

FAKE NEWS DETECTION USING NLP

PHASE - 4

INTRODUCTION:

- * We consume news through several mediums throughout the day in our daily routine, but sometimes it becomes difficult to decide which one is fake and which one is authentic.
- * Do you trust all the news you consume from online media
- * Every news that we consume is not real.
- * If you listen to fake news it means you are collecting the wrong information from the world which can affect society because a person's views or thoughts can change after consuming fake news which the user perceives to be true.
- * Since all the news we encounter in our day-to-day life is not authentic, how do we categorize if the news is fake or real.

PREPROCESSING THE DATA :

- * In data processing, we will focus on the text column on this data which actually contains the news part.
- * We will modify this text column to extract more information to make the model more predictable.

*To extract information from the text column, we will use a library, which we know by the name of 'nltk'.

*Here we will use functionalities of the 'nltk' library named Removing Stopwords, Tokenization, and Lemmatization.

*So we will see these functionalities one by one with these three examples. Hope you will have a better understanding of extracting information from the text column after this.

DIFFERENT ANALYSIS OF THE PROJECT :

* NLP employs algorithms to analyze language patterns and identify markers of misinformation.

* One critical aspect is sentiment analysis, which gauges the emotional tone of a piece.

* Fake news might employ sensational language to provoke reactions, while genuine news tends to maintain a more balanced tone.

* Another aspect is linguistic consistency.

* Fact-based news adheres to grammatical and logical structures, while fake news might exhibit inconsistencies or illogical reasoning.

* NLP algorithms can identify such anomalies, flagging content for further evaluation.

CONCLUSION :

- * We have classified our news data using three classification models.
- * We have analysed the performance of the models using accuracy and confusion matrix.
- * But this is only a beginning point for the problem.
- * There are advanced techniques like BERT, GloVe and ELMo which are popularly used in the field of NLP.
- * If you are interested in NLP, you can work forward with these techniques.