

FSE570 - Data Science
Capstone Project

SENTIMENT ANALYSIS OF AMAZON FINE FOOD REVIEWS USING VADER AND HUGGING FACE MODEL ROBERTA

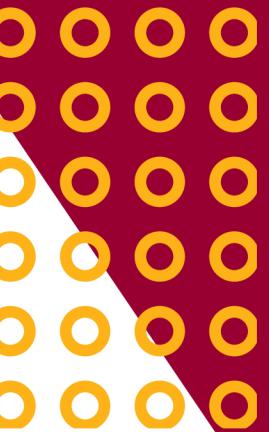
Team Pluto



ARIZONA STATE UNIVERSITY

ABOUT THE PROJECT

- Online product reviews provide valuable insights into customer experiences.
- However, analyzing sentiment in these reviews is challenging due to sarcasm, ambiguous language, and varying writing styles.
- Traditional methods struggle to capture the full meaning behind customer feedback.
- Objective: Improve decision-making for businesses by leveraging VADER (fast, rule-based sentiment analysis) and RoBERTa (deep learning-based contextual understanding).



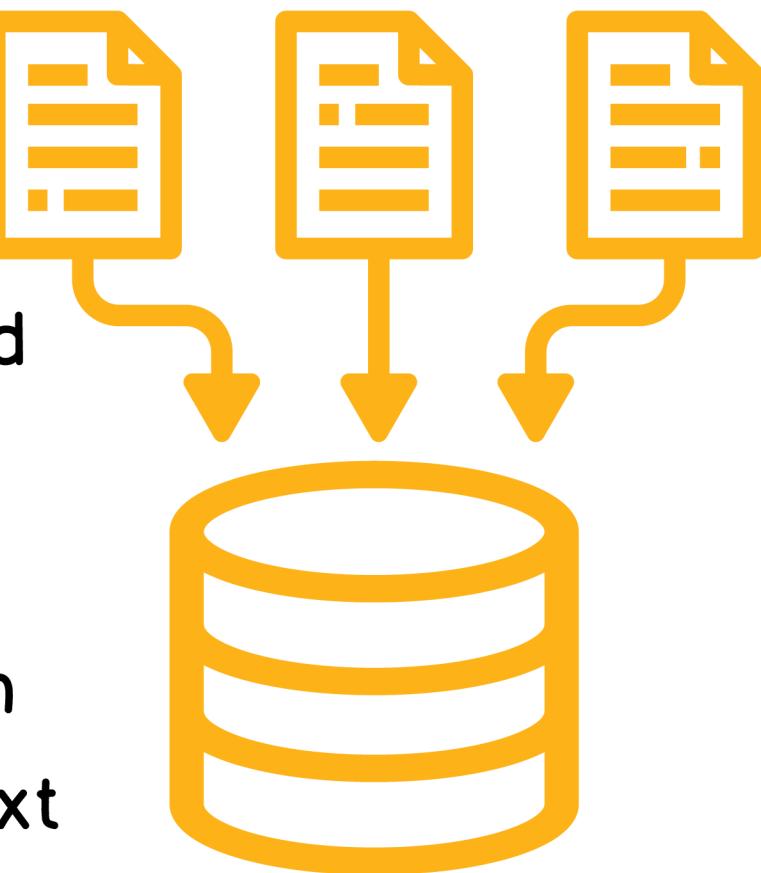
PROJECT OBJECTIVES

- **Sentiment Classification:** Classify Amazon Fine Food reviews as positive, neutral, or negative using NLP techniques.
- **Insight Extraction:** Analyze sentiment patterns to uncover customer preferences and trends.
- **Model Comparison:** Compare VADER (lexicon-based) and RoBERTa (transformer-based) for sentiment analysis accuracy and context.
- **Pattern Recognition:** Identify key trends to optimize product offerings and customer experience.
- **Recommendations:** Provide data-driven insights for improving marketing and customer engagement.

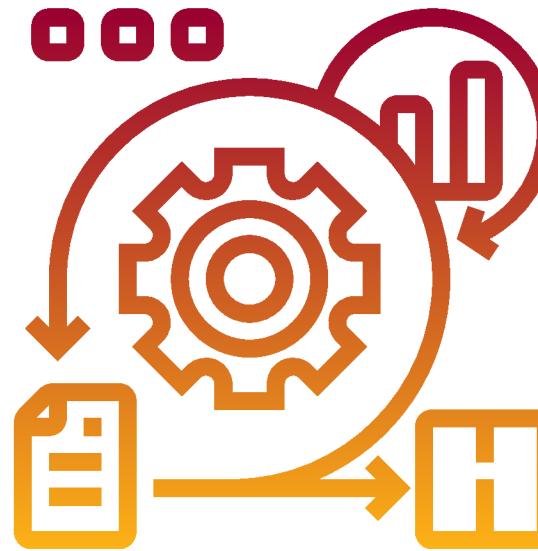


DATASET OVERVIEW

- **Dataset Overview:** Contains over 500,000 Amazon food product reviews, featuring unstructured text data alongside metadata such as product IDs, categories, and user identifiers.
- **Attributes:** Each review includes star ratings (1–5), helpfulness scores, and free-text reviews, offering a rich source for sentiment analysis, behavior modeling, and text classification.
- **Data Characteristics:** Represents a real-world, diverse dataset with varied writing styles, nuanced sentiment expressions, and different levels of text quality, suitable for training and evaluating machine learning models.



METHODOLOGY



1. **EDA:** Conducted review-level analysis using word clouds, n-grams, and sentiment distribution plots; identified sentiment trends across scores and time-based variations.
2. **Sentiment Models:**
 - a. **VADER:** Lexicon-based and fast; effective for general polarity but limited in contextual depth.
 - b. **RoBERTa:** Transformer model fine-tuned for multi-class sentiment; excels at capturing subtle semantic cues.
3. **Pipeline:** Designed a three-stage system: base VADER for speed, boosted VADER hybrid (XGBoost) for accuracy lift, and RoBERTa for deep contextual classification.
4. **Evaluation:** Employed stratified splitting and k-fold validation; tracked model accuracy, F1-score, and generated visual reports using Streamlit interface.

IMPLEMENTATION

I. Data Preparation-

Dataset Source: Amazon Fine Food Reviews – over 500,000 user-generated product reviews.

Objective: Prepare raw text data and review scores for sentiment classification.

Text: User-written review content.

Score: Numerical product rating (1 to 5 stars).

Transformed the Score into three sentiment classes:

- Negative (1–2 stars)
- Neutral (3 stars)
- Positive (4–5 stars)



	ID	ProductId	UserId	ProfileName	HelpfulnessNumerator	HelpfulnessDenominator	Score	Time	Summary	Text
0	1	B001E4KFG0	A3SGXH7AUHU8GW	delmartian	1		1	5	1303862400	Good Quality Dog Food I have bought several of the Vitality canned d...
1	2	B00813GRG4	A1D87F6ZCVE5NK	dll pa	0		0	1	1346976000	Not as Advertised Product arrived labeled as Jumbo Salted Peanut...
2	3	B000LQOCHO	ABXLMWJIXXAIN	Natalia Corres "Natalia Corres"	1		1	4	1219017600	"Delight" says it all This is a confection that has been around a fe...
3	4	B000UA0QIQ	A395BORC6FGVXV	Karl	3		3	2	1307923200	Cough Medicine If you are looking for the secret ingredient i...
4	5	B006K2ZZ7K	A1UQRSCLF8GW1T	Michael D. Bigham "M. Wassir"	0		0	5	1350777600	Great taffy Great taffy at a great price. There was a wid...

IMPLEMENTATION

2. Exploratory Data Analysis (EDA)-

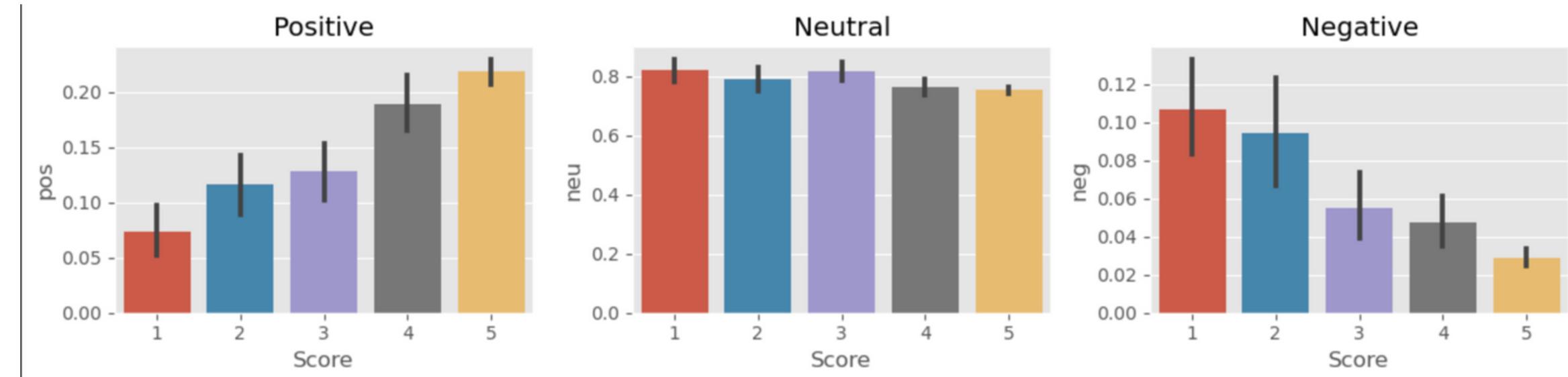
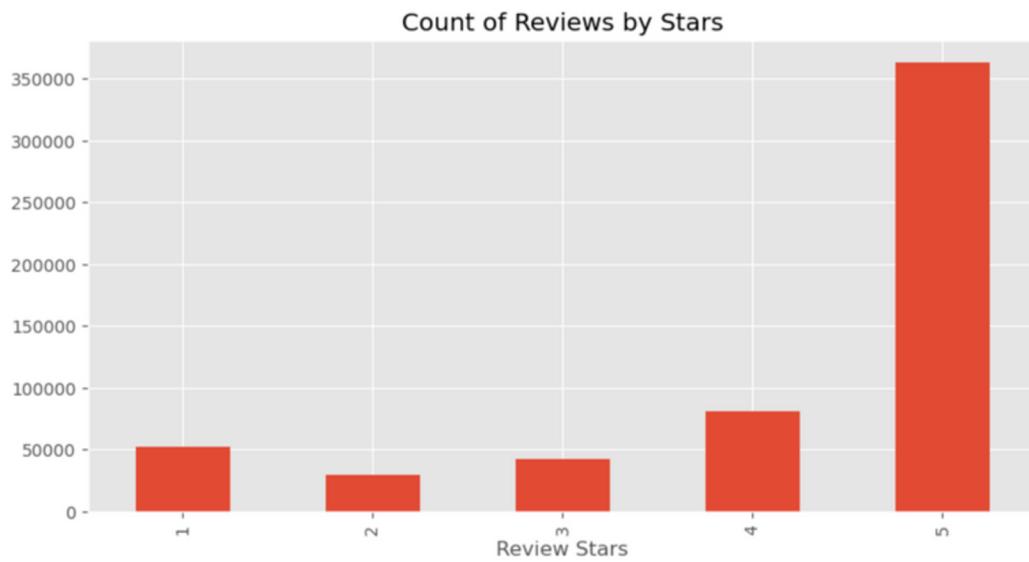
Goal: Understand the data structure, sentiment distribution, and common linguistic patterns.

Approach:

Analyzed class imbalances (e.g., more positive than negative/neutral reviews).

Extracted frequent keywords and phrases to understand vocabulary trends and user sentiment language.

Visualized distributions through bar plots, word clouds, and frequency charts.



```
sia.polarity_scores('I am so happy!')  
{'neg': 0.0, 'neu': 0.318, 'pos': 0.682, 'compound': 0.6468}  
  
sia.polarity_scores('This is the worst thing ever.')  
{'neg': 0.451, 'neu': 0.549, 'pos': 0.0, 'compound': -0.6249}
```

IMPLEMENTATION

3. Model Fine-Tuning-

a) **RoBERTa Fine-Tuning:** Fine-tuned transformer model on review texts.

Saved model, tokenizer, and validation accuracy plots.

b) **VADER Fine-Tuning:** Expanded lexicon with frequent words. Evaluated accuracy with custom lexicon.

c) **Boosted VADER Hybrid Training:** Enhanced lexicon with key phrases.

Trained XGBoost Classifier on VADER features. Saved XGBoost model and improved accuracy.

4. Streamlit Web App-

User Interaction: Input review for sentiment prediction.

Prediction Models: Normal VADER, Boosted VADER, RoBERTa.

Visualizations: Pie charts, Wordcloud, Accuracy of each model.



CHALLENGES FACED

- **Dataset:**
- Has more positive reviews than negative, Has noise, missing values, and redundant entries, some reviews are subjective and have sarcasm.
- **Roberta Model:**
- Computation problem as the dataset was very big to be fine tuned on.
- **Streamlit GUI:**
- Connecting the streamlit GUI to the fine tuned models of Roberta, Vader and boosted Vader

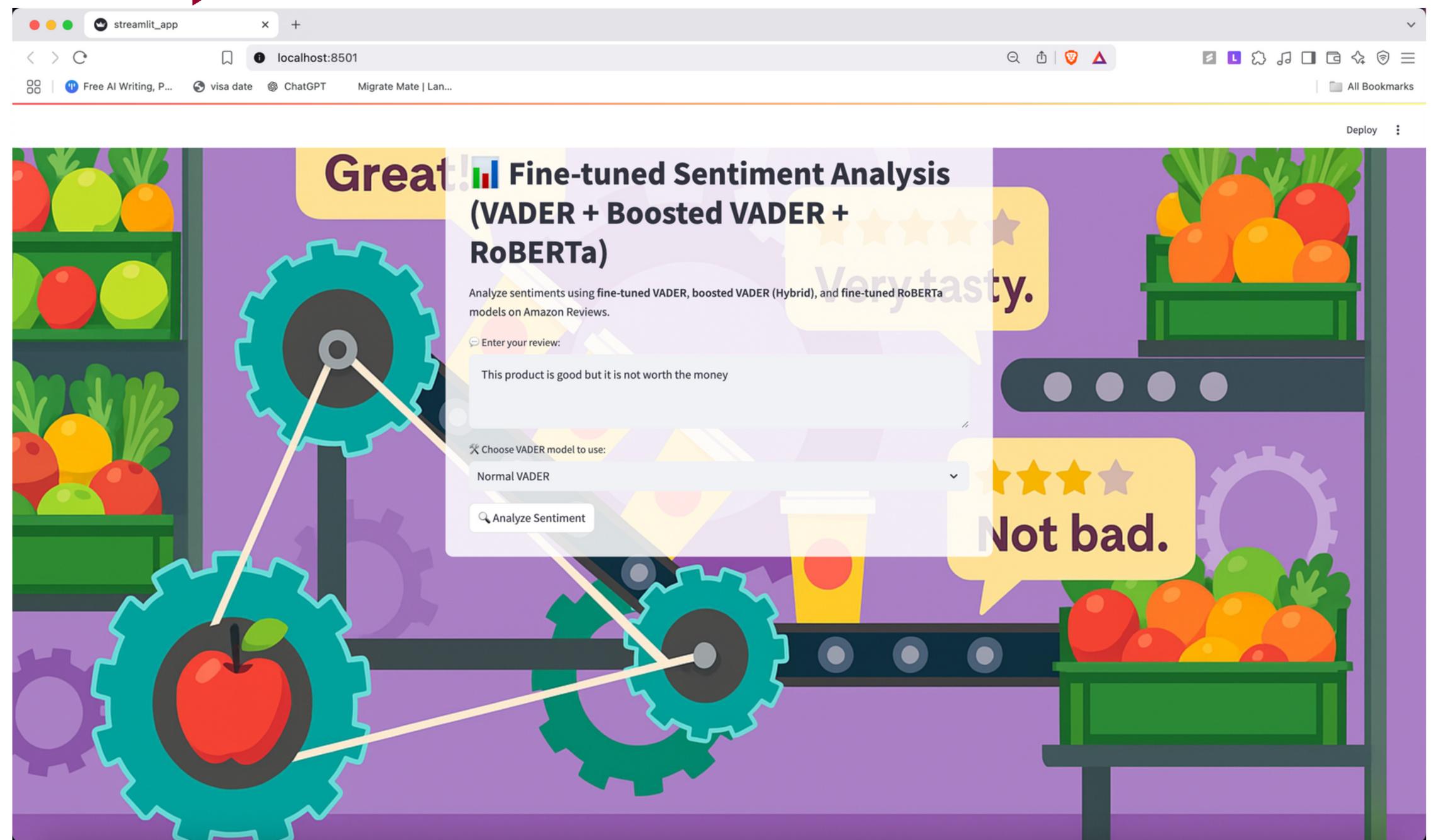


RESOLVING CHALLENGES

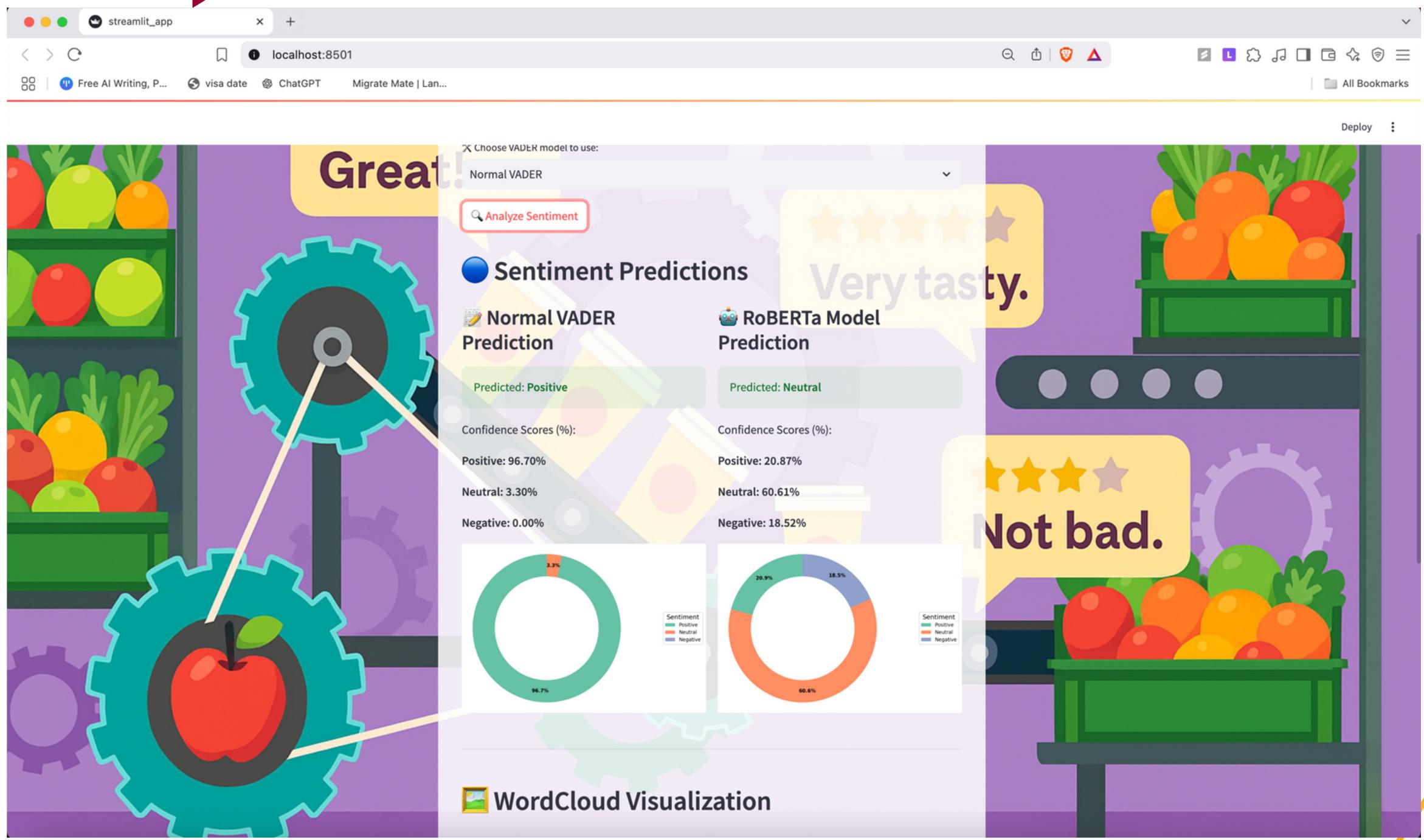
- **Dataset:**
- Cleaning of the Dataset to remove outliers and then did EDA on the dataset
- **Roberta Model:**
- As the dataset was very huge it took Roberta model 10+ hours to run 3 epochs for the dataset on a GPU.
- **Streamlit:**
- We created dump files for the models so they can be used in the applications without the need of training them again and again.



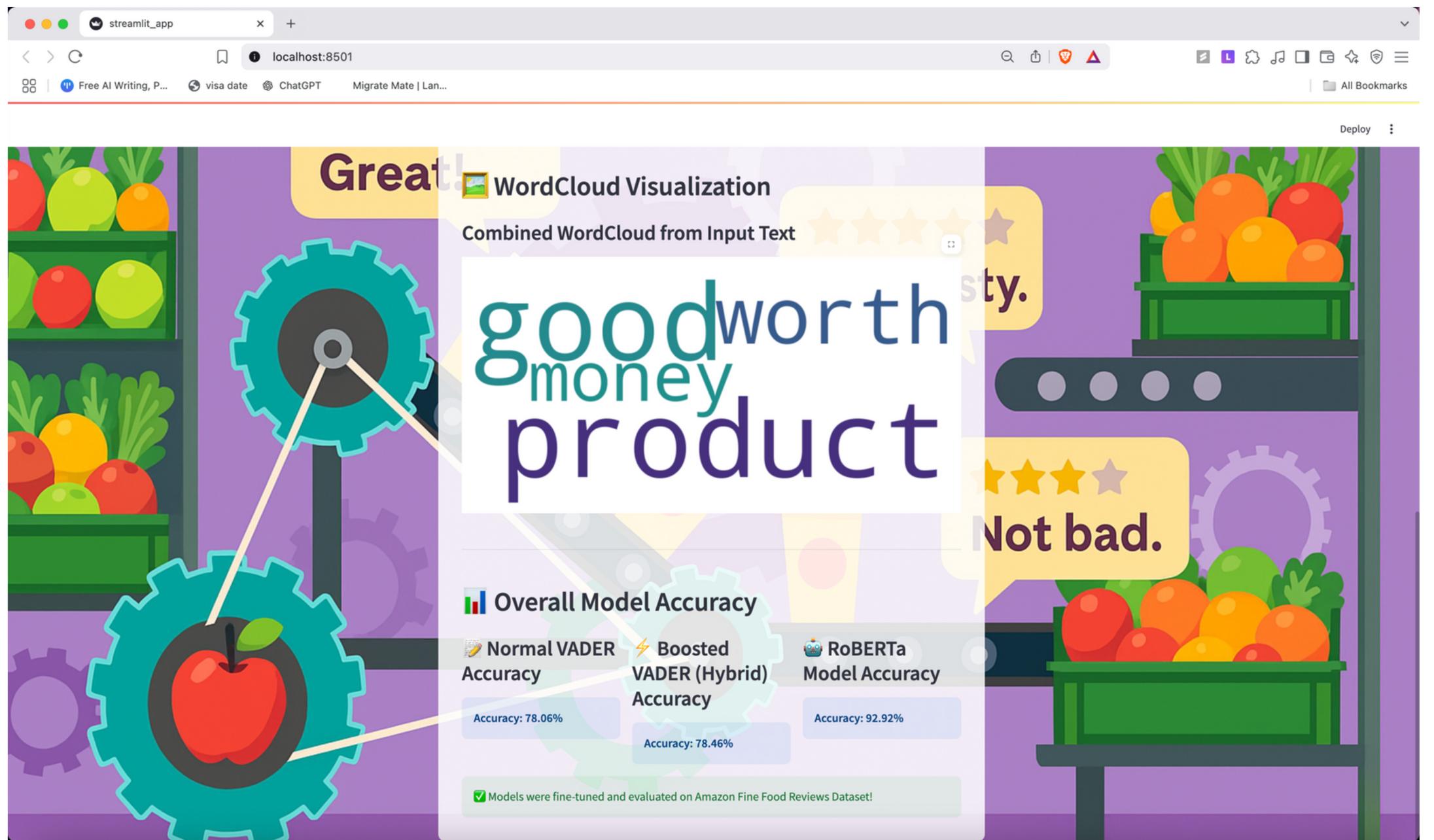
RESULTS



RESULTS



RESULTS



IMPACT

1. Industry Impact:

Empowers e-commerce platforms to understand customer sentiment at scale, refine product recommendations, improve user satisfaction, and make data-driven business decisions using real-time feedback analysis.

2. Societal Impact:

Enables consumers to make informed choices through transparent sentiment insights, while helping companies reduce unnecessary churn and focus on genuine customer needs fostering ethical AI-powered feedback systems.

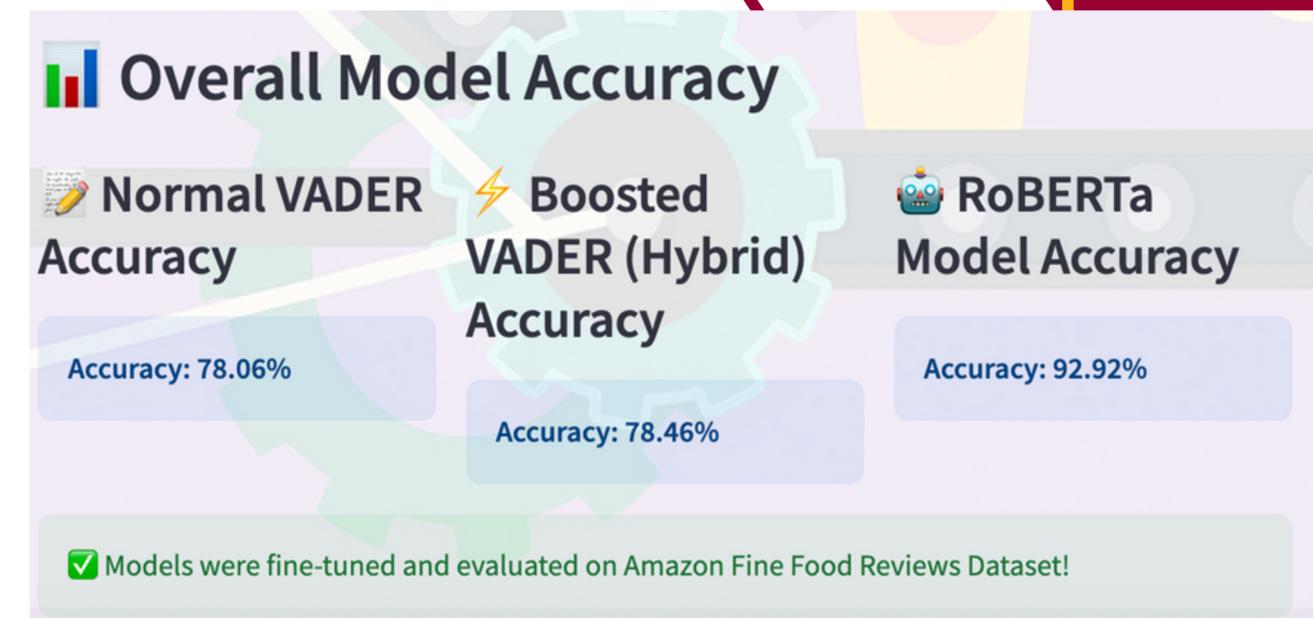
3. Research Impact:

Demonstrates the effectiveness of combining rule-based, hybrid, and transformer-based models (like RoBERTa) for sentiment analysis, contributing to advancements in explainable NLP, model interpretability, and scalable deployment via interactive Streamlit apps.



CONCLUSION

This project fuses advanced NLP models with efficient data pipelines to tackle a core challenge in e-commerce: extracting meaningful sentiment from customer reviews. By combining the speed of VADER (78.06%), the enhanced accuracy of a boosted hybrid model (78.46%), and the precision of fine-tuned RoBERTa (92.92%), it delivers real-time, actionable insights through a user-friendly Streamlit dashboard. The system supports business decisions that improve customer satisfaction, optimize product strategy, and reduce inefficiencies promoting both commercial and sustainable goals. Its scalable architecture also makes it applicable across diverse domains where unstructured text is abundant.



FSE 570

THANK YOU

TEAM PLUTO

Anand Khanna Sai Varun Alexis Myers Hetvi Shah