**Personality Prediction**

**B.E. Project Report**

Submitted in partial fulfillment of the requirements For the degree of

### Bachelor of Engineering in

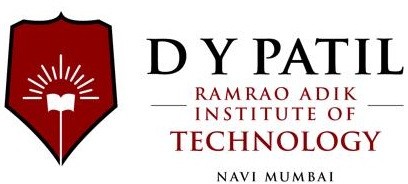
**Computer Engineering**

Submitted by

### Ms. Shital Nehete 17CE8008 Ms. Mansi Pable 17CE7001 Mr. Jeet Choudhari 17CE7027

Guided by

### Mrs. Smita Bharne



Department of Computer Engineering Ramrao Adik Institute Of Technology

Dr. D. Y. Patil Vidyanagar, Sector-7, Nerul, Navi Mumbai-400706. (Affiliated to University of Mumbai)

**November 2020**

# Ramrao Adik Institute of Technology

(Affiliated to the University of Mumbai)

Dr. D. Y. Patil Vidyanagar, Sector-7, Nerul, Navi Mumbai-400706.

**CERTIFICATE**

*This is to certify that, the project ’A’ titled*

### “ Personality Prediction ”

*is a bonafide work done by*

### Ms. Shital Nehete 17CE8008 Ms. Mansi Pable 17CE7001 Mr.Jeet Choudhari 17CE7027

*and is submitted in the partial fulfillment of the requirement for the degree of*

**Bachelor of Engineering**

in

**Computer Engineering**

to the

**University of Mumbai**



Supervisor

**(Mrs. Smita Bharne)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Project Co-ordinator |  | Head of Department |  | Principal |
| **(Mrs. Smita Bharne)** |  | **(Dr. Leena Ragha)** |  | **(Dr. Mukesh D. Patil )** |

Project Report Approval for B.E

This is to certify that the project ‘A’ entitled ***“ Personality Prediction”*** is a bonafide work done by ***Ms. Shital Nehete***, ***Ms. Mansi Pable***, and ***Mr. Jeet Choudhari*** under the supervision of ***Mrs. Smita Bharne*** This dissertation has been approved for the award of ***Bachelor’s Degree in Computer Engineering, University of Mumbai***.

Examiners :

1...............................

2...............................

Supervisors :

1...............................

2...............................

Principal :

. . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

Date :. . . /. . . /. . . . . .

Place :. . . . . . . . . . . .

## Declaration

We declare that this written submission represents my ideas in my own words and where others ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken whenneeded.

|  |  |  |
| --- | --- | --- |
| **Ms. Shital Nehete** | **17CE8008** |  |
| **Ms. Mansi Pable** | **17CE7001** |  |
| **Mr. Jeet Choudhari** | **17CE7027** |  |

Date :. . . /. . . /. . . . . .

# Abstract

In recent years, the problems of mental health of people are increasing day by day due to increase of the stress levels. Employment of less skilled people results is reducing of the outcome of the organization. To solve these problems, the personality of the person plays a major role. So, this project aims at predicting the personality of the person in Big-Five traits. The prediction can be done by 3 processes i.e. text, questionnaire and social media. The machine learning model used to predict the personality uses myPersonality dataset. It uses tf-idf vectorization to convert words to vectors and Random Forest Algorithm for all the prediction. Social media also helps to compare the personality of the person with the friends it has on it. Radar graph is used to show and compare the personality division on the web app. Web app provides a proper interface for the processes and display of the personality for easy access.

# Contents

|  |  |
| --- | --- |
| **Abstract** | **i** |
| **List of Figures** | **iv** |
| **List of Tables** | **v** |
| **List of Algorithms** | **vi** |
| **1 Introduction** | **1** |
| 1.1 Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 1 |
| 1.2 Objective . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 1 |
| 1.3 Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 1 |
| 1.4 Organization of report . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 1 |
| **2 Literature Survey** | **2** |
| 2.1 Survey of Existing System . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 2 |
| 2.2 Limitations of Existing System . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 2 |
| 2.3 Problem Statement . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 2 |
| 2.4 Scope . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 2 |
| **3 Project Proposal** | **3** |
| 3.1 Proposed Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 3 |
| 3.2 Proposed Methodology . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 3 |
| 3.3 Details of Hardware/Software Requirement . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 3 |
| **4 Planning And Formulation** | **4** |
| 4.1 Schedule for Project . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 4 |
| 4.2 Detail Plan of Execution . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 4 |

1. Design of System 5
   1. [Design Diagram with Explanation . . . . . . . . . . . . . . . . . . . . . . . . . 5](#_TOC_250004)
2. Expected Results 6

[6.1 Implementation Details . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6](#_TOC_250003)

[6.2 Project Outcomes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6](#_TOC_250002)

1. Conclusion and Future Work 7

References 7

[Appendices 10](#_TOC_250001)

1. Weekly Progress Report Project A 10
2. Plagiarism Report 11

[Acknowledgement 12](#_TOC_250000)

1. Fig 3.2.1 Proposed Methodology
2. Fig 4.1.1 Gantt Chart
3. Fig 5.1.1 Design of the System
4. Fig 6.1.1 myPersonality Dataset
5. Fig 6.1.2 Generation of Pickle files
6. Fig 6.1.3 Training
7. Fig 6.1.4.1 Text Predictor Example 1
8. Fig 6.1.4.2 Text Predictor Example 2
9. Fig 6.1.4.3 Text Predictor Example 3
10. Fig 6.1.4.3 Text Predictor Example 4
11. Fig 6.1.5 Dataset for Questionnaire
12. Fig 6.1.6.1 Questionnaire Example 1
13. Fig 6.1.6.2 Questionnaire Example 2
14. Fig 6.1.7.1 Facebook Login
15. Fig 6.1.7.2 Navigating to See your profile
16. Fig 6.1.7.3 Navigating to Friends tab
17. Fig 6.1.7.3 Extracting Friends
18. Fig 6.1.8.1 Questionnaire
19. Fig 6.1.8.2 Submitting Questionnaire
20. Fig 6.1.8.3 Questionnaire Results
21. Fig 6.1.8.4 Text Predictor
22. Fig 6.1.8.5 My Network tab
23. Fig 6.1.8.6 Comparison with friends

**1.** 2.1 Survey of Existing System

**Introduction**

### Overview

### The traits, set of etiquettes, emotional patterns define one's personality which plays an important role in taking major life decisions. The capability of personality features to forecast the life upshots have been interrogated due to the presumed repercussion of personality. Predicting personality with the help of OCEAN model can be obliging in order to screen a particular person or diagnose metal health, or to improve his/her own personality. As when we acknowledge our own personality, we can question our requirements also reform our behavior.

### Social media is the social interlinkage among people in which there is creation, sharing or exchange of information and ideas in virtual communities and networks. Nowadays, people use more and more time in social media, every time there is a flood of information creation and dissemination. This information includes user’s social behavior, user generated texts and language habit, to some extent reflects user's personality. This project tries to use user generated information on social network which is easy to get to predict user’s personality.

### This Big 5 Personality Traits, also known as OCEAN stands for Openness, Conscientiousness, Extroversion, Agreeableness, Neuroticism. According to analysts, every single person is born with these personality traits. Individuals who overlay high in the first trait are often eccentric. The more organized you are, more you rate in Conscientiousness. "I am the life of the party", can be the best example for Extroversion. Low Agreeableness shows logical behavior. Stress, mood comes under Neuroticism. Thus, with the help of these five personality traits listed out under OCEAN model we can easily predict individual’s personality.

* 1. **Objective**

The objective of this project is to predict personality based on the 5 features (BIG5 – O.C.E.A.N.) which stand for Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism in order to diagnose psychological problems as well as assessment of candidates. Specifically, the following were defined to achieve the general objective of the study:

1. Develop an app to screen candidates for college, employment.
2. To diagnose psychological problems.
3. Focusing on strengths, weaknesses, temperament, and style of leadership.
4. Predict personality at workplace as well as for personal inventories.
   1. **Motivation**

Personality Prediction with the help of BIG 5 model i.e. OCEAN model wouldhelp the young youth not only to predict their personality but also improve in that particular aspect. For eg, if a user gets a low rate in Openness implies that he/she is consistent and cautious. Thus, the user can improve this that particular personality feature. As social media is the best platform for social interaction, where people spend most of their time. Hence, comparing friend's personality features with user's traits can help screen that person with all aspects of OCEAN model.

Accordingly, the motivation can be summarized as:

* + Screen candidates based on predicted personality with the help of questionnaire
  + To help user compare his personality with others.
  + Diagnose mental health problems.
  1. **Organization of report**

Personality of a person plays a vital role in the response what a person gives when exposed to a situation. This study can be helpful not only with respect to psychology but also to screen candidates at workplace. Thus, there is need of something through which we can automate the process. This process can predict the common personality traits in people which are Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism and their probability. This is usually known as OCEAN model.

We did research on various papers for the Personality Prediction System and analyzed those systems. This was helpful with many respects to design our Personality Prediction model and overcome all the limitations which have been observed in previously developed systems. The limitations which have been overcome by our Personality Prediction System helped to improve the accuracy and get more precise results.

The approach which we followed to propose our work based on the Personality Prediction is well explained in proposed work and proposed methodology. Also the details for hardware and software used in the same which ease the work of implementation are mentioned.

The proper week-wise schedule for the development of this Personality Prediction system has been shared in the schedule section which also can be understood with the help of detailed explanation following the schedule.

The design of the system that we implemented easily explains the definite working of it. All the details have been properly shown in that respective diagram. Also, for all the 3 sections of the system, individual diagrams have been developed. All the diagrams are followed by proper explanation.

The result observed by the developed system can be found out in this report which is properly labeled with respective information. This helps in understanding the system better. The results also include the accuracy of the model which has been trained for predicting purpose.

Hence this Personality Prediction system can be used in various ways right from the recommendation system to recruitment process for screening candidates. Some future work has to be done for further improvement of the developed system.

**Literature Survey**

* 1. **Survey of Existing System**

|  |  |  |  |
| --- | --- | --- | --- |
| Sr. No | Author | Title | Summary |
| 1. | Majid Ramezani, Mohammad Reza Feizi Derakhshi, Mohammad-Ali Balafar, Meysam Asgari-Chenaghlu, 2020 | Automatic Personality Prediction; an Enhanced Method Using Ensemble Modeling | The major objective of this study is to reinforce the accuracy of Automatic Personality Prediction from the text by using ensemble modeling.  Automatic Personality Prediction is that the automated speculation of the  personality on differing types of human generated contents such as text, speech, image, video. Here five different basic methods, having various processing level which are to be applied in ensemble modeling are introduced which incorporates term frequency vector-based, ontology-based, enriched ontology-based,  latent semantic analysis-based, and deep learning-based methods.  Essays Dataset which are written by psychology students is employed altogether the methods  It contains 2,467 essays and a 1.9 million words. Further (hierarchical attention network) HAN is employed because the meta-model to implement the stacking algorithm. |
| 2. | Madhura Jayaratne, Buddhi Jayatilleke, 2020 | Predicting Personality Using Answers to open ended interview questions year | In this paper measure of once personality is completed to form better hiring decisions within the selection process  Here data from over 46,000 job applicants who completed a web chat  interview that also included a personality questionnaire was collected.  Using nlp and ml a regression model was build to infer HEXACO trait values from textual content  Further the performance of 5 different text representation methods were calculated and it had been found that term  frequency-inverse document frequency  (TF-IDF) with Latent Dirichlet Allocation (LDA) topics performed the simplest .  The training dataset for the experiment was collected from  PredictiveHire1 FirstInterview (™) product which is a web chatbased interview tool  that contains 7 open ended questions and 40 self-rating questions. |
| 3. | Aditi V. Kunte, Suja Panicker, 2019 | Using textual data for Personality Prediction: A Machine Learning Approach | Three algorithms have been used to predict the personality with the help of Machine Learning approach which are namely: Multinomial Naïve Bayes, AdaBoost and LDA. The relevances are compared and the higher value is taken. Multinomial Naïve Bayes is observed to be giving 73.43 accuracy, precision as 0.7, recall of 0.71 followed by 0.72 as F1-score. |
| 4. | Abdur Rahman, Asif Al Faisal, Tayeba Khanam, Mahfida Amjad, Saeed Siddik, 2019 | Personality Detection from Text using Convolutional Neural Network | The textual data is converted into numbers using Word2Vec model and then the data containing non-neural sentences is filtered then it is finally given to deep convolutional neural network for classification. It aims at comparing the performance of 3 activation function such as sigmoid, ReLu and leaky ReLu. |
| 5. | Guanqun Sun, Ao Guo, Jianhua Ma and Jianguo Wei, 2019 | Personal Trait Analysis Using Word2vec Based on User-generated Text | It proposes a Personal Trait Matrix model for the personality prediction. Words are extracted using LDA or lexions and then converted into vector using word2vec to form a matrix for different individuals. They scaled diversity on affect and social interaction with personal traits. |
| 6. | Raad Bin Tareaf, Seyed Ali Alhosseini, Philipp Berger, Patrick Hennig, Christoph Meinel, 2019 | Towards Automatic Personality Prediction Using Facebook Likes Metadata | It used Facebook’s hierarchy to categorical pages as features and given it as input to ML model for predicting Big-Five personality traits scores. It established a relation between likes metadata and mapped it to feature to calculate the big-five scores. |
| 7. | Mariam Hassanein Wedad Hussein Sherine Rady  Tarek F. Gharib, 2018 | Predicting Personality Traits from Social Media using Text Semantics | It aims at measuring similarities between user’s text and vectors representing different personality traits. The score is presented to how much user’s text is similar to the personality vector using vector space model. |
| 8. | Raad Bin Tareaf, Philipp Berger, Patrick Hennig, Christoph Meinel, 2018 | Personality Exploration System for Online Social Networks: Facebook Brands As a Use Case | It used Facebook’s hierarchy to categorical pages as features and given it as input to ML model for predicting Big-Five personality traits scores. It established a relation between likes metadata and mapped it to feature to calculate the big-five scores. |

* 1. **Limitations of Existing System**
* Smaller dataset - myPersonality sample dataset (250 users, 9917 status updates).
* Semantic level analysis (terms and topics) used. Other types of features, such as the use of parts of speech (POS), readability, formality, use of emojis can further increase the accuracy.
* Social networking sites like Facebook doesn’t directly provide needed data to researchers due to privacy reasons.
* The accuracy of the analysis can have great impact due to the millions, trillions fake users on Facebook.
* Dynamic changes in likes and number of friends can affect the accuracy.
* Only one way to predict the personality.
  1. **Problem Statement**

Personality of the person plays a vital role in the response what a person gives when exposed to a situation. It can be difficult to predict when we meet the person for the first time. Study of personality is considered as a psychology research based on the survey or questionnaire. Due to this, research data gets limited to fewer questions. Hence there is a need of something through which we can automate the process. The project proposes prediction of user’s personality with 3 ways namely Text prediction, Questionnaire, Facebook.

* 1. **Scope**

The main aim of the project is to predict the personality based on 5 traits i.e. O.C.E.A.N. The predicted personality can be used for recommender systems, job satisfaction and requirement details, educational purposes, psychologically and much more. So, this project aims to help the health care sector, educational sector, employment sector. Personality is the important aspect of the person’s traits and this project aims to predict it. The scope can be very large so our project focuses on the prediction through text, questionnaire and through social media so that it widens the methods of predicting it and outputs the percentage of the trait a person consists. This data obtained in numbers for the O.C.E.A.N. traits can be further used for multiple purposes.

**Project Proposal**

* 1. **Proposed Work**

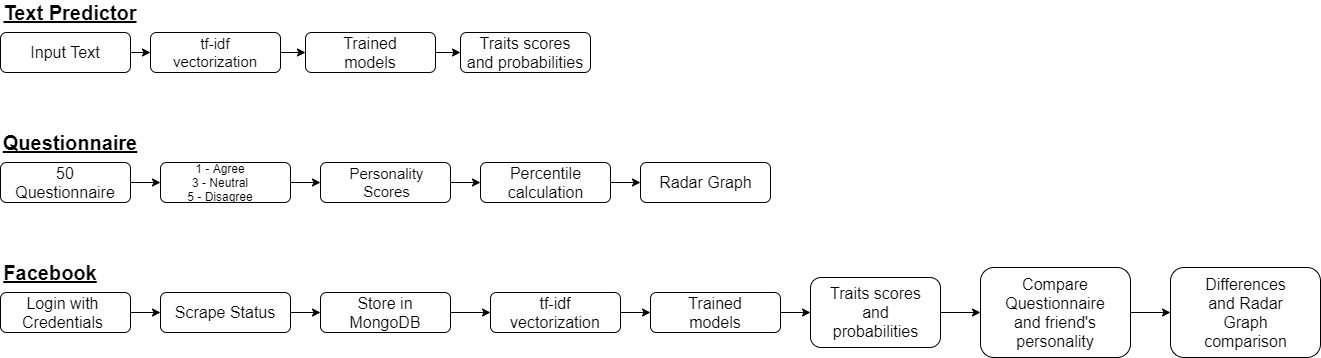
We propose a Personality Prediction web app which predicts personality based on the BIG 5 i.e. OCEAN model and rates the personality for his Openness, Conscientiousness, Extroversion, Agreeableness, Neuroticism of that person. This app consists of 3 sections namely Text Predictor, Questionnaire, Facebook.

In first section, user interface of the web app gives option to input a text which is provided as an input to the OCEAN model for predicting. This is followed by the scanning those text sentences in the myPersonality sample dataset (250 users, 9917 status updates).

In Questionnaire, User undergoes 50 questions regarding the O.C.E.A.N model, based on which radar graph is implemented. It can also be termed as Psychometric Assessment. The options given for the questions consists of Disagree, Neutral, Agree which correspond to 1, 3,5 respectively, based on which the personality score of a particular user is calculated. This can be further used for comparing the personality among social network.

For comparing the personality among the Facebook connections or any social media site network tab is helpful. Comparison between user’s personality score evaluated from the Questionnaire section and the model based prediction is used in this section. With the help of Facebook Web Scrapper the statuses of user’s Facebook network is extracted. For every connection personality is predicted and compared with the user’s personality based on the personality points.

* 1. **Proposed Methodology**



**Fig 3.2.1 Proposed Methodology**

1. Text Predictor:

The Personality of a person can be predicted by inputting any text with the help of this tab. The user interface of the web app gives option to predict the personality based on the text. The text is converted into vectors using tf-idf vectorization then, five models for the Big-Five personality traits are trained using RandomForest Algorithm. The vectors are given input to the models and the score as well as the classification is obtained.

1. Questionnaire:

In this section, user undergoes 50 questions regarding the O.C.E.A.N model, 10 questions for each personality traits. The score is evaluated based on the answers to the questions and the radar graph for user’s personality is implemented. It can also be termed as Psychometric Assessment.

1. Facebook:

For comparing the personality among the Facebook network or any other social media site network this tab is helpful. Comparison between user’s personality score evaluated from the Questionnaire section and the same models trained for text prediction are used here for evaluation of the scores of the Big-Five personality traits. The comparison is shown in the radar graph too on the user interface.

* 1. **Details of Hardware/Software Requirement**

**React** is used to design the user interface as it is a efficient and versatile JavaScript library. The Web App which can be created using React.js are going to be using backend as Flask and MongoDB.

**Python** is used as a programming language. Python is simple and its easy to find syntax emphasizes readability and thus reduces the value of program maintenance.

**MongoDB** is used to store Facebook statuses from web scraping and it is further used for prediction.

**Selenium** automated browser will scrape statuses from Facebook and add them to a database in MongoDB.

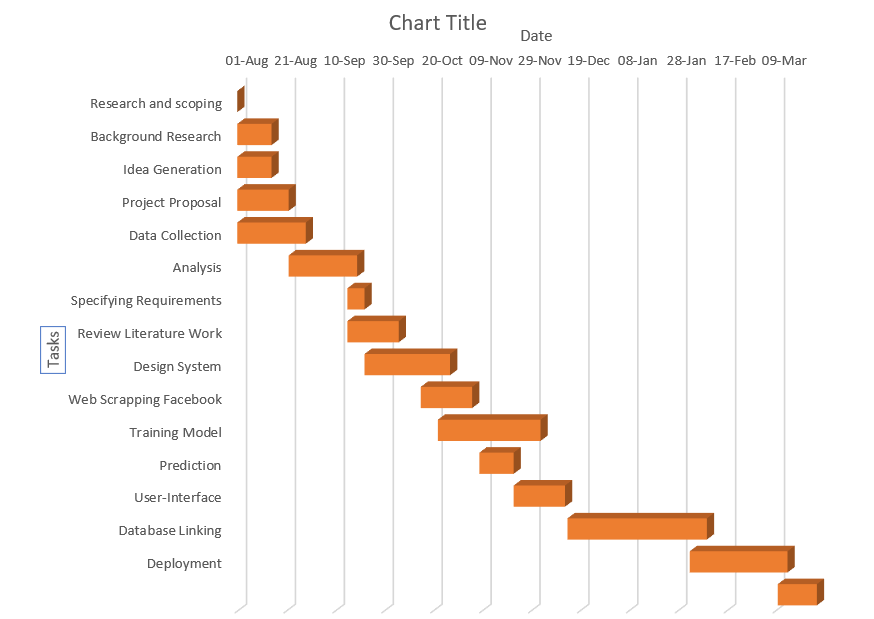
**Node.js** is used to run the web app.

**Random forest models** give a prediction of personality score, using the regression model, and a probability of the binary class, using the classification model, for every personality trait.

**Chapter 4**

**Planning And Formulation**

* 1. **Schedule forProject**



**Fig 4.1.1 Gantt Chart**

**Chapter 4**

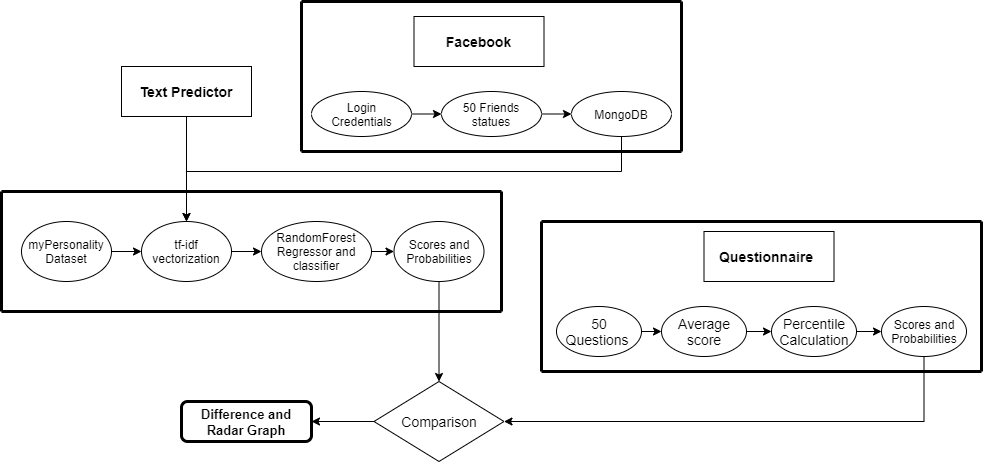
### Detailed Plan of Execution

The development of the concerned system starts right from the research and scoping based on the topic of personality prediction. This phase is followed by the background research needed for developing the system which did help in generating the basic idea of developing our system. The research based on the system which had to be generated with the help various published papers, their proposed work, etc. After proposed work, data for predicting the personality was collected followed by its analysis. After understanding the basic system design, the system requirements were found out. The whole system was designed based the requirements of the system. The implementation of the system was initialized with web scrapping the Facebook and extracting the network in MongoDB and giving it to the model trained with the help of Random Forest Algorithm. The output is predicted based on that model.

**Chapter 5**

# Design of System

### Design Diagram with Explanation



**Fig 5.1.1 Design of the System**

**Chapter 5**

The web app consists of 3 sections Text Predictor, Questionnaire and Facebook. The model is trained on myPersonality dataset. The Models produce a predicted personality score, using the regression model, and a probability of the binary class, using the classification model, for each personality trait. In the first section personality of a person can be predicted by inputting any text with the help of text predictor. This section is totally based on the models trained earlier. The user interface of the web app gives option to predict based on the text. The models trained will give a score on each of the 5 personality traits.

In the Questionnaire section user undergoes 50 questions Big 5 personality test regarding the O.C.E.A.N model, User will get options like agree, disagree and neutral on basis of that certain score will be generated, which then displays your corresponding percentile scores for each personality trait and based on which radar graph is implemented. It is like a psychometric assessment. The score generated in this section is used to compare personality obtained in the Facebook section.

Facebook is used for comparing the personality among the Facebook or any social media site network. Here users credentials are provided in a config file to access Facebook, through automation the statuses of user’s friends are extracted and stored in database. The model which is trained earlier is used to give a prediction on the data which is stored in database. This data is also used for comparison between user’s personality score evaluated from the Questionnaire section and the model-based prediction.

**Expected Results**

### Implementation Details

1. **Text Predictor:**

First, the text inputted by the user is converted into vectors using tf-idf vectorization. Second, the vectors of the text are provided as inputs to the five models trained on myPersonality dataset by RandomForest algorithm. Third, the scores and probabilities are evaluated from the models and are displayed on the screen according to the respective personality traits.

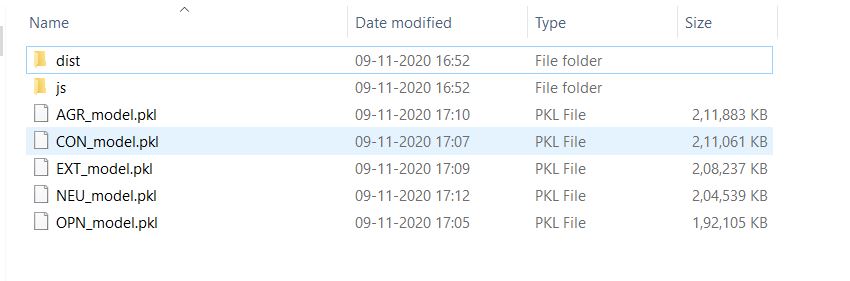
1. **Questionnaire:**

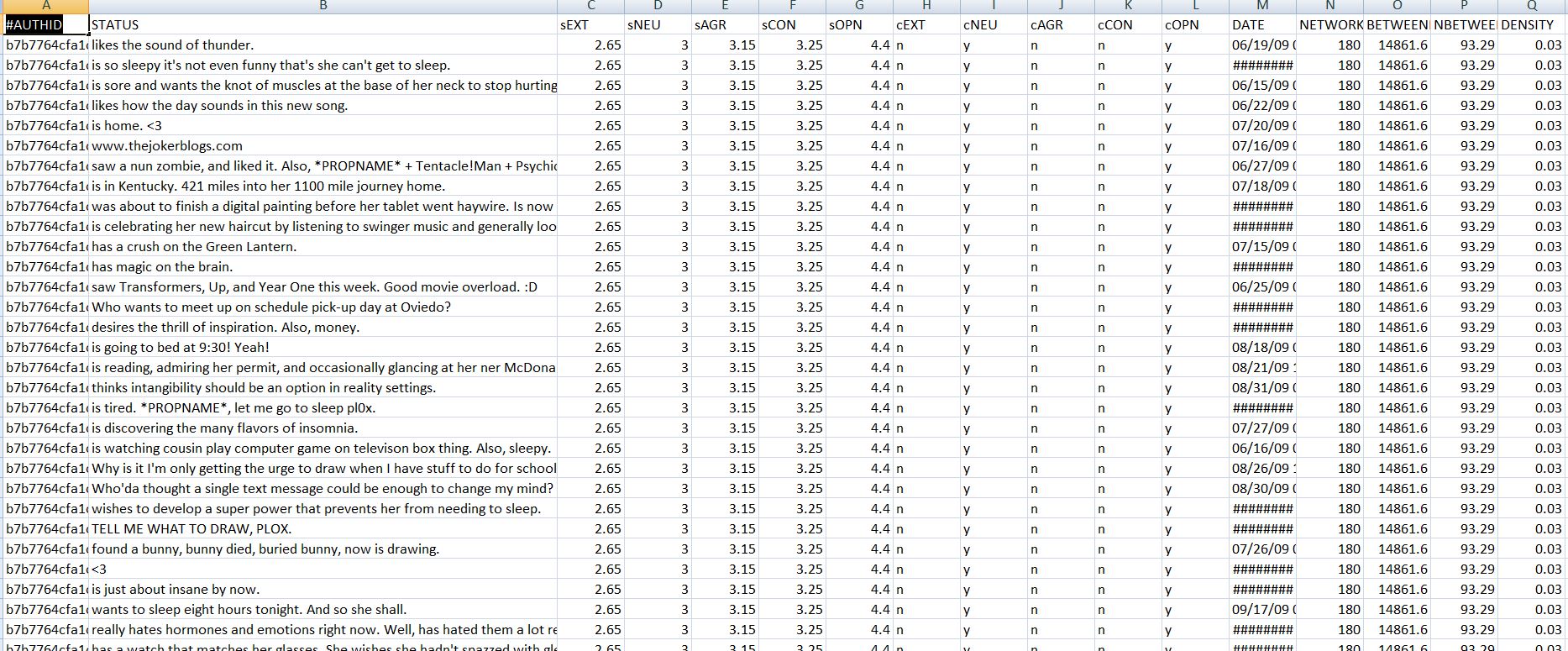
First, the questions are prepared according to the Big-Five personality traits, 10 questions for each personality trait. Thus, the user has to answer 50 questions for the evaluation of the personality. Second, the answer the questions are stored in the dictionary and the score is averaged according to the answer. Third, the percentile is also evaluated based on the user’s answer and the dataset of values for each personality trait. Forth, the score and the percentile is displayed on the user interface.

1. **Facebook:**

First, the user provides credentials for the facebook account. Second, selenium is used for automation for extracting the facebook statues of 50 friends from the friend list provided in the Friends section on the Facebook account. Third, the extracted statues are stored in MongoDB. Forth, the statues are used to evaluate the personality of each friend according to the RandomForest model. Fifth, the calculated personality is used to compare with the user’s personality in questionnaire section. Sixth, the visual comparison can be done user by clicking on the compare button on the user interface. The radar graph and the difference between the scores is clearly visible on the user interface.

Fig 6.1.1 myPersonality Dataset

****

****Fig 6.1.2 Generation of Pickle files

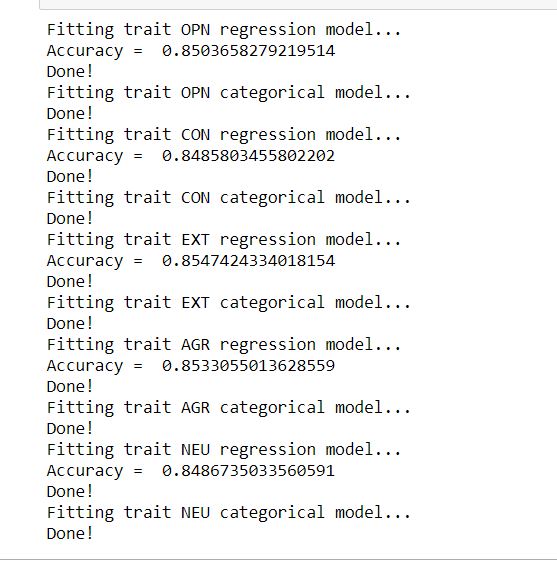
****

Fig 6.1.3 Training

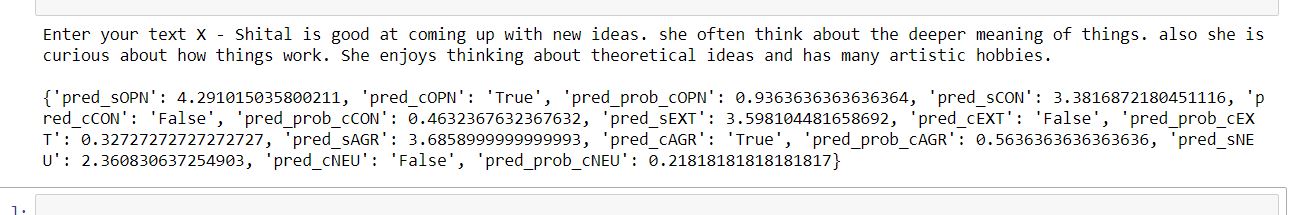
****

Fig 6.1.4.1 Text Predictor Example 1

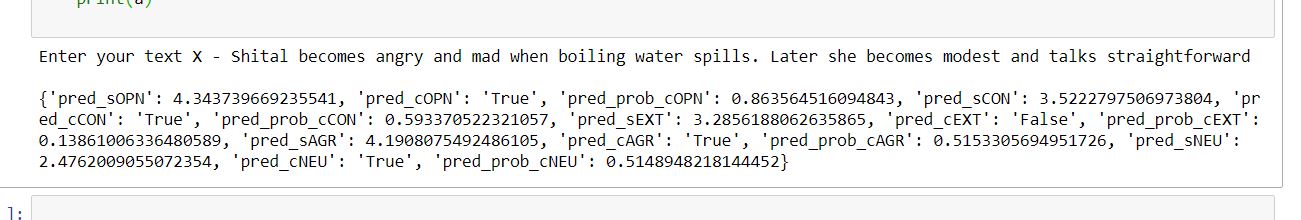
****

Fig 6.1.4.2 Text Predictor Example 2

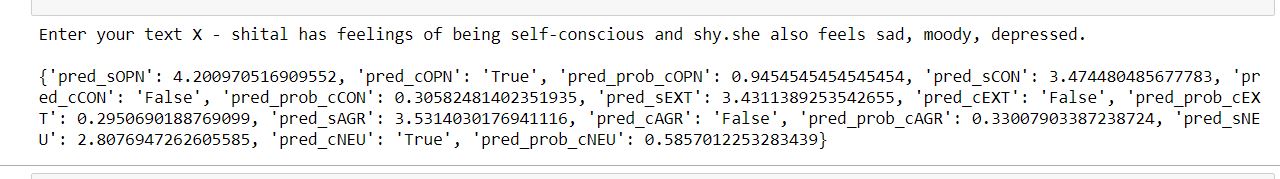
****

Fig 6.1.4.3 Text Predictor Example 3

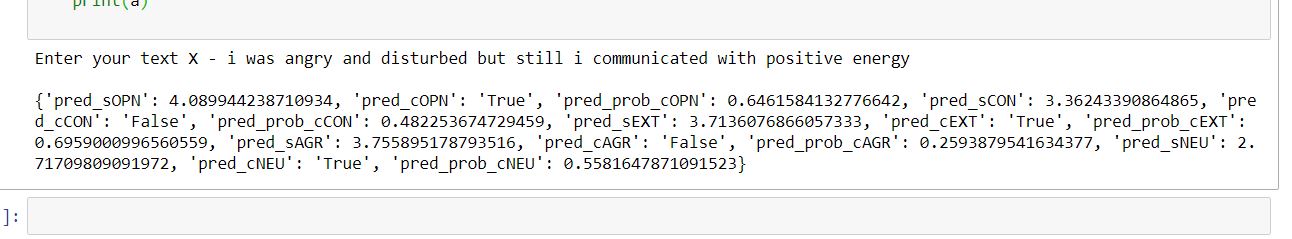
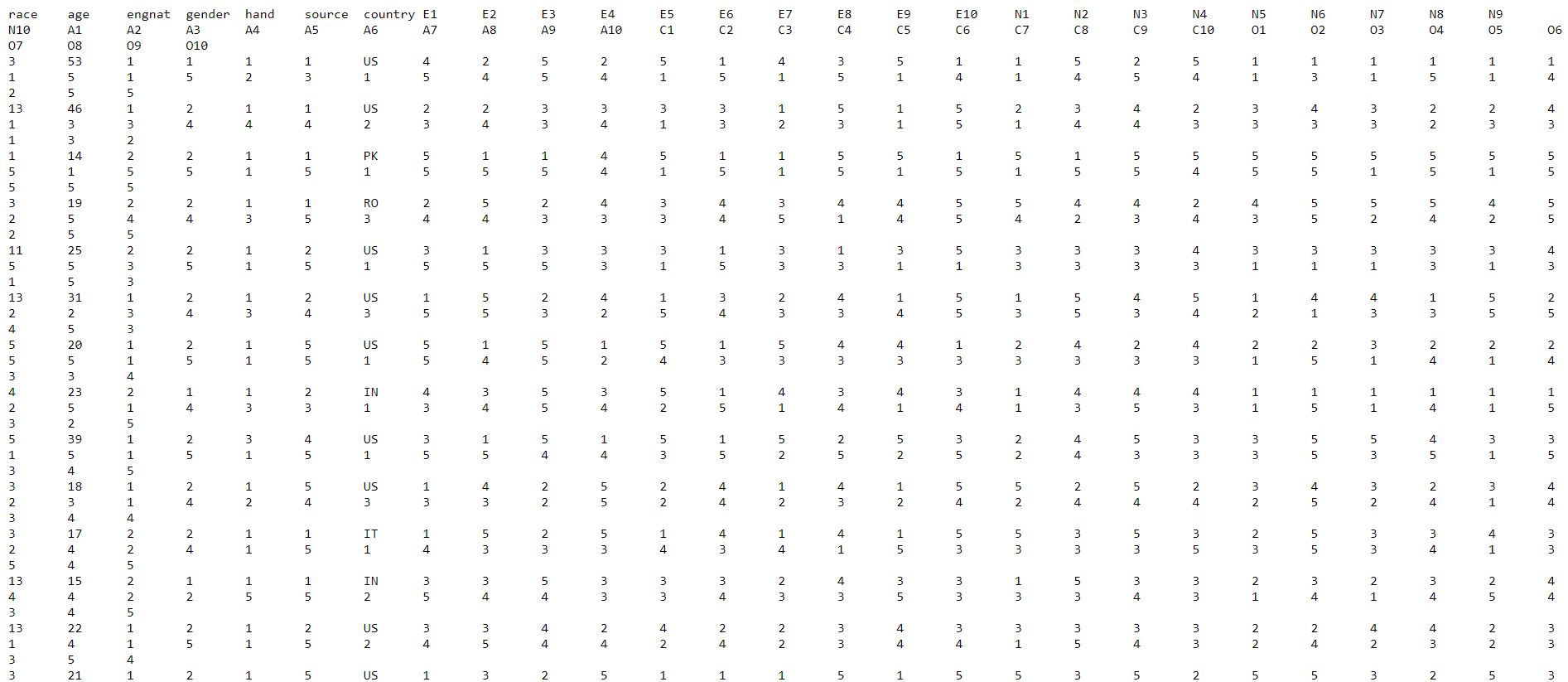
****

Fig 6.1.4.4 Text Predictor Example 4

****

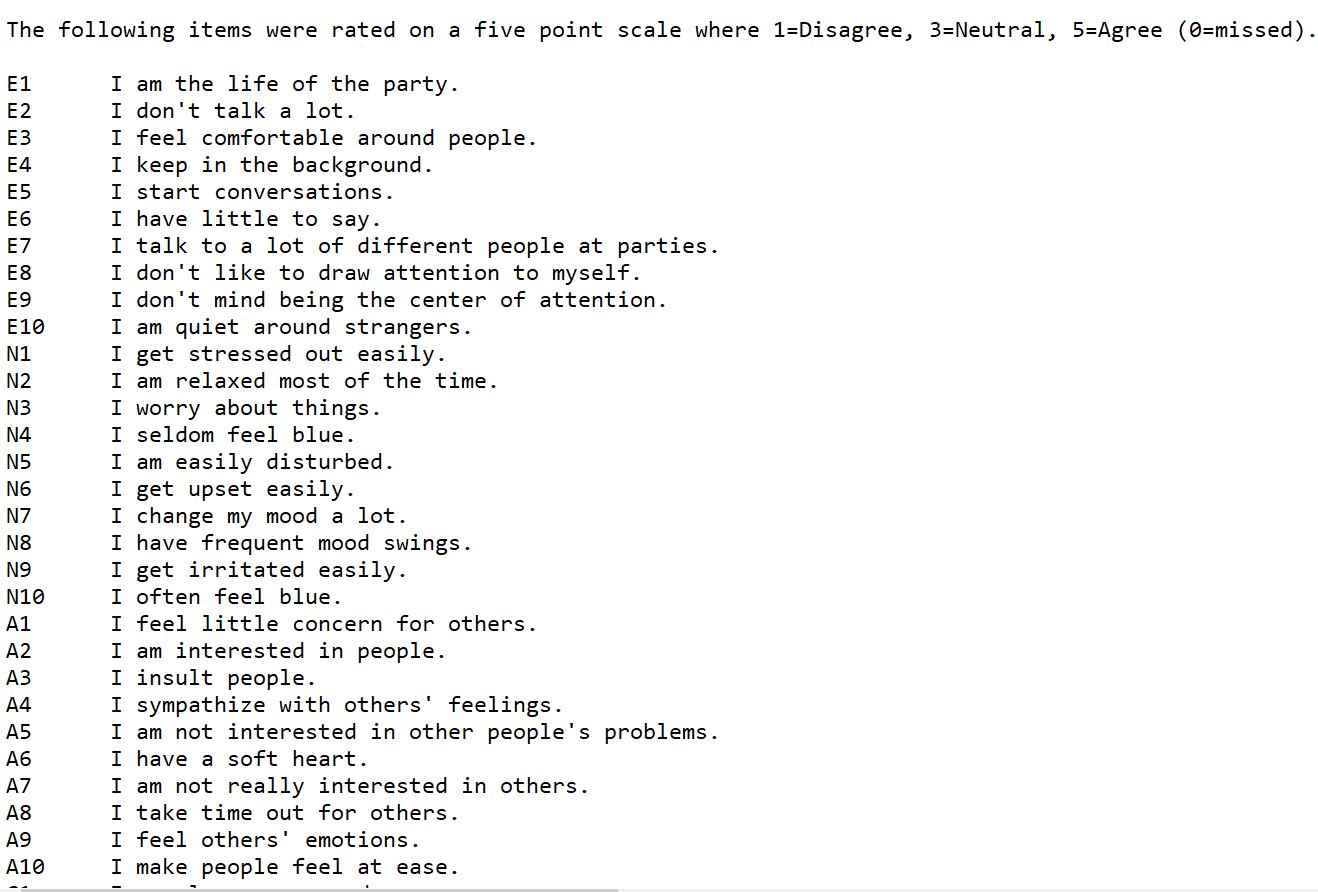
****

Fig 6.1.5 Dataset for Questionnaire

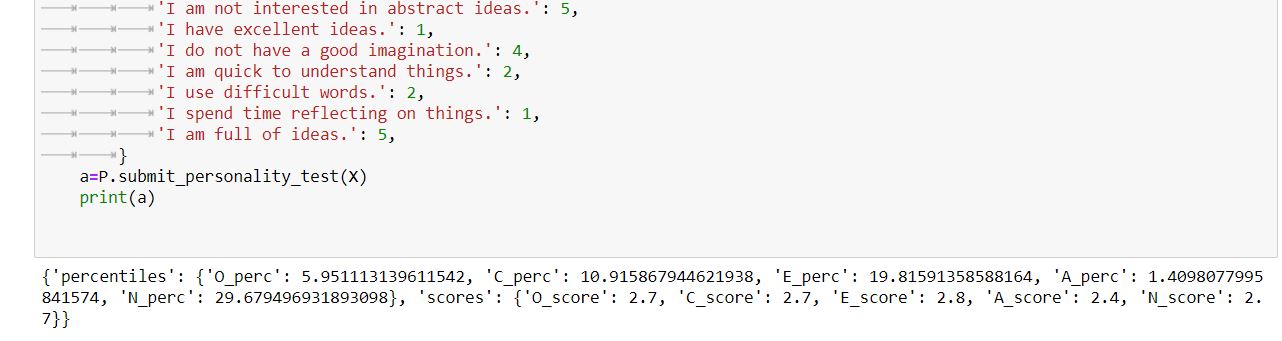
****

Fig 6.1.6.1 Questionnaire Example 1

****

Fig 6.1.6.2 Questionnaire Example 2

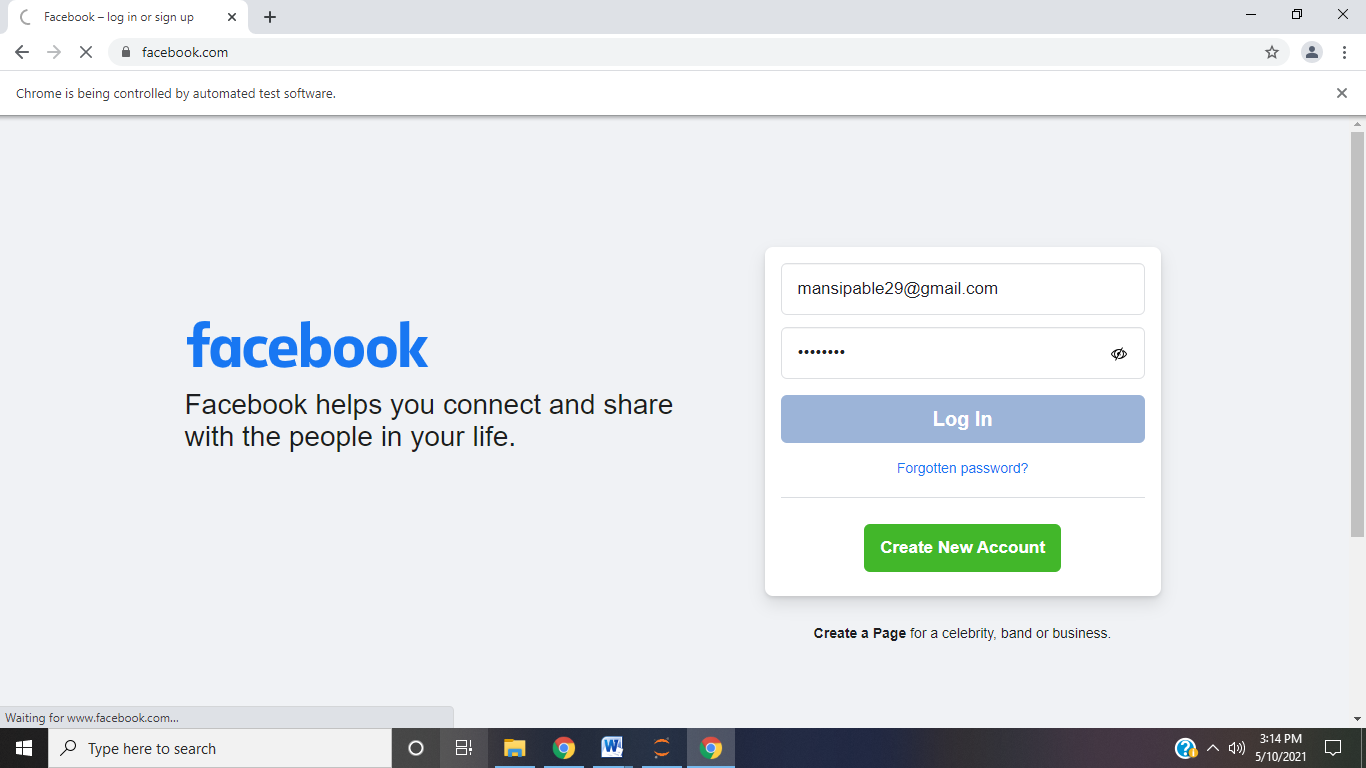


Fig 6.1.7.1 Facebook Login

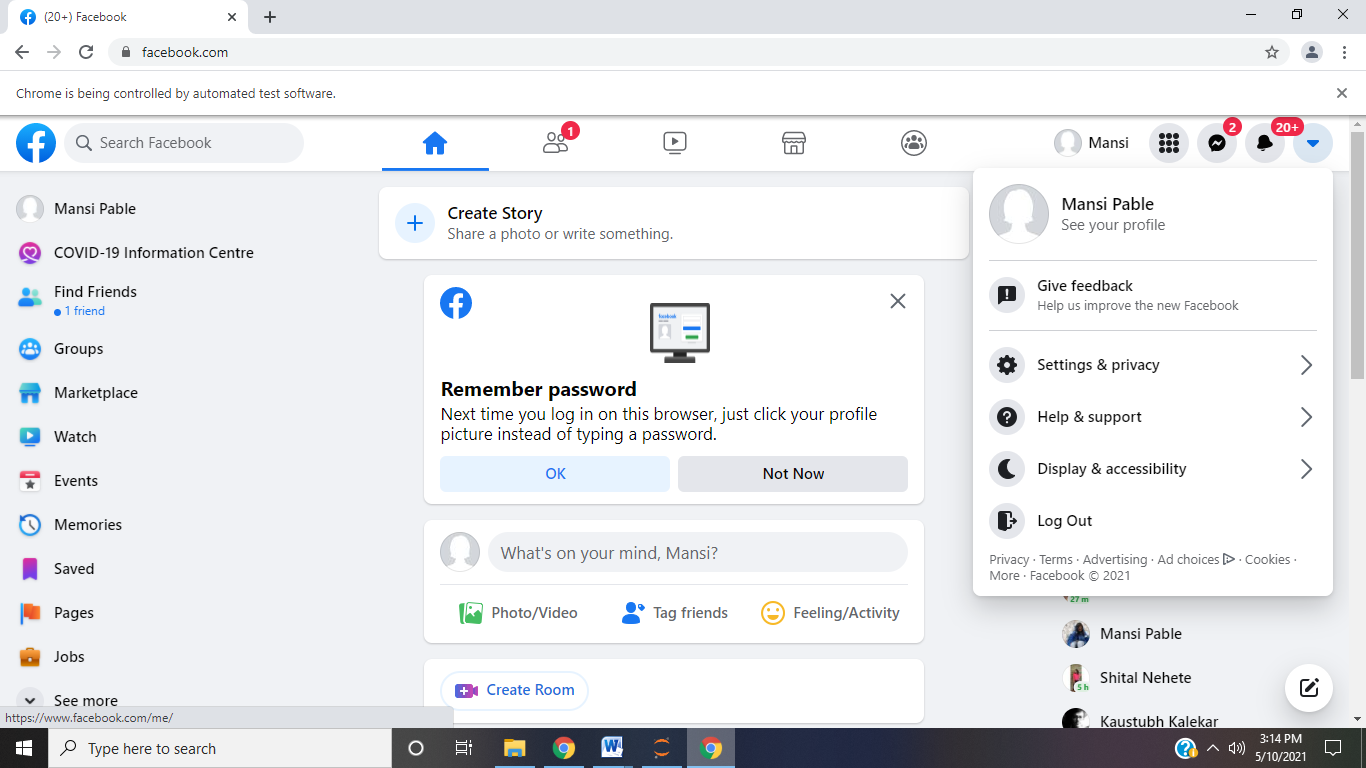


Fig 6.1.7.2 Navigating to See your profile

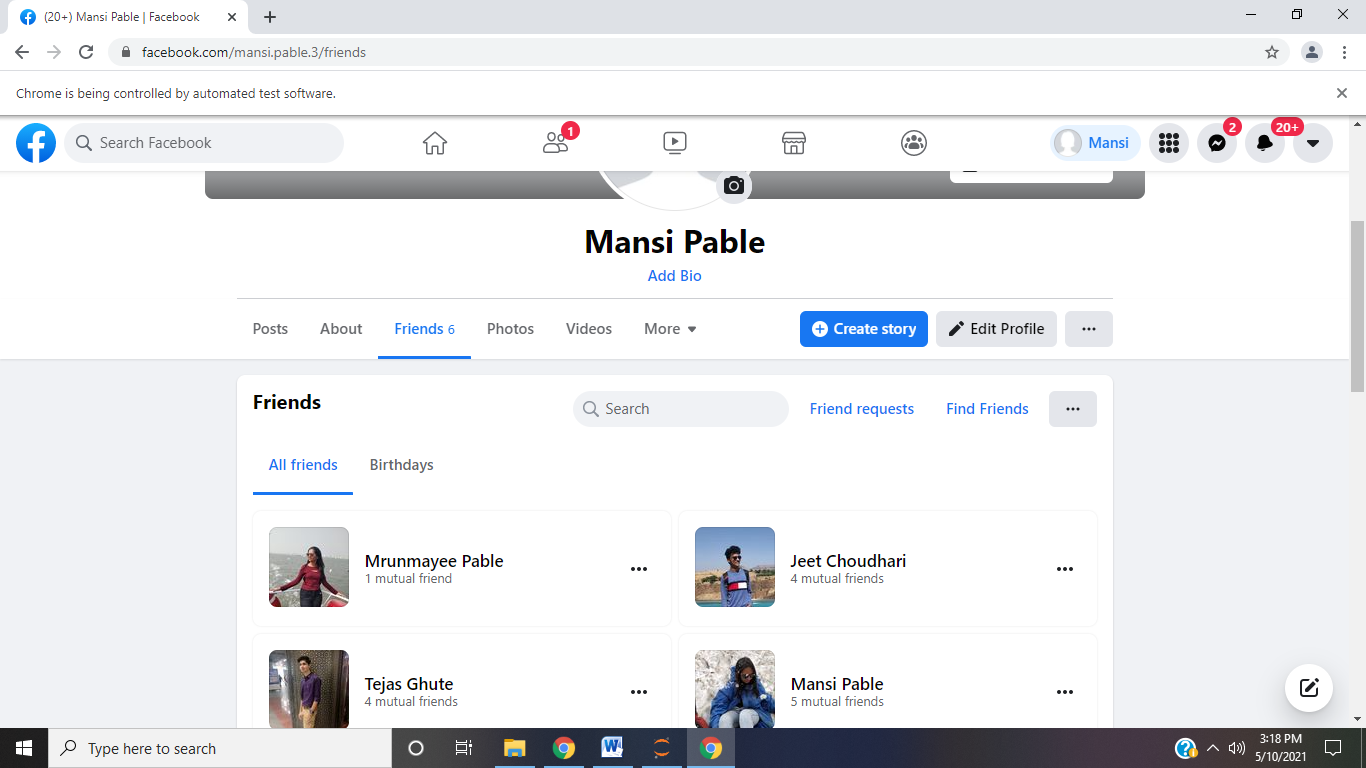


Fig 6.1.7.3 Navigating to Friends tab

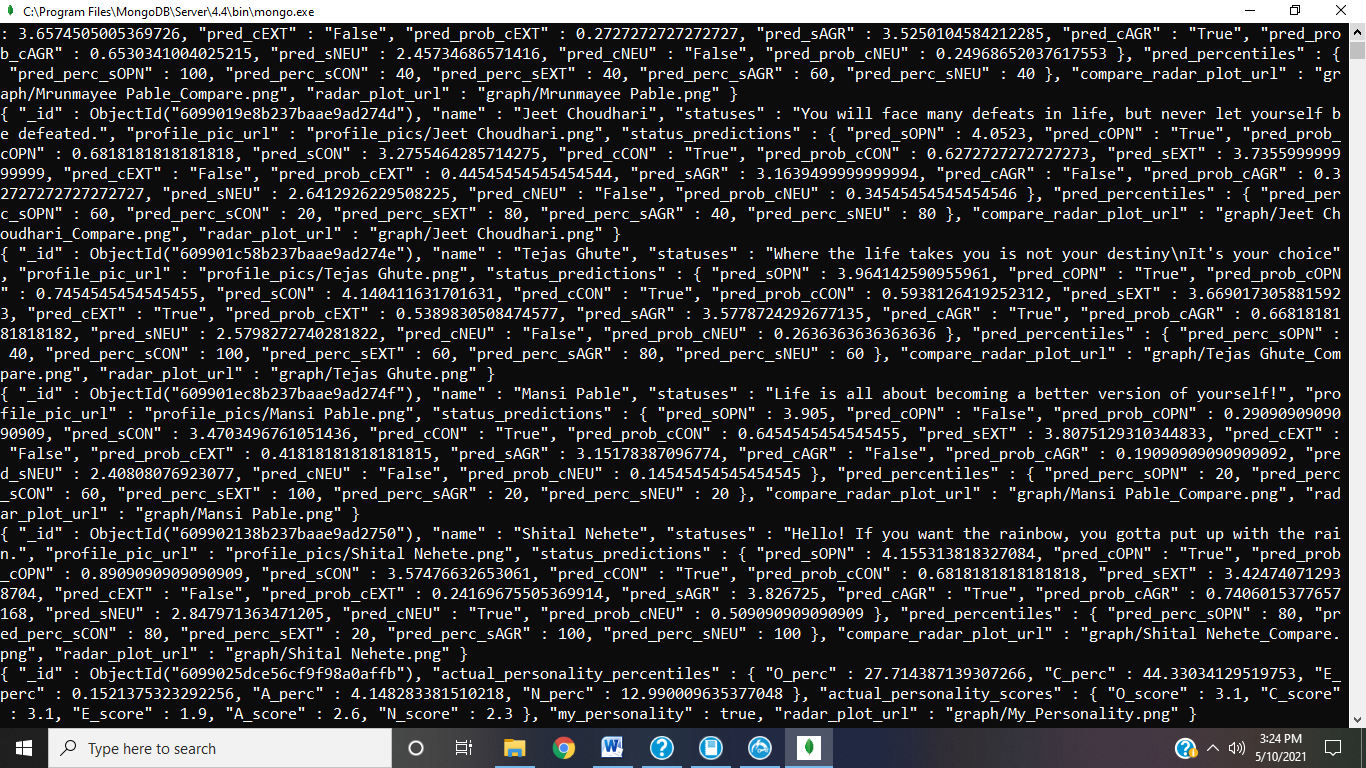


Fig 6.1.7.4 Extracting Friends

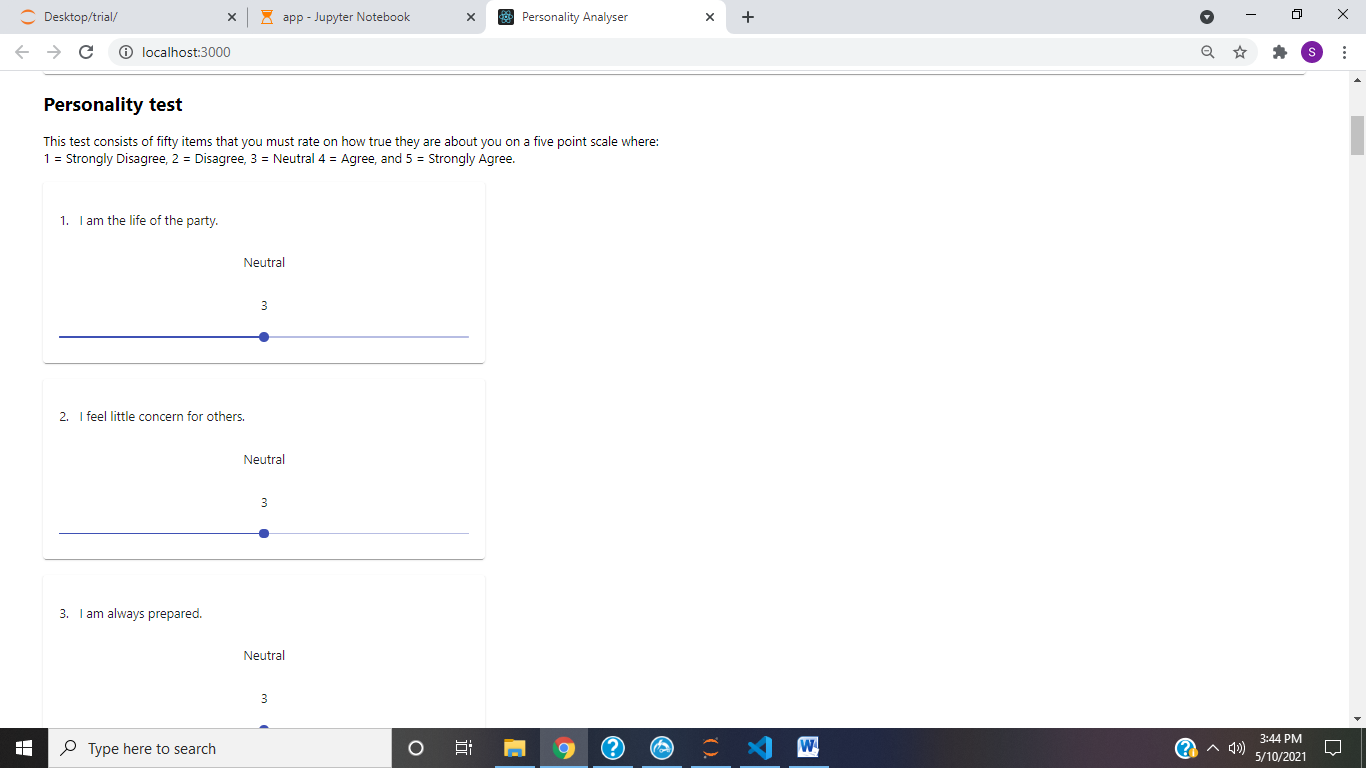


Fig 6.1.8.1 Questionnaire

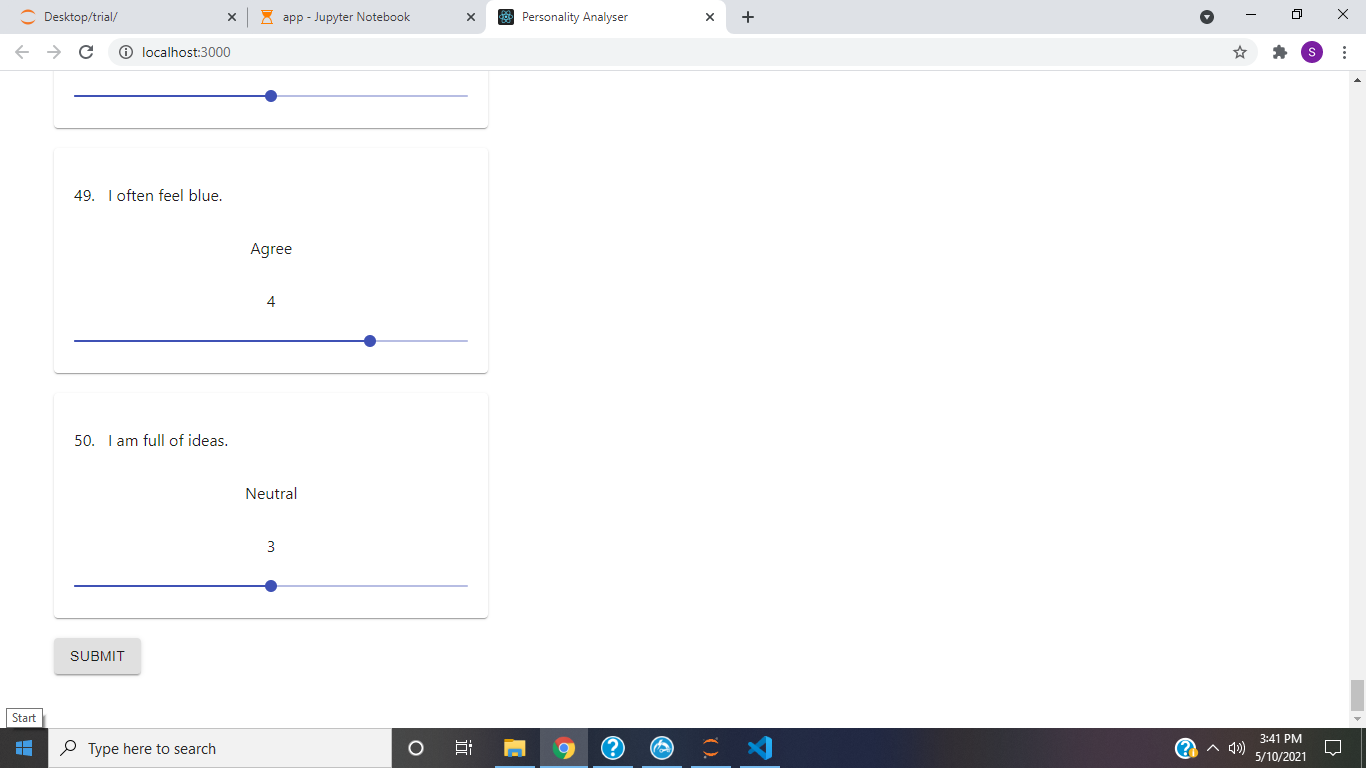


Fig 6.1.8.2 Submitting questionnaire

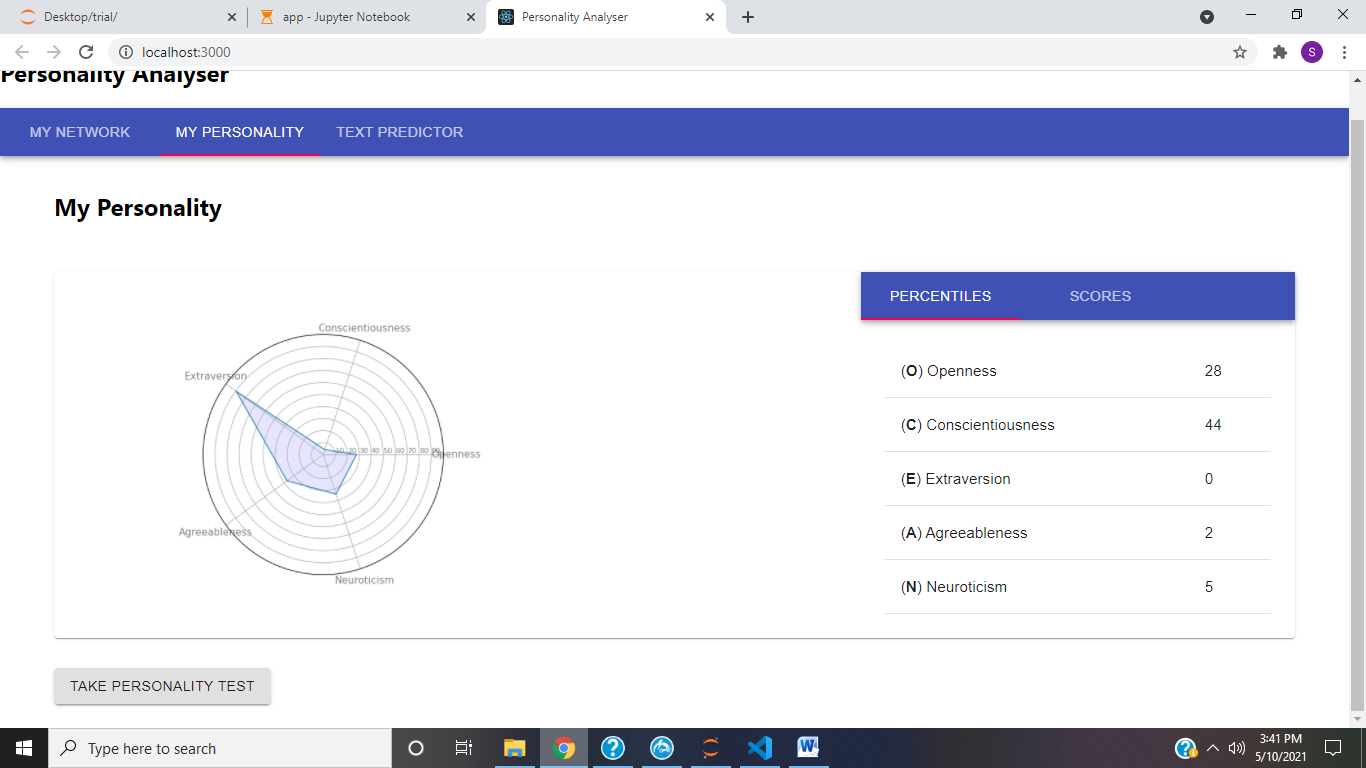


Fig 6.1.8.3 Questionnaire Results

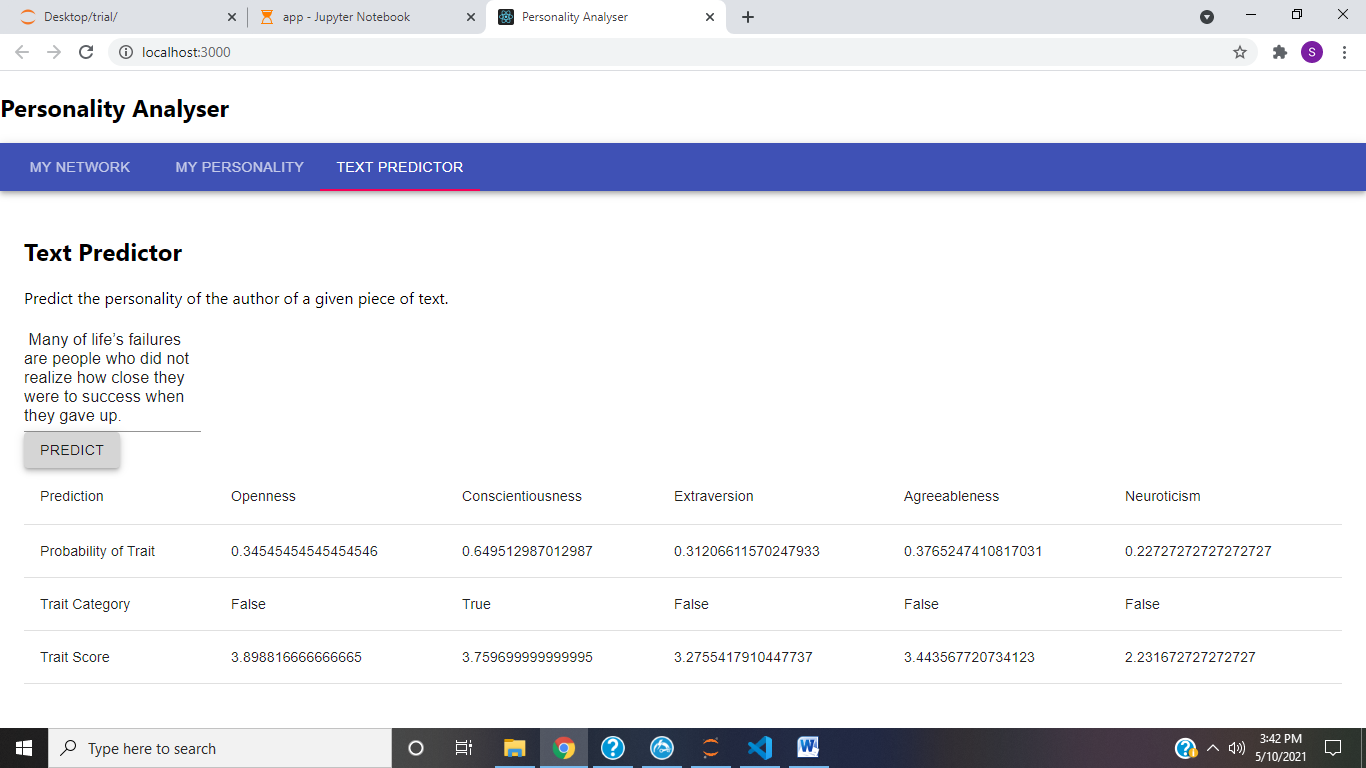


Fig 6.1.8.4 Text Predictor

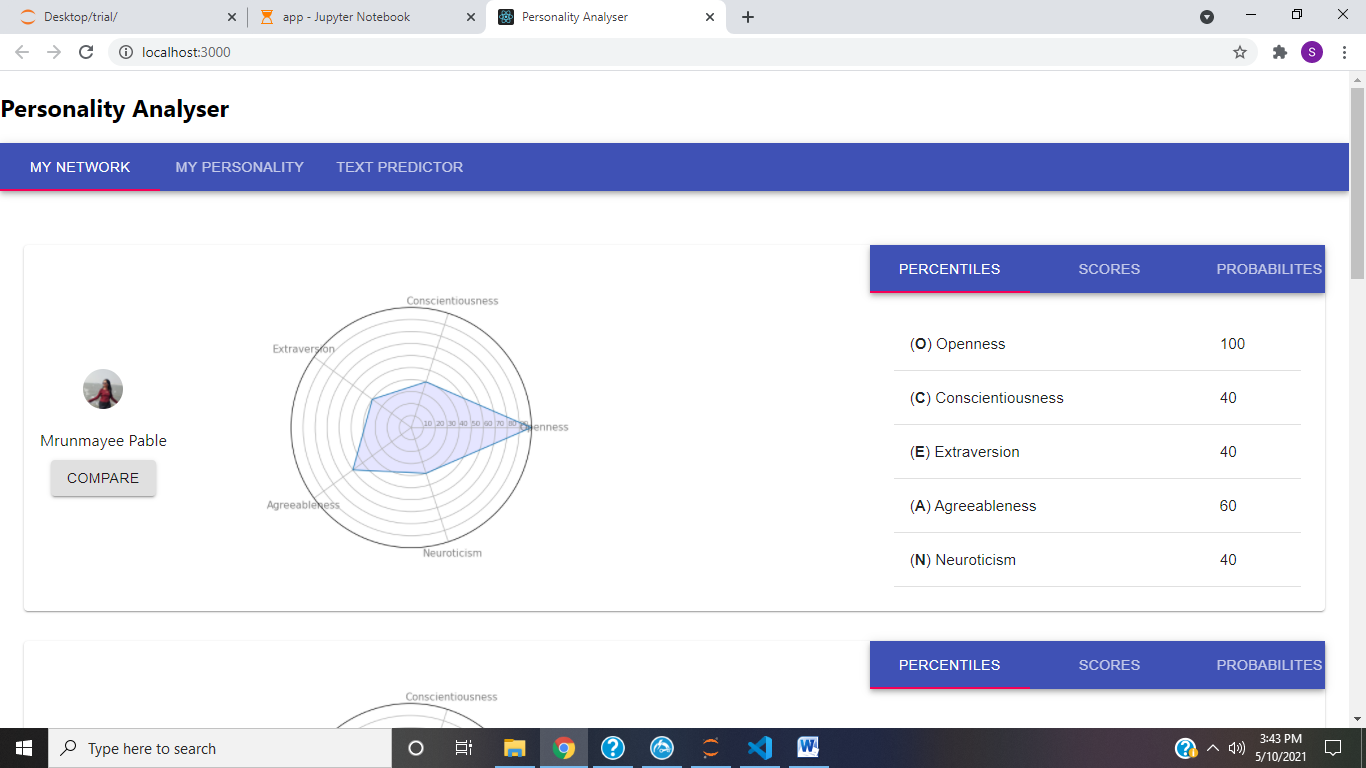


Fig 6.1.8.5 My Network tab

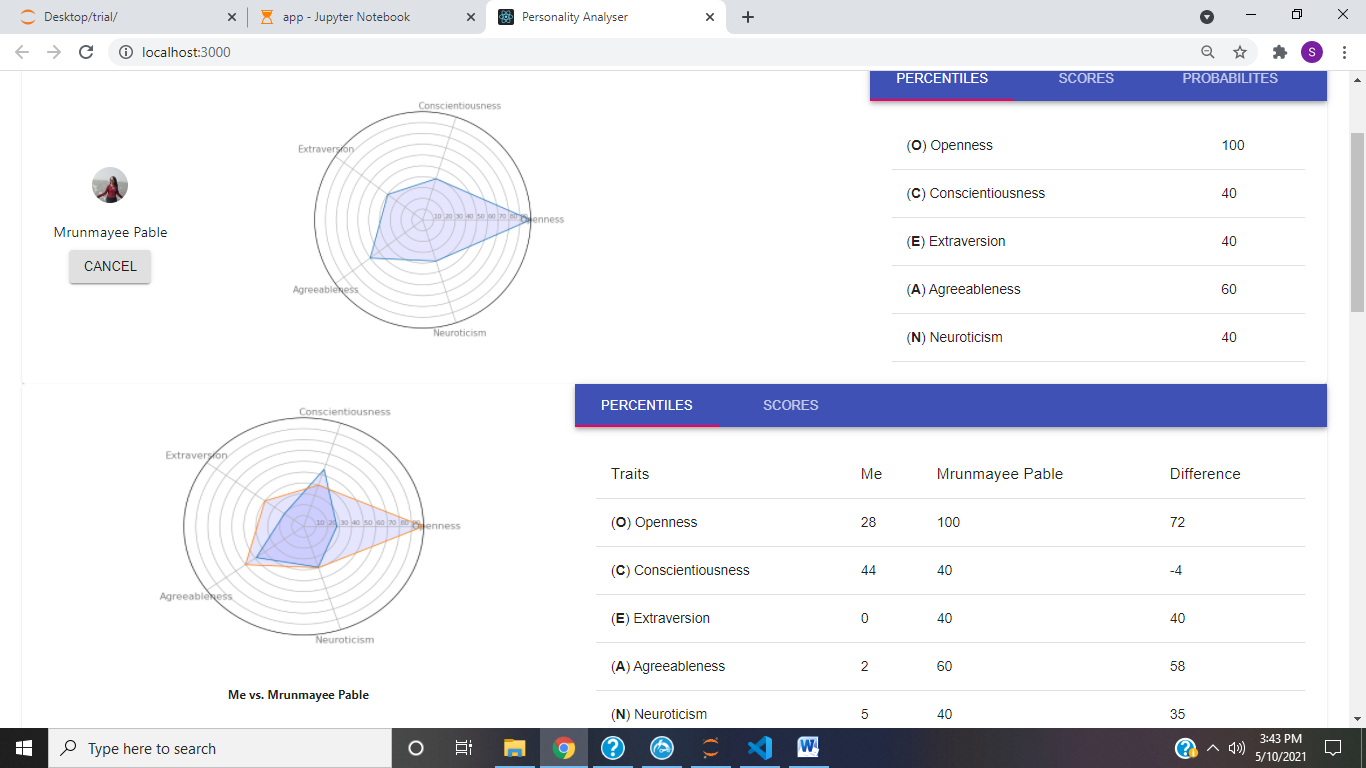


Fig 6.1.8.6 Comparison with friends

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Traits | | | | |
| Algorithm | OPN | CON | EXT | AGR | NEU |
| Naïve Bayes | 70.00% | 59.20% | 68.80% | 56.40% | 54.40% |
| SVM | 70.40% | 56.00% | 61.60% | 56.80% | 60.40% |
| Logistic Regression | 70.40% | 54.40% | 68.40% | 53.60% | 60.40% |
| Gradient Boosting | 63.20% | 56.40% | 68.00% | 63.20% | 59.20% |
| LDA | 70.00% | 58.40% | 68.00% | 58.00% | 60.80% |

### Project Outcomes

A web app is created to provide interface for the navigation and display of the personality scores and percentile. There are 3 tabs namely, text predictor, questionnaire and social network. The personality of the person is calculated through a machine learning model i.e. Random Forest algorithm. It predict the personality on the text provided in the textbox on the web app. The scores and percentiles are displayed in the corresponding tabs and a radar graph is formed. Social Network tab helps to compare the personality score of the person with the friends on the network and the visual difference is also showed on the radar graph. The compare options shows the personality scores and percentile of the friend and the comparison can be made. The web app shows in details scores of the Big-Five personality traits. There are 5 different models for predicting 5 different traits. The traits are namely, Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism. The scores and percentile is displayed for these traits and comparison can also be made with other people who they are friends with.

**Conclusion and Future Work**

Personality is a very important aspect nowadays. In this growing world, the person is identified by the personality he/she carries so, this project aims to decoding the personality in 5 essential traits. The score obtained from the prediction are the percentage of the traits present in the person and it can be used further for many purposes. One’s personality prediction can show the job satisfaction and performance in the corporation also. It can be used in the health care sector too for psychological analysis. The recommender systems also rely on the personality of the person and help as per his interests. There are many more applications such has marketing and advertisement. Hence, this project has growing demands and many aspects together.

Future Work is necessary for more accurate results as the dataset is very small. The limitations of the dataset affect the accuracy of the model. For questionnaire, the number of questions are limited and hence, the personality is accurately predicted. For social networks sites, there are large number of fake accounts which affects the accuracy and prediction and the access to the accounts requires the credentials that may be a concern. For text, the text the person enters might be influenced by the fact that it I getting evaluated for personality so, the chances of the person lying or not writing what he/she feels are very high which will result is improper prediction.

[1] Majid Ramezani, Mohammad Reza Feizi Derakhshi, Mohammad-Ali Balafar, Meysam Asgari-Chenaghlu, “Automatic Personality Prediction; an Enhanced Method Using Ensemble Modeling”, 2020

[2] Madhura Jayaratne, Buddhi Jayatilleke , “Predicting Personality Using Answers to open-ended interview questions”, 2020

[3] Aditi V.Kunte, Suja Panicker , “Using textual data for Personality Prediction: A Machine Learning Approach”, 2019

[4] Abdur Rahman, Asif Al Faisal, Tayeba Khanam, Mahfida Amjad, Saeed Siddik, “Personality Detection from Text using Convolutional Neural Network,” IEEE, 2019.

[5] Guanqun Sun, Ao Guo, Jianhua Ma and Jianguo Wei, “Personal Trait Analysis Using Word2vec Based on User-generated Text,” IEEE, 2019.

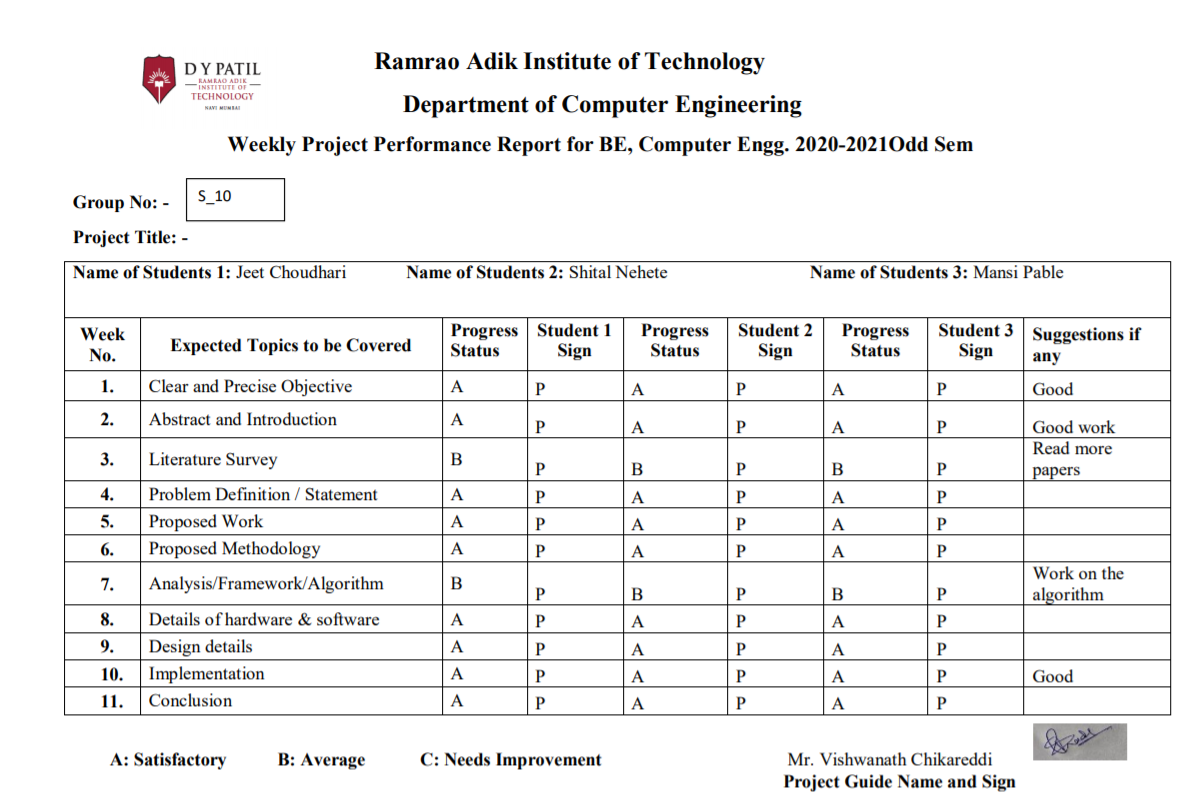
[6] Raad Bin Tareaf, Seyed Ali Alhosseini, Philipp Berger, Patrick Hennig, Christoph Meinel, “Towards Automatic Personality Prediction Using Facebook Likes Metadata,” IEEE, 2019.

[7] Mariam Hassanein, Wedad Hussein, Sherine Rady Tarek, F. Gharib, “Predicting Personality Traits from Social Media using Text Semantics,” IEEE, 2018.

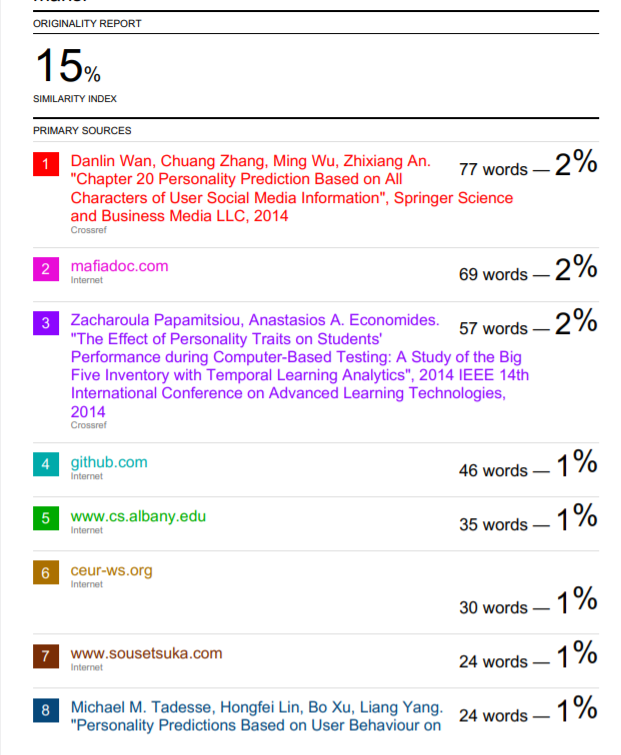
[8] Raad Bin Tareaf, Philipp Berger, Patrick Hennig, Christoph Meinel, “Personality Exploration System for Online Social Networks: Facebook Brands As a Use Case”, 2018

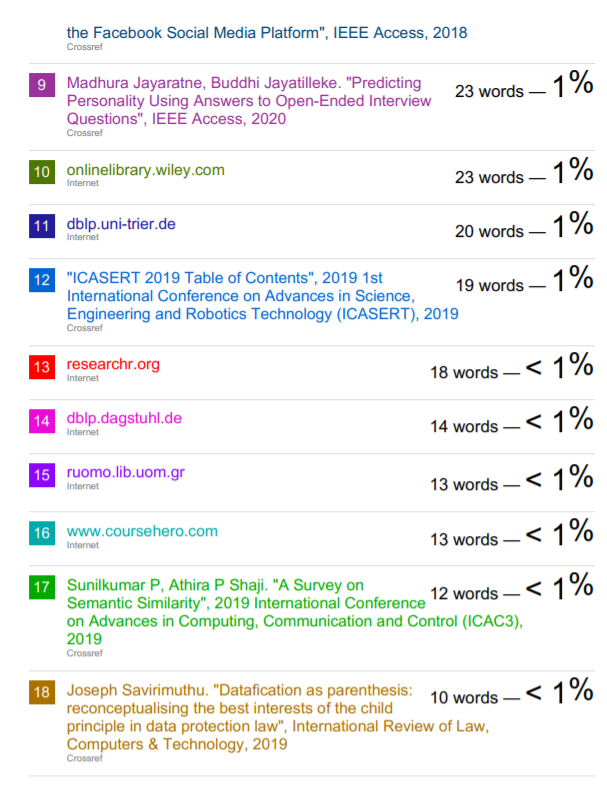
# Appendices

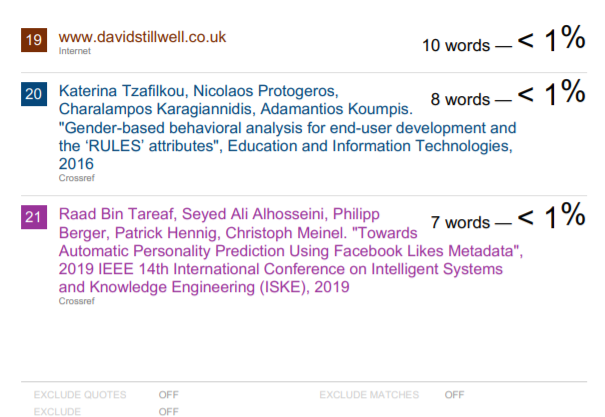
**Weekly Progress Report Project A**



**Plagiarism Report**







# Acknowledgement

We take this opportunity to express my profound gratitude and deep regards to our guide **Mrs. Smita Bharne** for his/her exemplary guidance, monitoring and constant encouragement throughout the completion of this report. We truly grateful to his/her efforts to improveourunderstandingtowardsvariousconceptsandtechnicalskillsrequiredinourproject. The blessing, help and guidance given by him/her time to time shall carry us a long way in the journey of life on which we are about toembark.

We take this privilege to express our sincere thanks to Dr. Mukesh D. Patil, Principal, RAIT for providing the much necessary facilities. We are also thankful to Dr. Leena Ragha, Head of Department of Computer Engineering, Project Co-ordinator Mrs. Smita Bharne and Project Co- coordinator Mrs. Bhavana Alte, Department of Computer Engineering, RAIT, Nerul Navi Mumbai for their generous support.

Last but not the least we would also like to thank all those who have directly or indirectly helped us in completion of this thesis.

**Ms. Shital Nehete**

**Ms. MansiPable**

**Mr. Jeet Choudhari**