

People's Behaviour Analysis in Chat Message using Natural Language Processing

SYNOPSIS

Nowadays, the mode of communication is mainly through messages. A lot of information has been conveyed through WhatsApp. WhatsApp is the most popular chat application with active users of more than 650 million. It has been widely used by all, especially among the business people and youngsters. Using several analysing tools, users can analyse the WhatsApp group chat or personal chat. Authentically users wish to analyse their chat for several purposes. This research work is intended to perform a flirt analysis and time analysis. This project has many use cases like the parent, who wants to analyse their child chat; the police, who want to get valuable information from culprit chat; the business people, who wants to know the status of the business in the group chat. Using the Deep Learning model (NLP), sentimental analysis has been performed for each text. This helps to find the state of mind of the chatters. Further, this research work calculates the number of positive and negative statements that are used by each person in the text by using the text mining concept. As now due to this pandemic situation, every conversation and also the important discussion has been done through the WhatsApp and it was highly needed for the person who wants to check their child's conversation and also for the higher authority for enquiry and for the business chair person who are needed to analyse their business wellbeing group can also be used for their personal usage of analyse using the algorithm in this method.

Data collection is the step in which we are going to collect the data from various users having the chat conversation with their friends and known people. WhatsApp has an option to export our chats as a text file. An exported text file will be converted into a CSV file for easy analysis.

By this we can collect the data and can store the details for the sentimental analysis

Collected data will be made pre-processed which means encoding the categorical information in the data, dropping unwanted parameters, scaling the parameter values to achieve normal distribution (Zero mean and Standard Deviation as one), handling missing values, and so on. Pre processing is the method in which the data will be undergone into certain steps which can be made for the future analysis procedures which will be efficient for the calculation procedures.

In the flirt analysis, the text message will be compared with the flirt list. The total number of flirt words used by a person will be divided by the total number of flirt words. The flirt percentage will be calculated for each person in the chat

Flirt Analysis is the step in which the major part of the analysis was done. The main aim of the module is to detect the flirt which was done between the peoples via chat by comparing the chat with the list

Natural language processing (NLP)

It is a subfield of computer science and artificial intelligence concerned with the interactions between computers and human language, in particular how to program computers to process and analyze large amounts of natural language data. The goal is a computer capable of "understanding" the contents of

documents, including the contextual nuances of the language within them. The technology can then accurately extract information and insights contained in the documents as well as categorize and organize the documents themselves.