**YouTube Trend Analyzer (Streamlit Web App)**

**Project Overview**

The YouTube Trend Analyzer is an interactive Streamlit web application designed to uncover insights from YouTube video metadata. This project leverages data analysis, visualization, and machine learning to help creators, marketers, and analysts identify what types of videos perform best on YouTube. By analyzing factors such as categories, tags, engagement metrics, and video metadata, the app provides actionable insights to optimize content strategies.

**Features**

* Data Upload & Processing
  + Supports CSV/JSON files containing YouTube video metadata.
  + Cleans, preprocesses, and structures data for analysis.

**Data Visualization**

* Bar charts and pie charts to highlight engagement across categories.
* Trend analysis for views, likes, comments, and subscriber impact.
* Keyword/tag analysis to identify high-performing topics.

**Machine Learning Model**

* Predicts whether a video may perform well based on input metadata.
* Uses supervised learning techniques for classification.

**Interactive Dashboard**

* Filter videos by category, creator, or engagement metric.
* Explore top-performing creators and content types.

**Deployment**

* Hosted on Streamlit Cloud for easy access via any web browser.

**Tech Stack**

Programming Language: Python

Libraries: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Streamlit

Version Control: Git & GitHub

Deployment: Streamlit Cloud

**Business Value**

* Helps content creators identify high-engagement categories and tags.
* Provides data-driven insights for marketers to design effective ad campaigns.
* Enables trend forecasting by analyzing engagement over time.
* Demonstrates end-to-end data science workflow from raw data to deployment.

**Project Process: YouTube Trend Analyze**r

**1. Problem Definition**

Identified the need to analyze YouTube video trends to understand which types of videos get higher engagement.

Defined project goals:

* Find high-performing categories, keywords, and creators.
* Build a predictive model to estimate video performance.
* Deploy an interactive dashboard for end users.

**2. Data Collection**

* Collected YouTube video metadata from a dataset (CSV/JSON).
* Data included features such as title, tags, category, views, likes, dislikes, and comment count.
* Ensured dataset variety across multiple video categories.

**3. Data Preprocessing**

* Cleaned missing values and removed duplicates.
* Processed categorical features (e.g., video categories).
* Extracted useful features such as:
* Engagement rate = (likes + comments) / views
* Word count in titles/tags
* Normalized numerical data for better model performance.

**4. Exploratory Data Analysis (EDA)**

* Used Matplotlib/Seaborn for visualizations:
* Bar charts showing most popular categories.
* Pie charts showing engagement share.
* Time-based trends for views and likes.
* Identified patterns: e.g., entertainment and music videos attract higher engagement.

**5. Machine Learning Modeling**

* Applied Supervised Learning (Classification/Regression) to predict video performance:
* Features: category, tags, views, likes, comment counts.
* Target: Engagement level (High / Low).
* Split data into train/test sets.
* Used Scikit-learn models (Logistic Regression, Random Forest, etc.).
* Evaluated performance using accuracy, F1-score, and confusion matrix

**7. Deployment**

* Created requirements.txt for dependencies.
* Uploaded project to GitHub for version control.
* Deployed the app on Streamlit Cloud, making it publicly accessible.

**End Result:**

A fully functional Streamlit web app that combines data science + machine learning + deployment to deliver actionable insights about YouTube trends.

The YouTube Trend Analyzer project demonstrates a complete end-to-end data science workflow, starting from data preprocessing and exploratory analysis, to machine learning modeling, and finally deployment as an interactive Streamlit application.

Through this project, I showcased my ability to:

* Work with real-world, unstructured data and extract meaningful insights.
* Apply data visualization techniques to highlight engagement patterns across categories and keywords.
* Develop and evaluate machine learning models to predict video performance.
* Deploy a fully functional web app on Streamlit Cloud, making insights accessible to non-technical users.

This project not only highlights technical expertise in Python, data analysis, visualization, and machine learning, but also emphasizes the importance of data storytelling and business value creation. Future improvements such as YouTube API integration, sentiment analysis, and recommendation engines will further enhance the app’s real-world applicability.