**Project Initialization and Planning Phase**

|  |  |
| --- | --- |
| Date | 15 JULY 2024 |
| Team ID | 739869 |
| Project Title | View count visionary:data driven approach to forecasting youtube videos views project |
| Maximum Marks | 3 Marks |

**Project Proposal (Proposed Solution) template**

This project proposal View Count Visionary - A Data-Driven Approach to Forecasting YouTube Video Viewse outline a solution to address a specific problem. With a clear objective, defined scope, and a concise problem statement, the proposed solution details the approach, key features, and resource requirements, including hardware, software, and personnel.

|  |  |
| --- | --- |
| **Project Overview** | |
| Objective | View Count Visionary - A Data-Driven Approach to Forecasting YouTube Video Views |
| Scope | - Collect and preprocess data  - Train and deploy machine learning model  - Develop user interface  - Evaluate model performance |
| **Problem Statement** | |
| Description | The problem addressed by "View Count Visionary" revolves around the uncertainty content creators, marketers, and analysts face in predicting the future popularity of YouTube videos shortly after their release. |
| Impact | -content creators  -marketers  -analysts  -industry competitiveness  -platform optimiztion |
| **Proposed Solution** | |
| Approach | -data collection and storage  -data processing and analysis  -real-time processing  -visualization and reporting  -machine learning integration |
| Key Features | -user interface and accessibility  -scalability and performance  -security and compliance  -automation and alerts  -integration capabilities |

**Resource Requirements**

|  |  |  |
| --- | --- | --- |
| **Resource Type** | **Description** | **Specification/Allocation** |
| **Hardware** | | |
| Computing Resources | Virtual machines with multiple cores (e.g., 8 cores) and sufficient processing power (e.g., Intel Xeon, AMD EPYC). | e.g., Virtual Machines: Each with 8 cores, 3.5 GHz Intel Xeon processor |
| Memory | Servers with 64GB or more of RAM to ensure smooth operation and quick access to data. | e.g., 64GB DDR4 RAM per virtual machine |
| Storage | Combination of fast storage (SSD/NVMe) for active data and larger capacity storage (HDD) for less frequently accessed data. | e.g., 1TB NVMe SSD for high-speed data access and processing |
| **Software** | | |
| Frameworks | -web framework  -big data framework | e.g., Use Django with PostgreSQL for building a robust web application to track and manage view counts, integrating with Apache Spark for data processing and analysis. |
| Libraries | -data analysis  -visualization  -machine learning | e.g., Utilize Pandas and Matplotlib to analyze and visualize historical view count data, identifying patterns and trends. |
| Development Environment | -IDEs(integrated development environments)  -version control  -containerization | e.g., Develop in Visual Studio Code, use Git for version control, and Docker for containerizing the application for seamless deployment. |
| **Data** | | |
| Data | -data storage  -data processing  -data visualization | e.g., Store user interaction and view count data in MongoDB for flexibility and scalability, process real-time data streams with Apache Kafka, and visualize insights with Tableau dashboards. |