A PRELIMENERY REPORT ON

COMPARATIVE ANALYSIS OF FEATURE SELECTION ALOGORITHMS FOR COMPUTATIONALPERSONALITY PREDICTION FROM SOCIAL MEDIA

SUBMITTED TO THE SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE
IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE OF

BACHELOR OF ENGINEERING (COMPUTER ENGINEERING)

SUBMITTED BY

GOURI GHADMODE	**Exam No: B150434267**
VIDYA CHVAN	**Exam No: B150434242**
SHREYASH SUL	**Exam No: B150434404**
AYUSH DHOKRIYA	**Exam No: B150434253**

Sinhgad Institutes

DEPARTMENT OF COMPUTER ENGINEERING

STES'S SINHGAD ACADEMY OF ENGINEERING

KONDHWA (BK.), PUNE 411048

SAVITRIBAI PHULE PUNE
UNIVERSITY

2018 -2019

Sinhgad Institutes

CERTIFICATE

This is to certify that the project report entitles

"COMPARATIVE ANALYSIS OF FEATURE SELECTION ALOGORITHMS FOR COMPUTATIONALPERSONALITY PREDICTION FROM SOCIAL MEDIA "

Submitted by

GOURI GHADMODE	**Exam No: B150434267**
VIDYA CHVAN	**Exam No: B150434242**
SHREYASH SUL	**Exam No: B150434404**
AYUSH DHOKRIYA	**Exam No: B150434253**

is a bonafide student of this institute and the work has been carried out by him/her under the supervision of **Prof. S.N.BAHIR** and it is approved for the partial fulfillment of the requirement of Savitribai Phule Pune University, for the award of the degree of **Bachelor of Engineering** (Computer Engineering).

(Prof. S.A. BAHIR)	**(Dr. S. N.SHELKE)**
Guide	Head,
Department of Computer Engineering	Department of Computer Engineering

(Dr. K. P. Patil)
Principal,
Sinhgad Academy of Engineering Pune – 48

Place : Pune
Date

ACKNOWLEDGMENT

Abstract

Personality is one feature that determines how people interact with the outside world. Personality can be defined as a necessary element of a person's behavior. The way people interact with other people determines their personality. This paper covers the topic of Automated Personality Classification – a system that analyses the personality of a user based on certain features using Data Mining Algorithms. In this paper, a system is proposed which analyses the personality of an applicant. This system will be helpful for organizations as well as other agencies who would be recruiting applicants based on their personality rather than their technical knowledge. The personality prediction results are based on Big Five Personality traits and the classification is done using Naive Bayes Algorithm and Support Vector Machine.

INDEX

List of Figures

CHAPTER 1

INTRODUCTION

OVERVIEW

Social media platforms such as Facebook, Twitter,Google, and Instagram has gained popularity due to ease access throughout the world and user-friendly interfaces to start communicating with others within a short period of time. Each user in these social networking sites (SNSs) is considered as an entity, and each entity is connected with other entities as friends, connections, or followers. While using these SNS's, users are facilitated by many activities, such as posting statuses/tweets, sharing others' posts/retweets, liking others' posts, commenting on others' posts, chatting directly with the friends, and playing online games with the friends. It is evident that from the activities performed by users, online behavior could be depicted . Understanding users' behavior may help to identify personality traits. Predicting users' personalities from digital footprints of social media is a challenging task as the context of identifying personality traits in social media is not trivial. positive or negative reviews/opinions about a political party. These types of statuses may have contextual trends, as other friends of the users may also be involved in posting similar statuses. Considering the trend, user may post his/her political views. Users behave differently in social media and real life.

Motivation

The analysis of the data revealed that almost all respondents used social media. Based on factor analysis results, their motivations for doing so are entertainment, information seeking, personal utility and convenience.

Objective

- The objective of this project is to analyse and classify personalities of a given set of people or an individual using advanced data mining concepts.

CHAPTER 2

LITERATURE SURVEY

2.1 STUDY OF RESEARCH PAPER

1.Paper Name: Social Media Text - A Source for Personality Prediction
Author: P.S.Dandannavar

Abstract ::-Social media usage has been on an ever increasing exponential rise. Usage of social media sites, such as Twitter and Facebook, for social interaction has also become a popular trend. It is estimated that on an average, around 6,000 tweets are tweeted on Twitter every second. With people spending on an average 35 minutes on Facebook each day, it is also estimated that there are about 317,000 status updates on Facebook per minute. These vast volumes of data have powerful information locked within them. This data can be analyzed and several purposes. The use of such social media data for predicting user personality is common. Prediction models have been successfully built that can predict several user attributes - age, gender, personality traits, occupation, political orientation etc. Standards in personality models such as the Big Five model, DISC and the Myers-Briggs Type Indicator have been the basis for all such personality prediction. A user's social media data can thus be used to predict his/her personality. The main objective of this work is to review the work carried out for personality prediction using social media data

2. Paper Name: :- Speaker Trait Prediction for Automatic Personality Perception using Frequency Domain Linear Prediction features

Author: S. Jothilakshmi and R. Brindha

Abstract : The aim of automatic personality perception is to predict the personality of the speaker perceived by the listener from nonverbal behavior. Extroversion, Conscientiousness, Agreeableness, Neuroticism and Openness are the speaker traits used for personality evaluation. In this work, a speaker trait prediction approach for automatic personality assessment is proposed to model the relationship between speech signal and personality traits using frequency domain linear prediction (FDLP) technique. Among several feature extraction techniques, FDLP features render increased performance. SSPNet Speaker Personality Corpus is used for experiments and evaluation. The proposed method predicts the speaker traits with 90-99classification accuracy.

3.Paper Name: Personality Prediction of Social Network Users

Author name: Chaowei Li, Jiale Wan, Bo Wang†

abstract : Through weibo users, we extract social data and questionnaire, and focus on how to use the user text information to predict their personality characteristics. We use the correlation analysis and principal component analysis to select the user information, and then use the multiple regression model, the gray prediction model and the multitasking model to predict and analyze the results. It is found that MAE values of the gray prediction are better than the multiple regression model Multitask model, the overall effect of the prediction between 0.8 and 0.9, the overall accuracy of good prediction. This shows that gray prediction in the user's personality prediction shows a good generalization and non-linear ability.

4. Paper Name: Personality Traits Prediction Based on Users' Digital Footprints in Social Networks via Attention RNN

Author:-Shipeng Wang, Lizhen Cui†, Lei Liu†, Xudong Lu†, Qingzhong Li

abstract : With the increasing popularity of social networks, massive digital footprints of individuals in online service platforms are generated. As a result, an emerging technology namely personality trait analysis has drawn much attention. The prediction and analysis of personality trait is an efficient way to voting prediction, review analysis, decision analysis and marketing. The existing studies generally employ classification models while ignore the temporal property of digital footprints, which may lead to unsatisfactory results. To make an improvement, this paper proposes an effective method to predict the personality traits by taking the temporal factors into account through the use of Attention Recurrent Neural Network (At-tRNN). The experimental results based on the dataset of 19000 Facebook volunteers suggest the proposed method is effective for predicting personality traits

5.Paper Name: Personality Classification System using Data Mining

Author: Sandhya Katiyar

Abstract: Personality is one feature that determines how people interact with the outside world. Personality can be defined as a necessary element of a person's behavior. The way people interact with other people determines their personality. This paper covers the topic of Automated Personality Classification – a system that analyses the personality of a user based on certain features using Data Mining Algorithms. In this paper, a system is proposed which analyses the personality of an applicant. This system will be helpful for organizations as well as other agencies who would be recruiting applicants based on their personality rather than their technical knowledge. The personality prediction results are based on Big Five Personality traits and the classification is done using Na¨ıve Bayes Algorithm and Support Vector Machine.

6. paper Name: Prediction of the Big Five personality traits using static facial images of college students with different academic backgrounds

Author: Jia Xu1,2*, Weijian Tian1 ,Guoyun Lv1 , Shiya Liu3 ,Yangyu Fan

Abstract: Appearance can affect social interaction, which in turn affects personality development. There is ample evidence that facial morphology and social cues provide information about human personality and behavior. In this study, we focused on the relationship between self-reported personality characteristics and facial features. We propose a new approach for predicting college students' personality characteristics (on the basis of the Big Five personality characteristics) with static facial images. First, we construct a dataset containing 13,347 data pairs composed of facial images and personality characteristics. Second, we train a deep neural network with 10,667 sample pairs from the dataset and use the remaining samples to test (1335 pairs) and validate (1335 pairs) self-reported Big Five personalities. We trained a series of deep neural networks on a large, labeled dataset to predict the self-reported Big Five personality trait scores. This novel work applies deep learning to this topic. We also verify the network's advanced nature on the publicly available database with obvious personality characteristics. The experimental results show that 1) personality traits can be reliably predicted from facial images with an accuracy that exceeds 70most accurate, and the prediction accuracy exceeded 90better than traditional manual features in predicting personality characteristics. The results strongly support the application of neural networks trained on large-scale labeled datasets in multidimensional personality feature prediction from static facial images. 3) There are some differences in the personality traits of college students with different academic backgrounds. Future research can explore the relative contribution of other facial image features in predicting other personality characteristics

7. Paper Name: Clustering based Personality Prediction on Turkish Tweets

Author name: Tutaysalgir

Abstract: In this paper, we present a framework for predicting the personality traits by analyzing tweets written in Turkish. The prediction model is constructed with a clustering based approach. Since the model is based on linguistic features, it is language specific. The prediction model uses features applicable to Turkish language and related to writing style of Turkish Twitter users. Our approach uses anonymous BIG5 questionnaire scores of volunteer participants as the ground truth in order to generate personality model from Twitter posts. Experiment results show that constructed model can predict personality traits of Turkish Twitter users with relatively small errors

8. Paper Name: Makine Ögrenimi ˘ ˙Ile MBTI Ki¸silik Tipi Tahmini MBTI Personality Prediction With Machine Learning

Author: Kaan Sönmezöz

Abstract:- —Personality traits, continuously affect our lives, from our behavior to our career decisions. It's possible to design more accurate recommendation systems and develop more efficient digital marketing strategies by the help of personality traits. In this work, people's MBTI personality traits were predicted according to their social media posts. Although it is the first study which uses Turkish language, the results show that personality type prediction can be applied in Turkish language as well. The best results were obtained when the MBTI dimensions were predicted. The F-Score values of those except one dimension are approximately 60

9.Paper name: Comparative Analysis of Feature Selection Algorithms for Compu-tational Personality Prediction From Social Media

Author: Ahmed Al Marouf , Md. Kamrul Hasan , and Hasan Mahmud

Abstract: With the rapid growth of social media, users are getting involved in virtual socialism, generating a huge volume of textual and image contents. Consideringthe contents such as status updates/tweets and shared posts/retweets, liking otherposts is reflecting the online behavior of the users. Predicting personality of a userfrom these digital footprints has become a computationally challenging problem. Ina profile-based approach, utilizing the user-generated textual contents could be use-ful to reflect the personality in social media. Using huge number of features of dif-ferent categories, such as traditional linguistic features (character-level, word-level,structural, and so on), psycholinguistic features (emotional affects, perceptions, so-cial relationships, and so on) or social network features (network size, betweenness,and so on) could be useful to predict personality traits from social media. According to a widely popular personality model, namely, big-five-factor model (BFFM),the five factors are openness-to experience, conscientiousness, extraversion, agreeableness, and neuroticism. Predicting personality is redefined as predicting eachof these traits separately from the extracted features. Traditionally, it takes hugenumber of features to get better accuracy on any prediction task although applyingfeature selection algorithms may improve the performance of the model. In this arti-cle, we have compared the performance of five feature selection algorithms, namelythe Pearson correlation coefficient (PCC), correlation-based feature subset (CFS),information gain (IG), symmetric uncertainly (SU) evaluator, and chi-squared (CHI)method. The performance is evaluated using the classic metrics, namely, precision,recall, f-measure, and accuracy as evaluation matrices.

10. Paper Name: Predicting Personality Using Answers to Open-Ended InterviewQuestion

Author : MADHURA JAYARATNE1,2, (Member, IEEE) AND BUDDHI JAYATILLEKE1

Abstract: One's personality is widely accepted as an indicator of job performance, job satisfaction and tenure intention. The ability to measure an applicant's personality in the selection process helps recruiters, hiring managers and the applicant make better hiring decisions. Our work shows that textual content of answers to standard interview questions related to past behaviour and situational judgement can be used to reliably infer personality traits. We used data from over 46,000 job applicants who completed an online chat interview that also included a personality questionnaire based on the six-factor HEXACO personality model to self-rate their personality. Using natural language processing (NLP) and machine learning methods we built a regression model to infer HEXACO trait values from textual content. We compared the performance of five different text representation methods and found that term frequency-inverse document frequency (TF-IDF) with Latent Dirichlet Allocation (LDA) topics performed the best with an average correlation of r=0.39. As a comparison, a large study of Facebook messages based inference of Big 5 personality found an average correlation of r=0.35 and IBM's Personality Insights service built using twitter text data reports an average correlation of r=0.31. We further validated our model with a group of 117 volunteers who used an agreement scale of yes/no/maybe to rate the individual trait descriptors generated based on the model outcomes. On average, 87.83the six traits. The ability of algorithms to objectively infer a candidate's personality using only the textual content of interview answers presents significant opportunities to remove the subjective biases involved in human interviewer judgement of candidate personality

CHAPTER 3

PROBLEM STATEMENT

3.1 PROBLEM STATEMENT

We have extracted over 150 features to analyze the predictive system over different types of features, such as traditional linguistic, psycholinguistic, and SN features. In the literature, many researchers have used few features to predict personality, but the overall outcome of those approaches is not quite satisfactory. We have considered several scenarios/cases of feature combinations based on psycholinguistic features to find the best subset of features to predict each personality trait differently. Hence, we have determined the accuracy with and without SN features that are reported in the experiments.

CHAPTER 4

PROJECT REQUIREMENT

EXTERNAL INTERFACE REQUIREMENT

User Interface

Application Based Personality Prediction

Hardware Interfaces:

RAM : 8 GB

As we are using Machine Learning Algorithm and Various High Level Libraries Laptop

RAM minimum required is 8 GB.

Hard Disk : 40 GB

Data Set of CT Scan images is to be used hence minimum 40 GB Hard Disk memory is required.

Processor : Intel i5 Processor

Pycharm IDE that Integrated Development Environment is to be used and data loading should be fast hence Fast Processor is required

IDE : Pycharm

Best Integrated Development Environment as it gives possible suggestions at the time of typing code snippets that makes typing feasible and fast.

Coding Language : Python Version 3.5

Highly specified Programming Language for Machine Learning because of availability of High Performance Libraries.

Operating System : Windows 10

Latest Operating System that supports all type of installation and development Environment

Software Interfaces

Operating System: Windows 10

IDE: Pycharm ,Spyder

Programming Language : Python

NON FUNCTIONAL REQUIREMENT

Performance Requirements

The performance of the functions and every module must be well. The overallper-formance of the software will enable the users to work eciently. Perfor-mance of encryption of data should be fast. Performance of the providingvirtual environment should be fastSafety Requirement•The application is designed in modules where errors can be detected and xedeasily. This makes it easier to install and update new functionality if required.

Safety Requirement

The application is designed in modules where errors can be detected and fixed easily. This makes it easier to install and update new functionality if required.

Software Quality Attributes

Our software has many quality attribute that are given below:-

Adaptability: This software is adaptable by all users.

Availability: This software is freely available to all users. The availability of the software is easy for everyone.

Maintainability: After the deployment of the project if any error occurs then it can be easily maintained by the software developer.

Reliability: The performance of the software is better which will increase the reliability of the Software.

User Friendliness: Since, the software is a GUI application; the output generated is much user friendly in its behavior.

Integrity: Integrity refers to the extent to which access to software or data by unauthorized persons can be controlled.

Security: Users are authenticated using many security phases so reliable security is provided.

Testability: The software will be tested considering all the aspects.

CHAPTER 5

SYSTEM ANALYSIS
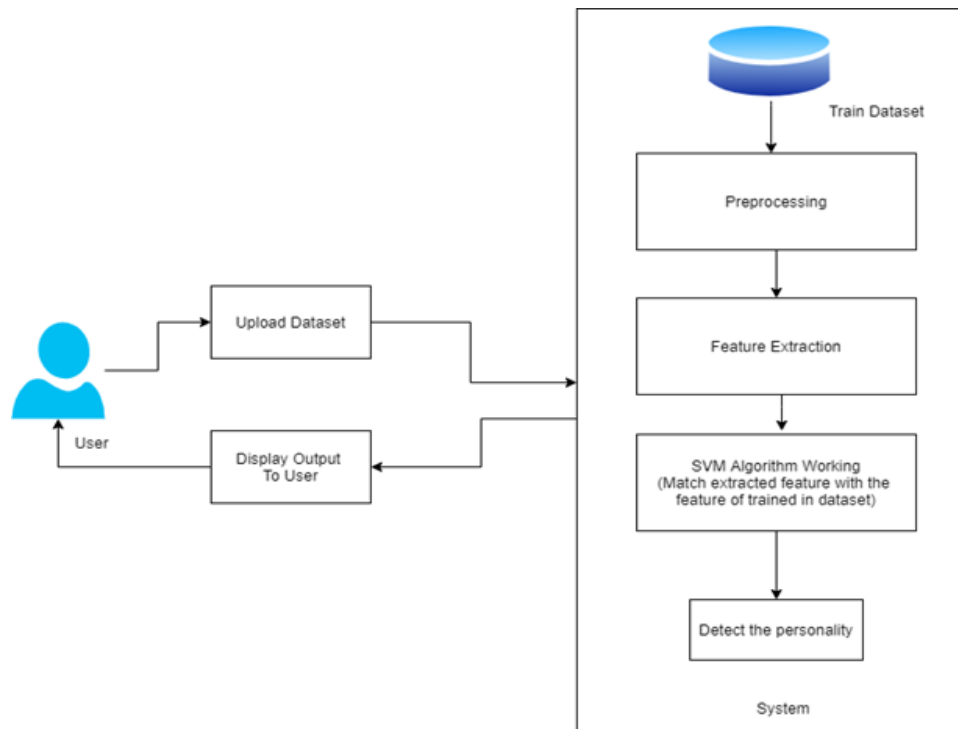
SYSTEM ARCHITECTURE

Figure 5.1: system Architecturel

Module

- Admin

- In this module,the Admin has to log in by using valid user name and password. After login successful he can do some operations such as View All Users and Authorize, View All E-Commerce Website and Authorize,View All Products and Reviews,View All Products Early Reviews, View All Keyword Search Details, View All Products Search Ratio, View All Keyword Search Results, View All Product Review Rank Results.

- View and Authorize Users

- In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and

admin authorizes the users.

- View Charts Results

- View All Products Search Ratio,View All Keyword Search Results,View All Product Review Rank Results.

- Ecommerce User

- In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful,he has to login by using authorized user name and password Once Login is successful user will do some operations like Add Products, View All Products with reviews, View All Early Product's reviews, View All Purchased Transactions.

- End User

- In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will best or to the database. After registration successful,he has to login by using authorized user name and password. Once Login is successful user will do some operations like Manage Account, Search Products by keyword and Purchase, View Your Search Transactions, View.

Data Flow Diagram

In Data Flow Diagram,we Show that flow of data in our system in DFD0 we show that base DFD in which rectangle present input as well as output and circle show our system,In DFD1 we show actual input and actual output of system input of our system is text or image and output is rumor detected like wise in DFD 2 we present

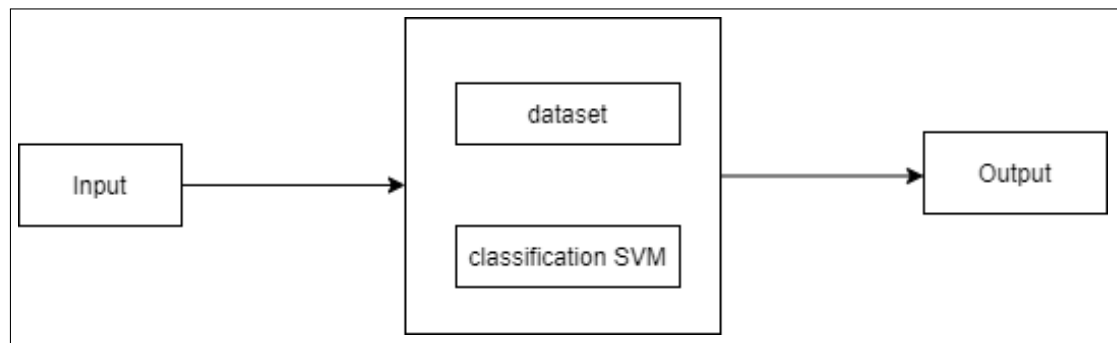operation of user as well as admin.

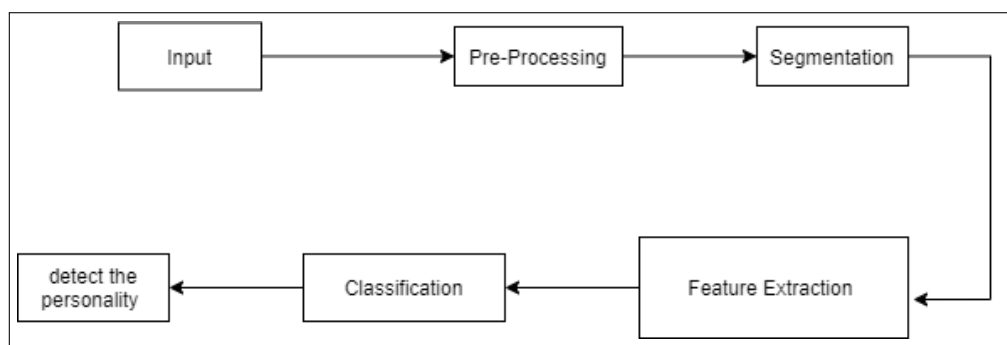Figure 5.2: Data Flow(0) diagram

Figure 5.3: Data Flow(1) diagram

Figure 5.4: Data Flow(2) diagram

UML DIAGRAMS

Unified Modeling Language is a standard language for writing software blueprints.The UML may be used to visualize,specify,construct and document the artifacts of a softwareintensive system.UML is process independent,although optimally it should be used in process that is use case driven,architecture-centric,iterative,and incremental.The Number of UML Diagram is available.

Class Diagram.

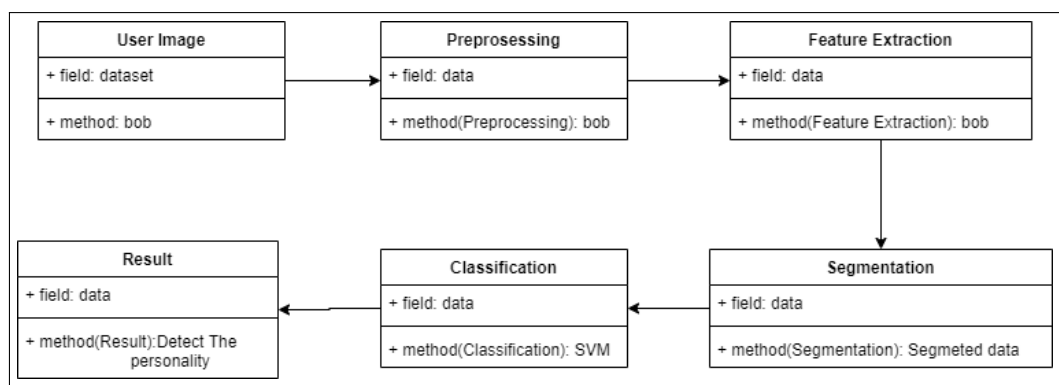Use case Diagram.

Activity Diagram.
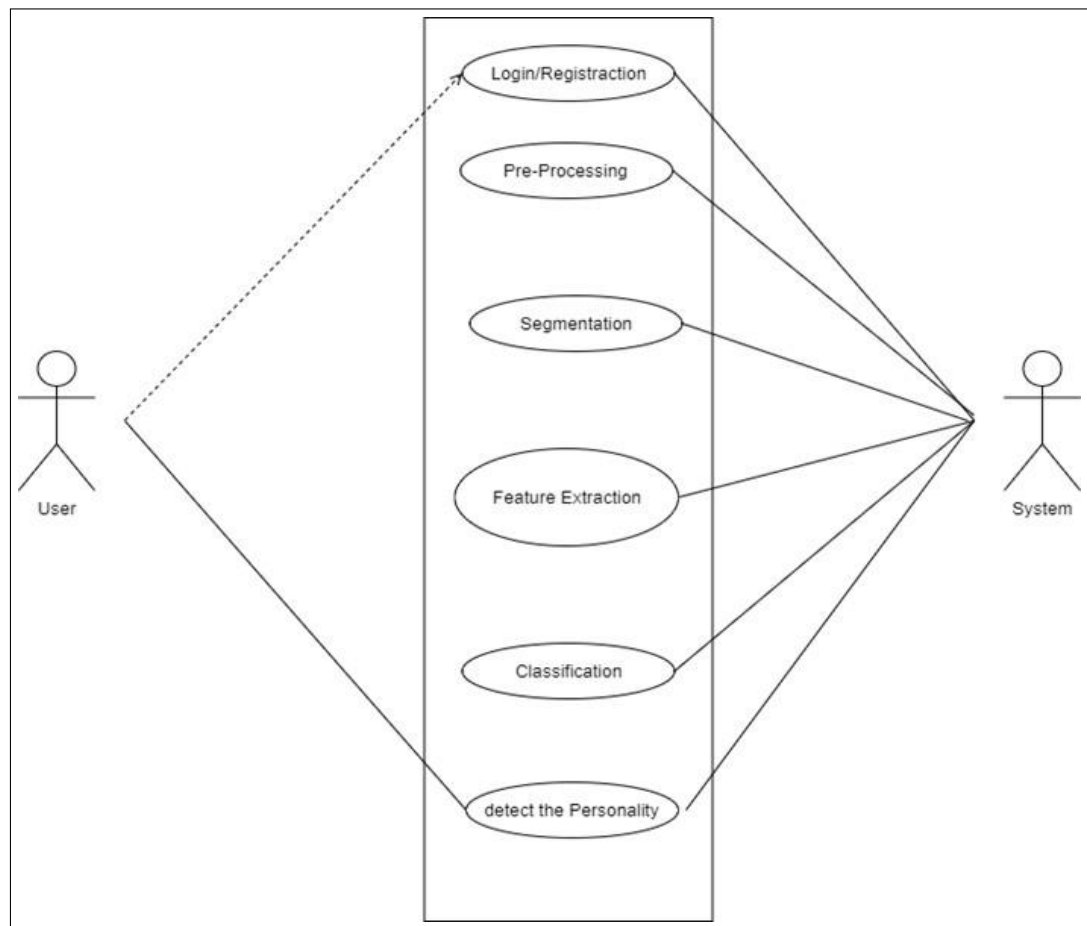
Sequence Diagram.

Figure 5.5: Class Diagram Diagram
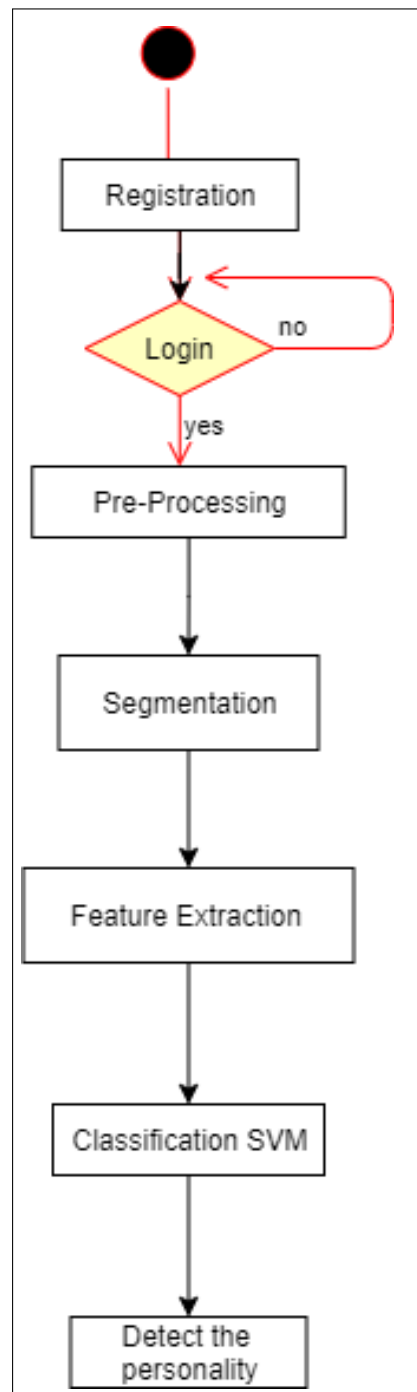
Figure 5.6: Use case Diagram

Figure 5.7: Activity Diagram
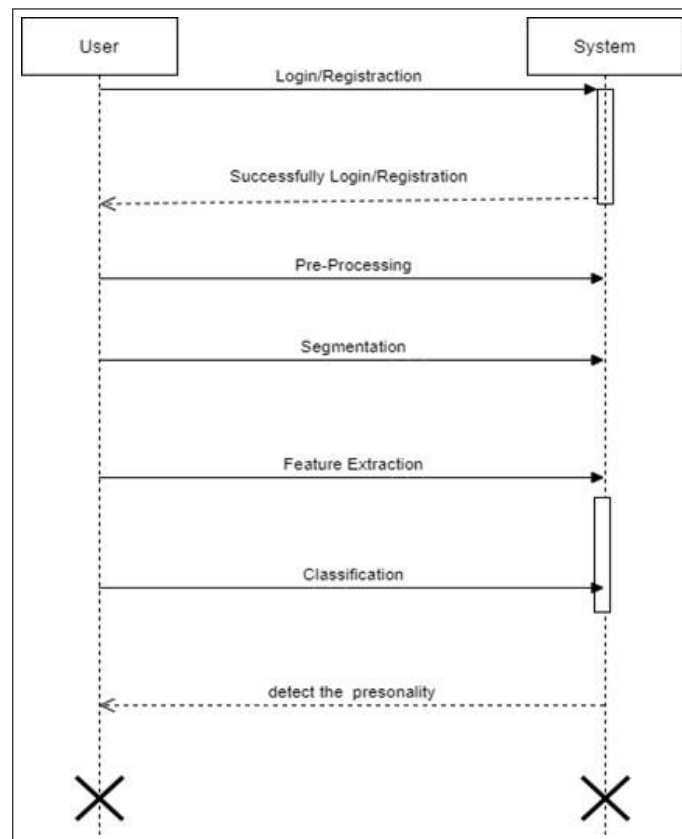
Figure 5.8: Sequence Diagram

CHAPTER 6

SOFTWARE INFORMATION

Python is an interpreted, high-level and general-purpose programming language. Created by Guido van Rossum and first released in 1991, Python's design philosophy emphasizes code readability with its notable use of significant whitespace. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.

Python is dynamically typed and garbage-collected. It supports multiple programming paradigms, including structured (particularly, procedural), object-oriented, and functional programming. Python is often described as a "batteries included" language due to its comprehensive standard library.

Python was created in the late 1980s as a successor to the ABC language. Python 2.0, released in 2000, introduced features like list comprehensions and a garbage collection system with reference counting.

Python 3.0, released in 2008, was a major revision of the language that is not completely backward-compatible, and much Python 2 code does not run unmodified on Python 3.

The Python 2 language was officially discontinued in 2020 (first planned for 2015), and "Python 2.7.18 is the last Python 2.7 release and therefore the last Python 2 release."[30] No more security patches or other improvements will be released for it. With Python 2's end-of-life, only Python 3.6.x and later are supported.

Python interpreters are available for many operating systems. A global community of programmers develops and maintains CPython, a free and open-source reference implementation. A non-profit organization, the Python Software Foundation, manages and directs resources for Python and CPython development.

Python was conceived in the late 1980s by Guido van Rossum at Centrum Wiskunde Informatica (CWI) in the Netherlands as a successor to the ABC language (itself inspired by SETL), capable of exception handling and interfacing with the Amoeba operating system. Its implementation began in December 1989. Van Rossum shouldered sole responsibility for the project, as the lead developer, until 12 July 2018, when he announced his "permanent vacation" from his responsibilities as Python's Benevolent Dictator For Life, a title the Python community bestowed

upon him to reflect his long-term commitment as the project's chief decision-maker. He now shares his leadership as a member of a five-person steering council. In January 2019, active Python core developers elected Brett Cannon, Nick Coghlan, Barry Warsaw, Carol Willing and Van Rossum to a five-member "Steering Council" to lead the project.

Anaconda: Anaconda is a free and open-source distribution of the Python and R programming languages for scientific computing (data science, machine learning applications, large-scale data processing, predictive analytics, etc.), that aims to simplify package management and deployment. The distribution includes data-science packages suitable for Windows, Linux, and macOS. It is developed and maintained by Anaconda, Inc., which was founded by Peter Wang and Travis Oliphant in 2012. As an Anaconda, Inc. product, it is also known as Anaconda Distribution or Anaconda Individual Edition, while other products from the company are Anaconda Team Edition and Anaconda Enterprise Edition, both of which are not free.

Package versions in Anaconda are managed by the package management system conda. This package manager was spun out as a separate open-source package as it ended up being useful on its own and for other things than Python. There is also a small, bootstrap version of Anaconda called Miniconda, which includes only conda, Python, the packages they depend on, and a small number of other packages. Anaconda distribution comes with over 250 packages automatically installed, and over 7,500 additional open-source packages can be installed from PyPI as well as the conda package and virtual environment manager. It also includes a GUI, Anaconda Navigator, as a graphical alternative to the command line interface (CLI).

The big difference between conda and the pip package manager is in how package dependencies are managed, which is a significant challenge for Python data science and the reason conda exists.

When pip installs a package, it automatically installs any dependent Python packages without checking if these conflict with previously installed packages[citation needed]. It will install a package and any of its dependencies regardless of the state of the existing installation[citation needed]. Because of this, a user with a working installation of, for example, Google Tensorflow, can find that it stops working having used pip to install a different package that requires a different version of the dependent numpy library than the one used by Tensorflow. In some cases, the package may appear to work but produce different results in detail.

In contrast, conda analyses the current environment including everything currently installed, and, together with any version limitations specified (e.g. the user

may wish to have Tensorflow version 2,0 or higher), works out how to install a compatible set of dependencies, and shows a warning if this cannot be done.

Open source packages can be individually installed from the Anaconda repository, Anaconda Cloud (anaconda.org), or the user's own private repository or mirror, using the conda install command. Anaconda, Inc. compiles and builds the packages available in the Anaconda repository itself, and provides binaries for Windows 32/64 bit, Linux 64 bit and MacOS 64-bit. Anything available on PyPI may be installed into a conda environment using pip, and conda will keep track of what it has installed itself and what pip has installed.

Custom packages can be made using the conda build command, and can be shared with others by uploading them to Anaconda Cloud, PyPI or other repositories.

The default installation of Anaconda2 includes Python 2.7 and Anaconda3 includes Python 3.7. However, it is possible to create new environments that include any version of Python packaged with conda

CHAPTER 7

PROJECT PLAN

In this chapter we are going to have an overview about how much time does it took to complete each task like- Preliminray Survey Introduction and Problem Statement, Literature Survey, Project Statement, Software Requirement and Specification, System Design, Partial Report Submission, Architecture Design, Implementation, Deployment, Testing, Paper Publish, Report Submission and etcetera. This chapter also gives focus on stakeholder list which gives information about project type, customer of the proposed system, user and project member who developed the system.

STAKEHOLDER LIST

SYSTEM IMPLEMENTATION PLAN

The System Implementation plan table, shows the overall schedule of tasks compilation and time duration required for each task.

Sr. No.	Name/Title	Start Date	End Date
1	Preliminary Survey		
2	Introduction and Problem Statement		
3	Literature Survey		
4	Project Statement		
5	Software Requirement And Specification		
6	System Design		
7	Partial Report Submission		
8	Architecture Design		
9	Implementation		
10	Deployement		
11	Testing		
12	Paper Publish		
13	Report Submission		

CHAPTER 8

RESULT

8.1 RESULT
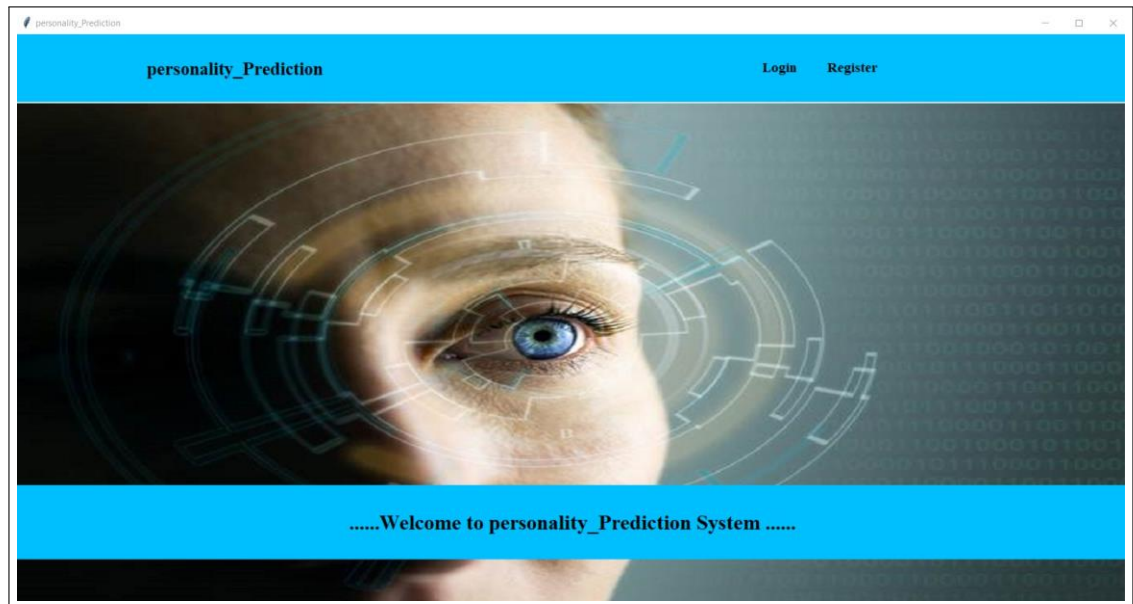
Figure 8.1: main File

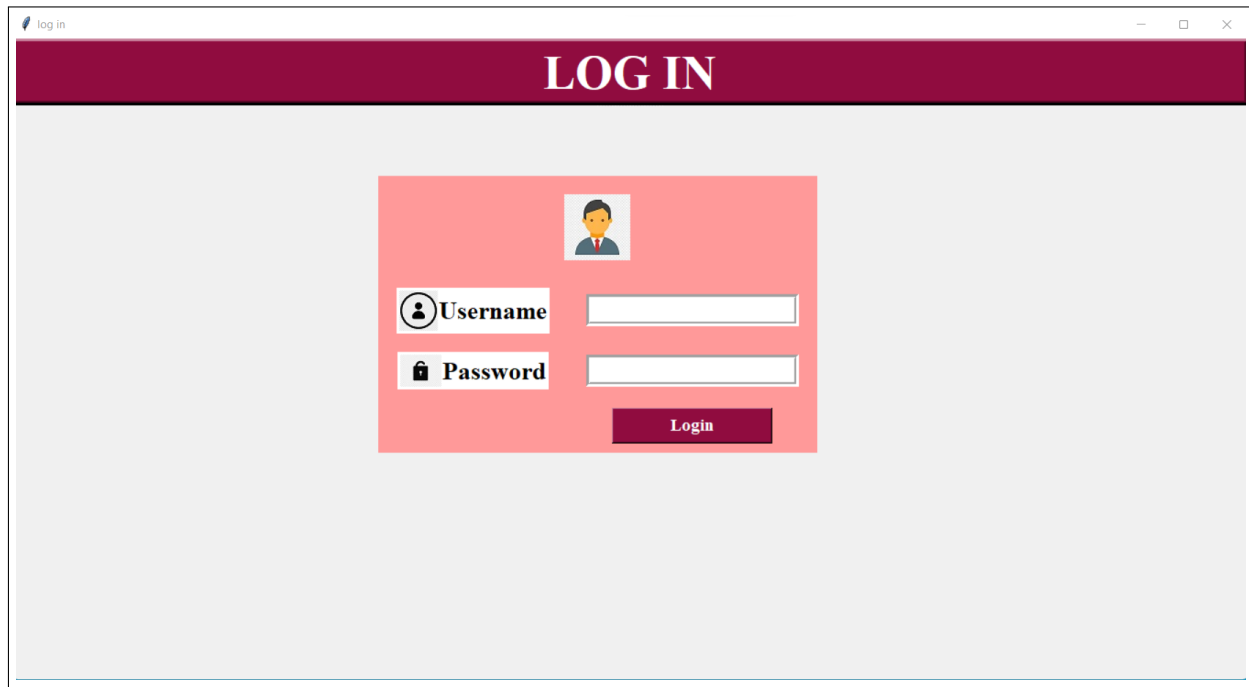Figure 8.2: Registration File

Figure 8.3: master

Figure 8.4: Check File

CHAPTER 9

TEST PLAN

9.1   TEST CASES

Test case	Login Screen- Sign up
Objective	Click on sign up button then check all required/ mandatory fields with leaving all fields blank
Expected Result	All required/ mandatory fields should display with symbol "*". Instruction line "* field(s) are mandatory" should be displayed
Test case	Create a Password >>Text Box Confirm Password >>Text Box
Objective	Check the validation message for Password and Confirm Password field
Expected Result	Correct validation message should be displayed accordingly or "Password and confirm password should be same" in place of "Password mismatch".

Figure 9.1: GUI TESTING

Test Case ID	Test Case	Test Case I/P	Actual Result	Expected Result	Test case criteria(P/F)
001	Enter the number in username, middle name, last name field	Number	Error Comes	Error Should Comes	P
001	Enter the character in username, middle name, last name field	Character	Accept	Accept	p
002	Enter the invalid email id format in email id field	Kkgmail,com	Error comes	Error Should Comes	P
002	Enter the valid email id format in email id field	kk@gmail.com	Accept	Accept	P
003	Enter the invalid digit no in phone no field	99999	Error comes	Error Should Comes	P
003	Enter the 10 digit no in phone no field	9999999999	Accept	Accept	P

Figure 9.2: Registration test case

Test Case ID	Test Case	Test Case I/P	Actual Result	Expected Result	Test case criteria(P/F)
001	Enter The Wrong username or password click on submit button	Username or password	Error comes	Error Should come	P
002	Enter the correct username and password click on submit button	Username and password	Accept	Accept	P

Figure 9.3: Login test case

Test Case ID	Test Case	Test Case I/P	Actual Result	Expected Result	Test case criteria(P/F)
001	Store Xml File	Xml file	Xml file store	Error Should come	P
002	Parse the xml file for conversion	parsing	File get parse	Accept	P
003	Attribute identification	Check individual Attribute	Identify Attributes	Accepted	P
004	Weight Analysis	Check Weight	Analyze Weight of individual Attribute	Accepted	P
005	Tree formation	Form them-Tree	Formation	Accepted	P
006	Cluster Evaluation	Check Evaluation	Should check Cluster	Accepted	P
007	Algorithm Performance	Check Evaluation	Should work Algorithm Properly	Accepted	P
008	Query Formation	Check Query Correction	Should check Query	Accepted	P

Figure 9.4: System Test Cases:

CHAPTER 10

CONCLUSION

CONCLUSION

As we know, the base of social media is basically a graph. The connections between the nodes and impact on them due to social media interactions could be reflected through the SN features. We have designed and performed experiments utilizing the linguistic features as well as SN features. To the best of our knowledge, we have used the most number of features to compare the performance of the feature selection algorithms to predict personality traits.

CHAPTER 11

REFERENCES

1 M. R. Barrick and M. K. Mount, "The big five personality dimensions and job performance: A meta-analysis," Personnel Psychology, vol. 44, no. 1, pp. 1–26, 1991. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-6570.1991.tb00688.x

2 S. Rothmann and E. P. Coetzer, "The big five personality dimensions and job performance," SA Journal of Industrial Psychology, vol. 29, no. 1, pp. 68–74, 2003. [Online]. Available: https://journals.co.za/content/psyc/29/1/EJC88938

3 J. F. Salgado, "The big five personality dimensions and counterproductive behaviors," International Journal of Selection and Assessment, vol. 10, no. 1, pp. 117–125, 2002. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-2389.00198

4 J. W. Lounsbury, R. P. Steel, L. W. Gibson, and A. W. Drost, "Personality traits and career satisfaction of human resource professionals," Human Resource Development International, vol. 11, no. 4, pp. 351–366, 2008. [Online]. Available: https://doi.org/10.1080/13678860802261215

5 J. W. Lounsbury, N. Foster, H. Patel, P. Carmody, L. W. Gibson, and D. R. Stairs, "An investigation of the personality traits of scientists versus nonscientists and their relationship with career satisfaction," RD Management, vol. 42, no. 1, pp. 47–59, 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/9310.2011.00665.x

6 V. Ariyabuddhiphongs and S. Marican, "Big five personality traits and turnover intention among thai hotel employees," International Journal of Hospitality Tourism Administration, vol. 16, no. 4, pp. 355–374, 2015. [Online]. Available: https://doi.org/10.1080/15256480.2015.1090257

7   T. A. Timmerman, "Predicting turnover with broad and narrow personality traits," International Journal of Selection and Assessment, vol. 14, no. 4, pp. 392–399, 2006. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-2389.2006.00361.x

8   P. T. Costa Jr. and R. R. McCrae, "The Revised NEO Personality Inventory (NEO-PI-R)." in The SAGE handbook of personality theory and assessment, Vol 2: Personality measurement and testing. Thousand Oaks, CA, US: Sage Publications, Inc, 2008, pp. 179–198.

9   K. Lee and M. C. Ashton, "Psychometric properties of the hexaco-100," Assessment, vol. 25, no. 5, pp. 543–556, 2018. [Online]. Available: https://doi.org/10.1177/107

10 N. Anderson, J. Salgado, and U. Hülsheger, "Applicant reactions in selection: Comprehensive meta-analysis into reaction generalization versus situational specificity," International Journal of Selection and Assessment, vol. 19, pp. 291–304, 08 2010