Intel Core i7-12700K Processor Sentiment Analysis

Analyzing Customer Sentiments from Online Reviews

Presented by Team Techashu

Mentor Name: Ms. Sarika Madavi
Email: Sarika.wctm@gmail.com

Problem Statement

- **Problem:** With the increasing number of online reviews for Intel products, it's challenging to manually analyze customer sentiments and extract meaningful insights.
- **Impact:** Understanding customer sentiments is crucial for Intel to improve product quality, customer satisfaction, and market strategies.
- **Goal:** Develop an automated system to analyze sentiments from online reviews, providing actionable insights for Intel.

Manual Sentiment Analysis System

- **Innovation:** Utilizes machine learning and natural language processing to analyze and classify sentiments from online reviews.
- **Efficiency:** Reduces manual effort and time required to understand customer feedback.
- **Insights:** Provides comprehensive insights into customer opinions, highlighting strengths and areas for improvement.

Key Features of the Sentiment Analysis System

- **Manual Sentiment Analysis:** Downloading HTML pages of reviews and extracting data from them.
- **Sentiment Classification:** Categorizes reviews into positive, negative, and neutral sentiments.
- **Keyword Extraction:** Identifies common keywords and phrases in reviews.
- **Insights Generation:** Provides valuable insights into customer opinions.

Process flow

01

Data Collection

- **Task:** Download HTML pages of Amazon and Best Buy reviews for the Intel Core i7-12700K processor.
- **Method:** Use web scraping tools to collect reviews from multiple country-specific Amazon sites and Best Buy.

02

Data Extraction

- **Task:** Extract relevant information (Content, Rating, Name, Date) from the downloaded HTML pages.
- **Method:** Utilize BeautifulSoup to parse the HTML and extract review data, saving it to a CSV file.

03

Data Translation

- **Task:** Translate non-English reviews to English.
- **Method:** Use the ‘langdetect’ ibrary to detect the language and Google Translate API for translation, saving results to ‘translated.csv’.

04

Data Cleaning

- Task: Clean the translated reviews data.
- **Method**: Remove duplicates, extract numeric ratings, filter unreadable text, and keep only relevant columns, saving cleaned data to ‘cleaned_reviews.csv’.

05

Sentiment Analysis, Evaluation and Visualization

- **Task:** Perform sentiment analysis on the cleaned reviews.
- **Method:** Train and apply machine learning models to classify sentiments, generate visualizations, and evaluate model performance.

Methodology

❖ **Data Collection**
- Sources: Collected user reviews from various platforms such as e-commerce websites (e.g., Amazon, Bestbuy)

❖ **Sample Size**
- Gathered a total of 737 reviews to ensure a comprehensive analysis.

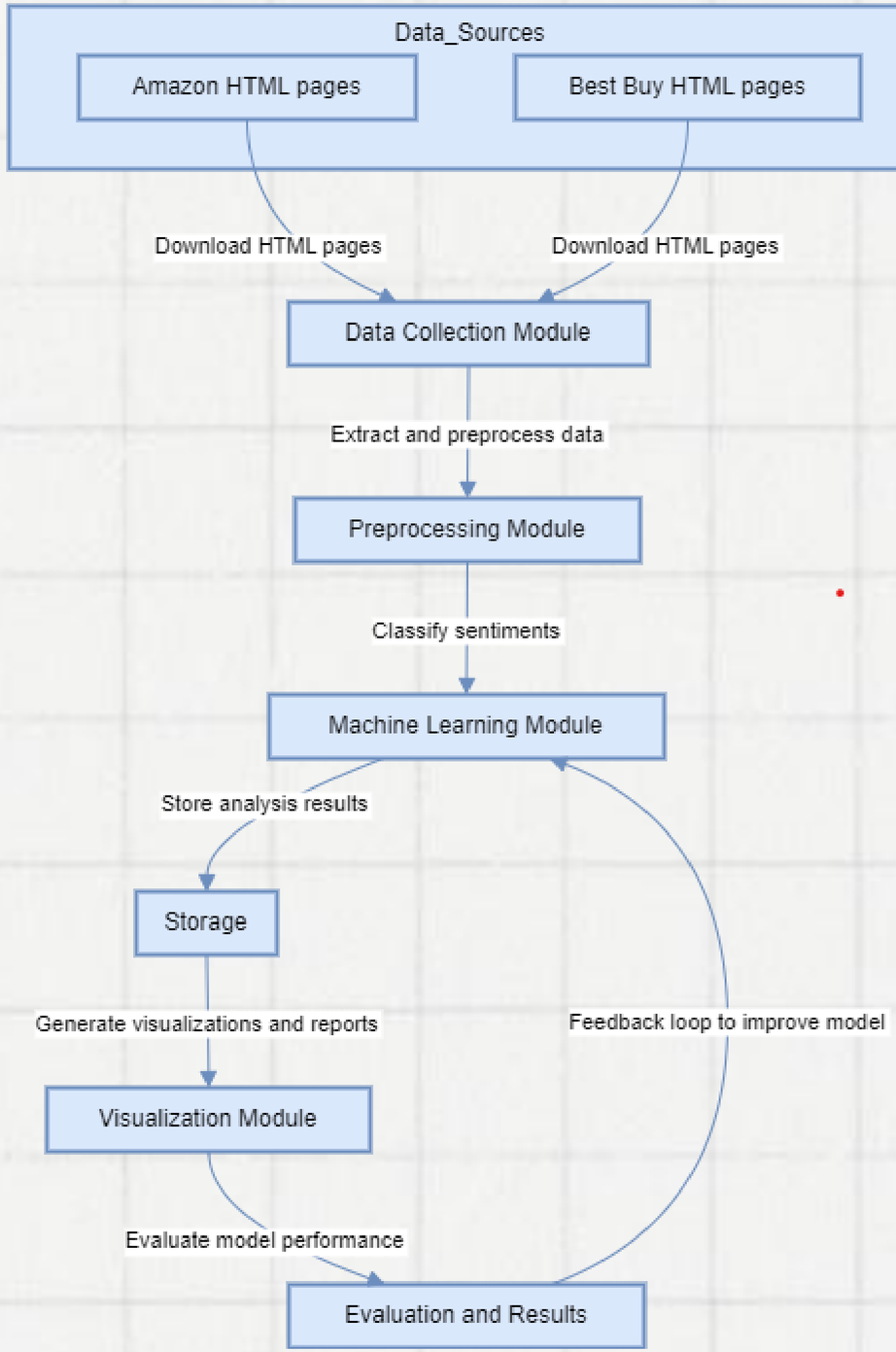
❖ **Data Preprocessing**
- Cleaning: Removed any duplicate reviews, non-English text, and irrelevant information such as advertisements or spam.
- Tokenization: Broke down the text into individual words or tokens to facilitate analysis.

❖ **Sentiment Analysis Tools**
- NLP Libraries: Used Natural Language Processing (NLP) libraries such as NLTK to perform sentiment analysis.
- Sentiment Scoring: Each review was assigned a sentiment score (e.g., positive, negative, neutral) based on the analysis.

❖ **Sentiment Classification**
- Threshold Setting: Defined thresholds for categorizing reviews into positive, negative, and neutral. For instance, scores above 0.5 were classified as positive, below -0.5 as negative, and in between as neutral.
- Manual Validation: Conducted a manual review of a subset of reviews to validate the accuracy of the sentiment analysis.

❖ **Data Visualization**
- Charts and Graphs: Created visual representations such as pie charts, bar graphs, and word clouds to illustrate the sentiment distribution and key themes.
- Highlighting Key Themes: Identified and highlighted the most common positive and negative aspects mentioned in the reviews

Architecture Diagram



Technologies and Tools

- **Programming Languages:** Python
- **Libraries:** pandas, numpy, BeautifulSoup, scikit-learn, imbalanced-learn, googletrans, langdetect, nltk, wordcloud
- **Machine Learning Models:** Logistic Regression
- **Visualization:** Matplotlib, Seaborn

Sentimental Analysis Breakdown

Positive Feedback

❖ Performance

- High praise for speed and multi-core capabilities.
- Excellent gaming and professional application performance.
- The multi-core performance of this processor make this outstanding.

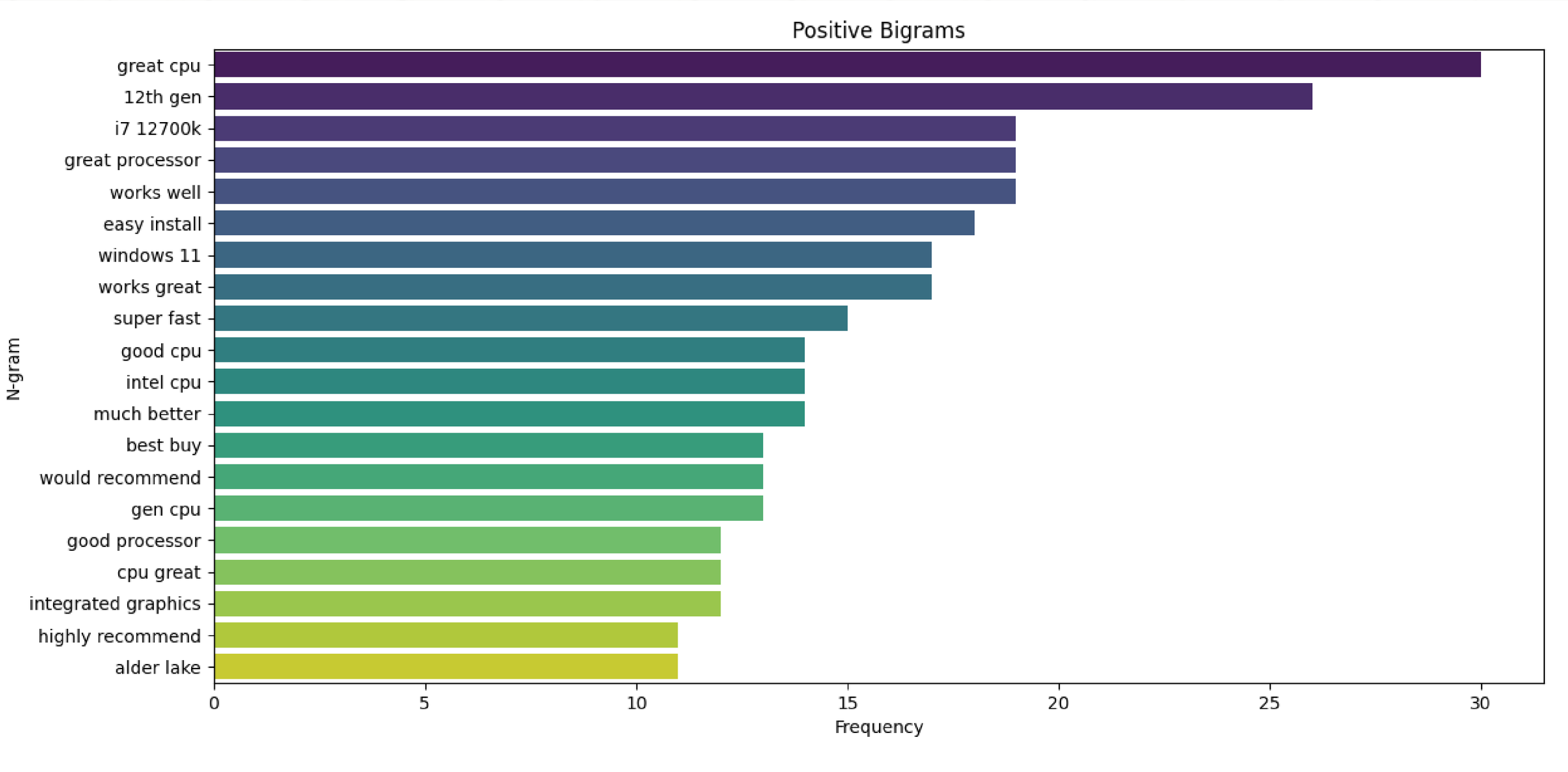
❖ Value for Money

- Perceived as offering great value compared to other high-end processors.
- Many users felt it offered good value relative to its capabilities

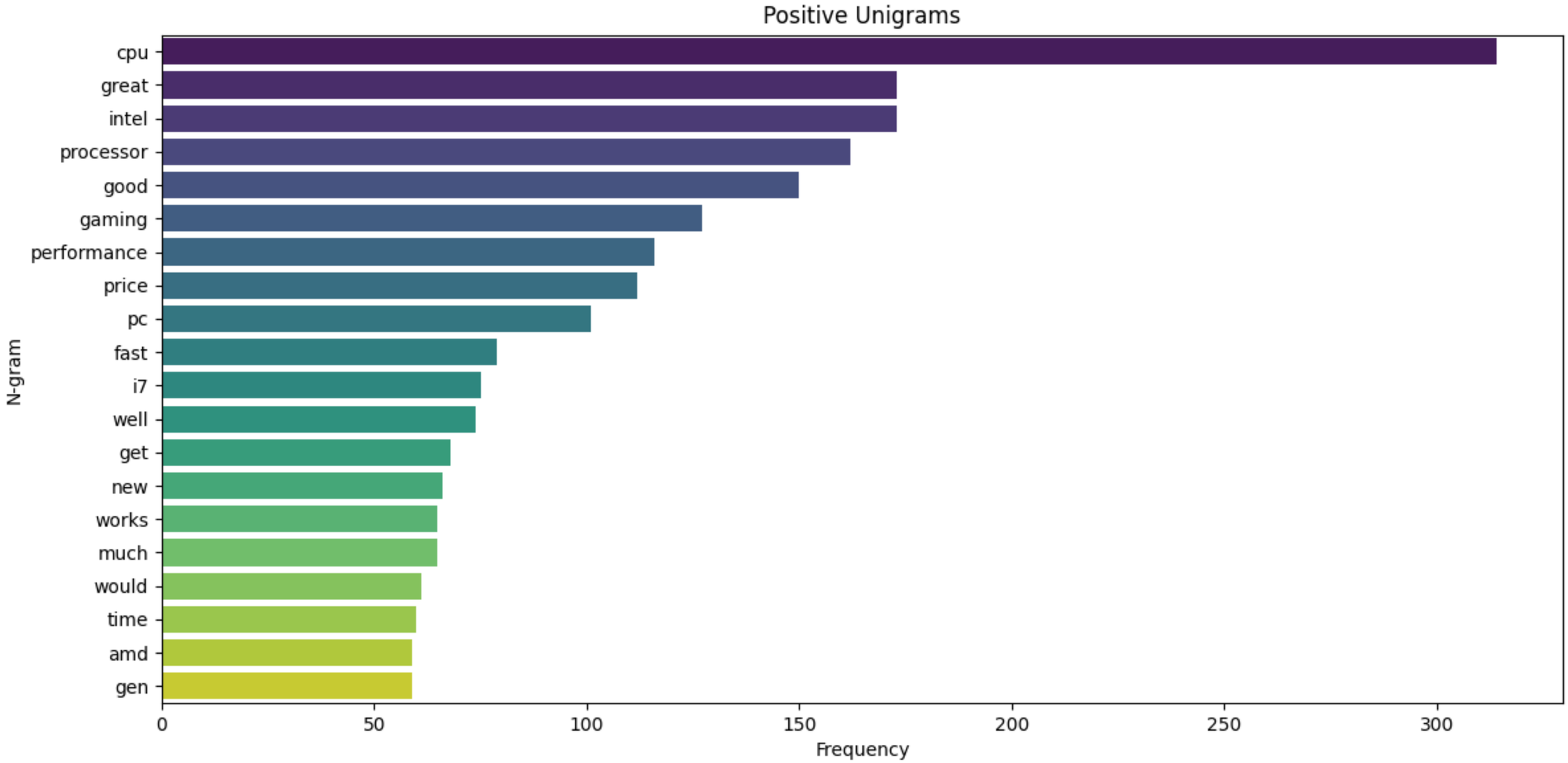
❖ User Experience

- Easy installation and compatibility with various systems.
- High overall satisfaction.

- **Visual Representation Of Most Frequent Words Used In The Reviews Of The User In The Form Of Bigrams**



- **Visual Representation Of Most Frequent Words Used In The Reviews Of The User In The Form Of Unigrams**



Negative Feedback

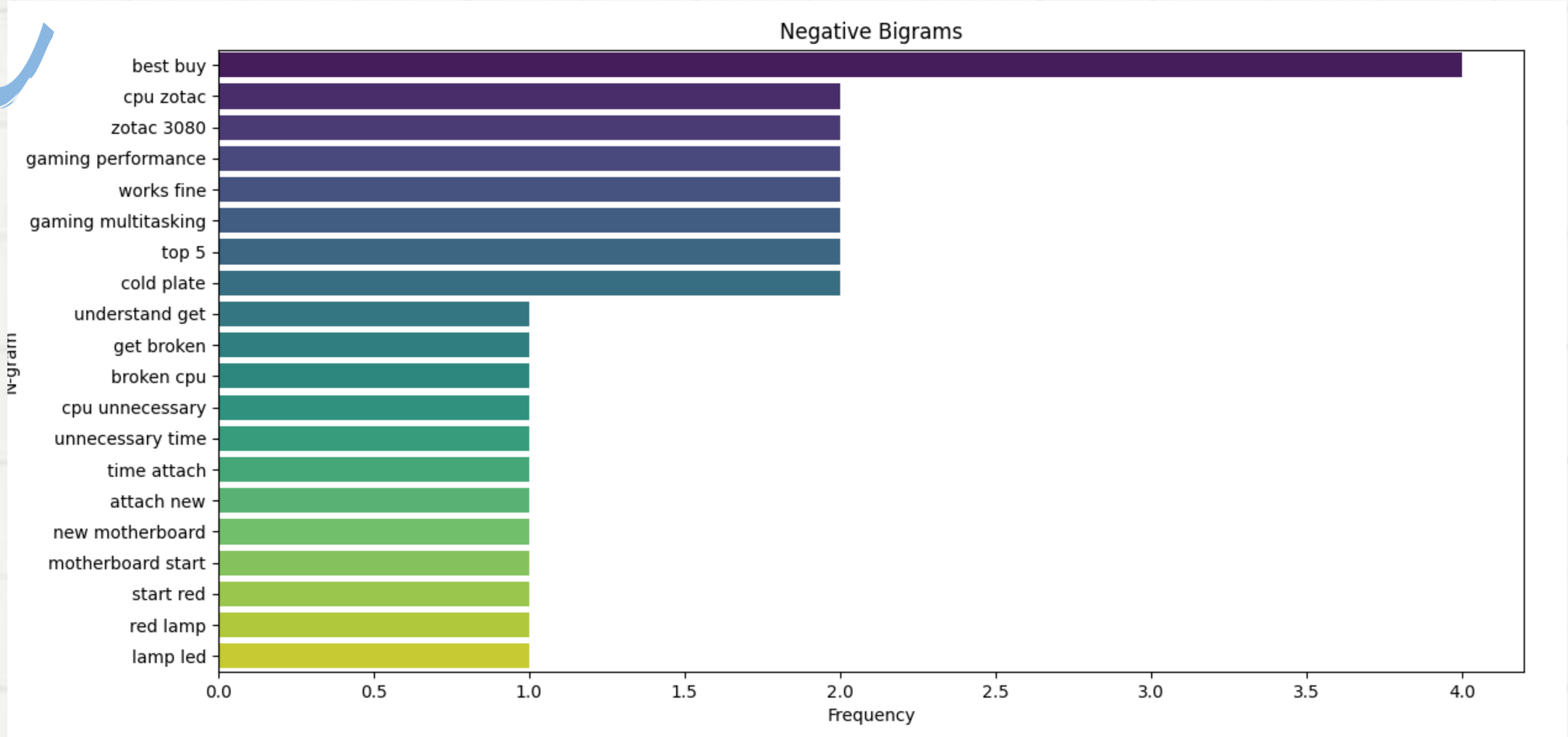
❖Heat Generation

- Reports of running hot under load.
- Necessity for additional cooling solutions.

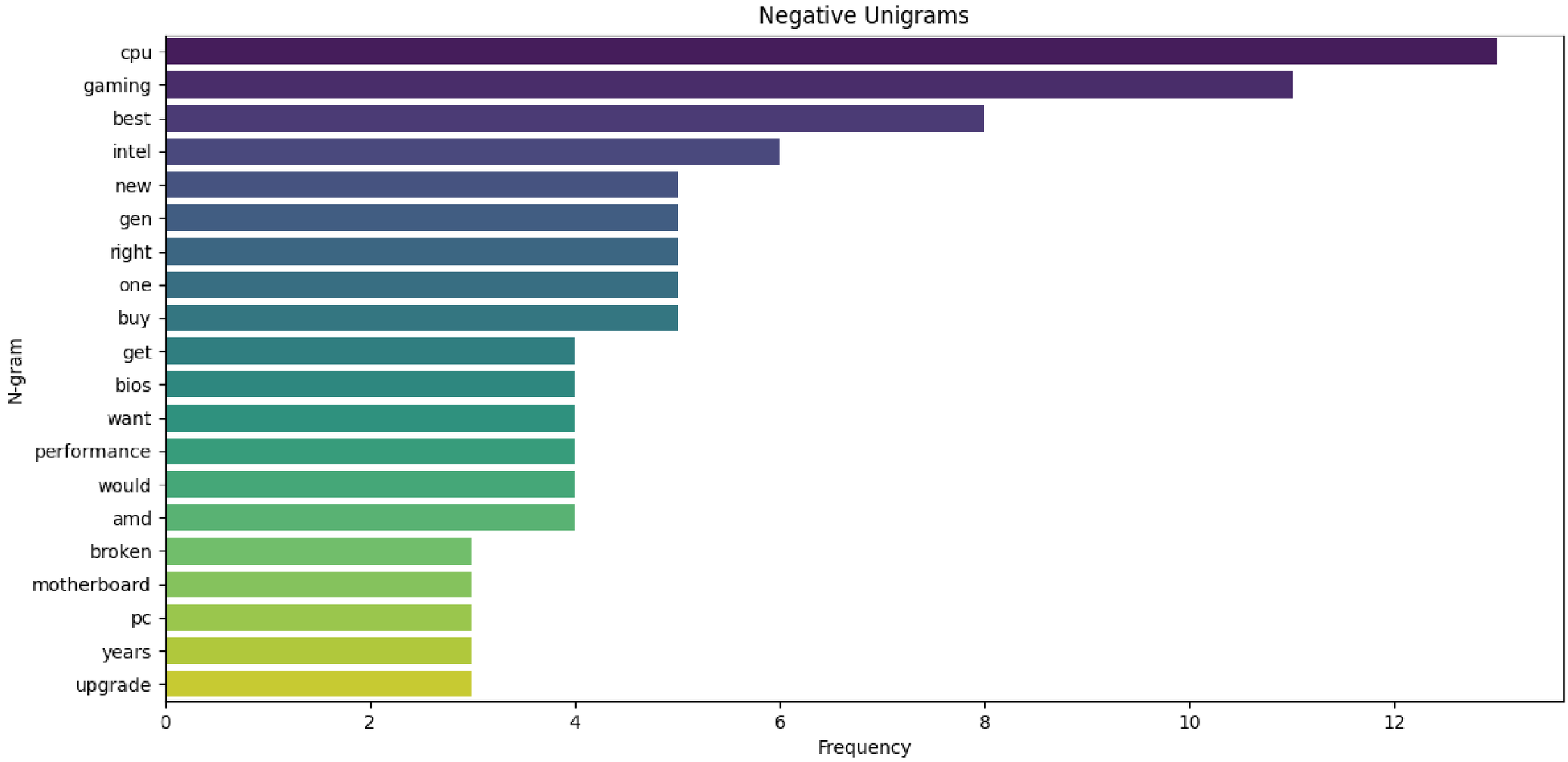
❖Other Issues

- Occasional faulty units and customer service complaints.

- **Visual Representation Of Most Frequent Words Used In The Reviews Of The User In The Form Of Bigrams**



- **Visual Representation Of Most Frequent Words Used In The Reviews Of The User In The Form Of Unigrams**



Neutral Feedback

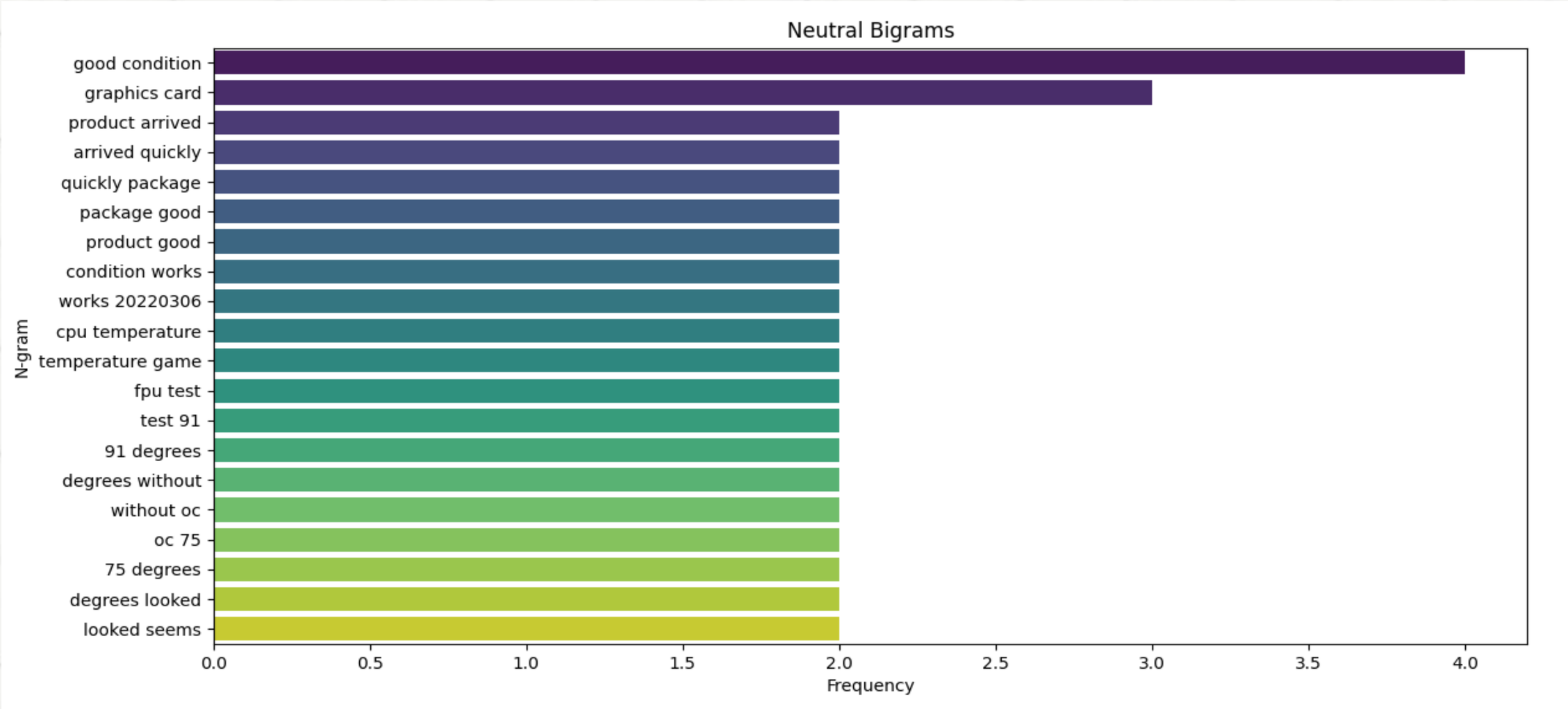
❖ **Balanced Views**

- Mixed feedback, noting both strengths and weaknesses.

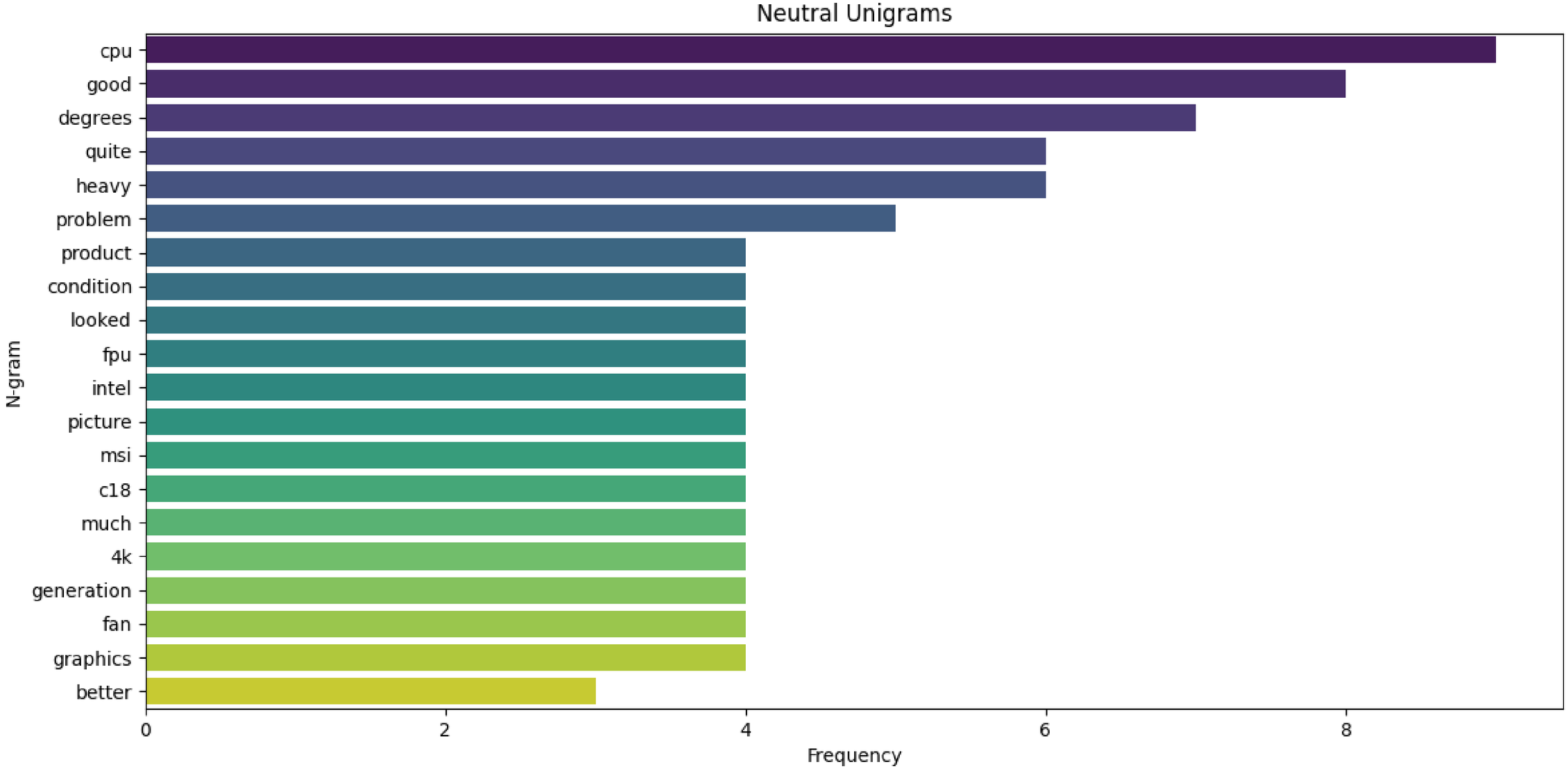
❖ **Common Themes**

- Performance praised but heat management and cost

- **Visual Representation Of Most Frequent Words Used In The Reviews Of The User In The Form Of Bigrams**

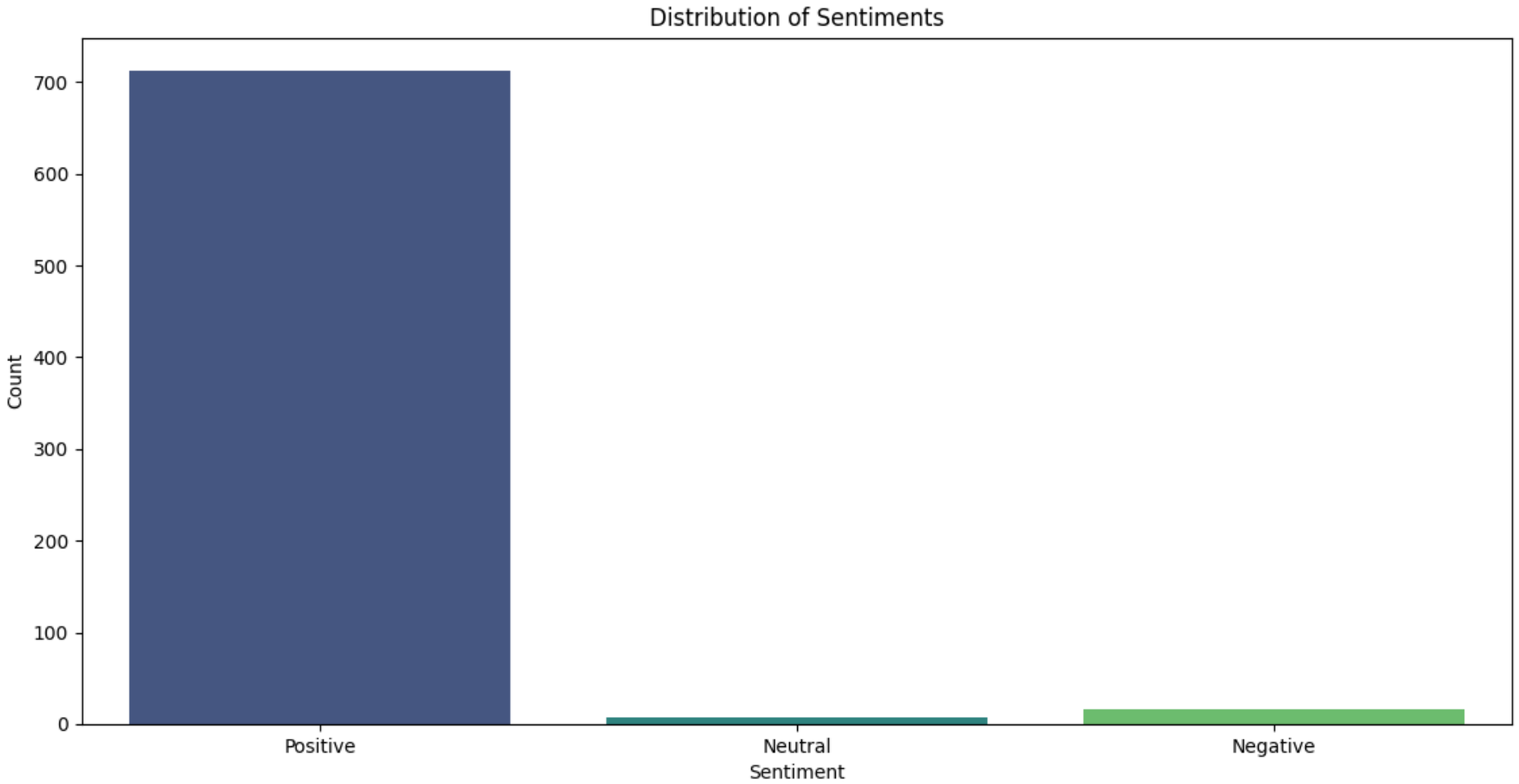


- **Visual Representation Of Most Frequent Words Used In The Reviews Of The User In The Form Of Unigrams**



Visual Representations

- **Distributions Of Positive, Negative And Neutral Reviews**

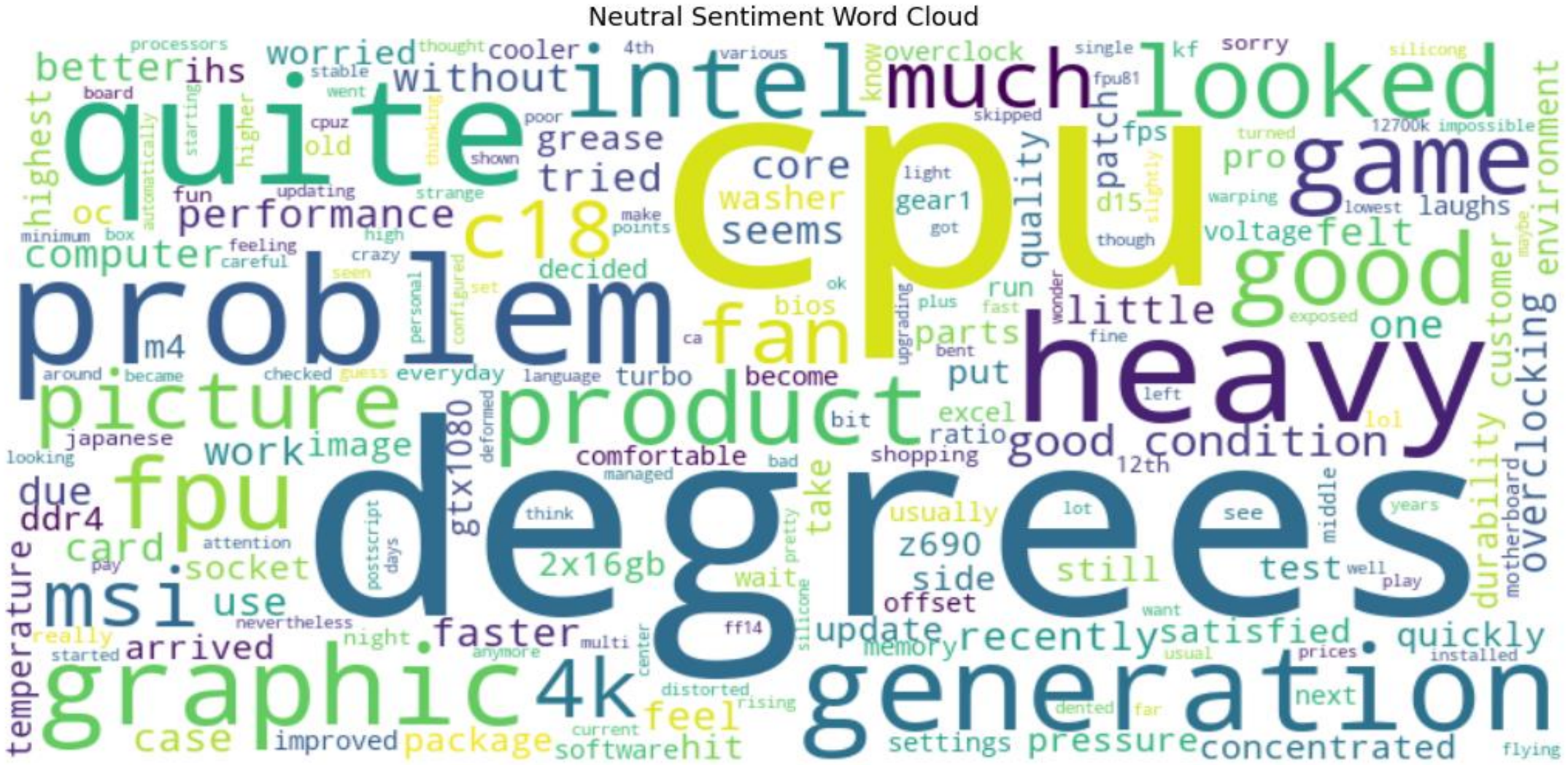


Visualising Using Word Cloud

Positive Review Word Cloud

Negative Review Word Cloud

Neutral Review Word Cloud



Team members and their contribution:

1. **Priyanshu Kumar (Team Lead):** Sentiment analysis, model training, visualization, and reporting
2. **Mayank Vaibhav:** Data Scraping, Data Collection
3. **Harshit Raj:** Data Translation and Cleaning
4. **Amit Kumar:** Resources Collection

Conclusion and Future Work

The sentiment analysis project for the Intel Core i7-12700K processor provides valuable insights into customer opinions and satisfaction levels across different regions. Through a comprehensive workflow involving data collection, translation, cleaning, and analysis, we were able to accurately categorize customer reviews into positive, neutral, and negative sentiments.

- **Positive Feedback:** A significant portion of reviews were positive, indicating high customer satisfaction with the Intel Core i7-12700K processor’s performance and features.
- **Common Themes:** The most frequent positive comments highlighted the processor’s speed, efficiency, and value for money. Negative comments were often related to pricing and occasional technical issues.
- **Model Performance:** The Logistic Regression model, optimized using GridSearchCV, demonstrated satisfactory accuracy in predicting review sentiments. The use of SMOTE effectively addressed class imbalance, enhancing the model's performance.
- **Insights for Improvement:** The analysis provides actionable insights for Intel to address common customer concerns and improve future product iterations.

Thank you very much!