International Engineering Journal For Research & Development
Vol.5
Issue 5

E-ISSN NO:-2349-0721

Impact factor : 6.549

AUTOMATED PERSONALITY CLASSIFICATION BASED ON DATA MINING TECHNIQUES

[1]Mr. Mandar S Joshi, [2]Ms.Saniya F Fadnaik, [3]Ms. Aishwarya S.Shetye, [4]Ms.Jyoti S.Nachankar
Department of Information Technology Hope Foundation's Finolex Academy of Management and Technology Ratnagiri, India[1, 2, 3, 4]
mandar.joshi@famt.ac.in[1] , Saniyafadnaik97@gmail.com[2], Shetyeaishwarya21@gmail.com[3], Jyotinachankar97@gmail.com[4]

ABSTRACT

Understanding character type can assist you with understanding preferences and the preferences of others. Character types are helpful for perceiving how individuals lead, impact, convey, work together, arrange business and oversee pressure. This proposed system is useful where we have information identified with individual behavior. This individual behavior data can be valuable for differentiating individual dependent on his character attributes. This behavior qualities will be then put away in database. Later when client enters his character attributes his character is analyzed in database and framework will recognize the character of client. This proposed system will utilize Naïve Bayes algorithm. The system is useful for association or organizations to recruit their employees by observing behavior of workers. It can also be used in different e-commerce business to distinguish the client behavior.

INTRODUCTION

Character is one kind of combination of patterns that impacts conduct, Motivation, thoughts and emption in person. The personality of an individual can be characterized as a set of character that actuates a tendency on the individual behavior; this tendency is steady through time and circumstances. This framework will distinguish the personality of a given individual which will give hints about how he would presumably respond when confronting various circumstances. Distinguishing a user character can help to know, for instance, his latent capacity needs in various contexts. Subsequently, this automated personality classification System may profit by having models of client's character to adjust their conduct as needed. There is a wide variety of areas wherein this can be helpful, i.e., assistive technologies, e-learning, e-commerce business, social insurance or recommender systems, among others. Techniques like review, interview are very tedious and as individuals are prone to biases which influence the exactness and can't give us the results as right outcome as this personality classification system.

LITERATURE SURVEY

1. Understanding Personality through Social Media

The relationship between human language on Twitter and personality traits is studied. Specifically, we want to know how linguistic features correlate with each personality trait and to what extent can we predict personality traits from language. Here they gather personality data from Myers-Briggs personality test which

contains thinking, feeling, sensation, intuition, introversion, extroversion, judging and perceiving. Also, they collect 200 most recent tweets from users with personality values. This system design three categories of feature, namely bag of n-grams, Twitter POS tags, and word vectors. Analysis of these features provide insights of language use for different personalities. [3]

2. Predicting personality types from user comments

This project explores the significance of personality types in determining the use of language, particularly online comments in forums. The machine learning task is to predict personality type of a user based on comments on an online forum. The project has far-reaching implications if you consider the ever-increasing availability of user data on online forums and social media platforms. The words that the feature selection ranked as significant are greatly insightful to the nature of what it means to be of a "Thinking (T)" or "Feeling (F)" type. It is amazing to see that the following words, which are less explicitly related to the "feeling" dichotomy but still one of its characteristics, are predictors of the dichotomy: love, beautiful, really, heart, thank, hope, haha and song. For the "thinking" dichotomy, the results are less interesting because the difference in the frequencies for each output class are not as significant.

3.Mining Facebook Data for Predictive Personality Modeling

Initial investigation in personality modelling based on Facebook data are encouraging evidence that by selecting the most indicative features the precision of the classifiers could be improved. Extracting qualitative knowledge from the large quantities of data is just the beginning of our search for meaning and plausible explanation of personality-determined social network activities. This paper explores the feasibility of modeling user personality based on a proposed set of features extracted from the Facebook data. [2]

4.Personality Classification Based on Twitter Text

E-ISSN NO:2349-0721

Social media is a place where users express themselves to the world. Posts made by users of social media can be analyzed to obtain their personal information. This experiment uses text classification to predict personality based on text written by Twitter users. The languages used are English and Indonesian. Classification methods implemented are Naive Bayes, K-Nearest Neighbors and Support Vector Machine. Testing results showed Naive Bayes slightly outperformed the other methods. [1] EXISTING SYSTEM

In Existing System, Facebook or Instagram profiles, Twitter language or content is utilized to distinguish/judge the character of client which may not give the right result. Making a decision about the personality via web-based networking media content can't give us the exact outcome. Anyway the individual can be misjudged or his real character can be overlooked. Techniques like review, meet are very tedious and as individuals are inclined to biases which influence the exactness and can't furnish us with precise outcome. While taking meetings the questioner might not have accurately understood a word or expression you utilized in view of your elocution, your pronunciation, or inflection—which may have driven them to (generally adversely) confuse your importance. Along these lines, such ways brings about misconception in the personality of individuals.

PROPOSED SYSTEM

Personality is the result of social collaboration in bunch life. Individuals have various kinds of characters since people are not the same. It alludes to the propensities, mentalities just as physical attributes of an individual which are not same but rather have fluctuate from gathering to gathering and society to society, everybody has character, which might be positive or negative, noteworthy or unremarkable. Character characterization is dictated by the examples of reasoning and conduct that create after some time. Five attributes of various people usually known as large five qualities specifically, receptiveness, neuroticism, scruples, appropriateness and extraversion are put away in a dataset and utilized for preparing. In light of this preparation, the character of people are anticipated utilizing information mining ideas. The character qualities will be put away in database. Afterward, when client enter his character qualities his character is inspected in enormous prior database and framework will identify the character of client. The framework will distinguish clients character dependent on past information put away in database. This framework will analyze character of client dependent on character characteristics referenced by the client. What's more, will furnish client with different highlights applicable to his character. The connection between the character and client conduct is tried. This framework will assist notice with peopling to showcase their item dependent on the character of the client which thusly gives pay to the firm who is utilizing this framework. This framework can be implanted with social destinations, the greatest number of client can purchase and sell their items utilizing these interpersonal organizations.

BACKGROUND OF DATA MINING SYSTEMS

Naive Bayes Classifier.

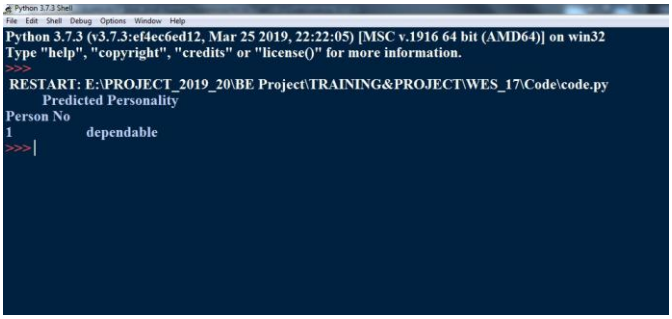
Naive bayes classifier is based on Bayes Theorem. Naive Bayes is a collection of classification algorithm.Bayes' Theorem Bayes' Theorem finds the probability of an event occurring given the probability of another event that has already occurred.

Bayes' theorem is stated mathematically as the following equation:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

The below image shows the questions that will classify the person based on different personality .In below image ,every personality type has four question related to it:

The output will be shown as:



We maintain an access database for every person uses the system so that it can be used in future scope:

	A	B	C	D	E	F	G	H
1	Gender	Age	openness	neuroticism	conscientiousne	agreeableness	extraversion	Personality (Class label)
2	Male	17	7	4	7	3	2	extraverted
3	Male	19	4	5	4	6	6	serious
4	Female	18	7	6	4	5	5	dependable
5	Female	22	5	6	7	4	3	extraverted
6	Female	19	7	4	6	5	4	lively
7	Male	18	5	7	7	6	4	lively
8	Female	17	5	6	5	7	4	extraverted
9	Female	19	6	6	7	5	4	extraverted
10	Male	18	5	7	5	6	7	dependable
11	Female	19	5	5	7	4	5	lively
12	Male	19	6	7	5	6	3	serious
13	Male	19	7	6	7	7	6	extraverted
14	Male	19	7	6	6	5	6	lively
15	Female	19	6	7	5	5	5	dependable
16	Female	19	5	5	4	5	4	responsible
17	Male	19	5	6	4	6	3	extraverted
18	Female	19	7	7	2	6	5	serious
19	Female	18	6	7	4	4	2	dependable
20	Female	19	6	6	6	4	3	responsible

CONCLUSION

This paper shows the investigation of connection among client and his character. This character arrangement framework is exceptionally useful in other informal organizations where character plays a crucial repetition like wedding destinations, E-trade locales. The methods utilized for location of character is AI based methodology like innocent Bayes and SVM. The accessible character information dataset is talked about. Along these lines the character of individual is consequently ordered by framework.

Further improvement should be possible by utilizing progressively exact dataset to improve the precision additionally this can be useful for vocation direction module where, if client has great talking and persuading abilities he can approach towards promoting.

REFERENCES

1. BAYU YUDHA PRATAMA,RIYANARTO SARNO, “Personality classification based on twitter text,” International Conference on Data and Software Testing ,2015.
2. DEJAN MARKOVIKJ, SONJA GIEVSKA, MICHALKOSINSKI, DAVID STILLWELL, “Mining Facebook data forPredictive personality Modeling,” Association for the Advancement of Artificial Intelligence,2013.

3. YILUN WANG, “Understanding Personality through social media,” International of computer Science standford University,2015.
4. MANASI OMBHASE, PRAJAKTA GOGATE, TEJAS APTIL, KARAN NAIR, PROF. GAYATRI HEGDE,“Automated Personality classification using Data Mining Techniques,” International Conference on Data and Software Testing , April 2017.
5. MAHESH KINI, SAROJA DEVI, PRASHANT G DESAI, NIRANJAN CHIPLUNKAR,”Text Mining approach toclassify Technical Research Document using naive bayes”, International journal of Advanced Research in computer and communication engineering,Vol $,issue 7, July 2015.
6. DURGESH K.SRIVASTAVA, LEKHA BHAMBHU, “DATA Classification using Support Vector Machine,” Journal of Theoretical and Applied Information Technology,2005.
7.