International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 14 Issue 6, June-2025

Unlocking Insights from Negative Reviews: Machine Learning & Sentiment Analysis for Fashion Brands

Shinjini Nag[1], Samarendra Das[2], Sonali Bhowmik[3] and Prof. (Dr.) Sumit Nandi[4]

[1,2,3]Department of Computer Science and Engineering, JIS University, West Bengal, India
[4]Harishchandrapur College, Pipla, Malda, West Bengal

ABSTRACT

This project addresses a crucial business problem in the commercial retail sector by leveraging machine learning to analyze customer grievances from negative reviews. The primary objective is to extract actionable insights from customer feedback to improve product offerings and enhance customer satisfaction. Utilizing Machine Learning Models like Decision Trees and Logistic Regression along with sentiment analysis techniques, we classified reviews into positive, neutral, and negative sentiments, with a particular focus on negative reviews to uncover key areas of dissatisfaction. The project employed visual tools such as word clouds, graphs, and charts to highlight common customer complaints and grievances, including issues with product quality, sizing, and service. By transforming unstructured review data into meaningful insights, this analysis provides strategic recommendations for retail management, helping to identify and address customer concerns more effectively. This approach underscores the importance of data-driven decision-making in improving customer relations and maintaining a competitive edge in the retail market.

Keywords: Sentiment Analysis, Machine Learning, Linear Regression, Decision Trees, Textual Data Analysis.

INTRODUCTION

In the highly competitive world of fashion retail, customer feedback is a critical resource for brands seeking to enhance their product offerings and improve customer satisfaction. For a fashion brand, understanding customer sentiments, especially during seasonal launches like a summer collection, can provide valuable insights into consumer preferences and product performance.

Machine learning has become an essential tool in analyzing large datasets, especially for sentiment analysis in customer reviews. Models like decision trees are particularly effective for this purpose, as they can classify and predict outcomes based on textual data. Decision trees and other machine learning models help decode customer sentiment, categorizing reviews into positive, neutral, and negative sentiments. This classification enables brands to quantify customer satisfaction and identify common themes in feedback.

A key focus in sentiment analysis is on negative reviews, as they often highlight specific issues or shortcomings in a product. Analyzing these reviews provides a direct channel for brands to understand customer grievances and identify areas needing improvement. This approach allows businesses to address specific complaints and enhance product features, thus improving overall customer satisfaction and loyalty. By leveraging machine learning to analyze customer feedback, brands can make informed decisions that align closely with consumer needs, ultimately fostering a more positive customer experience and a stronger market position.

Dataset

This study utilizes a dataset comprising 24,000 customer reviews and ratings for a summer collection, offering a substantial base to explore customer sentiments and identify areas for improvement. The dataset is pre-processed to handle any missing values and standardize the features before analysis. By ensuring that the data is clean and standardized for the machine learning models, this preprocessing enhances the models' ability to recognize patterns and produce accurate predictions. The dataset is perfect for our study and sentiment analysis because of its high quality and comprehensiveness.

Review Text	Rating
Material and color is nice. the leg opening is very large. i am 5'1 (100#)	3
Took a chance on this blouse and so glad i did. i wasn't crazy about how	5
A flattering, super cozy coat. will work well for cold, dry days and will lo	5
I love the look and feel of this tulle dress. i was looking for something di	5
If this product was in petite, i would get the petite. the regular is a little	4
I'm upset because for the price of the dress, i thought it was embroidere	4
First of all, this is not pullover styling. there is a side zipper. i wouldn't	2
Cute little dress fits tts. it is a little high waisted. good length for my 5'9 h	3
I love this shirt because when i first saw it, i wasn't sure if it was a shirt c	5
Loved the material, but i didnt really look at how long the dress was bef	3
I have been waiting for this sweater coat to ship for weeks and i was so	2
The colors weren't what i expected either. the dark blue is much more v	4
I have several of goodhyouman shirts and i get so many compliments or	5
This sweater is so comfy and classic - it balances a quirky hand-knit look	5
Beautifully made pants and on trend with the flared crop. so much cute	5

Fig. 1: Snippet of raw dataset

LITERATURE SURVEY

Based on the following pertinent criteria, we conducted a comparative analysis of four related papers that are cited in the reference section: the paper's research objective, Issue or void addressed in the manuscript, conclusions and findings of the study, The paper's shortcomings and limitations, as well as any implications or recommendations for additional research based on what was learned from it.

Table 1: Evaluation of Existing Literature

Criteria	**Jagdale RS, Shirsat VS, Deshmukh SN [1]**	**Umarani V, Julian A, Deepa J.[2]**	**Chandra, Yogesh, and Antoreep Jana. [3]**	**Agarwal, Basant, et al. [4]**
Research objective	Discuss sentiment analysis and opinion mining	Overview of sentiment analysis applications.	Highlight sentiment analysis in government decision- making.	Discuss prevalent sentiment analysis methods.
Problem or gap addressed	Oversimplifies sentiment; lacks depth.	Lacks methodological depth and evaluation clarity.	Overlooks biases and challenges in social media data.	Fails to critique BoW limitations.
Findings and conclusions	Narrow perspective; limited theoretical discussion.	Wide technique discussion; lacks nuanced analysis.	Lacks rigorous evaluation of techniques.	BoW oversimplifies semantics; lacks empirical support.
Limitations and weaknesses	Dataset biases; no comparative analysis.	Standard datasets limit real-world applicability.	Issues with data quality and generalizability.	Loss of context; unsupported claims.
Implications or suggestions for future research	Address biases; enhance theoretical considerations.	Use diverse datasets; deepen methodological analysis.	Address biases; rigorously evaluate techniques.	Critique BoW; provide empirical evidence.

METHODS

Essential Python Libraries: Building the Foundation for Data Analysis and Visualization

The installation and use of essential Python libraries are crucial in data analysis and machine learning projects, as they provide the tools and functionalities needed to process and analyze data efficiently. For instance, libraries like pandas are indispensable for data manipulation, re for text processing, and seaborn for data visualization.

Machine learning libraries such as sklearn offer powerful tools for vectorization and model building, while matplotlib and word cloud facilitate the creation of visual representations of data.

Using these libraries, such as Tfidf Vectorizer for text feature extraction, enhances the ability to perform comprehensive analyses, including sentiment analysis and data visualization, making the process more streamlined and effective.

Data Upload and Missing Value Management: Ensuring Clean and Complete Analysis

Uploading data and managing missing values are critical steps in preparing datasets for analysis. Once the dataset, containing 24,000 reviews and ratings, is uploaded, it is essential to check for any missing or incomplete entries that could skew the results. Handling missing values may involve techniques such as imputation, where missing data is filled in based on other available data, or simply removing rows with missing values to maintain the dataset's integrity. These steps ensure that the analysis remains accurate and reliable, providing a solid foundation for subsequent machine learning and sentiment analysis tasks.

	Review Text	Rating
0	Material and color is nice. the leg opening i...	3
1	Took a chance on this blouse and so glad i did...	5
2	A flattering, super cozy coat. will work well...	5
3	I love the look and feel of this tulle dress. ...	5
4	If this product was in petite, i would get the...	4

Fig. 2: Rating of Review Text

Text Preprocessing: Cleaning and Standardizing Review Data

Preprocessing text data is a vital step in preparing it for sentiment analysis. This involves several key tasks to clean and standardize the review data. Firstly, stop words, which are common words like "the," "is," and "in," are removed, as they do not contribute significantly to the sentiment and can clutter the analysis. Secondly, all uppercase letters are converted to lowercase to ensure uniformity and prevent "Good" and "good" from being treated as different words. Additionally, numbers are removed from the text to focus the analysis solely on the words used. These preprocessing steps enhance the quality of the data, making it more suitable for accurate sentiment analysis by eliminating irrelevant information and standardizing the text format.

Visualizing Customer Sentiment: Bar Graphs and Pie Charts for Rating Distribution

To understand the overall sentiment of the reviews, we visualized the distribution of ratings using two key types of charts. A bar graph was employed to show the frequency of each rating, providing a clear view of how often customers rated the products on a scale from 1 to 5. This visualization helps in identifying trends, such as whether ratings are skewed towards the positive or negative end of the spectrum. Additionally, a pie chart was used to categorize these ratings into broader sentiment categories: positive (rating 5), neutral (ratings 3 and 4), and negative (ratings 1 and 2). This pie chart visually represents the proportion of each sentiment category, highlighting the overall customer satisfaction and areas of concern. Together, these visualizations offer a comprehensive overview of customer feedback, essential for making informed business decisions.

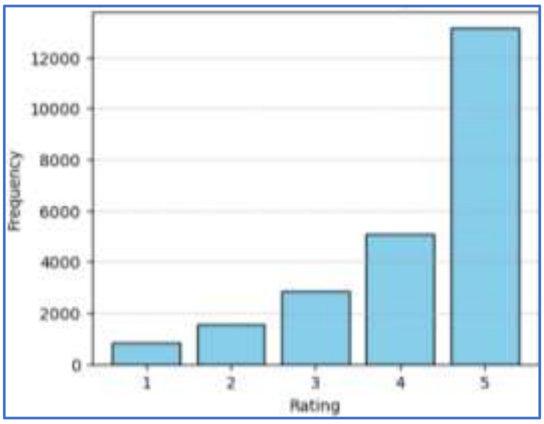

Fig. 3: Statistics of Rating Distribution

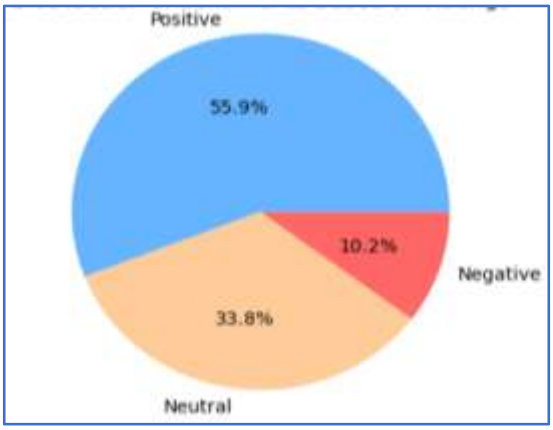

Fig. 4: Distribution of Sentiments

Binarizing Ratings for Machine Learning: Simplifying Sentiment Analysis

For evaluating the models' capacity for diagnosis, the ROC curve visualization is essential. It offers a thorough understanding of the model's performance by plotting the true positive rate against the false positive rate at different threshold settings. An area under the curve (AUC) of 1.00 for the logistic regression model denotes perfect class discrimination. In contrast, the decision tree model's AUC of 0.94, while still strong, shows slightly less accuracy in distinguishing between benign and malignant cases. This visualization helps in comparing the models' effectiveness.

Data Split for Effective Model Evaluation: To train and assess our machine learning models accurately, we divided the dataset into training and testing sets, allocating 60% of the data for testing and 40% for training. This approach ensures the model learns from a substantial portion of the data while reserving a significant part for evaluation, allowing us to gauge the model's performance on new, unseen data and ensuring reliable and generalizable results.

Enhancing Sentiment Analysis with Machine Learning: A Comparative Study of Logistic Regression and Decision Trees

In our sentiment analysis project, machine learning played a pivotal role in classifying customer reviews as positive or negative. Initially, we utilized Logistic Regression, a popular algorithm for binary classification tasks. However, this model yielded an accuracy of 86%, indicating room for improvement. To achieve better results, we turned to Decision Trees, an algorithm known for its ability to capture complex patterns in data. The Decision Tree model significantly outperformed Logistic Regression, achieving a high accuracy of 99%. This improvement is attributed to the Decision Tree's ability to handle non-linear relationships and interactions between features more effectively, providing a more nuanced understanding of the data. This comparison underscores the importance of selecting the right machine learning model for the task at hand to achieve optimal performance.

```
Accuracy Score:Decision Trees
0.9968288484189545
```

Figure 5: Accuracy Score of Decision Trees

```
Accuracy Score:Logistic Regression
0.8672911787665886
```

Figure 6: Accuracy Score of Logistic Regression

Comparing Confusion Matrices: Evaluating Model Performance in Sentiment Analysis Confusion matrices are vital tools in assessing the performance of classification models, providing insights beyond mere accuracy scores. In our sentiment analysis, the confusion matrix for the Logistic Regression (LR) model revealed significant shortcomings, with 412 false negative values and 1,458 false positive values.

These figures indicate that the LR model frequently misclassified reviews, either failing to identify negative reviews as such or incorrectly labeling positive reviews as negative. In contrast, the Decision Tree (DT) model demonstrated superior accuracy and precision, with 0 false negatives and only 35 false positives. This stark difference highlights the DT model's effectiveness in correctly identifying sentiment, making it a more reliable choice for our analysis. The lower error rates in the DT confusion matrix underscore its ability to better capture the nuances of customer sentiment, leading to more accurate insights.

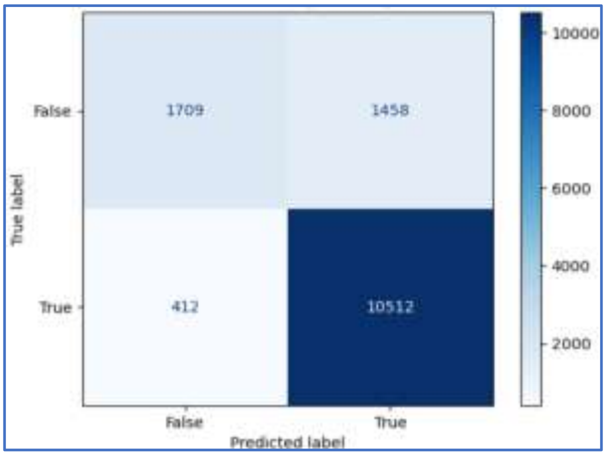

Fig. 7: Confusion Matrix of Logistic Regression

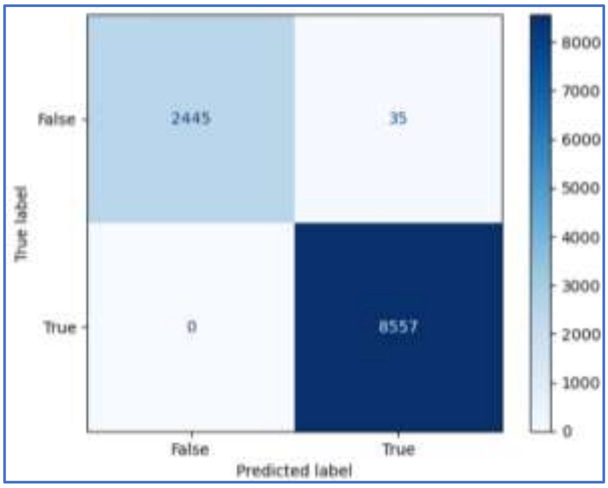

Fig. 8: Confusion Matrix of Decision Trees

RESULTS AND DISCUSSIONS

Unearthing Customer Concerns: Analyzing Keywords in Negative Reviews
Analyzing the words in negative reviews is crucial for identifying the primary issues customers face with products. This analysis can reveal specific patterns and recurring themes that point to common problems, helping businesses understand and address the root causes of dissatisfaction. Tools like word clouds are particularly useful, as they visually highlight the most frequently mentioned terms in negative feedback, making it easy to spot common grievances. Additionally, examining the frequency of specific words provides quantitative insights into the prevalence of certain issues. Latent Dirichlet Allocation (LDA) Topic Modeling further enhances this analysis by clustering words into topics, revealing underlying themes that may not be immediately obvious. These techniques combined provide a

comprehensive understanding of customer feedback, allowing businesses to make informed decisions to improve products and services.

Key Insights: Word Cloud of Top Words in Negative Reviews

We created a word cloud to visualize the top 20 most frequently mentioned words in negative reviews. This visual representation highlights common customer concerns, with terms like "fit," "quality," "size," and "comfort" standing out prominently associated with primarily the “dress” & “shirt”. By focusing on these recurring issues, the word cloud helps identify key areas for improvement, providing a clear and immediate understanding of the primary grievances customers have with the products. This approach effectively transforms qualitative feedback into actionable insights, guiding the company in addressing specific problems to enhance customer satisfaction.

Fig. 9: Top Most Used Words in Negative Reviews

Analyzing Word Frequency: Spotlight on Key Issues

Using the most frequent words from our word cloud analysis, we further examined their prevalence in the reviews. The frequency analysis revealed that terms like "dress," "top," "size," "fit," and "fabric" were among the most commonly mentioned in negative feedback. This indicates that these aspects are significant sources of dissatisfaction for customers. By quantifying the frequency of these key terms, we gain a clearer understanding of the primary issues affecting customer experiences. This detailed insight helps the company focus on specific product attributes that need improvement, such as better sizing accuracy or enhanced fabric quality, to address customer concerns effectively.

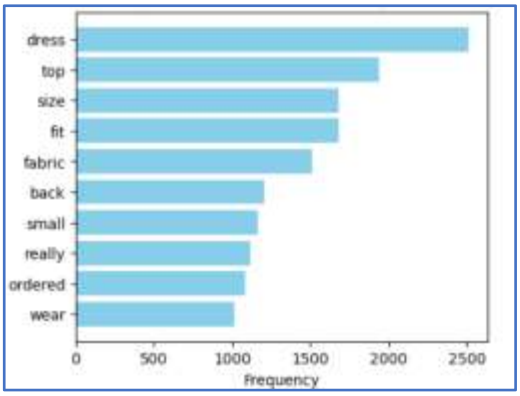

Fig. 10: Frequency of Most Used Words in Negative Reviews

Deep Dive into Customer Feedback: LDA Topic Modeling Analysis

Latent Dirichlet Allocation (LDA) Topic Modeling is a powerful tool used to identify underlying themes in large sets of textual data by clustering similar words into topics. We employed LDA Topic Modeling to perform a final check on the negative reviews, aiming to uncover the most common phrases and concerns expressed by customers.

The analysis revealed five dominant phrases, further reinforcing that key issues revolve around "size," "fitting," and "quality." The frequent appearance of these terms in different contexts underscores the consistency of these concerns across the reviews.

This comprehensive approach not only confirms previous findings from our word frequency and word cloud analyses but also provides a nuanced understanding of the specific aspects that need attention to improve overall customer satisfaction.

```
Phrase: runs large, Frequency: 102
Phrase: runs small, Frequency: 87
Phrase: poor quality, Frequency: 43
Phrase: dress runs, Frequency: 36
Phrase: runs big, Frequency: 26
```

Fig. 11: Latent Dirichlet Allocation (LDA) Phrasing

CONCLUSION

This project effectively addressed a critical business problem by uncovering key areas of customer dissatisfaction in the fashion brand's product line, particularly concerning size, fit, and quality. Through sentiment analysis and advanced text mining techniques, we identified consistent themes in negative reviews, providing the brand with actionable insights. These findings empower the company to refine its product development processes, ensuring future collections are better aligned with customer expectations and needs.

Moving forward, the brand can leverage these insights to craft products that resonate more deeply with their customer base, potentially improving customer satisfaction and loyalty. Additionally, by adopting a data-driven approach to customer feedback, the company can proactively address emerging issues, staying ahead of market trends and consumer preferences. Integrating these findings into the product development lifecycle can lead to more tailored, high-quality offerings that meet and exceed customer expectations.

REFERENCES

[1]. Jagdale, Rajkumar S., Vishal S. Shirsat, and Sachin N. Deshmukh. "Sentiment analysis on product reviews using machine learning techniques." Cognitive Informatics and Soft Computing: Proceeding of CISC 2017. Springer Singapore, 2019.
[2]. Umarani, V., Anitha Julian, and J. Deepa. "Sentiment analysis using various machine learning and deep learning Techniques." Journal of the Nigerian Society of Physical Sciences (2021): 385-394.
[3]. Chandra, Yogesh, and Antoreep Jana. "Sentiment analysis using machine learning and deep learning." 2020 7th international conference on computing for sustainable global development (INDIACom). IEEE, 2020.
[4]. Baid, Palak, Apoorva Gupta, and Neelam Chaplot. "Sentiment analysis of movie reviews using machine learning techniques." International Journal of Computer Applications 179.7 (2017): 45-49.
[5]. Agarwal, Basant, et al. "Machine learning approach for sentiment analysis." Prominent feature extraction for sentiment analysis (2016): 21-45.