**AIM:** Text classification for Sentiment analysis using KNN

**SOFTWARE/TOOLS USED:**

* Python
* Nltk
* Twitter text dataset
* Jupyter

**THEORY:**

Sentiment analysis is the process of using natural language processing, text analysis, and statistics to analyze customer sentiment. The best businesses understand the sentiment of their customers—what people are saying, how they’re saying it, and what they mean. Customer sentiment can be found in tweets, comments, reviews, or other places where people mention your brand. Sentiment Analysis is the domain of understanding these emotions with software, and it’s a must-understand for developers and business leaders in a modern workplace.

As with many other fields, advances in deep learning have brought sentiment analysis into the foreground of cutting-edge algorithms. Today we use natural language processing, statistics, and text analysis to extract, and identify the sentiment of words into positive, negative, or neutral categories.

K-Nearest Neighbors

K-Nearest Neighbor classifier is one of the introductory supervised classifiers, which every data science learner should be aware of. This algorithm was first used for a pattern classification task which was first used by Fix & Hodges in 1951. To be similar the name was given as KNN classifier. KNN aims for pattern recognition tasks.

K-Nearest Neighbor also known as KNN is a supervised learning algorithm that can be used for regression as well as classification problems. Generally, it is used for classification problems in machine learning.

KNN works on a principle assuming every data point falling in near to each other is falling in the same class. In other words, it classifies a new data point based on similarity.

**RESULT:**

Sentiment analysis using KNNs on twitter text dataset was attempted and model metrices recorded

**CONCLUSION:**

An approach towards sentiment analysis using KNN was understood and implemented in python