



PREDICCIÓN DE EMOCIONES A PARTIR DE CARACTERÍSTICAS DE AUDIO MUSICAL UTILIZANDO TÉCNICAS DE APRENDIZAJE MÁQUINA

Autor: Víctor Iglesias Cuevas

Tutora: Rebeca Goya Esteban

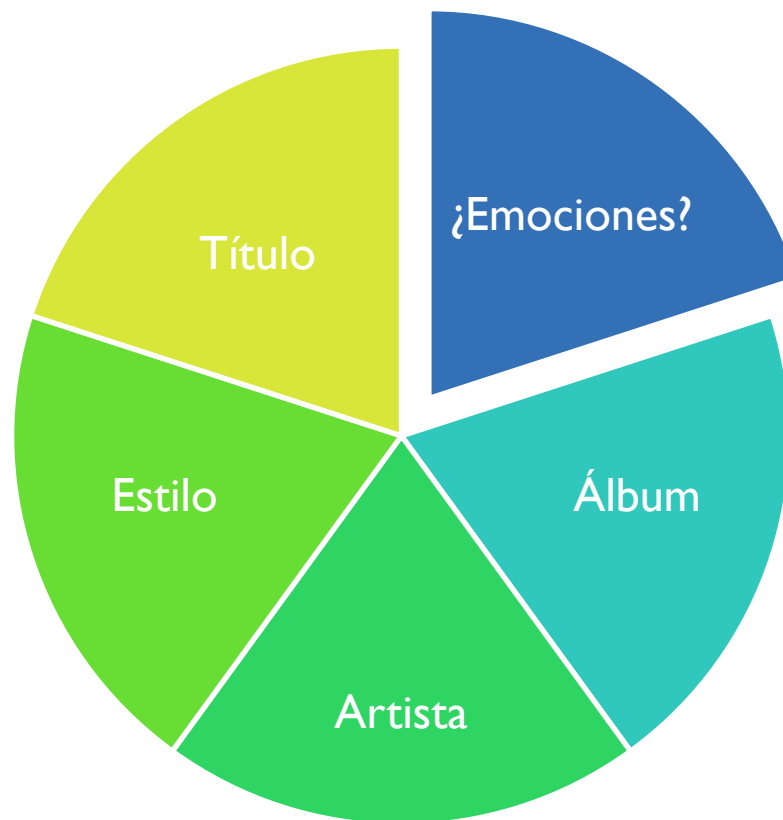
Índice de contenidos

1. Motivación y contexto
2. Objetivos
3. MER
4. Machine Learning
5. Conjunto de datos
6. Desarrollo
7. Resultados
8. Conclusiones

Índice de contenidos

1. Motivación y contexto
2. Objetivos
3. MER
4. Machine Learning
5. Conjunto de datos
6. Desarrollo
7. Resultados
8. Conclusiones

Motivación y contexto



Índice de contenidos

1. Motivación y contexto
2. Objetivos
3. MER
4. Machine Learning
5. Conjunto de datos
6. Desarrollo
7. Resultados
8. Conclusiones

Sistema de reconocimiento de emociones en pistas de audio

MER

Machine Learning

Python

Índice de contenidos

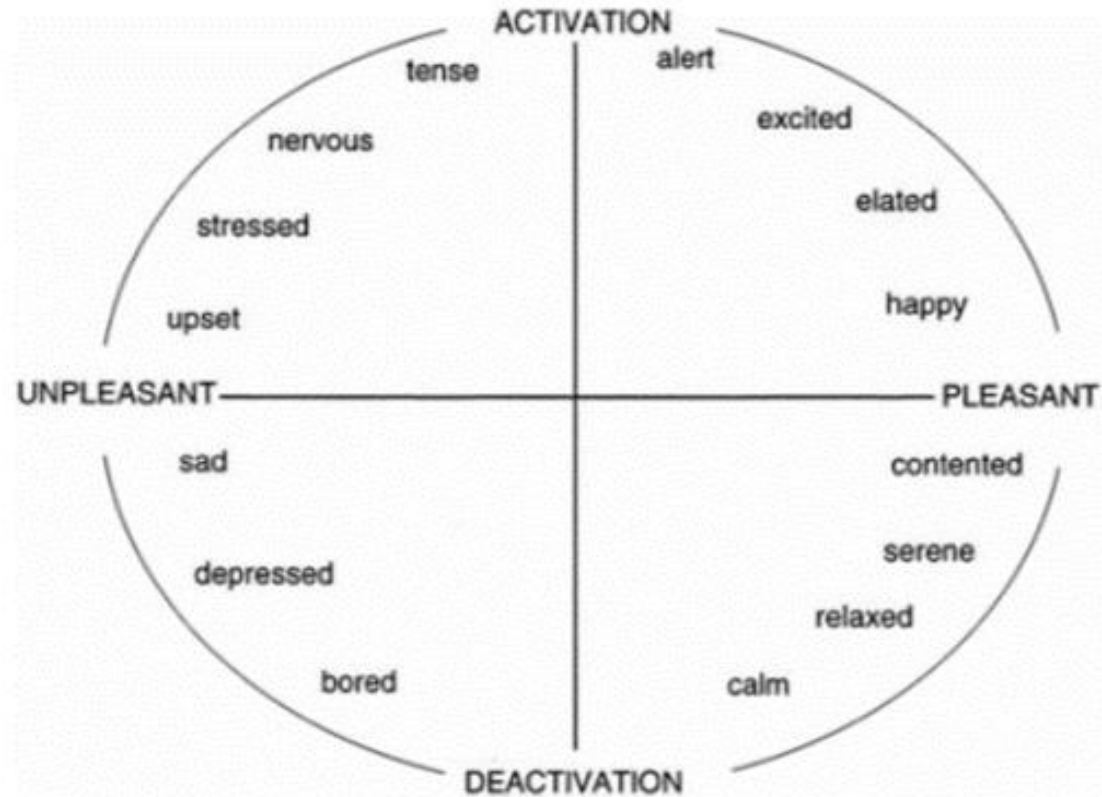
1. Motivación y contexto
2. Objetivos
3. MER
4. Machine Learning
5. Conjunto de datos
6. Desarrollo
7. Resultados
8. Conclusiones

- *Music Emotion Recognition*
- Identifica y clasifica las emociones que la música evoca en el oyente
- Disciplina de *Music Information Retrieval* (MIR)

Emociones

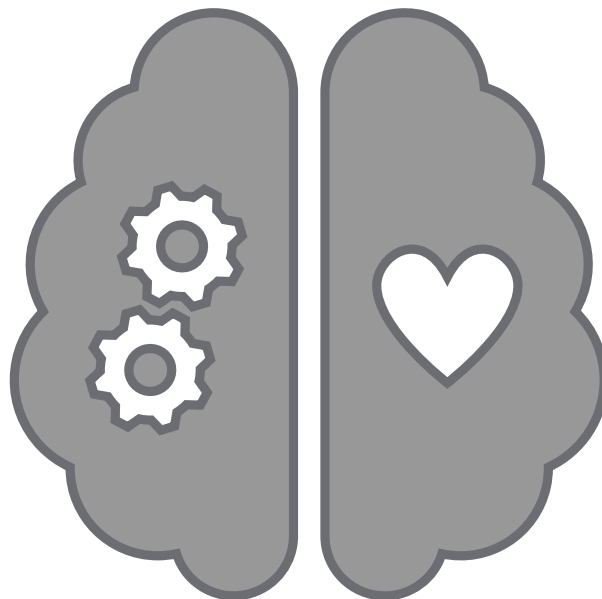
- Modelos teóricos de la emoción:
 - Discretos
 - Dimensionales
 - Misceláneos
 - Específicos de la música

Modelo dimensional de Russell



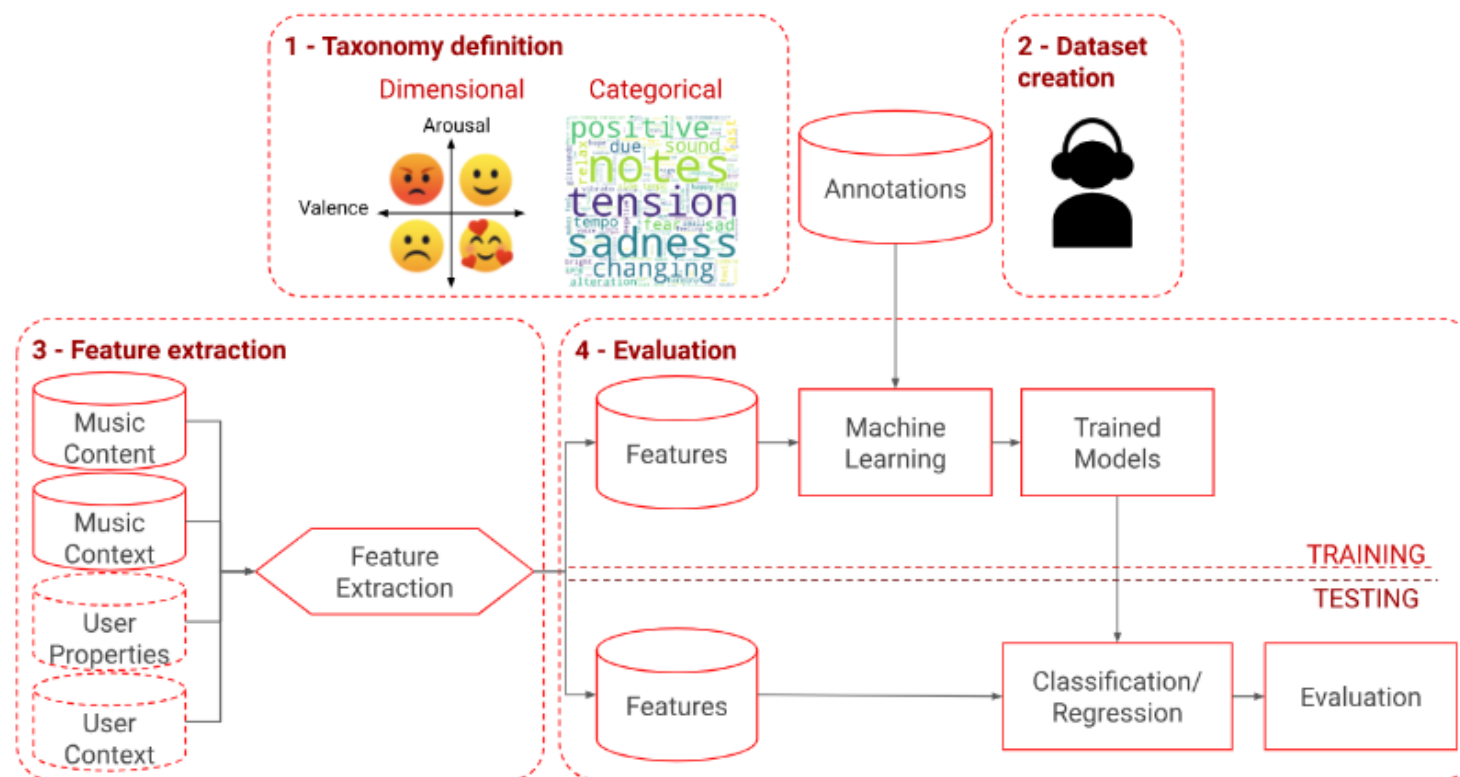
Emociones

Percibidas



Inducidas

Sistema tradicional



Índice de contenidos

1. Motivación y contexto
2. Objetivos
3. MER
4. Machine Learning
5. Conjunto de datos
6. Desarrollo
7. Resultados
8. Conclusiones

Machine Learning

Regresión

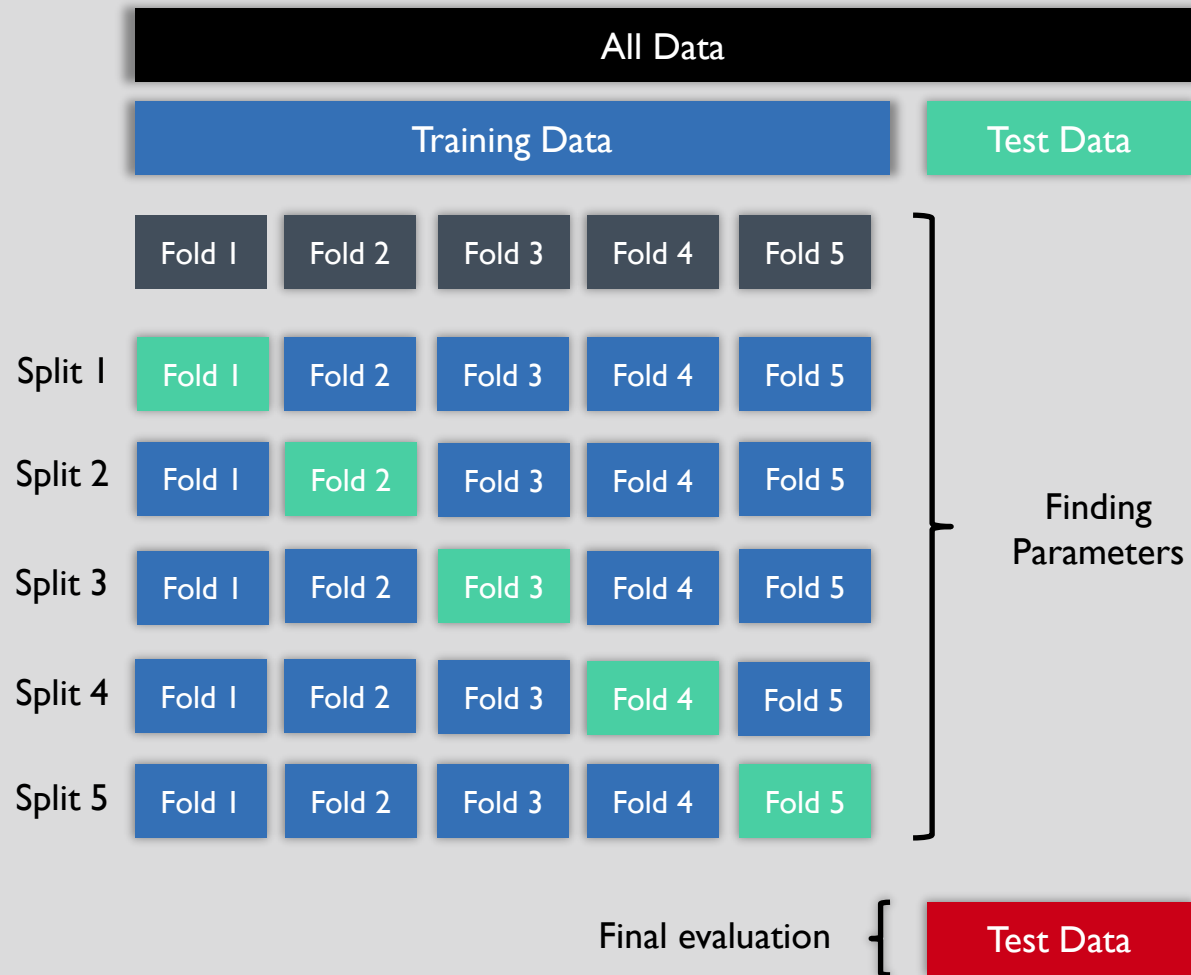
Árbol de decisión

Bosque aleatorio

Regresión lineal

Regresión ridge

Machine Learning



Cross-validation

- K-fold
- Evalúa capacidad predictiva del modelo
- Evita sobreajuste (overfitting)

Índice de contenidos

1. Motivación y contexto
2. Objetivos
3. MER
4. Machine Learning
5. Conjunto de datos
6. Desarrollo
7. Resultados
8. Conclusiones

Conjunto de datos

DEAM

- DEAM dataset - The MediaEval Database for Emotional Analysis of Music
- 1800 extractos de canciones
- 45s en ventanas de 500ms (descartando los primeros 15s)

Características

Metadatos

**Valencia
y
activación**

Conjunto de datos

Características

☐ openSMILE

☐ Media y desviación estándar

☐ Más de 250 valores

- Energía
- Intensidad de fotograma
- Coeficientes cepstrales Mel y Bark
- Sonoridad
- Coeficientes predictivos lineales (LPC)
- Pares espectrales de líneas
- Frecuencia fundamental
- Coeficientes perceptivos lineales predictivos
- Jitter
- Shimmer
- Tasa de cruce por cero
- Frecuencia fundamental
- Armonía espectral
- CROMA
- Relaciones de armónicos
- Probabilidad de emisión de voz

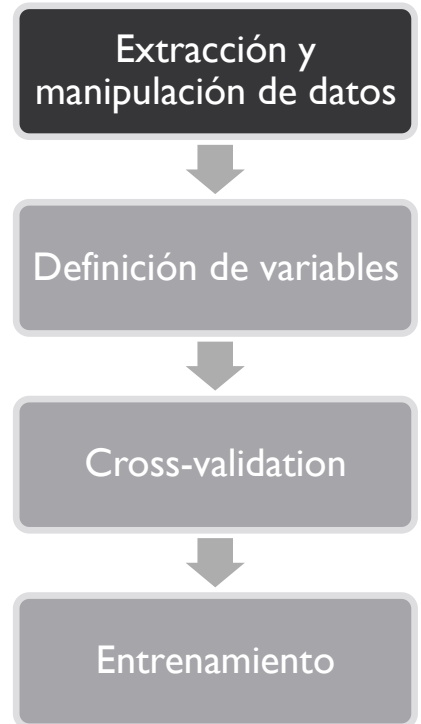
Índice de contenidos

1. Motivación y contexto
2. Objetivos
3. MER
4. Machine Learning
5. Conjunto de datos
6. Desarrollo
7. Resultados
8. Conclusiones

Extracción y manipulación de datos

Se unifican los datos para crear una tabla con todas las características junto con los valores de valencia y activación

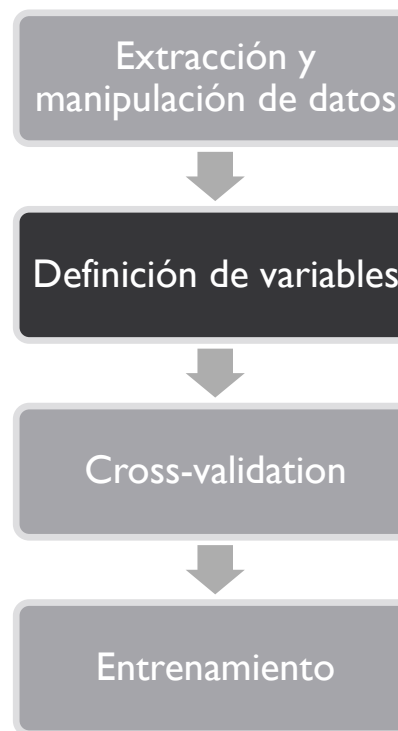
pcm_fftMag_mfcc_sma_de[13]_stddev	pcm_fftMag_mfcc_sma_de[13]_amean	pcm_fftMag_mfcc_sma_de[14]_stddev	pcm_fftMag_mfcc_sma_de[14]_amean	valence	arousal
1.848157	0.000444	1.709386	0.000424	5.5	4.6
1.238022	0.000212	1.116202	0.000100	2.7	3.7
2.053561	0.001272	1.702499	0.001127	5.0	4.3
1.802331	-0.000547	1.680737	-0.001378	3.5	2.8
1.909060	0.001739	1.753582	-0.001443	5.1	6.1
...
1.671897	-0.000347	1.513882	-0.000573	4.8	6.5
1.501074	-0.002797	1.444453	-0.005020	3.9	2.5
2.564815	-0.000453	2.606689	-0.003860	4.0	6.1
2.023299	0.002849	1.714624	0.000608	5.5	5.1
1.759458	-0.003256	1.899548	-0.003618	6.2	6.8



Definición de variables

```
1 explicative = matrix.drop(columns=['valence', 'file', 'Unnamed: 0', 'arousal'])
2 objective_valence = matrix.valence
3 objective_arousal = matrix.arousal
4
```

	Valencia	Activación
Variables entrenamiento	X_train_val, y_train_val	X_train_ar, y_train_ar
Variables test	X_test_val, y_test_val	X_test_ar, y_test_ar
Datos de entrenamiento	90%	90%
Datos de test	10%	10%

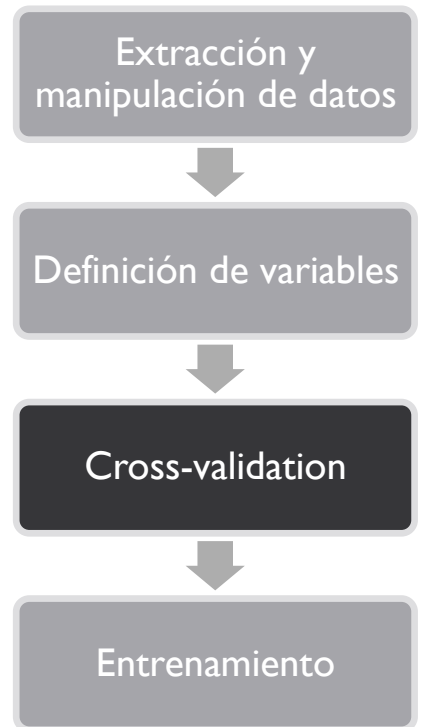


GridSearchCV

```

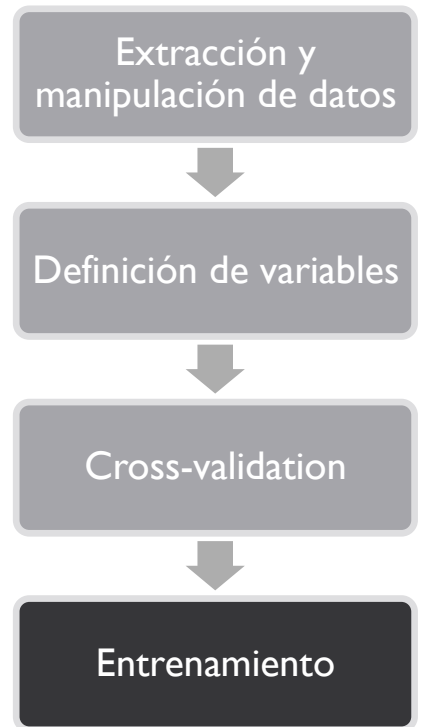
15 tree_regr_model_valence = GridSearchCV(
16     estimator = DecisionTreeRegressor(),
17     param_grid = grid_hp_tree_reg,
18     cv = 5,
19     n_jobs = -1,
20     verbose = 10,
21 )

```



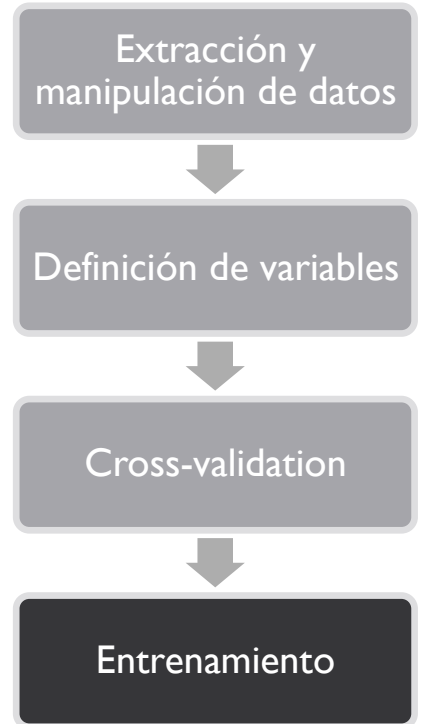
Hiperparámetros para árbol de decisión

Hiperparámetro	Valencia	Activación
criterion	'squared_error'	'squared_error'
splitter	'best'	'best'
max_depth	3	3
min_samples_split	4	2
min_samples_leaf	1	1
max_features	None	None
random_state	None	None
max_leaf_nodes	None	None



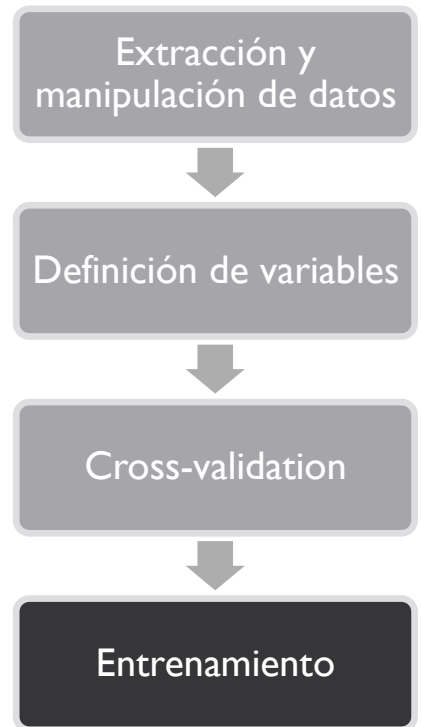
Hiperparámetros para bosque aleatorio

Hiperparámetro	Valencia	Activación
n_estimators	200	200
criterion	'squared_error'	'squared_error'
max_depth	None	None
min_samples_split	4	2
max_features	None	None
random_state	None	None
bootstrap	True	True



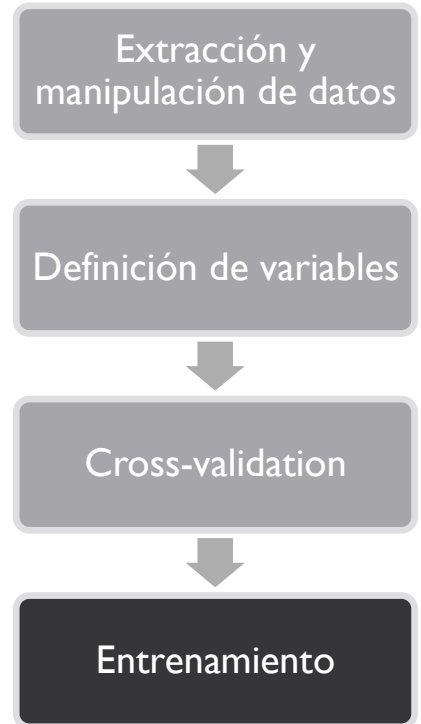
Hiperparámetros para regresión lineal

Hiperparámetro	Valencia	Activación
copy_X	True	True
fit_intercept	False	False
n_jobs	None	None
positive	True	True



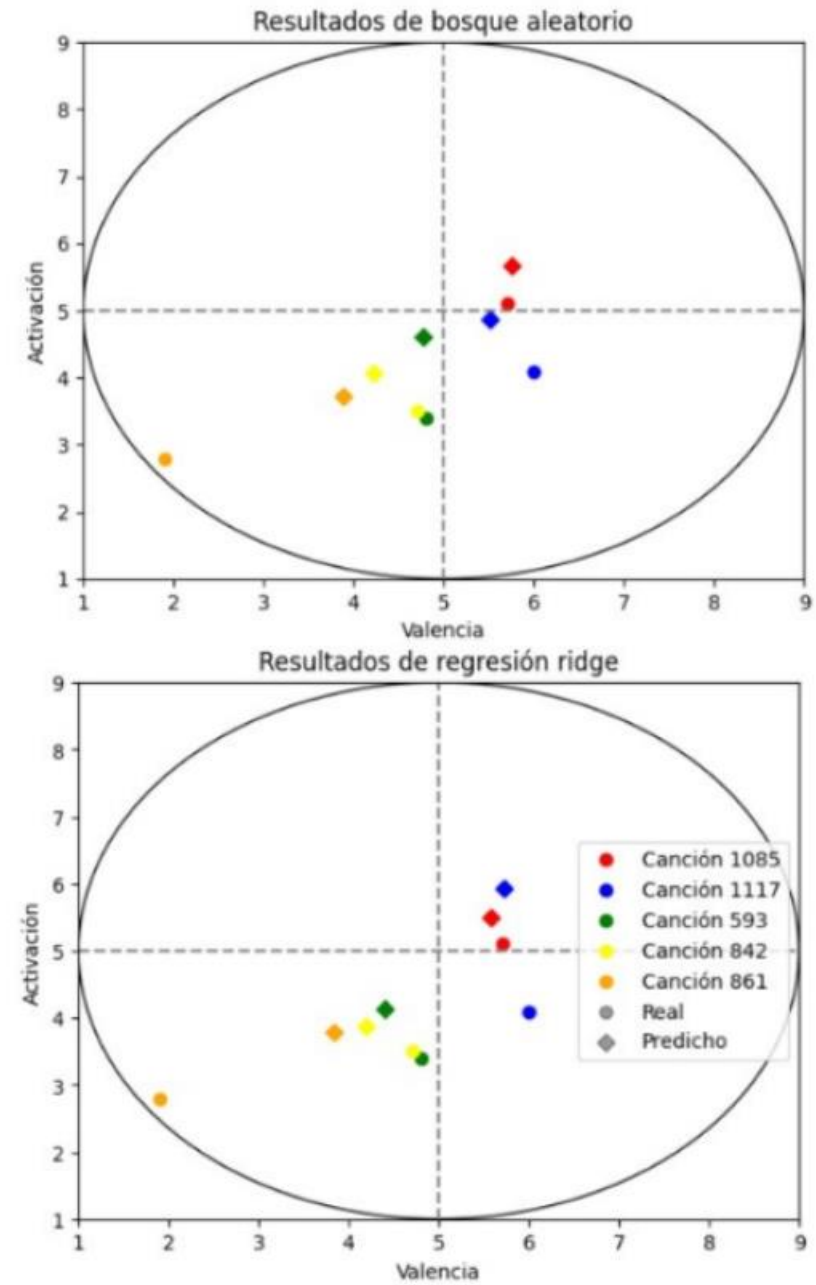
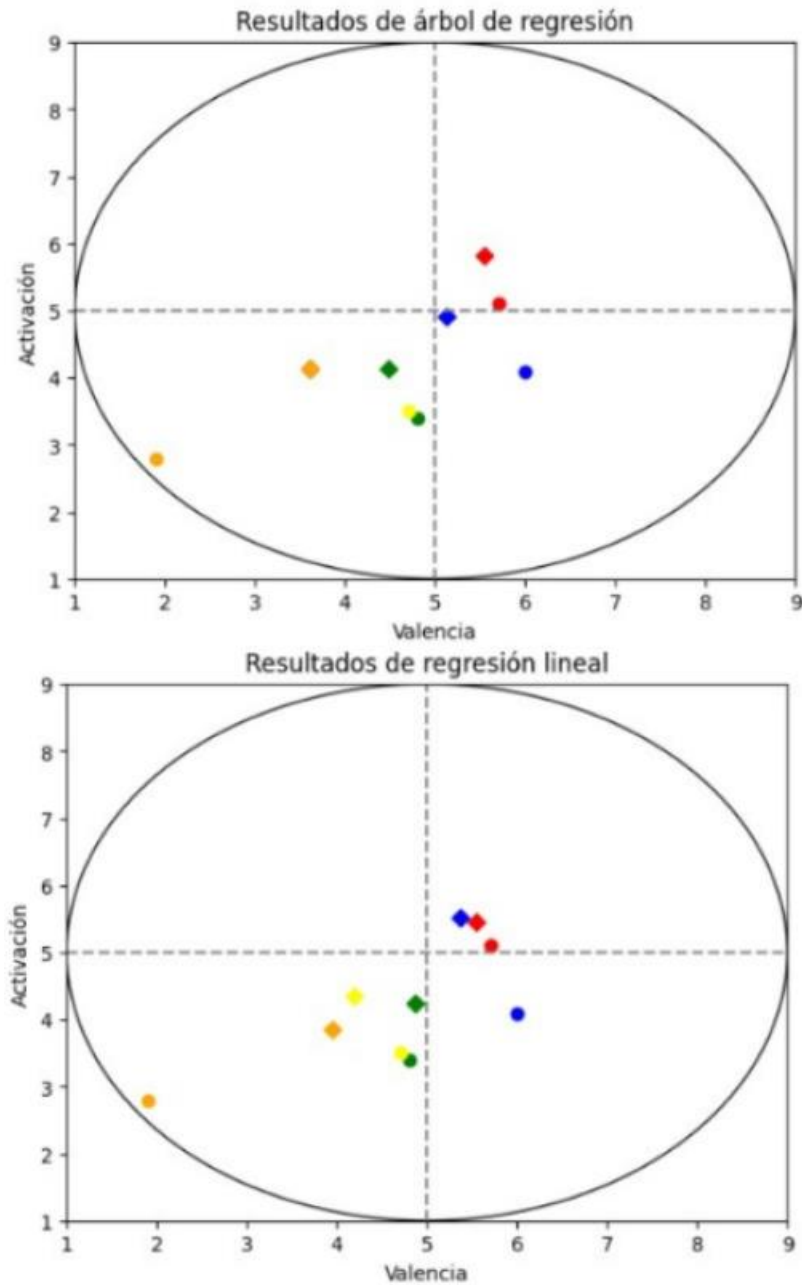
Hiperparámetros para regresión ridge

Hiperparámetro	Valencia	Activación
alpha	5.0	5.0
max_iter	None	None
positive	False	False
random_state	None	None
solver	'svd'	'svd'



Índice de contenidos

1. Motivación y contexto
2. Objetivos
3. MER
4. Machine Learning
5. Conjunto de datos
6. Desarrollo
7. Resultados
8. Conclusiones



Resultados

	Importancia
audspec_lengthLlnorm_sma_de_stddev	0.590566
pcm_fftMag_mfcc_sma[1]_amean	0.283756
jitterLocal_sma_de_stddev	0.063279
audspec_lengthLlnorm_sma_amean	0.062399
F0final_sma_stddev	0.000000
audSpec_Rfilt_sma_de[13]_amean	0.000000
audSpec_Rfilt_sma_de[13]_stddev	0.000000

Árbol de decisión

	Importancia
shimmerLocal_sma_de_amean	27340.029645
jitterDDP_sma_de_amean	24311.199612
jitterDDP_sma_de_stddev	22115.540627
shimmerLocal_sma_de_stddev	7369.382856
voicingFinalUnclipped_sma_de_stddev	4621.678970
jitterLocal_sma_de_amean	1136.452039
pcm_fftMag_spectralSlope_sma_de_amean	598.417183

Regresión lineal

	Importancia
audspec_lengthLlnorm_sma_de_stddev	0.192966
pcm_fftMag_mfcc_sma[1]_amean	0.182091
pcm_fftMag_spectralFlux_sma_de_stddev	0.031868
audspec_lengthLlnorm_sma_amean	0.030294
pcm_fftMag_spectralFlux_sma_amean	0.019323
shimmerLocal_sma_de_stddev	0.016924
jitterLocal_sma_de_stddev	0.010951

Bosque aleatorio

	Importancia
pcm_fftMag_fband250-650_sma_de_stddev	0.640207
audspec_lengthLlnorm_sma_amean	0.626606
audspec_lengthLlnorm_sma_stddev	0.583462
pcm_fftMag_mfcc_sma_de[4]_stddev	0.575928
pcm_fftMag_spectralFlux_sma_stddev	0.550687
audspec_lengthLlnorm_sma_de_stddev	0.493948
pcm_fftMag_spectralFlux_sma_de_stddev	0.313460

Regresión ridge

Resultados

	Importancia
audspec_lengthLlnorm_sma_amean	0.577854
pcm_fftMag_spectralEntropy_sma_de_stddev	0.147803
audspec_lengthLlnorm_sma_de_stddev	0.147049
logHNR_sma_amean	0.037789
pcm_fftMag_mfcc_sma[1]_amean	0.030927
pcm_fftMag_spectralVariance_sma_amean	0.030360
shimmerLocal_sma_de_stddev	0.028218

Árbol de decisión

	Importancia
shimmerLocal_sma_de_amean	6200.551602
jitterLocal_sma_de_amean	1012.118046
pcm_fftMag_spectralSlope_sma_de_amean	581.074969
pcm_fftMag_fband1000-4000_sma_de_amean	141.952809
pcm_fftMag_fband250-650_sma_de_amean	134.747001
pcm_fftMag_spectralEntropy_sma_de_amean	65.647901
pcm_fftMag_spectralHarmonicity_sma_de_amean	37.686683

Regresión lineal

	Importancia
audspec_lengthLlnorm_sma_amean	0.242928
pcm_fftMag_spectralEntropy_sma_de_stddev	0.106660
audspec_lengthLlnorm_sma_de_stddev	0.055677
pcm_fftMag_mfcc_sma[1]_amean	0.047834
pcm_fftMag_spectralVariance_sma_amean	0.023260
pcm_fftMag_psySharpness_sma_de_stddev	0.014910
shimmerLocal_sma_de_stddev	0.014439

Bosque aleatorio

	Importancia
audspec_lengthLlnorm_sma_amean	0.643724
audspec_lengthLlnorm_sma_de_stddev	0.533155
audspec_lengthLlnorm_sma_stddev	0.505169
pcm_fftMag_mfcc_sma_de[4]_stddev	0.480647
pcm_fftMag_mfcc_sma_de[2]_stddev	0.472288
pcm_fftMag_mfcc_sma_de[1]_stddev	0.425346
pcm_fftMag_spectralFlux_sma_stddev	0.422188

Regresión ridge

Resultados

Métricas

- ☐ Mean absolute error (MAE)
- ☐ Mean squared error (MSE)
- ☐ Coeficiente de determinación (R2 Score)
- ☐ Mean absolute percentage error (MAPE)

Resultados

Mean absolute error (MAE)

Valencia

Algoritmo	MAE
Árbol de decisión	0.811
Bosque aleatorio	0.680
Regresión lineal	0.649
Regresión ridge	0.696



Regresión lineal

Activación

Algoritmo	MAE
Árbol de decisión	0.842
Bosque aleatorio	0.718
Regresión lineal	0.795
Regresión ridge	0.852



Bosque aleatorio

Resultados

Mean squared error (MSE)

Valencia

Algoritmo	MSE
Árbol de decisión	0.904
Bosque aleatorio	0.688
Regresión lineal	0.646
Regresión ridge	0.708



Regresión lineal

Activación

Algoritmo	MSE
Árbol de decisión	1.035
Bosque aleatorio	0.806
Regresión lineal	0.946
Regresión ridge	1.094



Bosque aleatorio

Resultados

Coeficiente de determinación (R²)

Valencia

Algoritmo	R ²
Árbol de decisión	0.303
Bosque aleatorio	0.469
Regresión lineal	0.502
Regresión ridge	0.454



Regresión lineal

Activación

Algoritmo	R ²
Árbol de decisión	0.363
Bosque aleatorio	0.504
Regresión lineal	0.417
Regresión ridge	0.326



Bosque aleatorio

Resultados

Mean absolute percentage error (MAPE)

Valencia

Algoritmo	MAPE
Árbol de decisión	0.170
Bosque aleatorio	0.143
Regresión lineal	0.137
Regresión ridge	0.149



Regresión lineal

Activación

Algoritmo	MAPE
Árbol de decisión	0.181
Bosque aleatorio	0.160
Regresión lineal	0.174
Regresión ridge	0.186



Bosque aleatorio

Resultados

Comparativa de métricas

	Valencia	Activación
MAE	Regresión lineal	Bosque aleatorio
MSE	Regresión lineal	Bosque aleatorio
R2	Regresión lineal	Bosque aleatorio
MAPE	Regresión lineal	Bosque aleatorio

Índice de contenidos

1. Motivación y contexto
2. Objetivos
3. MER
4. Machine Learning
5. Conjunto de datos
6. Desarrollo
7. Resultados
8. Conclusiones

Conclusiones

Objetivos

Sistema de reconocimiento de emociones en pistas de audio



MER



Machine Learning



Python



Conclusiones

Competencias y conocimientos

- Aprendidas durante el estudio del Grado:
 - Python (Anaconda y Jupyter)
 - Técnicas Machine Learning
 - Características acústicas de las señales de audio
- Adquiridas durante el desarrollo del trabajo:
 - Latex (TeXstudio)
 - MER
 - Algoritmos Machine Learning

Conclusiones

Futuras líneas de trabajo

- Implementar otros algoritmos de Machine Learning para mejorar las puntuaciones de precisión del sistema
- Trabajar con un conjunto de datos con categorización discreta de emociones
- Análisis de la letra de las canciones
- Aplicación con interfaz de usuario amigable



Muchas gracias



Universidad
Rey Juan Carlos