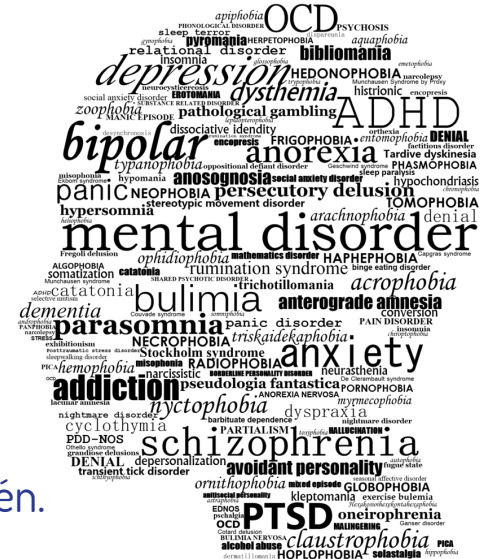


Clasificación de estados mentales con datos de electroencefalografía (EEG)

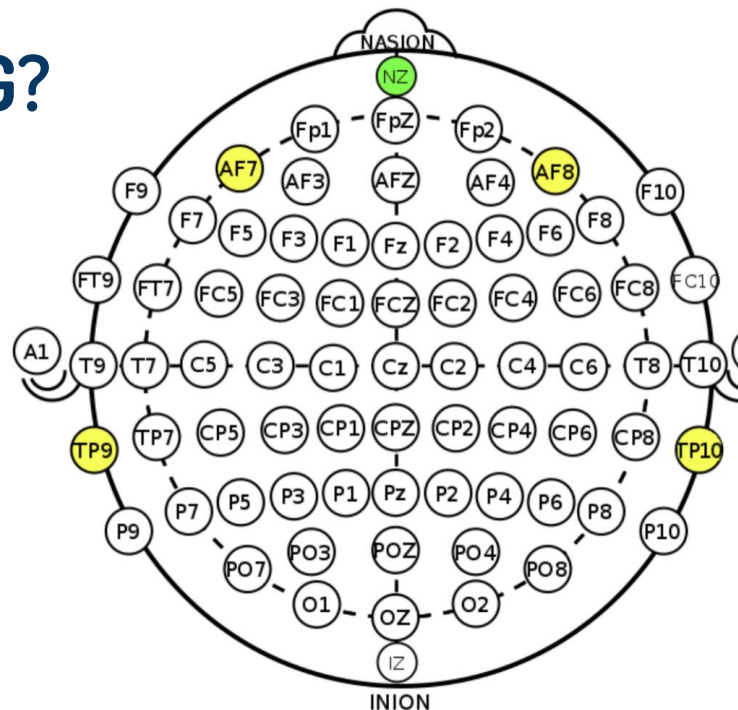
Laboratorio de Datos 1C - 2021



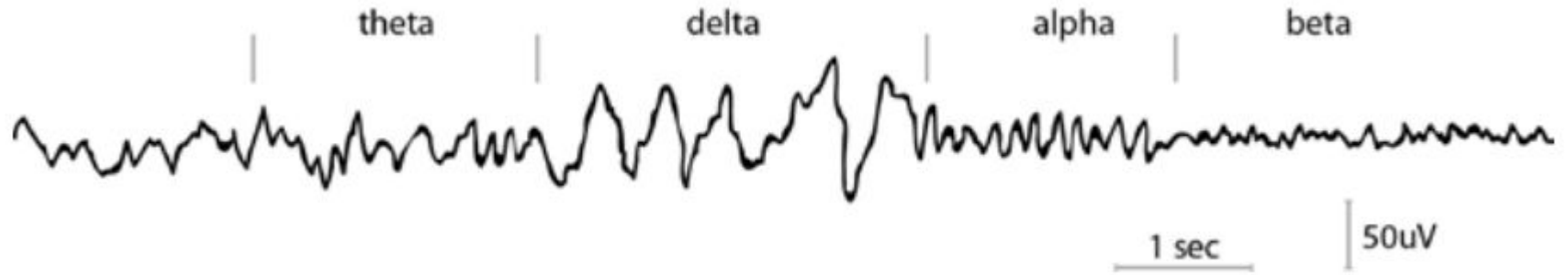
Integrantes: Devia, Amalia. Galvan, Hugo; Palmucci, Matias; Ticona, María Belén.

¿Qué es y para qué sirve un EEG?

- Registro actividad bioeléctrica cerebral
- Sirve para obtener información objetiva sobre el estado de integridad funcional de la corteza cerebral
- Aplicaciones en el diagnóstico, monitoreo y evaluación de numerosas patologías neurológicas, psiquiátricas y neuroquirúrgicas
- Permite evaluaciones del **estado cognitivo de sujetos experimentales** o pacientes, estudios de trastornos del sueño, desórdenes de la conciencia y muerte cerebral.



Señales de EEG



4 - 7 Hz.

Up to 4 Hz.

7 - 14 Hz.

15 - 30 Hz.

Objetivos

- Generar modelos simples de clasificación de estados mentales a partir de señales EEG
- Tratar de obtener resultados similares al trabajo publicado por Birds et al (2018)
 - Selección de features relevantes para la clasificación y comparar resultados
 - Creación de distintos modelos y comparar el accuracy
- Aplicar técnicas y herramientas vistas en la materia para poder caracterizar las señales de cada uno de los estados
- Analizar desempeño de modelos tomando solamente datos de individuo vs distintos individuos



Dataset

- 4 individuos
- Estados mentales:
 - Relajado
 - Neutral
 - Concentrado
- 4 electrodos secos
- 5 frecuencias
- ~900 features por cada ventana de 1 segundo.



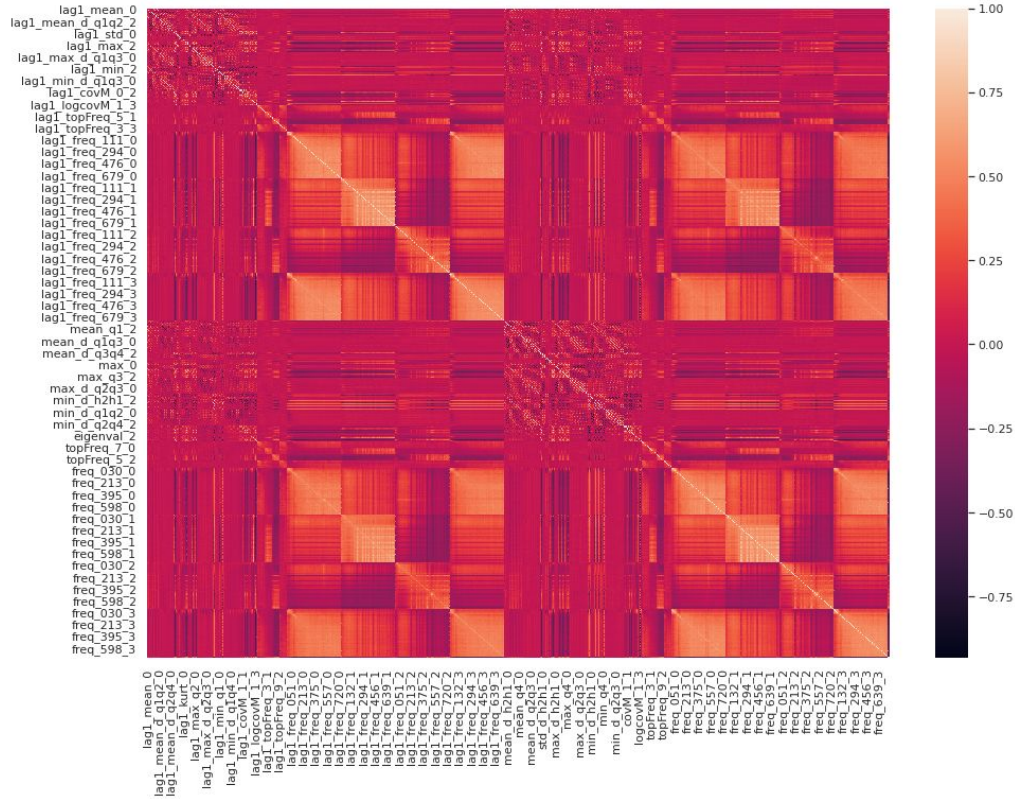
Muse Headband - Dispositivo con sensores de señales cerebrales

1. Feature Selection

- Desafío: lidiar con gran cantidad de features
 - Método del paper original seleccionado: corte según correlación de Pearson
Es el método utilizado en el paper debido a su simplicidad y a su bajo costo computacional.
 - Método alternativo usado: por Regresión Logística por cada feature

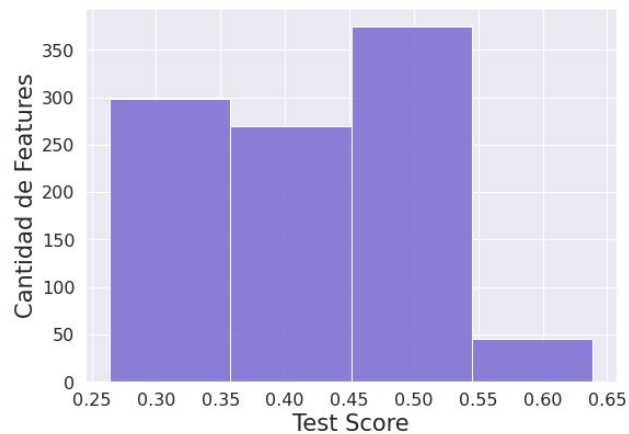


Correlación de Pearson entre variables

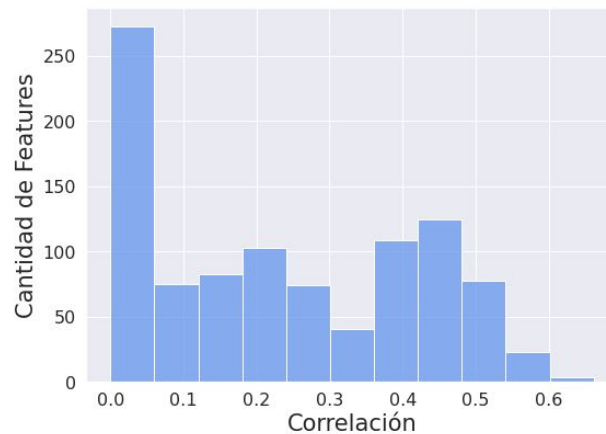


1. Feature Selection

Distribución Performance individual en LR



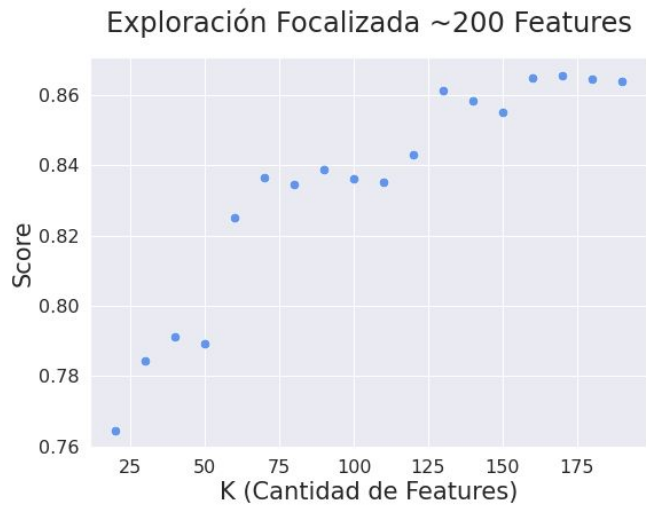
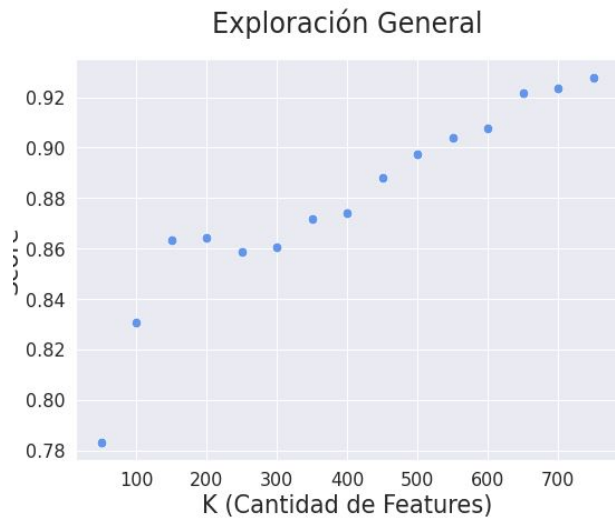
Distribución Feat. Corr. Pearson con *target*



1. Feature Selection

K-mejores según score individual de cada feature

LR: Performance Según K-Mejores Features

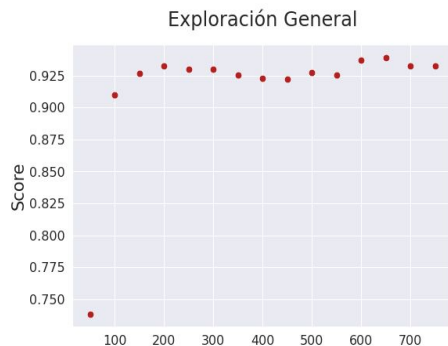


1. Feature Selection

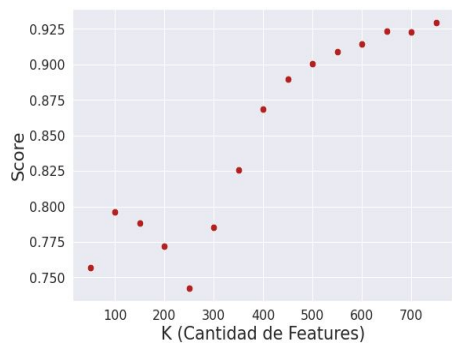
K-mejores según corr. Pearson con el *target*

LR: Performance Según K-Mejores Features

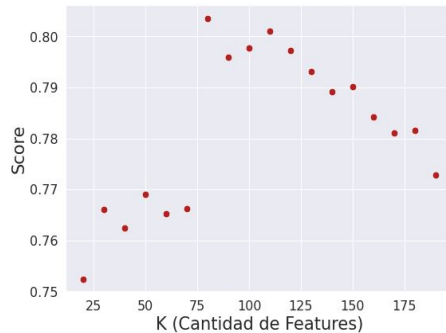
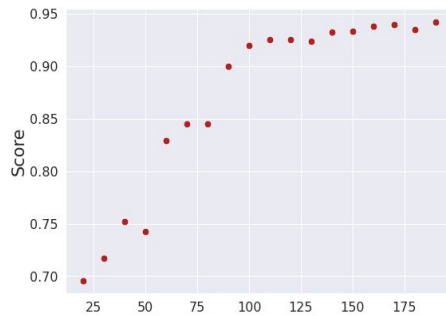
No Abs



Valor Abs

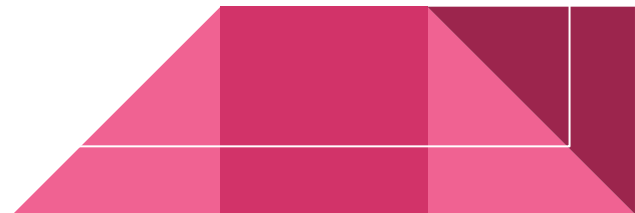


Exploración Focalizado -200 Features



2. Elección de los Modelos Predictivos

- Se decidió aplicar una regresión logística y un modelo de máquinas de vectores de soporte para predecir la variable estado. La variable estado es una variable binaria que indica si después de la ventana de tiempo seleccionada se produjo el cambio en la señal evaluada. Las variables independientes son las frecuencias alfa (α), beta (β), delta (δ), theta (θ) y gamma (γ) En el desarrollo de los modelos se aplicó una ventana de tiempo fija de 1 segundo.
- Comprender cantidad de componentes



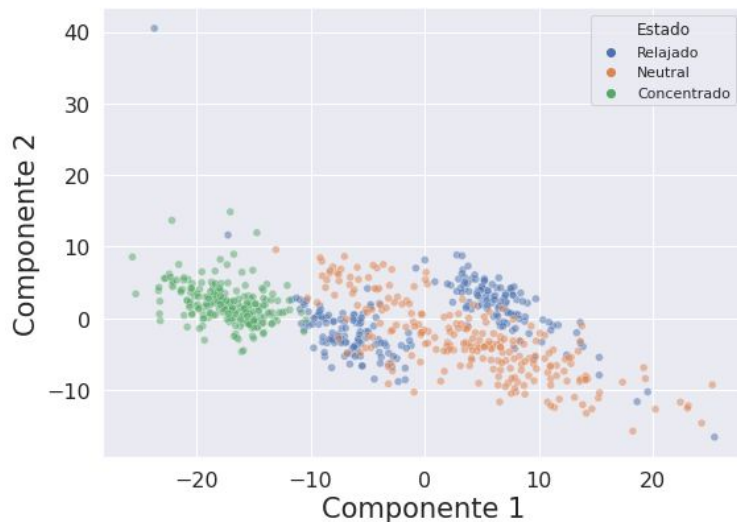
Objetivos

- Categorizar 3 estados mentales posibles: relajación, neutralidad y concentración.
- Crear un conjunto de datos con 4 sujetos y sesiones de un minuto para cada clase de estado mental con el fin de entrenar y probar diferentes métodos.
- Probar una combinación de diferentes algoritmos de selección de características y modelos de clasificador para comparar su rendimiento en términos de precisión de reconocimiento y número de características necesarias.

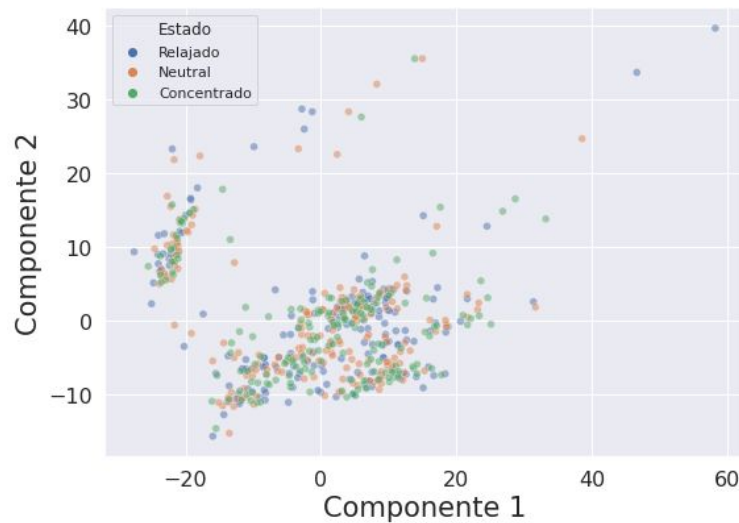


Visualización de Estados Mentales por PCA

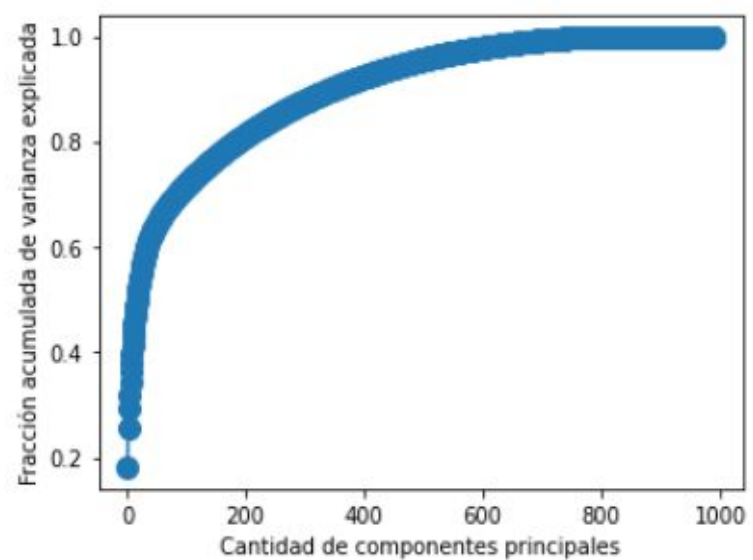
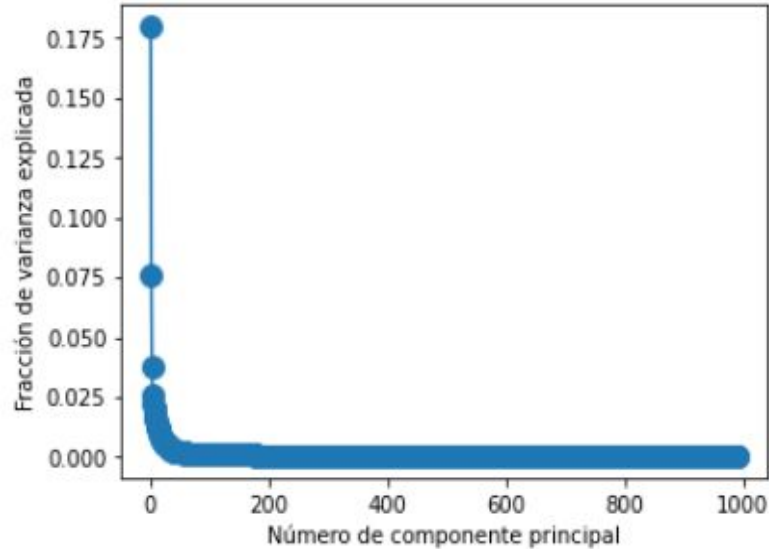
Sujeto con estados diferenciables



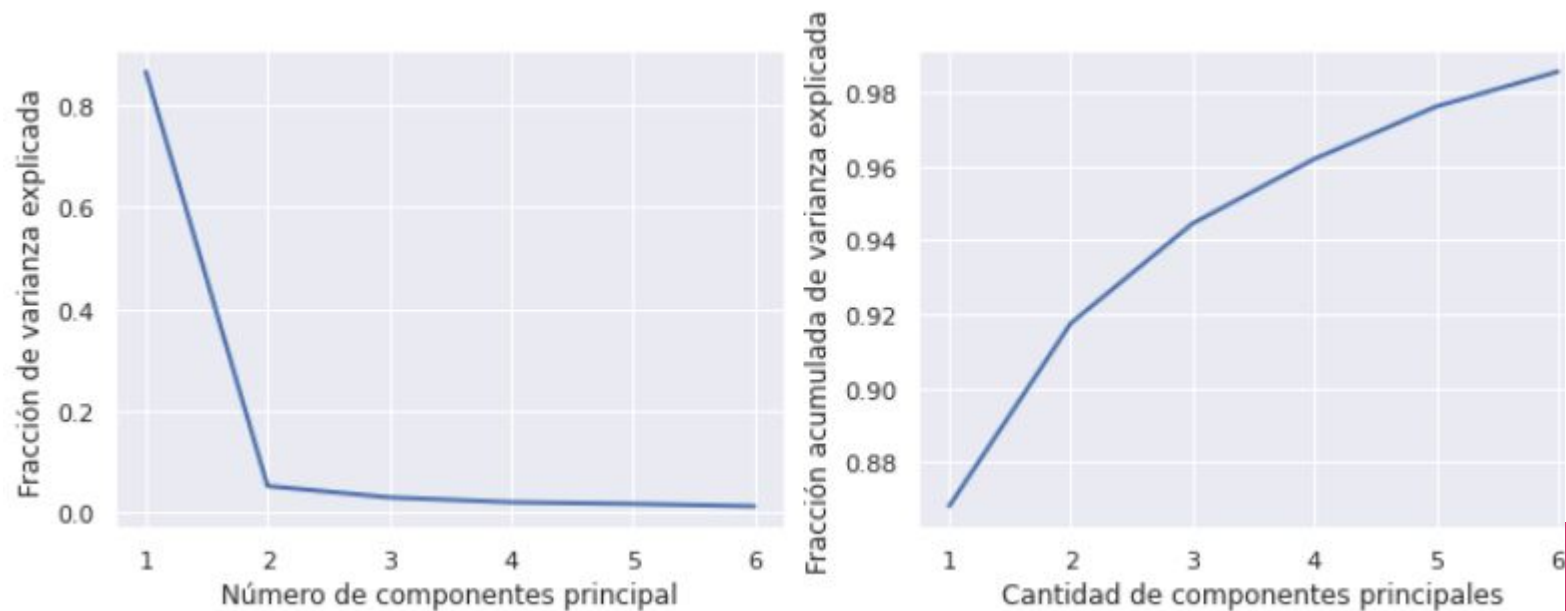
Sujeto con estados no tan diferenciables



Modelo PCA sin Correlación de Pearson

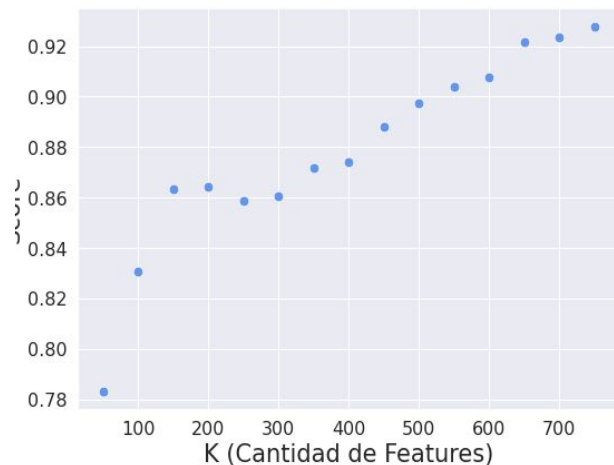


Modelo PCA con Correlación de Pearson

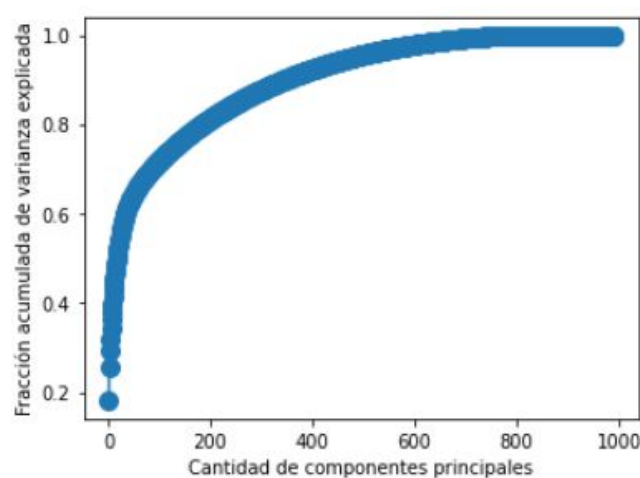


Dimensión features necesarias:

Comparación entre PCA y Regresión Logística

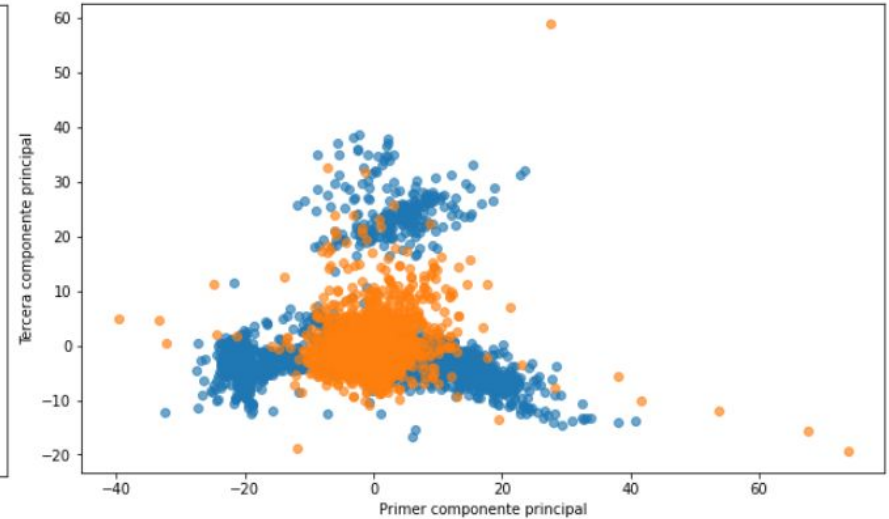
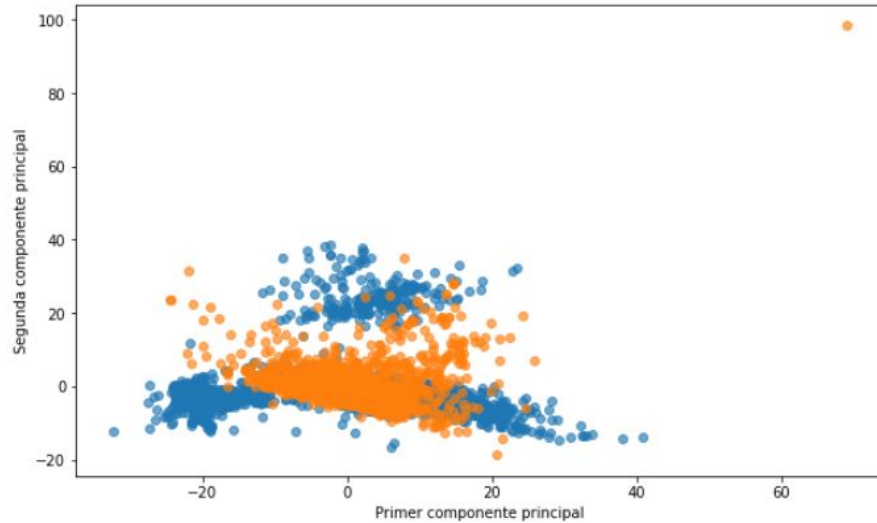


Regresión Logística

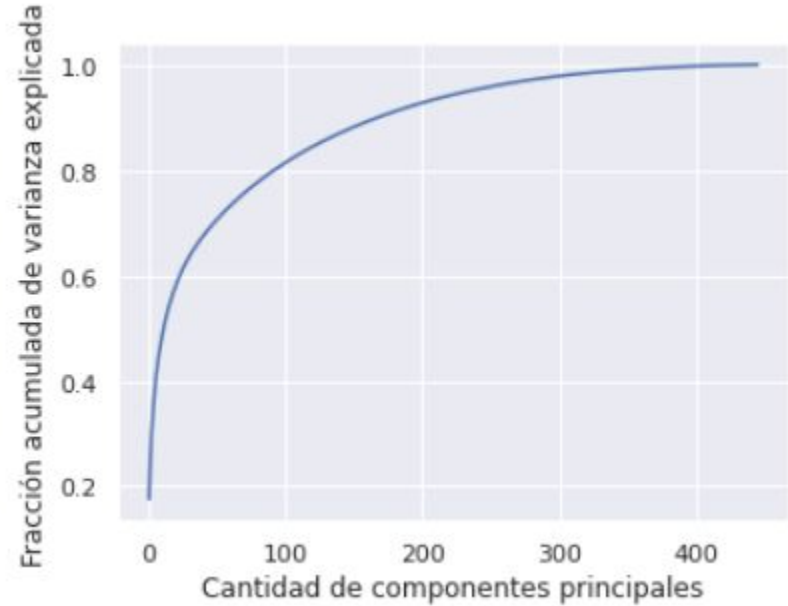
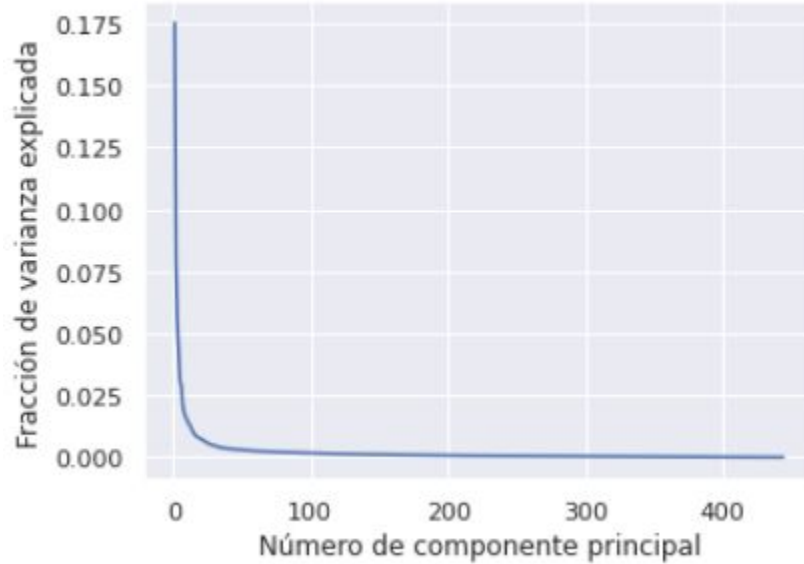


PCA

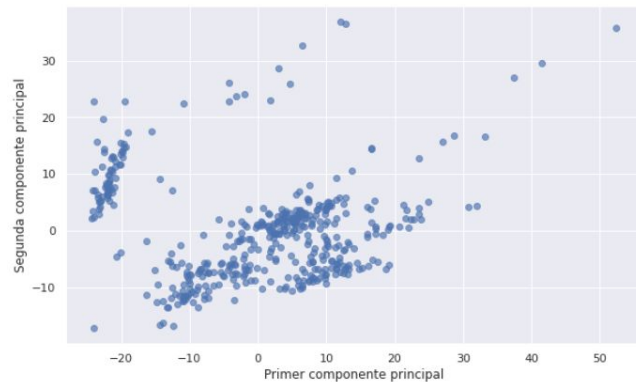
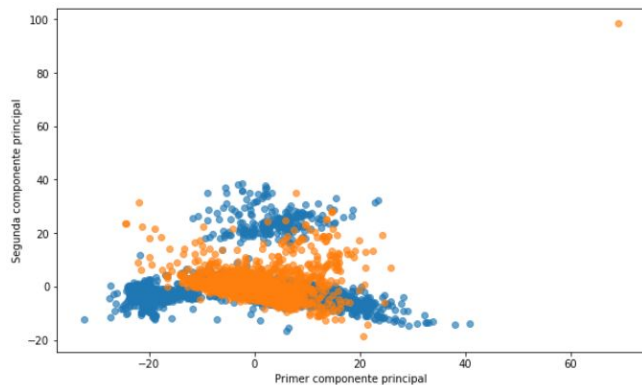
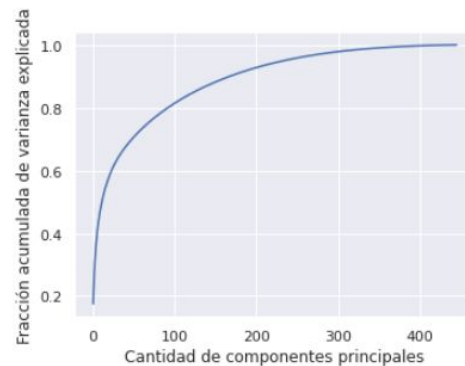
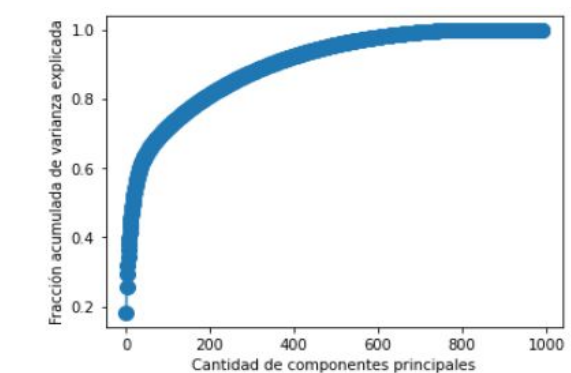
Modelo PCA con las 3 componentes principales



Modelo PCA para individuo único



Modelo PCA comparación



Modelo SCV para individuo único

Modelo SCV para dataset completo

con 4 individuos:

```
score SVC Lineal: 0.6826196473551638
```

```
score SVC Gama: 0.3501259445843829
```

Modelo SCV para dataset con

1 individuo:

```
score SVC Lineal: 1.0
```

```
score SVC Lineal: 1.0
[[45  0  0]
 [ 1 47  2]
 [ 0  0 16]]
[[1.  0.  0. ]
 [0.02 0.94 0.04]
 [0.  0.  1. ]]
Matriz de confusion del modelo es:
[[45  0  0]
 [ 1 47  2]
 [ 0  0 16]]
Sensibilidad del modelo es de: 1.0
Especificidad del modelo es de: 0.9783
BA del modelo es de: 0.9891
```

Random Forest

std_2 <= 60.085
gini = 0.667
samples = 1586
value = [539, 523, 524]

lag1_std_2 <= 7.744
gini = 0.595
samples = 1155
value = [537, 484, 134]

gini = 0.173
samples = 431
value = [2, 39, 390]

gini = 0.189
samples = 224
value = [201, 8, 15]

freq_111_3 <= 0.058
gini = 0.592
samples = 931
value = [336, 476, 119]

lag1_freq_649_2 <= 0.015
gini = 0.576
samples = 688
value = [178, 395, 115]

gini = 0.466
samples = 243
value = [158, 81, 4]

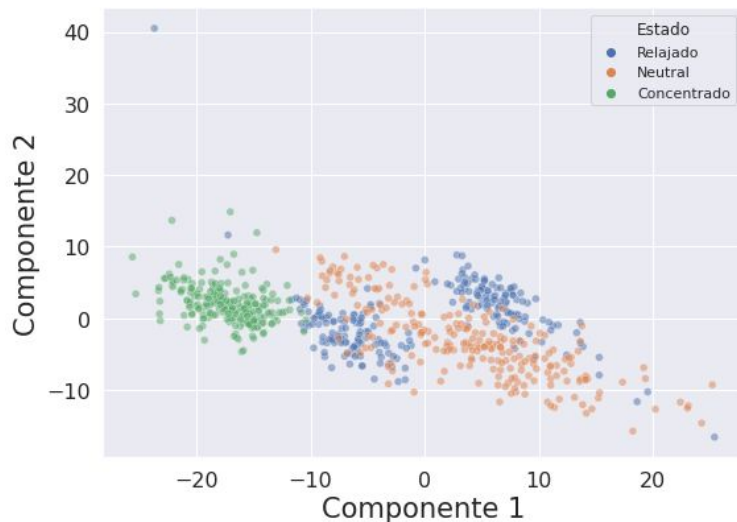
gini = 0.651
samples = 399
value = [114, 174, 111]

gini = 0.366
samples = 289
value = [64, 221, 4]

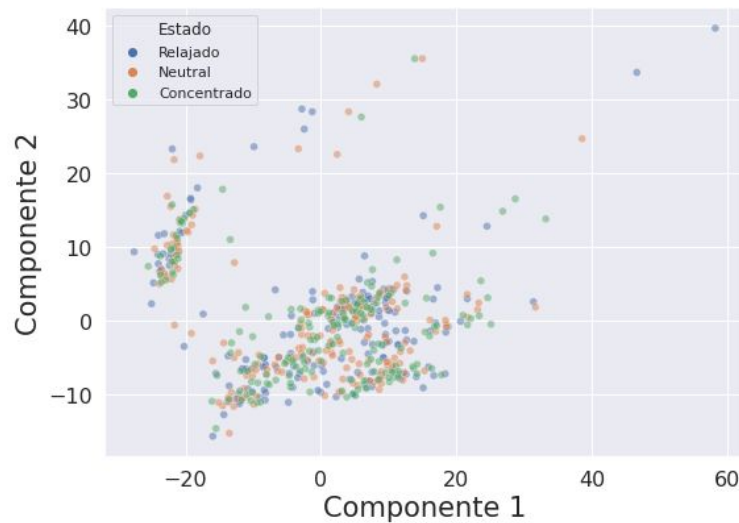


Visualización de Estados Mentales por PCA

Sujeto con estados diferenciables



Sujeto con estados no tan diferenciables



Conclusiones

- Se comprendieron las dimensiones de la cantidad de features necesarias para realizar distintos modelos
- Al trabajar con muchas features correlacionadas se requiere un tratamiento adecuado para no cometer overfitting
- El método de PCA y analizar el aporte de cada feature en LR pueden ser útiles para el procesamiento y análisis de los features sobre datos de estados mentales obtenidos con EEG.
- RF resultó útil para reducir las dimensiones de las variables pero hay pérdida de interpretación.

