Report on Naive Bayes Classifier for Sentiment Analysis

Overall Performance of the Classifier

The Naive Bayes classifier was evaluated on a test set containing 200 reviews, consisting of 100 positive and 100 negative samples. The following metrics were computed to assess the model's performance:

- **Confusion Matrix**:

[[75 25]

[18 82]]

    - True Positives (TP): 75 (correctly predicted positive reviews)
    - False Negatives (FN): 25 (incorrectly predicted negative reviews)
    - False Positives (FP): 18 (incorrectly predicted positive reviews)
    - True Negatives (TN): 82 (correctly predicted negative reviews)
- **Precision**:
    - Positive Class (pos): 0.8065
    - Negative Class (neg): 0.7664
- **Recall**:
    - Positive Class (pos): 0.7500
    - Negative Class (neg): 0.8200
- **F1 Score**:
    - Positive Class (pos): 0.7772
    - Negative Class (neg): 0.7923

The F1 scores indicate a balanced performance, with the negative class showing slightly better recall compared to the positive class. Overall, the model demonstrates effective classification of movie reviews, particularly in identifying negative sentiments.

Challenges Encountered

While implementing the Naive Bayes classifier, several challenges were faced by me:

1. **Text Preprocessing**: Proper text cleaning and tokenization were crucial for feature extraction. The removal of stop words and punctuation needed careful handling to avoid losing meaningful context.
2. **Handling Imbalance**: There was a need to ensure that the dataset had an equal representation of positive and negative samples to avoid bias in predictions.
3. **Computational Efficiency**: As the dataset grows, calculating probabilities for a large vocabulary can become computationally intensive, which required optimization considerations.
4. **Evaluation Metrics**: Understanding and implementing appropriate evaluation metrics, such as precision, recall, and F1 score, were essential for assessing the classifier's performance accurately.

Ideas for Improving the Model

Several strategies could enhance the performance of the Naive Bayes classifier for sentiment analysis:

1. **Using N-grams**: Instead of relying solely on unigrams (individual words), incorporating bigrams or trigrams could capture more context in the text, leading to better feature representation and improved classification.
2. **Improved Text Preprocessing**: Implementing more advanced text preprocessing techniques, such as stemming or lemmatization, could reduce vocabulary size and improve classification accuracy by focusing on the root forms of words.
3. **Feature Selection**: Using techniques such as TF-IDF (Term Frequency-Inverse Document Frequency) could provide more meaningful features by weighing the importance of words based on their frequency across the dataset.

Conclusion

The Naive Bayes classifier provided a solid foundation for binary sentiment classification, demonstrating reasonable accuracy and performance on the IMDB dataset. With further refinements and enhancements, its performance could be improved for more nuanced sentiment analysis tasks.