AASD 4016 - Full Stack Data Science Systems

Amazon Review Insights

Created & presented by (Group 6):

Jai Vigneshwar Aiyyappan (101448497)
Vignesh Baskaran(101435847)
Ali Guneysel (101486766)
Madhoumithaa Veerasethu (101471815)

April 18, 2024

Table of Contents

Problem Statement

Product sellers face challenges in effectively analyzing and understanding customer feedback and sentiments regarding their products. They lack efficient tools to extract actionable insights from customer reviews, identify product strengths and weaknesses, and track sentiment trends over time. This limits their ability to optimize product offerings, address customer concerns, and enhance overall sales performance on the platform.
Amazon users encounter difficulties in quickly accessing comprehensive information and unbiased reviews about products of interest. They struggle to make informed purchasing decisions due to the overwhelming amount of data available and the lack of easily digestible summaries. They seek efficient tools that can provide concise summaries of product features and reviews.

Our solution

Our solution will automate the extraction of meaningful insights from product reviews

Key features include

1. Action items
2. Pros and Cons for each feature of the product
3. Sentiment analysis over time.
4. Visualizations such as word clouds and rating trends

Literature Review

Helium 10, AMZScout, and Sellics are softwares tailored specifically for Amazon sellers. These tools offer a wide range of features aimed at optimizing various aspects of selling on the Amazon marketplace. Key functionalities include product research, keyword research, listing optimization, and sales estimation. Additionally, they provide review analysis tools to help sellers understand customer feedback, monitor reviews, and make data-driven decisions to improve their products and enhance their overall performance on Amazon.

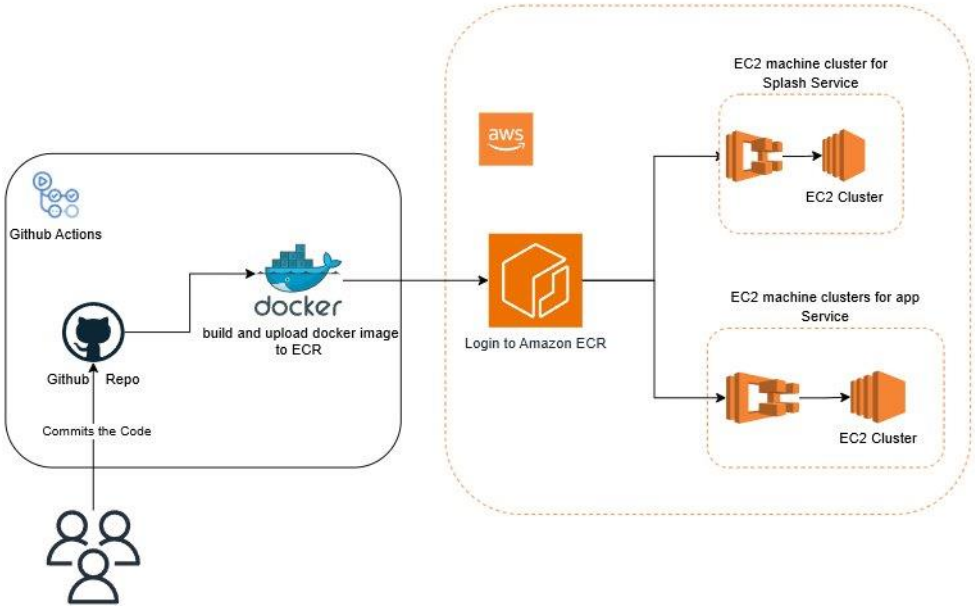
Drawbacks of existing solution:

Complexity
Seller-Centric Focus
Paid features

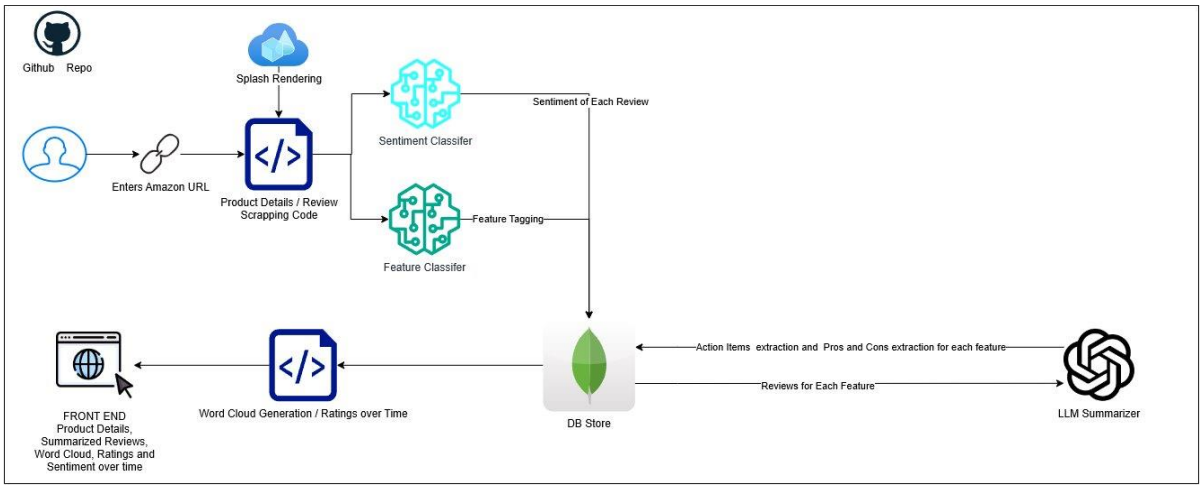
Benefits of Our Solution:

Condensed Insights
Feature-Based Summaries
User-Centric and Seller-Centric Focus

Architecture Diagram



Dataflow Diagram



Challenges and Issues faced

- Showing updates in real-time was needed, so we solved it using async-update model instead of traditional request-response model using Redis.
- Multi-container deployment was needed as we were using splash as our page rendering microservice.
- Latency problems exists as there were heavier processing involved. So, we incorporated both threading and concurrency using threading module and asyncio module thorough out all the steps.
- Issues with word cloud package while building docker image (Instead created word cloud in js)

- Limited Coverage and small dataset (500 reviews maximum) might lead to inadequate insights
- Product should have features to get insights
- Token size limitation in sentiment classification model. (512 characters)
- Asynchronous scrapping limitation in Amazon leading to increased scrapping time

GitHub Repository

Link to Amazon Review Insights Repository

Cloud deployed solution

Link to Cloud deployed solution

- We have mentioned a few links that can be tried out (available in db so it's faster)
- https://www.amazon.com/Sony-PlayStation-DualSense-Wireless-Controller-5/dp/B097YYGXY5/ref=sr_1_6?sr=8-6
- https://www.amazon.com/Lighters-Multipurpose-Windproof-Refillable-Fireplaces/dp/B0B58TB46B/ref=sr_1_45?sr=8-45
- https://www.amazon.com/iPhone-13-128GB-Pink-Unlocked/dp/B0BGQS8YV4/ref=sr_1_10?sr=8-10
- If using any new link : Please note that it might take a few minutes to generate the insights. Please use links available from Amazon.com and ensure that the product has aspect tagged to it above reviews
  ex: