

Personality-prediction-system

Simple demonstration 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