**REPORT**

This report is for the assignment of the Automated text summarization.

The aim of this project was to develop an automated text summarization model that can extract key points from large volumes of text data, which used various aspects of Natural Language Processing (NLP).

I have made a model that uses the Bidirectional Auto-Regressive Transformers method and the well-known PEGASUS method for making a model that can summarize the text as I was most comfortable using these models and, they had a good reputation for being the best models for the given problems.

I have also implemented the BLEU metric and ROUGE metric to evaluate the models and their efficiency.

I have given text headings to all the parts of my python notebook which explains my model and what I was trying to achieve as the output at every important step.

**CHALLENGES**:

DATA PRE-PROCESSING: Cleaning and preprocessing the data to ensure consistency and quality added complexity to the project.

Model Training and Fine-tuning: Training and fine-tuning the BART and Pegasus models on the available dataset were time-consuming and resource-intensive tasks. Ensuring convergence and achieving satisfactory performance demanded careful hyperparameter tuning and experimentation.

Implementing different metrics: I had to learn from scratch the way to use rouge metric and understand its working and implement it successfully in my assignment.

My approach to the problem was to implement the two best known techniques of text summarization and crosscheck their performance and then proceed with the best method at the time.

The result of my model shows that the Pegasus model is the best technique that can give the summary of large texts and maintain good efficiency and quality of output as the rouge scores of Pegasus were very high compared to BART.

**APPLICATIONS OF MY MODEL**

The potential application for a model as such as mine would be to summarize textbooks into a smaller condensed notes mainly for students to learn from instead of going through their whole textbooks.

Another application of my model would be to summarize a project plan which tend to go for many months with various stages of developments etc. into a small summary which can be presented to senior executives so as to not waste time and also make them understand the project workflow.