

SENTISCAPE

AI POWERED SENTIMENT ANALYSIS TEST FOR TEST DATA

AI HACKTHON

TEAM ID : 25KTJAI N17843

N V K A SRI VARSHINI (25KTJN V451953)

MOYYI PRASHANTHI(25KTJMOY759813)

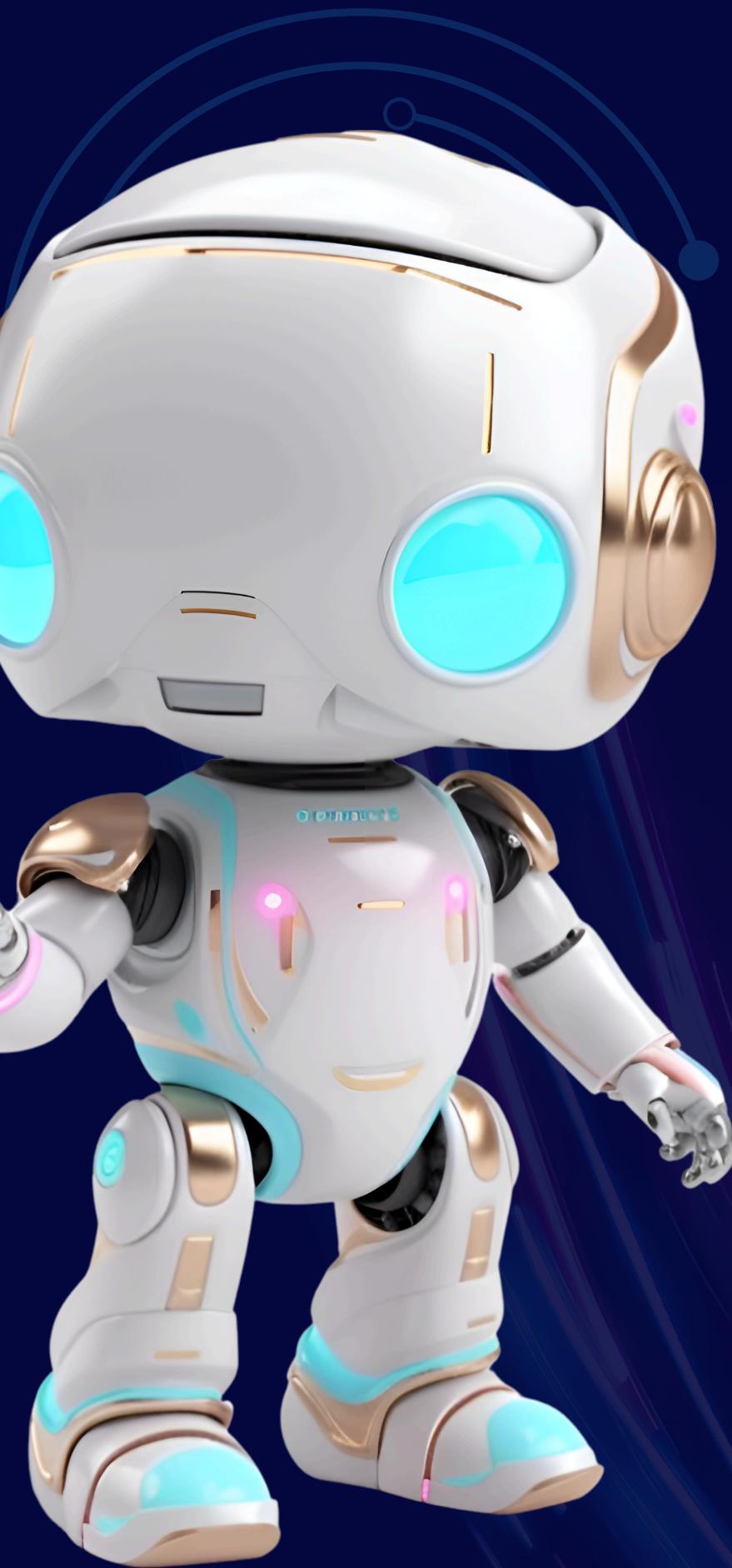
JADA SUKANYA (25KTJJ S526191)

BALLA SAI SAMPATH (25KTJSA1273966)

INTRODUCTION

Sentiment analysis, often referred to as opinion mining, is a field within natural language processing (NLP) that focuses on analyzing and understanding the emotional tone behind a piece of text. It aims to determine whether the sentiment expressed is positive, negative, or neutral, often with varying degrees of intensity.

In today's digital age, people express their opinions widely across platforms like social media, reviews, and forums. Analyzing this data manually is impractical due to its sheer volume. Sentiment analysis automates the process, providing actionable insights for businesses, researchers, and decision-makers.



ABSTRACT

In today's data-driven society, it becomes crucial for organizations to understand customers sentiments better in order to improve satisfaction and loyalty. Below is a presentation of the new innovative sentiment analysis tool that looks forward to automatizing the process of how feedback is being classified to ultimately give actionable information on customer experiences. The main technologies used herein include the HuggingFace Pipeline, which will accurately and in real-time analyze any form of feedback from customers.

PROBLEM STATEMENT

Businesses receive vast amounts of customer feedback daily via surveys, reviews, and social media, making manual sentiment analysis inefficient and error-prone. To address this, develop an AI solution for real-time sentiment analysis that classifies feedback as positive, neutral, or negative and provides actionable insights to enhance customer satisfaction and loyalty. The solution must utilize a provided labeled dataset with "comments" and "sentiment" columns to train a machine learning or deep learning model. Participants will deliver a trained model and a CSV file with predictions on a test dataset, enabling businesses to automate feedback analysis and drive data-driven improvements.



METHODOLOGY

The project uses

- **Data Preprocessing**: Loads the reviews and sentiment labels from a CSV file, cleans the text by removing special characters, and normalizing it using Python libraries like pandas and re.
- **TF-IDF Analysis**: Uses TF-IDF for term extraction, implemented through scikit-learn to identify relevant terms across sentiment categories (Positive, Negative, Neutral).
- **Sentiment Analysis**: Leverages transformers and Hugging Face models for sentiment classification and calculates the sentiment distribution while identifying distinctive terms for each sentiment group.
- **Insight Generation**: Uses Python to generate representative reviews, and summarizes insights in both a natural language paragraph and a detailed report for actionable business decisions.





CONCLUSION

In conclusion, the AdvancedReviewAnalyzer successfully processes and analyzes customer reviews using TF-IDF to extract key insights from sentiment-labeled data. It categorizes reviews into positive, negative, and neutral sentiments, identifies important terms, and highlights distinctive features specific to each sentiment group. The tool also selects representative reviews based on TF-IDF terms to illustrate common customer feedback. The results are presented through a natural language summary and a detailed report, providing businesses with actionable insights to enhance their products or services. This methodology offers a robust approach to sentiment analysis, enabling data-driven decision-making and improving customer experience.

THANKYOU