1. Project Overview

The goal of this project is to automate the processing of customer reviews by comparing traditional machine learning (ML) methods with a deep learning approach (transformers). Additionally, the project aims to use Generative AI to summarize reviews based on ratings and product categories and to create a visualization dashboard.

2. Project Scope

1. Sentiment Classification:  
    - Classify reviews as \*Positive, Neutral, or Negative\* based on text content.  
    - Compare traditional ML methods with a transformer-based model.  
     
   Generative AI (Bonus Task):  
    - Summarize reviews per rating (0-5 stars) and top product categories.  
     
   Dashboard & Visualization:  
    - Create a dynamic, interactive dashboard using Plotly, Tableau, or another visualization tool.

3. Data Collection

- Amazon customer reviews dataset

A graph of a box plot

Description automatically generated

Average Sentence Length: 30.40 words

Max Length: 1858.00 words

Min Length: 1.00 words  
  
Preprocessing Rules:

- Convert rating scores:  
 - 1, 2, 3 → Negative  
 - 4 → Neutral  
 - 5 → Positive

4. Methodology

1. Traditional NLP & ML Approach

|  |  |
| --- | --- |
| **Overall Model Accuracies** | |
|  |  |
| **Model** | **Accuracy** |
| **Naive Bayes** | 0.7001 |
| **Logistic Regression** | 0.7274 |
| **Support Vector Machine** | 0.7216 |
| **Random Forest** | 0.7109 |

Model Evaluation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Naive Bayes** |  |  |  |  |
| **Class** | **Precision** | **Recall** | **F1-Score** | **Support** |
| **0 (Negative)** | 0.8 | 0.01 | 0.02 | 462 |
| **1 (Neutral)** | 0.48 | 0.08 | 0.13 | 1709 |
| **2 (Positive)** | 0.71 | 0.99 | 0.83 | 4755 |
| **Accuracy** | - | - | **0.7** | 6926 |
| **Macro Avg** | 0.66 | 0.36 | 0.33 | 6926 |
| **Weighted Avg** | 0.66 | 0.7 | 0.6 | 6926 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Logistic Regression** | |  |  |  |
| **Class** | **Precision** | **Recall** | **F1-Score** | **Support** |
| **0 (Negative)** | 0.63 | 0.27 | 0.38 | 462 |
| **1 (Neutral)** | 0.53 | 0.24 | 0.33 | 1709 |
| **2 (Positive)** | 0.76 | 0.95 | 0.84 | 4755 |
| **Accuracy** | - | - | **0.73** | 6926 |
| **Macro Avg** | 0.64 | 0.49 | 0.52 | 6926 |
| **Weighted Avg** | 0.69 | 0.73 | 0.68 | 6926 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Support Vector Machine** | |  |  |  |
| **Class** | **Precision** | **Recall** | **F1-Score** | **Support** |
| **0 (Negative)** | 0.62 | 0.26 | 0.37 | 462 |
| **1 (Neutral)** | 0.55 | 0.15 | 0.23 | 1709 |
| **2 (Positive)** | 0.74 | 0.97 | 0.84 | 4755 |
| **Accuracy** | - | - | **0.72** | 6926 |
| **Macro Avg** | 0.64 | 0.46 | 0.48 | 6926 |
| **Weighted Avg** | 0.68 | 0.72 | 0.66 | 6926 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Random Forest** |  |  |  |  |
| **Class** | **Precision** | **Recall** | **F1-Score** | **Support** |
| **0 (Negative)** | 0.69 | 0.1 | 0.17 | 462 |
| **1 (Neutral)** | 0.56 | 0.12 | 0.2 | 1709 |
| **2 (Positive)** | 0.72 | 0.98 | 0.83 | 4755 |
| **Accuracy** | - | - | **0.71** | 6926 |
| **Macro Avg** | 0.65 | 0.4 | 0.4 | 6926 |
| **Weighted Avg** | 0.68 | 0.71 | 0.63 | 6926 |

A graph of negative and negative matrix

Description automatically generated

**Classification Report for Logistic Regression (SMOTE)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Class** | **Precision** | **Recall** | **F1-Score** | **Support** |
| **0 (Negative)** | 0.3 | 0.65 | 0.41 | 462 |
| **1 (Neutral)** | 0.36 | 0.43 | 0.39 | 1709 |
| **2 (Positive)** | 0.82 | 0.68 | 0.74 | 4755 |
| **Accuracy** | - | - | **0.61** | 6926 |
| **Macro Avg** | 0.5 | 0.58 | 0.52 | 6926 |
| **Weighted Avg** | 0.67 | 0.61 | 0.63 | 6926 |

1. Transformer-Based Approach (Hugging Face)

'bert-base-multilingual-uncased-sentiment'

https://huggingface.co/nlptown/bert-base-multilingual-uncased-sentiment

*This is a bert-base-multilingual-uncased model finetuned for sentiment analysis on product reviews in six languages: English, Dutch, German, French, Spanish, and Italian. It predicts the sentiment of the review as a number of stars (between 1 and 5).*

*This model is intended for direct use as a sentiment analysis model for product reviews in any of the six languages above or for further finetuning on related sentiment analysis tasks.*

\*\*Bonus: Fine-Tuning\*\*  
- Train the model on the dataset to adapt it for sentiment classification.  
- Optimize batch size, learning rate, and training epochs.

\*\*Step 3: Model Evaluation\*\*  
- Compare pre-trained vs. fine-tuned model performance.  
- Metrics: Accuracy, Precision, Recall, F1-score.  
- Confusion Matrix: For deeper insights.

5. Bonus: Generative AI Summarization

- Summarization Objective: Generate review summaries:  
 - Grouped by review score (0-5 stars).  
 - Grouped by product categories.  
- Approach: Use GPT-based models (e.g., T5, BART) for summarization.

6. Dashboard & Visualization

- \*\*Tool:\*\* Use Plotly, Tableau, or other visualization tools.  
- \*\*Features:\*\*  
 - Sentiment distribution over time.  
 - Breakdown of reviews per rating category.  
 - Summarized insights for product categories.

7. Deliverables

1. \*\*PDF Report:\*\* Documenting the methodology, results, and analysis.  
2. \*\*Reproducible Code:\*\* Jupyter notebooks or Python scripts.  
3. \*\*PowerPoint Presentation:\*\* Summarizing the findings.  
4. \*\*Bonus:\*\* Deploy a web app for real-time querying.