**CSE523 Machine Learning**

**Prof. Mehul Raval**

**Product Classification using their Ingredient**

**Week 6 Report**

| **Name** | **Enrolment Number** |
| --- | --- |
| Aharnish Pithva | AU2040022 |
| Jevin Jivani | AU2040051 |
| Astha Bhalodiya | AU2040067 |
| Yug Patel | AU2040181 |

**1) Tasks Performed in the week.**

During the past week, our primary focus was on data collection for our ML project, which involved utilising an API key to interact with ChatGPT. We started by understanding the documentation and guidelines provided by the API provider, and then we successfully obtained the API key required for authentication. We then integrated the API key into our project code and implemented the necessary API calls to fetch the data. This involved writing code to send requests, handle responses, and parse the data in a suitable format for further processing.

In addition to the technical aspects of data collection, we also researched and identified the relevant data sources that would provide the most valuable information for our ML project. We evaluated different data sources based on their reliability, relevance, and suitability for the project's objectives. This involved reviewing documentation, exploring data source options, and deciding which sources to utilise.

**2) Outcomes of the tasks performed.**

The outcomes of the data collection tasks successfully obtained the required data for our ML project. We retrieved data from the API using the provided API key and integrated it into our project code. The collected data was structured, allowing for further processing and analysis. We also identified and evaluated relevant data sources that provided valuable insights and information for our ML project. The outcomes of these tasks laid the foundation for the next steps in our project, which involve implementing cosine similarity.

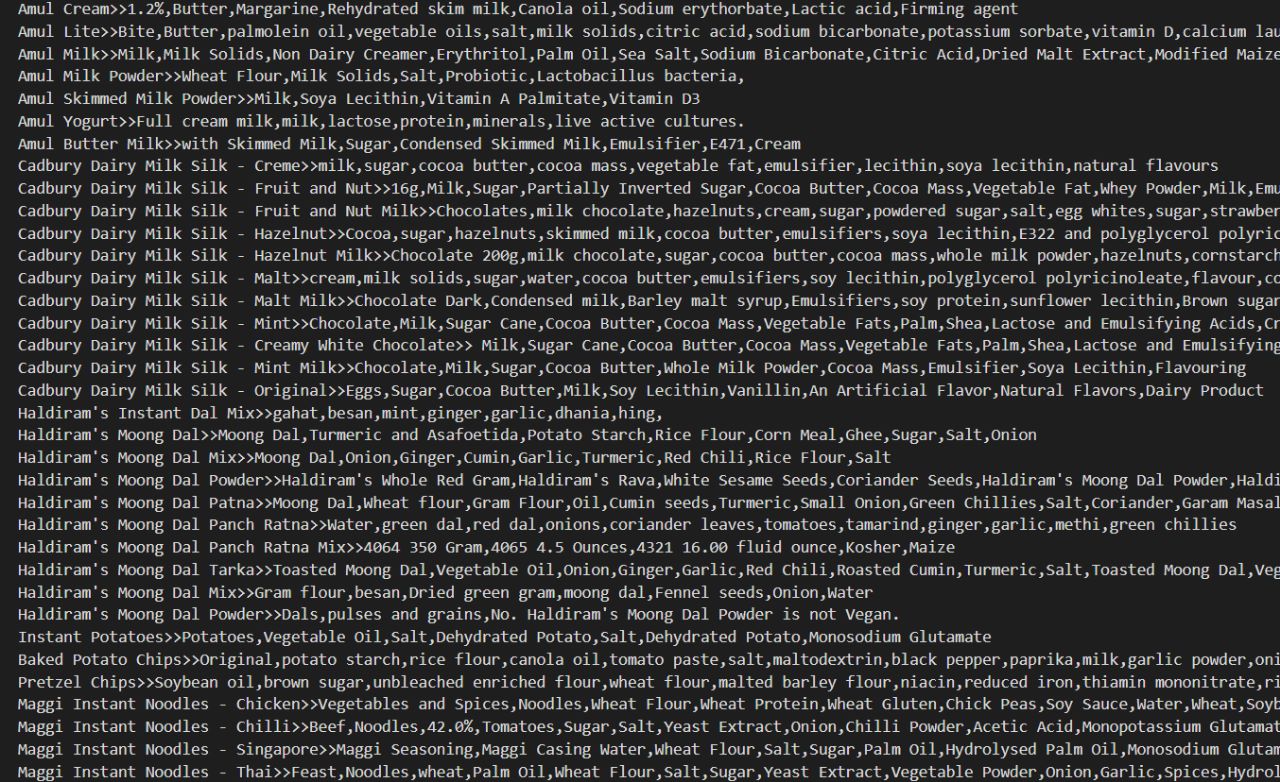
We then decided to pre-process the collected data to clean and transform it into a suitable format for similarity calculations. This may involve tasks such as data cleaning, data normalisation, and feature extraction, depending on the specific requirements of our ML project.

Results obtained:

Number of food products obtained: 565

Number of beauty products obtained: 539





**3) Tasks to be performed in the upcoming week.**

The next step in our ML project involves implementing cosine similarity. We plan to utilise this technique to measure the similarity between different data points in our dataset. This will allow us to identify patterns, trends, and relationships within the data, which will be valuable for our project's objectives.