Feedback and Reflection:

During the completion of this assignment, I gained valuable insights into working with Elasticsearch and data preprocessing. However, there were certain aspects where I believe I could have optimized my workflow for better efficiency.

Areas of Improvement:

1. Data Cleaning Priority: Before initiating the indexing process, dedicating time to clean the dataset thoroughly would have been beneficial. This practice would have ensured that the data indexed in Elasticsearch was uniform and consistent.

2. Proper Salary Field Conversion: Accurately converting the salary field to a float type proved to be a challenging task. This issue consumed significant time and effort. In hindsight, implementing a robust approach for this conversion, coupled with effective debugging techniques, could have expedited the process.

3. Utilizing Multiple Datasets: While time constraints influenced my decision to use a single dataset, I acknowledge that incorporating all available datasets would have provided a more comprehensive representation of the compensation data. This could have led to more robust insights and a more thorough analysis.

4. Mapping Considerations: I initially underestimated the importance of mapping in Elasticsearch. It became evident that Elasticsearch requires explicit mapping, especially when dealing with various data types such as floats, strings, and null values. A more proactive approach to mapping could have prevented potential issues.

Time Management Reflection:

Completing the assignment took longer than anticipated. Approximately four extra days were needed to finish. Two days were spent on setting up Elasticsearch due to installation difficulties, and another two days were dedicated to resolving an SSL error. In hindsight, these challenges could have been efficiently addressed through team collaboration or prior experience with Elasticsearch.

I believe that addressing these areas of improvement will enhance my proficiency in future projects involving Elasticsearch and data preprocessing. This experience has underscored the significance of thorough preparation and consideration of data nuances.

Message To Recruiter:

Thank you for providing me with the opportunity to undertake this task. I greatly appreciate the chance to challenge myself and extend my capabilities. I am fully aware of the areas where my submission may have fallen short, and I am committed to overcoming these obstacles. My goal is to become a valuable asset to the company, contributing effectively to its success.

I am eager to collaborate with the team and leverage my knowledge and passion for technology to create valuable solutions at Sentisum. Once again, I express my gratitude for this opportunity and I look forward to potentially working together.

Thank you.

Hrithik Vashishtha