


The background is a dark blue gradient. On the left, there are two overlapping geometric shapes: a blue parallelogram and a light green parallelogram. In the bottom left corner, there is a circular inset showing a close-up of a circuit board with various components. In the top right corner, there is a faint, stylized pattern of interconnected lines and squares, resembling a circuit or a data network.


Deanononymizing Quora Answers

Arunika Yadav :1601CS56

Harshika :1601CS14



Quora is a knowledge sharing website where users can ask/answer questions with the option of anonymity .We investigate the problem of Author-identification for Quora answers using deep learning techniques in Natural Language Processing.



Authorship identification has been a very important and practical problem in Natural Language Processing. The problem is to identify the author of a document from a given list of possible authors. It enables us to identify the most likely author of articles, news or messages. Authorship identification can be applied to tasks such as identifying anonymous author, detecting plagiarism or finding ghost writer.




Problem Statement:

We will solve the problem of deanonymizing quora answers using the different techniques of authorship identification which are available and also study a combination of models to improve the accuracy.



Why is this problem
exciting??



Quora, currently, does not have a system to deanonymize it's answers. Hence it is vulnerable to cyber crime, plagiarism and other such content related crime. Thus, in order to curb the anonymity of answers and offer a better experience to the users we offer this solution.



AIMS AND OBJECTIVES



At the end of the project we want to realise an author if we feed his answer into our model ,hence achieving deanonymization.



WORK FLOW:

Data Set Collected :

Quora (From Kaggle.com):

- Collected 50 answers of 50 authors each
- Top writer's answers
- Split the dataset into 8:2 ratio(train : test)



Data Preprocessing:

- Importing the dataset in the form of text file itself (unlike the regular way of importing it in the form of .csv file using the panda library).
- Replaced all the 'newline' characters with a space in order to make the full answer as a whole chunk of words rather .
- CLEANING the Data:
 - # removing punctuations
 - # converting into lower case
 - # removing the stop words



Data Preprocessing :

stemming

- Joining the answers of one particular author to form a list of words



Data Preprocessing:

The **bag-of-words** model is one of the feature extraction algorithms for text.

The bag of words model ignores grammar and order of words.

We defined a collection of strings called a corpus. Then we used the CountVectorizer to create vectors from the corpus(collection of strings).

The number of elements in the feature vector is called the **dimension**.

This made the data ideal for splitting the dataset into a ratio of 8:2 to be taken as the training data and test data respectively.



Naive Bayes' Classifier :

- The Naïve Bayes classifier attempts to simplify the problem by computing the probability of some value given the set of all attributes.
- It is one of the first methods applied to a classification problem.
- It is a supervised classification algorithm



There are three popular Naive Bayes algorithms:

- **Gaussian Naive Bayes,**
- **Multinomial Naive Bayes.**
- **Bernoulli Naive Bayes.**

We have used the Gaussian Naive Bayes approach to classify our test dataset. We have tested for 50 authors and achieved an accuracy of 69 percent.

The link to our github code :

https://github.com/arunikayadav42/Deanonymizing_quora_answers/tree/master




TIMELINE FOR FUTURE WORK

PROPOSED MODELS:

#1 : STYLE MARKER FEATURES MODEL : Stylometric features or style markers try to capture the writing style of the author. We have used a large number of stylometric features in our model.

#2 : DOCUMENT FINGER-PRINTING : Fingerprint identification is a well-known technique in forensic sciences. The basic idea of identifying a subject based on a set of features left by the subject actions or behavior can also be applied to other domains. Identifying text authorship based on an author fingerprint is one such application. Finger-printing has mainly been used in the past for plagiarism detection.



#3 : WORD FREQUENCY MODEL :The term frequency-inverse document frequency (also called tf-idf), is a well known method to evaluate how important is a word in a document.tf-idf is a very interesting way to convert the textual representation of information into sparse features.

#4 : LSTM Model(Long Short-Term Memory):The LSTM network is a type of recurrent neural network used in deep learning because very large datasets can be succesfully trained.

At, the end of the day ,improving the accuracy is all that matters!!



THANK YOU