COSE474-2024F: Final Project Proposal
Llama-based Text Summarization Model

Park Subeen

1. Introduction

In recent years, large-scale pre-trained language models have demonstrated powerful performance across various natural language processing tasks. This project aims to explore automatic article summarization with Llama, using CNN/Daily Mail dataset. Additionally, its performance will be compared with other state-of-the-art (SOTA) models such as BART and Pegasus.

2. Problem Definition & Challenges

Text summarization is the task of generating brief and concise summaries from lenghty documents while preserving the key information. The challenge in article summarization is to maintain the core content without distorting the fact delivered by the original article.

3. Related Works

BART and Pegasus are two popular SOTA models for summarization tasks. BART is a denoising autoencoder model designed for sequence-to-sequence tasks, while Pegasus is specifically optimized for abstractive summarization. Although these models achieve strong performance, this project will investigate how Llama performs in comparison to these two models.

4. Datasets

The CNN/Daily Mail dataset will be used for training and evaluating the model. It contains over 300,000 news articles along with human-written summaries, offering diverse categories and varying levels of complexity are also diverse. Therefore, it is well-suited dataset for evaluating the performance of Llama on article summarization tasks.

5. State-of-the-art Methods and Baselines

We will compare the performance of Llama with that of SOTA models such as BART and Pegasus. Metrics such as ROUGE will be used to assess the quality of the generated summaries in terms of precision and recall. Additionally, the efficiency of each model in terms of computational cost and parameter size will be analyzed.

6. Schedule

- **Week 1-2**: Literature review and understanding existing SOTA models (BART, Pegasus) and their methodologies.
- **Week 3**: Data collection and preparation of the CNN/Daily Mail dataset.
- **Week 4**: Implementing the Llama model for summarization tasks.
- **Week 5-7**: Training the Llama model and evaluating its performance.
- **Week 8-9**: Comparing Llama's results with SOTA models and analyzing efficiency of all models.
- **Week 10**: Finalizing the project report.