International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 11 | Nov 2019 www.irjet.net p-ISSN: 2395-0072

Personality Prediction System using AI

Dhanashree Sonaje[1], Shrutika Alladwar[2], Neha Bhise[3], Vishakha Dhote[4], Ketan Desale[5]

[1]Dhanashree Sonaje, Dept. of Computer Engineering, PCCOE, Maharashtra, India
[2]Shrutika Alladwar, Dept. of Computer Engineering, PCCOE, Maharashtra, India
[3]Neha Bhise, Dept. of Computer Engineering, PCCOE, Maharashtra, India
[4]Vishakha Dhote, Dept. of Computer Engineering, PCCOE, Maharashtra, India
[5]Professor Ketan Desale, Dept. of Computer Engineering, PCCOE, Maharashtra, India

---

Abstract - *In this system, to make the recruitment process more effective and efficient, we are building the personality prediction model. In this system, we are going to implement the model which is based on CV, aptitude and personality test. The concept behind this model is to shortlist the candidates according to the admin requirement. So that the workload and time required for selection of the best CV will be reduced. Firstly, the candidate will be shortlisted on the basis of CV and then aptitude and personality test link will be given to him. According to the score and HR requirement candidate will be selected. For implementing this system, Random Forest Algorithm, Support Vector Machine, Weighted Majority Voting Algorithms are studied.*

Key Words: personality prediction, random forest algorithm, SVM, Weighted Majority Voting Algorithm aptitude, recruitment, big five personality traits

1. INTRODUCTION

Personality prediction can be used in employee recruitment process. For personality prediction, there are various ways to find out like, self-descriptive report, by taking psychometric tests, etc. But these techniques are not always reliable to predict personality. Personality can be predicted by matching the most frequent used words. This proposed system consists of two modules mainly candidate module and HR module. In candidate module, candidate has to register firstly on the company's website. In this registration, candidate has to complete his profile first, then upload his CV. In HR module, HR will add the job description, shortlist the top candidates according to uploaded CV, add the aptitude and personality questions, provide links to the shortlisted candidates.

1.1 Project Objectives

1. To shortlist the candidate through CV.
2. To take the online test (CV and aptitude test).
3. To rank the candidate according to CV and online test.

1.2 Problem Statement

To enable a more effective way to short list submitted candidate CVs from a large number of applicants providing a consistent and fair CV ranking policy, which can be legally justified. System will rank the experience and key skills required for particular job position. Then system will rank the CV's based on the experience and other key skills which are required for particular job profile. This system will help the HR department to easily shortlist the candidate based on the CV ranking policy.

2. LITERATURE SURVEY

2.1 Big Five Personality traits:

Every human being is unique , but this uniqueness originates from the combination of 5 personality traits we all share. From a organizational perspective, information about how an individual personality can provide valuable information about that person, what task and job is he suitable for. The big five model is most widely used model and more recognized model today. A brief description of big five traits is provided below.

1. Openness - these are people who like to learn some new things and enjoy doing new things score more in openness. Openness includes traits such as insightful and imaginative, deep understanding.

2. Conscientiousness – people who do ones work or duty thoroughly, person who do his job seriously and they are achievement oriented.

3. Extraversion- Extravert are people who get the energy by interacting with other people, they engage themselves with external environment. Traits include more energetic, chatty, and confident.

4. Agreeableness – a person here is friendly and interactive, more cooperative, these people have optimistic view of human nature. Traits include caring, cooperative, and attentive.

5. Neuroticism - People who score more on neuroticism have negative emotions and they are emotionally .Traits include anger, fear, envy, jealousy, and depressed mood.

There are many traits which have been defined and studied for these five one, which are mentioned in above description. Similarly, some other have identified a set of facts for each of the big five personality traits based on different data. Some claims about the universality of the big five model traits, some have criticized the lack of conceptual validation.

However, other studies have shown the model is good, and comprehensively subsumes nearly all English traits. Accordingly, the big five model was used to analyze individual personality in this study.

2.2 Random forest Algorithm

Random forest can be used as classification as well as regression tasks. Random forest algorithm is the combination of "Bootstrap aggregating method" and "random subspace method" to build various decision trees [3]. Random forest is an ensemble classifier based on large number of decision trees. Each decision tree gives the individual outcome by analyzing the conditions. The class having the more votes consider as a model's prediction. It is a supervised machine learning algorithm.

To check the generalized error in random forests, an upper bound is derived. This error is consisting of two parameters mainly,

1) The accuracy of individual classifiers

2) The dependency between the individual classifiers[2]

Random Forest Algorithm for online Streaming data:

1. Steaming Random Forest Algorithm:
   It is a stream classification algorithm. This algorithm uses Gini index for splitting. It is based on two attributes:
   1) Number of trees to be built
   2) Tree window
2. Dynamic Steaming Random Forest Algorithm:
   It is a self-adjusting steam classification algorithm. As the concept changes, it reflects. Difference between the Steaming Random Forest Algorithm and Dynamic Steaming Random Forest Algorithm is the value of tree window. The size of tree window is not constant for all the trees.
3. Online Bagging and Boosting:
   It guarantees for strong experimental results. Previous technique requires the whole training data set available at once. So, to overcome this issue online bagging and boosting were proposed by Oza and Russell. When the data is arriving continuously, it is used[2].

2.3 Support Vector Machine

Support Vector Machine (SVM) is a type of supervised learning algorithm. It analyses the data and recognizes pattern for classification. SVM is based on a linear division. For making the point linearly separable, SVM is used to map the low dimensional space point to high dimensional space. It takes the set of training data and marking it as part of a category and predicts whether the test document is a member of an existing class. Support Vector Machine is a linear model for classification and regression problems. It can solve linear and non-linear problems and work well for many practical problems.

The idea of SVM is simple: The algorithm creates a line or a hyperplane which separates the data into classes.

2.4 Weighted Majority Voting Algorithm

Weighted Majority Voting Algorithm is a Meta learning algorithm which is being used for constructing a compound algorithm from n-number of prediction algorithm which may be the type of learning algorithm. Weighted Majority Voting checks for the shares i.e. the person having more number of shares will have more number of votes.

In Weighted Majority Voting there are few terms they are:- Players, Quota, Dictator, Veto Power, Dummy.

- Players: - Each individual person is called player (eg. P1,P2,P3,…,Pn) and every player has its own weight. This weight tells us how many vote does a player has.
- Quota:- It checks for minimum weight among all players which is needed for vote to pass need. (eg. Quota(Q):w1,w2,w3,…,wn)
- Dictator:- Dictator is the one who can block any of the proposal from getting passed. The player who is having weight greater than or equal to the quota.
- Veto Power: - If any of the player's weight is much more necessary for the quota to be reached then that player has Veto Power. More than one player can have Veto Power. (eg.[30:19,15,11] here 19 has veto power)
- Dummy: - If any of the player's vote is never being essential for a group to reach the quota then that is dummy.

3. SOLUTION

3.1 Mathematical Model

Let S be a system which predicts the personality from the CV and tests and accordingly gives rank to the candidate. It is defined as,

S={ I, Cnst, O}

where,

I = CV which is uploaded to the system and Aptitude test

Cnst = constraints specified by the HR

O = rank of the candidate

Detail flow of the system:

1. Upload CV while registration. Candidate will give the aptitude test and personality test.

$I = \{ D_t, D_{t+1}, D_{t+2}, ... D_{t+n} \}$

where,

$Dt = \{ X_0, X_1...X_n \}$

2. Apply the constraints on the CV and check the answers of the test.

   $Cnst = \{ C, A \}$

   where,

   C = selection criteria according to CV specification

   $C = \{ C_1, C_2...C_n \}$

   A = answers of the test

   $A = \{ A_1, A_2...A_n \}$

3. O = Output of the system

   where, $O = \{ O_1, O_2, ...O_n \}$

   On the basis of this rank candidate will be shortlisted according to company policy.

3.2 Working

3.2.1 Admin Module:

This module consists of add aptitude test questions, add personality test questions, view details of all selected questions and corresponding job detail, job search will give the information about job requirements and specifications.

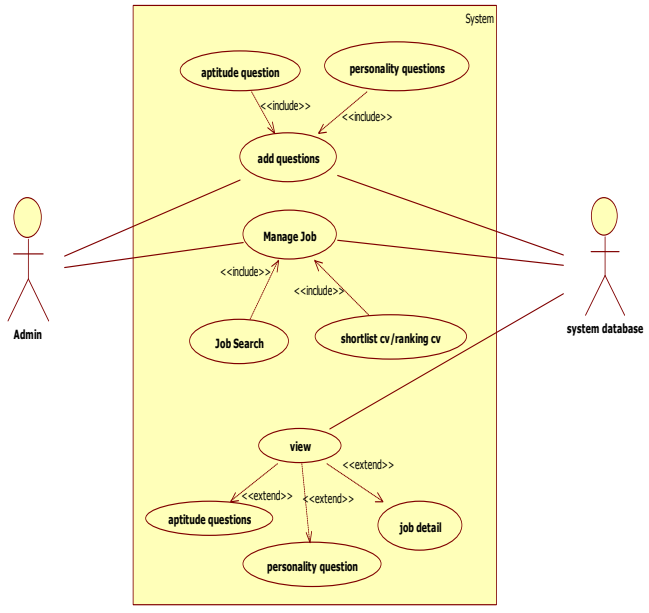

Fig.1-Use case diagram: Admin Module

3.2.2 Candidate Module:

This module consists of create profile and upload CV. If CV is shortlisted then only the candidate will be called in office and test link will be given to him.

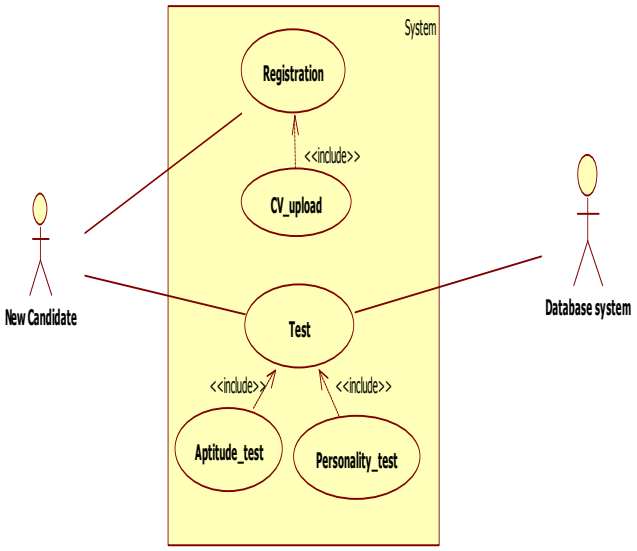

Fig.2-Use case diagram: Candidate Module

4. CONCLUSION

In this proposed system, we have implemented system which is useful for any recruiter for recruitment process where the applied candidates are in a huge number. This system will reduce the workload of recruiter or Human Resource Department.

5. REFERENCES

[1] Allan Robey, Kaushik Shukla, Kashish Agarwal, Keval Joshi, Shalmali Joshi, "Personality Prediction System through CV Analysis", Volume: 06 Issue: 02, 2019

[2] Eesha Goel, Er. Abhilasha,"Random Forest: A Review", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 7, Issue 1, 2017

[3] Qiong Ren, Hui Cheng, Hai Han "Research on Machine Learning Framework Based on Random Forest Algorithm", AIP Conference Proceedings, 2017

[4] Amir Saffari, Christian Leistner, Jakob Santner, Martin Godec, Horst Bischof, "On-line Random Forests", IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, 2009

[5] Harsh Valecha "Prediction of Consumer Behaviour using Random Forest Algorithm", 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), 2018

[6] Jennifer Golbeck, Cristina Robles, Karen Turner, "Predicting Personality with Social Media", *CHI 2011*, May 7–12, 2011, Vancouver, BC, Canada. ACM 978-1-4503-0268-5/11/05.

[7] B.Gayathri, V.Kavitha, "Personality Prediction Based on Big Data",International Journal of Research and Advanced Trends in Engineering & Science, Sep 2018.

[8] Niels van de Ven and Aniek Bogaert, "Personality Perception based on LinkedIn profiles", Emerald Publication Limited, 2017.

[9] Bayu Yudha Pratama, Riyanarto Sarno, "Personality Classification Based on Twitter Text Using Naïve Bayes, KNN and SVM", 2015 International Conference on Data and Software Engineering, 2015.