A Machine Learning & NLP Based Credit Scoring System

Contact Us

Key Details/Highlights Of The Solution

CHALLENGE

Creating a robust credit scoring system with the help of Machine Learning (ML) and Natural Language Processing (NLP) to help SMBs like our client as well as the financial institutions to process loans in real-time without much of a hassle and with minimum default rate.

SOLUTION

Developing a data source and introducing Machine Learning and Natural Language Processing to automate the generation of credit scores.

Client

We created this system for an African-based SMB that provides small loans to mostly the people without banks. The company asked for a simple streamlining system for loan approvals based on the data source that is more credible than the credit bureau’s data. This streamlined system allowed them to expand their business without worrying about the credibility of loan approval.

Challenge

The most challenging part of this project was to find a reliable data source for the client that is better than Credit Bureau's data. The second challenge we faced was defining the set of parameters for calculating the credit score in the system.

Solution

The basic solution to what they needed was to build a data source that is not available on any traditional channel like Credit Bureau. So, we built a system on the basis of mobile device data which included extraction of data like income, investments, expenses, type of mobile devices, installed apps, etc with the help of NLP. This information accumulated by the NLP is then passed to the credit score system where the parameters were set. Due to the high volume of the data, this process did not work for long. To resolve this issue we introduced Machine Learning to automate the generation of credit scores.

The Approach We Used –

Here is how we implemented the plan we had for the company. We created this system in 4 steps –

- **Data collection from mobile devices**

  We fetched the information from the mobile SMS data of the customer using mobile SDK after getting the consent of the customer. After fetching the information, we stored the data on the server database of the system.

- **Data Warehousing**

  After collecting data, we then stored the data into a common data warehouse that was built for the system. After the successful warehousing, the data then gets transformed in the way so that the NLP engine can pass it to the credit scoring engine.

- **NLP engine implementation to fetch the information from the data**

  This is the engine designed to read the raw SMS data being collected from the customer's mobile devices and fetching the desired information like expenses, income, investments, and more depicting the financial strength of the customer.

- **Credit score engine installation**

  After this, the information including the financial strength and the type of device, contact list, and more are fed to the ML-based credit score engine. At first, this system was created on a rule-based algorithm in which the data gets weighed and is tagged to different data points. Once we had enough data points, the algorithm was upgraded and was trained to calculate credit scores automatically.

Conclusion

As per the requirement of the client, they needed a system to calculate the credit score on the basis of a more reliable data source than Credit Bureau Data. For this, we designed an efficient system that collected the data from customers' mobile devices and calculated the credit score on the basis of that. This entire process helped them speed up the process of client conversion.

The System Architecture We Built For Them –

Input Data

Once, the user opens the app they are provided with the following options - Choose the image from a database or click the image through the camera.

Image Processing

After that, the cropped image is passed to the processor algorithm where the image is converted in the set format compatible with the classifier as well as the device.

Detection

Once, you select the option and the image is clicked or selected, the image will be passed to a face detector algorithm.

Classifier

The processed image is passed to the classifier and the result is then stored in the local SQLite storage.

Output

Whenever the user uses the system to mark the attendance, the image is processed through the complete pipeline in which the output of the classier is compared to the database

Conclusion

We helped the client establish an offline hands-free attendance system for their company. While they were concerned about the internet facility in their areas of operations.

Want to Build a ML Based Credit Scoring System

GET STARTED