Stack Overflow Keyword Identification

Team Members: -

- Achyutanand Mishra (11542607)
- Manohar Kowthavarapu (11554625)
- Rama Mamidi (11551417)
- NaveenRaja Yalagandula (11521505)

Goals and Objectives: -

Motivation: -

The goal of this project is to help determine the precise corpus from which to find the relevant keyword given in the input, making it easy for the user to acquire results. As a student or developer, we require assistance from websites to fix errors and learn new things. If you simply search for just any keywords on the Stack Overflow program, you will be redirected to the precise output for the reference.

Significance: -

A large amount of data is available on the Stack Overflow website, where numerous individuals are looking for answers to their queries. We apply this tagging technique to the user-posted questions in order to improve the effectiveness and ease of this search. This decreases overall user interaction with the application, making it more user-friendly.

Objective: -

- Each person has a unique style of expressing a query, which may not be understood by another user. Therefore, we may summarize the query into a brief, understandable statement.
- User searches are made simple using tags.
- Through tags, a member of the application who want to respond to the question can gain some understanding of it.

Example: - If a query related to python was posted, it is simpler for a member to decide whether to focus on a question when they receive a message that a question based on Python has been posted.

Features: -

The process of locating pertinent words and phrases inside unstructured material is made easier by the use of keywords. You may save a ton of time by automating activities like labeling incoming survey replies or reacting to urgent customer inquiries using keyword extraction.
For automated keyword extraction, you can employ a variety of methods. Using straightforward statistical methods that count word frequencies to identify keywords.
Word frequency, Word collocations and co-occurrences, TF-IDF, Linguistic approaches, Machine learning approaches.

References: -

https://www.kaggle.com/code/niyamatalmass/texts-summarizing-with-the-help-of-spacy

https://towardsdatascience.com/keyword-extraction-process-in-python-with-natural-language-processing-nlp-d769a9069d5c

https://scikit-learn.org/stable/modules/classes.html#module-sklearn.feature_extraction

https://scikit-learn.org/stable/modules/classes.html#sklearn-metrics-metrics

GitHub links: -

https://github.com/Achyut2995/StackOverflowKeywordidentification
https://github.com/mkowthavarapu
https://github.com/Srujana390
https://github.com/NaveenrajaYalagandula/NaveenrajaYalagandula.git