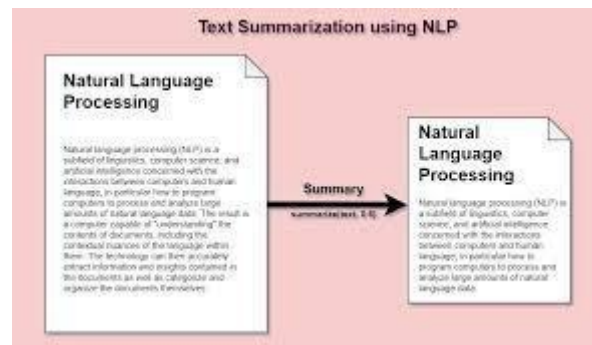


**Title:**

Text Summarization using NLP.

**Project Statement:**

Text Summarization focuses on converting large bodies of text into a few sentences summing up the gist of the larger text. There is a wide variety of applications for text summarization including News Summary, Customer Reviews, Research Papers, etc. This project aims to understand the importance of text summarization and apply different techniques to fulfill the purpose.

**Outcomes:**

The model should be able to summarize the text given in the documents.

**Modules to be implemented**

- Data Collection – Images related to any activity
- Data Exploration (EDA) and Data Preprocessing
- Explore types of Text Summarization and applications
- Create models using different NLP approaches
- Performance metrics
- Presentation and Documentation

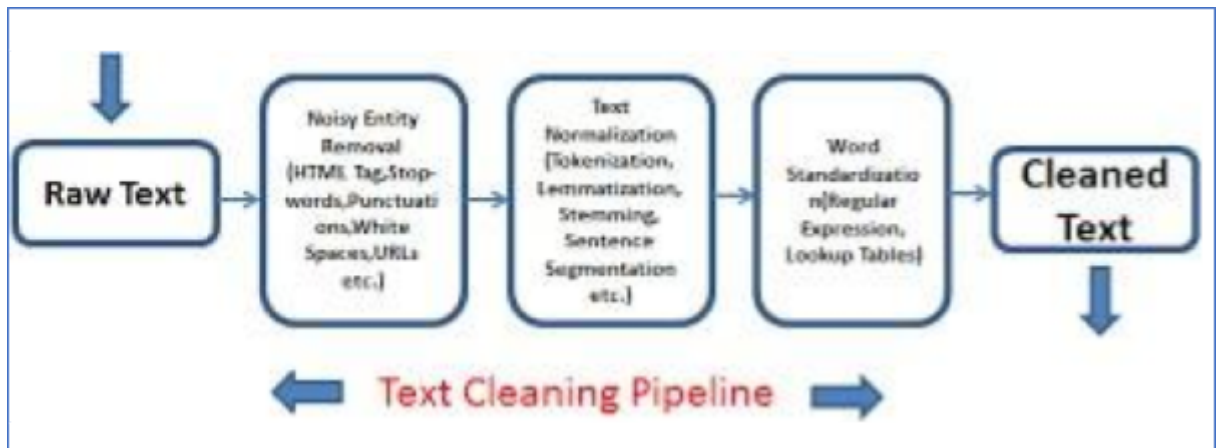
**3. Week-wise module implementation and high-level requirements****Milestone 1: Weeks 1-2****Module 1: Data Collection**

- Understand the problem statement
- Collect data from multiple sources and collect them.
- Finalize the dataset to be used

**Module 2: Data Exploration and Data Pre-processing**

- Cleaning text data
- Tokenization
- Normalization (Stemming & Lemmatization)
- Removing stop words
- Parts of speech tagging

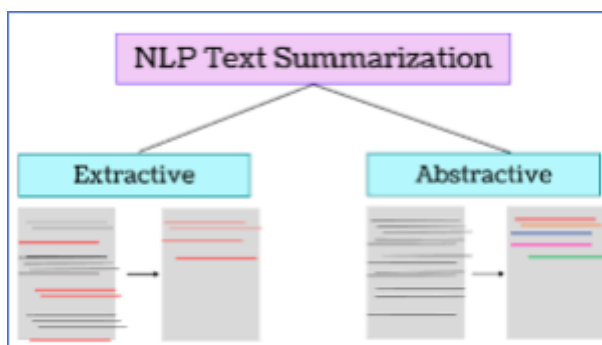
- Clean the dataset



## Milestone 2: Weeks 3-5

### Module 3: Explore Text Summarization types and its applications

- Extractive Summarization.
- Abstractive Summarization



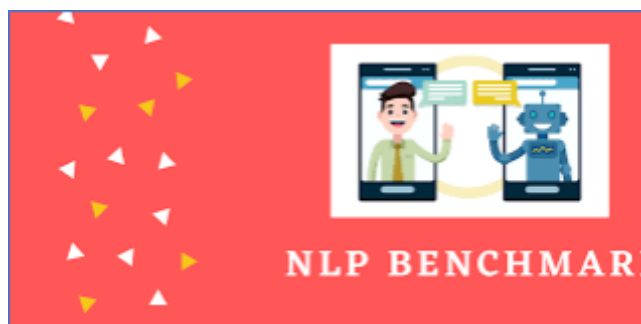
## Milestone 3: Weeks 6-7

### Module 4: NLP Models for Text Summarization.

- Explore ways to generate summarized text.
- Modules like nltk, spacy, huggingface, etc

### Module 5: Performance Metrics

- Evaluate models created using performance metrics like BLEU score
- Finalize the model with the best metrics



## **Milestone 4: Weeks 8**

### **Module 5: Presentations and Documentation**

- Prepare a presentation that must include the details of the problem statement, details of the data collected, data preprocessing methods and their outcomes, model-building methodology, Hyperparameters, performance metrics, and recommendations based on the outcome.
- The project document should capture the same topics mentioned above in a more detailed format.

### **Evaluation Criteria:**

#### **Milestone 1 Evaluation (Week 1-2):**

- Approval on the master dataset to be used.
- Analysis performed on the master data.
- Approval of the data preprocessing techniques.
- Approval of data treatments performed on the data.

#### **Milestone 2 Evaluation (Week 3-5):**

- Approval on deep learning architectures and Models to be used on the master dataset.

#### **Milestone 3 Evaluation (Week 6-7):**

- Approved approach to create Text Summarization
- Completion - Performance Metrics on all the built Models.

#### **Milestone 4 Evaluation (Week 8):**

- Approved Final Model.
- Approved Presentation and Project Documentation.
- Approved Remediation/Action plans for the Business.
- Final Code Submissions on GitHub.