Hybrid Deep Learning for Document Summarization & Non-Factoid Q&A

The study tries to address the problem of giving consumers solutions to complicated questions (non-factoid queries) that are brief, cohesive, and coherent. The issue originates from the present constraints of available question-answering systems, which sometimes struggle to combine data from several sources to provide a single thorough response. Additionally, it might be difficult for automatic text summary systems to provide clear and succinct summaries from extensive materials.

The research suggests a unique strategy that combines artificial text summarization with non-factoid question responding to address this issue. The objective is to develop a hybrid deep learning system that incorporates extractive, abstractive, and compressive methods for document summarization. Using pre-trained models like BERT, GPT, and Longformer, this system will be fine-tuned. It will consider the user's inquiry, find pertinent materials, and provide a well-organized summary that immediately responds to the user's question.

The research will involve the following objectives:

- Create the architecture, employ BERT for extractive summarizing, GPT for abstractive summarization, and Longformer for compressive summarization in the design and implementation of the hybrid deep learning system for document summarization.
- Analyze the hybrid summarization system's effectiveness and accuracy: Along with human review for coherence, relevance, and information completeness, evaluation will be done using metrics like Rouge and BLEU ratings.
- Create a deep learning-based system for answering questions that aren't factoids: This will include incorporating the summarization system into the question-answering architecture while utilizing DPR for document retrieval and BERT for question processing.
- Analyze the non-factoid question-answering system's performance, precision, and effectiveness systematically. Benchmarking against cutting-edge models and doing qualitative analyses are both part of the review process.

Designing the architecture, putting the models into use with Python and PyTorch, and fine-tuning the models utilizing MS MARCO datasets will all be part of the technique. To enhance model performance, hyperparameter adjustment will be done. Standard evaluation criteria will be used for quantitative assessment, and human evaluators will be used for qualitative evaluation.

Summary: By developing a hybrid deep learning system for document summarization and non-factoid question answering, this research intends to advance the area of deep learning-based text analysis. The objective is to improve the accuracy, brevity, and utility of automated systems in responding to complicated user inquiries by integrating several methodologies and thoroughly analyzing the system.

Project Specifications:

- OS: MacOS
- RAM: 8GB
- Processor: Apple Silicon M1
- Development IDE: Visual Studio Code & Google Colaboratory
- Python Environment: PyTorch, HuggingFace Transformers, NumPy, pandas, Matplotlib, NLTK, and Scikit-learn
- Version Control: git (Github)
- Cloud Computing (If necessary): Google colab GPU/ AWS Sagemaker Studio

Timeline:

Week 1-3:
- Data Collection and Preprocessing:
  - Identify and obtain suitable datasets for document summarization and non-factoid question answering.
  - Preprocess the data, including tokenization, padding, and handling special tokens, to prepare it for model training.
  - Implementing Hybrid Summarization Architecture:
  - Design and implement the hybrid deep learning architecture for document summarization (Objective 1).
  - Fine-tune the pre-trained models (BERT, GPT, Longformer) on the CNN/Daily Mail dataset or a similar dataset for document summarization.
  - Evaluate the performance of the hybrid summarization model using ROUGE and BLEU scores.

Week 4:
- Evaluating the Hybrid Summarization Model:
  - Conduct a human evaluation to assess the coherence, relevance, and information completeness of the generated summaries.
  - Calculate inter-annotator agreement (Fleiss's kappa) to validate the human evaluation results.

Week 5-8:
- Designing and Implementing Non-factoid QAS Architecture:
  - Design the QAS architecture, incorporating the hybrid summarization model from Week 5-6 (Objective 3).
  - Implement the question processing module, document retrieval module (DPR), and the summarization module using PyTorch and HuggingFace Transformers.

Week 9-10:

- Evaluating the Non-factoid QAS System:
  - Evaluate the performance of the QAS system on the MS MARCO dataset or a similar dataset.
  - Use metrics like Precision@K, Recall@K, F1-Score, and BLEU for quantitative evaluation.
  - Conduct a human evaluation to assess the quality and correctness of the answers generated by the QAS system.
  - Analysis and Comparison with Benchmark Models:
  - Compare the performance of the developed QAS system with state-of-the-art benchmark models based on quantitative and qualitative evaluation results.
  - Analyze areas of improvement and potential future directions for the research.
  - Make final adjustments to the project based on any feedback received during the evaluation and review process.