Project Title:

Youtube video title classification

Project Description:

We will be classifying titles based on categories and will predict category base of input of a title.

Dataset Source:

- **Youtube**
- **Kaggle**

Data Scraping Tools:

Data collection will be performed using the following tools:

- **Python Libraries**:
    - `Selenium: For extracting youtube titles.`

- # Preprocessing:

    Cleaning: Removing irrelevant characters, standardizing format.

    Tokenization: Breaking titles into words or phrases for analysis.

    Normalization: Converting titles to a standard form (e.g., lowercasing).

● Classification Model:

Model Selection:
Exploring different models like Naive Bayes, SVM, Neural Networks.
Training:
Splitting data into training and test sets.
Training the model on labeled data.
Validation and Testing:
Cross-validation techniques.
Performance metrics (accuracy, precision, recall).