**Overview**

It's critical to keep up with the most recent news and developments in the fast-paced world of today. However, the sheer volume of information available can often be overwhelming, making it challenging for individuals to efficiently consume relevant news content. The AI News is a comprehensive solution for staying informed.

The AI News Summarizer is a web application designed to streamline the process of accessing and digesting news articles. By leveraging AI and LLM capabilities, the application offers users a range of features to facilitate easy and efficient news consumption. This includes accessing top articles by category, obtaining articles based on a search term or summarizing article content given their URLs.

**Problem Statement**

The rapid growth of information on the internet has led to information overload for many individuals. As a result, keeping up with the latest news and developments in various fields has become increasingly challenging. Traditional news consumption methods, such as browsing multiple websites or reading lengthy articles, can be time-consuming. Additionally, the abundance of news sources and articles makes it difficult for users to identify the most relevant piece of information. As a result, there is a pressing need for a solution that can aggregate, summarize, and deliver news content in a concise and accessible format.

The AI News Summarizer project addresses these challenges by providing users with a user-friendly platform to access and digest news articles quickly and efficiently. By leveraging AI algorithms and natural language processing techniques, the application aims to streamline the news consumption process and enables users to stay informed without information overload.

**Methodology**

The AI News Summarizer project utilized a variety of tools and APIs to implement its functionalities, including the Azure OpenAI API and the Groq API for leveraging the capabilities of Large Language Models (LLMs). The implementation involved several key components:

1. **Data Scraping and Cleaning**:

Python scripts were developed to scrape data from major newspapers and ensure the obtained data is clean and formatted properly. These scripts interface with the newspapers' websites and extract relevant news articles. The clean data is then passed to the LLM for further processing.

1. **Streamlit Web Application**:

The **app.py** Python script serves as the backbone of the Streamlit web application. It integrates the functionalities of the data scraping scripts and provides a user-friendly interface for accessing news content. The individual newspaper scraping scripts are imported as classes and utilized within the web application.

1. **Category and Keyword Search**:

Users can browse news articles based on broad categories such as sports or technology. The application routes the user to one of the existing newspapers to provide summaries of the top 5 articles from the selected category. Additionally, users can search for specific keywords or topics using the Google News scraper, which fetches the top 5 results from Google News.

1. **URL Summarizer**:

A separate Python script handles the URL summarization functionality using the Groq API. Users can input the URL of a news article, and the script analyzes and summarizes the content, providing users with a concise overview of the article.

**The Groq API** utilizes Language Processing Unit (LPU) technology. This LPU is known for its exceptional speed and efficiency, particularly when working with large language models like Mixtral-8x7b-Instruct-v0.1 and Llama-70b. By incorporating the Groq API into the project, the goal is to significantly improve the effectiveness of AI processing. This will enable real-time experiences with minimal delay, which is crucial for tasks that involve understanding and generating natural language

**The Azure OpenAI** Service offers REST API access to a range of language models, including GPT-4 and GPT-3.5-Turbo, enabling tasks such as content generation, summarization, and semantic search. In this project, Azure OpenAI API is utilized for text summarization tasks and used for generating embeddings from text content.

1. **FastAPI**:

The FastAPI framework is utilized to deploy the URL summarizer functionality, which enables the summarization of articles based on URLs and returns the summary in JSON format, enabling seamless integration with other applications and services.

1. **The Chainlit chatbot:**

Utilizes various libraries and APIs to develop a chatbot capable of summarizing text from URLs and answering questions. It utilizes Langchain, Chainlit, Groq, and FAISS that is a vector storage library, to efficiently handle large-scale text embeddings and facilitate fast retrieval of relevant information during question answering. The bot extracts text from a URL, generates embeddings via Azure OpenAI, sets up a retrieval-based question answering system, and enables users to interact with the content.