Abstract:

This Project explores the field of text summarization by comparing the efficacy of two different models: spaCy and optimized T-5 models, using deep learning techniques. The main goal is to solve the problem of reducing large amounts of textual content while maintaining important context and information. The Cnn_dailymail dataset is the one used for fine-tuning. Torch, Transformers, and NLTK are used in the implementation, with special attention paid to the built-in text summarizing features of spaCy and the training of T-5 models for the same objective.

Evaluation metrics such as Rouge scores are employed to assess the effectiveness of each model, providing a detailed analysis that showcases the strengths and weaknesses of each approach. The project underscores the significance of text summarization in information retrieval, content processing, and knowledge extraction. Insights gained from understanding the nuances and effectiveness of different deep learning models contribute to advancements in natural language processing and comprehension.

In conclusion, this project anticipates providing valuable insights into the comparative performance of spaCy and fine-tuned T-5 models in text summarization. The findings contribute to the broader understanding of text summarization technologies, offering practical guidance for practitioners and researchers in the field.
Project URLs:

[Link to 15-minute video] : https://youtu.be/DkSgzKY1ZUk

[Link to 3-minute video]: https://youtu.be/ZRLSyxZ1OMA