Personality Prediction Using Social Media : A Neural Network Approach

Submitted in the partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY
IN
COMPUTER ENGINEERING

Under the Supervision of
Mr. Mohd. Zeeshan Ansari

Submitted By
Akshay Kumar (10-CSS-06)
Niyas C (10-CSS-44)

DEPARTMENT OF COMPUTER ENGINEERING
FACULTY OF ENGINEERING AND TECHNOLOGY
JAMIA MILLIA ISLAMIA, NEW DELHI – 110025
Year – 2014

CERTIFICATE

This is to certify that the project (CEN-891) entitled "Personality Prediction Using Social Media" done by Akshay Kumar (10-CSS-06) and Niyas C (10-CSS-44) respectively is an authentic work carried out by them under my guidance.

The matter embodied in this project work has not been submitted earlier for the award of any degree or diploma to the best of my knowledge and belief.

Dr. Tanveer Ahmad
Department of Computer Engg.
Faculty of Engineering and Technology
Jamia Millia Islamia
New Delhi-25

Mr. Mohd. Zeeshan Ansari
Department of Computer Engg.
Faculty of Engineering and Technology
Jamia Millia Islamia
New Delhi-25

ACKNOWLEDGEMENT

We are greatly indebted to our supervisor and guide, **Mr. Mohd. Zeeshan Ansari**, for his invaluable technical guidance and support, great innovative ideas and overwhelming moral support, great course of our project. We are also grateful to our HOD, **Prof. Tanveer Ahmad**, for his valuable support through out the project. We also express our gratitude to the Department Of Computer Engineering and the entire faculty members, for their teaching and guidance and encouragement. We are also thankful to our classmates and friends for their valuable suggestions and support.

Akshay Kumar (10-CSS-06)
Department of Computer Engg.
Faculty of Engg. & Tech.
Jamia Millia Islamia,
New Delhi-25

Niyas C (10-CSS-44)
Department of Computer Engg.
Faculty of Engg. & Tech.
Jamia Millia Islamia,
New Delhi-25

Table of Contents

1 Introduction

1.1 Personality

Personality, a person's characteristic behavior pattern. Many people think of personality as being made up of observable traits such as shyness, friendliness, and initiative. However, such traits are only the outward expression of various inner conditions and processes such as intelligence, attitudes, interests, and motives. Many psychologists include these inner elements in their definitions of personality.

Normal persons develop relatively consistent personalities that are distinguished by certain dominant traits. Persons belonging to the same cultural group—such as a family or nation—have many traits in common. But each person exhibits different traits under different circumstances, and each has unique traits as well as those he or she shares with others. It is also known that personalities develop throughout life, and may even undergo fundamental changes as a result of psychotherapy or other treatment. For these reasons, psychologists do not believe that personalities can be rigidly classified according to type.

Personality is produced by the interaction of heredity and environment. Inborn qualities affect the individual's response to the outside world, and the environment influences the way in which inborn capacities develop. However,

the precise reasons why one person develops certain personality traits, while another develops other traits, are not known.

Studies of personality origins are limited by the difficulty of making controlled experiments on human behavior. Except for identical twins, no two persons have exactly the same biological inheritance, and even identical twins do not share exactly the same environment. Parents and others respond to each twin differently, providing each with a unique emotional setting.

Inherited traits such as bone structure, skin color, type of hair—play a part in personality only if given a meaning by the environment. For example, an extremely tall boy may develop either shyness or self-confidence, according to whether he is teased because of his height or praised for using it to advantage in playing basketball.

The structure and function of the nervous and glandular systems are inherited qualities having a more direct effect upon personality. Intelligence, talents, and skills are largely dependent upon these systems. However, the effects of even these qualities can be modified by the environment.

External influences include both physical environment (climate, geography) and social environment (other individuals, and human institutions). Social influences are considered to be of greater importance in personality formation. Most psychologists believe that basic personality traits are

acquired in early childhood, and that the family is therefore of primary importance in forming an individual's personality.

Wider cultural groups such as tribes and nations set up rules, values, and goals, and thus influence personality formation. Diversity within large groups is produced by subcultures such as social and economic classes and religious groups.

1.2 Traits theory

In psychology, trait theory (also called dispositional theory) is an approach to the study of human personality. Trait theorists are primarily interested in the measurement of traits, which can be defined as habitual patterns of behavior, thought, and emotion. According to this perspective, traits are relatively stable over time, differ across individuals (e.g. some people are outgoing whereas others are shy), and influence behavior.

Gordon Allport was an early pioneer in the study of traits, which he sometimes referred to as dispositions. In his approach, central traits are basic to an individual's personality, whereas secondary traits are more peripheral. Common traits are those recognized within a culture and may vary between cultures. Cardinal traits are those by which an individual may be strongly recognized. Since Allport's time, trait theorists have focused more on group statistics than on single individuals. Allport called these two emphases

"nomothetic" and "idiographic," respectively.

There is a nearly unlimited number of potential traits that could be used to describe personality. The statistical technique of factor analysis, however, has demonstrated that particular clusters of traits reliably correlate together. Hans Eysenck has suggested that personality is reducible to three major traits. Other researchers argue that more factors are needed to adequately describe human personality including humor, wealth and beauty. Many psychologists currently believe that five factors are sufficient.

Virtually all trait models, and even ancient Greek philosophy, include extraversion vs. introversion as a central dimension of human personality. Another prominent trait that is found in nearly all models is Neuroticism, or emotional instability.

1.3 The Big 5 Personality traits

Today, many contemporary personality psychologists believe that there are five basic dimensions of personality, often referred to as the "Big 5" personality traits. Previous trait theorist had suggested a various number of possible traits, including Gordon Allport's list of 4,000 personality traits, Raymond Cattell's 16 personality factors and Hans Eysenck's three-factor theory.

However, many researchers felt that Cattell's theory was too complex

and Eysenck's was too limited in scope. As a result, the five-factor theory emerged to describe the basic traits that serve as the building blocks of personality.

The "big five" are broad categories of personality traits. While there is a significant body of literature supporting this five-factor model of personality, researchers don't always agree on the exact labels for each dimension. However, these five categories are usually described as follows:

1. **Extraversion:** This trait includes characteristics such as excitability, sociability, talkativeness, assertiveness and high amounts of emotional expressiveness.
2. **Agreeableness:** This personality dimension includes attributes such as trust, altruism, kindness, affection, and other prosocial behaviors.
3. **Conscientiousness:** Common features of this dimension include high levels of thoughtfulness, with good impulse control and goal-directed behaviors. Those high in conscientiousness tend to be organized and mindful of details.
4. **Neuroticism:** Individuals high in this trait tend to experience emotional instability, anxiety, moodiness, irritability, and sadness.
5. **Openness:** This trait features characteristics such as imagination and

insight, and those high in this trait also tend to have a broad range of interests.

1.4 Social Media and Personality

Communication methods have developed substantially over the past two decades and these changes have influenced consumption patterns. Extensive use of information technology has become prevalent and the emergence of the World Wide Web and Internet-based social media allows consumers to interact easily both with other individuals en masse and with commercial entities (Mangold & Faulds, 2009). In particular, Web 2.0 is an important technological infrastructure, with its customer-based content properties and collective social media platforms, which enable the creation and distribution of social information (Berthon, Pitt, Plangger, & Shapiro, 2012). This interactive technology has enabled blogs and other social networking forums to develop rapidly. Content sharing, personal commentary, and private life broadcasting have emerged as the distinct elements of this new media. Accordingly, social networking websites allow their users to find individuals with similar interests for social and emotional purposes (Ross et al., 2009). McKenna and Bargh (2000) suggested that there are two main drivers behind the tendency to interact with other people on the Internet: individuals' characteristics and their social backgrounds.

Social media users are most commonly young adults (aged 18 to 31); three-quarters of adult Internet users under age 25 have a profile on a social media site (Lenhart, 2009). Social media websites are designed to be widely accessible and initially attract homogeneous populations, so it is not uncommon to find groups using websites to segregate themselves according to nationality, age, educational level, or other factors that typically segment society, even if that was not the intention of the designers (Boyd & Ellison, 2008). There are gender differences among personality traits. Correa, Hinsley, and Zuniga (2010) found that extraversion was positively related to social media use, while emotional stability was negatively related to the use of these online social applications. In the study by Correa et al., the relationship between social media use and openness to experience was not found to be statistically significant for males. However, men who were more extraverted and anxious were more likely to use social media. For women, extraversion and openness to experience were positively related to social media use. In contrast to men, emotional stability was unrelated to social media use. There was a relationship between demographic variables (including age, gender, level of education, and level of income) and personality traits.

2 Previous works/ Related Works

Data mining techniques play a fundamental role in extracting correlation patterns between personality and variety of user's data captured from multiple sources. Generally, two approaches were adopted for studying personality traits of social network users. The first approach uses a variety of machine learning algorithms to build models based on social network activities only. The second one extends the personality-related features with linguistic cues (Mairesse et al. 2007; Oberlander and Nowson 2006).

Several classification and regression techniques were used to build predictive personality models along the five personality dimensions using the linguistic features of a dataset comprised of few thousand essays solicited from introductory psychology students (Mairesse et al. 2007). The Linguistic Inquiry and Word Count – LIWC (http://www.liwc.net) was used as a tool for linguistic analysis. The reported precisions of the classifiers were in the range of 54-62% for all traits. In (Oberlander and Nowson 2006), SMO and Naïve Bayes were used for modeling four out of five personality dimensions by extracting n- gram features from a corpus of personal web-blogs. Their results point out to the importance of the process of feature selection in increasing the classifiers precision yielding 83%-93% for automatic feature selection. We would like to point out the differences in the datasets used in

these studies compared to ours, namely different solicitation methods and the sources from which they were collected.

The correlation between users' social network activity and personality has been the focus of several studies in the last decade (Bai, Zhu, and Cheng 2012; Golbeck, Robles, and Turner 2011; Bachrach et al. 2012). Personality traits of the Chinese most popular social network RenRen users were analyzed in (Bai, Zhu, and Cheng 2012). C4.5 Decision Trees have shown the best results, yielding 69-72 percent accuracy, for a combination of features related to users' network activity along with affective linguistic features extracted from statuses and blog posts.

Two regression techniques, namely m5sup/Rules and Gaussian Processes, were applied to build predictive personality models. The authors consider users' Facebook data through parameters such as structural characteristics, personal info, activities and preference, in addition to the linguistic attributes extracted with LIWC from the users' statuses. The lack of demographic diversity in participant sampling was one of the major drawbacks for generalizing the results of the last two studies, Chinese population and authors' Facebook friends respectively.

Few studies using considerably larger number of instances from the same dataset under our investigation have a rather different objective from

ours, namely to examine the correlations between the personality traits and Facebook activity data (Bachrach et al. 2012) and the associations between personal attributes and Facebook Likes (Kosinski, Stillwell, and Graepel 2013). These studies were not meant to look at the rich linguistic patterns that occur in the language use on social networks, which is in the focus of this research.

3 Working Method

3.1 Data Collection

In order to train the proposed neural network system, we must collect the data. Here the data means, the actual values of individual's personality traits and their facebook information. The actual values of individuals personality information is collected using 44-item inventory that measures an individual on the Big Five Factors (dimensions) of personality which was proposed by Gold Berg in 1993. The inventory contain 44 question and the answer of each question can be

1. Strongly Disagree - 1
2. Disagree – 2
3. Neither Agree Nor Disagree – 3
4. Agree – 4
5. Strongly Agree - 5

The number given on right side of each answer indicates the score of each answer. For some questions score may be in reverse order. Along with these questions we request user's permission to access relevant information from their facebook account.

The questions in 44-item inventory is as follows .

1. I see as my self as one who is talkative

2. I see as my self as one who tends to find fault with others

3. I see as my self as one who does a through job.

4. I see as my self as one who is depressed, blue.

5. I see as my self as one who is original, come up with new ideas.

6. I see as my self as one who is reserved

7. I see as my self as one who is helpful and unselfish with others

8. I see as my self as one who can be somewhat careless.

9. I see as my self as one who is relaxed, handles stress very well.

10. I see as my self as one who is curious about many different things.

11. I see as my self as one who is full of energy

12. I see as my self as one who starts quarrel with others

13. I see as my self as one who is a reliable worker

14. I see as my self as one who can be tense

15. I see as my self as one who is ingenious a deep thinker

16. I see as my self as one who generates a lot of enthusiasm

17. I see as my self as one who has a forgiving nature

18. I see as my self as one who tend to be disorganized

19. I see as my self as one who worries a lot

20. I see as my self as one who has an active imagination

21. I see as my self as one who tend to be quiet.

22. I see as my self as one who is generally trusting

23. I see as my self as one who tends to be lazy

24. I see as my self as one who is emotionally stable not easily upset

25. I see as my self as one who is inventive

26. I see as my self as one who has an assertive personality

27. I see as my self as one who can be cold and aloof

28. I see as my self as one who preserves until the task is finished

29. I see as my self as one who can be moody

30. I see as my self as one who values artistic, aesthetic experience

31. I see as my self as one who is sometimes shy, inhibited

32. I see as my self as one who is considerate and kind to almost everyone

33. I see as my self as one who does things efficiently

34. I see as my self as one who remains calm in tense situations

35. I see as my self as one who prefers to work that is routine

36. I see as my self as one who is outgoing, sociable

37. I see as my self as one who is sometimes rude to others

38. I see as my self as one who makes plans and follows through with them.

39. I see as my self as one who gets nervous easily

40. I see as my self as one who likes to reflect plays with ideas.

41. I see as my self as one who has a few artistic interests

42. I see as my self as one who likes to cooperate with others

43. I see as my self as one who is easily distracted.

44. I see as my self as one who is sophisticated in arts, music or literature.

These questions are evaluated using following scale (R indicates that score is in reverse order)

Extraversion: 1, 6R, 11, 16, 21R, 26, 31R, 36

Agreeableness: 2R, 7, 12R, 17, 22, 27R, 32, 37R, 42

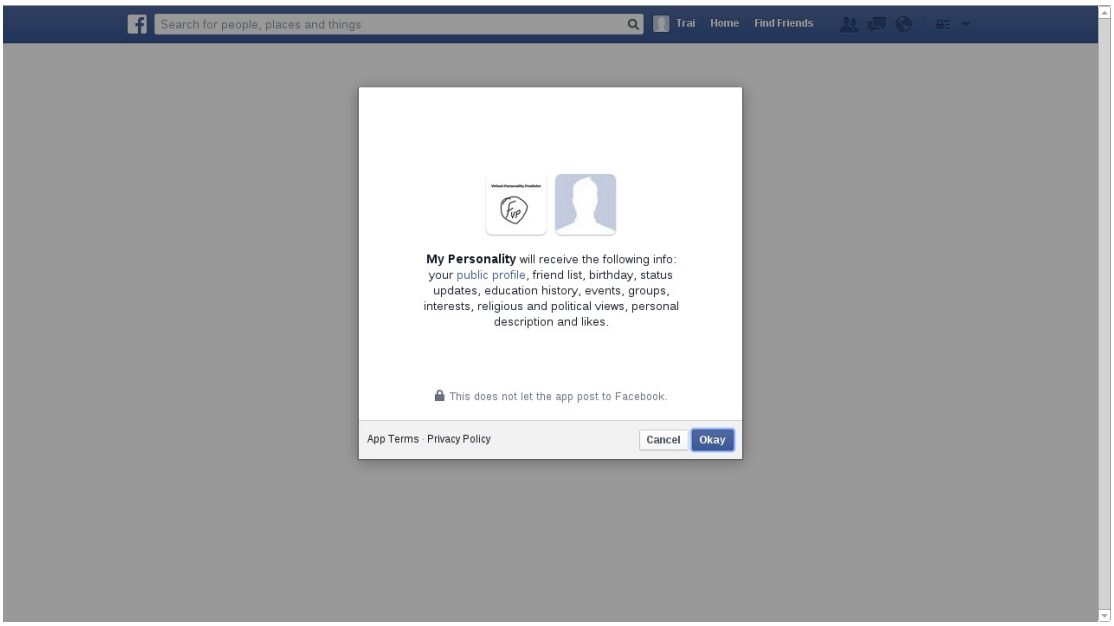
Conscientiousness: 3, 8R, 13, 18R, 23R, 28, 33, 38, 43R

Neuroticism: 4, 9R, 14, 19, 24R, 29, 34R, 39

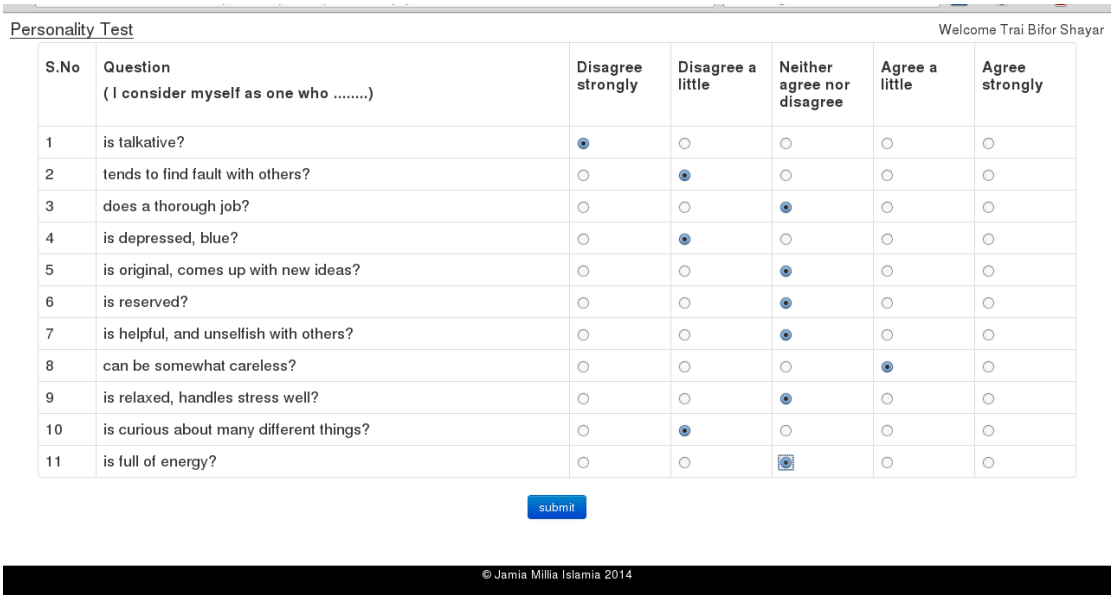
Openness: 5, 10, 15, 20, 25, 30, 35R, 40, 41R, 44

Average of sum of these scores are taken to find the personality trait value of individuals.

Following are few of the screen shots of data collection system developed for the purpose of project.



Requesting user permission to access relevant facebook information



Part of questionnaire

3.2 Feature Extraction

After the first phase, we have permission to access user's import facebook features. We extract following features from their account

3.2.1 Basic Statistical Details

It includes the followings.

1. Number of friends
2. Ratio of friends in same gender to total number of friends
3. Ratio of friends in other gender to total number of friends
4. Ration of same gender friends and other gender friends
5. Number of groups
6. Number of Likes
7. Number of statuses

3.2.2 Category wise fraction of page likes

In this section we find fraction of page likes in each category using the formula

$$fraction(category) = \frac{\text{Number of page likes in category}}{\text{Total Number of Pages likes}}$$

We have taken following 160 categories in to consideration.

['Political party', 'Non-profit organization', 'Website', 'Community', 'Education',

'Journalist', 'Just for fun', 'Computers/internet website', 'Politician', 'Political organization', 'Organization', 'Tv channel', 'App page', 'Media/news/publishing', 'Non-governmental organization (ngo)', 'Community organization', 'News/media website', 'Local business', 'Public figure', 'Education website', 'Entertainment website', 'Magazine', 'Company', 'School', 'Entertainer', 'Product/service', 'Interest', 'Computers', 'Studio', 'Computers/technology', 'Software', 'Cause', 'Personal blog', 'Religion', 'Electronics', 'University', 'Book', 'Games/toys', 'Tv show', 'Movie', 'Government organization', 'Food/beverages', 'Travel/leisure', 'Professional sports team', 'Cars', 'Radio station', 'Recreation/sports website', 'Fictional character', 'Author', 'Telecommunication', 'Camera/photo', 'Music chart', 'Clothing', 'Athlete', 'Retail and consumer merchandise', 'Video game', 'Book series', 'Music', 'Musician/band', 'Book genre', 'Actor/director', 'Sports/recreation/activities', 'Tv network', 'Sports league', 'Sport', 'State/province/region', 'City', 'Tv genre', 'Teacher', 'Artist', 'Society/culture website', 'Writer', 'Personal website', 'News personality', 'Public places', 'Consulting/business services', 'Education/work status', 'Outdoor gear/sporting goods', 'Automobiles and parts', 'Restaurant/cafe', 'Food', 'Sports venue', 'Movie character', 'Government official', 'Arts/humanities website', 'Bank/financial institution', 'Song', 'Internet/software', 'Household supplies', 'Attractions/things to do', 'Record label', 'Comedian', 'Arts/entertainment/nightlife', 'Airport', 'Science website', 'Doctor', 'Waterfall',

'Album', 'Bar', 'Health/beauty', 'Regional website', 'Reference website', 'Musical genre', 'Tv', 'Tv/movie award', 'Health/medical/pharmaceuticals', 'Food/grocery', 'Business person', 'Jewelry/watches', 'Professional services', 'Amateur sports team', 'Shopping/retail', 'Library', 'Movie theater', 'Literary editor', 'Other', 'Monarch', 'Aerospace/defense', 'Publisher', 'Work position', 'Book store', 'Language', 'Small business', 'Music video', 'Community/government', 'Church/religious organization', 'Teens/kids website', 'Country', 'Field of study', 'Movie general', 'Profession', 'Legal/law', 'Automotive', 'Real estate', 'Pet services', 'Engineering/construction', 'Home improvement', 'Home/garden website', 'Transport/freight', 'Tours/sightseeing', 'Hotel', 'Chef', 'Vitamins/supplements', 'Phone/tablet', 'Local/travel website', 'Business services', 'Neighborhood', 'Government website', 'Hospital/clinic', 'Health/wellness website', 'Landmark', 'Diseases', 'Club', 'Drugs', 'Event planning/event services', 'Museum/art gallery', 'Photographer', 'Producer', 'Concentration or major', 'Health/medical/pharmacy',]

3.2.3 Linguistic features

In this section we perform natural language processing on user's status updates and extract different features. PHP natural language processing library named NLP Tools was used for this purpose. Following features were extracted in this section

1. Number of words referencing singular first person.

2. Number of words referencing plural first person

3. Number of words referencing second party

4. Number of words referencing singular third party

5. Number of words referencing plural third party

6. Number of relative words

7. Number of singular indefinite pronouns

8. Number of plural indefinite pronouns

9. Number of symbols

10. Number of indefinite sop words

11. Number of words/status

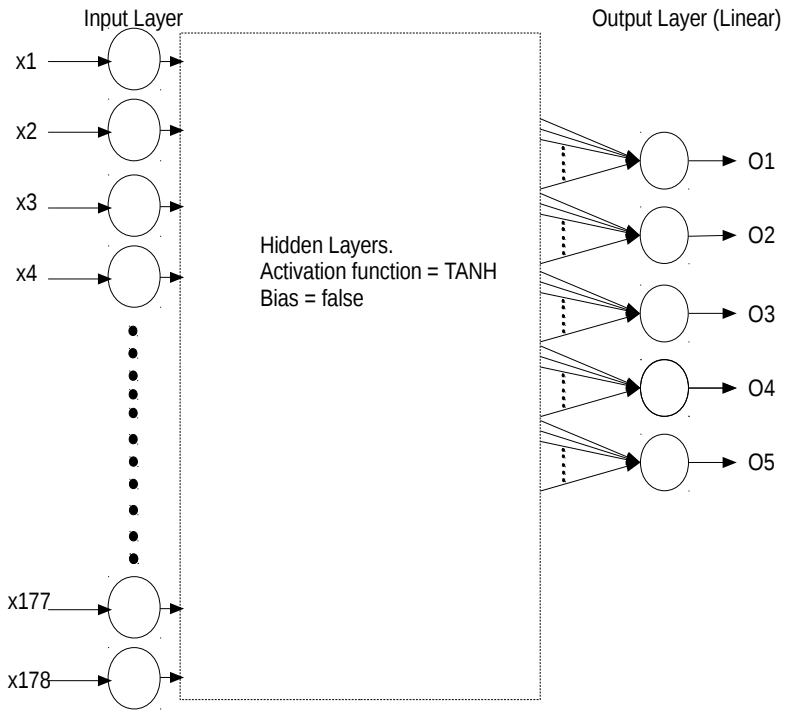
3.3 Training the neural network

Once we have the personality traits of individuals and their facebook features, we can use these information to train our neural network. The features extracted will serve as input and the values of personality traits will serve as output for training of neural network.

We have used python library named PyBrain for training the neural network. PyBrain is a modular Machine Learning Library for Python. Its goal is to offer flexible, easy-to-use yet still powerful algorithms for Machine

Learning Tasks and a variety of predefined environments to test and compare algorithms.

PyBrain is short for Python-Based Reinforcement Learning, Artificial Intelligence and Neural Network Library. In fact, we came up with the name first and later reverse-engineered this quite descriptive "Backronym".



4 Experimental Results

In a straight forward method, out of 24 total available samples, 20 samples were used for training the neural network and 4 samples to test the trained system. Following were the results of testing obtained.

Test set 1

Personality Trait	Predicted Value	Actual Value	Error %
Extraversion	2.31	3.5	29.65
Agreeableness	3.15	3.67	12.94
Conscientiousness	2.82	2.0	20.54
Neuroticism	2.64	2.25	9.78
Openness	3.17	4.3	28.24

Test set 2

Personality Trait	Predicted Value	Actual Value	Error %
Extraversion	3.11	2.0	27.88
Agreeableness	3.71	4.33	15.63
Conscientiousness	3.29	2.67	15.66
Neuroticism	2.7	3.0	7.40
Openness	3.52	3.33	5.72

Test set 3

Personality Trait	Predicted Value	Actual Value	Error %
Extraversion	3.11	2.625	12.25
Agreeableness	3.71	4.44	18.41
Conscientiousness	3.29	3.22	1.77
Neuroticism	2.70	3.5	19.91
Openness	3.53	3.0	13.22

Test set 4

Personality Trait	Predicted Value	Actual Value	Error %
Extraversion	3.37	2.25	28.17
Agreeableness	4.63	4.66	0.77
Conscientiousness	2.00	3.56	38.83
Neuroticism	3.68	3.375	7.65
Openness	2.76	3.8	25.88

We also performed random sub set sampling method, in which we for a fixed number of training samples, say k, we randomly select k training samples from the dataset and the remaining data samples in dataset will be served as test set. We performed this test for number of training samples, ie k ranging from 3 to 21 and 20 iteration were performed for each. The average of accuracy obtained in all 20 iteration were considered as overall accuracy for number of training samples k. The graph is drawn for different number of hidden layers.

In figure 1 we are showing graph for hidden layers ranging from 1 to 6.

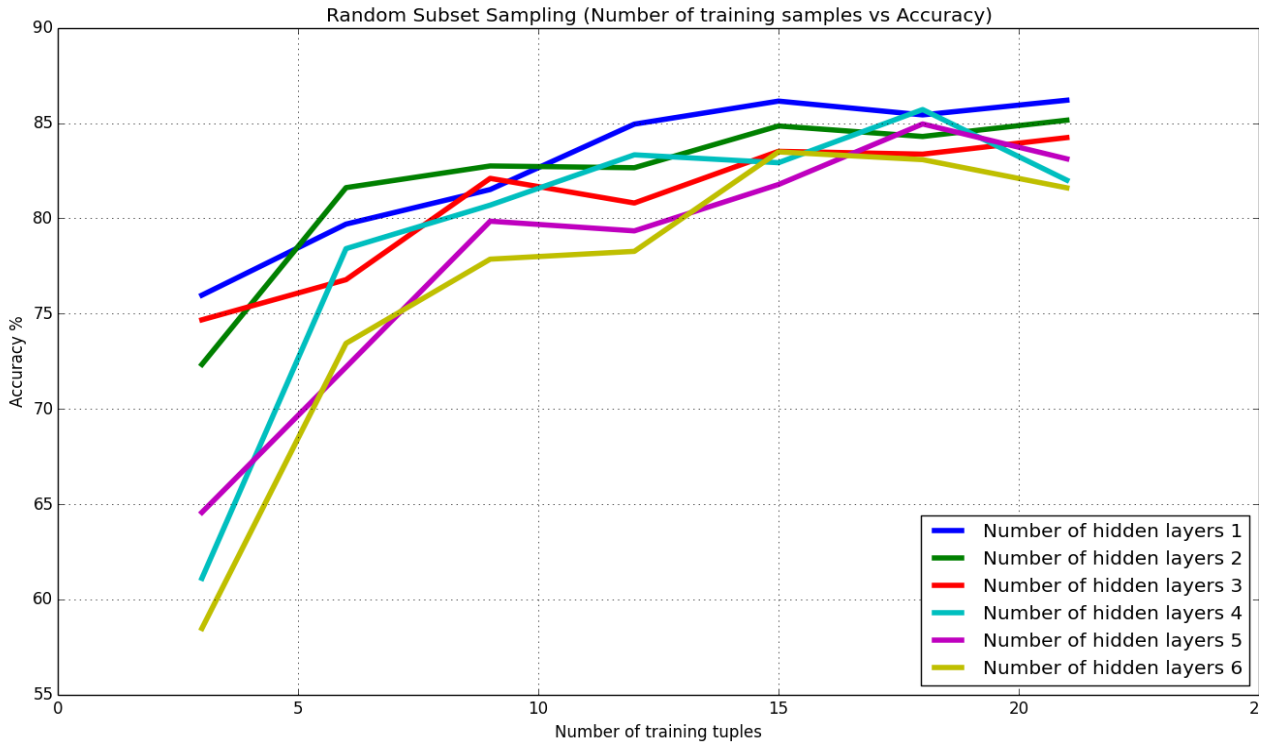


Figure 1 : Number of training samples v/s Accuracy

From the graph it is evident that accuracy of prediction increases as number of training tuples increases. This trend is visible for any number of hidden layers.

5 Conclusions and Future scope

Experimental results shows that it is possible to develop a neural network based system which can predict the personality of individuals using their social media information after proper training. The important thing observed is that as the number of training samples increases the accuracy of system also increases. If we train the system for people living in diverse culture the system can be improved more and more. In our experiment, we were able to get a maximum accuracy of 85%. Which is sufficient enough for the applications like online product promotion, advertisement preferences .. etc.

In our project we are using a lot of features. It is not necessary that all of them are important. We may follow a feature selection method which will help us to select most important features only and it will save a lot of computational time and improve accuracy. Other demographic details such as age of individuals, location can also be used to predict the value of traits. Linguistic feature extraction process may be used for other textual contents like comments, album/photo descriptions ..etc and more linquistic features can be extracted using tools like LIWC.

6 Area of Applications

1. Online marketing system can recommend product in personalized way by analyzing individuals profile in social media

2. Advertisement can be personalized as per personality of individuals

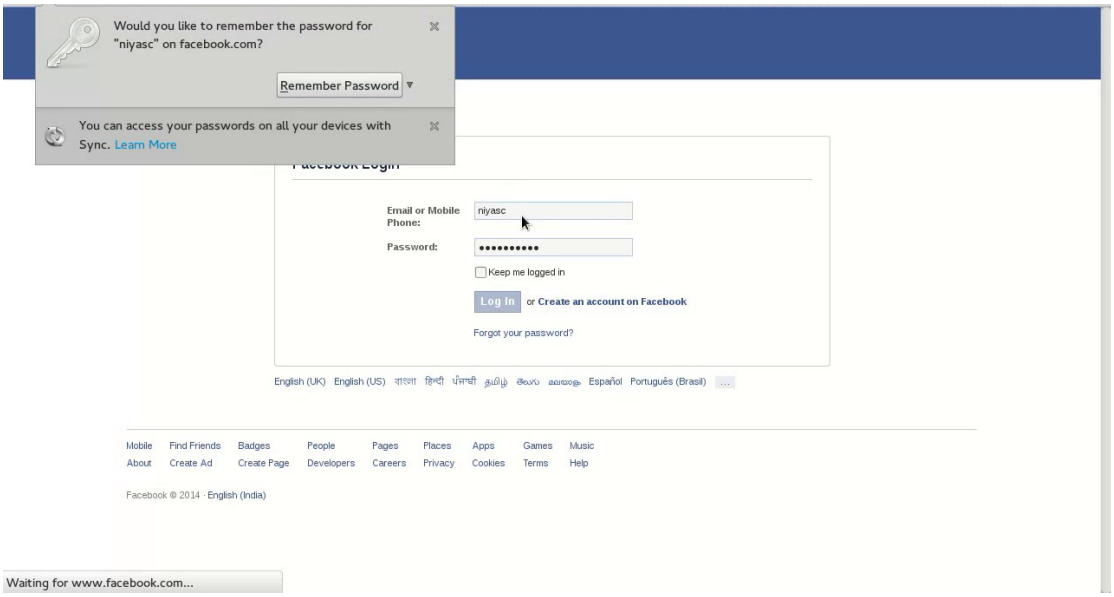
3. Instant feedback : there is no questions at all and we get a score in matter of seconds

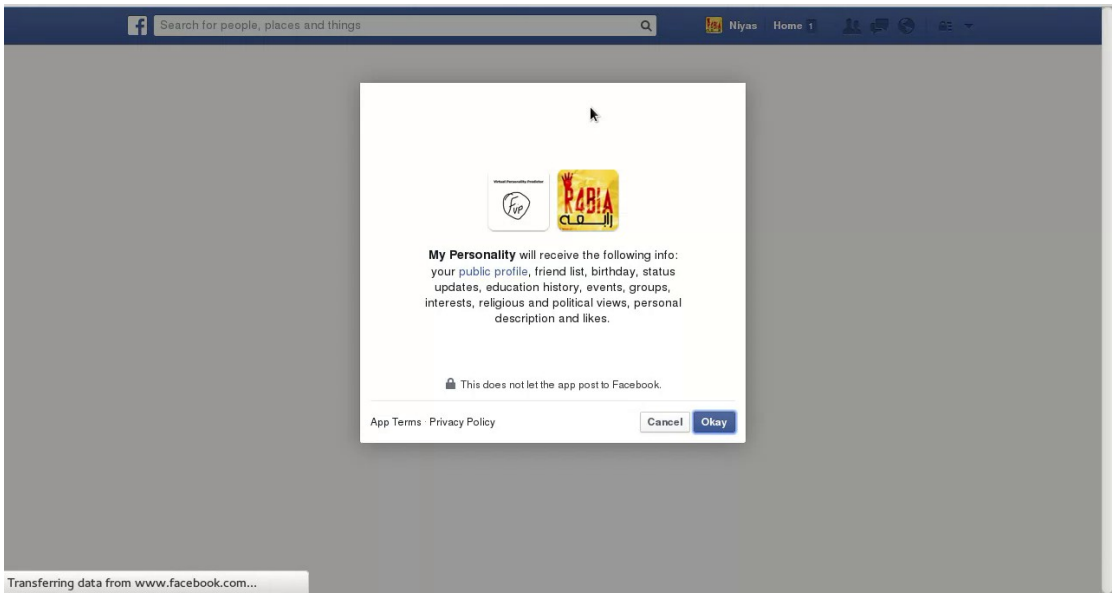
4. Real world validity : It look at individuals actual social media behavior rather than asking them about their real world behavior.

5. Difficult to fake : in traditional testing, people often misrepresent who they really are (even if they do not mean to). By observing actual record of behavior and choices individuals made to 'cheat' the assessment.

7 APPENDIX

7.1 Web app screen shots





2.4808634564 2.44055201671 3.03548665752 4.64765154188 2.0960077817

Personality Test

Niyas Chirayakuth

Traits	Predicted Value
Extraversion	2.48
Agreeableness	2.44
Conscientiousness	3.04
Neuroticism	4.65
Openness	2.1

You may attend Personality Quiz to compare your actual and predicted personality trait values

Take Personality Quiz

You may find significance of above traits here

Personality Test

Welcome Niyas Chirayakuth

S.No	Question (I consider myself as one who .....)	Disagree strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree strongly
1	is talkative?	○	○	○	○	○
2	tends to find fault with others?	○	○	○	○	○
3	does a thorough job?	○	○	○	○	○
4	is depressed, blue?	○	○	○	○	○
5	is original, comes up with new ideas?	○	○	○	○	○
6	is reserved?	○	○	○	○	○
7	is helpful, and unselfish with others?	○	○	○	○	○
8	can be somewhat careless?	○	○	○	○	○
9	is relaxed, handles stress well?	○	○	○	○	○
10	is curious about many different things?	○	○	○	○	○
11	is full of energy?	○	○	○	○	○

submit

Personality Test

Niyas Chirayakuth

Traits	Predicted Value	Actual Value	Error
Extraversion	2.48	1.88	15.15%
Agreeableness	2.44	2.11	8.24%
Conscientiousness	3.04	2.33	17.55%
Neuroticism	4.65	4.25	9.94%
Openness	2.1	2.3	5.1%

Retake Personality Quiz

You may find significance of above traits here

8 References

1. Ensemble methods for personality prediction | Ben, Walter, Tom | AAAI Technical Report|2013
2. Mining facebook data for predictive personality modeling | Denjan, Sonja, Michal, David
3. Predicting personality from twitter | Jennifer, Christina, Michona, Karen |2011 IEEE international conference
4. Machine prediction of personality from facebook profiles | Randal and Taghi(Florida, Atlantic University) , Chris (online privacy foundation)
5. Predicting personality with social behavior | Sibel, Jenniffer | 2012 IEEE/ACM conference
6. Our twitter personality ourselves. Predicting personality with twitter | Daniel, Michal, David, John
7. Personality traits recognition on social network-facebook | Firoj, Evgeny, Stepanov, Giuseppe
8. Towards automatic personality identification using speech acts | Scott, Ericca, Heather, Rudolph, Mappus
9. Personality and patterns of facebook usage | Yoram, Michal, Thore | 2012 acm
10. Recognizing personality traits using facebook status updates | Golnoosh, Susanna, Marie, Martine | 2013 WCPR
11. Too Neurotic, Not Too Friendly: Structured Personality Classification on Textual Data | Fransisco, Aron |2013 WCPR
12. The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. |John, O. P., & Srivastava, S. (1999). | L. A. Pervin & O. P. John (Eds.), Handbook of personality: Theory and research (Vol. 2, pp. 102–138). New York: Guilford Press.
13. Facebook Graph API | https://graph.facebook.com
14. NLP tools for php | http://php-nlp-tools.org
15. Python library pybrain | http://pybrain.org