SVKM'S NARSEE MONJEE INSTITUTE OF MANAGEMENT STUDIES

SCHOOL OF TECHNOLOGY AND MANAGEMENT ENGINEERING

SUBJECT: NATURAL LANGUAGE PROCESSING

PROJECT REPORT/DOCUMENTATION

SUBMITTED BY:
Name: Nishka Vijayvargiya
Class: B.Tech CE Sem VII
Batch: Batch 1
Roll no: T046
SAP Id: 70022200461

SUBMITTED TO:
Dr. Raj Gaurav Mishra
STME, NMIMS

1. Introduction

Natural Language Processing (NLP) enables machines to understand and process human language.
In today's information-heavy world, text summarization and keyword extraction are vital for quickly identifying the core meaning of large documents, news articles, or reports.

This project presents an NLP-based web application that automatically generates concise summaries and extracts relevant keywords using transformer models and semantic embeddings.

2. Objectives

1. Develop a user-friendly NLP web app for summarization and keyword extraction.
2. Use transformer-based models for high-quality summaries.
3. Use KeyBERT for semantically meaningful keyword extraction.
4. Enable user controls for summary and keyword tuning.
5. Visualize keyword importance through charts.

3. System Overview

System Architecture:
User Input → Preprocessing → Transformer Model (Summarization) → Summarized Text → KeyBERT Model → Keywords + Visualization

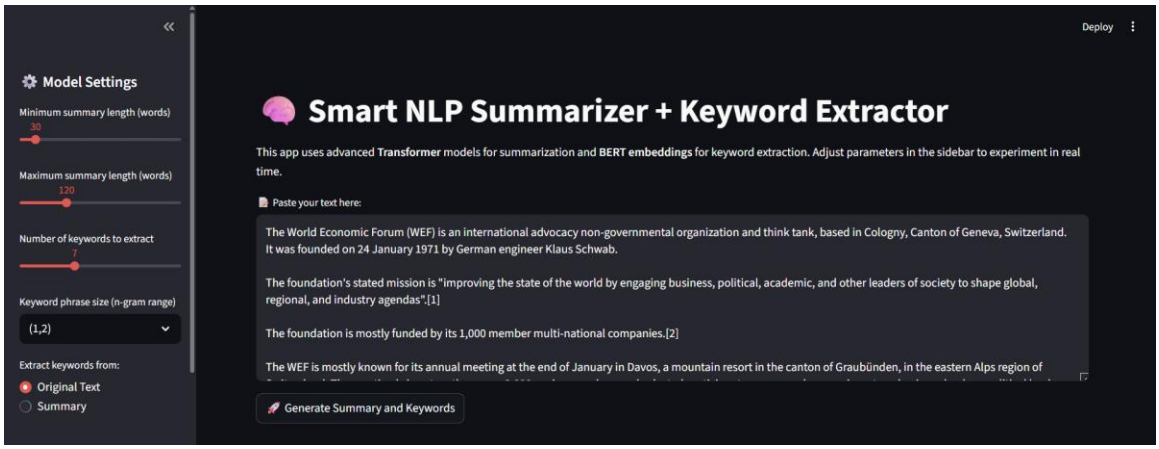
Tools and Technologies:
- Streamlit
- Transformers (BART-large-cnn)
- KeyBERT
- PyTorch (CPU)
- Matplotlib

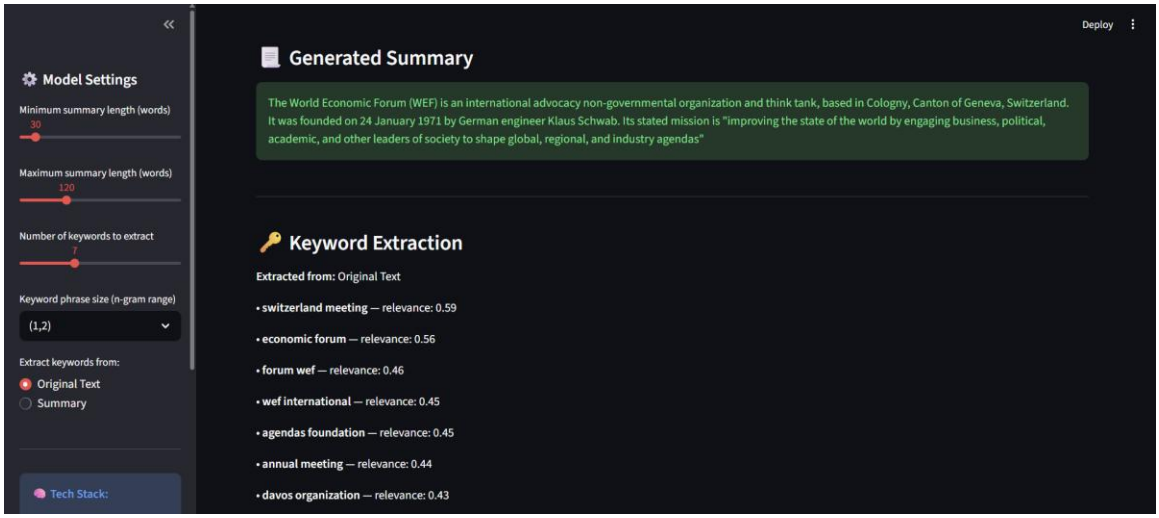
4. Methodology

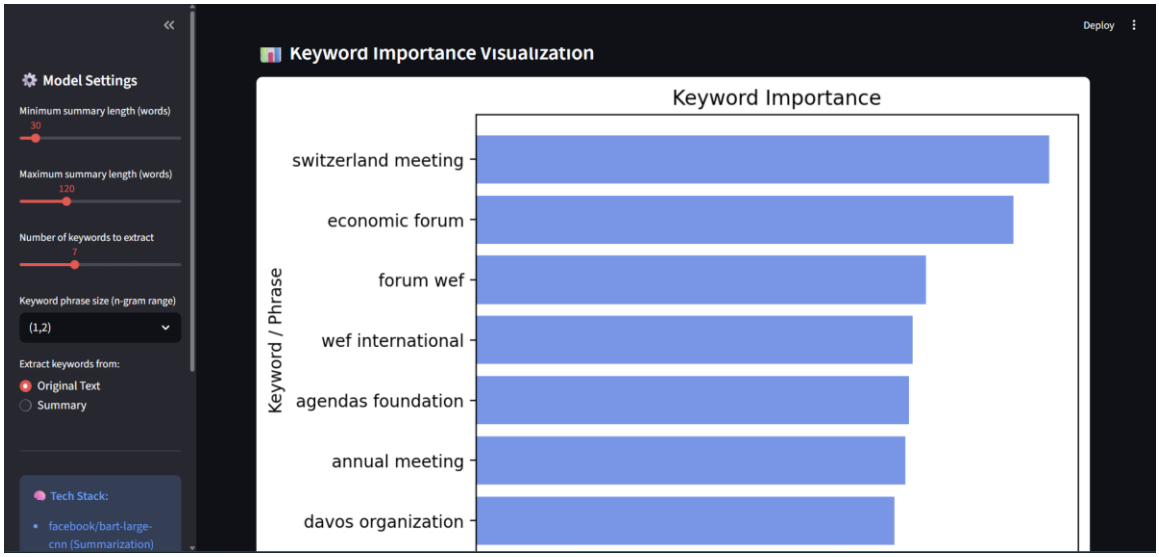
Step 1: User enters text into the Streamlit app.
Step 2: Text is summarized using facebook/bart-large-cnn (abstractive summarization).
Step 3: KeyBERT extracts semantically meaningful keyphrases.
Step 4: Results and relevance chart displayed, exportable as text file.

5. Results

Input: Long article
Output: Summarized text and top keywords
Visual Output: Bar chart showing keyword importance
File Output: summary_keywords.txt







6. Applications

- Research and news summarization
- Academic abstract generation
- SEO and content indexing
- Business intelligence text analysis

7. Conclusion

The Smart NLP Summarizer and Keyword Extractor demonstrates the potential of Transformer-based NLP applications for efficient and interpretable text understanding.