

VarthaGenie

A Telugu News Summarizer



Problem Statement

"Despite the growing demand for convenient access to news content in regional languages, there is a notable scarcity of Natural Language Processing (NLP) tools tailored specifically for Telugu speakers. This project addresses this gap by developing a Telugu news reader chatbot equipped with summarization and question-answering capabilities, aiming to enhance accessibility and engagement with news articles in the Telugu language."



Feedback from 1st Evaluation

- 01 Our project was ambitious – We wanted to do both Summarizing and Q&A tool
- 02 We were asked to look into contextual datasets and links – Relevant data links

Proposed Timeline

Completion of Dataset creation
and Data Pre-processing

31st March

Fine-tuning and improving
the chosen language model

30th April

15th April

Implement text summarization
and deploy chosen language
model

15th May

Deployment and Evaluation
of the final product

Github Link:

https://github.com/raghu-prashanka/VarthaGenie_Team_22

Reference Model:

<https://huggingface.co/Helsinki-NLP/opus-mt-en-de>

Data Sets:

<https://huggingface.co/datasets/SuryaKrishna02/aya-telugu-news-articles?row=77>

<https://github.com/AnushaMotamarri/Telugu-Newspaper-Article-Dataset>

Team Members

- 01 Charan Sai Panuganti - SE21UCSE142
- 02 Raghu Prashanka - SE21UCSE039
- 03 Somishetty Badreesh - SE21UCSE211
- 04 C Dhanush Reddy - SE21ECM016
- 05 Varun Pepeti - SE21UECM040



Thank You