Transformer-based Text Summarization: Scoring and Limitations

Team Members

David Gary	801325583
Michael Smalenberger	800984973
Matthew Seman	801156143
Ryan Marsh	800552800

Models & Datasets

XLNet Model
Reddit TIFU Dataset
Multi-News Dataset
S2ORC Dataset

Introduction

Transformers allow us to vectorize text and analyze the relationships between words and sentences. This technique allows for more powerful and larger language models to be created. Accuracy assessments through expert review are often unavailable for highly-contextual text summarization, so standardized evaluation schemes like Pyramid or ROUGE are used. A growing body of literature focuses on assessing these metrics' efficacy and limitations. Our goal is to review the literature and summarize the findings of these studies. In addition, we aim to apply the XLNet transformer model to various datasets from distinct contexts and thoroughly review the cost and limitations of generalizable text summarization.

Our proposed deliverables and the subsequent system will perform text analysis and summarization, following the current standards of generalizability. Anyone can use this system to summarize or assess the tone of text input. Since the point here is to show the limitations of generalized text summarization models, perfect accuracy is not the goal of our system. Users should beware of this and, as a result, only utilize this tool for interpretation guidance.

Deliverable 1: Length-Agnostic Sentiment Analysis

Our first deliverable will have a foundational ability to assess the general sentiment of an input text. We will utilize the XLNet transformer model as a backbone and employ some of the length-agnostic methods shown in the Guan, Smetannikov, and Tianxing paper[1].

Use Cases:

1. Automate customer feedback interpretation from collected surveys.
2. Assessing the emotional intensity of messages without context.
3. Mining general public opinion on a specific topic from public forums such as Reddit.
4. Score attempts at persuading individuals after targeted advertising.

Deliverable 2: Text Summarizer

Next, we will build on the XLNet framework to produce a generalized text summarizer. This will include a score report that shows how well the model performs according to the Pyramid and ROUGE schemes.

Use Cases:

1. Create a highlights reel of current news articles.
2. Automatically summarize chapters from a book.
3. Distill user messages for chatbots to understand more easily.
4. Extract key information from a medical patient's electronic health records.

References

[1] Guan, W., Smetannikov, I., and Tianxing, M. Survey on automatic text summarization and transformer models applicability. In *2020 International Conference on Control, Robotics and Intelligent System* (New York, NY, USA, 2020), CCRIS 2020, Association for Computing Machinery, p. 176–184.