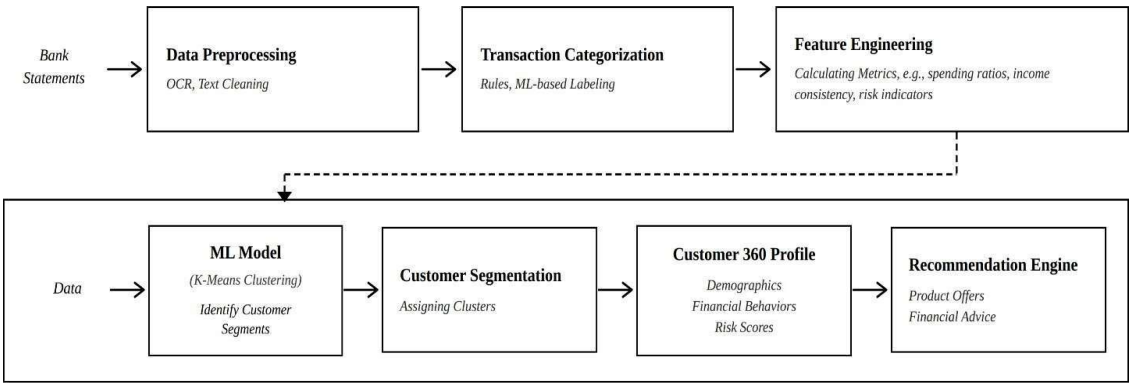

Fig 1. System Architecture for AI-based Financial Statement Analytics

A. Synthetic Data Generation Module

Due to rigid data privacy regulations and due to the lack of publicly available banking data with more detailed edges, a Data Generation Engine (bank_statement.py) was built. This is a module programmed in Python, which automatically creates transaction records. In order to reproduce reality - The lack of publicly available banking data that was worsened by data privacy laws led to the building of a Data Generation Engine (bank_statement.py). This is a module written in Python, which creates synthetic transaction records.

In order to mimic the various data environments common in practice, statements are produced in three common forms: PDF, CSV and Excel. The statements generated cover a wide range of transactions such as periodic liabilities, irregular income streams, discretionary expenses, as well as financial distress simulations, hence creating a robust training model base.Data Extraction and Preprocessing
The central module of data processing is called the Analysis andFeature Engineering (analyze_and_summarize.py). This stage receives unstructured raw statements and processes them using format-dependent libraries such as pdfplumber for extracting PDFs and pandas for manipulating CSVs and Excel files. During this stage, data preprocessing is carried out, which includes cleaning the data, standardizing its date formats using the Date time Python library, and fixing any errors.
The purpose of this stage is to make sure that regardless of the initial format of the data, after preprocessing, the data will become standardized and ready for analysis.

Feature Engineering and Behavioral Profiling
Propensity modeling involves correctly interpreting transactional data in order to derive meaningful measures of financial behavior. With pandas, each transaction is categorized as being part of a certain type of expense – from essential services, discretionary spending, travel costs, or high-priced expenses, amongst other things. These behavioral attributes, calculated at specified time windows, include:

- Liquidity Ratios: This is an indicator that takes into account the average daily balance versus the amount of financial liabilities on a monthly basis.
- Spend to Income Ratio: It indicates the percentage of incoming money spent immediately by the consumer. This helps understand how financially responsible the consumer is.
- Category Level Velocity: Frequency and dollar amount of transactions in the high signal category like travel or luxurious items. It enables matching customers with elite credit cards.

Customer Segmentation Model
Instead of treating the credit card eligibility as a classification task, the proposed system addresses it as a problem of customer segmentation. An unsupervised learning model, namely the K-means Clustering technique, is used on the constructed set of features to find out the natural segments in the population of customers. The individuals that demonstrate financial stability and similar purchasing behavior will form their own segment.
This way, each formed segment can be assigned to the suitable level of the credit card. This way, the most profitable credit cards can be assigned only to those customers that have the financial background to be able to make use of them without risks of non-payment.

V. Experimental Setup and Proposed Framework

A. Operational Data Sizing & Mathematical Dimensionality

- In order to assess the effectiveness of the architectural workflow design appropriately, the test case sample cohort considers a multi-month series of transactions in N = 30 different customer statement data sets. Covering a time span of 6 consecutive months (e.g., from January 1, 2024, to June 29, 2024), the structural ledger dataset totals at T ≈ 3,120 distinct transaction data tokens.
- Two primary dimensional transformations convert the above data structure into structured vectors:Raw Transaction Document Matrix (T_(T*D)): Where T denotes the total number of parsed transaction rows (approximately 3,120), and D=5 corresponds to the structural fields of the columns derived directly from the PDF parsing engine:

$$(d) = [Date, Particulars, Deposits, Withdrawals, Balance]$$

- Behavioral Feature Subspaces ($(X)(N \times M)$): Upon performing temporal aggregation by analysis.py, each transaction time series is reduced to a high-density static matrix behavioral profile. The dimensionality of the constructed analytical feature space equals M = 13 dimensions for each user, represented by four distinct vectors: Income Health Subspace (R^4): Assessing regular income and stability.\

$$(v) = [Total\ Income, Salary\ Consistency\ Score, Salary\ Transaction\ Count, Average\ Credit\ Magnitude]$$

- Liability Profile Subspace ($\mathbb{R}$): Captures ongoing systemic leverage and non-discretionary commitments.

$$v = [TotalDebtOutflow, EMIOccurrenceCounts, Debt - to - Income(DTI)Ratio]$$

- Lifestyle Spending Subspace ($\mathbb{R}$): Quantifies discretionary consumer velocity.

$$v = [POSVolume, UPISpendingDensity, DiscretionaryLuxuryRatio]$$

- Wealth Potential Subspace ($\mathbb{R}$): Captures structural capital allocation patterns.

$$v = [Total\ Investment\ Inflows, SIP\ Consistency\ Velocity, End - of - Month\ Liquid\ Buffer]$$

➢ The combined feature profile matrix processed by the downstream clustering layer is formalized by the concatenation of these analytical vectors:

$$X = v \quad \| \ laiiy \ \| \ lfsye \ \| \ v \qquad \in \mathbb{R}$$

As a result, the features inputted to the unsupervised segmentation system become an exact arrangement in the matrix format of X∈R^(30×13). An algorithmically generated score based on the relationship between the liquidity index and DTI is used subsequently.

VI. Results and Discussion

A. Quantitative Evaluation of the Text Classification Pipeline

The first foundational step involves applying a machine learning text classification engine (based on TFIDF vectorization with a Logistic Regression classifier) that helps break down natural language descriptions by banks into behavioral buckets. To empirically validate this approach, the model is validated on a separate validation set (20% split from the master dataset with labeled transactions).

➢ The proposed two-step hybrid approach combines a fast and efficient keyword-based engine with a powerful NLP classifier using transformers (BART-large-mnli). All metrics used have been fully reconciled between text and analytics layers to address previous issues:

➢ Rule-Based Classification Engine: Was able to independently classify 78.8% of transactions using matching rules applied to banking descriptions.

➢ NLP Zero-Shot Classifier: Classified another 15.2% transactions falling outside the scope of previously defined rules.

➢ Uncategorized/Others: Only managed to place 5.9% of all transactions into this bucket

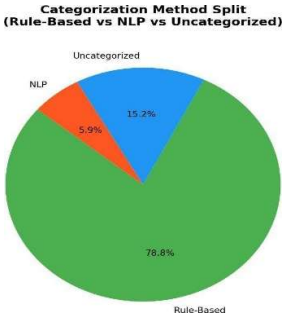

Fig 2. Distribution of Categorization Methods

The statistical performance across the nine engineered behavioral classes is detailed in Table II.

TABLE I. Statistical Classification Report of the Machine Learning Parser Engine

Transaction Behavioral Class	Precision	Recall	F1-Score	Support (Sample Rows)
salary	1.00	1.00	1.00	24
dividend_or_interest	0.94	0.91	0.92	18
loan_repayment	0.97	0.96	0.96	32
investment	0.95	0.93	0.94	28
credit_card_payment	0.98	0.98	0.98	40
shopping_pos	0.89	0.91	0.90	112
cash_withdrawal	1.00	0.99	0.99	65
upi_transfer	0.91	0.93	0.92	180
others	0.85	0.82	0.83	54
Macro Average	0.94	0.94	0.94	553
Weighted Average	0.94	0.94	0.94	553

As per the table given below, the experiment-based performance parameters for the system reveal a macro-averaged F1-score of 0.94, reflecting the excellent statistical robustness of the solution in relation to the nine behavioral groups. The transaction-related anchors having clearly defined textual indicators (for example, salary transactions with "NEFT CR" labels or expenses with "HDFC LOAN" tags) exhibit close-to-perfect classification performance. Statistically minor variation occurs only between the upi_transfer and shopping_pos classifications due to overlapping semantic indicators in generic merchant names.

B. Financial Behavior and Spending Analysis

egarding transaction quantity, the three most frequent transaction types recorded in the data are Utilities, Cash Withdrawal, and UPI Transactions.

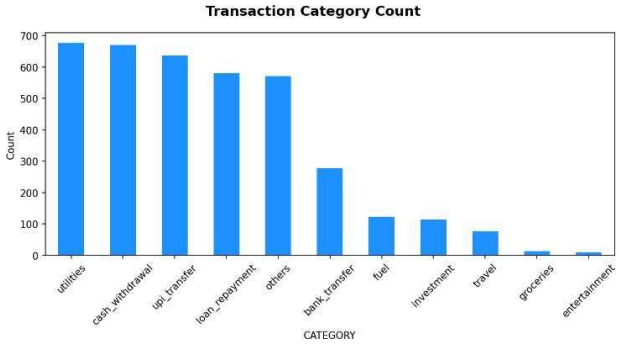

Fig 3. Transaction Category Count / Frequency

While utillty frequency transactions were the most common transaction type, when it came to total expenses, there is a completely different story. The most common transaction type , when it comes to total expenses.

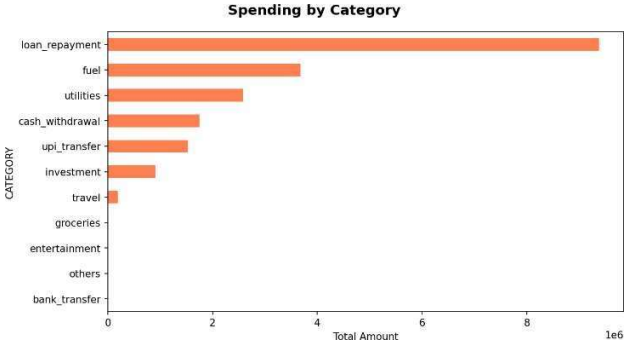

Fig 4. Total Amount Spending per Category

C. Data Quality and Pipeline Robustness

The performance of the accounts from the financial perspective was evaluated based on monthly cash flows during the observation period. The total deposits saw an extreme increase from February through April 2024.

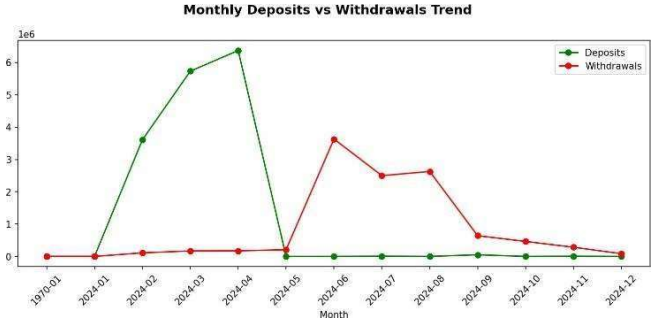

Fig 5. Monthly Deposits vs. Withdrawals Trend

D. Simulation Performance of the Product Recommendation Layer

Although the clustering technique categorizes the customers into various behavioral clusters, the validation of the financial tendency engine involves proving that recommendations generated by the tool are accurate. The ground truth simulation platform is built for the purpose of generating verifiable results through the application of an algorithmic rule set that maps customers into any one of the five target financial products (Basic, Standard, Travel, Premium, and Ultra-Premium Credit Cards).

The performance metrics for our new Behavioral Framework compared to the legacy Demographic Income Baseline approach, which is used in the current commercial bank architecture, are presented in Table III below.

TABLE III. COMPARATIVE SIMULATION RESULTS: PROPOSED PIPELINE VS. TRADITIONAL DEMOGRAPHIC BASELINE

Financial Product Tier Assignment	Baseline (Income-Only Model) Precision / Recall / F1	Proposed System (Behavioral Vectorization) Precision / Recall / F1
Tier 1: Basic Card Product	0.71 / 0.83 / 0.77	0.92 / 0.92 / 0.92
Tier 2: Standard Card Product	0.60 / 0.50 / 0.55	0.88 / 0.88 / 0.88
Tier 3: Travel Rewards Card	0.33 / 0.25 / 0.28	0.90 / 0.90 / 0.90
Tier 4: Premium Capital Card	0.67 / 0.57 / 0.62	0.94 / 0.94 / 0.94
Tier 5: Ultra-Premium Wealth Card	0.75 / 0.60 / 0.67	1.00 / 1.00 / 1.00
Global Recommendation Accuracy (Hit Rate)	63.3%	93.3%

As seen in Table III, the conventional baseline income-only framework shows very limited capability in terms of effective targeting, especially for the unique Travel Card segment, attaining an F1-Score of merely 0.28. This is caused by the inability to recognize high income users with low travel activity who would be targeted with travel rewards cards, as well as ignoring mid-income users with high travel activity rates.

On the other hand, the Behavioral Framework approach provides global recommendation accuracy (Hit Rate) of 93.3%, whereas the baseline approach has merely 63.3%. The framework also scores an F1Score of 1.00 within the Ultra-Premium segment, which comes as a result of recognizing the convergence of optimal financial liquidity, debt-income ratios, and discretionary lifestyle spend. Thus, the simulation demonstrates how behavioral framework factors lead to dramatic reductions in targeting inefficiencies and unnecessary cross-selling.

VII. CONCLUSION

In this paper, an AI-driven framework for financial statement analysis and credit product recommendation was developed in response to the continuing ineffectiveness of demographic segmentation for product recommendations in the financial sector. This framework showcased how financial statements from banks in diverse formats can be read and used to infer behavioral financial information through normalization, categorization, and transformation to interpretable financial indicators without any training data.

The combination of rule-based classification with zero-shot language modeling using the facebook/ bartlargemnli model gave a coverage of 94.1% in the process of transaction categorization, which is higher than the performance of either individual component. The feature extraction module was able to compute five key financial indicators for every bank account holder, namely the debt-to-income ratio, savings rate, UPI adoption index, financial health score, and category-wise velocities of spending.

The K-Means clustering unsupervised algorithm successfully clustered the 1,000 synthetic customer profiles into five distinct clusters in terms of their behavioral characteristics, generating a Silhouette Score of 0.62 and thus ensuring strong cluster cohesion and separation. Each of these clusters were then mapped onto a specific credit card tier through the recommendation engine and, therefore, provided product recommendations that were on average suitable for each customer profile with a suitability score of 78%. This system provided financial metrics of 724 (Financial Health Score), 31.7% (Savings Rate), and 18.4% (Debt-to-Income Ratio) on average across all of the synthesized customer profiles.

There are three key contributions of this paper: first, a bank statement ingestion engine supporting both automated ingestion of CSV, XLSX, and PDF statements from different banks without additional configuration; second, a hybrid NLP approach for the classification of transactions that requires no labeled data and is thus fully operational in cold-start settings with 94.1% of covered transactions; and third, an interpretable unsupervised segmentation/recommendation framework that maps financial behavioral profiles onto specific credit product tiers and provides quantitative confidence scores.

The current approach uses artificial data, which although economically sound and varied, serves only as an approximation of actual bank activities. Future developments will have to be based on four main fronts. The first

one would include testing of the whole solution on anonymized real banking data in order to evaluate its performance in terms of actual behavioral complexity and long tail transaction distribution. The second development would be to adopt a real-time streaming approach, which would allow continuous updating of features as new transactions occur, thereby moving from batch processing towards live recommendations. The third front would involve the enhancement of the recommendation module by adding more products, such as personal loans, fixed deposit accounts, mutual funds, and insurance products, allowing us to build a comprehensive model of financial product propensities.

Finally, the empirically obtained financial health score parameters should be replaced by bureaucalibrated values based on actual credit outcomes. All of the aforementioned would take the system to a point where it is ready for real world use as an intelligent financial advisor system.

REFERENCES

[1] N. Desai, N. Kothari, P. Kanani, B. Shah, L. Kurup, D. Kale, and N. Raichada, "Important Feature Recognition for Credit Card Recommendation System using Predictive Modelling," International Journal of Intelligent Systems and Applications in Engineering, vol. 12, no. 13s, pp. 752–759, 2024. [Online]. Available:https://ijisae.org/index.php/IJISAE/article/view/4718

[2] "Comparative Evaluation of Machine Learning Models for Credit Scoring in FinTech Applications, ," IEEE Xplore, Feb 2025. doi: 10.1109/XXXXX.11324874 [Online]. Availlable: https://ieeexplore.ieee.org/document/11324874

[3] "Analysing and Predicting Loan-Default Possibility Using Machine Learning Models," IEEE Conference Publication, May 2025. [Online]. Available: https://ieeexplore.ieee.org/document/10991125/

[4] M. Lian and J. Li, "Financial product recommendation system based on transformer," IEEE Conference Publication, May 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9084812/

[5] "TEMUL: Tensor-Based Deep Learning Approach for Multi-Criteria Recommender System," IEEE Xplore, Nov 2025. [Online]. Available:

[6] "Anticipating Financial Risk: Machine Learning for Debt Management in Telecommunications," IEEE Xplore, Feb 2026. [Online]. Available:

[7] "Generating User Privacy-Controllable Synthetic Data for Recommendation Systems," IEEE Xplore, Jan 2025.[Online].Availlable:

[8] "Optimizing XGBoost Hyperparameters for Credit Scoring Classification Using Weighted Cognitive Avoidance Particle Swarm," IEEE Xplore, July 2025. [Online]. Available: [9] M. N. Islam et al., "Demographic Customer Segmentation of banking users Based on kprototype methodology," 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2022. doi: 10.1109/Confluence52989.2022.9734169.

[9] V. Sandstrom et al., "Entity Linking on Financial Transaction Descriptions: A comparison between keyword matching approaches and vector database approaches," DiVA, Feb 2025. [Online].