Sentiment Analysis of Customer Reviews Using Machine Learning

Authors

Anurag Kumar Singh (22SCSE1010130)
Harsh Singh (21SCSE1011070)

Supervisor: Dr. Suveg Moudgil Sir (Department of CSE, Galgotias University)

Abstract

Sentiment analysis is a key area in natural language processing (NLP) that helps interpret customer opinions, feedback, and emotions expressed in textual data. This research explores multiple machine learning techniques for sentiment analysis of customer reviews. Various supervised learning models, including Naïve Bayes, Support Vector Machines (SVM), and deep learning methods like Long Short-Term Memory (LSTM) networks, are examined. The dataset comprises real-world customer reviews sourced from e-commerce platforms. Our study indicates that deep learning models, particularly LSTMs, outperform traditional machine learning models in sentiment classification. These insights can help businesses enhance customer experience through automated review analysis.

Keywords

Sentiment Analysis, Machine Learning, Natural Language Processing, Customer Reviews, Deep Learning, Text Classification

1. Introduction

With the rapid expansion of e-commerce and digital platforms, customers frequently share their opinions through reviews. Sentiment analysis plays a crucial role in extracting meaningful insights from these reviews. This research aims to compare multiple machine learning models in classifying customer sentiments as positive, negative, or neutral. Businesses use sentiment analysis to gauge customer satisfaction, enhance product offerings, and make datadriven decisions. Various machine learning and deep learning techniques have been employed to tackle sentiment classification challenges effectively.

2. Related Work

Several studies have investigated sentiment analysis using traditional machine learning models such as Naïve Bayes and SVM. More recent research has focused on deep learning techniques like LSTMs and transformers, which have demonstrated higher accuracy in text classification tasks. Previous research suggests that traditional machine learning models, when combined with feature engineering techniques like TF-IDF and word embeddings, yield promising results. However, deep learning approaches, particularly recurrent neural networks (RNNs) and transformers, provide superior performance due to their ability to capture sequential dependencies in text.

3. Methodology

3.1 Data Collection

The study utilizes a publicly available dataset containing customer reviews from various e-commerce platforms. The dataset includes labeled reviews categorized as positive, negative, or neutral. The dataset is preprocessed to remove noise and standardize text representation.

3.2 Preprocessing

To enhance model performance, the following preprocessing steps are applied:

- **Tokenization:** Splitting text into individual words or phrases.
- **Stopword Removal:** Removing frequently occurring words that do not contribute to sentiment (e.g., "the", "is").
- **Lemmatization:** Converting words into their base form (e.g., "running" → "run").
- **Feature Extraction:** Using TF-IDF and word embeddings (Word2Vec, GloVe) to convert text into numerical representations.

3.3 Machine Learning Models

The following models are implemented and compared:

- **Naïve Bayes (NB):** A probabilistic model based on Bayes' theorem, commonly used for text classification.

- **Support Vector Machine (SVM):** A classification model that finds an optimal hyperplane for separating sentiment classes.
- **Random Forest (RF):** An ensemble learning technique utilizing multiple decision trees to improve classification accuracy.
- **Long Short-Term Memory (LSTM):** A deep learning model designed for sequential text data, capturing long-range dependencies effectively.
- **Transformer-based models (e.g., BERT):** Advanced NLP models leveraging attention mechanisms for sentiment analysis.

3.4 Model Evaluation

Model performance is assessed using multiple evaluation metrics:

- **Accuracy:** Overall correctness of sentiment classification.
- **Precision, Recall, and F1-score:** Measures of classification performance considering false positives and false negatives.
- **Confusion Matrix:** Visual representation of model predictions versus actual sentiment labels.
- **ROC Curve & AUC Score:** Indicators of model robustness in distinguishing sentiment classes.

4. Results and Discussion

The experimental findings reveal that deep learning models, particularly LSTMs and transformer-based models like BERT, achieve superior accuracy compared to traditional machine learning models. While Naïve Bayes and SVM perform well with TF-IDF features, LSTMs utilizing word embeddings provide more accurate sentiment classifications. Transformer models like BERT further enhance sentiment analysis performance by leveraging attention mechanisms to capture contextual meanings in reviews. The results demonstrate that:

- **Naïve Bayes and SVM models** perform well for simple sentiment classification tasks but struggle with complex sentence structures.
- **LSTMs outperform traditional models** by capturing sequential dependencies in customer reviews.
- **Transformer models (BERT)** achieve the highest accuracy by understanding deep contextual representations of text.

5. Conclusion and Future Work

This study emphasizes the effectiveness of machine learning models in sentiment analysis of customer reviews. Deep learning models, especially LSTMs and transformer-based architectures, outperform traditional models in sentiment classification tasks. Future research may explore more advanced NLP techniques, such as reinforcement learning for sentiment adaptation and hybrid deep learning approaches. Additionally, integrating sentiment analysis with recommendation systems can enhance personalized customer experiences. Real-time sentiment monitoring using cloud-based AI models is another promising direction for further investigation.

6. References

1. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*.
2. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*.
3. Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
4. Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*.
5. Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. *arXiv preprint arXiv:1408.5882*.
6. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*.
7. Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. *Proceedings of EMNLP*.
8. Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global vectors for word representation. *Proceedings of EMNLP*.
9. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.

10. Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*.

11. Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011). Learning word vectors for sentiment analysis. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*.

12. Turney, P. D. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *Proceedings of ACL*.