



الجامعة الإسلامية العالمية ماليزيا
INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA
يُونِيسْكَتِي إِسْلَامُ إِنْتَارَا بَغْسَا مِلْسِيَا

Garden of Knowledge and Virtue

KULLIYAH OF INFORMATION & COMMUNICATION TECHNOLOGY

**CSCI 4342 NATURAL LANGUAGE PROCESSING
SEMESTER 1, 2020/2021**

**Project Proposal
(Text Summarization)**

PREPARED BY: GROUP Alpha

NAME	MATRIC NO.
Abdella Mame Abdo	1714883
Islam MD Shariful	1720601
Mohamed Moubarak Mohamed Misbahou Mkouboi	1820705

LECTURER

DR. SURIANI BT. SULAIMAN

Project Title: Text summarization Using Natural Language Processing

- **Project Background**

Text summarization is one of the NLP is the process of summarizing the information in large texts for quicker consumption and it's one of the applications of natural language processing and is becoming more popular for information condensation. Text summarization is a process of reducing the size of the original document and producing a summary by retaining important information of the original document. The process of text summarization includes reducing the size of an original document while preserving its information content and its summary is less than half of the main text.

- **Objective of Project**

- To understand the concepts of natural language processing and creating a tool for text summarization.
- To understand past works related to suspicious activities.
- To apply the Tf-Idf model to the problem of multi -document summarization.
- To understand, concentrate on creating a tool which automatically summarizes the document.

- **Proposed Method and Dataset**

In this project we are trying to summarize the text by using real data. So that summarization can give more or less accuracy to analyze the real outcomes.

Data collection: we will collect some articles and news from online sources.

Tools: we will use google colab and Jupyter notebook.

Model: we will use some approaches such as position of sentence, top down approaches, retain corpus, tf-idf method etc, for text summarization.