NARASARAOPETA ENGINEERING COLLEGE (AUTONOMOUS)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

2025-2026

Batch Number	DB5
Team Members	R. Nandu (23475A0521) SK. Faheem Ahmed (22471A05N9) K. David Shalem (23475A0506)
Guide	K.V. Narasimha Reddy M. Tech
Title	Text Summarization Using NLP Techniques
Domain/Technology	DEEP LEARNING
Base Paper Link	https://xplorestaging.ieee.org/document/10838534
Dataset Link	https://www.kaggle.com/datasets/shineucc/bbc-news-dataset/data
Software Requirements	Browser: Any latest browser like Chrome Operating System: Windows 7 Server or later Python (COLAB)
Hardware Requirements	SystemType: Intel Core i5 or above RAM: 8 GB Number of cores:5 Number of Threads: 4
Abstract	The most common traditional approaches to summarizing large texts while retaining their importance are TF-IDF and TextRank. However, This study's improved summarization methodology overcomes these limitations by combining the linguistic and semantic resources. For example, various textual features such as Named Entity Counts, Noun Counts, and Sentence Position contribute to the summarization quality appropriately. This study's summarization algorithm was tested using the CNN, XSum and BBC Summarization datasets, which aggregate documents from different areas. The methodology was checked against traditional methods using ROUGE-1 and ROUGE-2, ROUGE-L and BERTScore. The last one, BERTScore, evaluates the semantic similarity of the generated summaries and the references. **ROUGE Scores** reflect how much overlap there is between the model-generated summary and a human-written one. ROUGE-1 and ROUGE-L being over 60%. **BERT Score**'s **semantic similarity** 88.57%

Signature of the student(s) **Signature of the Guide** **Signature of the project coordinator**