Define Problem / Problem Understanding

|  |  |
| --- | --- |
| Date | 09 October 2023 |
| Team Id | NM2023TMID05059 |
| Project Name | Data Titans: Unearthing Trends from LinkedIn Influencers |

**Business Requirements**

**1. Data Collection and Integration:**

* **LinkedIn Data Access:** Obtain necessary permissions and API access from LinkedIn to collect public data from influencers' profiles, posts, comments, and connections.
* **Data Integration:** Integrate LinkedIn data with Data Titans' analytics platform for seamless processing and analysis.

**2. Influencer Identification:**

* **Criteria Definition:** Define criteria (such as follower count, engagement rate, industry relevance) for identifying influencers within specific sectors.
* **Algorithm Development:** Develop algorithms to automatically identify and classify influencers based on defined criteria.

**3. Natural Language Processing (NLP) and Text Mining:**

* **Text Processing:** Implement NLP techniques to process textual data, including sentiment analysis, entity recognition, and keyword extraction from influencers' posts and comments.
* **Language Support:** Ensure language compatibility for multilingual influencers to cover a diverse range of content.

**4. Trend Analysis:**

* **Trend Detection:** Develop algorithms to identify trending topics, keywords, and themes within influencers' content.
* **Temporal Analysis:** Perform trend analysis over time to identify patterns, spikes, and recurring trends.

**5. Data Privacy and Security:**

* **Compliance:** Adhere to LinkedIn's terms of service and privacy policies while accessing and processing user data.
* **Anonymization:** Ensure that personally identifiable information (PII) is anonymized and not stored or shared without explicit consent.

**6. User Interface and Experience:**

* **Interactive Dashboard:** Create a user-friendly dashboard allowing clients to explore influencer trends visually.
* **Customization:** Provide options for clients to customize their dashboard views based on specific influencers, industries, or time frames.

**7. Reporting and Visualization:**

* **Visual Representation:** Generate visual reports, charts, and graphs to represent trends and insights clearly.
* **Export Functionality:** Allow clients to export reports and visualizations for further analysis or presentation purposes.

**8. Scalability and Performance:**

* **Scalability:** Design the system to handle a large volume of data and scalable to accommodate increasing data sources and clients.
* **Performance Optimization:** Optimize algorithms and processing pipelines for efficiency and real-time or near-real-time analysis.

**9. Documentation and Support:**

* **Documentation:** Create comprehensive documentation outlining the data sources, processing methods, and algorithms used for transparency and understanding.
* **Client Support:** Provide client training and support to interpret the generated insights and make informed decisions.

**10. Feedback Mechanism:**

* **Client Feedback:** Implement a mechanism for clients to provide feedback on the insights, ensuring continuous improvement based on user experience and requirements.