Poster for Natural Language Processing (202047809)
on
Gujarati News Translator & Summarizer Web App

Supervised by : DR . Priyang Bhatt
Prepared by: Prashant Kansara (12302130503003),Parth Kansara(12302130503008)
Academic Semester:7, Year: 2025-26

Abstract

A lightweight backend service that ingests Gujarati news (URL or raw text), removes boilerplate (e.g., copyright/footer), translates content to English, and produces a concise extractive summary. Built with FastAPI and designed for serverless deployment (e.g., Vercel) without heavy ML dependencies.

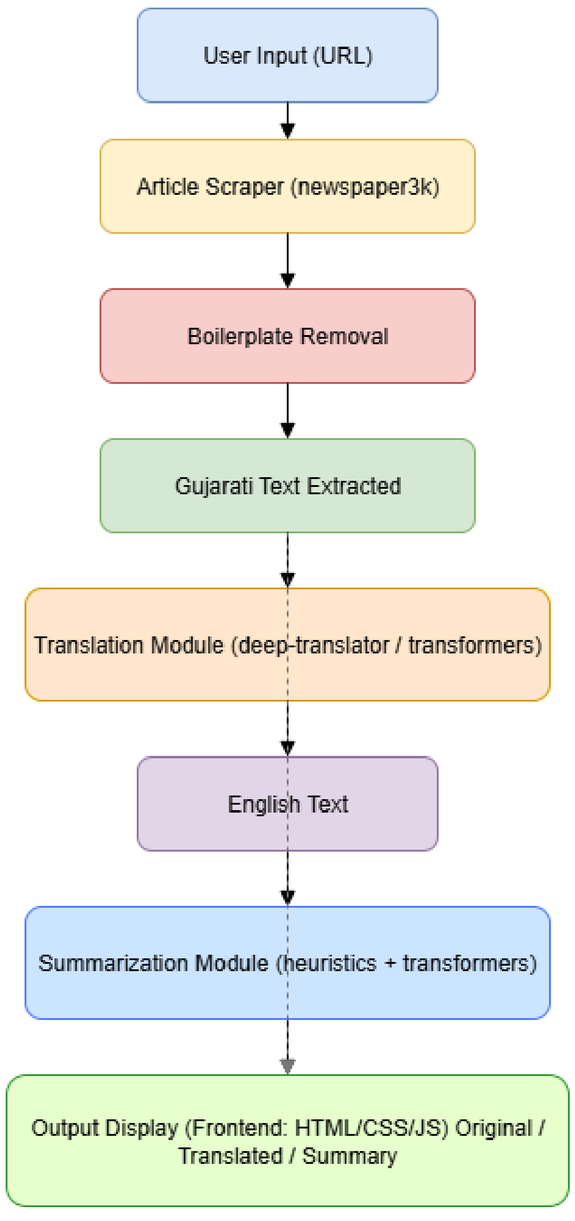
Introduction

Non-Gujarati audiences lack a fast, reliable way to extract Gujarati news articles, remove boilerplate, translate them to English, and obtain concise summaries. This project provides a scalable and lightweight solution suitable for serverless deployment on minimal resources.

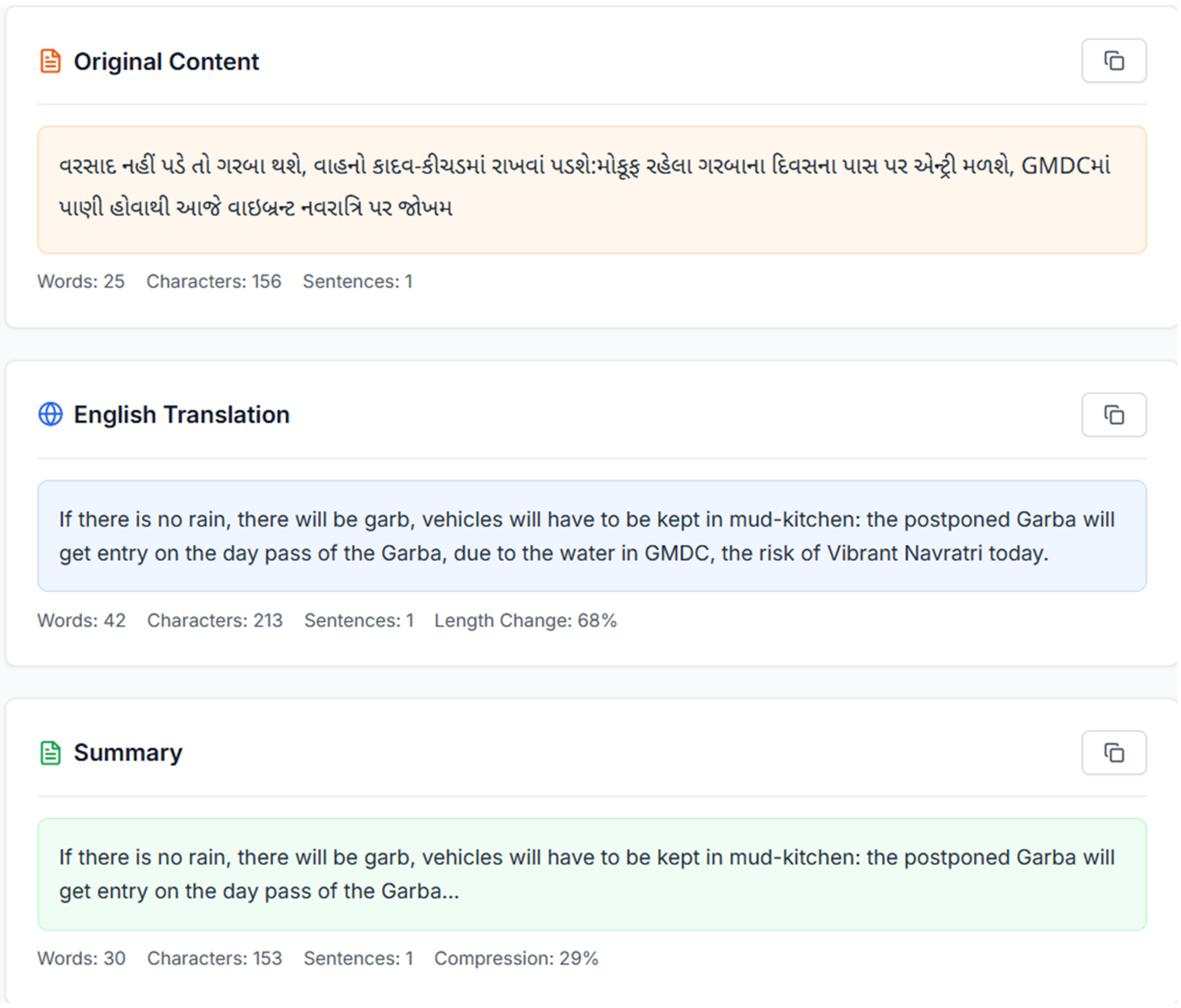
Tech Stack

Frontend: React
Backend: FastAPI, Flask
NLP & ML: transformers, torch, newspaper3k, deep-translator.
Deployment: Vercel

System Architecture



Result



Conclusion & Applications

The project delivers a lightweight, serverless-ready system for Gujarati news translation and summarization, enabling non-Gujarati audiences to quickly access regional news. By integrating article scraping, boilerplate removal, translation, and summarization into a single pipeline, it ensures efficiency, low dependency overhead, and real-world usability.

Applications:

Cross-lingual News Monitoring: Track Gujarati news in English for journalists and organizations.

Media Analytics & Research: Provide quick digests for media houses, researchers, and policymakers.

News Aggregation & Alerts: Integrate into real-time pipelines for concise English updates.

Educational & Accessibility Use: Support students, non-native speakers, and global stakeholders in understanding Gujarati content.

References

[1] T. Wolf *et al.*, “Transformers: State-of-the-Art Natural Language Processing,” in *Proc. 2020 Conf. Empirical Methods in Natural Language Processing (EMNLP)*, 2020. [Online]. Available: https://huggingface.co/transformers

[2] A. Paszke *et al.*, “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” in *Advances in Neural Information Processing Systems*, vol. 32, 2019. [Online]. Available: https://pytorch.org

[3] Newspaper3k, “Article scraping and curation library.” [Online]. Available: https://newspaper.readthedocs.io

[4] Deep-Translator, “Python library for translations using multiple APIs.” [Online]. Available: https://pypi.org/project/deep-translator/

[5] S. Ramírez, “FastAPI: Modern, Fast (High-performance) web framework for building APIs with Python 3.6+.” [Online]. Available: https://fastapi.tiangolo.com