Final Report:

Customer Sentiment & Opinion Mining on Women’s Clothing Reviews using NLP

---

1. Executive Summary

Customer reviews contain valuable insights about product quality, customer satisfaction, and service issues. However, due to their unstructured nature and large volume, manual analysis is inefficient. This project applies Natural Language Processing (NLP) and machine learning techniques to analyze customer reviews from a women’s clothing e-commerce dataset in order to understand sentiment patterns, identify recurring topics, and generate actionable business insights.

The dataset was preprocessed using standard NLP techniques such as text cleaning, tokenization, stopword removal, and lemmatization. Exploratory text analysis was performed to examine frequently used words, phrases, and sentiment distribution. Multiple machine learning models were evaluated for sentiment classification, including Logistic Regression, Naive Bayes, and Linear Support Vector Machine (SVM).

To address class imbalance in customer sentiment, a fine-tuned Logistic Regression model with class-weight balancing was implemented. This model demonstrated improved recall for negative and neutral reviews, making it more suitable for real-world customer feedback analysis. Topic modeling and trend analysis further revealed key factors influencing customer satisfaction and dissatisfaction. The findings from this project provide insights that can support product improvement, customer engagement, and proactive issue management.

---

2. Text Analysis & Sentiment Insights

2.1 Data Preprocessing and Cleaning

The raw review text was preprocessed to reduce noise and improve model performance. This included:

- Removal of HTML tags, special characters, and extra whitespace
- Conversion of text to lowercase
- Tokenization of words
- Removal of stopwords
- Lemmatization to normalize word forms

The cleaned and processed text was then converted into numerical form using TF-IDF vectorization to capture the importance of words across reviews.

2.2 Exploratory Text Analysis

Exploratory analysis revealed that the dataset is dominated by positive reviews, with a smaller proportion of neutral and negative reviews. Word frequency and word cloud analysis showed that positive reviews frequently mention terms related to *fit*, *comfort*, *fabric*, and *style*, while negative reviews commonly include words associated with *size issues*, *poor quality*, and *stitching problems*.

N-gram analysis (bigrams and trigrams) highlighted commonly occurring phrases such as “fits perfectly,” “good quality,” and “poor stitching,” providing deeper insight into customer opinions beyond individual words.

2.3 Topic Modeling Insights

Latent Dirichlet Allocation (LDA) was applied to extract major topics from customer reviews. The identified topics broadly corresponded to:

- Fit and sizing
- Fabric quality and comfort
- Product design and appearance
- Quality and durability issues
- Price and value for money

Topics related to fabric comfort and design were generally associated with higher ratings, while topics focusing on sizing inconsistencies and quality issues were associated with lower ratings. This indicates that fit and quality are critical drivers of customer satisfaction.

---

3. Classifier Performance & Findings

3.1 Model Comparison

Three machine learning models were evaluated for sentiment classification:

- Logistic Regression
- Logistic Regression with class-weight balancing

- Multinomial Naive Bayes
- Linear Support Vector Machine (SVM)

Naive Bayes showed strong bias toward the majority positive class and performed poorly on negative and neutral sentiments due to class imbalance. Standard Logistic Regression performed better but still showed bias toward positive sentiment.

To mitigate this issue, a balanced Logistic Regression model was trained using class weights. This significantly improved recall for negative and neutral reviews while maintaining reasonable performance for positive reviews.

3.2 Final Model Selection

The balanced Logistic Regression model demonstrated the best trade-off between overall accuracy and fair representation of all sentiment classes. The confusion matrix and normalized evaluation showed a substantial reduction in misclassification of negative reviews as positive, which is critical for identifying dissatisfied customers.

While Linear SVM also performed competitively, the balanced Logistic Regression model was selected as the final model due to its interpretability, stable performance, and ability to provide probability-based outputs useful for business decision-making.

3.3 Key Performance Observations

- Positive sentiment was identified with high accuracy
- Recall for negative and neutral sentiments improved significantly after class balancing
- Neutral reviews remained the most challenging due to overlapping language patterns
- The final model provided a more realistic and business-safe sentiment classification

---

4. Key Recommendations

4.1 Product Improvement

Analysis of negative sentiment and low-rated topics indicates recurring issues related to size inconsistency and product quality. Improving size standardization, enhancing quality checks, and ensuring accurate product descriptions can help reduce dissatisfaction.

4.2 Customer Communication

Neutral reviews represent customers who are neither fully satisfied nor dissatisfied. These customers can be targeted through follow-up communication, feedback requests, or personalized offers to improve their experience and convert them into positive customers.

Negative reviews should be prioritized for timely customer support responses to address complaints and reduce potential churn.

4.3 Crisis Management and Monitoring

Trend analysis of sentiment over time can help detect sudden spikes in negative sentiment, which may indicate product defects, supply chain issues, or service disruptions. Implementing automated sentiment monitoring and alert systems can enable proactive issue resolution before problems escalate.

4.4 Business Value

By combining sentiment analysis, topic modeling, and trend analysis, this system enables organizations to transform unstructured customer feedback into structured insights. These insights support data-driven decision-making for product development, marketing strategy, and customer relationship management.

Conclusion

This report demonstrates the end-to-end application of NLP techniques, machine learning modeling, and analytical reasoning to solve a real-world business problem. The approach balances technical rigor with practical interpretability, making the solution suitable for both academic evaluation and real-world deployment.