Sentiment and Theme Analysis in Taylor Swift's Lyrics Using NLP and Machine Learning Models

Abhrajit Das
Seidenberg School of Computer Science and Information Systems, Pace University, USA

Overview

- Music lyrics are a powerful medium for exploring emotions and societal trends.
- Taylor Swift's diverse discography, spanning genres and themes, serves as a rich dataset for analysis.
- This study focuses on sentiment trends and motifs like "day," "night," "love," and "hate."
- The analysis reveals how Taylor Swift's lyrical sentiment and themes evolve, reflecting her personal growth and changing artistic narrative.

Goal

- Understand sentiment evolution (positive, neutral, negative) in Taylor Swift's lyrics.
- Examine thematic patterns and their emotional alignment.
- Evaluate the effectiveness of machine learning models in lyrical sentiment classification.

Literature Review

- Sentiment analysis in music lyrics explores how lyrics reflect individual emotions and societal trends.
- Machine learning models are increasingly used to classify songs based on mood, genre, and commercial success.
- Studies on artists like Bob Dylan and The Beatles reveal how sentiment shifts mirror personal evolution and societal changes.
- Limited research exists on Taylor Swift's lyrics, despite her cultural influence and genre transitions.
- This study applies sentiment analysis to Taylor Swift's lyrics to understand her thematic evolution and emotional expression.

Workflow Diagram

Dataset Overview

- **Lyrics Dataset:** Taken from Kaggle which contains lyrics from all Taylor Swift albums up to October 20, 2022.
- Covers 8 Albums or 154 Songs.
- **Album Metadata Dataset:** Contains album names, release years, and versions (standard and deluxe).
- Tracks career progression and aligns lyrical data with corresponding years for analysis.

Results

Sentiment Patterns:

- Predominantly positive, with *1989 (Deluxe Version)* being the most optimistic (~0.8) and *Evermore (Deluxe Version)* the most melancholic (~0.01).

Themes of Time:

- "Day" appears more frequently (387) than "night" (287), reflecting themes of light and positivity, especially in earlier albums.
- Recent works show a shift toward darker and introspective themes, with mentions of "night" rising post-2022.

Key Observations

- **"Day" Sentiment**: Songs mentioning "day" have higher average sentiment scores (~34.36) compared to "night" (~20.20).
- **Common Themes**: Frequent words like "I" "you" and "we" underscore the relational and personal nature of her lyrics.

Correlations and Findings

- **Sentiment over Time**: A negative correlation
- (-0.57) between album release year and sentiment suggests a gradual shift toward more introspective tones.
- **Song Length and Sentiment**: Negligible correlation (0.0259) implies emotional tone is independent of song length.
- **Vocabulary Complexity**: No significant relationship (-0.0158) with sentiment scores

Machine Learning Models

Model	Accuracy	MSE
Logistic Regression	83.30	0.25
Naïve Bayes	76.41	0.32
SVM	88.37	0.21
Random Forest	90.45	0.15
Random Forest (After Hypertuning)	75.82	0.29
LSTM	88.04	0.06

Logistic Regression and SVM

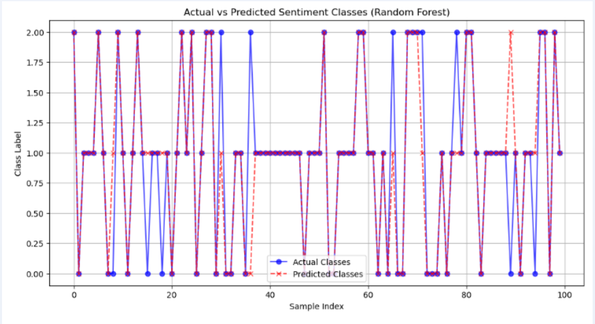
- Logistic Regression achieved 83.30% accuracy, serving as a reliable baseline but struggled with nuanced emotional shifts.
- Support Vector Machine (SVM) achieved 88.37% accuracy but had difficulty capturing negative sentiments, suggesting a need for better data representation.

Naïve Bayes and LSTM

- Naïve Bayes achieved 76.41% accuracy but struggled with sentiment imbalance, limiting its effectiveness in capturing complex emotional patterns.
- LSTM excelled with 95.16% training accuracy, capturing contextual nuances and long-term dependencies, making it ideal for analyzing Taylor Swift's lyrics.

Random Forest

- Random Forest achieved 90.13% accuracy, showing strength in handling non-linear relationships and feature interactions.
- Hyperparameter tuning with GridSearchCV caused Random Forest's performance to drop to 75.82%, likely due to overfitting.



Conclusion & Future Work

- The sentiment analysis shows Taylor Swift's lyrical evolution from optimism to introspection and maturity.
- Random Forest effectively classified sentiment categories, revealing thematic changes in her music.
- Future research could incorporate advanced deep learning models like Transformers for more nuanced sentiment analysis.