Plagiarism Scan Report

Report Generation Date: May 26,2020 **Words:** 873 **Characters:** 5472

Exclude URL :

7%	93%
Plagiarism	Unique
4	50
Plagiarized Sentences	Unique Sentences

Content Checked for Plagiarism

3.4.1 Algorithm The algorithm that we use here is Naive Bayes Classier Algorithm.This is a classication technique which is based on Bayes Theorem with an assumption of independence among the given data values/set. In our project , text given in several posts on social media platforms are the data values which forms our data set . This works well on the text/categorical data instead of numeric data. A classier under the supervised learning based on probabilistic logic(bayes theorem). In lay-man language, we can say that a existence of the number of posts of a particular person/ individual is unrelated / independent from the existence of number of posts of a another person /individual and that's the assumption in naives bayes classier. For each attribute from each class set ,it uses probability to make predictions. fX1; X2; : : : : : : : : : : : : ; Xng fC1; : : : : : : :; Ckg (1) In our project, X1 denotes the no of posts of a particular person/ individual. X2 denotes the no of posts of another person/ individual. Xn denotes the no of posts of nth person/individual . C1is the probability of the number of posts describing a particular trait per- sonality trait. Ck is the probability of the number of posts describing a kth { personality trait . The data model which is yielded is called as Predictive model with probabilis- tic problems at foundation. Bayes theorem gives us a way to calculate posterior probability by the given equation:- 7 Posterior probability= Likelihood* Class Prior Probability /Predictor Prior Probability Input=50 posts Trained Data 58.354826823876195 Test Data 10.4463235294117645 Output:- The language used in the number of posts (twitter posts) re ect- ing/describing each 16 types/ traits of personality and predicting the personal- ity trait people possess. To increase the accuracy of our model, we took 4 classiers of personality traits to classify the individual's personality(using MBTI) Input=50 posts Data Introvert-Extrovert Intuition-Sensing Thinking-Feeling Prospecting-Judging Trained 57.401437 67.658438 79.504422 73.613670 Test 49.705882 73.566176 53.198529 48.768382 Output:- The language used in the number of posts (twitter posts) re ecting/describing each(4)classiers of personality and predicting the MBTI Trait using the clas- siers. 8 Algorithm 1 1: split [] 2: for i in range 16 do 3: split += [len(features[i])*0.6] 4: split np:array(split; dtype = int)end for 5:6: train [] 7: for i in range 16 do 8: train +=features[i][:split[i] 9: 10: end for 11: sentimentclassifier = NaiveBayesClassifier:train(train) 12: nltk.classify.util.accuracy(sentimentclassifier; train)100 13: test [] 14: for i in range 16 do 15: test +=feautures[i][split[i]:] 16: 17: end for 18: nltk.classify.util.accuracy(sentimentclassifier; test)100 4 Experiment Setup & Results We are splitting our data set into training and test data. We are calculating accuracy in 4 trials (50:50, 60:40, 70:30 ,80:20) . This splitting is done on complete dataset where we have 16 classes each class representing a personality type. The accuracy through this method turned out to be 10% approximately as shown in Figure 3. Hence, instead of selecting all 16 personalities as a unique feature, we decided to simplify the dataset. The MBTI personality type divides

everyone into 16 personality types across 4 axis. 1. Introversion(I) or Extroversion(E) 2. Intuition(N) or Sensing(S) 3. Thinking (T) or Feeling (F) 4. Prospecting (P) or Judging (J) Now we have 4 classes, we create 4 classiers (Naive Bayes Classier to classify the person into a particular personality) . We got approximately 53% accuracy after classifying the personality types into 4 classes rather than 16 types, as shown in Figure 4. In Figure 5, the graph shows which trait has higher percentage than the other and thus chooses the higher trait to predict the personality type. 9 Figure 3: Splits vs Accuracy In Figure 6,we tried predicting the personality of Barack Obama based on his tweets and we got INFJ as shown in Figure 6 which is dierent from his original personality which is ENFJ. 5 Conclusion and Future Work There is a slight dierence between the personality predicted by the model and the personality predicted by 16personalities.com.This might be because: 1. We have not scraped the prole but have copied few posts of the user into the test le. 2. We are using Naive Bayes classier, the accuracy of which is 50%,so according to the accuracy of the model,we are getting a good result. 3. We didn't proportionalise the data and thus it's more likely that our code predicts INFP or traits related to INFP as it has the highest number of posts. Our data is very imbalanced. 10 Figure 4: After classication into 4 classes Figure 5: Model Classifying trait References [1] Mihai Gavrilesku. Study on determining the myers briggs personality type based on individual's handwriting. The fth IEEE International Conference on E-Health and bioengineering, 11,2015. [2] Champa H N and Dr. K R Anandakumar. Articial neural network for human behaviour prediction through handwriting analysis. International Journal of Computer Applications, 2010. 11 Figure 6: Personality prediction of Barack Obama 6 Appendix: Similarity Report In this soft copy, attach snapshot of rst page of similarity report which clearly indicates percentage plagiarism. When submitting this report in print, attach print-out of entire similarity report at the end of project report print-out. 12

Sources

Machine Learning | Naive Bayes Classifier - YouTube

naive bayes algorithm is a method set of probabilities. for each attribute from each class set, it uses probability to make predictions. #machinelearning...

3%

https://www.youtube.com/watch?v=AkPSgAh3q4g

Progress in Computing, Analytics and Networking: Proceedings ...

... the posterior probability can be computed as Posterior Probability = (Likelihood × Class Prior Probability)Predictor Prior Probability For a fixed prior probability ...

2%

https://books.google.com.au/books?id=UFBVDwAAQBAJ

Study on determining the Myers-Briggs personality type based on...

results show an accuracy of 86.7% in determining the personality type, with highest accuracies for extravert vs. introvert and thinking vs. feeling personality primitives.the classic work on the 16 major personality types as identified in the myers-briggs type indicator.

2%

https://www.researchgate.net/publication/301563870_Study_on_determining_the_Myers-Briggs_personality_type_based_on_individual's_handwriting

(PDF) automated handwriting analysis for human behavior prediction

12. "automated human behavior prediction through handwriting analysis" by champa h n and k r anandakumar.champa h n, dr. k r ananda kumar, "artificial neural network for human behaviour prediction through handwriting analysis", international journal of computer applications(0975...

2%

https://www.researchgate.net/publication/321294888_AUTOMATED_HANDWRITING_ANALYSIS_FOR_HUMAN_BEHAVIOR_PREDICTION