

# Exploración Emocional de Canciones a Través de Visual Analytics y Procesamiento Semántico

## Motivación y Fundamentación

En el contexto actual de consumo digital, las plataformas de música buscan constantemente ofrecer experiencias más personalizadas. Sin embargo, la mayoría de los sistemas de recomendación se enfocan exclusivamente en atributos acústicos, sin considerar el contenido lírico, a pesar de que las letras representan un medio poderoso para expresar emociones. Este proyecto surge del interés por aprovechar las letras de canciones como fuente primaria para analizar sentimientos, mediante herramientas modernas de procesamiento de lenguaje natural (NLP) y visual analytics.

Investigaciones previas han demostrado que el análisis de letras puede ser más eficaz que el análisis de audio para clasificar el estado de ánimo en la música (Hu & Downie, 2010). Además, estudios como el de Hutto y Gilbert (2014) muestran que bibliotecas como VADER son eficientes para extraer sentimientos en textos breves y coloquiales, como ocurre con las letras musicales. Complementariamente, herramientas como spaCy permiten un análisis más estructurado y profundo del texto (spaCy, 2020), mientras que el uso de modelos emocionales bidimensionales como el de Thayer (Yeh et al., 2009) permite una clasificación más granular de emociones.

Este proyecto también responde a la necesidad de representar visualmente los resultados del análisis de sentimientos. El uso de visual analytics —la combinación de visualización interactiva con algoritmos de ciencia de datos— ofrece una manera eficaz y comprensible de explorar grandes volúmenes de datos textuales (Cartwright, 2017). De esta manera, los usuarios pueden entender de forma intuitiva cómo se distribuyen las emociones en diferentes canciones, artistas o géneros musicales.

---

## Problema

Los actuales sistemas de análisis musical se centran predominantemente en características acústicas, dejando de lado la riqueza semántica y emocional que ofrecen las letras. A pesar de los avances en NLP y visual analytics, existen pocas herramientas que permitan analizar y visualizar sentimientos expresados en letras de canciones de forma interactiva y accesible. Esta falta de integración entre análisis textual, clasificación emocional y representación visual limita las capacidades de exploración, descubrimiento musical y personalización en plataformas digitales.

---

## Objetivo General

**Desarrollar una herramienta basada en visual analytics que permita realizar análisis de sentimientos en letras de canciones, utilizando bibliotecas de procesamiento de lenguaje natural y técnicas de clustering, y que represente los resultados mediante visualizaciones interactivas que apoyen la exploración y comprensión emocional del contenido musical.**

---

### Objetivos Específicos

1. **Aplicar técnicas de procesamiento de lenguaje natural utilizando VADER y spaCy para clasificar emociones en letras de canciones.**
  2. **Implementar un modelo de clasificación emocional multiclase basado en dimensiones de valencia y activación.**
  3. **Usar algoritmos de agrupamiento como K-Means para mejorar la asignación de etiquetas emocionales.**
  4. **Desarrollar una interfaz visual interactiva que permita explorar y comparar sentimientos por canción, artista o género.**
  5. **Evaluar la concordancia entre los resultados automáticos y etiquetas humanas para medir la efectividad del análisis**
- 

### Referencias

1. Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. (2011). Sentiment Analysis of Twitter Data. *Proceedings of the Workshop on Language in Social Media*, 30–38.
2. Cartwright, M. (2017). *Seeing Sound: Investigating the Effects of Visualizations and Complexity on Crowdsourced Audio Annotations*. Proceedings of the ACM on Human-Computer Interaction.
3. Gonçalves, P., Araújo, M., Benevenuto, F., & Cha, M. (2013). Comparing and combining sentiment analysis methods. *Proceedings of the ACM COSN*.
4. Hu, X., & Downie, J. S. (2010). When Lyrics Outperform Audio for Music Mood Classification: A Feature Analysis. *Proceedings of the International Society for Music Information Retrieval (ISMIR)*.
5. Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225.
6. Jamdar, A., Abraham, J., Khanna, K., & Dubey, R. (2015). Emotion Analysis of Songs Based on Lyrical and Audio Features. *International Journal of Artificial Intelligence & Applications (IJAIA)*, 6(3), 1–12.
7. Maas, A., Daly, R., Pham, P., Huang, D., Ng, A., & Potts, C. (2011). Learning Word Vectors for Sentiment Analysis. *Proceedings of the HLT Conference*, 142–150.
8. Shanmugapriya, K., & Srinivasan, B. (2015). An Efficient Method for Determining Sentiment from Song Lyrics Based on WordNet Representation Using HMM. *International Journal of Innovative Research in Computer and Communication Engineering*, 3(2).

9. Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45(4), 1191–1207.
10. Yeh, C.-H., Lin, H.-H., & Chang, H. (2009). An Efficient Emotion Detection Scheme for Popular Music. *International Journal of Computer and Information Science*.
11. spaCy. (2020). *Industrial-strength Natural Language Processing in Python*.  
<https://spacy.io/>