Project Proposal

Project : Article Summarization and Categorization

Team Members :

Divya Geethanjali Birudharaju

Jaya Prakash Reddy Gade

Ayyappa Ankishetty

Goals and Objectives :

- **Motivation :** The motivation behind this project is, we find few articles on Instagram, Blog or any social platform but it would be difficult for us to find what is in the article before reading it. Sometimes, after seeing the length of the article, we lose interest and stop reading it. But, if the article is about one of our interests, we might miss reading it. So, if there is a tool which categorizes and summarizes articles, it would be a time saving for most of the people. We can read the article if it is about our interested topic. We can skip reading the article if it is not in our interested category.

- **Significance :**
  - Summarization helps individuals to ignore unimportant information in the text.
  - Categorization helps people to identify what is the article about.
  - Summarization and Categorization both can help lot of people, especially students to read the articles which are related to the topic they want to learn about.

- **Objectives :**
  1. To create a model which summarizes and categorize articles.
  2. Train the model with the existing data of article summarization and categorization
  3. To achieve a model with a good accuracy.
  4. To create a user-friendly UI which is connected to the backend using API's.

- **Features :**
  - Text Classification
  - Text Summarization
  - Stop word Removal
  - Text lemmatization or stemming
  - Keyword Extraction

GitHub Link : https://github.com/jayaprakashreddy007/NLP

References :

- https://medium.com/analytics-vidhya/automated-keyword-extraction-from-articles-using-nlp-bfd864f41b34
- https://monkeylearn.com/text-classification/
- https://towardsdatascience.com/understand-text-summarization-and-create-your-own-summarizer-in-python-b26a9f09fc70