# BEHAVIOURAL SCORING BASED ON SOCIAL ACTIVITY AND FINANCIAL ANALYTICS

## 15IT496L MAJOR PROJECT

### PROJECT REVIEW REPORT - 3

*Submitted by*

**Subham Pramanik [Reg No: RA1611008010376]**
**Sanidhya Pandey [Reg No: RA1611008010320]**
**Harsh Krishna [Reg No: RA1611008010348]**
**Anmol Gupta [Reg No: RA1611008010364]**

*in partial fulfilment of the requirements for the degree of*

**BACHELOR OF TECHNOLOGY**


**DEPARTMENT OF INFORMATION TECHNOLOGY**

**FACULTY OF ENGINEERING AND TECHNOLOGY**

**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**

**KATTANKULATHUR- 603 203**

**March 2020**

# BONAFIDE CERTIFICATE

Certified that this project report titled "**BEHAVIOURAL SCORING BASED ON SOCIAL ACTIVITY AND FINANCIAL ANALYTICS"** is the Bonafide work done by Subham Pramanik [Reg No: RA1611008010376], Harsh Krishna [Reg No: RA1611008010348], Sanidhya Pandey [Reg No: RA1611008010320], Anmol Gupta [Reg No: RA1611008010364], of IT B.Tech Degree course in the theory 15IT496L – Major Project during the academic year 2019-20.

**Name:** Mr P. Gouthaman

**Designation:** Assistant professor (O.G)

**Department:** IT

**Signature of Guide:**

**Date of Submission:** 10th March 2020

# TABLE OF CONTENTS

# ABSTRACT

Credit scoring is without a doubt one of the oldest applications of analytics. In recent years, a multitude of sophisticated classification techniques has been developed to improve the statistical performance of credit scoring models. Instead of focusing on credit scoring itself, this paper leverages alternative data sources to enhance and incorporate other factors to determine the overall persona of a person. This project identifies unique factors through a person's online presence and his financial record to provide them with a score that signifies his behaviour.

The project demonstrates how a person's online activity on social media sites, such as Facebook and Twitter determines the nature and behaviour of the person. Some factors that are included for the social scoring are, types of posts shared, comments added, posts posted, pages followed and liked. These data are plotted against a graph signifying the time and we obtain a social score.

There is a financial scoring model that will determine the person's financial fitness and likelihood to engage in criminal activities due to financial deformity. Combining both social scoring and financial scoring at a specific weight will provide us with the behavioural score.

This score will classify the subjects and help determine good citizens among the rest. Subjects with higher behavioural scores will show the promise and practice of a good citizen. This can be used to engage and provide added incentives to good citizens in order to promote good citizenship.

# CHAPTER 3

# PROPOSED WORK

This project, "**Behavioural Scoring Based on Social Activity and Financial Analytics**" presents an outline for the construction of a so-called 'Social Credit System'.

The aim of this project is to focus on the development of a scientific and methodological approach to determine a scoring mechanism to assess and score human behaviour based on online social media activity and financial activity. The scores will help reward people with good behaviour and hence promote good citizenship.

This project will provide a credit score to individuals based on various factors. Personal factors include sincerity, honesty, and integrity, which are key determinants of an individual's personality. Financial factors include a person's spending, how he handles the money, timely payment of loans, EMI's. The financial aspects are evaluated under the conventional banking framework, by analysing accounting statements and financial projections.

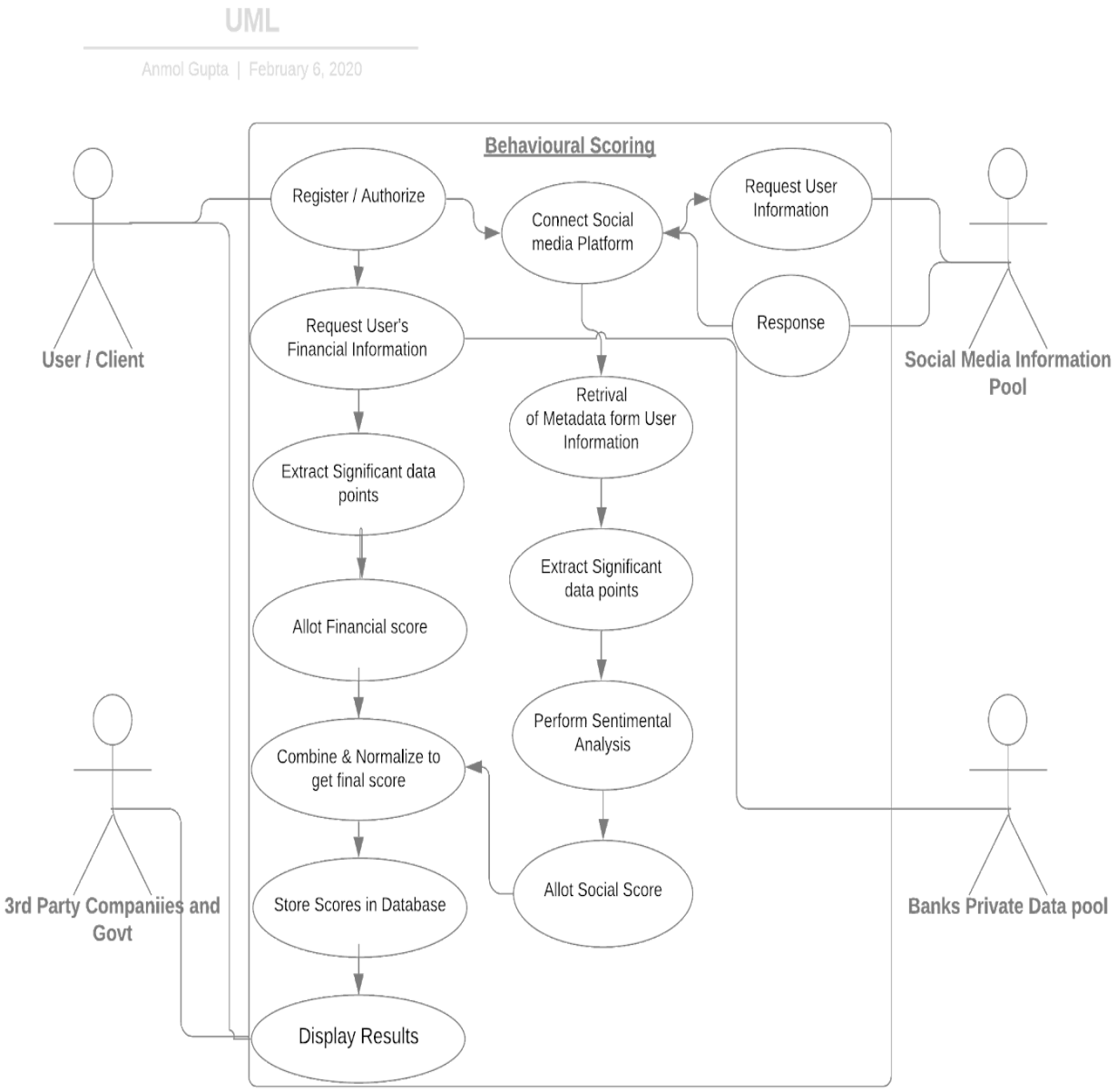
In a nutshell, its main innovation, once fully implemented, could be that each citizen will be given a score measuring their sincerity, honesty, and integrity and that this score will then be a major determinant for their lives, for instance, whether to be able to get credit, rent a flat, or buy a plane ticket, or being given preferred access to hospitals, universities and government services.

Social Credit score will consider a wide range of personal factors. It also resembles, but goes further than, a range of systems that are intended to increase the prominence of reputation in relation to transactions, online platforms and in the 'sharing economy'. Thus, we focus here on rating systems concerning individual persons.

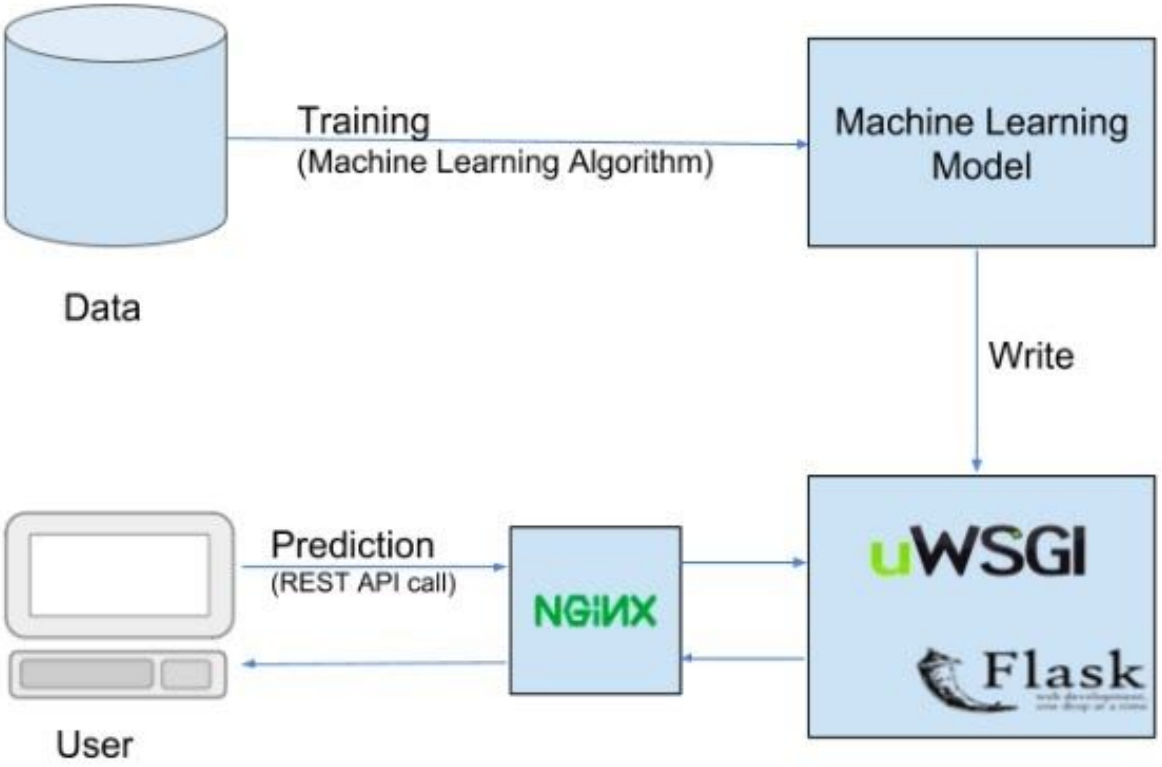
The financial aspects are evaluated under the conventional banking framework, by analyzing accounting statements and financial projections. The social aspects try to quantify the loan impact on the achievement of Millennium Development Goals such as employment, education, environment, health or community impact. The social credit score model should incorporate the lender's know-how and should also be coherent with its mission.

Scoring alone on the basis of financial aspect may risk the institution to let a socially bad person getting loans and other financial benefits. A socially bad person may tend to be a defaulter or use financial benefits for unethical purposes. Therefore keeping this view in mind, we proposed a methodology which will score a person based on both social and financial aspects.

## 3.1 Proposed System Workflow



## 3.1 System Architecture Diagram



# CHAPTER 4

**Step 1: Input Access Token**

**Step 2: Enter Financial Information**

**Step 3: View Result**

## 4.1 Implementation and Tools Used

Our project has been implemented in Python. As discussed before, we had implemented the project in 3 modules - Financial module and behavioural module. The web framework which we have used is Flask.

For the implementation of the project we have used-

**Monkeylearn API** - MonkeyLearn requires that you authenticate by sending an API Key with each request to grant you access to the API. Every MonkeyLearn account has its own API Key.

**Graph API** - The Graph API is the primary way for apps to read and write to the Facebook social graph.

**Random Forest Classifier** - A random forest is a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting.

# Modules

The project is divided into two major modules.

## 4.2 Behavioural Scoring -

The objective of the social scoring module is to classify and provide a score to a person based on his social media activity.

**Inspection of Parameters**.

- Feed

-Posts

-Likes

**Fetching data using Facebook Graph API**

-Generation of Permanent Access key

-JSON/CSV files for Sentiment Analysis

-Cleaning data, removing noise

**Sentiment Analysis** - MonkeyLearn API

- Weight of Evidence and assignment to parameters.

WoE = (ln(Distr Goods/ Distr Bads)) * 100

**Allot ranks**

-Posts are weighted by time

-Returns rank of every respective series passed

**Ranking based on Time**

-This is plotted between Weight and Time.

-Max weight is set at 25 and rank is dynamic.

-This helps in giving more priority to data which is at the median rather than the most recent. This is because old data is of very less use. New information can be fished or faked.

**Calculation of Scope Parameter (ScopeP)**

For the parameters which are not time dependent

ScopeP = Senti_Value.

For Posts & Likes, it is given by,

ScopeP = Summation of (Weighted time * Senti_Value)

**Final Calculation of Scores**

This is calculated as,

Summation of (Weight_parameter * ScopeP) / Total number of observations

## 4.3 Financial Scoring -

The objective of this module is to run an analysis on a person's financial history and a lot of scores based on his financial activity.

**Selection of Credit Scoring Model**

-Choosing the desired model as per constraints

**Training and Analysis of Data**

-Train model to satisfy all the factors

-Run data with model

-Finally, obtain a credit score

**Score Normalization**

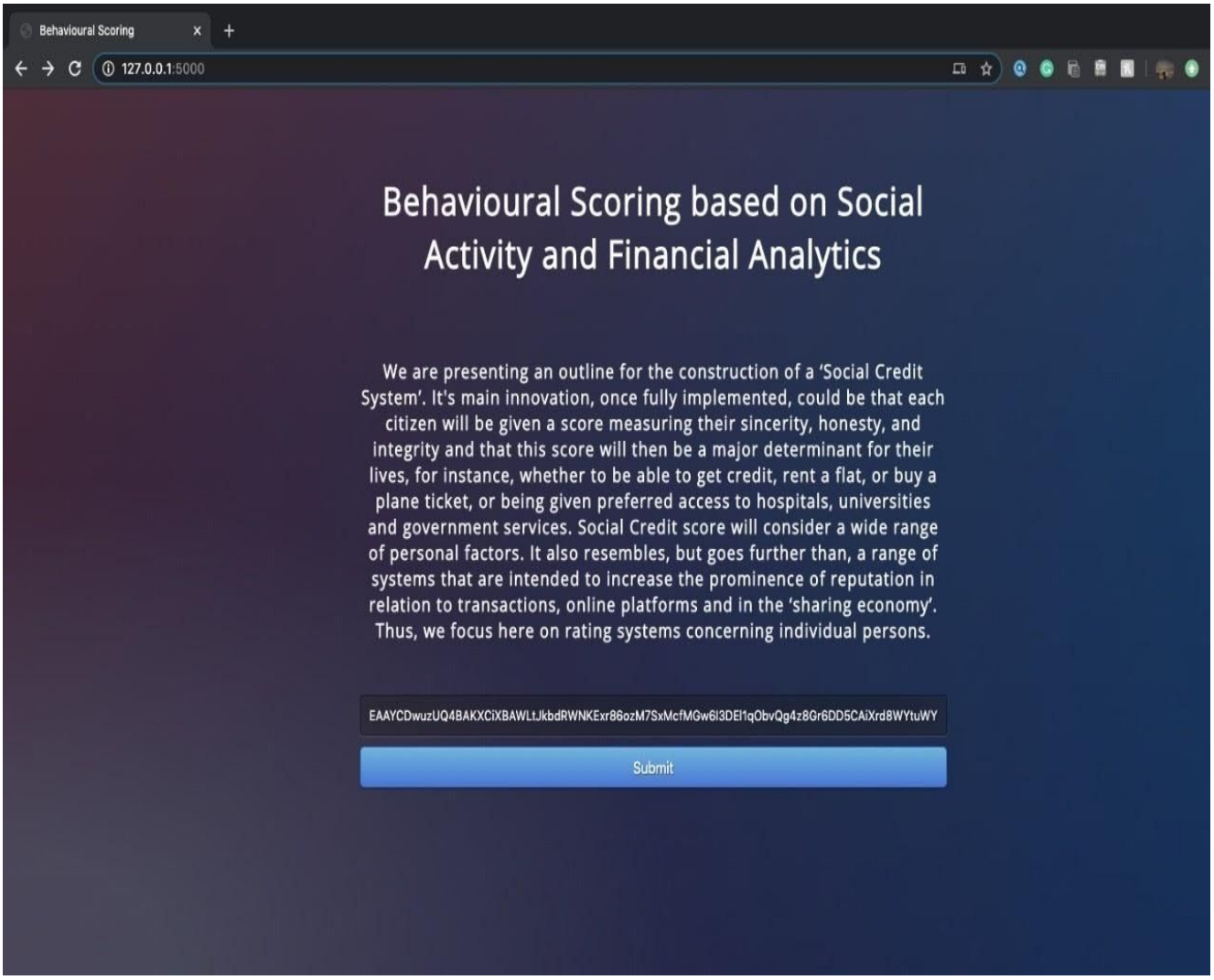
-Combining of Financial and behavioural scores

-Store final scores in Database

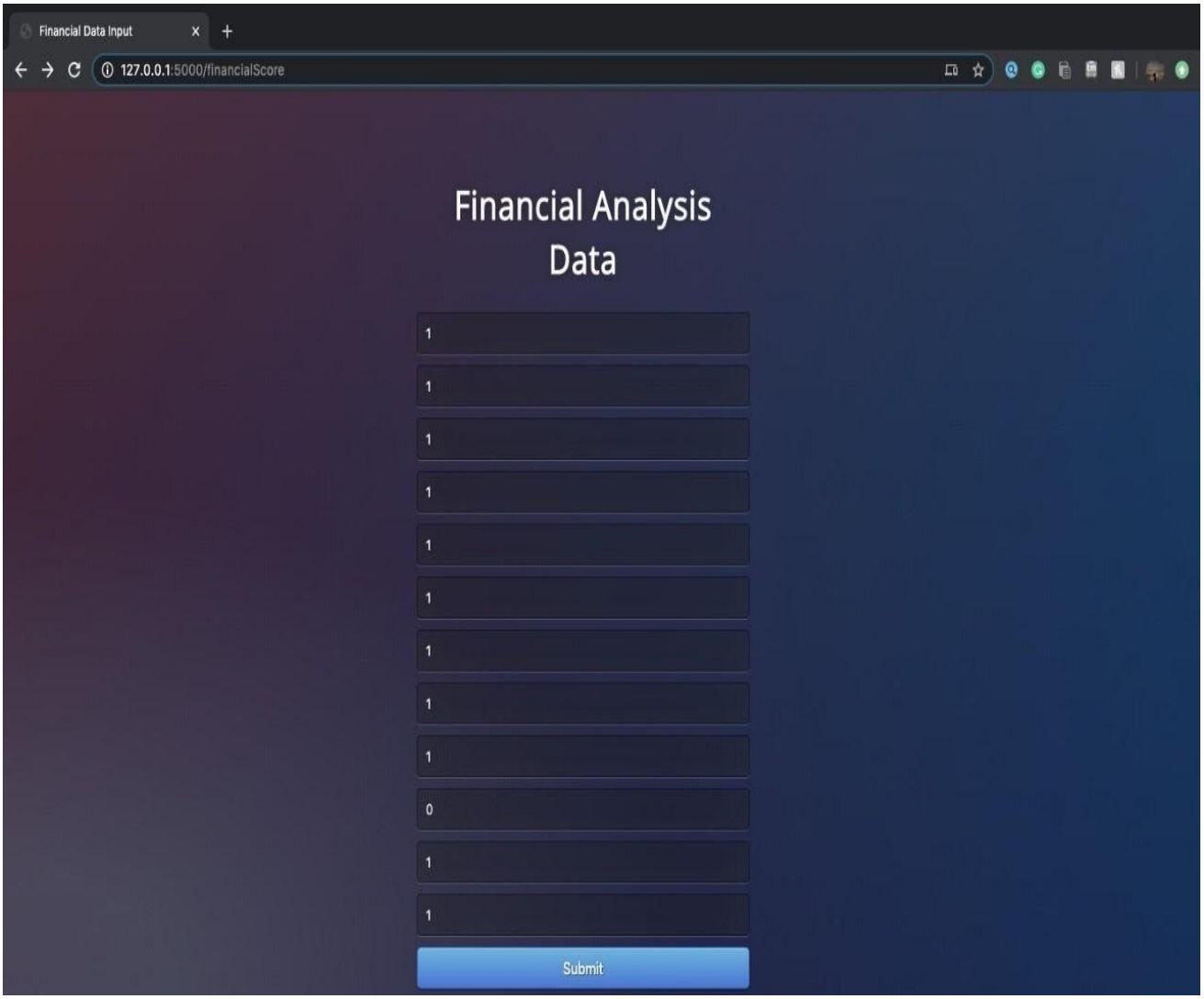
-Provide scores to users and third party partners

# 4.4 RESULTS

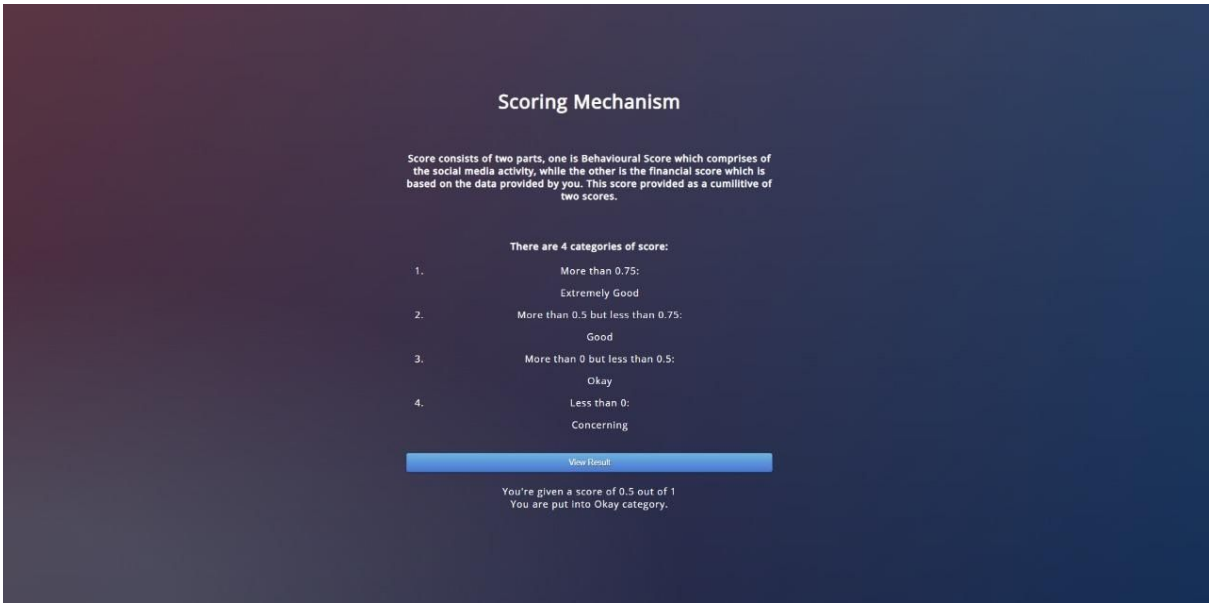
**Entering Access Token**



## Entering Financial Information



**View Result**




# 4.5 CONCLUSION

**Score groups-**

0.75-1: “Excellent”

0.5-0.75: “Good”

0-0.5: “Okay”

-1-0: “Concerning”

Based on the above scoring groups, an individual can be classified into various groups.