**RK18XCB40 : Aman Kumar(11814305)**

**RK18XCB41 : Aditya Kumar(11814112)**

**RK18XCB38 : Rahul Nainala(11813397)**

**GitHub Link : https://github.com/kr-aman24/prsonality\_prediction**

**Personality Prediction System**

**# Simple demostration of using data mining concepts and python data science libraries to analyse and classify personalities of a given set of people or an individual.**

**## Abstract**

The project uses learning algorithms and advanced data mining concepts to mine user characteristics data and learn from the patterns.

This project will come across areas where it has access to large amounts of person behavioral data. This data can be helpful to classify persons using Automated personality prediction and classification. There are areas where there is access to large amounts of person behavioral data. This data can help us classify persons using automated personality classification.

**## Approach**

Five characteristics of different individuals commonly known as big five characteristics namely, openness, neuroticism, conscientiousness, agreeableness and extraversion are stored in a dataset and used for training. Based on this training, the personality of individuals are predicted using data mining concepts. Before testing the dataset, it is pre-processed using different data mining concepts like handling missing values, data discretization, normalisation etc.This pre-processed data can then be used to classify/predict user personality based on past classifications. The system analyses user characteristics and behaviors. System then predicts new user personality based on personality data stored by classification of previous user data.

Model used to predict test dataset is “Logistic Regression” because Logistic regression is an effective model to predict output class labels for dependent categorical data.

**## Dataset description**

Attribute Description:

No. of attributes are 7 as listed below.

**|S.NO | ATTRIBUTE | TYPE |RANGE|**

| --- | --- | --- | --- |

|1 |Gender |nominal| Male / Female|

|2 |Age | numeric| 17-28|

|3 |Openness | numeric| 1-8|

|4 |Neuroticism |numeric| 1-8|

|5 |Conscientiousness |numeric |1-8|

|6 |Agreeableness |numeric |1-8|

|7 |extraversion | numeric |1-8|

Class label description:

No. of class labels: 5

Type: Nominal

**Values:**

**● Extraverted**

**● Serious**

**● Responsible**

**● Lively**

**● dependable**

**T**his will 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. Than 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. This system will focus not only in qualification and in experience but also focuses on other important aspects, which are required for particular job position. This system will help the human resource department to select right candidate for particular job profile, which in turn provide expert workforce for the organization.

Candidate here will register him/herself with all its details and will upload their own CV into the system, which will be further used by the system to shortlist their CV. Candidate can also give an online test, which will be conducted on personality questions as well as aptitude questions. After completing the online test, candidate can view their own test results in graphical representation with marks.

**Features:**

* This system will automatically determine the key skill characteristic by defining each expert's preferences and ranking decisions.
* The presented system automates the processes of requirements specification and applicant's ranking.
* The proposed system produces ranking decisions that were relatively highly consistent with those of the human experts.
* This system will 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.

**Neuro Fuzzy Logic Explanation**

* We use neural network
* System will assign weight age for each requirement
* Resume will be shortlisted based on overall weight age

**Modules:**

The system comprises of 2 major modules with their sub-modules as follows:

**Advantages:**

* This system can be used in many business sectors that may require expert candidate.
* This system will reduce workload of the human resource department.
* This system will help the human resource department to select right candidate for particular job profile, which in turn provide expert workforce for the organization.
* Admin or the concern person can easily shortlist a candidate based on their online test marks and can select an appropriate candidate for desired job profile.

**Disadvantage:**

* This system requires large memory space as it stores data related to CV’s.
* Requires an active internet connection.
* May provide inaccurate results if data not entered properly.

**Application:**

* This system can be used in many organizations in order to shortlist expert candidate.