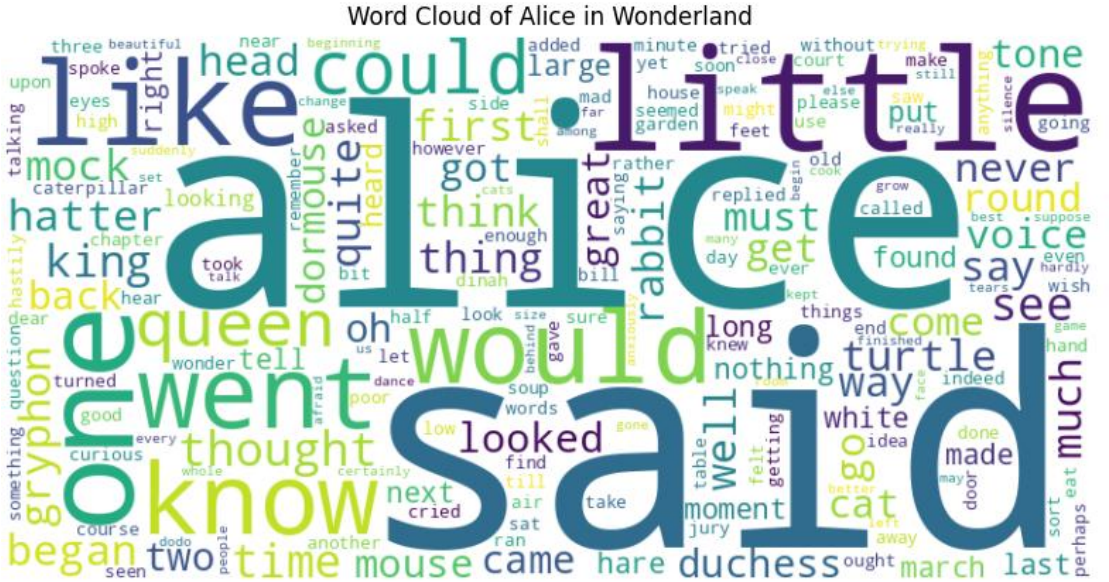
Alice in Wonderland NLP Analysis Report

1. Word Cloud

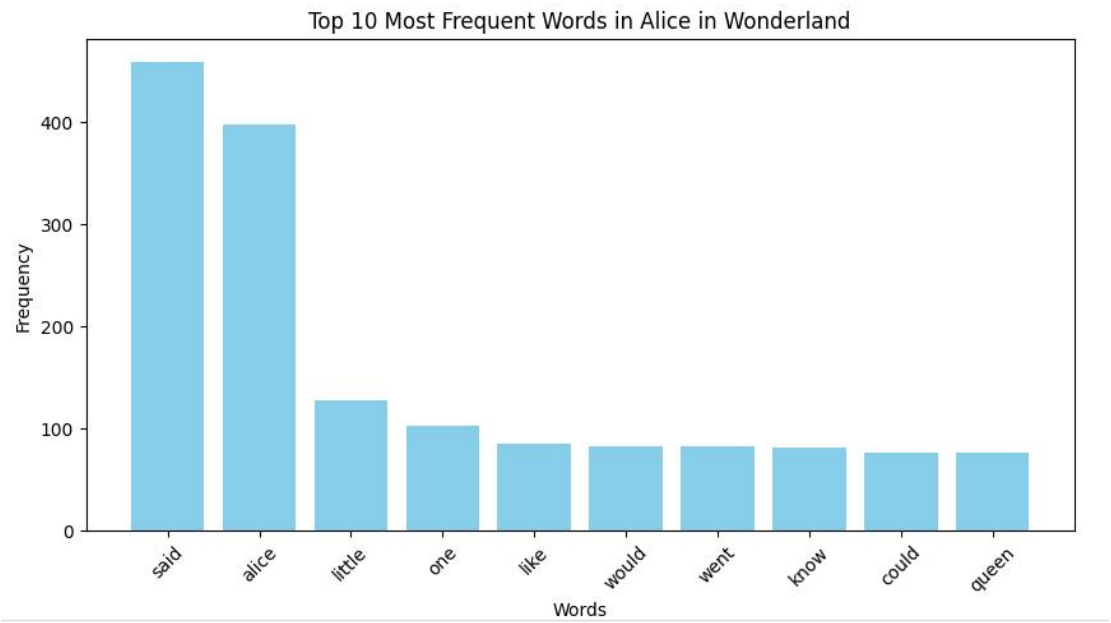
The word cloud provides a visual representation of word frequencies in *Alice's Adventures in Wonderland*. Words with higher frequencies appear larger. Key terms such as "Alice," "said," "queen," and "little" dominate, reflecting their central roles in the narrative and dialogue-driven nature of the text.



2. Word Frequency Bar Chart

Top 20 Most Frequent Words

This bar chart ranks the 20 most frequent words in the text. "Said" and "Alice" are the most prominent, underscoring the importance of dialogue and the protagonist. Other frequent words like "queen," "rabbit," and "king" highlight key characters integral to the story.

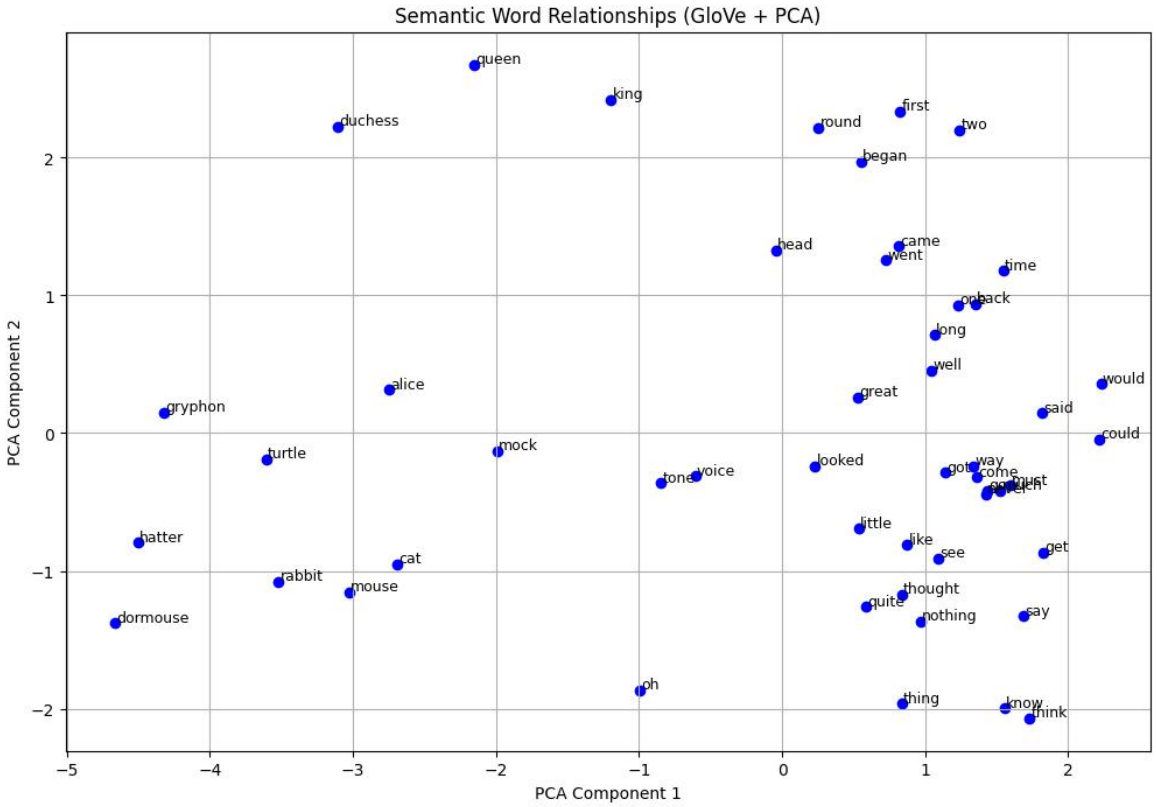


3. Semantic Word Relationships (PCA)

Using GloVe word embeddings reduced to two dimensions via PCA, this plot visualizes semantic relationships among words. Words with similar meanings or contexts cluster together. Notable clusters include:

- **Animals**: e.g., "rabbit," "turtle," "mouse"
- **Characters**: e.g., "queen," "king," "duchess"
- **Actions**: e.g., "said," "looked," "began"

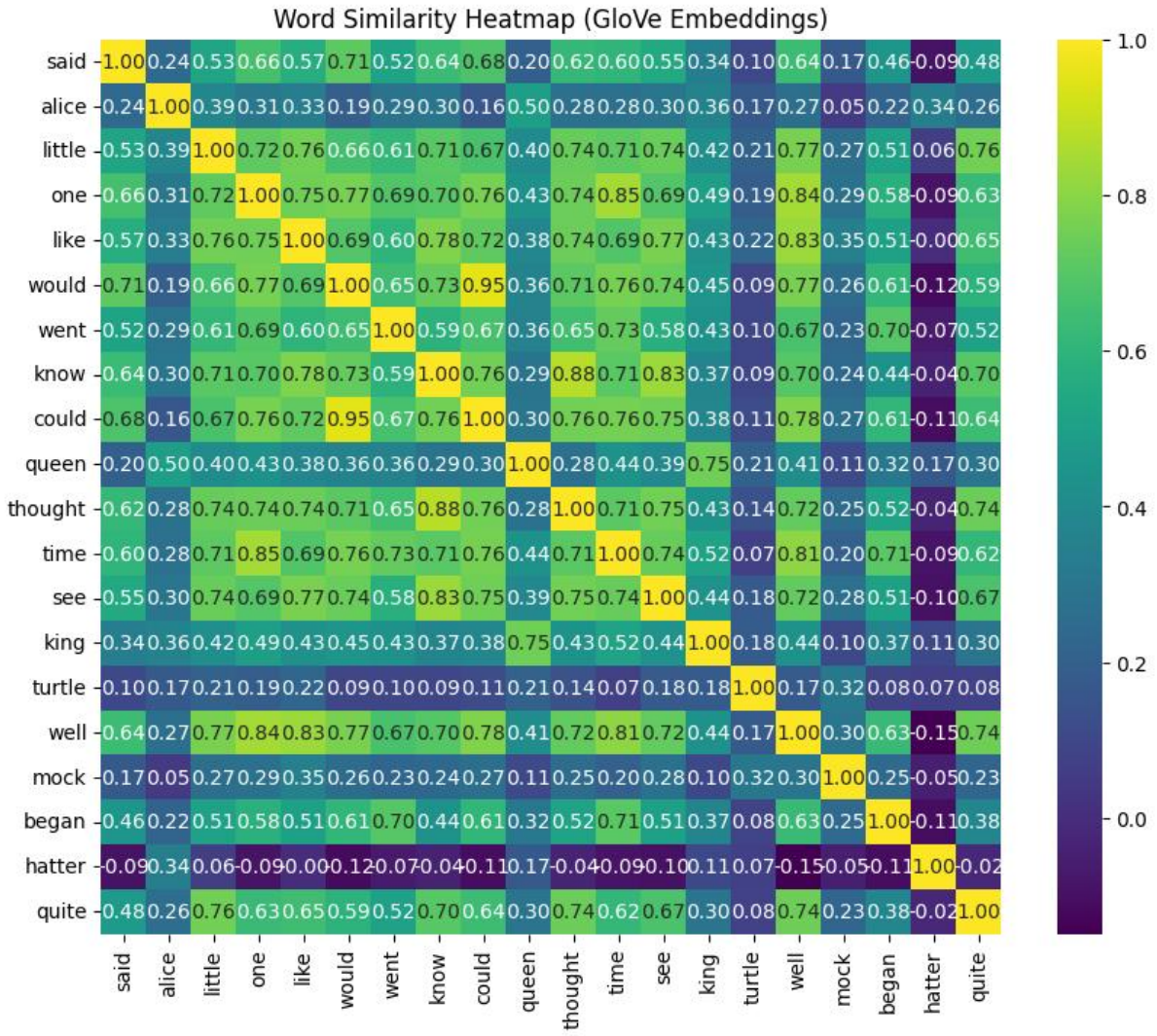
This clustering demonstrates how word embeddings capture thematic and contextual relationships in the text.



4. Cosine Similarity Heatmap

Cosine Similarity of Top 25 Words

The heatmap illustrates the cosine similarity between the top 25 most frequent words, derived from their GloVe embeddings. High similarity scores indicate close semantic relationships. For instance, "queen" and "king" exhibit strong similarity, reflecting their related roles. The heatmap helps identify synonyms, thematic connections, and character associations based on their contextual usage.



5. Discussion & Insights

The word cloud and frequency analysis confirm that "Alice" and dialogue (marked by "said") are central to *Alice's Adventures in Wonderland*. The PCA visualization reveals meaningful groupings of words, supporting the ability of word embeddings to capture semantic nuances. The cosine similarity heatmap further validates these findings by quantifying relationships, such as the close association between "queen" and "king" or "rabbit" and "mouse."

Collectively, these NLP techniques highlight the power of computational methods in extracting structure, themes, and relationships from unstructured text, providing deeper insights into the narrative and linguistic patterns of the book.