

Natural Language Processing

PROJECT

**MARATHI TWEET
SENTIMENT ANALYSIS**

Introduction

Sentiment analysis, a key component of Natural Language Processing (NLP), plays a crucial role in understanding the sentiments expressed in text data. In today's digital age, the prevalence of social media has made sentiment analysis even more significant. In this presentation, we delve into the domain of Marathi tweet sentiment analysis, highlighting the need to decipher the emotions, opinions, and attitudes conveyed through tweets in the Marathi language.





Semester 5

Motivation

The motivation behind embarking on this project lies in recognizing the importance of understanding Marathi sentiments in the context of social media. By gaining insights into the sentiments of Marathi-speaking users, we can make informed decisions in various fields, including market research, brand management, and social listening. The project aims to bridge the gap between the Marathi-speaking population and NLP applications.





Semester 5

Problem Statement

Our project is centered on addressing the challenge of sentiment analysis in Marathi text. We confront the complexities associated with this regional language, working with a dataset of Marathi tweets to categorize sentiments. Our goal is to develop a model that can accurately classify the sentiments expressed in Marathi social media content.





Semester 5

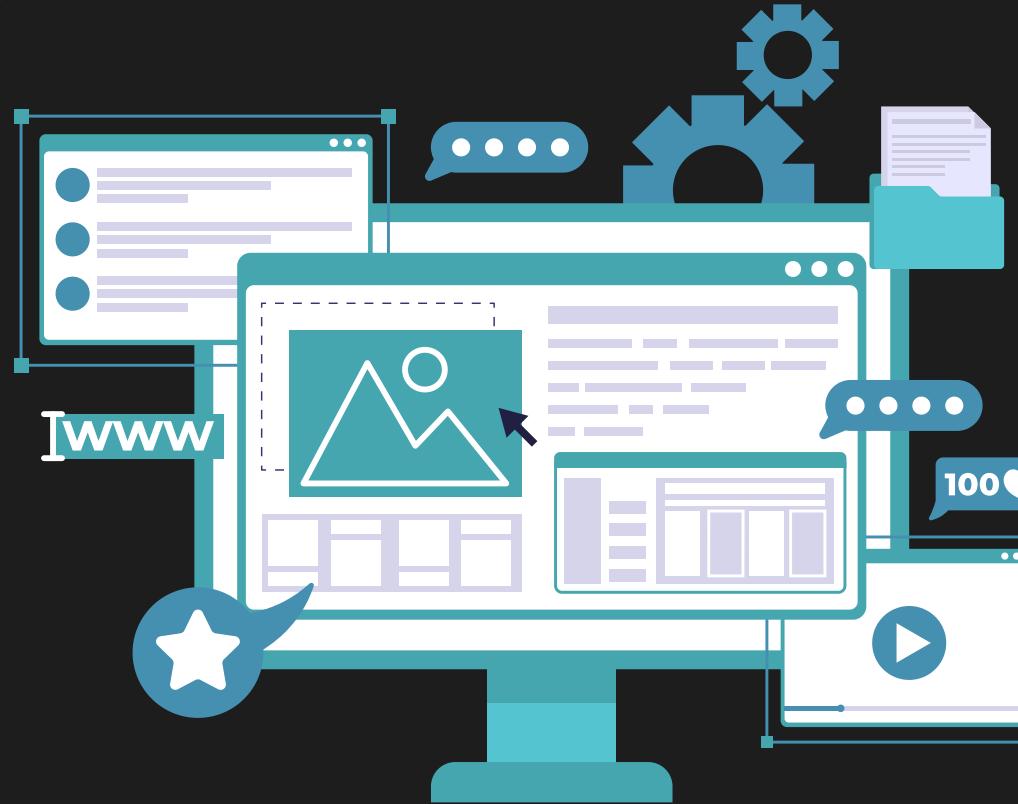
Objective

Our objectives encompass sentiment classification in Marathi tweets, rigorous model evaluation, and exploring the practical applications of our analysis. By achieving these objectives, we aim to enhance our understanding of Marathi sentiments and contribute to the broader field of NLP.



Methodology

- 1. Data Collection
- 2. Data Preprocessing
- 3. Data Encoding and Tokenization
- 4. Model Selection: Indic BERT
- 5. Fine-Tuning
- 6. Sentiment Analysis
- 7. Model Evaluation: Accuracy , F1 Score, Precision, Recall





Model's WorkFlow

extracted tweet data and eliminated unnecessary symbols.

1

2

3

4

Data encoding and tokenization for the pre-trained model.

Sentiment analysis of tweets in Marathi.

altered the n architecture using a pre-trained model.



Semester 5

Results

Validation	Score
Recall	76%
F1 Score	75%
Precision	74%
Accuracy	75%





Semester 5

Limitations

- **Limited Resources:** There aren't many tools and data available for understanding Marathi sentiment, which can make it harder to build accurate models.
- **Regional Variation:** Marathi speakers from different regions express feelings differently, affecting analysis accuracy.





Semester 5

Conclusion

In conclusion, our project has successfully navigated the complexities of Marathi tweet sentiment analysis. The understanding of Marathi sentiments and the impact of this analysis cannot be understated. Our work paves the way for a deeper exploration of regional languages in NLP, contributing to a more inclusive digital landscape.



Thanks
to you