

---

# Análisis de letras de canciones en función de su país de origen

---



Minería de datos y el paradigma Big Data  
Curso 2018–2019

Autor

Alejandro Díaz Román

Jaime Del Rey García

Paula Munoz Lago

Facultad de Informática

Universidad Complutense de Madrid



# Índice

<b>1. Introducción</b>	<b>1</b>
<b>2. Objetivos</b>	<b>3</b>
<b>3. Metodología</b>	<b>5</b>
<b>4. Implementación</b>	<b>7</b>
4.1. Desarrollo de la aplicación . . . . .	7
4.2. Creación del conjunto de datos final . . . . .	7
4.3. Limpieza y preprocesamiento de los datos . . . . .	7
4.4. Transformación de los datos . . . . .	7
4.5. Elección de la tarea de minado adecuada . . . . .	7
4.6. Elección del algoritmo de minería de datos . . . . .	7
4.7. Aplicación del algoritmo elegido . . . . .	7
4.8. Interpretación de los resultados . . . . .	7
4.9. Aplicación de los resultados obtenidos . . . . .	7
<b>5. Resultados y Evaluación</b>	<b>9</b>
<b>6. Conclusiones y Trabajo Futuro</b>	<b>11</b>



# Capítulo 1

## Introducción

*“Frase célebre dicha por alguien inteligente”*

— Autor

El auge de la producción musical en los últimos años ha propiciado un aumento en el volumen de los datos, de los cuales se puede extraer el contexto social y cultural del país de procedencia del autor.



## Capítulo 2

# Objetivos

En esta práctica, nuestro objetivo es extraer, además de las palabras más utilizadas de cada autor, y de cada país, es extraer el sentimiento que prima en dicho país. Partimos de la hipótesis de que la temática principal de las letras de las canciones españolas es el amor, y lo comprobaremos como parte de esta práctica. El alcance de nuestro trabajo lo hemos establecido en función al tiempo del que disponemos, por lo que no pretendemos cubrir la totalidad de los países, sino los más significativos o de los que dispongamos la información suficiente.





## Capítulo 3

# Metodología

Para realizar la investigación hemos barajado las distintas metodologías y hemos seleccionado KDD (knowledge discovery in databases). Este modelo está directamente relacionado con la extracción de información y conocimiento de los datos disponibles. Sigue un proceso iterativo que, explorando los datos, extrae las distintas relaciones. <http://www.webmining.cl/2011/01/proceso-de-extraccion-de-conocimiento/>



## Capítulo 4

# Implementación

- 4.1. Desarrollo de la aplicación
- 4.2. Creación del conjunto de datos final
- 4.3. Limpieza y preprocesamiento de los datos
- 4.4. Transformación de los datos
- 4.5. Elección de la tarea de minado adecuada
- 4.6. Elección del algoritmo de minería de datos
- 4.7. Aplicación del algoritmo elegido
- 4.8. Interpretación de los resultados
- 4.9. Aplicación de los resultados obtenidos



## Capítulo 5

# Resultados y Evaluación



## Capítulo 6

# Conclusiones y Trabajo Futuro

Conclusiones del trabajo y líneas de trabajo futuro.





*–¿Qué te parece desto, Sancho? – Dijo Don Quijote –  
Bien podrán los encantadores quitarme la ventura,  
pero el esfuerzo y el ánimo, será imposible.*

*Segunda parte del Ingenioso Caballero  
Don Quijote de la Mancha  
Miguel de Cervantes*

*–Buena está – dijo Sancho –; fírmela vuestra merced.  
–No es menester firmarla – dijo Don Quijote–,  
sino solamente poner mi rúbrica.*

*Primera parte del Ingenioso Caballero  
Don Quijote de la Mancha  
Miguel de Cervantes*

