



PERSONALITY PREDICTION USING MACHINE LEARNING AND DJANGO

Srinivasarao D

School of Computing and Information Technology
REVA University, Bengaluru
sforce2009@gmail.com

G. Parthasarathy

School Of Computing and Information Technology
REVA University, Bengaluru
Parthasarathy.g@reva.edu.in

Abstract: Personality can be defined as a set of characteristics which makes a person unique. The study of personality is of central importance in psychology. There are various conventional ways of assessing one's personality which either costs too much of manual efforts or cannot be done in real time. To solve these problems, this research aims to measure the Big-Five personality from a set of questions. The user is asked to answer a set of few questions and according to the questions answered by the user the personality of the user is predicted using logistic regression model.

Keywords: LR algorithm, Django, Dataset.

I. Introduction

In today's world, the personality of a person is of more importance. For example, for hiring any employee in a company it is important to know the personality of a person according to the job role. They use different strategies to know one's personality. The ways of life have radically changed the ways in which people convey their opinions and feelings. The analysis of social media data or a makes it possible to identify important personality traits, that is, characteristics or qualities which describes his/her personality. These personality reveals only what a person's wants other people to know. Therefore, it is important to know the actual personality of a person. For this reason, we make use of Big five personality test. The Big five personality test includes five different personality traits namely, Openness, Neuroticism, Conscientiousness, Extroversion and Agreeableness. Logistic Regression, a machine learning algorithm is used in this project for the prediction of personality of a person.

II. Literature Survey

Big data is now a popular term used to describe the exponential growth and availability of data, both structured and unstructured data. It varies from conventional cloud services, one of the main features of big data services is coupled between data and computation. This can be conducted only when the related data is available. The social networking sites like Twitter, Face book,

LinkedIn and YouTube allow the users to create and share content related to different subjects, reflect their activities, feelings, thoughts and opinions. This data provides the information about human behavior and social interactions. It makes it possible to understand the user's interests and their needs. This information may be used to survey the consumer's opinions in conceptualizing different business strategy. This paper reviews the techniques used in analyzing social media data to identify important of personality traits. The personality trait is the characteristics and quality of a person. It can be used in different areas such as psychology, marketing and sociology. A parallelism among individual's personality traits and linguistic information are processed for analytics.

Prediction refers to classification of unknown data or to forecast trends. Predicting categorical values is referred to as classification, but if the goal is to model values or continuous functions it is referred to as estimation. Different machine learning prediction techniques are used for mining social media data. Machine learning includes three strategies: supervised, unsupervised or semi-supervised. In to associate personality scores to Twitter users, they gathered data from a Facebook application called my Personality. My Personality users can give their consent to share their personality scores and profile information, and around 40% of them choose to do so. They performed the Big Five personality test on those users. They studied the relationship between the

personality traits of the Big Five Model and five types of micro blog users: listeners and highly read and two types of influence indice Using these, the authors created a correlation table and then performed regression by the M5 Rules algorithm to predict personality of profiles.

III. Objectives

In today's world, the personality of a person is of more importance. For example, for hiring any employee in a company it is important to know the personality of a person according to the job role. This system will help us to find out the mental ability of the user.

IV. Problem Definition (Exiting System)

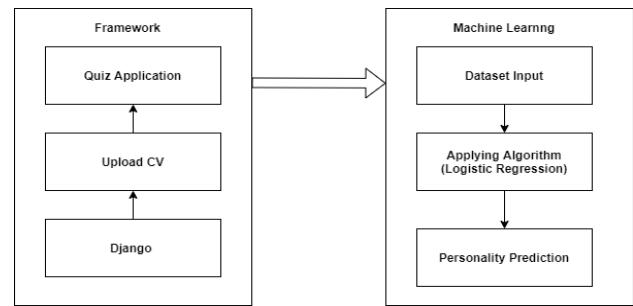
Conventional personality assessment depends on self-report inventory, which costs a lot to collect information. Various models make use of Big five personality test which uses various machine learning techniques. In our system we expect to predict the personality of any person using logistic regression as well as K-Nearest Neighbor algorithm.

V. Proposed System

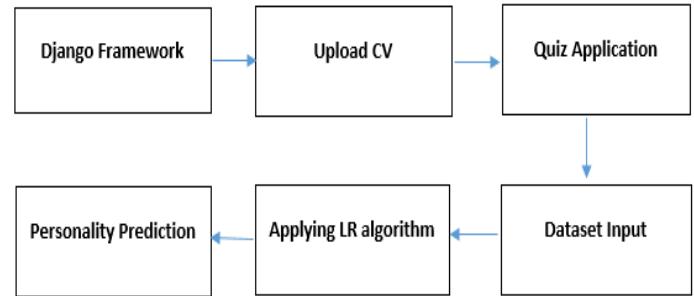
The proposed system aims at using Web framework for personality prediction. Once the Django server starts the CV is to be uploaded. We upload CV to the Django framework. Then the quiz application is given to the user and based on that the performance is checked. We use Logistic Regression to input to the algorithm. The data is preprocessed using Logistic Regression algorithm. Finally prediction of the user personality is done.

VI. System Architecture

The system architecture is shown below for the Frame(Django) and ML(LR algorithm) is applied for the prediction of the user personality. The figure below shows the flow.



VII. Block Diagram



The Django framework is used to upload the CV. User uploads the CV. Then the Quiz application starts for the questions to ask the user. After the answers are given the ML process starts for the prediction. The dataset is imported to the Logistic Regression algorithm. Finally the user who uploaded the CV's personality is predicted.

VIII. Working

We use a web framework for personality prediction. The cv of the user is to be uploaded using Django framework. Then user is given a quiz and based on that the performance is checked. Then a dataset in input to the algorithm. Logistic Regression algorithm is applied to the machine. Finally prediction of the user personality is done.

IX. Django Framework

[Django](#) is a high-level Python Web framework that encourages rapid development and clean pragmatic design. A Web framework is a set of components that provide a standard way to develop websites fast and easily. Django's primary goal is to ease the creation of complex database-driven websites. Some well known sites that use Django include PBS, Instagram, Disqus, Washington Times, Bitbucket and Mozilla.

X. Algorithm

Logistic Regression algorithm: is a supervised classification machine learning algorithm which uses Logit Function to predict the output. Therefore the outcome must be a categorical or discrete value.

It can be either Yes or No, 0 or 1, true or False, etc. But instead of giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1. Logistic regression is used for binary classification(Yes/No, A /B). Therefore, it falls under the classification algorithm. In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1). The curve from the logistic function indicates the likelihood of something such as whether the cells are cancerous or not, a mouse is obese or not based on its weight, etc. Logistic Regression is a significant machine learning algorithm because it has the ability to provide probabilities and classify new data using continuous and discrete datasets.

Table1: Hardware requirements:

I. COMPUTER	II. LAPTOP /DESKTOP
III. RAM	IV. 1 GB
V. ROM	32 B.

Table2: Software requirements:

VI. OPERATING SYSTEM	VII. WINDOWS 10
VIII. TECHNOLOGY	IX. MACHINE LEARNING
Framework	Django
IDLE	Python 3.7 or higher.

X1. Results

The personality prediction starts with the Django server as shown below

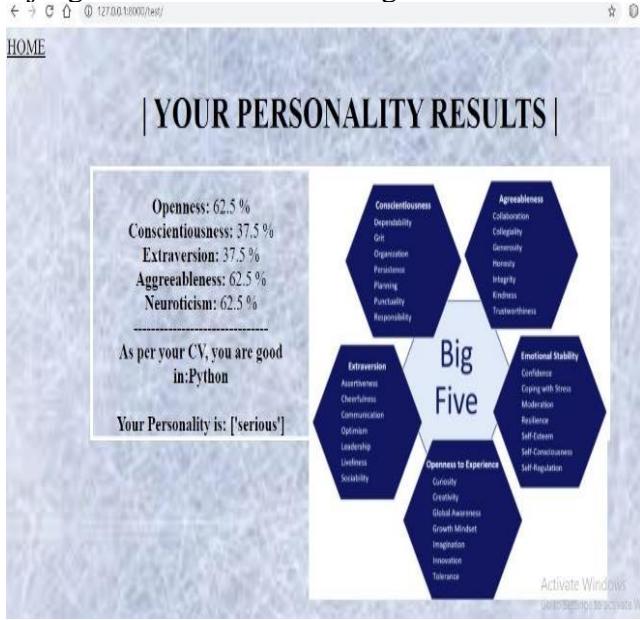


The user details(name age) is entered and CV is uploaded as shown below.

After successful click of the take test then, the user will be asked few questions as shown below.

Finally the user personality is predicted as shown below. It shows the personality type i,e serious. Then the percentage(Openness-62.5%, Neuroticism -62.5%, Conscientiousness-37.5%, Extraversion-37.5% and Agreeableness-62.5%) of Big5 personalities is also displayed. It also shows you that you are good in the which skill i,e Python as per the CV analzed. Hence

these are the output results displayed using Django and Machine Learning.



XII. Conclusion

The project is designed using Big-Five personality from a set of questions. The user is asked to answer a set of few questions and according to the questions answered by the user the personality of the user is predicted using logistic regression model. Hence the model reaches the accurate prediction of the user. Thus the user personality is predicted using ML and Django framework.

XIII. REFERENCES

- [1] J. Staiano, B. Lepri, N. Aharony, F. Pianesi, N. Sebe, and A. Pentland, "Friends don't lie: inferring personality traits from social network structure," in Proceedings of the 2012 ACM conference on ubiquitous computing. ACM, 2012, pp. 321–330.
- [2] J. W. Pennebaker, R. L. Boyd, K. Jordan, and K. Blackburn, "The development and psychometric properties of liwc2015," Tech. Rep., 2015.
- [3] S. Argamon, S. Dhawle, M. Koppel, and J. Pennebaker, "Lexical predictors of personality type," 2005.
- [4] R. R. McCrae and P. T. Costa, "Personality, coping, and coping effectiveness in an adult sample," Journal of personality, vol. 54, no. 2, pp. 385–404, 1986.
- [5] B. R. Karney and T. N. Bradbury, "The longitudinal course of marital quality and stability: A review of theory, methods, and research." Psychological bulletin, vol. 118, no. 1, p. 3, 1995.
- [6] Ashwinkumar.U.M and Dr. Anandakumar K.R, "Predicting Early Detection of cardiac and Diabetes symptoms using Data mining techniques", International conference on computer Design and Engineering, vol.49, 2012.
- [7] P. T. Costa and R. R. McCrae, "Normal personality assessment in clinical practice: The neo personality inventory." Psychological assessment, vol. 4, no. 1, p. 5, 1992.
- [8] S. Adali and J. Golbeck, "Predicting personality with social behavior," in Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM International Conference on. IEEE, 2012, pp. 302–309