**Rida Amir (21L–5184)**

**Hamza Amer (21L-5368)**

**Haisem Naeem (21L-7528)**

**AI-Based Text Summarizer**

**Introduction:**

The purpose of this project is to develop an artificial intelligence-based text summarizer capable of condensing lengthy documents or articles into concise summaries while preserving the main points and key information. This tool will assist users in quickly extracting relevant content from large volumes of text, saving time and improving efficiency in information consumption.

**Data Collection:**

Obtain a diverse dataset consisting of various types of documents, articles, and texts from different domains, such as news articles, research papers, blog posts, and product descriptions. Ensure that the dataset includes both short and long texts to train the summarization model effectively.

**Data Preprocessing:**

Clean and preprocess the text data by removing special characters and punctuation marks. Tokenize the text into sentences or paragraphs and perform lemmatization or stemming to normalize words. Additionally, remove stop words and irrelevant phrases to improve the quality of the summaries.

**Building the Model:**

Choose an appropriate machine learning or deep learning architecture for text summarization, such as Recurrent Neural Networks (RNNs), Transformer models like BERT, or sequence-to-sequence models with attention mechanisms. Train the model on the preprocessed text data using techniques like extractive or abstractive summarization, depending on the chosen architecture.

**Evaluating the Model:**

Evaluate the performance of the text summarizer using standard evaluation metrics such as ROUGE (Recall-Oriented Understudy for Gisting Evaluation) scores, which measure the overlap between the generated summary and reference summaries. Additionally, conduct human evaluation tests to assess the readability, coherence, and informativeness of the generated summaries.

**Deployment:**

Save the final trained model to a file format compatible with deployment frameworks. Develop a user-friendly web application or API that accepts input text and generates a summary output. Deploy the application or API on a cloud platform for public use.