**Project Design Phase-I**

**Solution Architecture**

|  |  |
| --- | --- |
| Date | 03 October 2023 |
| Team ID | NM2023TMID05047 |
| Project Name | The Future of Work : Glassdoor Jobs |
| Maximum Marks | 4 Marks |

**Technical architecture :**

Data collection: The first step in the process will be to identify key influencers in a particular industry or niche. This can bedone using various tools suchas LinkedIn search, third-party influencer databases ,and social media monitoring tools. Once the influencers have been identified, their posts, comments, and engagement metrics will be collected using Glassdoor API.

Data processing: The next step in the process will be to preprocess the collected data. This will involve cleaningandtransformingthedata,removingduplicates,andconvertingthedatainto aformat suitablefor analysis.ThepreprocessingstepwillbedoneusingPythonlibrariessuchas PandasandNumPy.

Data storage: The collected data will be stored in a database such as MongoDB or MySQL. The database will be optimized and efficient data retrieval.

Data visualization: Finally, the insights gained from the analysis will be visualized using tools such as Tableau or Power BI. The visualizations will be used to communicate the findings to stakeholders and informmarketing strategiesandcontentcreation.

Dataanalysis: The preprocessed data will thenbe analyzed using various data analytics techniques such as sentiment analysis, topic modeling, and predictive analytics. The analysis will be done using Python librariessuchasScikit-learnandNLTK.

Overall, the technical architecture of the project will involve a combination of tools and technologies to collect,store,process,analyze,andvisualizedatafromLinkedIninfluencers

The technical architecture of the project will involve several components, including data collection, data storage,data processing,and datavisualization.

**Architecture Diagram:**

