



**INFORMATICS
INSTITUTE OF
TECHNOLOGY**

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

**A Novel Abstractive Text Summarization approach for Lecture
Content**

A Dissertation by

Mohamed Boosary Fathima Farheen

Supervised by

Mr. Prasan Yapa

Submitted in partial fulfillment of the requirements for the BEng (Hons) in Software
Engineering degree at the University of Westminster.

July 2022

ABSTRACT

In the current technological era, with the advancement of information and communication technology, textual information can be found in abundance in a variety of formats for different topics such as news articles, educational transcripts, medical documents, legal documents, etc making it difficult for user(s) to comprehend and extract the significant information with ease in a short period. Hence it needed a technique to be able to condense this many textual information into a summary that concisely includes the salient contents of the main source. The field of automatic text summarization (ATS) which has been in existence since the 1950's is anticipated to find solutions to this issue by automating the process of concise summary creation. However, the machine-generated summaries are still far from human-generated summaries as the machine-generated summaries mostly suffered from issues such as repetition of words, not being able to handle out of vocabulary words (OOV), less semantic, etc.

The research study focused on designing and developing a system to generate concise coherent summaries for lecture video content by utilizing lecture video transcripts for that the system employed an adversarial process for abstractive text summarization which adopts a Generative Adversarial Network (GANs) architecture that utilized a transformer-based pointer generator network for the generator component and CNN based text classifier for the discriminator to address the research gap that identified within the existing literature. The author was able to achieve satisfactory result when compared to other international researchers in abstractive text summarization domain within given the limited resources and time frame.

Keywords: Natural Language Processing, Automatic Text Summarization, Extractive Text Summarization, Abstractive Text Summarization, Pointer Generator Network, Transformers, Generative Adversarial Network.