STAGE 1 : Project Proposal

I. Describe what data is stored in the database. (Where is the data from, and what attributes and information would be stored?)

https://www.kaggle.com/datasets/rsrishav/youtube-trending-video-dataset?select=US_youtube_trending_data.csv

Dataset Name: US YouTube Trending Video Data

Source: Kaggle Description: The dataset comprises data from several months of daily trending YouTube videos specific to the United States. It encompasses a diverse range of content with up to 200 trending videos identified daily.

Key Attributes Include:
Video Title: The title of the trending video.
Channel Title: The name of the channel publishing the video.
Publish Time: The timestamp indicating when the video was uploaded.
Tags: Associated metadata tags for the video.
Views: Total number of views the video has received.
Likes & Dislikes: The number of likes and dislikes the video has garnered.
Description: A brief overview of the video content.
Comment Count: The total number of comments made on the video.

II. What are the basic functions of your web application? (What can users of this website do? Which simple and complex features are there?)

The simple features of our website will consist of functions the user can interact with. Users will be able to add filters to sort trending videos on features including comments, comments disabled, rating disabled, and our website will display a data grid of all the trending Youtube videos sorted by the filters added. Users will also be able to search for keywords or tags. Searching keywords/tags will show a list of the most popular videos and what is written in their description/tags + other information as well as how commonly used those tags/keywords are used in trending videos.

The complex feature of our website will allow users to save any keywords they find are very popular or commonly used. Another complex feature is a real time visualization of the data. The visualization will consist of graphs that show distributions of the most popular tags/keywords as well as things like the average number of comments and average percentage of likes trending videos receive. The dataset is updated daily, so our visualization will also be updating everyday.

III. What would be a good creative component (function) that can improve the functionality of your application? (What is something cool that you want to include? How are you planning to achieve it?)

A good creative component can be the use of a Machine learning model to predict the number of views a data might get. To elaborate, we will let users put down information related to their videos.

Users may provide two different types of data for analysis:

True Data: Real-time measurements allow established content producers to assess expected performance and gain understanding of how their work is received by a wider audience.
Hypothetical Data: This feature enables users to enter simulated data to estimate possible traction, assisting in goal-setting and tactical planning for aspiring YouTubers or those testing content strategy.

By creating a ML model, our web application will not only cater the basic needs, but will serve as a useful approach to gather traffic on our website.

Project Title

YouPredict: Predicting YouTube Video Trending Dynamics

IV. Project Summary: It should be a 1-2 paragraph description of what your project is.

Our project YouPredict: Predicting YouTube Video Trending Dynamics is a web application that will help deliver the user with data and insights on how the past videos that have trended on YouTube in various different categories within the United States of America. It leverages robust data sourced from Kaggle. This data includes various information on categories, publishing time, the date the video trended on, the tags associated with the video, the view count, number of likes, dislikes, comments, description and other video related features.

Our Platform offers a detailed view of daily trending videos encompassing diverse content across various different categories. Users can use advanced filtering options and search functionalities to discover the most trending videos, along with the description and tags. Our application's use of cutting-edge machine learning techniques sets it apart from others. Our algorithm will forecast the prospective viewership based on the user-inputted preliminary information

about their videos, providing a priceless tool for content producers and strategists alike.

V. **Description** of an application of your choice. State as clearly as possible what you want to do. What problem do you want to solve, etc.?

Our application will allow the use filters to sort the most popular trending videos to see what kinds of videos are popular. The application will also let users to search up tags and keywords to see how often they show up on trending videos. The user will also be able to store tags/descriptions that they like or see are popular. Users will also be able to enter their own description, tags, and upload time, and our application will predict how many views the user will get based on past trending videos.

Problem Solving: Our application aims to solve the problem of clarity in Youtubers. As we mentioned earlier, our application will be a great source for Youtubers to get insight regarding their videos. By providing them search and advanced filters, users can see the entire data of top trending videos in a tabular form. In addition to this, our creative component, the Machine Learning prediction model will help aspiring Youtubers to get an idea on how many “views” he/she can achieve within a given time frame.

VI. **Usefulness**. Explain as clearly as possible why your chosen application is useful. Make sure to answer the following questions: Are there any similar websites/applications out there? If so, what are they, and how is yours different?

Our project will allow users to see what types of Youtube show up most of the trending page based on things like category, tags, descriptions, and other metrics. Our website will also take in an example description/tag and predict how many views, likes, and comments the video will get. This is useful because it will show users what type of Youtube videos are the most successful and can help them modify their Youtube videos to become more popular. Our project is most useful for Youtubers and aspiring Youtubers, who want to get on the trending page and become more popular.

There are some similar websites out there such as https://ahrefs.com/youtube-keyword-tool which gives information on how often a keyword was used. Our website is different from this one because ‘ahrefs’ gives information on all videos whereas we will only provide information on keywords used in trending/popular videos. Another notable difference is our website will

give more information on keywords such as view count, likes, number of comments, etc. Our website will also give predictions on how many views we can get based on a complete description or tags.

VII. **Realness**. Describe what your data is and where you will get it.

This dataset is a daily updated record of trending YouTube videos in different regions, where the “trending status” is determined by several features including likes, comments, shares and views. The dataset is derived from Youtube’s official data via the Youtube API, thus it comes from an authentic and reputable source.

Because there are sub-datasets from different countries - including the United States, India, and Korea - the data is comprehensible and extremely reflective of the trends in different nations. Additionally, the real-time aspect ensures that the data remains current and relevant. The availability of column metadata provides extra information about each column, which further aids users in understanding the dataset’s structure and content.

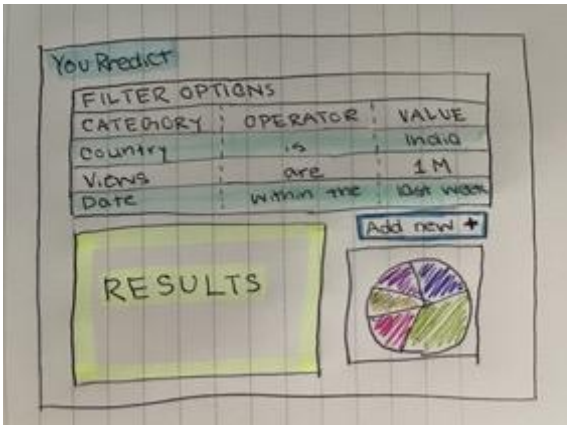
VIII. Description of the **functionality** that your website offers. This is where you talk about what the website delivers. Talk about how a user would interact with the application (i.e., things that one could create, delete, update, or search for). Read the requirements for stage 4 to see what other functionalities you want to provide to the users. You should include:

The website will help deliver the user with a variety of analytics regarding YouTube video trendings. This can be done using various data points and features such as number of likes, comments, among others. The user can add different filters and sort the trending videos based on their desired way of usage. The user can also search for particular items using a search box. Users also have the option to go through visualizations provided for better readability and interactions.

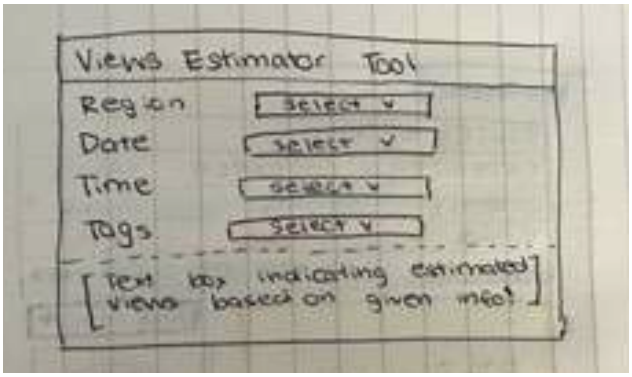
The user will also be able to save the keywords and descriptions they think that will be helpful to their own usage. This will help them for their own future usage. On top of this, users will be able to use our pretrained model that predicts the number of views the user's video can get based on a complete description, tags and different features. This will help them to get a better idea of what they are doing correctly with their own content and what can be done better in their future content.

A. **A low-fidelity UI mockup**: What do you imagine your final application's interface might look like? A PowerPoint slide or a pencil sketch on a piece of paper works!

Main webpage sketch



Functional tool sketch



B. **Project work distribution**: Who would be responsible for each of the tasks or subtasks?
List of the person responsible for which exact functionalities in section 6. Explain how backend systems will be distributed across members. Be as specific as possible as this could be part of the final peer evaluation metrics.

Our project work contribution would be 25% each:

Data acquisition and database Implementation: David & Daivaksh
Machine Learning Implementation: Daivaksh & Priyansh
UI/UX Design: Navya & David
Backend Implementation: Priyansh & Navya