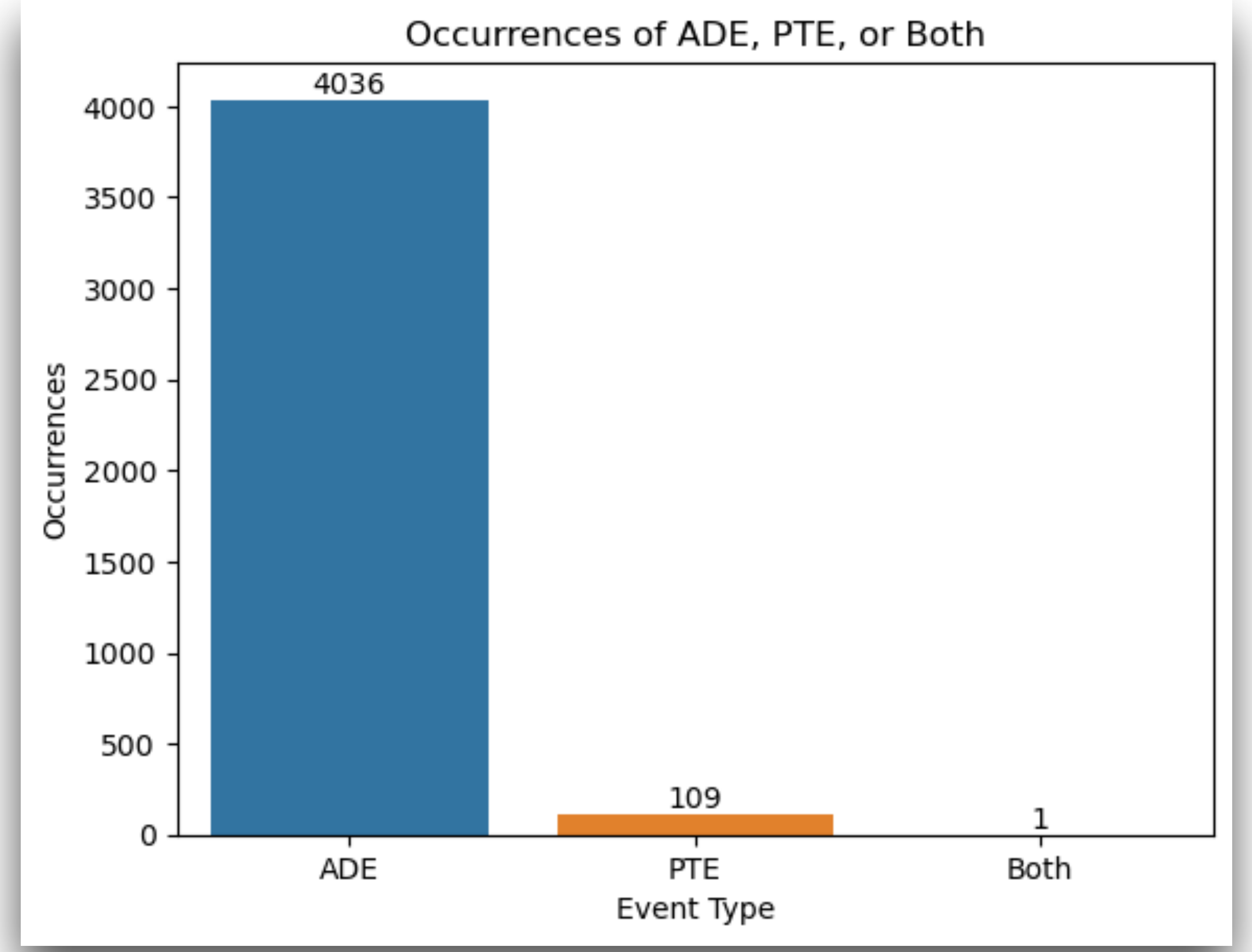
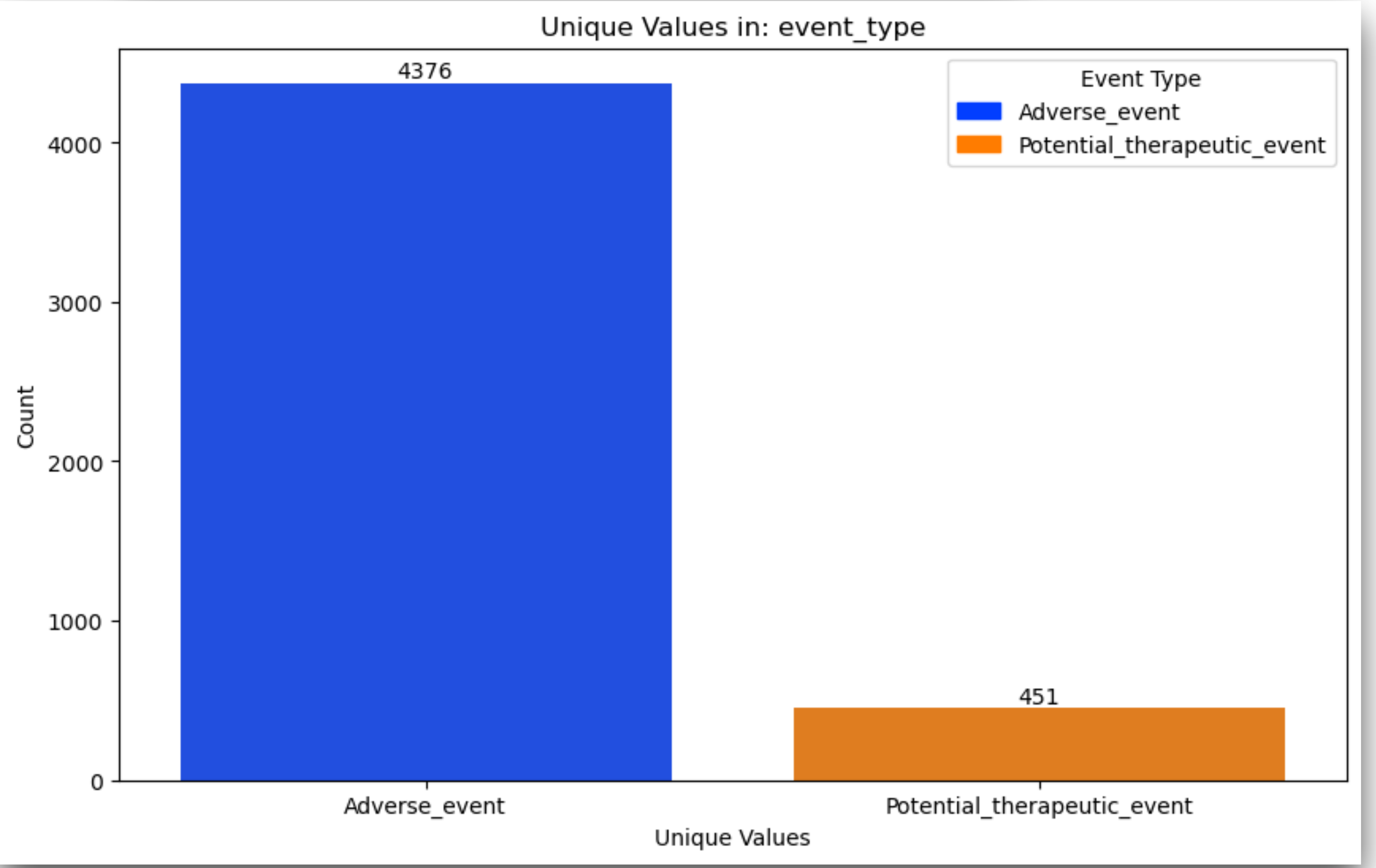
Pharmacovigilance NLP Sentiment Analysis

Exploring Sentiment Patterns in Adverse Drug Event Reports

Exploring Pharmacovigilance Data - Distribution Analysis

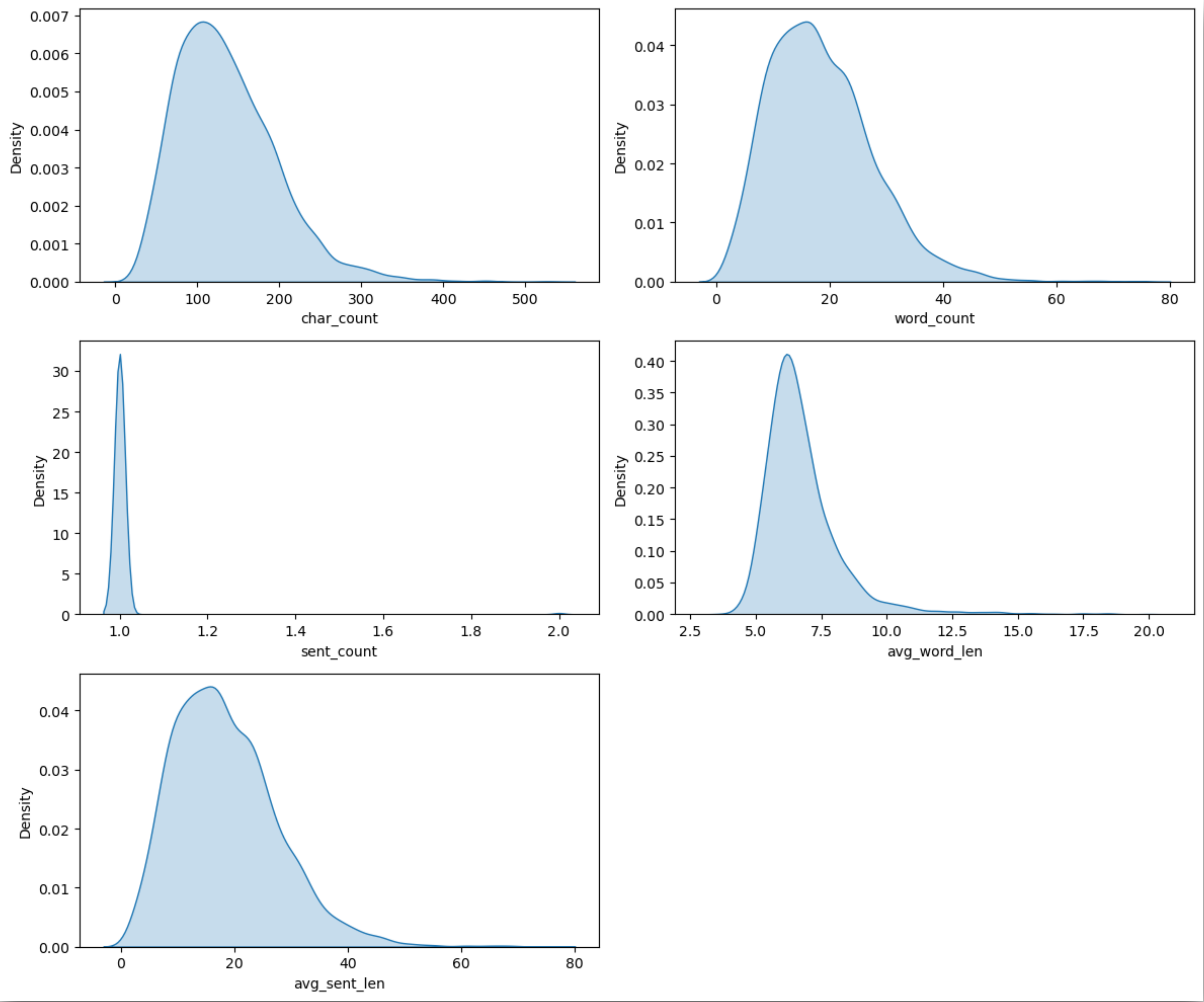
- **ADEs Prevalent**: With 4036 instances, Adverse Drug Events (ADEs) significantly outnumber Potential Therapeutic Events (PTEs), indicating potential safety concerns.
- **PTEs Scarce**: Only 109 instances of Potential Therapeutic Events were recorded, suggesting limited beneficial impacts of the drugs examined.
- **Dual Effects**: A single case showed both adverse and therapeutic effects simultaneously, pointing to complex drug interactions.
- **Need for Detailed Analysis**: The data calls for a comprehensive study, including considerations like event severity, medical conditions, dosage, and patient characteristics.
- **Class Imbalance**: The data exhibits a striking imbalance with 4376 ADEs against 451 PTEs, which could skew sentiment analysis towards ADEs.
- **Imbalance Impact**:The imbalance might lead to misleading conclusions about drug safety and efficacy, disproportionately classifying events as adverse.





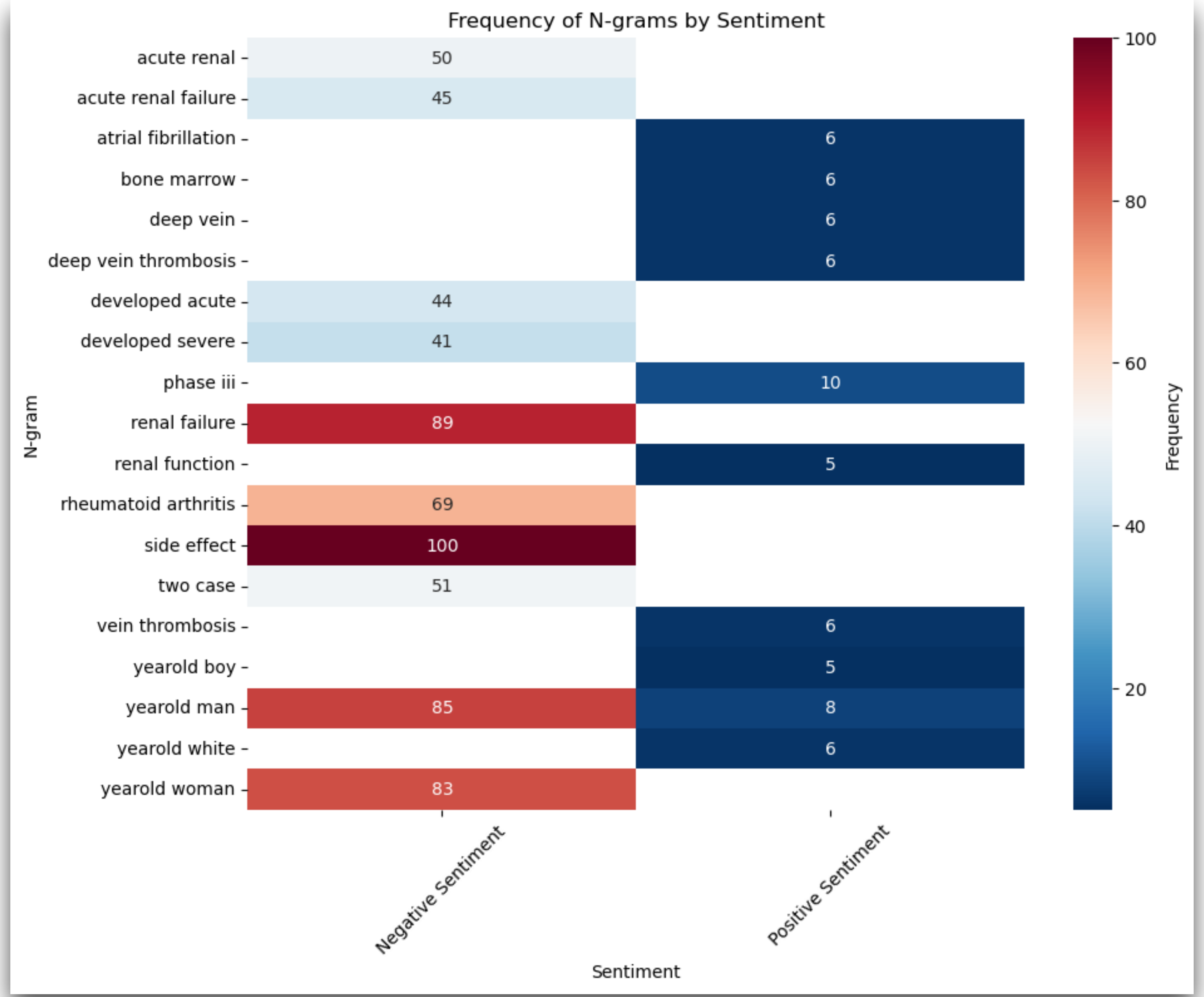
Insights from Sentiment Analysis and Linguistic Patterns

- **Sentiment:** Dataset is primarily negative with a smaller portion of positive sentiments.
- **Character Count:** Text length ranges from ~20 to 550 characters, showing variation.
- **Word Count:** Number of words ranges from 2 to 75, with moderate skewness.
- **Sentence Count:** Most texts have a single sentence, indicating consistent structure.
- **Avg. Word Length:** Average word length is ~6.75 characters, revealing linguistic patterns.
- **Avg. Sentence Length:** Average sentence length is ~18 words, with some longer sentences.



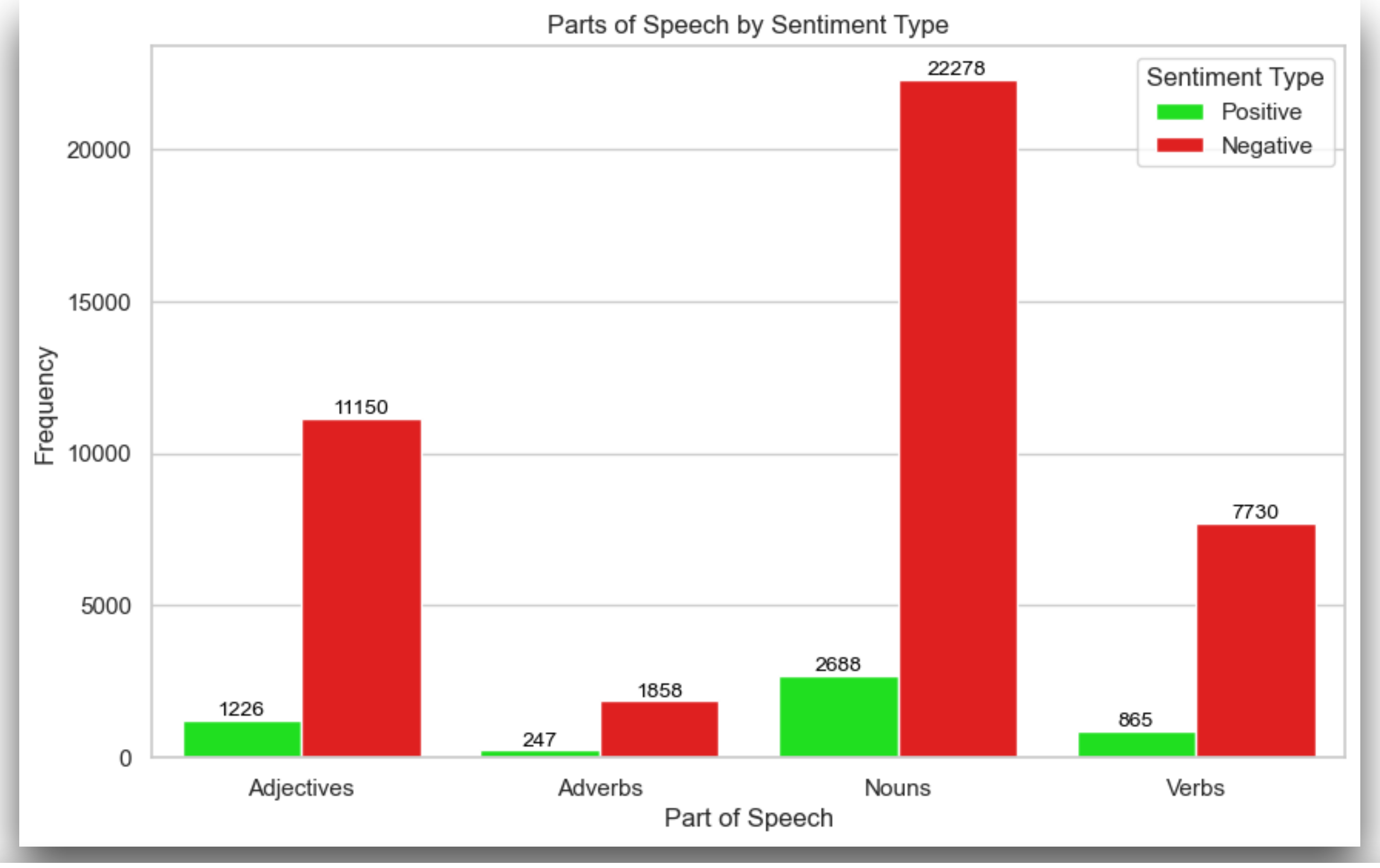
Insights from Sentiment Analysis and N-gram Patterns

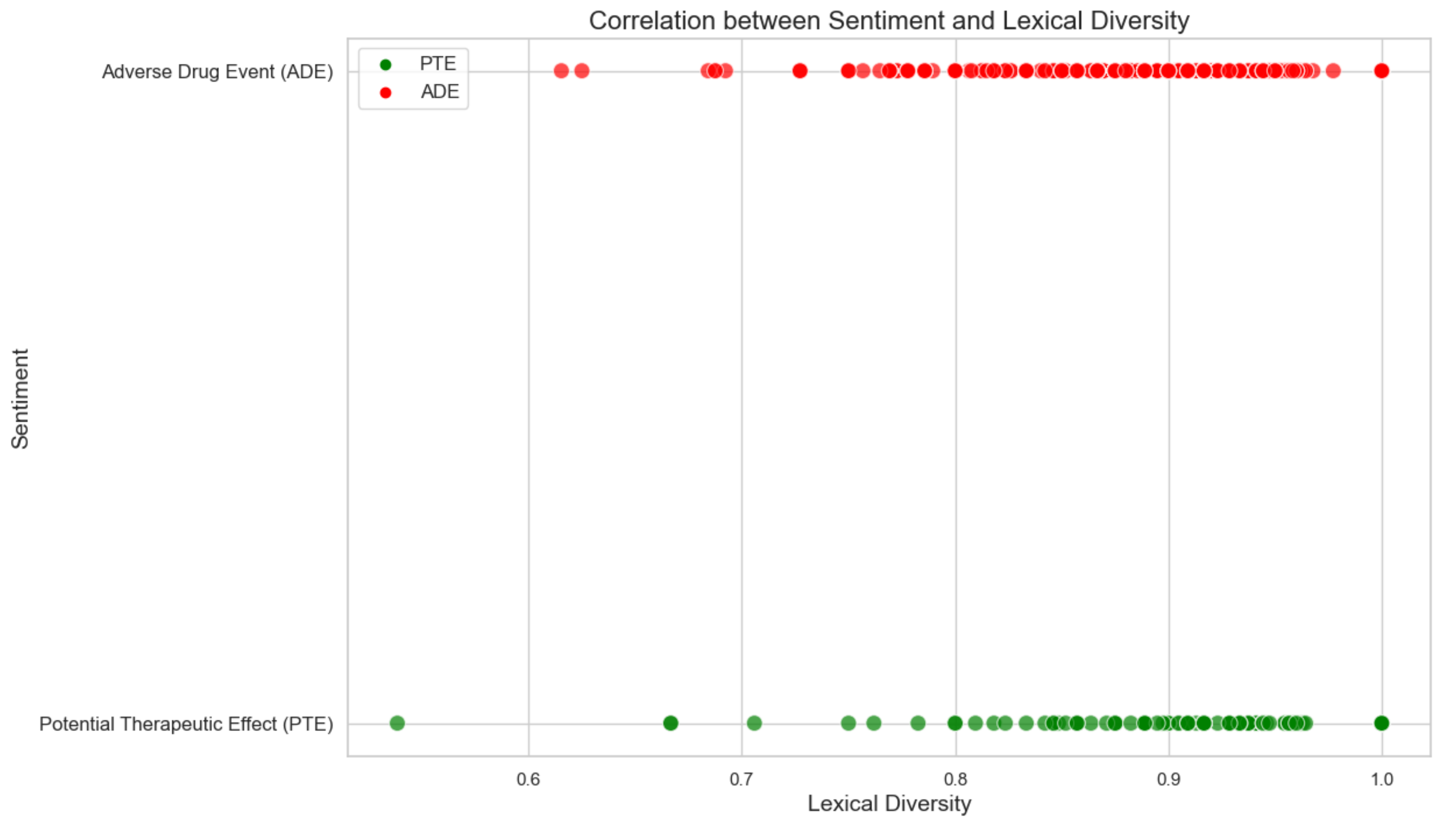
- **Top N-grams:** Negative - 'Side effect', 'Renal failure', 'Yearold man/woman', 'Rheumatoid arthritis'. Positive - 'Phase III', 'Yearold man', 'Yearold white', 'Vein thrombosis'.
- **High Frequency Negative N-grams:** 'Side effect' dominates negative sentiment, along with other N-grams indicating specific health conditions or demographics.
- **Low Frequency Positive N-grams:** 'Phase III' is the most frequent positive N-gram, but positive sentiments are less commonly discussed.
- **N-gram Overlap:** 'Yearold man' appears in both positive and negative sentiments, highlighting the importance of context analysis.
- **N-grams and Health Conditions:** N-grams like 'Renal failure' and 'Rheumatoid arthritis' are associated with negative sentiment, suggesting challenges in those areas.



Visual Patterns in Pharmacovigilance Reporting

- **Adjectives:** Low frequency in both positive and negative reports, indicating limited subjective descriptions.
- **Adverbs:** Slightly higher count in negative reports, emphasizing event intensity.
- **Nouns:** Higher frequencies overall, broader range in positive reports for adverse events.
- **Verbs:** Highest frequencies, especially in positive reports, highlighting active expression.
- **Lexical Diversity:** ADE sentiments show higher diversity, while PTE sentiments have even distribution.
- **Implication:** Accurate reporting of adverse events using specific nouns and verbs is crucial in pharmacovigilance.

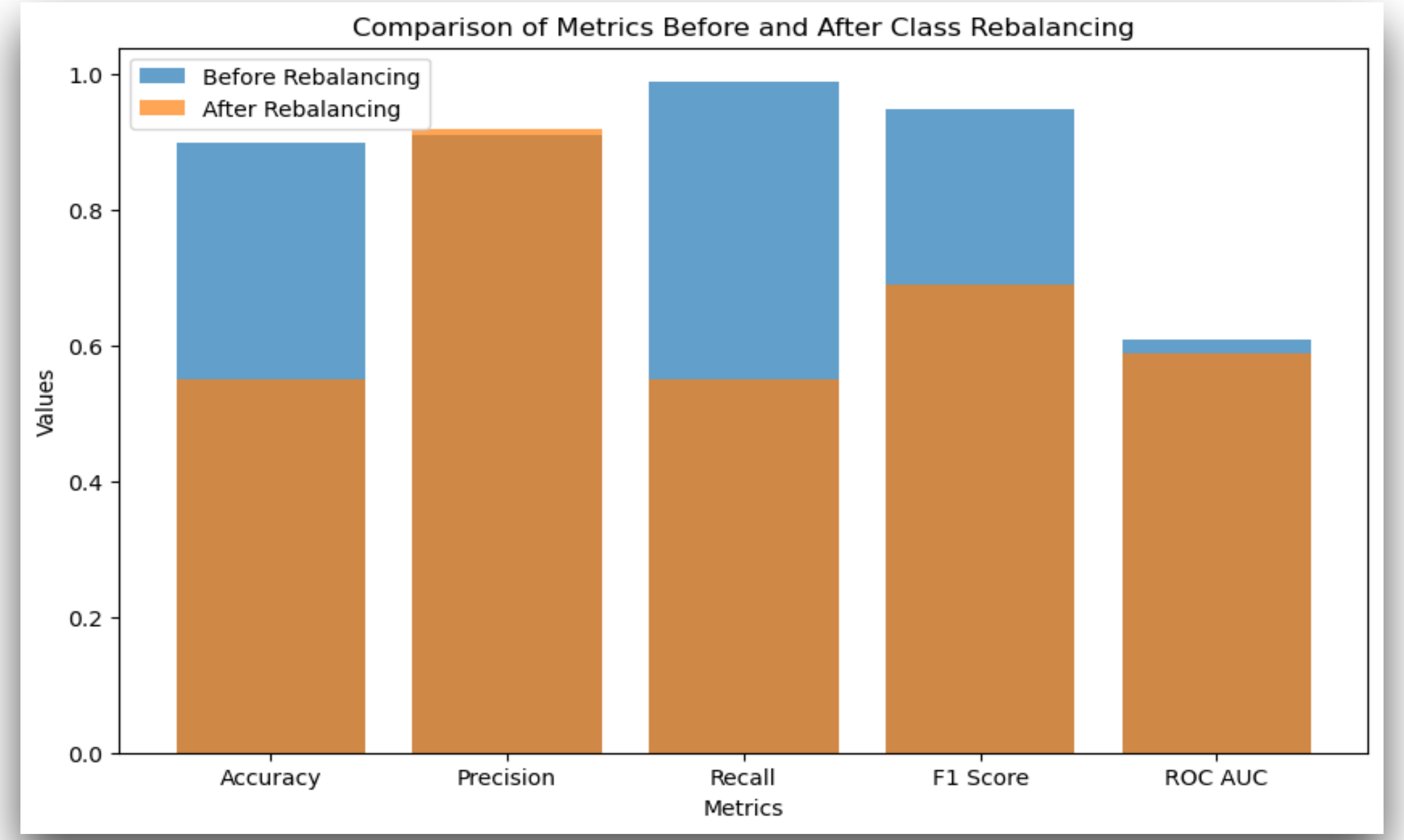


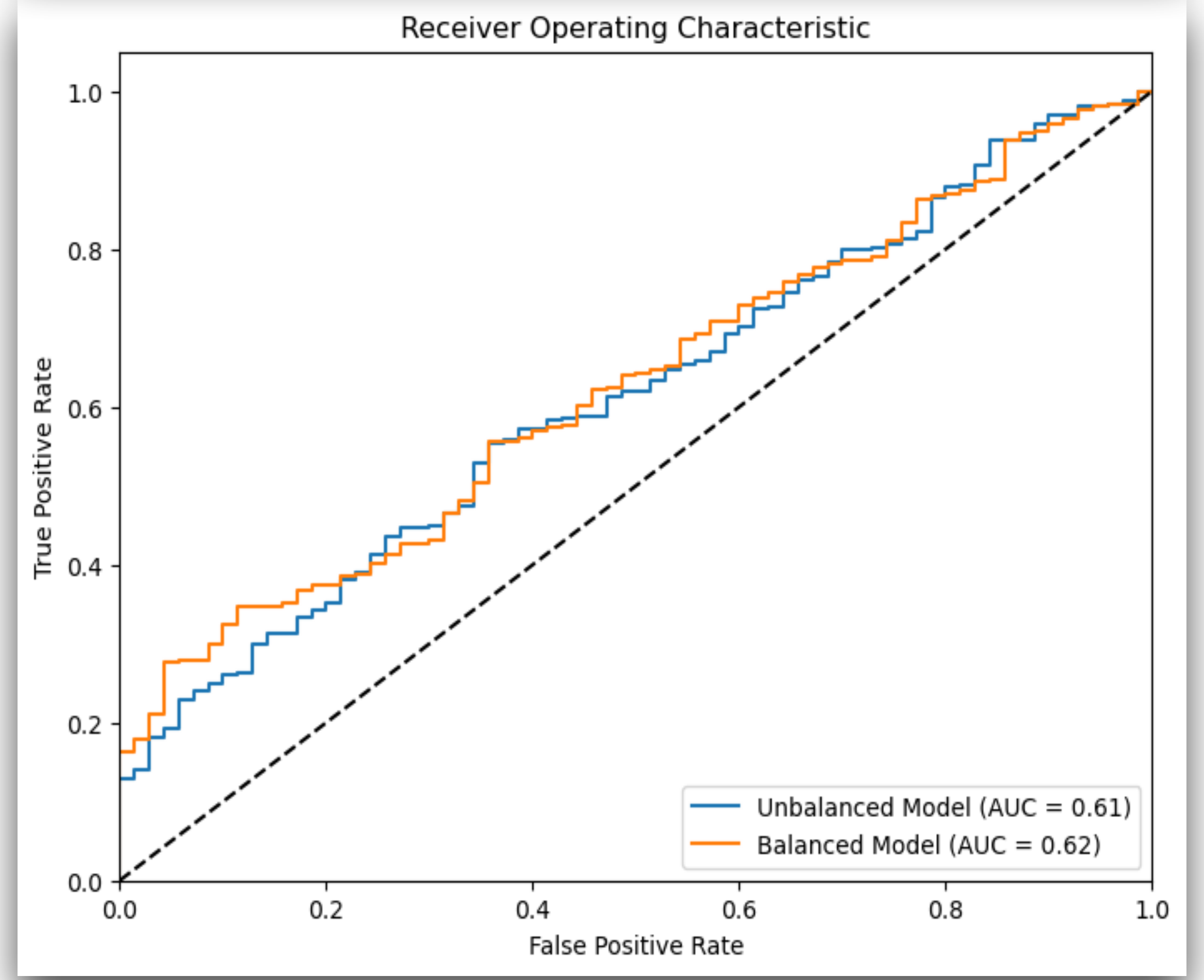


Higher lexical diversity indicates a greater variety of unique words within that document.

Unveiling Patient Insights for Safer Healthcare

- **Modeling**:
  - ***MultinomialNB*** was chosen for its effectiveness in handling text data and limited labeled data.
  - Initial Evaluation: Unbalanced model achieved a mean CV score of approximately 91%.
- **Evaluation**:
  - Class Imbalance Challenge: Unbalanced model struggled to identify positive sentiments (PTE) while performing well in identifying negative sentiments (ADE).
  - Applied SMOTE technique to rebalance the dataset.
- **Modeling (Continued)**:
  - Rebalanced Model Evaluation:
    - Decreased overall accuracy, recall, F1 score, and slight decline in ROC AUC compared to the unbalanced model.
    - Improved sensitivity in capturing PTE sentiments.
- **Conclusion**:
  - Unbalanced Model: Showcased superior overall predictive performance with higher accuracy and F1 score.
  - Rebalanced Model: Provided better sensitivity in identifying adverse events (ADE) at the expense of decreased overall accuracy.
- **Recommendations**:
  - Use the *unbalanced model* for high overall accuracy in predicting sentiments.
  - Choose the *rebalanced model* when accurate identification of adverse events (ADE) is crucial.





Pharmacovigilance Analysis: A Study on Cosine Similarity and XGBoost

- **Outstanding Training Results:** The model performed flawlessly on the training data, accurately predicting each instance. However, these perfect results might suggest the model is too finely tuned to the training data - overfitting, and might not generalize well to new data.

- **Solid Test Performance:** Despite potential overfitting, the model did very well on the test data, correctly identifying most instances of both categories. This shows the model's strong ability to handle new, unseen data.

- **Effective Techniques for Imbalanced Data:** The approach of using cosine similarity and XGBoost worked well for handling data with unequal class representation. It was able to clearly distinguish between classes and prioritize correctly identifying the minority class.

- **High Confidence for Real-World Use:** The model's high performance on both training and test data gives us strong confidence for its use in a real-world pharmacovigilance scenario, where accurate identification is crucial.

- **Validation and Improvement:** When tested with unseen data, the model accurately identified most instances (about 90.53% accuracy). However, it was better at identifying adverse drug events (ADEs) than positive therapeutic effects (PTEs), which was due to an imbalance in the training data favoring ADEs. This highlights the need to improve our model's ability to identify PTEs.

- **Importance of More Data:** Collecting more data, especially for underrepresented classes like PTEs, is key to improving the model's performance. Rather than adjusting the data statistically, gathering more real-world examples will give the model a more accurate understanding of the problem, leading to better performance in pharmacovigilance tasks.

Evaluating CNN: High Accuracy and Perfect Recall in Pharmacovigilance

- **Model Performance:** The Deep CNN model achieves high accuracy rates between 90.42%-91.75% in training and 87.88%-93.20% in validation, showcasing effective sentiment identification in the pharmacovigilance dataset.

- **High Precision & Recall:** The model boasts precision and recall rates of 0.91 and 1.00 respectively, indicating strong performance in minimizing false positives and successfully detecting all relevant cases - critical factors in pharmacovigilance.

- **Impressive F1 Score:** With an F1 score of 0.95, the model ensures a balanced performance between precision and recall, demonstrating overall effectiveness.

- **AUC-ROC - Area of Enhancement:** An AUC-ROC score of 0.63 suggests potential room for improvement in balancing false positives and negatives at varied thresholds, enhancing model reliability.