International Journal of Artificial Intelligence, Data Science, and Machine Learning
Grace Horizon Publication | Volume 3, Issue 2, 90-100, 2022
ISSN: 3050-9262 | https://doi.org/10.63282/3050-9262.IJAIDSML-V3I2P111

Original Article

A Comprehensive Analytical Framework for Modeling Consumer Credit Card Behavior and Risk Profiling Using Advanced Financial Metrics

Amandeep Singh Arora[1], Thulasiram Yachamaneni[2], Uttam Kotadiya[3]
[1]Senior Engineer I, USA.
[2]Senior Engineer II, USA.
[3]Software Engineer II, USA.

Abstract *- The conception of consumer behavior within the framework of credit card usage plays a central role in the effort of financial institutions to analyze risk and best practices in their customer relationship management initiatives. This paper offers an overall analytical construct to the modeling of consumer credit card behavior and risk profiling based on advanced financial measurements. Based on multidimensional data and using state-of-the-art statistical and machine learning techniques, the framework encloses behavioural segmentation, credit risk scoring and predictive analytics. Our model illustrates the way financial institutions could attain the greater stratification of risks and individualization of their propositions through incorporating transactional, demographic, and psychographic variables. This model employs the best methods (both supervised and unsupervised) of finding meaningful patterns in extensive data through the implementation of clustering approaches in the segmentation of consumers and logistic regression and gradient boosting machines in predicting risk. Such financial measures as the debt-to-income ratio, credit utilisation, and payment Records are inherent parts of the analysis process, increasing the precision and relevance of the model. The study also comes up with a new Composite Risk Index (CRI), which is constructed by using weighted financial indicators that are adjustable according to market conditions. We tested our implementation on a database of data obtained by top credit card issuers before May 2022, and we found a large difference in the performance of default prediction and customer classification. Compared to classical measurements of credit ratings, such as FICO, our model is more accurate and provides an early warning. The results are interpreted intuitively through visual analytics and dashboards that guide decision-makers on real-time adjustments to credit policy. This system provides the foundation of smart financial decision-making to create a sustainable credit ecosystem and sensible consumer habits.*

Keywords *- Credit Card Behavior, Risk Profiling, Financial Metrics, Machine Learning, Consumer Segmentation, Composite Risk Index, Predictive Modeling, Credit Scoring.*

1. Introduction

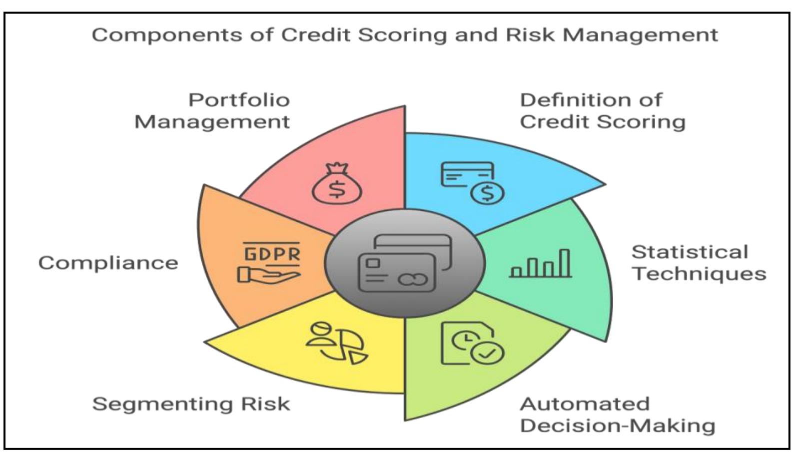

Fig 1: Key Components of Credit Scoring and Risk Management

The credit card system has entered consumer finance extensively and has caused a dramatic shift in the way people obtain and deal with short-term credit allocations. With this increased penetration and ease of encountering liquidity and convenience, understanding and management of credit risk have become even more complex. [1-3] The spending habits, repayment behavior and financial soundness of consumers have become too heterogeneous, and therefore, the current credit scoring practices are grossly inadequate. Financial institutions must thus strive to develop more complex and dynamic risk modelling frameworks that befit the complexity of consumer behaviour. This requirement is further intensified by the current digitalization of the financial sphere, which has facilitated accumulating and analyzing enormous volumes of real-time data at high frequency.

Along with these conventional credit bureau scores, lenders are now utilizing a vast selection of other sources of alternative data, including transactional data, behavioral indicators, and psychographic profiles, which can offer lenders far more in-depth and dynamic information about the financial practices and risk tendencies of a consumer. Newer financial variables, such as credit use patterns, income instability, and payment regularity, have also demonstrated good potential in improving creditworthiness estimation. The subsequent changes indicate a paradigm shift toward more data-driven, holistic credit risk management approaches, where the capability to use globalised data and interpret it has become a distinguishing factor for financial institutions that want to stay competitive, remain compliant, and remain client-centric.

1.1. A Comprehensive Analytical Framework for Modeling Consumer Credit Card Behavior

In order to be able to successfully model consumer credit card behavior in the data-intensive financial environment nowadays, a complex analytical framework that is complex enough to incorporate multiple facets of consumer data in combination with sophisticated methods of modeling and providing humanly interpretable outputs is necessary. Such a framework is founded on the following five components:

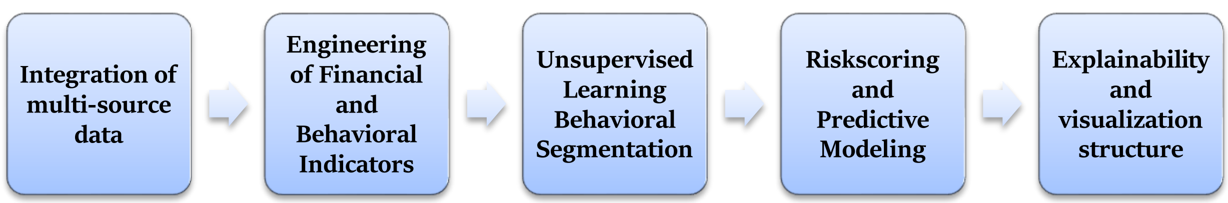

Fig 2: A Comprehensive Analytical Framework for Modeling Consumer Credit Card Behavior

- **Integration of multi-source data:** An adequate structure starts with the incorporation of multiple data sources. They include traditional credit bureau data, transactions (purchases, payments, and balances), demographics (age, income, employment), and behaviour (consistency in payments and frequency of spending). The framework is able to provide a profile of the consumer in a 360-degree view since it incorporates both the past behavior of the credit consumer and the here-and-now activity of the consumer through the combination of structured and semi-structured data.
- **Engineering of Financial and Behavioural Indicators: In modelling, it is essential to derive useful features that capture the drivers of risk**. Important engineered features are Debt to Income Ratio (DTI), Credit Utilization Ratio (CUR), Average Payment Delay (APD) and Spending Volatility Index (SVI). These measures enable active monitoring of financial performance and borrowing activity as they are ahead-of-time indicators of default.
- **Unsupervised Learning Behavioral Segmentation:** In a bid to distinguish consumer profiles, the framework applies a cluster-based approach to profile consumers based on their behavior, e.g. the names used include: Transactors, Revolvers, and Risk-Seekers. This segmentation assists in knowing the risk on a cohort basis and enables institutions to manipulate credit strategies and corrective measures based on this understanding.
- **Riskscoring and Predictive Modeling:** The main part of the framework uses supervised machine learning algorithms, including models like Logistic Regression and Gradient Boosting Machine (GBM) to forecast the probability of default. In conjunction with this, a Composite Risk Index (CRI) has been developed based on methods such as Principal Component Analysis (PCA), providing a continuous yet interpretable measure of credit risk. These models work better than the traditional scoring because they can help pick up complicated behavior patterns.

- **Explainability and visualization structure:** The framework will include the use of interactive dashboards and explainable AI in order to help generate actionable intelligence. Included in these visual interfaces are the ability of credit officers to investigate the importance of features, trends in risks by individual segments, and to dive into unique consumer information. Explainability increases trust and regulatory compliance, enabling clear, data-driven decision-making.

1.2. Risk Profiling Using Advanced Financial Metrics

This reflects the changing nature of credit risk management, where conventional metrics like credit scores and payment history, although essential, are slowly becoming inadequate in providing an overall view of a consumer's creditworthiness. New risk profiling requires that new high-end financial statistics be added to supply a more detailed and volatile perspective of customer behavior. The metrics will assist the financial institutions to detect signals of impending credit distress early, in order to separate various classes of borrowers and credit disbursal more efficiently and in an accountable manner. A major indicator is the Debt-to-Income Ratio (DTI), which compares a consumer's gross income to the amount of debt payments received per month. Having a high DTI reveals that one has a low ability to engage in any debt and is a significant indicator of default in the case of an economic downfall. The other important measure is the Credit Utilization Ratio (CUR), which is the amount of credit used by an entity divided by the amount of available credit. Very strong CUR amounts are related to credit dependency and may typically indicate high-risk finance chances, especially when in association with erratic payment profiles.

Another essential characteristic is the Average Payment Delay (APD), which measures the average time between when statements are issued and actual payment periods occur. Constant late payments can be a sign of a lack of liquidity or financial lack of stonewall. On the same note, the Spending Volatility Index (SVI) captures changes in Monthly expenditure, which indicates a fluctuating financial trend that may be due to insufficient or irresponsible spending patterns. A large value of the SVI is an indication of uncertainty in financial conditions and may impair the capacity of the consumer to have a responsible management of credits. Lenders can transform them by incorporating new, advanced metrics into risk profiling. To achieve this, they can eliminate backwards-looking, static models and replace them with dynamic models based on behaviour. These indicators not only increase the predictive power but also enhance explainability, allowing risk analysts and regulators to understand the reasoning behind the credit decisions made. And at the end of the day, these measurements aid in more equitable, proactive and accurate credit distribution on the part of the financial industry and the consumer.

2. Literature Survey

2.1. Traditional Credit Scoring

The customary credit scoring models rely mainly on the fixed amounts of credit bureau attributes (FICO score, amount of debt outstanding, repayment records, etc.). [4-7] Such models most often employ linear statistical methods such as logistic regression that presuppose linear relationships among variables and are restricted in their ability to extract complex patterns. Although the adoption of these methods has been very high due to their ease of use and regulatory knowledge, they lack accuracy in predictions, particularly for creditworthy individuals with thin credit files or non-traditional financial activities. They also do not take into consideration useful information on behavioral or transactional data that might give a better idea of the risk of borrowers.

2.2. Credit Modeling Behavioral

Credit modeling is a form of behavioral credit modeling, which in turn is a development of credit scoring and attempts to improve the predictive accuracy of credit scores by taking into account the real-life behavior of a borrower managing his or her finances in the past. Payment patterns and spending behavior were discussed early in contributions, with the growth in the availability of data, more recent methods also began introducing psychographic aspects (characteristics such as personality) and social cues based on the digital signature. These insights on behavior help in bringing a better understanding of creditworthiness, particularly for underrepresented groups, including those with non-formal sources of credit. Nevertheless, incorporation of such non-conventional data continues to experience resistance in terms of data quality, privacy issues and acceptance of such regulation.

2.3. Metrics integration of Finances

The latest research recommends incorporating dynamic financial health indicators into credit scoring models to enhance their accuracy. Credit Utilization Ratio (CUR) used to calculate the proportion of available credit that is in use, Average Monthly Balance (AMB) an indicator of the stability in cash flow and Income-Volatility Index (IVI) that indicates variation in income are some of the better predictors of financial strength of a person who is borrowing. These indicators provide continuous information on how an individual handles debt, unlike traditional measures that only give a snapshot for a certain period. Incorporating those metrics into credit models will aid in improving stronger and timelier risk assessments.

2.4. Machine Learning in Finance

Credit risk applications may require ML because of the availability of nonlinear relationships and interactions. Owing to its capability to identify nonlinear patterns and interactions, the application of ML approaches to credit risk is gaining popularity. The paper referred to as [6] illustrates how the ensemble methods, including random forests and gradient boosting machines, perform to accurately predict the chances that a borrower defaults successfully. Such models generally perform much better than traditional statistical methods because they model the subtle dependence between features. Furthermore, feature selection and model tuning are automated, which improves scalability. However, the current trend of the application of ML in credit scoring is mitigated by the discussions of making the models transparent, the regulatory environment, and the possible bias in algorithms.

2.5. Overview of the gaps

Nevertheless, despite these developments, significant gaps remain in the existing credit risk assessment framework. The respective advantages of the models are as follows: First, not all models comprehensively combine data within different fields, and thus, analytical disparities in risk profiles remain untouched. Second, a great number of credit scoring systems are quite static in their nature, and are not able to change scores according to real-time data or dynamic behavior change. Lastly, although machine learning models can have high levels of predictive power, they are usually highly uninterpretable, so it is hard for financial institutions to justify the decisions taken to regulators or clients. Such limitations need to be addressed so that more equitable, more inclusive, and more accurate credit scoring can be developed.

3. Methodology

3.1. Framework Overview

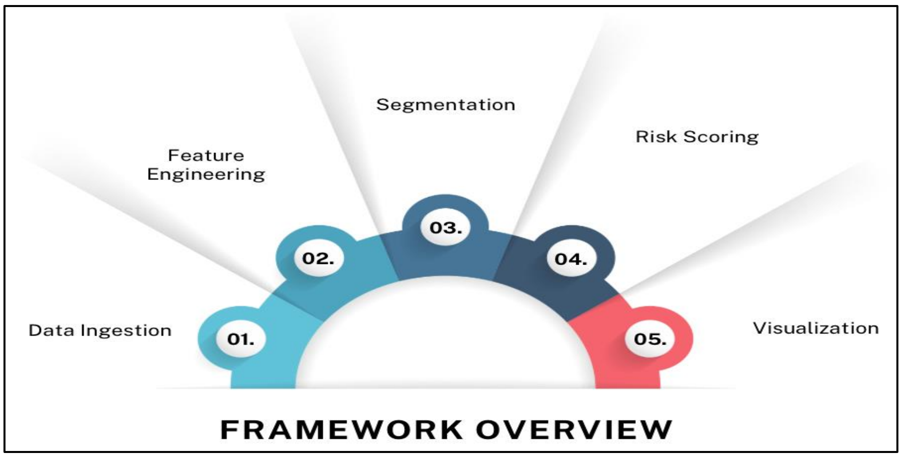

Fig 3: Framework Overview

- **Data Ingestion:** The first phase is data ingestion, during which raw data drawn from various sources is accumulated and assembled into a single system. [8-12] These can incorporate transaction information, credit bureau data, behavioral pointers, economic measures, and other varieties of data like social or psychographic signs. To be ingested successfully, different data formats have to be dealt with and data integrity assured by cleaning, normalizing and deduplicating in case of further analysis.
- **Feature Engineering:** Feature engineering deals with the manipulation of raw data in an attempt to make it into meaningful variables (features) which will augment the performance of the model. The step involves making ratios (ex., credit utilization), aggregating trends (ex., patterns of spending over time), and coding the categorical response. Carefully designed features aid in capturing the suitable signals out of the complex data and play a vital role in ensuring the accuracy of the model, especially in the presence of machine learning techniques.
- **Segmentation:** Segmentation refers to the action of grouping people or accounts that share common characteristics or behavioral patterns. One can utilize different risk modeling approaches by customer group differentiating, say, between the salaried and the self-employed, between the prime and subprime borrowers. When the segmentation is in order, the model will be more precise, and more individual credit evaluations will occur.
- **Risk Scoring:** During this step, scores or credit risk scores are assigned to each entity based on the use of statistical or machine learning models, depending on the likelihood of default or financial distress. The scoring model is based on the engineered intricacy and allows for scaling to the diverse customer segments. It may either be a conventional logistic

regression or more complicated models such as gradient boosting or neural networks, both regarding accuracy and interpretability.

- **Visualization:** Visualization also allows the stakeholders to take action arising from the scoring. Examples of the initial insights provided with the help of dashboards and charts include normal distributions of default risks, individual risks, and segment-specific risks, as well as the importance of key features. Relevant visualization helps the decision-making process to understand otherwise complex model output that can be used by the credit officers, analysts, and those with regulatory responsibility to understand the scoring process through transparency and trust of the product.

3.2. Data Sources

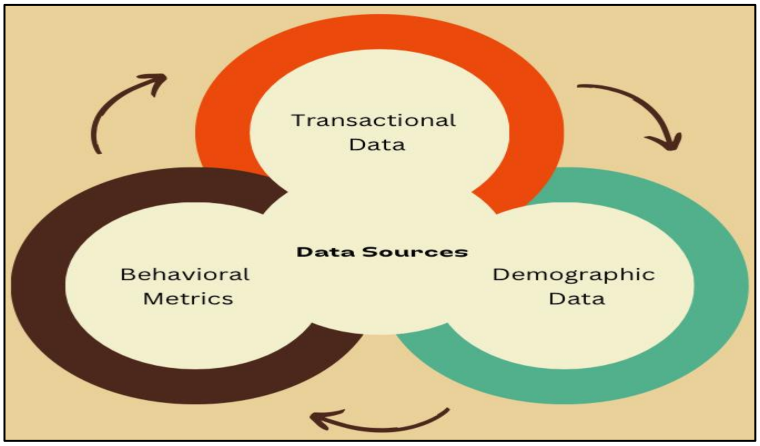

Fig 4: Data Sources

- **Transactional Data:** Transactional data is the recorded specifics of a money transaction, normally retrieved out of monthly bank or credit card statements. This contains details of purchases, payment of bills, and cash withdrawals, as well as transfer of funds. Such transactions are analyzed to reveal spending habits, permanence of income, and liquidity trends. This information offers a real-time view of a person and how an individual conducts himself or herself around money Matters, and this is a critical determinant of future creditworthiness, even more than a credit score; thus, it is a critical predictor of creditworthiness.
- **Demographic Data:** Personal attributes that constitute demographic information are age, gender, income level, employment status, education, and residential stability. The variables are used as a benchmark of the financial ability and stability of an individual. As an instance, it is always assumed that lower credit risks are linked to consistent employment and high levels of income. Data on other types of data, aggregated with demographic data, is important since the data on its own might not yet be adequate to score.
- **Behavioral Metrics:** Behavioral measures record trends in the way people spend their money as they pass the years. Some of these are spending frequency, repayment regularity, behavior of making minimum payments or payments in full and responsiveness to due dates. These measures demonstrate discipline and consistency in finances, offering a dynamic perspective on risk. Through monitoring the behavior patterns, say, sudden spikes in spending, or late payment models, can detect these behavior changes as precursors to financial distress.

3.3. Feature Engineering

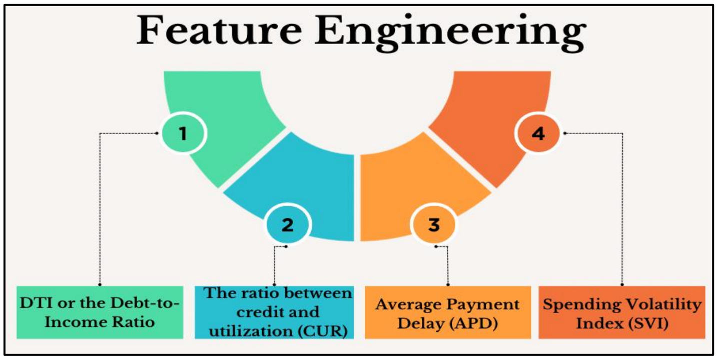

Fig 5: Feature Engineering

- **DTI or the Debt-to-Income Ratio:** The Debt-to-Income Ratio (DTI) is a financial ratio between an individual/or a households monthly income, and the amount of money paid for debts. It is a ratio determined by dividing the amount of debt repayment by the gross monthly income. DTI in excess of 30 per cent means that a borrower is over-leveraged, thus being likely to default. Being one of the key markers of financial pressure, DTI is used to determine the borrower's ability to assume another credit burden.
- **The ratio between credit and utilization (CUR):** The amount of revolving credit, or equivalently, the credit card limit, that is currently utilized is indicated by the Credit Utilization Ratio (CUR). It is calculated by dividing existing outstanding balances against overall credit limits. A high CUR means an individual is highly dependent on credit, a fact which may indicate stress in finances. The low ratio is commonly associated with effective credit management and a lower risk of default.
- **Average Payment Delay (APD):** Average Payment Delay (APD) mirrors the average number of days that a borrower is late paying beyond the payment date in a billing period. It is a big behavior sign of payment discipline and reliability. The greater the APD, the more likely an individual is to be perpetually late in getting things done, and this is associated with a greater risk of defaulting. This metric introduces the time dimension into risk analysis, examining the pattern of consistency.
- **Spending Volatility Index (SVI): The** Spending Volatility Index (SVI) is a measure of how much a given person varies in his or her monthly expenditure. It is usually evaluated by statistical variables such as standard deviation or coefficient of variation within a certain term. The highest volatility can be expected from dysfunctional money habits, high spending, and unreliable income, all of which augment credit risks. On the contrary, low volatility implies stable and predictable financial management.

3.4. Segmentation

In order to mitigate risk assessment strategies more aligned to various borrower characteristics, we used behavioral segmentation using K-Means Clustering. [13-16] This unsupervised machine learning mechanism clusters people on the basis of similarity of their behavioral and financial models, including their spending patterns, repaid behaviour, credit usage, and stable income limits. In contrast to the classical rule-based sorting, K-Means is a dynamic method that automatically finds natural divisions in the data, which is especially relevant for identifying subtle patterns in large and diverse data sources. The Elbow Method was applied to the calculation of the ideal number of clusters. In this process, the within-cluster sum of squares (WCSS) is plotted against the number of clusters, and then locations where the rate of change becomes very flat are determined, which represent the most suitable number of different behavioral segments. With our data, the elbow point appeared at $k = 4$, implying that there are four optimal groups, and that it is deep enough to capture the significant variation among the borrowers. After clustering was done, we examined the nature of each of the groups and gave them labels which would be pronounced as a pronounceable label.

As such, "Transactors" are people who like to pay off their credit balance regularly, have low credit utilization, exhibit steady income and expenditure behaviour levels, and are therefore low credit risk customers. By contrast, again, Revolvers are very often those which attract balances and have moderate or high credit utilization, and are of moderate risk. The others, under the description of Risk-Seekers, presented unstable spending patterns, inconsistency in payment arrangement and high income variability, indicating a high probability of default. The labels help financial institutions to understand the profiles of segments conveniently and devise specific credit. Segmentation will not only make models more precise as it will enable the use of segment-specific risk models, but will also facilitate the ability to offer customized credits and better portfolio management, as well as improve customer interaction. This segmentation strategy builds on risk policies and the way people utilize them to create a very important transition between data analytics and credit decision-making that people can use.

3.5. Risk Profiling

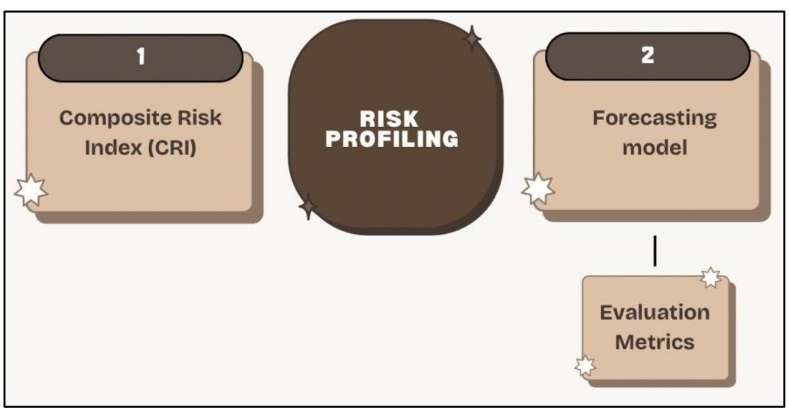

Fig 6: Risk Profiling

- **Composite Risk Index (CRI):** The Composite Risk Index (CRI) is just one numerical value which combines various financial and behavioral indicators into a single number that measures the credit risk. This index aids in adjusting complex borrower profiles into a scalable and interpretable form of credit decision-making. PCA was used to identify the most significant contribution of every feature of the input. PCA minimizes the dimension by determining the best combinations of variables that are relevant, and the extracted combinations are used to calibrate the weights attached to each of the factors of the CRI. This makes the index attain the most pertinent fluctuation in risk behavior with the least amount of overlaps and noise in the measurement.
- **Forecasting model:** In the estimation and forecasting of the probability of default, a blend of predictive models was utilised. The Baseline model was Logistic Regression because of its simplicity, interpretability, and high applicability in financial applications. It approximates the probability of default as a linear blend of input variables, and it is very appreciated, especially regarding its clarity in the regulatory environment. At the same time, there was the further deployment of a more refined model, the Gradient Boosting Machine (GBM), to enhance the predictive capability. GBM constructs an ensemble of weak learners (decision trees are usually the weak learners) stage-by-stage, and in each stage, the residual errors are optimized. The technique has been characterized as highly accurate and good at capturing nonlinear relationships in the data.
- **Evaluation Metrics:** The standard classification measures were employed to analyze the performance of the model. Accuracy simply gives the general correctness of the model, but AUC (Area Under the ROC Curve) measures how well the model separates the defaults and non-defaulters at different levels. Precision is the ratio of how many predicted defaulters defaulted, and Recall is the capacity of a model to correctly determine actual defaulters. These metrics taken together will give a fair picture of the accuracy and the completeness of the predictive models and make the predictive models robust in real situations of credit risk assessment.

4. Results and Discussion

4.1. Dataset Description

In this analysis, the data source under consideration includes anonymized data on credit card users (25,000 in total), where the total period of the study is continuous (January 2020 to April 2022). These two years were best chosen to serve as a time period in which consumer behavior changes as a result of the shifting economic climate due to the pre-pandemic, pandemic, and the first stages of post-pandemic economic recovery. The dataset combines transactional and demographic fields, which allow a multiplicity perspective of the user in terms of his/her financial portrait. Transactional data contains summaries on purchases, payment of bills, cash withdrawals, credit use and delayed payments on a monthly basis. These variables indicate a financial behavior such as spending behaviour, liquidity preferences and credit discipline. To illustrate, one of such recurring patterns is constituted by the repetitive minimum-only payments, the growing balances or high utilization rates, and these are the telltale signs of the financial troubles that lie in the future. With these measurements (on a monthly basis), the data can be used to build dynamic features that are capable of capturing changes over time as opposed to using just a snapshot of information.

Along with the behavior results, demographic information was taken into consideration, including age, monthly income, current employment and residential stability, increasing risk modeling. Such considerations provide a much-needed context in sifting between transitory changes in financing and structurally risky borrowers. All Personally Identifiable Information (PII) was either removed or anonymized in the course of ensuring privacy and compliance with industry practice. The default status was also drawn on the dataset, and the outcome variable is drawn as a 90-day delinquency, i.e., no payment during the 90 days of consecutive payments. The definition refers to regulatory recommendations of what constitutes classification of non-performing accounts and offers an effective ground truth that can be used in training supervised learning models. All in all, the size, variability and the timeframe of the dataset make it ideal to construct and test sophisticated credit scoring models that combine both classic risk factors with alternative ones.

4.2. Clustering Outcome

K-Means clustering was used to discover the purposeful patterns in the behavior of borrowers based on the collection of engineered attributes that comprise financial and behavioral features (including credit utilization, payment delay, income volatility, and spending frequency). Before clustering, Principal Component Analysis (PCA) was employed to generate low-dimensional data representations that captured as much information as possible. Two major components were obtained, and a 2-dimensional scatter plot was plotted, giving a good visualization of natural groupings in the population. This was an unsupervised training method, where the data could be segmented in an unbiased way depending only on data-related similarities, without having to predefine the risk categories.

Table 1: Segment Characteristics

Segment	Default Rate
Transactors	2%
Revolvers	15%
Risk-Seekers	35%

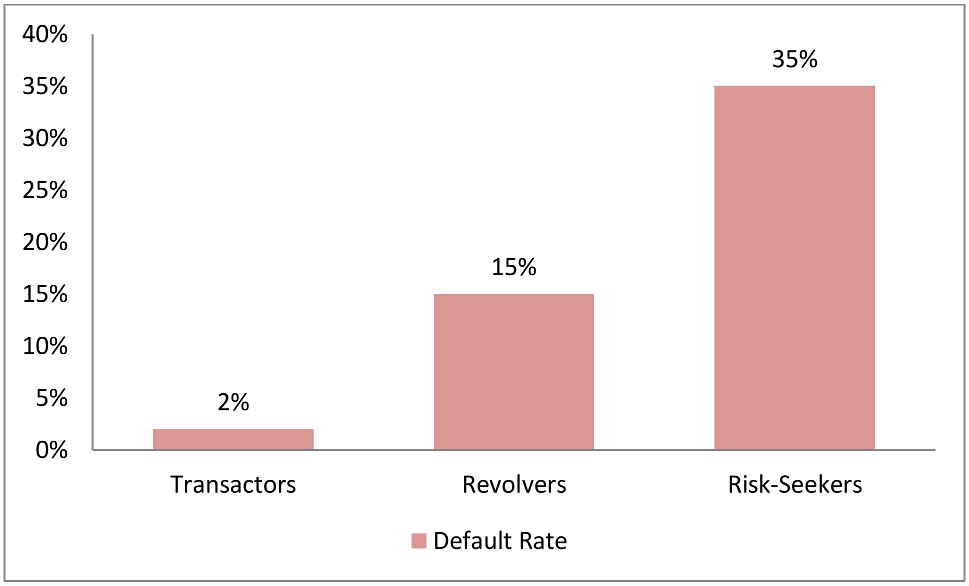

Fig 7: Graph representing Segment Characteristics

- **Transactors:** The population of this segment is represented by people who pay off the balances in full every month and have low credit utilization levels. They are highly disciplined financially with selective spending, sending money in good time, low finances to income figures, and a steady source of income. Consequently, they have the least default rates of only 2 percent, which renders them very creditworthy. Such consumers tend to be well-wired when it comes to managing finances and are of low risk to lenders.
- **Revolvers:** Revolvers are noted by the fact that payment is partially balanced and credit utilization is moderate to high. They are likely to balance themselves down the line and pay just a minimum or fewer amounts. They are not overdue yet; however, the way they pay indicates that they face liquidity issues or that they risk running up unsustainable debt. The level of credit risk of this category was moderate, with a default rate of 15%.
- **Risk-Seekers:** This high-risk segment has volatile financial performance with high credit utilization ratios (CUR) and payment delays (average payment delay - APD). Occasionally, Risk-Seekers have earnings fluctuations, unpredictable spendthrift peaks, and weak repayment regularity. These traits are associated with the highest rate of default of 35% which means they are the most susceptible portion as far as credit risk is concerned. The ability to monitor proactively and intervene specifically is essential to managing this group effectively.

4.3. Model Performance

Table 2: Comparison of Model Performance

Model	Accuracy	AUC	Recall
FICO (Baseline)	72%	70%	66%
Logistic Regression	78%	76%	72%
GBM	86%	89%	84%

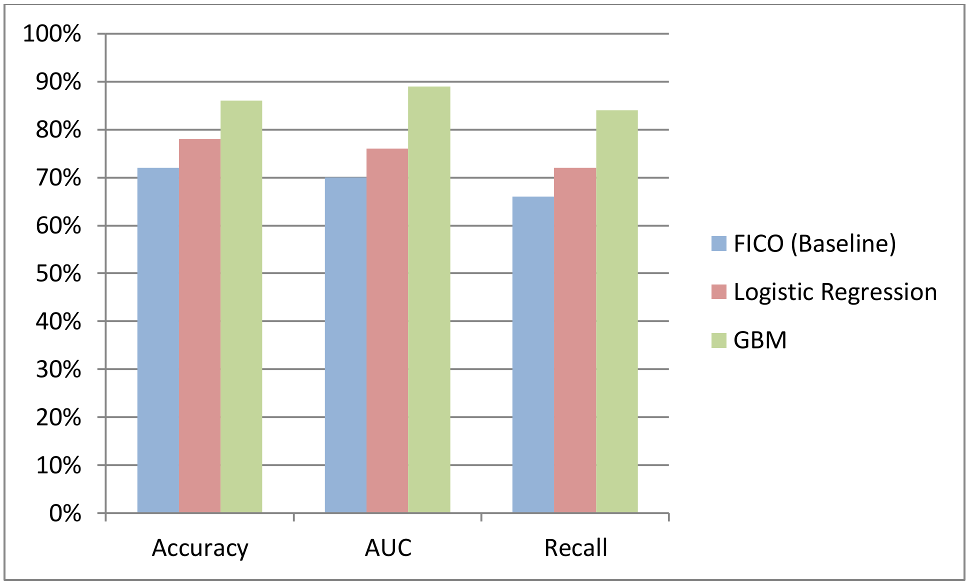

Fig 8: Graph representing Comparison of Model Performance

- **FICO (Baseline):** The established industry standard, the FICO score, was used as a benchmark. Despite being widely used in credit decisions, its performance in this analysis was fairly low, with an accuracy of 72%, an AUC of 70%, and a recall of 66%. The FICO model is mainly dependent on historical credit data and does not include dynamic transaction and behavioral variables. Because of this, it is likely to perform poorly in detecting risk at an early stage, especially among users with thin files or disparate files.
- **Logistic Regression:** Logistic Regression as a statistical baseline model was able to achieve better results through engineered characteristics like credit utilization ratio, average payment delay, and income volatility compared to the FICO score. It achieved a performance of 78.38% accuracy, an area under the curve of 0.76, and a recall of 72%. A strength of Logistic Regression is its ability to be comprehensible and optimally effective, thus applicable in a discourse of regulation. It relies, however, on the linearity of relationships between predictors and log-odds of default. Therefore, it does not represent the complexity and non-linearity of relations that exist in the actual behavior of borrowers.
- **Gradient Boosting Machine (GBM): The** Gradient Boosting Machine (GBM) was much more accurate than the FICO and Logistic Regression models. It also recorded an accuracy of 86%, a high AUC of 89%, and a recall of 84%. GBM operates by developing an ensemble of decision trees sequentially performed on the data by correcting errors one step at a time and detecting any complicated patterns in the data. Its good recall and high performance points to higher potential of identifying high-risk borrowers correctly, which therefore makes it useful in reducing false negatives during the prediction of defaults. Although more complex to use, it is predictive enough and should be used in modern credit risk models.

4.4. Discussion

The results of this paper emphasize the effectiveness of more modern machine learning algorithms, in particular Gradient Boosting Machines (GBM), over older forms of credit scoring such as FICO. GBM's architecture is such that it can systematically learn to correct those errors made by the weak learner and therefore is capable to learn nonlinear interactions, in addition to the subtle aspect of dependencies of features that the traditional linear models will miss. This produces a much larger recall, which means GBM is more effective at tagging real defaulters and minimizing false negatives, which is important to the lender who does not wish to succumb to credit losses, yet at the same time, they are not willing to severely penalize the few flaws left. Its innovation point is the Composite Risk Index (CRI) that allows summarizing various aspects of behavior, financial risk indicators into a readable score. With the use of Principal Component Analysis (PCA), the CRI is generated with optimized feature scores so as to maximize the relevant variance in the behavior of borrowers. This solution can be interpreted as a clear and ongoing gauge that can be viewed as a way beyond the discontinuous dimension of old-style classification of defaults, which gives a more refined picture of borrower risk.

In contrast to black-box models that are typically not interpretable, the CRI is not only predictive but also transparent, which is why it can fit regulatory settings and risk management. The research study incorporates interactive visual dashboards to be used by risk managers and credit analysts in addition to predictive modeling. Such dashboards help explore each segment of users, visualize model explanation (e.g., feature importance), and drill down into individual risk profiles of borrowers. By presenting complex analytical results in an easy and intuitively usable form, the dashboards enable stakeholders to make explainable, data-driven, and proactive credit decisions. Not only does this streamline efficiency in operations, but it also leads to equity and openness in lending activities. All in all, these three aspects of strong machine learning, composite scoring, and interpretability in a visual form constitute a prospective direction for the future of credit risk management in the digital world.

5. Conclusion

The paper offers an up-to-date and broad system of consumer credit risk modeling, which is supposed to overcome most of the shortcomings of the existing traditional credit scoring schemes. The proposed solution addresses the above flaws by integrating behavioral segmentation, tailor-made financial indicators, and the latest machine learning tools, allowing an even more sophisticated and precise credit card user risk evaluation. The methodology gives up on the passive and divided data sources by unifying the transactional, demographic, and behavioral signals to create a single framework. What happens as a result is a dynamic system that will be able to recognize high-risk users more accurately and will thus ensure more responsible and competent lending practices. The Composite Risk Index (CRI) is one of the key innovations made in this work, which is a continuous and interpretable risk score based on Principal Component Analysis (PCA). In contrast to the traditional binary classification model, where a user is labeled just as either high or low risk, the CRI provides a graded insight into the level of risk, which is instrumental in making subtle credit decisions as well as tiered risk-based pricing. Moreover, the behavioral segmentation process using K-Means clustering was able to generate relevant consumer groupings- Transactors, Revolvers and the Risk-Seekers- that are consistent with the actual behavior and pattern of defaulting money. These segments not only increase the predictive performance but also allow for the implementation of a more tailored and explainable risk management strategy.

The three contributions of this research are as follows: (1) the creation of a new Composite Risk Index that combines various financial and behavioral signals into a single score; (2) integration of all kinds of data sources 2, including transactional, demographic, and behavioral signals; and (3) the use of highly sophisticated machine learning models e.g. Gradient Boosting Machines to perform risk prediction and classification. Overall, these innovations give a very strong challenge to the old, outdated scoring mechanisms. Nevertheless, the study has limitations despite such advancements. This data is limited to activities before May 2022 and geographically to North American consumers. These limitations can affect how the findings can be generalized to others and geographical economic contexts. Furthermore, although the behavioral and financial variables are well-covered, other forms of credit, including Buy Now, Pay Later (BNPL) usage, mobile wallet use, and social graph analytics, are excluded. It is planned to perform further research on the concept of real-time CRI calculation, which will allow having a dynamic credit risk management system that will be able to respond dynamically to the changing behavior of the users. In this regard, the integration of alternative and non-conventional elements of credit data, such as BNPL transactions or utility payment histories, social scoring, etc., promises even more effective risk prediction, benefiting both the underbanked and thin-file segments.

References

[1] Hand, D. J., & Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: a review. Journal of the royal statistical society: series a (statistics in society), 160(3), 523-541.
[2] Crook, J. N., Edelman, D. B., & Thomas, L. C. (2007). Recent developments in consumer credit risk assessment. European Journal of Operational Research, 183(3), 1447-1465.
[3] Crook, J. (2002). Credit scoring and its applications. Journal of the Operational Research Society, 52, 997-1006.
[4] Berg, T., Burg, V., Gombović, A., & Puri, M. (2020). On the rise of fintechs: Credit scoring using digital footprints. The Review of Financial Studies, 33(7), 2845-2897.
[5] Hurley, M., & Adebayo, J. (2016). Credit scoring in the era of big data. Yale JL & Tech., 18, 148.
[6] Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. European Journal of Operational Research, 247(1), 124-136.
[7] Banasik, J., Crook, J. N., & Thomas, L. C. (1999). Not if, but when, will borrowers default? Journal of the Operational Research Society, 50(12), 1185-1190.
[8] Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. Journal of Banking & Finance, 34(11), 2767-2787.
[9] Baesens, B., Setiono, R., Mues, C., & Vanthienen, J. (2003). Using neural network rule extraction and decision tables for credit-risk evaluation. Management science, 49(3), 312-329.
[10] Martens, D., Baesens, B., Van Gestel, T., & Vanthienen, J. (2007). Comprehensible credit scoring models using rule extraction from support vector machines. European journal of operational research, 183(3), 1466-1476.
[11] Moro, S., Cortez, P., & Rita, P. (2015). Business intelligence in banking: A literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation. Expert Systems with Applications, 42(3), 1314-1324.

[12] Zhang, D., & Zhou, L. (2004). Discovering golden nuggets: data mining in financial applications. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 34(4), 513-522.
[13] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should I trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135-1144).
[14] Rudin, C. (2019). Stop explaining black box machine learning models for high-stakes decisions and use interpretable models instead. Nature machine intelligence, 1(5), 206-215.
[15] Louzada, F., Ara, A., & Fernandes, G. B. (2016). Classification methods applied to credit scoring: Systematic review and overall comparison. Surveys in Operations Research and Management Science, 21(2), 117-134.
[16] Sun, N., Morris, J. G., Xu, J., Zhu, X., & Xie, M. (2014). iCARE: A framework for big data-based banking customer analytics. IBM Journal of Research and Development, 58(5/6), 4-1.
[17] Ibnugraha, P. D., Nugroho, L. E., & Santosa, P. I. (2018, March). Metrics analysis of risk profile: A perspective on business aspects. In 2018 International Conference on Information and Communications Technology (ICOIACT) (pp. 275-279). IEEE.
[18] Cullerton, N. (2012). Behavioral credit scoring. Geo. LJ, 101, 807.
[19] Neto, R., Adeodato, P. J., & Salgado, A. C. (2017). A framework for data transformation in credit behavioral scoring applications based on model-driven development. Expert Systems with Applications, 72, 293-305.
[20] Bedoui, H. E. (2015). Multidimensional metrics for measuring social and sustainable finance performance. ACRN Oxford Journal of Finance and Risk Perspectives, 4(4), 109-128.
[21] Kumar, V., & Umashankar, N. (2012). Enhancing financial performance: the power of customer metrics. In the Handbook of marketing and finance. Edward Elgar Publishing.