

## Costco Wholesale – MSDS 25.4

### *Voice of Market (VoM): Unlocking Market Insights from Social Media*

#### **Project Description**

We are Costco's IT Data Science team, a dedicated group of data enthusiasts with a shared commitment to driving data-driven innovation. Our diverse backgrounds and expertise in machine learning, statistics, and data engineering converge to create a dynamic and collaborative environment. We are united by a common goal: to unlock the potential of data for Costco and deliver exceptional value to our members.

**Project Overview:** This project aims to analyze customer sentiment about Costco's products and services shared on public social media platforms. By leveraging sentiment analysis and advanced AI summarization techniques, the project will provide actionable recommendations through a web application to help Costco merchandise new products and services aligned with customer expectations. A summarization chatbot will also be developed to present key insights interactively to stakeholders.

#### **Outcomes:**

1. Identification of customer sentiment trends (positive, negative, neutral) for Costco products and services on social media.
2. Recommendations for new products and services based on customer demand and sentiment insights.
3. An intuitive and interactive tool that provides a knowledge graph view of products, product trends, key attribution tags with propensity scoring, social graphs and impressions, etc
4. Development of an AI-driven summarization chatbot for interactive reporting of VoC insights.

#### **Deliverables:**

1. Cleaned and segmented dataset of public social media data relevant to Costco.
2. Sentiment analysis report categorizing Costco products and services feedback.
3. Recommendations for new product/service opportunities with Knowledge graph with propensity scoring for each product and product family.
4. An AI-powered summarization chatbot prototype for presenting VoC insights.
5. Expecting code and end to end solution as a deliverable. (As a notebook)
6. We are looking for a reusable DS model/framework

#### **Stretch Goals**

1. Real-time sentiment analysis and monitoring dashboard.
2. Inclusion of trend analysis to identify emerging customer/market preferences.
3. Advanced topic modeling for deeper insights into customer expectations.
4. Building an interactive/Intuitive web application/tool to show insights.
5. Integration of the summarization chatbot with live social media streams.

#### **Required Student Skills**

##### **Data Collection and Preprocessing:**

- Web scraping and public API integration.
- Data cleaning and segmentation (NLP preprocessing).

##### **2. Sentiment Analysis and Modeling:**

- Proficiency in machine learning and natural language processing (NLP).
- Experience with sentiment analysis tools like VADER, TextBlob, or advanced transformers like BERT.

**Visualization and Reporting:**

- Experience in data visualization tools like Tableau, Power BI, or Python libraries (Matplotlib, Seaborn).
- Chatbot development using frameworks like Rasa or Dialogflow.

**4. AI and Automation:**

- Familiarity with text summarization techniques (extractive/abstractive).
- Integration of chatbot with interactive reporting mechanisms.

**Data Availability**

- **Source:** Public social media data (X, Reddit, Instagram, Facebook, etc.).
- **Collection Method:** Web scraping or public APIs for relevant platforms.

**Processing Steps:**

- Scrape or collect public posts mentioning Costco Products and Services. (Excluding Travel)
- Clean the data by removing duplicates, noise, and irrelevant content.
- Segment data by product/service mentions and geography if possible.

**Hardware/Software Requirements (Suggested)**

- **Data Collection:** Python libraries such as BeautifulSoup, Scrapy, Selenium.
- **Sentiment Analysis:** NLP libraries like NLTK, spaCy, Hugging Face Transformers.
- **Data Analysis and Visualization:** Pandas, NumPy, Matplotlib, Seaborn, Plotly.
- **Chatbot Development:** Rasa, Dialogflow, or OpenAI APIs for AI summarization.

**Project Management:** Agile tools like Jira, Trello.