SENTIMENT ANALYSIS

DATA MINING-II PROJECT

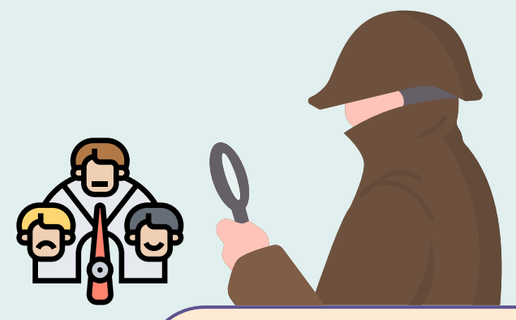
NATURAL LANGUAGE PROCESSING

Identifying emotions has become an integral part of many NLP and data science projects. With the help of this dataset, one can train and build various robust models and perform emotional analysis. This dataset is having 724 rows and 3 variables.

DATASET LINK: https://github.com/Harsh6063/Data_science/blob/main/emotional_dataset.csv

PROBLEM STATEMENT:-

- **The increasing volume of text data demands advanced methods to understand emotional and sentiment patterns. This project focuses on developing a system using NLP and machine learning to detect emotions (happiness, sadness, anger, etc.) and classify sentiments (positive, negative, neutral), contributing to research in emotional and sentiment analysis.**



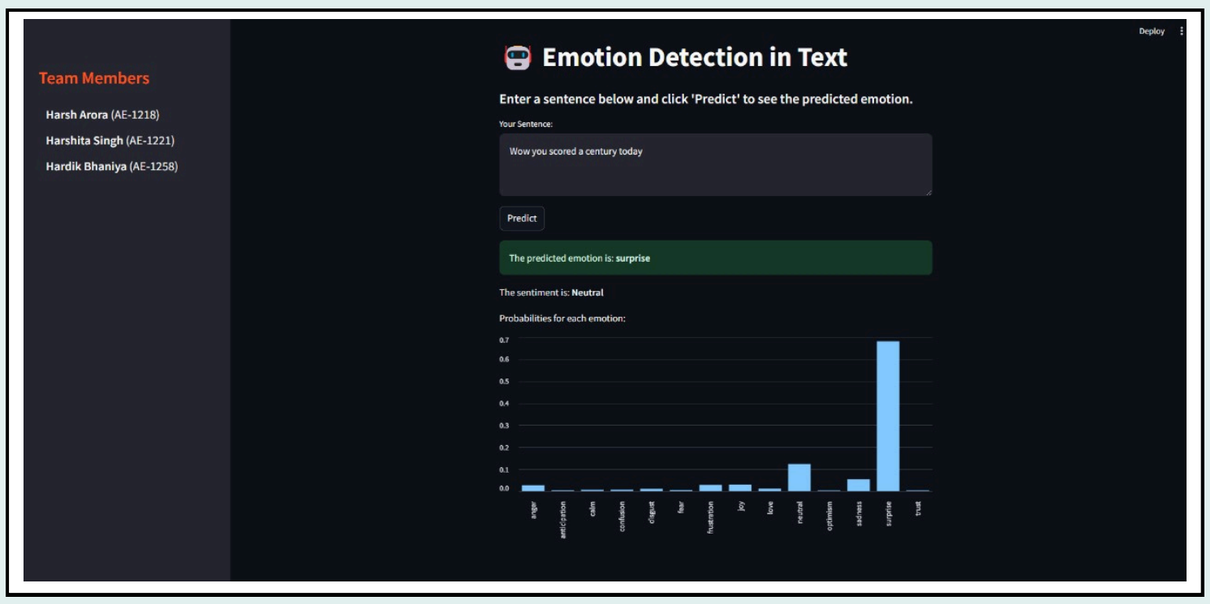
OBJECTIVE:-

To develop a model using NLP and machine learning techniques for detecting emotions (e.g., happiness, sadness, anger, etc.) and classifying sentiments (positive, negative, neutral) in text data, enabling deeper analysis of emotional patterns for research in various domains like mental health and social media.

HOME PAGE

	precision	recall	f1-score	support
anger	0.56	0.45	0.50	11
anticipation	0.78	1.00	0.88	7
calm	0.90	1.00	0.95	9
confusion	0.88	1.00	0.93	7
disgust	0.56	0.77	0.65	13
fear	0.86	0.67	0.75	9
frustration	0.82	0.60	0.69	15
joy	0.33	0.33	0.33	12
love	0.67	0.77	0.71	13
neutral	0.64	0.78	0.70	9
optimism	0.75	0.90	0.82	10
sadness	0.75	0.67	0.71	9
surprise	1.00	0.60	0.75	10
trust	0.89	0.73	0.80	11
accuracy			0.71	145
macro avg	0.74	0.73	0.73	145
weighted avg	0.73	0.71	0.71	145

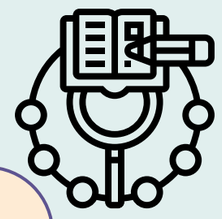
ACCURACY OF EVERY EMOTIONS



SENTENCE TO PREDICT EMOTION & SENTIMENT

	Classifier	Accuracy	Precision
0	Random Forest	0.586207	0.685890
1	Decision Tree	0.503448	0.631448
2	SVM	0.703448	0.736531

COMPARISON BETWEEN SVM, RANDOM FOREST, AND DECISION TREE



CONCLUSION:-

- **Accurate Emotion Detection:** The system effectively predicts emotions (e.g., anticipation, calm, sadness) with an overall accuracy of 71%.
- **Superior Classifier Performance:** SVM, optimized via bagging, outperformed Random Forest and Decision Tree classifiers with the highest accuracy (70.34%) and precision (73.65%).
- **Emotion-wise Performance:** Emotions like "calm" and "confusion" achieved high precision and recall, while "joy" had lower performance, indicating areas for improvement.
- **Practical Applications:** The model supports research in analyzing emotional patterns in various domains like mental health and social media analysis.
- **Future Scope:** Incorporating larger, balanced datasets and multilingual support could further improve the system's performance and applicability.

SUMMARY:-

This project develops a system to detect emotions and classify sentiments in text using NLP and machine learning. Among tested classifiers (SVM, Random Forest, Decision Tree), SVM with bagging achieved the best performance, providing accurate emotion predictions and analyzing sentiment patterns for research in mental health, feedback, and social media.

HARSH ARORA (AE-1218) B.SC.(H) COMPUTER SCIENCE

HARSHITA SINGH (AE-1221) B.SC.(H) COMPUTER SCIENCE

HARDIK BHANIYA (AE-1258) B.SC.(H) COMPUTER SCIENCE