INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

Emotion-Aware Summarization of Fiction Books

Final Report

by

Dilshan Boteju

W1850877/20210085

Supervised by

Ms. Lakna Gammedda

Submitted in partial fulfillment of the requirements for the MSc in Advanced Software Engineering degree at the University of Westminster.

Abstract

Fiction books are not just a straightforward telling of events; they are the author's imagination of those events, whose sentiments and ideas frequently appear in the book's story. Emotion-aware summarization of fiction books is a challenging task that aims to condense the narrative content of a book while preserving its emotional essence. The emotional tone in the summary can increase the book's selling rate.

Initially, This approach employs a robust variant of the BERT architecture(RoBERTa) to identify emotional sentences in fiction text. The Propose model is trained on emotion-labelled data to identify sentences that convey diverse emotional nuances, ensuring that the extracted emotional content conforms to the reader's expectations. The extracted emotive sentences are then fed into a Sumy pre-train summarization model for extractive summarization. These selected segments are combined to form a coherent and emotionally aware summary.

The author has posited that augmenting the quantity of emotional sentences may enhance the system's ability to generate emotional summaries. However, it was observed that this modification did not yield any discernible improvement in the final system. While the evaluation of this experiment did not demonstrate an enhancement in book summarization based on several established scoring methodologies, it succeeded in generating summaries that incorporated a more significant number of emotive terms while upholding the overall summary quality.

Keywords: Emotion-Aware, Extractive Summarization, Abstractive Summarization, Natural language processing

Subject Descriptors

- Computing methodologies → Artificial intelligence → Machine learning
- Computing methodologies → Machine learning → Natural language processing
- Computing methodologies → Machine learning → Natural language processing → Classification algorithms