



Empowering Financial Institutions

Leverage Advanced Analytics and Predictive Insights to unlock maximum efficiency

Executive Summary

As data becomes a privilege instead of commodity, C-suit executives and operation teams look for methods to leverage it and make insightful decisions for improving their operational efficiencies, finances, and overall performance.

This white paper takes an example of credit card product and one of the key metrics 'Cost-to-income Ratio' that measures the operational efficiency. We discuss the levers available to banks and financial institutions like advanced data analytics, machine learning (ML), and decision sciences to gain insights about customer behavior and business operations and take proactive decisions to have a positive influence on their performance metrics.

We will also cover the role of predictive analytics plays for institutions to stay competitive by leveraging near real-time information from their internal data sources as well as external sources, such as social media, market trends, news feeds, etc.

This white paper delves into:

- Insights of 'Cost-to-income Ratio'
- Customer Data Platform (CDP)
- Nagarro POV through a reference implementation
- Recommendations on improving cost-to-income ratio

Table of Contents

1.	Introduction	4
2.	Insights on 'Cost-to-income Ratio'	4
3.	Factors affecting the 'Cost-to-income Ratio' of Credit Cards	5
4.	Challenges in tracking the factors of the cost-to-income ratio	5
5.	How can banks manage their cost-to-income ratio effectively	6
5.1.	Customer Data Platform (CDP)	6
5.2.	CDP implementation overview	6
5.2.1.	Direct integration of CDP with data sources	6
5.2.2.	CDP integration with data sources through Data warehouse	7
6.	POV – Near real-time insights on the 'Cost-to-income ratio' for credit card	8
6.1.	Imperatives to manage cost-to-income ratio	8
6.2.	Solution approach to predicting credit card balances and defaults	8
6.2.1.	ML model for 'Credit Default Analysis'	9
6.2.2.	ML model for 'Credit Balance Prediction'	10
6.3.	Example of ML Modelling methodology and algorithm	11
6.4.	Model Maintenance and continuous operations	12
7.	Recommendations for improving cost-to-income ratio	12
7.1.	By managing cost	12
7.2.	By Improving Income	13
8.	Endnote	14
9.	Nagarro's take	15



1. Introduction

The banks are looking for cost-effective solutions to deal with available data that enables them to get near real-time insights about their customers, operations, and market conditions. Today, with the explosion of digital services across the banking industry, there is the availability of 'always fresh-always on' for banks about customers, market, competitors, etc. This availability has enabled better access to insights into banking products' cost structure and helped identify areas for cost optimization, enhance operational efficiency, and ultimately boost profitability.

Moreover, data engineering, analytics, ML capabilities, and data sciences are helping banks by providing scalable, efficient, and near real-time digital solutions. Let's explore these through a solution-driven approach for banks' 'Cost-to-income Ratio' performance metrics of their credit card product.

2. Insights on 'Cost-to-income Ratio'

Several parameters influence the overall performance of a bank, for instance customer experience, quality of assets, operational efficiency, risk management, competitive advantage and capital & liquidity requirements. Banks that excel in these parameters have a competitive edge and enjoy high profitability. These parameters are measured through different metrics – Return on Assets (ROA), Return on Equity (ROE), Risk adjusted performance measures (RARORAC, RAROC, RORAC), Cost-to-income Ratio, etc. The aim is to monitor a bank's performance and take proactive measures to improve.

The 'Cost-to-income Ratio' is one of the key metrics that banks use and it measures its operational efficiency. Some of the facets include:

- An inverse relationship between cost-to-income ratio and profitability. The lower the cost-to-income ratio, the higher the profitability, and vice-versa.
- A 50% or lower ratio is considered good, while above 60% ratio may indicate that the bank is facing operational challenges or inefficiencies.
- A comparison between the bank's cost-to-income ratio over the years can indicate their performance and forecast the future returns on their investments.
- It's calculated as:
$$\text{Cost-to-income} = \frac{\text{Operating expenses}}{\text{Operating income}}$$
Where, $\text{Operating income} = (\text{Interest earned} - \text{Interest expended}) + \text{Other income}$
- Credit cards for example, have following cost and income components
 - Cost – credit card processing cost, marketing expenses, collection cost, credit card servicing cost, fraud investigation cost and charge-offs, printing cost, rewards, outsourcing cost, litigation cost, and infrastructure cost.
 - Income – Interest charges (outstanding credit, loans, EMI transactions), fees (annual fee, late fee, cash advance fee, foreign currency transaction fee, processing fee, etc.), merchant discount rate (MDR) charges, partnership income with merchants.

3. Factors affecting the Cost-to-income Ratio of Credit Cards

Several factors have a ripple effect on the Cost-to-income Ratio and thereby impact overall profitability. Banks need to track these metrics closely as well as look for opportunities to make near real-time decisions based on these factors.

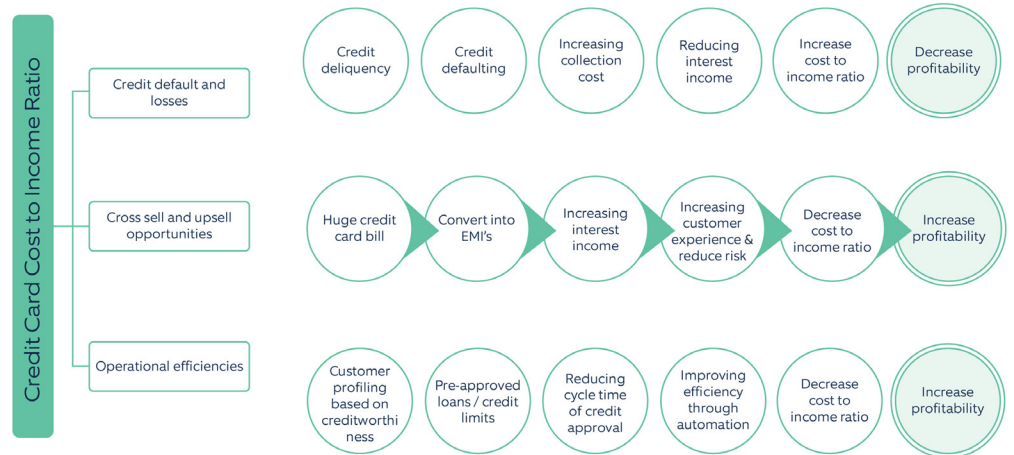


Figure 1: Depicting how factors have a ripple effect on the cost-to-income ratio as well as profitability of credit cards.

4. Challenges: Tracking factors of cost-to-income ratio

While tracking the factors affecting the cost-to-income ratio, a bank's main challenge is in terms of data blending and its engineering governance. Some of the key reasons include:

- Banks need a variety of data related to customers. For instance, for credit cards, banks need credit history to identify creditworthiness, transaction history to identify spending patterns and payment histories, customer surveys & social media to identify customer satisfaction, sentiments & preferences, market data to analyze economic conditions, etc.
- All this data is stored in different internal & external sources/applications and each contains a vast amount of data. For example, credit history information from credit bureaus, transaction information from bank credit card applications, market information from financial news outlets & journals, etc.
- Data from all these sources must be integrated, structured, and analyzed to evaluate the factors impacting the cost-to-income ratio.
- The quality, completeness, and integrity of integrated data are critical; otherwise, it leads to inaccuracy/ incomplete analysis.
- Also, the unavailability of near real-time data delays strategic decision-making as per the evolving market situations.

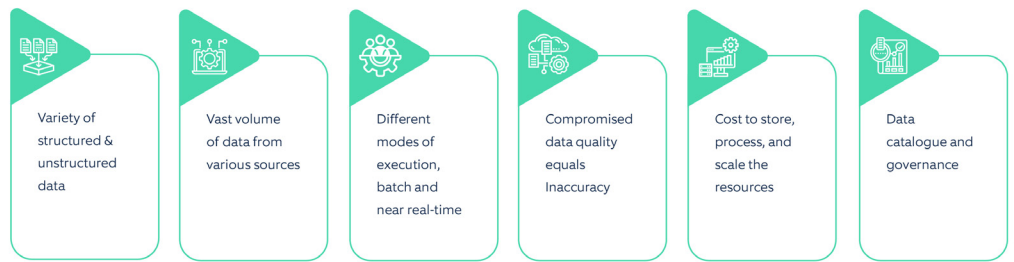


Figure 2: Key challenges while dealing with data

5. How can banks manage their cost-to-income ratio effectively?

5.1. Customer Data Platform (CDP)

The data challenges mentioned in the previous section can be handled through a Customer Data Platform (CDP) - a more holistic approach to customer data management. This approach can help financial institutions improve customer engagement, drive revenue growth, and reduce operational costs. Besides this, the following are some key benefits:

- **360-degree customer view:** It collects and integrates data from all data sources, including batch and streaming data, which helps create customer golden records with a 360-degree view, and allows for a more comprehensive understanding of customer behavior and preferences.
- **Data Integrity:** It ensures that data security and governance are in check and the data integrity in terms of data quality and accuracy is not compromised.
- **Personalization:** It provides a unified view of each customer across all the channels and touchpoints that enables better personalization and targeted marketing campaigns.
- **Real-time analytics:** A built-in analytical platform provides near real-time updates and insights on customer interactions that can help financial institutions respond quickly to changing customer needs and preferences.
- **Regulatory compliance:** It is designed to support compliance with regulatory requirements such as GDPR and CCPA, which is critical in the highly regulated banking industry.

5.2. CDP implementation: Overview

There are various ways to integrate a CDP platform into an existing IT infrastructure of any organization. The most efficient integration depends on the specific requirement of the banking entity and its existing IT infrastructure capabilities. A CDP platform is quite robust and can be integrated through APIs, data connectors embedded on various channels, or direct database integrations. The integration is done in a hybrid mode using multiple types of integrations to connect with various data sources.

In terms of data sources also, we can have all sources feeding into CDP directly via API/batch and customize the integrations wherein some of the key data sources feeds directly into CDP via API and other data sources can be sourced through a central system Data Warehouse/Data Lake.

The next section covers two possible depictions of a CDP integration in an existing banking system. In both depictions, the customer data journey overlaps, wherein the data is sourced through various channels and then unified in CDP. Then this data is sent to downstream channels/planning per the use cases.

5.2.1. Direct integration of CDP with data sources

In direct integration, all data sources are integrated with CDP via API. This has its advantages over other types of configurations. These include:

- Near real-time data ingestion to make the data available for analysis and activation immediately.
- Data consistency to reduce the quality issues and also help in improving the accuracy of the customer insights.
- Flexibility for data source to have its own mapping and transformation processes.

Scalability to add new data sources easily and scale up the volume of data being ingested.

- Lowering the overall cost to eliminate the need of a separate data warehouse or ETL process that can be expensive to develop, maintain, and scale.

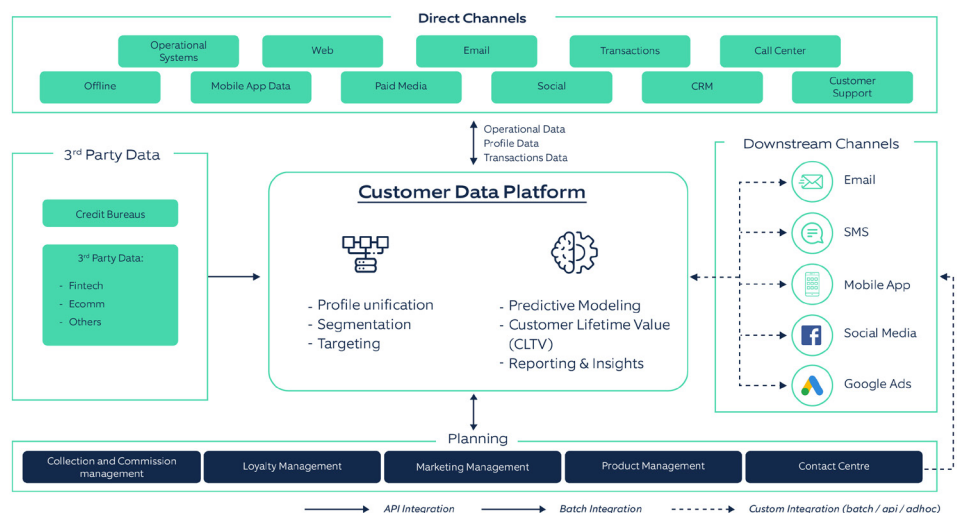


Figure 3: Depiction #1 – CDP is integrated directly with all data sources via API integration. Preferred for near real-time results

5.2.2. CDP integration with data sources through data warehouse

This hybrid depiction of CDP implementation has its own set of advantages. We integrate it with a few key channels directly via API and the remaining data information is being pulled into CDP from the existing data warehouse system.

- It does not require much data cleansing and transformation since in majority cases, the data available in the warehouse is structured and cleansed.
- The data warehouse gives access to the customer’s historical data that can be quite useful for CDP to analyze the customer behavior and trends over the period.
- It can simplify the data integration process for CDP that means we might not need too many systems to integrate with CDP separately.
- The data warehouse can be optimized for performance and scalability that makes it easier to handle large volumes of data and complex queries.

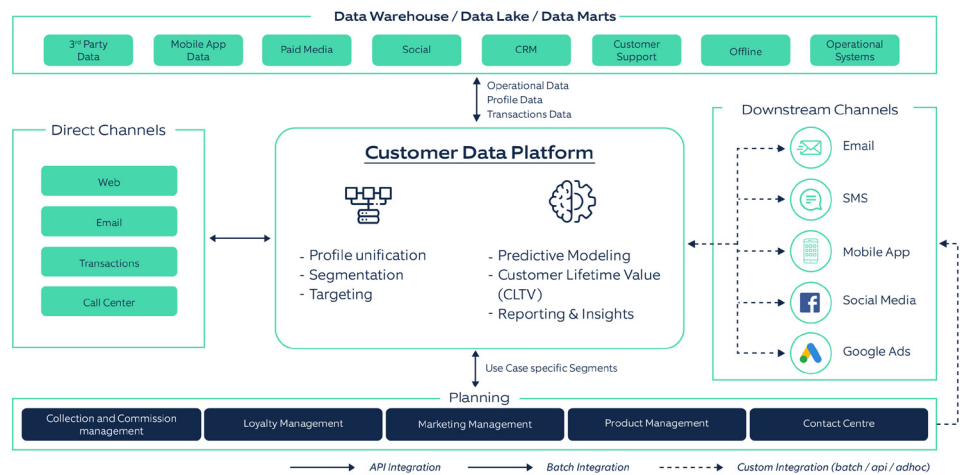


Figure 4: Depiction #2 – CDP is integrated with key data sources via API and other data sources via batch integration from existing data warehouse.

6. POV – Near real-time insights on the Cost-to-income ratio for credit cards

6.1. Imperatives to manage cost-to-income ratio (predicting credit card balances and defaults)

The credit card business is a highly competitive and cost-sensitive space. Being an unsecured lending product, banks must manage their cost-to-income ratio effectively. ML algorithms coupled with decision sciences can be used to develop predictive models that forecast credit balances and losses. With accurate predictions, banks can manage risks more effectively, reduce the costs associated with loan defaults, and thereby improve financial performance.

Predicting credit card balances and defaults can help banks in making informed decisions regarding credit risk and profitability.

- **Managing Credit Risk:** Predictive analytics helps with forecasting credit card balances and defaults, which can help assess credit risk. By understanding the real-time creditworthiness of their customers, banks can make informed decisions regarding credit limits, interest rates, and lending policies. Moreover, they can minimize losses due to default and maintain a healthy credit portfolio.
- **Optimizing Revenue:** Predictive analytics can optimize revenue streams by identifying customers likely to carry high balances and pay interest on their credit cards. By offering targeted promotions, such as balance transfer offers or low-interest rate cards to incentivize customers to use their cards more frequently, banks can boost revenue and enhance customer experience.
- **Improving Efficiency:** Predictive analytics helps streamline bank's operations by automating credit decisions and identifying customers requiring additional support or assistance. Banks can improve their efficiency and optimize their overall operating costs by reducing manual processing and focusing on high-risk customers.

6.2. Solution approach to predicting credit card balances and defaults (using CDP, ML & Data Analytics)

In the banking, financial services, and insurance (BFSI) sector, predicting credit card balances and defaults is crucial for assessing credit risk and making informed decisions regarding lending policies, credit limits, and interest rates. Banks must anticipate both variables to manage the associated losses and costs.

To address this issue, the following use case flow for credit cards using ML predicts loan default and loan balance utilization.

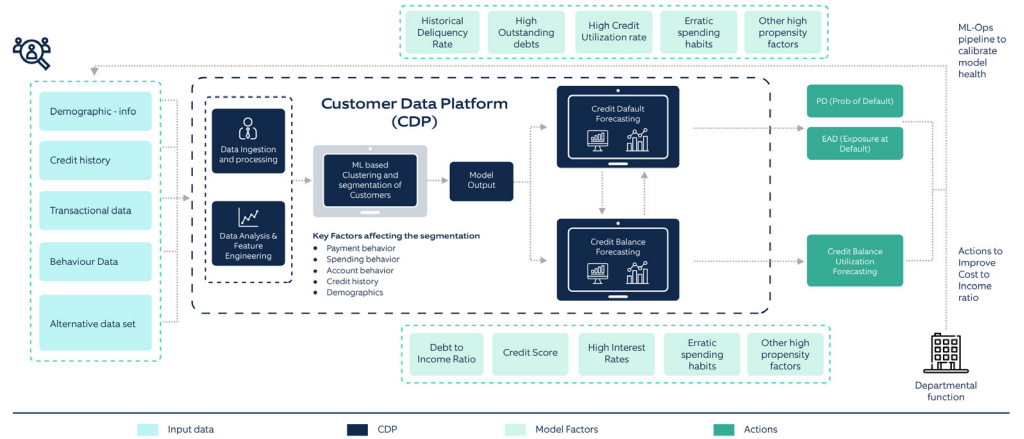


Figure 5: Solution approach to perform Credit Card Balances & Loss Forecasting and improve 'Cost-to-income Ratio' by improving cost

In this solution approach, banks collect the critical data attributes from credit bureaus, credit card applications, financial news outlets & journals, social media platforms, and other channels on a near real-time basis into the CDP platform. We have already shown a couple of CDP implementation approaches in sections 5.1.1 and 5.1.2, highlighting the approach to integrating CDP via API. CDP platforms are robust enough to process this near real-time captured data for analytics. Next, we use business criteria to divide customers into buckets and feed the output into an ML model that utilizes historical data to predict which customers are likely to default in the event of a loan default model, and predict the balance utilization of the customer's credit card.

This model generates a list that the concerned team can use to take appropriate action. By doing so, banks can better assess credit risk, make informed decisions about lending policies, credit limits, and interest rates, and proactively manage potential risks. This approach enables banks to improve customer service and enhance overall performance while helping customers achieve financial goals

6.2.1. ML model for 'Credit Default Analysis'

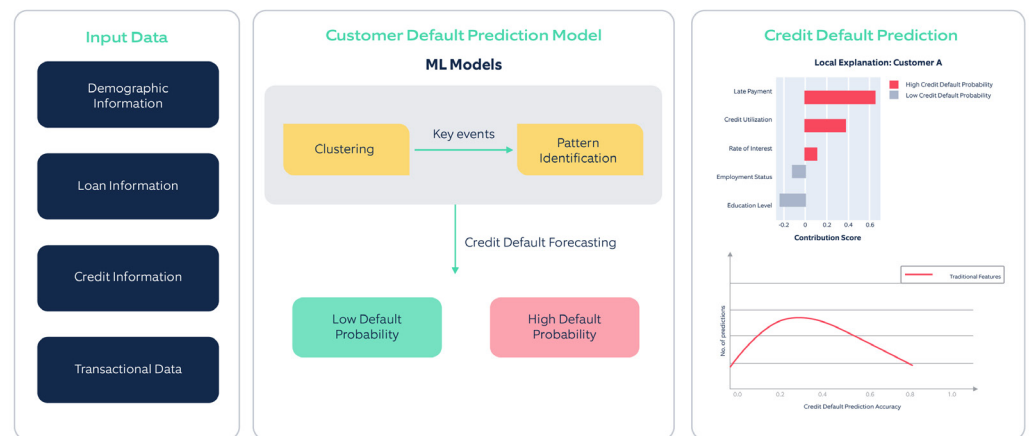


Figure 6: ML model to perform 'Credit Default Prediction' (without alternative data source)

Figure 6 depicts the flow of an ML model for predicting credit card defaults. Given that the predictor variable is of the classification type, we employ classifier models. The data is pulled from various sources and processed before being fed into the classifier that outputs the probability of default per customer. However, business requirements dictate the need for explainability, which can be achieved by employing multiple explainable models such as LIME, SHAP, etc.

As depicted in the figure, the 'Late Payment' and 'Credit Limit Utilization' features contribute the most to default and loss prediction. Valuing the model's performance is crucial, as shown by the evaluation criteria "Accuracy" in the diagram. However, the evaluation criteria may vary based on the use case.

As the number of features increases, the model's accuracy increases, as depicted in Figure 7.

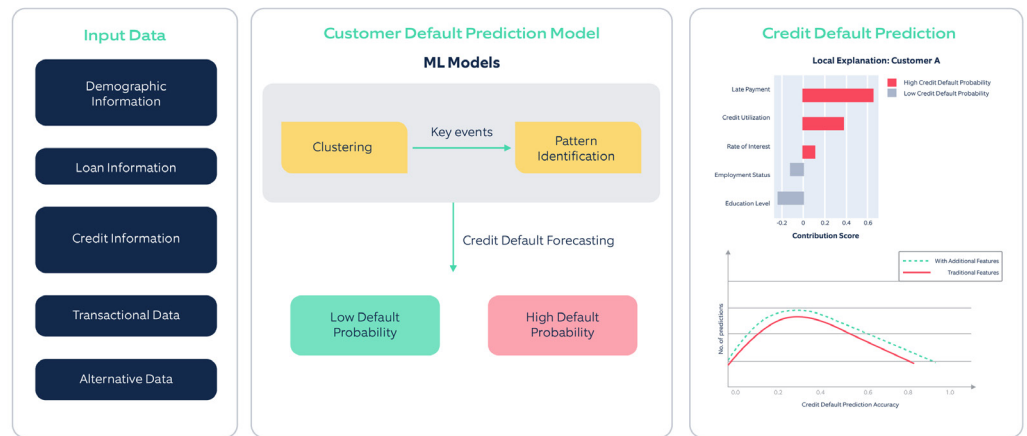


Figure 7: ML model to perform 'Credit Default Prediction' (with alternative data source)

6.2.2. ML model for 'Credit Balance Prediction'

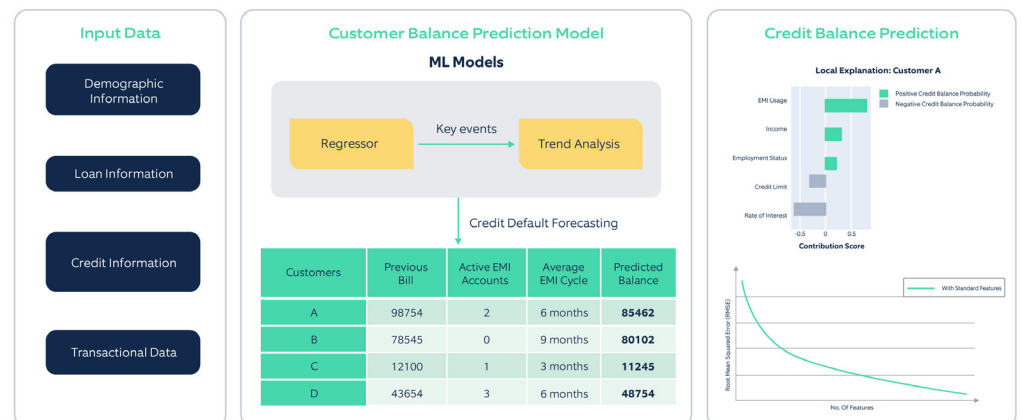


Figure 8: ML model to perform 'Credit Balance Prediction' (without alternative data source)

Figure 8 outlines the flow of an AI/ML model for predicting credit card balance utilization. Since the predictor variable is continuous, we utilize regression models. The data is sourced from various sources and processed before being fed into the regressor, which outputs the credit card balance utilization per customer. However, business requirements dictate the need for explainability, which can be achieved by employing multiple explainable models such as LIME, SHAP, etc.

As the figure shows, the 'EMI (Equated Monthly Instalment) usage' feature contributes the most to large credit card utilization. At the same time, 'Rate of Interest' has a negative impact, along with other features. It is crucial to validate the model's performance, as shown by the evaluation criteria "RMSE (Root Mean Square Error)" in the diagram. However, the evaluation criteria may vary based on the use case. As the number of features increases, the RMSE decreases, as depicted in the next slide.

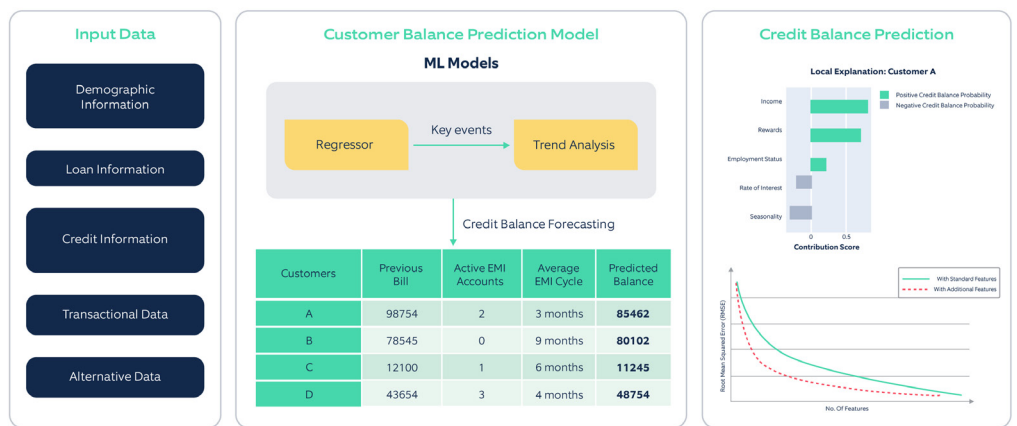


Figure 9: ML model to perform 'Credit default prediction' (with alternative data source)

6.3. Xgboost: An example of ML Modeling methodology and algorithm (for predicting credit card balances and defaults)

Xgboost is one of the powerful ML algorithms that can be used to predict credit card defaults or balance utilization. It is a tree-based algorithm designed to handle classification and regression tasks with high accuracy and efficiency. Leveraging an ensemble classifier, the model can identify the most important features contributing to credit card defaults or balance utilization and make predictions based on them. It can also handle missing data and non-linear relationships between variables, making it a useful tool for solving high complexity problems in the BFSI industry.

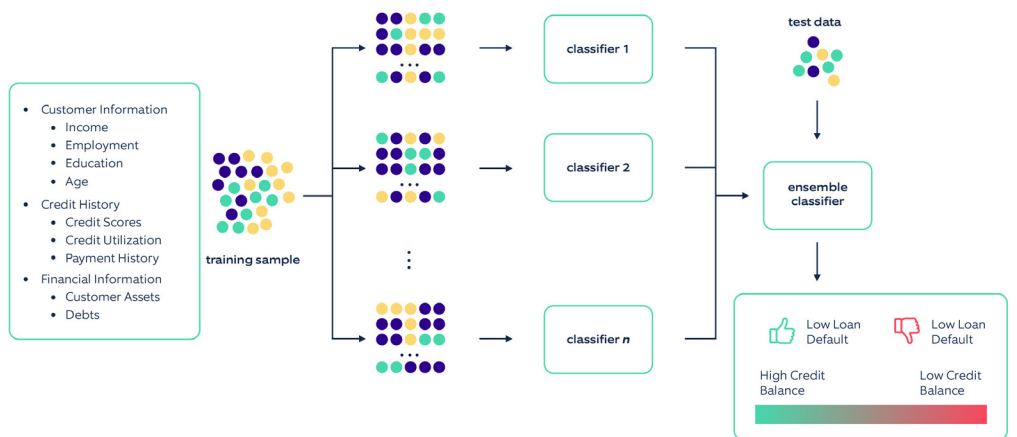


Figure 10: An overview of Xgboost algorithm

6.4. Model Maintenance and continuous operations

MLOps (Machine Learning operations) can be implemented to ensure continuous improvement in model accuracy and seamless production deployment. The MLOps pipeline will continuously monitor the model’s performance, retrain the model with new data, and deploy updated versions in production environments. This will increase the accuracy and continuity of model predictions, ultimately leading to better business decisions. Additionally, the MLOps pipeline can ensure that the model remains compliant with regulatory requirements and can be audited for transparency and accountability.

MLOps can also help uncover issues like data drift and model drift, triggering alerts for corrective action. By leveraging MLOps, banks can ensure their forecasting and loss prediction models are accurate, reliable, and scalable while reducing the time and cost associated with model deployment and management.

7. Recommendations for improving cost-to-income ratio

7.1. Manage cost

Now the first question that arises: How will credit default forecasting benefit your banking business?

The answer becomes clear when you examine the flow diagram (Figure 11). Once the model predicts potential defaulter and the reason for their default, the risk department is notified, and automated messages are sent to the users based on the identified reason. Depending on the user’s response, the bank can take any actions illustrated in the diagram. By following this process, the bank can reduce its cost-to-income ratio and improve the overall profitability of the business.



Figure 11: Recommended actions & benefits based 'Credit Default Prediction.

Now the second question: How credit card balance forecasting will be beneficial for Business?

The answer is clear in the flow diagram (Figure 12) where once the model predicts credit card balance utilization, the bank can compare the current liquidity to the forecasted demand based on the model's predictions. Based on this comparison, the bank can take any actions illustrated in the diagram. By following this process, the bank can reduce its cost-to-income ratio and improve the company's overall profitability.

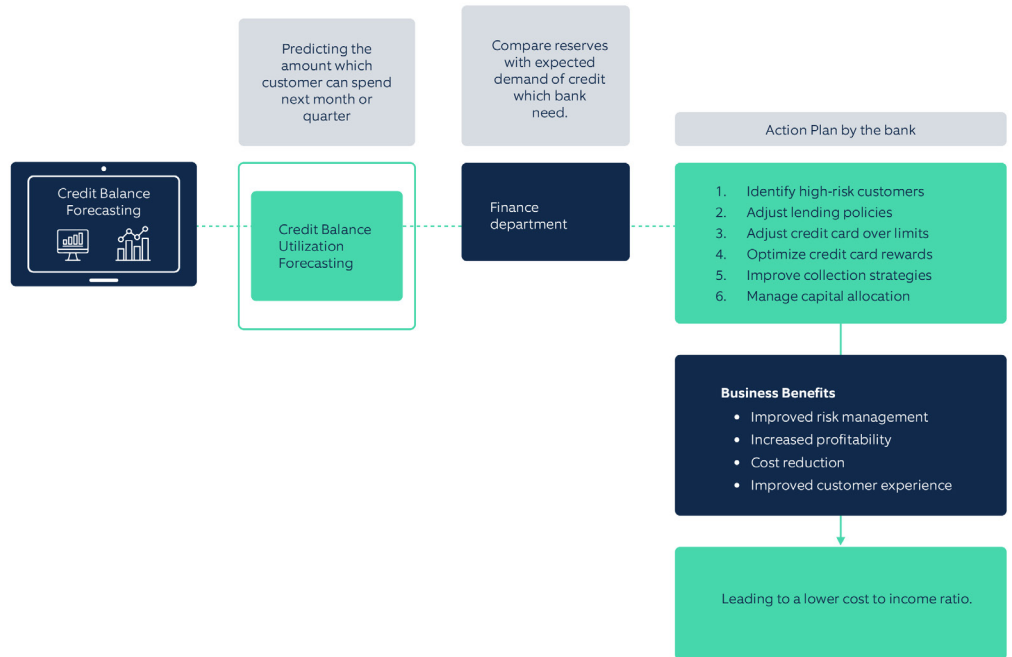


Figure 12: Recommended actions & benefits based 'Credit Balance Forecasting.'

7.2. Improve Income

Like credit card default and balance forecasting, banks can leverage the solution approach for cross-sell and upsell opportunities. These opportunities help banks to increase their customer engagement & revenue from credit cards and thus improve the cost-to-income ratio of credit card products.

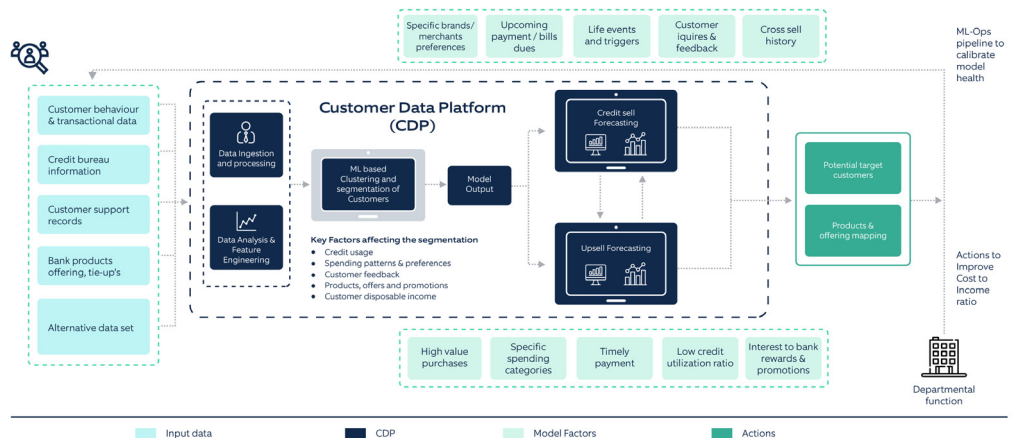


Figure 13: Solution approach to perform upsell & cross-sell forecasting and improve 'Cost-to-income Ratio' by improving income

By leveraging customer data and advanced machine learning algorithms, model can predict probability of customer acceptance of products & offers and prescribe actions to increase chances of successful cross-selling or upselling. Some of the recommendations are:

- The suggestion of preferred & personalized bank products and services (both credit card & others) to individual customers based on their transaction history, browsing behavior, and other attributes.
- Personalized marketing campaigns to send customers targeted marketing messages, rewards, and offers based on their profiles, preferences, and behavioral data. Also, prioritization of high-value customers in the marketing campaigns.
- Automated events activate cross-selling or upselling actions based on specific customer life events and milestone information (birthdays, marriage anniversary, job promotion, etc.)
- Customer sentiment analysis through feedback, reviews, and social media data to identify pain points, and uncover unmet needs, allowing for targeted cross-selling or upselling initiatives that address these areas.
- Dynamic/ optimized pricing strategies for different credit card products or features based on customer data, market trends, and competitor information.
- Real-time analytics and decision-making capabilities to identify cross-selling or upselling opportunities during customer interactions with service representatives through customer support or branch visit by offering personalized and timely recommendations.

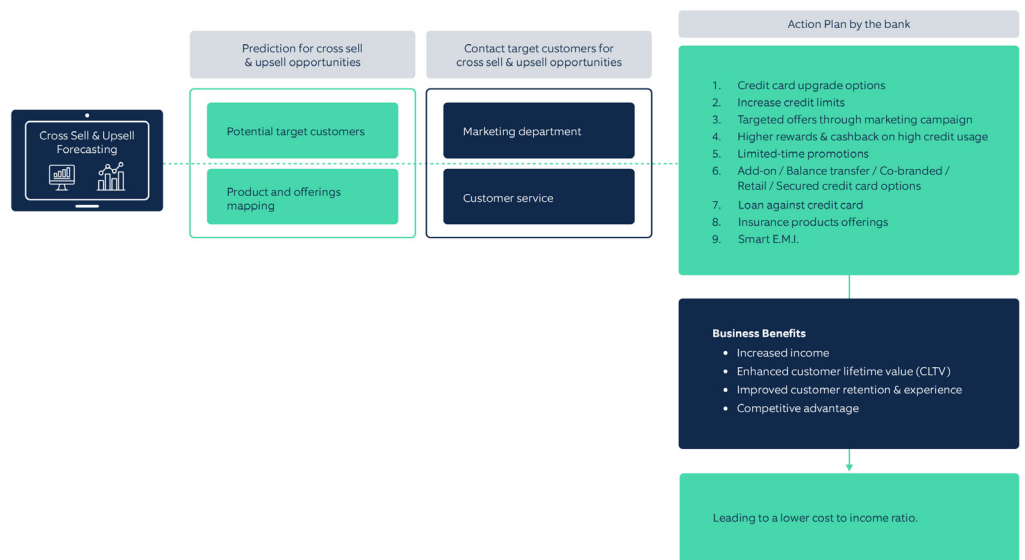


Figure 14: Recommended actions & benefits based 'Cross Sell & Upsell forecasting'

8. Endnote

Banks can improve their profitability, risk management, and capital adequacy. For example:

- **Lending automation** – Improves operational efficiency by 50% through pre-approved loans and increases revenue by 5-10%.
- **Improved default prediction** – An improvement by Gini coefficient of one percentage point saves up to \$10 million annually for every \$1 billion in underwritten loans.
- **Reduced risk-weighted assets** (by ~ \$100 billion for a global bank) – Reduces billions in capital reserves and improves **capital efficiency by 10-15%**.
- Customer segmentation and cross-selling – Increases interest income by 5-15%.
- Effective origination and underwriting processes – Reduce the sales & operating costs by 15-50%.
- Improved risk management and early warning indicators – Reduce the risk cost by 10-30%
- Improved collection processes – Increase the collector capacity by 5 to 10 percent through early self-cure customer identification and reduce charge-offs by 10-20% through customized pre-charge-off offers for customers.

Sources: Improvement in bank's return - McKinsey & Company; The analytics-enabled collections model - McKinsey & Company

9. Nagarro's take

Nagarro is a global digital engineering leader with a full service offerings including AI and ML-based solutions, Cloud data engineering, and Enterprise data & analytics. In the banking, financial services, and insurance (BFSI) industry, we enable 200+ customers around the globe in 18+ countries with our services.

Using our Thinking Breakthroughs framework, we take our design thinking approach with a product mindset in our point of view in this white paper, that

- Optimizing their financial performance by providing proactive recommendations for similar key metrics such as "Efficiency Ratio", "Loan Loss Provisions", "Capital Adequacy ratio", "Return on Assets", and "Non-Performing loans.", etc.
- Assessing risk-adjusted returns across investments, product lines, and business units to – make investment decisions, capital diversification, risk-based pricing, evaluate individual business unit contributions unit to overall profitability, etc.
- Optimize costs by identifying areas of inefficiency and enabling automation in their processes and recommending solutions
- Obtain competitive advantage by meeting demand for insights from enterprise data.

The authors



Vipul Chopra



Rahul Trehan



Sandeep Shukla



Lakshay Verma



Kamal Kumar

About Nagarro

Nagarro is a global digital engineering leader with a full-service offering, including digital product engineering, digital commerce, customer experience, AI and ML-based solutions, Cloud, immersive technologies, IoT solutions, and consulting on next-generation ERP. We help our clients become innovative, digital-first companies through our entrepreneurial and agile mindset, and we deliver on our promise of thinking breakthroughs.

We have a broad and long-standing international customer base, primarily in Europe and North America. This includes many global blue-chip companies, leading independent software vendors (ISVs), other market and industry leaders, and public sector clients.

Today, we are over 19,500+ experts across 35 countries, forming a Nation of Nagarrians, ready to help our customers succeed.

For more information, visit www.nagarro.com