International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 9, Issue 1 - February 2021 - Pages 26-30

Personality Classification from Online Text Using Machine Learning

Jisna Jenson, Preeshma George, Reeny Zackarias, Dr. T John Peter
Dept. of Computer Science & Engineering, IES College of Engineering, Chittilappilly,Kerala,India.

Abstract
The corporate world today focus on the skills a potential employee possesses and also on their personality. This system will focus on qualification, inexperience and also on other important aspects. The major objective of the system is to select the right candidate for the desired job profile. Here we propose a way in which the process of short listing gets streamlined and faster by personality prediction. This project examines different machine learning approaches for efficiently predicting personality through CV analysis and it uses KNN algorithm. The proposed system can be used in many organizations which may require expert candidates. This system will reduce the workload of HR Department of an organization. This system will helps the hiring authorities to select the right candidate for the desired job profile. Admin can easily select the appropriate candidate for a particular job profile. This system will focus on qualification, inexperience and also focuses on other important aspects, which are needed/ demanded for a particular job position. Admin can store the data for further comparison and sorting of data.

Keywords—Natural Language Processing, KNN,

INTRODUCTION

The personality of a person is very important in life. During the recruitment process, the personality of a candidate is tested in HR or final rounds, at that stage the HR won't be clear about the candidate. He will be confused to asking questions about the candidate. Nowadays, all companies want the best candidate for their profile. The personality Prediction system that we proposed can be used to identify the personality of a person and it will show which job is best suitable for the person. The system enables an effective way of short listing the candidates from a large number of applicants. The recruiter judge a candidate by only his aptitude and technical knowledge and this system will test the person's aptitude, and technical level as well as the personality of a person. Machine Learning is used to analyze the data given by the candidate during interview process. After analysis, the results will be shown to the candidate and the recruiter, which will help both interviewer and candidate. By knowing about the candidate's interest which will be easy for HR to ask questions in the interview and select the right candidate for a particular job profile which in turn provides perfect candidate for the organization.

Existing System:

In the existing system, the employer chooses appropriate candidate only by CV analysis and face to face interview. In this method they focused on only qualification, based on qualification people were selected which led to low performance as compared to other people who were technically, logically and emotionally sound. The drawback of this system are it requires large memory space as it stores data related to CV's and may provide incorrect data due error in analysis.

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 9, Issue 1 - February 2021 - Pages 26-30

Proposed system

The personality Prediction system that we proposed can be used to identify the personality of a person and it will show which job is best suitable for the person. The system enables an effective way of short listing the candidates from a large number of applicants. The recruiter judge a candidate by only his aptitude and technical knowledge and this system will test the person's aptitude, and technical level as well as the personality of a person. Machine Learning is used to analyze the data given by the candidate during interview process. After analysis, the results will be shown to the candidate and the recruiter, which will help both interviewer and candidate. By knowing about the candidate's interest which will be easy for HR to ask questions in the interview and select the right candidate for a particular job profile which in turn provides perfect candidate for the organization.

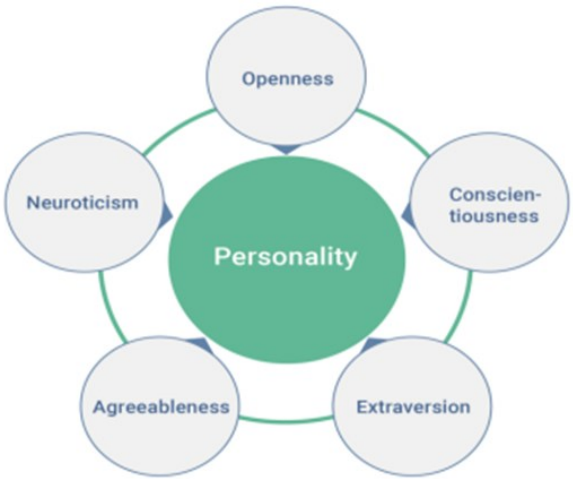
.

Background of Personality Perception

The Big Five Personality Traits model is based on findings from several independent researchers, and it dates back to the late 1950s. But the model as we know it now began to take shape in the 1990s. Lewis Goldberg, a researcher at the Oregon Research Institute, is credited with naming the model "The Big Five." It is now considered to be an accurate and respected personality scale, which is routinely used by businesses and in psychological research. The Big Five Personality Traits Model measures five key dimensions of people's personalities:

Openness: sometimes called "Intellect" or "Imagination," this measures your level of creativity, and your desire for knowledge and new experiences.

Conscientiousness: this looks at the level of care that you take in your life and work. If you score highly in conscientiousness, you'll likely be organized and thorough, and know how to make plans and follow them through. If you score low, you'll likely be lax and disorganized. Extraversion/Introversion: this dimension measures your level of sociability. Are you outgoing or quiet, for instance? Do you draw energy from a crowd, or do you find it difficult to work and communicate with other people?



Agreeableness: this dimension measures how well you get on with other people. Are you considerate, helpful and willing to compromise? Or do you tend to put your needs before others'?

Natural Reactions: sometimes called "Emotional Stability" or "Neuroticism," these measure emotional reactions. Do you react negatively or calmly to bad news? Do you worry obsessively about small details, or are you relaxed in stressful situations?

System Requirements & Technologies Used:

For the project we are using different methods for analyzing behavior of peoples,

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 9, Issue 1 - February 2021 - Pages 26-30

HARDWARE	SOFTWARE
● Desk top /Lap top	Pandas
● OS - Android 8.0 or more	NumPy
● Ram Capacity : 4GB	Tkinter
● Memory - 120 MB	Functools
	Scikit-learn
	Pyresparser-OMKAR PATHAK

Module Description:

The system built in this project predicts the personality of people by using their gender, age, the score of openness, conscientiousness, extraversion, agreeableness, neuroticism, and experience. It parses all the data from the CV/resume and on the result page, it shows all the information from the entered data and uploaded resume. This system uses logistic regression for training the model and pyresparser module for parsing the information from resume which is built using NLTK and spaCy module in python.

Description of Methods and Flow in the System:

train_model class: It contains two methods that train the model and predict the result by giving the various values.

a. train method: It read the dataset for training the model from a CSV file and builds a model using Logistic Regression. It uses different 7 values for training the model

b. test method: It predicts the personality of a person by passing an array of values that contains gender, age, and other 5 personality characteristics.

Main method: We start with creating an object of the train_model class and train the model by calling the training method of the class. Then we initialize a variable with a Tk object and design the landing page of the system using labels and buttons. A button with the name Predict Personality is designed which calls the predict_person method.

predict_person method: We withdraw the root Tkinter window and create a new top-level window and configure its size and attributes. We label the heading of the window followed by various labels and their entries. For selling a resume file, use the r needs to press choose file button which then calls the Open file method that takes an argument of the button. In the predict_person method, various entries are taken for predicting the personality. Submit button passes all the values to prediction_result.

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 9, Issue 1 - February 2021 - Pages 26-30

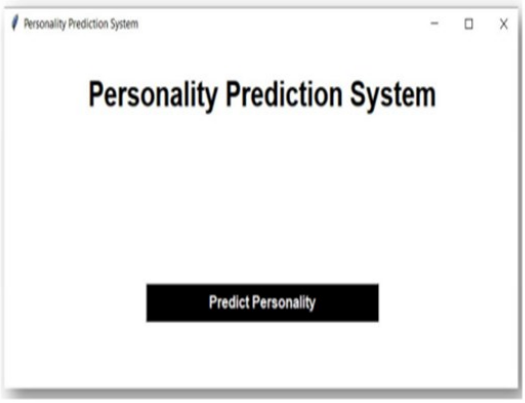
prediction_result method: This method firstly closes the previous Tkinter window which was used to take the data from the user. After this, it calls the test method of the model object and stores the result returned by the method. After this , it parse all the information from the resume and stores it in a variable followed by a try-except block which try to delete name and validate mobile number from fetched information from resume. Then it prints all the data submitted by user on console. After this, the method popup a full screen window which shows all the parsed information and predicted personality on GUI window along with the definition of each personality characteristic's definition.

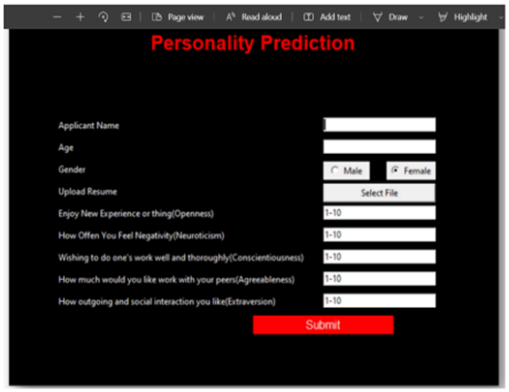
check_type method: It converts various strings and numbers into desired format and converts lists and tuples in string.

Results and Discussions

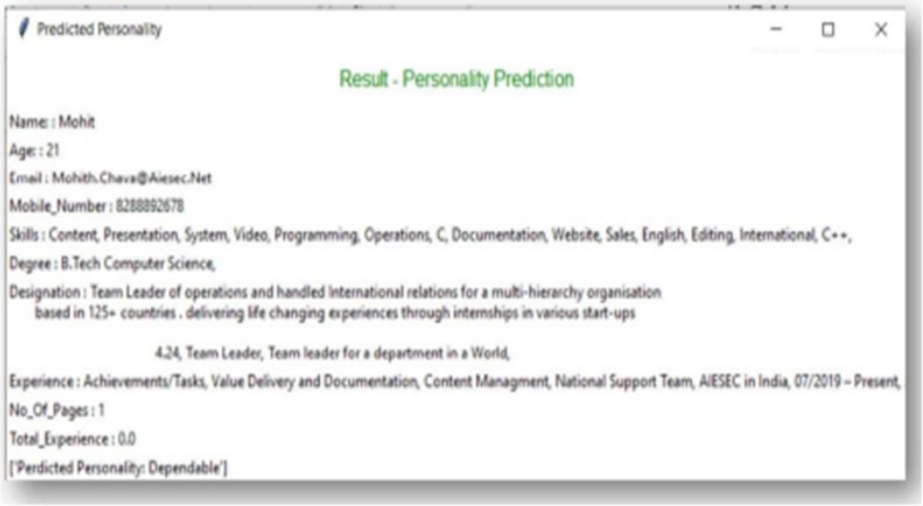
On the landing page, the 'Predict Personality' button pops up a new window for taking various inputs from users and submitting it prediction model which will predict the personality

APPLICATION INTERFACE

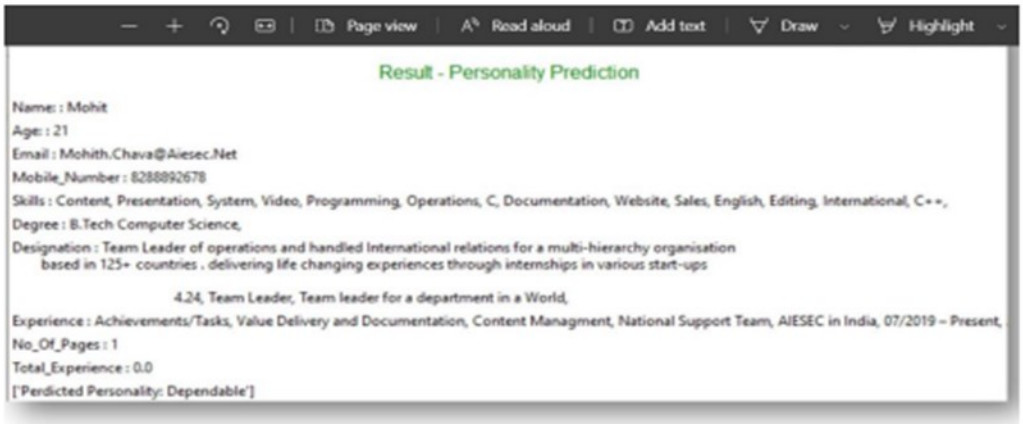




On the result page, all the manipulated information and predicted result will be displayed



International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 9, Issue 1 - February 2021 - Pages 26-30



8.Conclusion

We have presented in this paper, the prediction of human personality by using the KNN algorithm and logistic regression.The CV is shortlisted for the recruitment process and a fair and appropriate decision with some of questions is made by the HR department. Interactive and easy to use. Extract all the important features of a resume in seconds. It can be extended for commercial uses. In future we can increase the efficiency and performance of the proposed system in order to predict personality for recruitment analysis more perfectly. It can improve the training model for various addition features that help us to predict more accurate result. Instead of direct asking the five characteristic values we can add questionaries' which ask some multiple-choice questions and auto calculate the various values

9.References

1. A Demetriou, L. Kyriakides, and C. Avraamidou, The missing link in the relations between intelligence and personality in Journal of Research in Personality, vol. 37 issue 6,December 2003, pp 547- 581.
2. M. Kalghatgi, M Ramannavar, and Dr. N. S. Sidnal, Neural Network approach to personality prediction based on the Big-Five Model in IJIRAE, vol2 issue 8, August 2015, pp 56-63.
3. A..Robey, K. Shukla, K. Agarwal, K. Joshi, Professor S. Joshi Personality prediction system through CV Analysis, in IRJET vol 6, issue 02, February2019.
4. J. Zubeda, M. Shaheen, G. Narsayya Godavari, and S. Naseem Resume Ranking using NLP and Machine Learning,
5. Md Tanzim Reza, and Md. Sakib Zaman, Analyzing CV/Resume using natural language processing and Machine learning,