4.8 Experiment No. 7

Aim: Emotion Detection from Text (NLP + Classification)

Build a model to detect emotions (joy, anger, sadness, etc.) from user-generated text like tweets, comments, or messages using Natural Language Processing (NLP) and classification techniques.

Objective:

- To preprocess and clean textual data for analysis.
- To convert text into numerical features using TF-IDF vectorization.
- To train a machine learning model (Naive Bayes) for emotion classification.
- To evaluate the model's performance using accuracy, confusion matrix, and classification report.
- To predict the emotion for unseen or user-input text.

Theory:

Emotion Detection from text is a Natural Language Processing (NLP) task that involves analyzing textual content to determine the underlying emotional tone. It is treated as a text classification problem, where each input sentence is categorized into a predefined set of emotion classes (e.g., joy, anger, sadness, fear).

The process involves:
1. Text Preprocessing: Cleaning data by removing punctuation, converting to lowercase, and removing stopwords.
2. Feature Extraction: Transforming text into numeric form using TF-IDF (Term Frequency–Inverse Document Frequency), which measures the importance of words in a document relative to the dataset.
3. Model Training: A Naive Bayes classifier is used, as it performs efficiently with text data assuming independence between features (words).
4. Model Evaluation: The trained model is evaluated using metrics such as accuracy, precision, recall, and F1-score.

Applications:

- Sentiment and emotion analysis on social media platforms (Twitter, Reddit, etc.).

- Customer feedback and review analysis.
- Chatbots and virtual assistants for emotion-aware responses.
- Mental health monitoring through text-based inputs.
- Market analysis and brand reputation management.

Input:
A CSV dataset containing user-generated text and corresponding emotion labels.
Example:

Text	Emotion
I am feeling so happy today!	joy
I am really upset with what happened.	anger
I miss you so much.	sadness

Output:
- Model accuracy, confusion matrix, and classification report.
- Predicted emotion for user-input text.

Conclusion:

The Naive Bayes-based emotion detection model successfully classifies textual data into appropriate emotion categories with good accuracy.
 TF-IDF vectorization effectively converts text into meaningful numeric representations, enabling efficient model learning and prediction.

Outcome:
The outcomes of this experiment are:
- Gained understanding of how NLP techniques are applied to emotion detection.
- Learned to preprocess text data and convert it into machine-understandable format.
- Implemented a text classification model to predict human emotions based on written content.
- Understood model evaluation using various classification metrics.