Original Article

Building a Cash Flow Underwriting System: Insights from Implementation

Amol Gote[1], Vikas Mendhe[2]

[1]Solutions Architect, New Jersey, USA.
[2]Senior Consultant, Austin, Texas, USA.

[1]Corresponding Author : aamolgote@hotmail.com

Abstract - *In the evolving landscape of financial services, a pioneering implementation of a cash flow underwriting system is aimed at expanding credit access and enhancing risk assessment. This approach leverages the convergence of open banking, machine learning analytics, and dynamic underwriting rules to provide a more nuanced evaluation of creditworthiness beyond traditional credit scoring methods. The system consists of four critical components: (1) a partnership with a financial data aggregator to securely link customers' bank accounts, utilizing both open banking APIs and traditional integration techniques; (2) collaboration with a specialized third-party vendor that employs machine learning models to analyze transaction data, generating a comprehensive cash flow score and identifying verifiable income; (3) the development and implementation of underwriting rules that integrate the cash flow score and income data to make informed lending decisions; and (4) the seamless integration of these components into a mobile application and backend infrastructure, ensuring a user-friendly experience and efficient loan processing. This paper discusses the cash flow underwriting system's design, implementation, and strategic significance, highlighting its potential to democratize credit access by transcending traditional credit scores and opening new avenues for financial inclusion and risk assessment. This paper introduces a technical implementation approach for cash flow underwriting through collaborative specialized partners, providing a viable alternative to traditional credit evaluation methods.*

Keywords - *Cash Flow Underwriting, Financial Technology (Fintech), Open Banking, Credit Risk Analysis, Machine Learning in Finance.*

1. Introduction

In recent years, the financial services industry has witnessed a shift driven by technological advancements and a growing recognition of the limitations of traditional credit scoring methods. Standard models, heavily reliant on credit scores, often fail to accurately represent an individual's creditworthiness, especially for those with thin credit files or variable income streams. This has led to a significant portion of the population being excluded from access to credit.

The advent of open banking and the rise of machine learning technologies have paved the way for alternative approaches to credit underwriting. One such innovation is implementing a cash flow underwriting system, which differs from traditional credit evaluation.

The cash flow underwriting harnesses the power of open banking to securely access real-time financial transaction data directly from consumers' bank accounts. By collaborating with specialized third-party vendors, the system employs advanced machine learning models to analyse these transactions, generate a cash flow score, and identify income. This approach allows for evaluating an individual's financial health and repayment capacity beyond traditional credit scores. Implementing dynamic underwriting rules based on the insights derived from cash flow and income data further refines the lending decision process from the risk management perspective. It opens new avenues for financial inclusion. This article delves into the cash flow underwriting system's design, implementation, and strategic significance. It explores the technical implementation that powers cash flow underwriting. By embracing technology and data analytics, this system sets a new standard for responsible lending and financial innovation, marking a pivotal step towards a more inclusive financial ecosystem.

2. Literature Review

This literature review explores the genesis and limitations of traditional credit scoring, the imperative for financial inclusion, the pivotal role of technology in reshaping

financial services, and the emergence of alternative credit scoring models, with a particular focus on cash flow underwriting.

2.1. Traditional Credit Scoring Methods

Traditional credit scoring has been the cornerstone of credit risk assessment for decades, relying heavily on historical financial behaviours to predict future creditworthiness. FICO scores, for example, have become synonymous with credit evaluation, determining eligibility for loans, credit cards, and other financial products.

2.2. Need for Alternative Methods

The World Bank's Global Findex Database highlights the role of access to credit in achieving financial inclusion, yet millions remain marginalized from traditional financing offerings due to inadequate credit histories. Studies (Demirgüç-Kunt et al., 2018) argue for developing alternative credit assessment methods that consider broader financial behaviours and capacities. This includes examining non-traditional data points, such as utility payments, rent history, and social media activity, to provide a more holistic view of an individual's financial health.

2.3. Alternative Credit Scoring Models

Alternative credit scoring models are emerging due to technological advancements and regulation changes. The use of machine learning algorithms to analyse non-traditional data sources for credit scoring purposes has been explored in various studies (Agarwal and Zhang et al., 2020), highlighting their potential to enhance the accuracy of credit risk assessments significantly. Moreover, cash flow underwriting, which evaluates an individual's real-time financial transactions and income stability, has emerged as a promising approach. Research (Y. L. Toh et al., 2023) suggests that cash flow data can be a more reliable indicator of creditworthiness than traditional credit scores, particularly for those with fluctuating incomes or minimal credit history.

2.4. Summary

The literature underscores a clear trajectory towards more inclusive and technologically sophisticated credit assessment methods. The shift from traditional credit scoring to models that leverage open banking data and machine learning, such as cash flow underwriting, represents a pivotal evolution in pursuing financial inclusion.

3. Methodology

This section outlines the methodologies for developing and implementing a cash flow underwriting system. The cash flow underwriting system was conceptualized to leverage real-time financial transaction data obtained via open banking APIs and then use that data to run against a machine learning model to get cash flow score and income. Based on cash flow score and income, utilize underwriting rules to identify the subject's approval amount.

3.1. Open Banking Integration

With the customer's consent, we partnered with a leading financial data aggregator to access applicants' banking data securely through open banking APIs. This approach ensured a broad coverage of financial institutions and adherence to data privacy and security standards.

3.2. Machine Learning Model

In collaboration with a data analytics firm, we leveraged their model, which was trained on their dataset. The model was trained to identify patterns of income stability and expenditure behaviour, generating a cash flow score and estimated cash flow income.

3.3. Underwriting Rules

Based on the cash flow scores and identified income patterns, we formulated a set of dynamic underwriting rules. These rules were designed to align with varying levels of credit risk. These rules are the matrix of the cash flow score and the custom credit score.

3.4. Evaluation and Metrics

The evaluation of the cash flow underwriting system began with its deployment on an initial set of 50 loan applications. Early analysis revealed a less-than-anticipated approval rate, prompting a closer examination of our underwriting rules. It became apparent that the model, trained initially on credit card transaction data, did not align well with our target use case of consumer finance, leading to inaccuracies in assessing applicants' repayment capabilities. This discrepancy prompted adjustments aimed at refining false negatives and positives. Post-adjustment, the system saw improved approval rates through continuous monitoring, ensuring that our refinements successfully balanced risk management with accessibility and underscored the importance of aligning model training with specific financial contexts.

4. System Design and Implementation

The cash flow underwriting system is a complex ecosystem that optimizes credit assessment through real-time financial data analysis. This system integrates several critical components listed below:

- Consumer application (Mobile/Web)
- Underwriting Platform
- Aggregator Platform
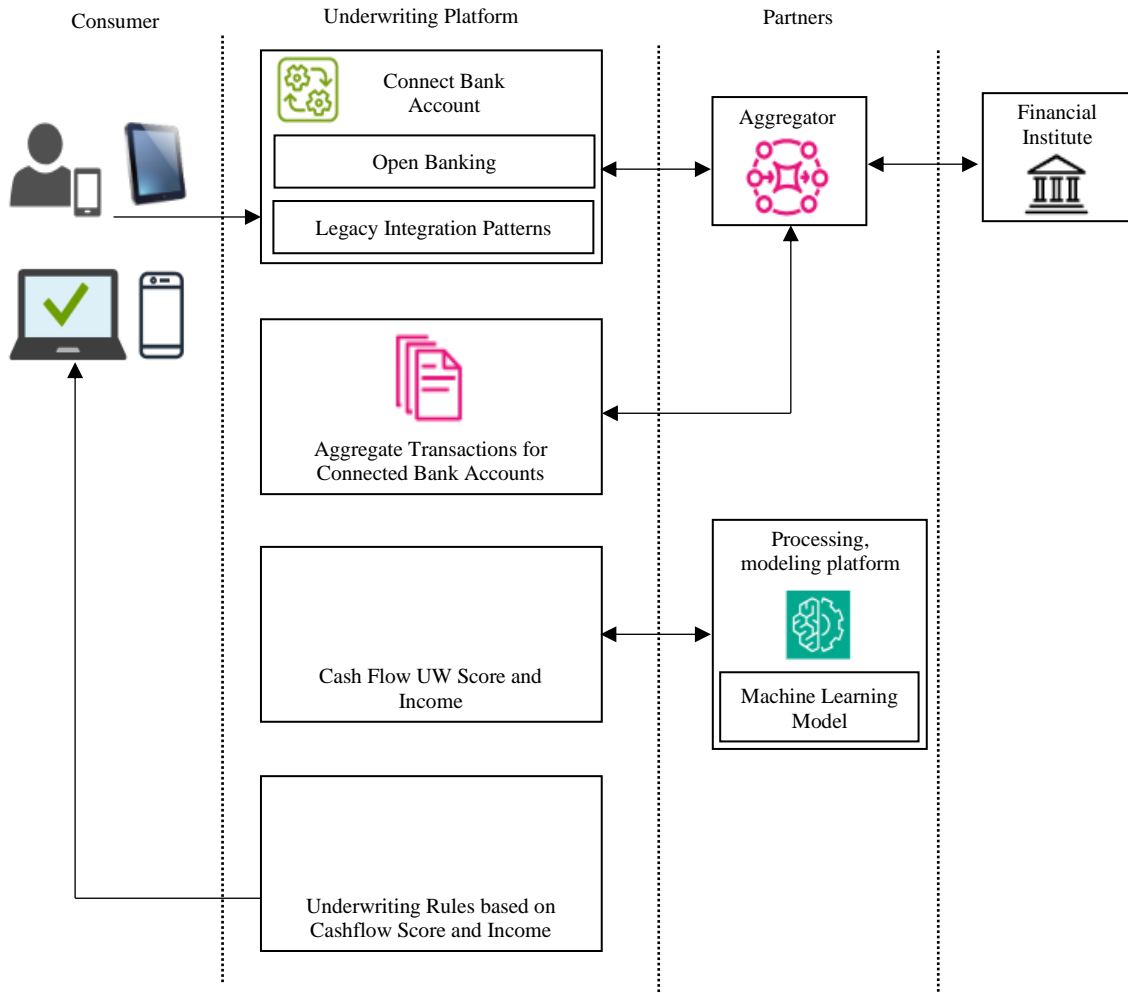- Processing and Modelling Platform

Fig. 1 Cash flow underwriting logical components

4.1. Consumer Application

The consumer application is the primary interface for users to interact with the underwriting system. Designed for mobile and web platforms, it ensures accessibility and convenience for users across different devices. Mobile users download native applications for the Apple Store or Play Store. It enables users to apply for financing, link their bank accounts via open banking APIs, and receive instant credit decisions. It interacts with the underwriting platform through a secure HTTPS channel, ensuring data is encrypted over the transport channel; this helps protect user information and financial transactions.

4.2. Underwriting Platform

The underwriting platform functions as the core of the cash flow underwriting system, orchestrating the intricate decision-making processes central to credit assessment. At its foundation, the underwriting platform employs a dual-layered assessment approach to evaluate loan applications.

4.2.1. Initial Traditional Credit Scoring

Initially, applicants undergo evaluation through a traditional credit scoring model. This step assesses applicants based on their credit history, outstanding debts, repayment history, and other traditional credit metrics. Applicants demonstrating robust credit profiles are immediately presented with optimal loan offers.

4.2.2. Secondary Cash Flow Analysis

Applicants who do not qualify for the best offers under traditional criteria or those facing potential decline are considered for cash flow underwriting.

4.2.3. Eligibility Criteria for Cash Flow Analysis

The platform employs a set of cash flow underwriting rules to pre-screen applicants. This proactive measure aims to filter out individuals with significantly adverse credit histories or those exhibiting multiple financial red flags, thereby maintaining the integrity of the underwriting process.

4.2.4. Informed Consent and Data Security

Applicants deemed eligible for cash flow analysis are invited to link their bank accounts securely using the consumer application. This process is conducted with explicit consent, ensuring transparency and adherence to data privacy standards.

4.3. Aggregator

In our system, using a third-party aggregator is a critical component for facilitating the secure connection to bank accounts. This aggregator bridges the gap between the underwriting platform and the financial institutions, offering two primary integration methods: open banking and traditional integration. When users connect their bank account, our system determines whether the chosen financial institution supports Open Banking. If supported, the user is presented with the option to link their bank account using this method or else fall back to traditional integration.

4.3.1. Open Banking Integration

The consumer application hosts a widget in this process, facilitating a direct and secure connection to users' bank accounts. The aggregator provides this widget. This widget enables users to log in via a familiar interface, mirroring the look and feel of their financial institution's web application, thereby maintaining a consistent user experience. Throughout this process, explicit consumer consent is obtained, with clear notifications that financial data is being shared with the aggregator. The integration adheres to Open Banking standards, prioritizing secure data sharing and user authentication and enhancing the user's comfort and confidence by providing a seamless and protected environment for financial data access.

4.3.2. Traditional Integration

For financial institutions that do not participate in Open Banking, the system interacts solely with the aggregator's APIs, which handle complex integrations with financial institutions. When employing these methods, the aggregator's APIs facilitate a secure login process for the user. This process often involves two-factor authentication or other robust authentication mechanisms. After successfully logging in, the user can select their bank account to link with our system. Throughout this traditional integration journey, it is crucial to note that users interact with the native UI of our consumer application. This user interface dynamically adapts based on the API responses from the aggregator.

4.2.3. Key Aggregators

In the realm of cash flow underwriting, there are a variety of aggregators, with Plaid, Ninthwave, Yodlee, and FinCity emerging as some of the most prominent ones, offering extensive capabilities for linking bank accounts and extracting transaction data.

4.4. Processing and Modeling Platform

The system incorporates a Processing and Modelling Platform to accurately evaluate an applicant's financial health. This component is essential for analysing cash flow transactions to derive two key metrics: the cash flow score and cash flow income. For this purpose, we collaborate with an external partner with a machine learning model designed explicitly for financial analysis. The machine learning model is adept at processing transactional data to assess cash flow patterns. The model calculates a cash flow score that reflects the applicant's financial stability and liquidity by analysing deposits, withdrawals, and other transaction types. Beyond scoring, the model employs algorithms to identify and quantify the applicant's cash flow income.

4.5. Implementation Summary

In summary, the cash flow underwriting system's comprehensive design integrates advanced technology and machine learning to provide a sophisticated approach to credit assessment. Each component, from the consumer-facing application to the backend processing and modelling platform, plays a crucial role in streamlining the lending process, ensuring data security, and enhancing the accuracy of credit decisions.

5. Challenges and Limitations

A significant challenge with the system was the inconsistency in the amount of transaction data provided by various financial institutions. While some banks offered up to two years of transaction history, others limited data access to as little as two months. This variability directly impacted the machine learning model's effectiveness, as a minimum dataset of 25 transactions over at least 90 days is critical for accurate analysis and creditworthiness assessment.

The second limitation arose from inconsistencies within the transaction data, notably missing attributes such as transaction dates and posted dates. To mitigate the impact of these discrepancies on the model's accuracy, we implemented measures to handle incomplete data records. Specifically, transactions lacking both date attributes were excluded from the analysis to maintain data integrity. In cases where only one date attribute was available, we assigned its value to the missing attribute, ensuring a consistent dataset for the model to process effectively.

The third limitation involves the system's integration across the full spectrum of banking institutions. While we successfully connect with most large banks via open banking or traditional methods, smaller credit unions and emerging fintech banks, which often serve as interfaces layered over conventional banks, present unique challenges. The latter's reliance on underlying banking infrastructure complicates direct data aggregation, highlighting the need for alternative strategies to ensure integration.

The fourth limitation is that the system faced a challenge with the immediate availability of transaction data post-bank account linkage. We observed delays ranging from a few seconds to up to three minutes before transaction data became accessible, necessitating the introduction of artificial delays and retry mechanisms in our process.

6. Future Work

In future investigations, we plan to explore integrating more diverse data sources to refine the cash flow scoring model, enhancing its predictive accuracy for creditworthiness. We plan to explore further enhancements to our system by continuously modelling existing results and refining the underwriting rules based on cash flow score and income. This iterative approach will allow us to adjust our system dynamically, ensuring it remains responsive to evolving financial behaviours and market conditions, thereby improving its accuracy and reliability in credit assessment.

In addition to our ongoing efforts, we plan to enhance our collaboration with the aggregator to upgrade our connections with financial institutions to premium endpoints. This upgrade aims to access more comprehensive data, thereby enriching our analysis. We also intend to transition to an advanced cash flow evaluation model provided by our partner, which utilizes less stringent criteria for transaction data, potentially increasing the model's applicability and accuracy. Furthermore, we will conduct in-depth analyses of customer drop-off points within the application process. By identifying and understanding these critical junctures, we aim to develop alternative strategies and solutions to improve user engagement and completion rates, ultimately enhancing the overall effectiveness of our system.

7. Conclusion

This research paper presents an innovative cash flow underwriting system designed to leverage real-time financial data for more nuanced creditworthiness assessments. By integrating open banking, machine learning, and dynamic underwriting rules, we offer an approach that enhances financial inclusion and access to credit. Despite facing challenges such as data variability and integration complexities, the system's development and deployment demonstrate its potential to transform traditional credit scoring.

Future work will focus on expanding data sources, refining models, and improving system integrations to increase accuracy and reliability further. This study underscores the critical role of technological advancements in evolving the landscape of financial services, promising a future where credit assessment is more inclusive, efficient, and reflective of an individual's financial reality.

References

[1] Finreglab, "The Use of Cash-Flow Data in Underwriting Credit," 2019. [Online]. Available: https://finreglab.org/wp-content/uploads/2019/07/FRL_Research-Report_Final.pdf

[2] A. Demirgüç-Kunt, L. Klapper, D. Singer, S. Ansar, and J. Hess, "The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution," The World Bank, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[3] S. Agarwal, S. Alok, P. Ghosh, and S. Gupta, "Financial Inclusion and Alternate Credit Scoring: Role of Big Data and Machine Learning in Fintech," Indian School of Business, Dec. 21, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[4] Y. L. Toh, "Addressing Traditional Credit Scores as a Barrier to Accessing Affordable Credit," Economic Review, vol. 108, no. 3, pp. 1-22, 2023. [CrossRef] [Google Scholar]

[5] J. H. Choi, "Reducing the Black-White Homeownership Gap through Underwriting Innovations," Urban Institute, 2022. [Online]. Available: https://www.urban.org/research/publication/reducing-black-homeownership-gap-through-underwriting-innovations.

[6] D. Antosz, "Cash Flow Underwriting: 5 Ways Lenders Can Drive Growth," Plaid. [Online]. Available: https://plaid.com/resources/lending/cash-flow-underwriting.