PREDICTIVE MODEL IN DETECTION OF CREDIT CARD DEFAULTERS – V S KEERTHANA

ABSTRACT

This study aims to explore the feasibility and efficacy of machine learning (ML) techniques in predicting consumer eligibility for credit cards as a means of mitigating credit risk and ensuring the financial stability and credit performance of banks and financial institutions. Credit cards are a ubiquitous form of credit facility offered by banks and financial organizations worldwide, and they inherently carry credit risk due to the uncertainty of repayments, which often lead to non-performing credit facilities (NPLs). Traditional credit scoring methods are utilized by banks to evaluate applicants' creditworthiness and eligibility, but their accuracy is not always reliable. This study seeks to leverage predictive models based on ML algorithms to assist banks and financial institutions in identifying and engaging with creditworthy customers more effectively.

Commercial banks receive a significant number of credit card applications, many of which are rejected due to a variety of factors, including excessive credit record inquiries, high loan balances, and low-income levels. These evaluations are typically performed manually, which can be arduous, prone to error, and time-consuming, resulting in potential revenue loss. To address this challenge, machine learning algorithms can automate the evaluation process, a technique that is already widely adopted by most commercial banks. This notebook presents an automated credit card approval predictor utilizing machine learning techniques, similar to those used by leading financial institutions.

INTRODUCTION:

NPAs pose significant challenges for the banking sector, and credit decisions regarding whether or not to extend credit to customers can have a direct impact on reducing potential NPAs. Credit cards are a credit facility that banks use to manage credit effectively, and this project aims to leverage machine learning (ML) techniques to predict customer eligibility for credit cards.

The banking sector is currently experiencing a fundamental shift in how it engages with and builds trust with customers. Innovative business models in market lending, advances in blockchain technology, and emerging trends in investment management are driving industry dynamics. Banks, especially public sector banks in emerging markets, are under pressure to deliver superior customer service to reduce margins and build customer trust. Digitization of branches has been made possible by advances in technology. Allied Market Research reports provide valuable insights into key industry dynamics, major trends, top bank portfolio mixes, and factors that shape the competitive landscape scenario.

Credit is defined as risk, and banks face potential losses if borrowers fail to meet their obligations (interest and principal). Continuous monitoring of customer payments can reduce the likelihood of accumulating risk. Granting a credit card to a customer requires several manual steps, including assessing the applicant's creditworthiness and eligibility. However, banks recognize that credit ratings are not always accurate, and machine learning applications can eliminate manual work, time-consuming processes, and enable data-driven decisions regarding credit card approvals for customers.

Previously, banks had to rely on applicants to provide background and credit history information to determine their creditworthiness. Matching application data with reference documents was often insufficient, resulting in customers and banks struggling to authorize credit cards. By leveraging ML techniques, this project aims to mitigate credit risk and automate credit scoring by accurately predicting customer eligibility using their demographics and historical transaction data. ML enables banks to make smarter, data-driven decisions for their customers.

COMPONENT OF A CREDIT CARD

Issuer Logo: The front side of the credit card features the logo of the credit card network (e.g., Visa, Master) as well as that of the issuing bank.

EMV Chip: This chip securely stores encrypted card data to prevent the theft of credit card numbers.

Magnetic Strip: The card's magnetic strip contains account information and can be read by specific machines for monetary transactions.

Card Holder Name and Card Number: The cardholder's name and credit card number are displayed on the front of the card.

Credit Card Expiration Date: Credit cards feature an expiration date that shows the month and year, which helps merchants verify the card's validity.

Signature Box: The signature box is where cardholders are expected to sign their names.

CVV Code: The three-digit CVV code, located on the back of the card, provides additional security and protects customers' financial transactions from fraud and theft.

Hologram: On the backside of the credit card, a unique three-dimensional hologram displays the credit card network's logo, such as Visa's dove hologram or Mastercard's world map hologram.

Credit Card Industry Growth and Evolution

The credit card industry has experienced significant growth and transformation over the last decade. With a strong focus on innovative product offerings, customer loyalty, and enhanced user experiences through technology solutions, the industry is poised for exponential growth in the coming years.

Diversification of Customer Segments

Traditionally, credit cards have been marketed to high-income individuals, salaried professionals, and those living in tier I cities with good credit scores. However, to broaden the customer base and increase credit card adoption among the masses, it has become necessary to target new customer segments.

Line of Credit

A Line of Credit (LOC) is the amount of credit extended by the card issuer to the cardholder when the credit card account is opened. Until the credit limit is reached, cardholders can withdraw funds from their credit card account and repay them later. The credit line may be increased based on the cardholder's repayment capacity.

Types of Credit Cards

Most popular credit card networks/brands are Visa, MasterCard and American Express. These cards were issued by banks and financial institutions. Different types of credit cards categories are in a particular brand as well such as for low net worth, medium net worth and high net worth customers. To attract more customers, different incentives are offering such as airline miles, hotel room booking, restaurant dine-in, super market grocery buying, gift certificates to major retailers and cash back on purchases. Furthermore, in some banks have established rewards system for credit card usage. At the end of year these rewards points can be redeemed.

Branded versions of credit cards are issued to generate customer loyalty with store's name/ organization name emblazoned on the face of the cards. These credit cards called co-branded credit cards.

The **main types of credit cards:**

Standard credit cards are the most traditional type of credit card; they allow you to spend up to a pre-set credit limit, with no interest charged if you pay off the full balance by the due date every month.

Rewards credit cards operate similarly, but offer rewards such as travel points or cash back (usually in the form of a statement credit).

Balance transfer credit cards have low introductory interest rates to make it more attractive to transfer balances from other credit cards.

Charge cards may allow you to spend without limit, but you generally have to pay the balance in full each month.

Student credit cards are for college students seeking their first card.

Business credit cards can be used by business owners and employees, and may offer business-focused rewards.

Secured credit cards require a cash security deposit that reduces risk for the credit card company.

Limited-purpose credit cards include store cards and gas cards that can only be used at specific retail locations.

Prepaid cards are not true credit cards; you have to load them with cash before you can use them for purchases.

Credit Card Issuing Process

A relationship must be established before a credit card can be offered to a customer, customers and banks.

Applying for a credit card for the first time takes time. Filling, completing an application form is compulsory and most banks these days allow online application by filling it out Application Form. Choosing the right card can be done after self-study or consultation sales department executives.

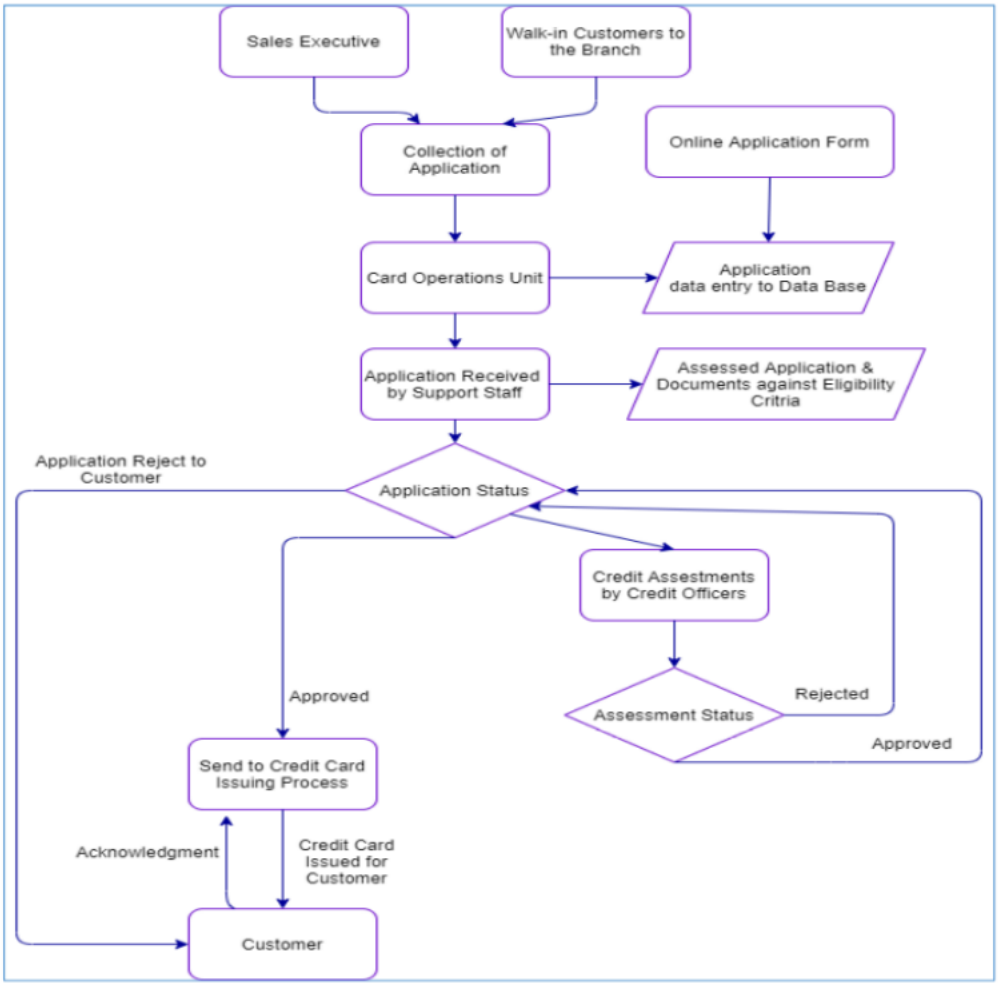


Fig: - The credit card issuing process

The credit card application process typically begins with the sales department sending an application form and required documents to potential customers. These leads may come from various sources, including branches or current customers. Upon receipt, all applications are forwarded to the credit card operating unit where the application data is entered into a card application database.

Many banks now offer an online application option, where applicants can enter their data directly through the bank's designated online portal. Once an application is received, the review team evaluates and assesses it along with accompanying documents and eligibility criteria in accordance with internal policies and procedures.

Applications meeting the bank's criteria are sent to the credit bureau and loan officer for final review and approval, based on the customer's Line of Credit (LOC). Eligible applications are not rejected, but instead added to the credit card issuance process. Finally, once approved, the applicant receives a credit card, and the bank sends a confirmation letter.

Overall, this process involves careful scrutiny and analysis of each application, to ensure only eligible customers receive a credit card, while adhering to internal policies and regulatory requirements.

Problem Statement

This problem involves a binary classification task of predicting credit card approval status. The objective is to determine whether a given credit card application will be approved or not. Another perspective on this problem is to predict the likelihood of a loan default, where an approval would indicate a low likelihood of default, and a rejection would indicate a high likelihood of default. The target variable, Approval Status, is binary, while the rest of the variables are independent or explanatory variables. Our goal is to build a model that can leverage these features to accurately predict the target variable.

Need Assessment of Machine Learning in the Credit Card Business

The role of machine learning in the credit card industry has grown significantly, providing improved accuracy in identifying patterns and predicting creditworthiness. Here's a brief assessment of the impact of machine learning on the credit card business.

Fraud Detection: Detecting fraudulent transactions is one of the significant applications of machine learning in credit cards. Machine learning algorithms can identify patterns in fraudulent transactions and flag them for further investigation, mitigating financial losses.

Credit Scoring: Machine learning can also help credit card companies make more accurate credit decisions by analyzing vast amounts of data. By identifying patterns and predicting the likelihood of a customer defaulting on their credit card payments, machine learning algorithms allow credit card companies to make more informed decisions about lending and setting interest rates.

Customer Segmentation: Machine learning can help credit card companies segment their customers based on spending habits and behaviours. This allows companies to personalize their marketing efforts, offer targeted promotions and rewards, and tailor their product offerings to meet the specific needs of different customer segments.

Personalized Offers: Machine learning algorithms can analyze customer data to offer personalized recommendations and offers to customers. This improves customer satisfaction, increases loyalty, and generates revenue for the credit card company.

Risk Management: Machine learning enables credit card companies to manage risk more effectively by identifying trends and patterns in credit card usage. This enables companies to proactively identify and manage potential risks, such as high-risk transactions or customers.

Overall, machine learning has the potential to transform the credit card industry by enabling companies to make better-informed decisions, reduce fraud and risk, and provide personalized services to customers. However, it is crucial to note that machine learning is not a panacea and should be used in conjunction with other methods and strategies to achieve the best results.

EXTERNAL SEARCHES: -

https://cloud.google.com/blog/products/data-analytics/how-to-build-a-fraud-detection-solution

https://github.com/ashishktripathi/Credit-Card-Approval/blob/master/CreditCardApproval.ipynb

https://archive.ics.uci.edu/ml/index.php

https://www.hindawi.com/journals/ddns/2021/5080472/

Github.com

Kaggle.com

Stackoverflow.com

sci-hub.hkvisa.net/

Researchnet.com

Googlescholar.com

UCI repository

TARGET SPECIFICATIONS –

The target specifications of a credit card approval prediction model may vary depending on the specific requirements of the application, but here are some general specifications that could be considered:

Accuracy: The model should be accurate in predicting whether a credit card application will be approved or rejected. The accuracy level should be high enough to minimize false positives and false negatives.

Speed: The model should be able to make predictions quickly, especially if it is being used in a real-time or near real-time application.

Robustness: The model should be robust enough to handle missing data, outliers, and other issues that may arise in real-world data.

Accessibility : The model should be explainable so that it is possible to understand why a certain prediction was made. This is important both for understanding the factors that affect credit card approval decisions and for building trust with users.

Scalability: The model should be scalable to handle large volumes of data and to work in different environments, such as cloud computing or edge devices.

Privacy: The model should protect the privacy of the applicants' sensitive information and comply with relevant data protection regulations.

Integration: The model should be easy to integrate with other systems, such as credit scoring or customer relationship management (CRM) systems.

Cost-effectiveness: The model should be cost-effective to develop, deploy, and maintain, while still meeting the required specifications.

BENCHMARKING: -

Rule-based vs ML-based Fraud Detection Systems

Rule-based fraud detection	ML-based fraud detection
Catching obvious fraudulent scenarios	Finding hidden and implicit correlations in data
Requires much manual work to enumerate all possible detection rules	Automatic detection of possible fraud scenarios
Multiple verification steps that harm user experience	The reduced number of verification measures
Long-term processing	Real-time processing

APPLICABLE GOVERNMENT AND ENVIRONMENTAL REGULATIONS: -

In the context of credit card approval prediction models, there are several applicable regulations that financial institutions, banks, and credit card companies must comply with. These regulations can be broadly classified into two categories: government regulations and environmental regulations.

Government Regulations:

Fair Credit Reporting Act (FCRA): The FCRA regulates the use of consumer credit information by credit reporting agencies, including the use of credit scores in making credit decisions. Financial institutions and credit card companies must comply with the FCRA to ensure that they are using consumer credit information in a fair and accurate manner.

Equal Credit Opportunity Act (ECOA): The ECOA prohibits discrimination in credit decisions based on factors such as race, gender, and marital status. Financial institutions and credit card companies must comply with the ECOA to ensure that they are making credit decisions based on objective factors, such as credit history and income.

Consumer Financial Protection Bureau (CFPB): The CFPB is a federal agency that regulates consumer financial products and services, including credit cards. Financial institutions and credit card companies must comply with CFPB regulations to ensure that they are offering credit products that are fair, transparent, and easy to understand.

Environmental Regulations:

Privacy Regulations: Financial institutions and credit card companies must comply with privacy regulations to ensure that they are protecting consumer information from unauthorized access or disclosure. These regulations include the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA).

Security Regulations: Financial institutions and credit card companies must comply with security regulations to ensure that they are protecting consumer information from cyber threats and other security breaches. These regulations include the **Payment Card Industry Data Security Standard (PCI DSS) and the New York State Department of Financial Services (NYDFS) Cybersecurity Regulation.**

Overall, financial institutions, banks, and credit card companies must comply with a variety of government and environmental regulations when developing and deploying credit card approval prediction models. By complying with these regulations, they can ensure that their credit products are fair, transparent, and in compliance with applicable laws and regulations.

Applicable budget constraints for the credit card approval prediction model

In building a credit card approval prediction model, the applicable budget constraints could include:

Data collection costs: Collecting the relevant data on credit history, income, employment status, and other variables that are used to assess creditworthiness may require paying fees to credit bureaus or other data providers.

Computing infrastructure costs: Running and maintaining the predictive model will require computing resources, which may involve costs such as server rental fees or cloud computing charges.

Personnel costs: Building and maintaining a credit card approval prediction model may require the expertise of data scientists, software developers, and other personnel, whose salaries and benefits will need to be factored into the budget.

Training and education costs: If the organization lacks the necessary expertise to build a predictive model in house, training or hiring outside consultants may be necessary, which may involve additional costs.

Model monitoring and maintenance costs: Once the model is deployed, ongoing monitoring and maintenance will be required to ensure that it continues to provide accurate predictions. This may involve additional personnel costs or computing infrastructure costs.

Overall, the budget for a credit card approval prediction model will depend on a variety of factors, including the complexity of the model, the amount and quality of data available, and the organization's internal resources and expertise.

Applicable expertise constraints for the credit card approval prediction model

In building a credit card approval prediction model, the applicable expertise constraints could include:

Data science expertise: Developing a credit card approval prediction model requires a strong understanding of statistical and machine learning techniques, as well as experience in data pre-processing, feature engineering, and model selection and evaluation.

Domain expertise: Understanding the credit industry, the factors that determine creditworthiness, and the regulatory and legal requirements surrounding credit approval is essential for building a predictive model that is both accurate and compliant.

Software engineering expertise: Building and deploying a credit card approval prediction model will require software development skills, including proficiency in programming languages such as Python or R, experience with version control systems, and knowledge of software design patterns and best practices.

Data governance expertise: Ensuring that the data used to train and evaluate the model is accurate, complete, and appropriately secured requires expertise in data governance and privacy regulations.

Model monitoring and maintenance expertise: Once the model is deployed, ongoing monitoring and maintenance will be required to ensure that it continues to provide accurate predictions. This may involve expertise in anomaly detection, performance optimization, and model updating.

Overall, building a credit card approval prediction model requires a diverse range of expertise, spanning data science, domain knowledge, software engineering, and data governance. The availability of these skills will be a significant constraint on the organization's ability to develop a successful predictive model.

Business Models:

The Credit Card Prediction Model can be monetized through various means, such as:

Licensing: The model can be licensed to credit card companies, financial institutions, or other businesses that use credit data to evaluate customer risk. The licensing fee can be based on model usage or a fixed fee.

Consulting Services: Consulting services can be provided to help financial institutions and credit card companies improve their credit risk assessment strategies using the prediction model. These services can include credit data analysis, model customization, and implementation support.

API Access: An API (Application Programming Interface) can be developed for the credit card prediction model, allowing other companies to integrate the model into their software. This can be offered as a paid service with a subscription fee or usage-based pricing.

Data Analysis: Credit card prediction models generate a vast amount of data that can be analyzed to gain insights into credit trends and behaviours. This data can be used to create reports or dashboards that can be sold to credit card companies or other businesses interested in the credit industry.

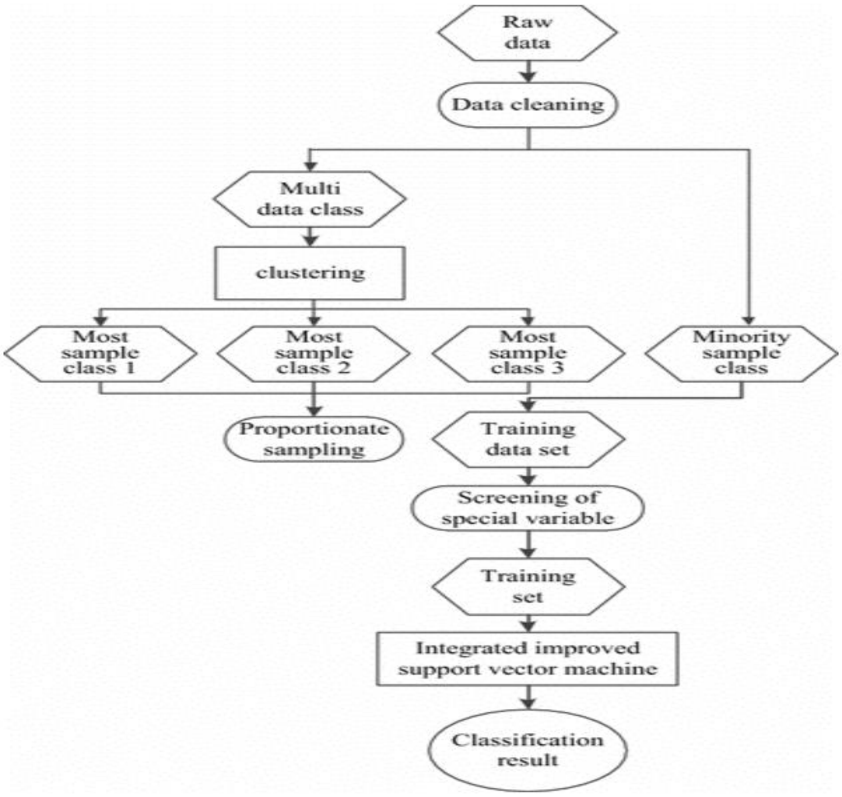
Training: Training services can be provided to financial institutions and credit card companies to help them understand the credit prediction model and how to utilize it to its full potential. These services can be offered as a one-time fee or as an ongoing subscription service.

It is crucial to note that any monetization strategy for a credit card prediction model must comply with all applicable regulations and data privacy laws. Additionally, the model should be transparent and explainable to ensure that users can understand how the model works and how its predictions are generated.

Overview:

Credit card approval prediction models are mathematical algorithms that use historical data to predict the likelihood of a credit card applicant being approved or rejected. The models typically use traditional statistical techniques and machine learning algorithms to analyse a range of variables, such as credit score, income, employment status, and payment history, to make predictions. The model typically involves data collection, pre-processing, feature engineering, model training, evaluation, and deployment. Credit card approval prediction models can help automate and streamline the credit approval process, reduce risk, and improve the customer experience, leading to more informed credit decisions, reduced losses due to defaults and fraud, and broader access to credit products.

SCHEMATIC DIAGRAM OF THE MODEL:



DATA SOURCES: -

The data set has been taken from UCI Machine Learning Repository. The structure of this notebook is as follows:

We have loaded the dataset and on displaying the data set we can see that the dataset has a mixture of numeric and non-numeric characteristics, it contains various ranges of values, and contains many missing entries.

Therefore, the dataset needs to be pre-processed so that the selected machine learning model can make good predictions.

After the pre-processing is complete, we have done exploratory data analysis to build intuitive machine learning model that can predict whether an individual's credit card application will be approved.

Data Description: -

This file concerns credit card applications. All attribute names and values have been changed to meaningless symbols to protect confidentiality of the data. This dataset is interesting because there is a good mix of attributes continuous, nominal with small numbers of values, and nominal with larger numbers of values. There are also a few missing values.

Number of Instances: 690

Number of Attributes: 15 + class attribute

Attribute Information:

A1: b, a.

A2: continuous.

A3: continuous.

A4: u, y, l, t.

A5: g, p, gg.

A6: c, d, cc, i, j, k, m, r, q, w, x, e, aa, ff.

A7: v, h, bb, j, n, z, dd, ff, o.

A8: continuous.

A9: t, f.

A10:t, f.

A11: continuous.

A12:t, f.

A13: g, p, s.

A14: continuous.

A15: continuous.

A16: +, - (class attribute)

Missing Attribute Values:

37 cases (5%) have one or more missing values. The missing values from particular attributes are:

A1: 12

A2: 12

A4: 6

A5: 6

A6: 9

A7: 9

A14: 13

9. Class Distribution

+: 307 (44.5%)

-: 383 (55.5%)

ALGORITHMS: -

Machine learning algorithm that can best solve the problem of predicting credit card approval. Some of the popular algorithms for this type of problem include logistic regression, decision trees, random forest, support vector machines, and neural networks.

Team Required: -

Building a predictive model for the credit card approval process requires a diverse team of professionals with different skills and expertise. Here are some of the roles that may be required:

• Data Analyst

• Data Scientist

• Business Analyst

• Domain Expert

• Project Manager

• Software Engineer

• Quality Assurance Engineer

• User Experience Designer

FINAL EQUATION:

Assuming that the cost of developing the credit card approval prediction model is the salary of the team members involved in the project, including a data analyst, a data scientist, and a software engineer. Let the salaries of the data analyst, data scientist, and software engineer be represented by 'da', 'ds', and 'se', respectively. So, the total cost c = da + ds + se.

Let's assume that the credit card company charges a fee of Rs.500 for each approved application. So the revenue generated per unit time (let's say one month) can be represented by r = 500*y(t), where y(t) is the number of approved applications in that month.

Then the profit or financial equation can be represented as:

p = r - c

p = 500*y(t) - (da + ds + se)

This equation shows that the profit generated by the credit card approval prediction model is dependent on the number of approved applications and the cost of developing the model. The credit card company can use this equation to estimate their potential profits and make decisions on pricing and resource allocation.

Here's an example of the data that could be used for credit card fraud detection:

Transaction ID	Amount	Location	Time of Day	Fraudulent
1	50.00	New York	08:00	0
2	200.00	Los Angeles	13:30	0
3	5000.00	New York	20:15	1
4	100.00	Chicago	11:00	0
5	2500.00	Miami	18:45	1
6	75.00	San Francisco	09:30	0
7	3000.00	New York	22:00	1
8	150.00	Miami	15:00	0
9	450.00	Chicago	10:45	0
10	1000.00	San Francisco	17:30	1

Assuming that the ML model takes 2 weeks to develop and the salaries of the team members are $5,000 for the ML engineers and $7,000 for the full-stack web developer, the total cost of developing the model would be $17,000.

Let's assume that with the implementation of the fraud detection system, the company can prevent losses of up to $500,000 per year due to fraudulent transactions. With an inflation coefficient of 0.98, the projected profits over the next 20 years would be:

$500,000 * 20 * 0.98 = $9,800,000

Of course, these numbers are just for example purposes and would vary based on the specific circumstances and data of the company.