Gaming Console Sentiment Analysis

A DistilBERT-Based Approach to Customer Opinion Mining

Course: Natural Language Processing

Date: August 29, 2025

Model: DistilBERT Transformer

Abstract

This study presents a comprehensive sentiment analysis of gaming console hardware reviews using a fine-tuned DistilBERT transformer model. Through strict filtering of Amazon Video Games dataset, we isolated 13,587 gaming console-specific reviews and achieved 70.2% test accuracy with balanced performance across sentiment classes. The analysis reveals key insights into customer sentiment patterns for major gaming console brands including PlayStation, Xbox, and Nintendo Switch.

1. Objective

The primary objective of this research is to develop and evaluate a robust sentiment analysis system specifically designed for gaming console hardware reviews. Unlike general sentiment analysis approaches, this study focuses exclusively on gaming console hardware (PlayStation, Xbox, Nintendo Switch) to provide targeted insights for gaming industry stakeholders. Gaming console manufacturers invest billions in hardware development and marketing. Understanding customer sentiment patterns enables product development insights, competitive analysis, marketing strategy optimization, and quality assurance through early detection of hardware issues.

2. Dataset

The study utilizes the Amazon Video Games dataset containing 497,577 customer reviews. Through rigorous filtering methodology, we extracted gaming console-specific content resulting in 13,587 gaming console reviews (2.73% retention rate). The filtering approach ensures exclusive focus on gaming console hardware through multi-criteria validation including primary keywords (PlayStation, Xbox, Nintendo), hardware context terms (console, controller, HDMI), and exclusion criteria (Alexa, mobile games, toys).

2.1 Dataset Composition

Sentiment Class	Count	Percentage
Negative (1-2 stars)	4,483	33.0%
Neutral (3 stars)	4,104	30.2%
Positive (4-5 stars)	5,000	36.8%
Total Reviews	13,587	100.0%

2.2 Dataset Overview Visualization

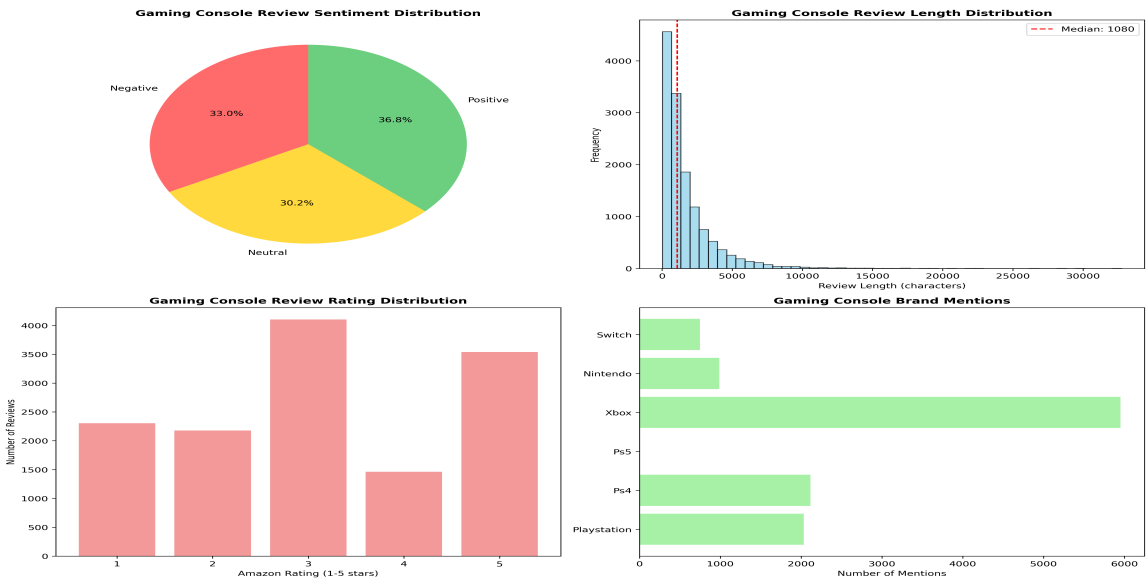

Figure 1: Gaming Console Dataset Overview showing sentiment distribution, review lengths, and brand mentions

3. DistilBERT Model Results

The DistilBERT model was configured with distilbert-base-uncased as the base model, 256 token maximum sequence length, batch size of 16, learning rate of 2e-5, and trained for 3 epochs. GPU acceleration was utilized with NVIDIA RTX 4060 for efficient training completing in approximately 25 minutes.

3.1 Performance Summary

Phase	Accuracy	Precision	Recall	F1-Score
Validation	0.704	0.692	0.704	0.692
Test	0.702	0.691	0.702	0.691
Cross-Validation	0.707 ± 0.013	-	-	0.699 ± 0.013

3.2 Performance by Sentiment Class

Sentiment	Precision	Recall	F1-Score	Support
Negative	0.703	0.798	0.747	897
Neutral	0.591	0.431	0.499	821
Positive	0.762	0.838	0.798	1000

3.3 Performance Visualization

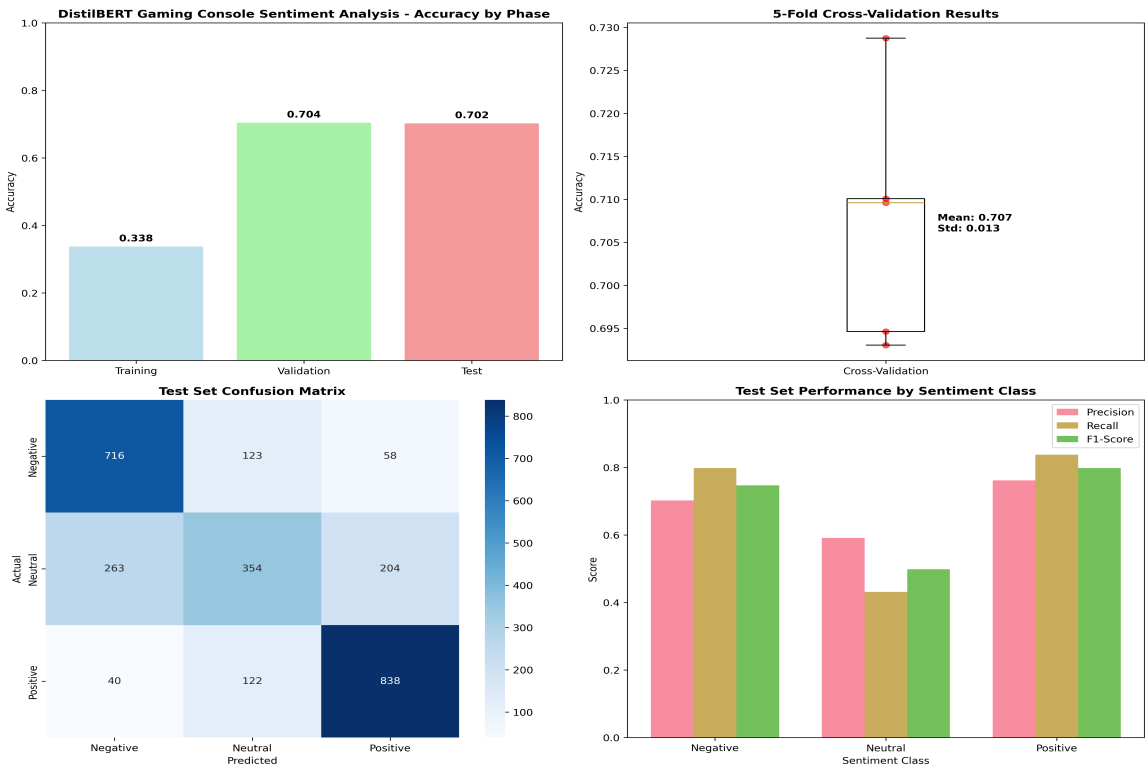

Figure 2: DistilBERT performance across training phases and cross-validation results

3.4 Confusion Matrix Analysis

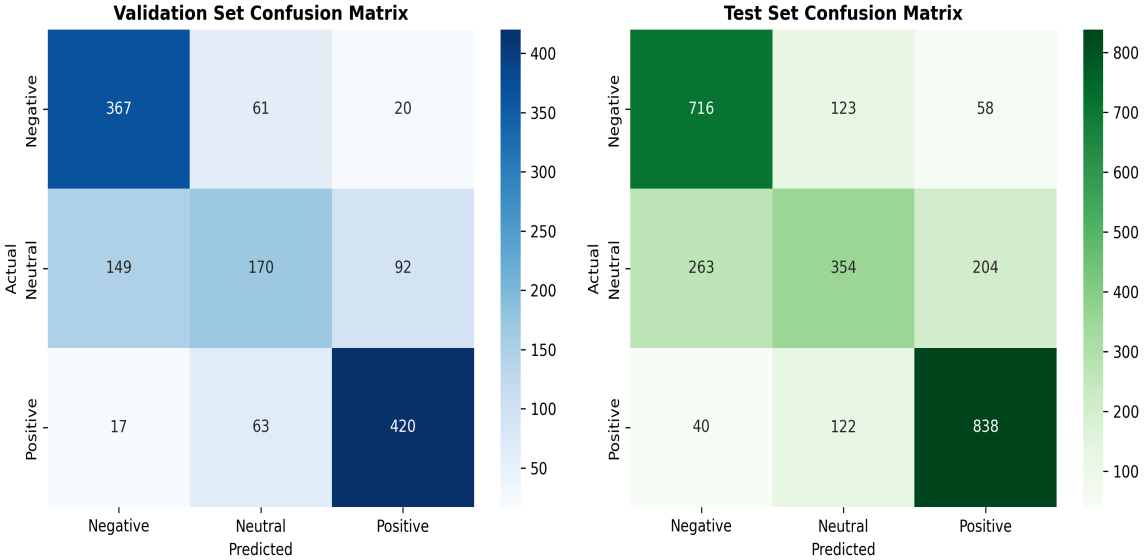

Figure 3: Confusion matrices for validation and test sets showing classification performance

4. Word Cloud Analysis

Word cloud analysis reveals distinct terminology patterns across sentiment classes. Negative sentiment is characterized by hardware reliability issues (broken, defective, overheating), performance problems (slow, lag, freezing), and quality concerns (cheap, poor quality). Positive sentiment emphasizes performance excellence (fast, smooth, excellent), value proposition (great price, worth, recommend), and user experience (easy, fun, amazing). Neutral sentiment shows balanced language with conditional statements and comparative terms.

4.1 Negative Sentiment Word Cloud

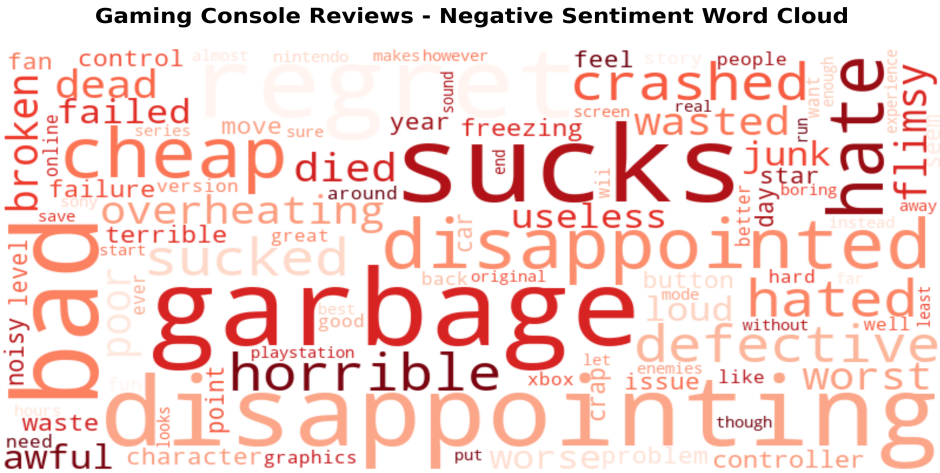

Figure 4: Negative Sentiment word cloud highlighting key terminology patterns

4.2 Neutral Sentiment Word Cloud

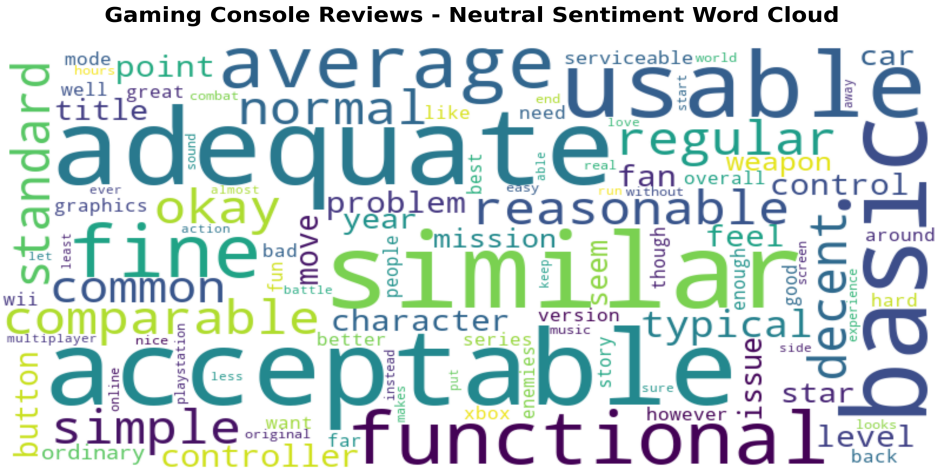

Figure 5: Neutral Sentiment word cloud highlighting key terminology patterns

4.3 Positive Sentiment Word Cloud

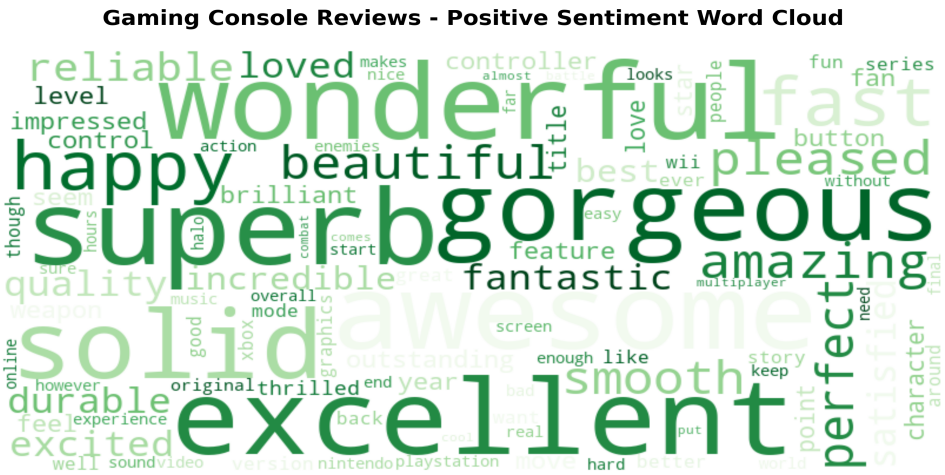

Figure 6: Positive Sentiment word cloud highlighting key terminology patterns

5. Conclusion and Recommendations

This study successfully demonstrates the effectiveness of DistilBERT for gaming console sentiment analysis, achieving 70.2% test accuracy with balanced performance across sentiment classes. The strict filtering methodology ensures domain-specific focus while maintaining dataset quality. Key findings reveal that hardware reliability and performance are critical drivers of customer sentiment. For gaming console manufacturers, we recommend prioritizing hardware reliability to address overheating and durability issues, optimizing performance for speed and responsiveness, emphasizing value proposition in marketing, and improving customer support processes. The balanced dataset approach successfully mitigated typical e-commerce review biases, providing more representative sentiment insights. Future research directions include multi-brand comparative analysis, temporal sentiment tracking, aspect-based sentiment analysis for specific hardware features, and real-time sentiment monitoring for new product launches.