

Proyecto Final - Neurociencia Computacional

Inferencia de redes en EEG de epilépticos



**Universidad
Nacional
de San Martín**

Alumno: Ing Valentin Barco

1. Resumen

Este trabajo se realizó en calidad de proyecto final de la materia de neurociencia computacional brindada por la Universidad Nacional de San Martín. Se procesaron diferentes muestras electroencefalográficas de pacientes neurotípicos y epilépticos con Python, brindadas por un Centro de Neurociencia de la Ciudad de Bahía Blanca con el objetivo de encontrar características que permitan diferenciar y encontrar patrones en los pacientes que padecían epilepsia.

2. Marco Teórico

2.1 Epilepsia

La epilepsia es una enfermedad crónica en la que las neuronas se descargan repentinamente y lleva al cerebro a tener una disfunción.

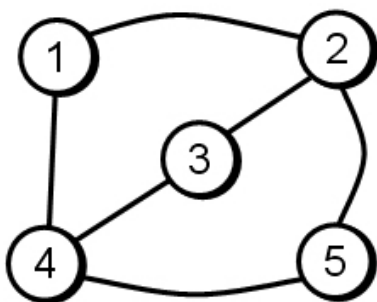
Esta puede provocar deterioro cognitivo, convulsiones, lesiones o incluso la muerte. De acuerdo con el grado de afectación de las áreas cerebrales en el ataque epiléptico, la epilepsia se puede dividir en dos clases, convulsiones generalizadas y convulsiones parciales. Las convulsiones generalizadas afectan a los hemisferios cerebrales bilaterales, mientras que las convulsiones parciales se originan en neuronas parciales de un hemisferio. Se puede detectar mediante:

- Un examen físico general y exploración neurológica
- Tomografía computarizada craneal
- Resonancia magnética craneal
- Electroencefalograma.

Como mencionamos anteriormente en este trabajo se utilizaran Electroencefalografías.

2.2 Grafo

Un grafo es una composición de un conjunto de objetos conocidos como nodos que se relacionan con otros nodos a través de un conjunto de conexiones conocidas como aristas. Los grafos permiten estudiar las relaciones que existen entre unidades que interactúan con otras.



M	1	2	3	4	5
1	0	1	0	1	0
2	1	0	1	0	1
3	0	1	0	1	0
4	1	0	1	0	1
5	0	1	0	1	0

Podemos representar diversas situaciones o elementos con grafos. Estos son extraordinariamente útiles en situaciones complejas, es por esto, que es común conseguir la implementación de análisis de grafos en estudios de ciencias exactas, ciencias sociales y en aplicaciones informáticas.

Existen formas de caracterizar los grafos con diferentes métricas:

- La distribución de grado en una red se representa habitualmente como $P(k)$ y es definida como la fracción de nodos en la red con un cierto grado k .
- Coeficiente clustering en un grafo se refiere a la medida de agrupamiento o cohesión de los nodos en una red. El coeficiente de clustering es una métrica utilizada para calcular esta distribución. Indica la probabilidad de que los vecinos de un nodo estén conectados entre sí.
- La longitud media del camino es la longitud promedio de todos los caminos más cortos entre todos los pares de nodos en el grafo. Proporciona una medida de la eficiencia de la comunicación en la red.
- El grado medio de un grafo es una métrica que proporciona la cantidad promedio de conexiones de los nodos en la red. Se calcula sumando los grados de todos los nodos en el grafo y dividiendo esta suma entre el número total de nodos en el grafo.

2.3 Coeficiente de correlación

El coeficiente de correlación es una medida estadística que indica el grado de relación lineal entre dos variables. Mide la fuerza de la relación entre las variables, permitiendo determinar si existe una asociación entre ellas.

$$Coef_{Corr} = \frac{\sum[(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum(x_i - \bar{x})^2 * \sum(y_i - \bar{y})^2}}$$

2.4 Matriz de correlación

Una matriz de correlación no es más que una tabla con los coeficientes de correlación (Coef-Corr) de distintas variables. La matriz muestra cómo se relacionan entre sí todos los posibles pares de valores de una tabla. Es una poderosa herramienta para resumir un gran conjunto de datos y encontrar y mostrar patrones en ellos.

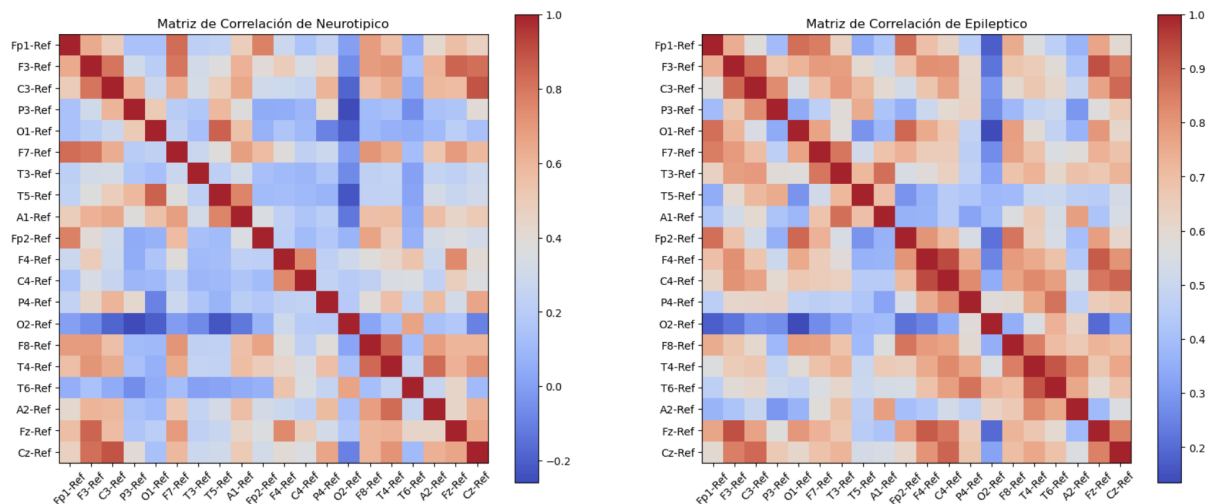
1	0.32	1	0.99	0.17
0.32	1	0.33	0.32	-0.023
1	0.33	1	0.99	0.21
0.99	0.32	0.99	1	0.18
0.17	-0.023	0.21	0.18	1

Luego a partir de estas matrices generadas, es posible inferir diferentes redes utilizando grafos, y a partir de estos, se puede realizar una caracterización utilizando diferentes métricas, como la distribución de grado, el coeficiente de clustering y el grado medio.

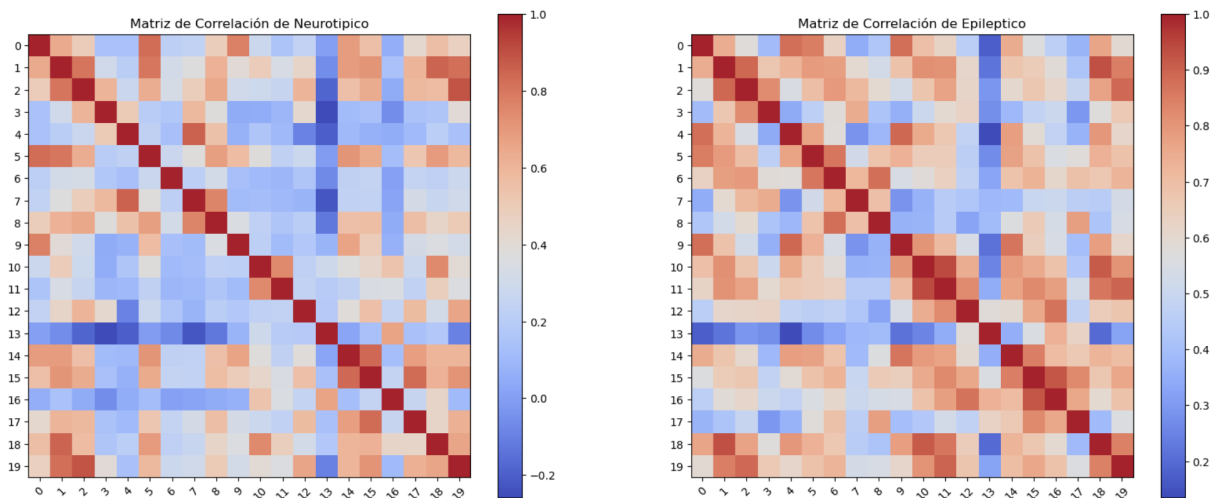
4. Desarrollo

4.1 Matrices de correlación aplicado al EEG

Primero se realizó la matriz de correlación de los 21 canales de todos los pacientes, como se puede observar en la siguiente figura que compara las matrices de dos pacientes, uno neurótico y otro epiléptico.



Para poder hacer el procesamiento sin necesidad de trabajar con etiquetas string, se utilizó una codificación de los canales del 1 al 21.

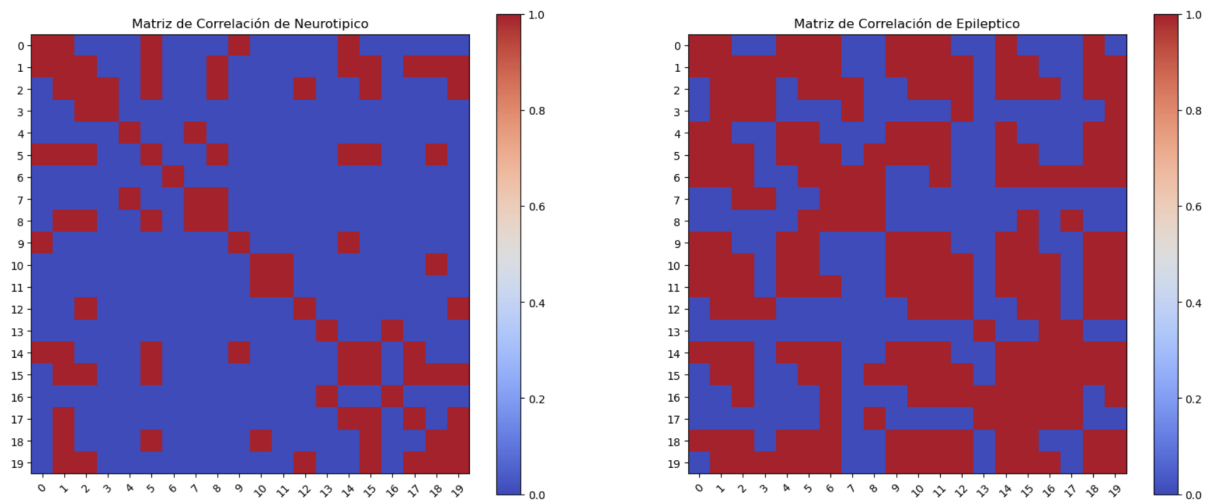


Desde el inicio del análisis, es evidente que la matriz de correlación del paciente con epilepsia muestra una correlación más alta entre sus canales. Este hallazgo sugiere fuertemente que el análisis podría ser acertado. La presencia de una mayor correlación entre los canales indica una mayor coherencia en la actividad eléctrica cerebral y respalda la idea de que la epilepsia implica una sincronización anormal de la actividad neuronal en el cerebro del paciente.

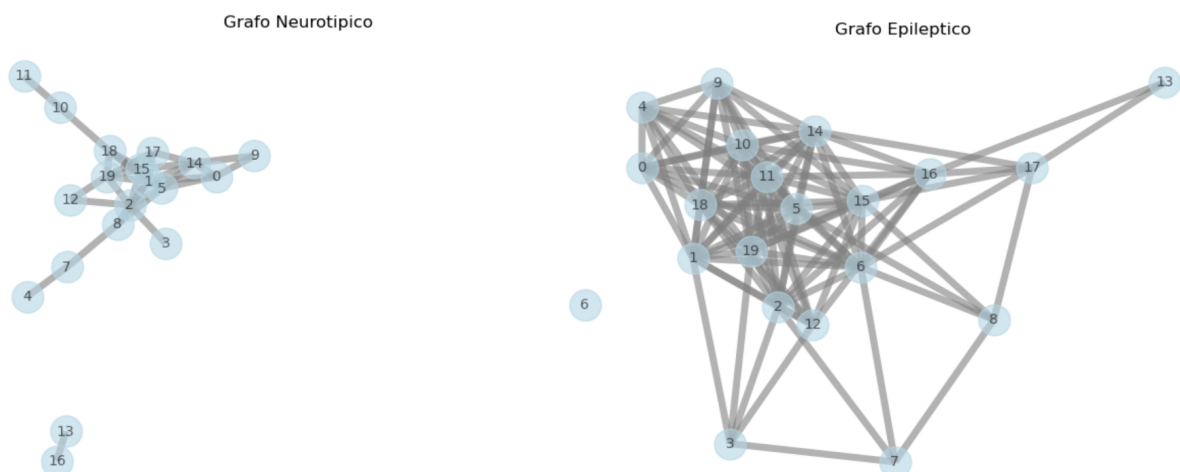
4.2 Binarización por umbral

Luego aplicando un umbral y binarizado en función de este, se puede obtener una matriz de conexiones más simplificada que se puede tomar como matriz de adyacencia para inferir una red.

En el caso del ejemplo mostrado en la sección 4.1 se obtienen las siguientes matrices binarizadas utilizando un umbral de 0.6.



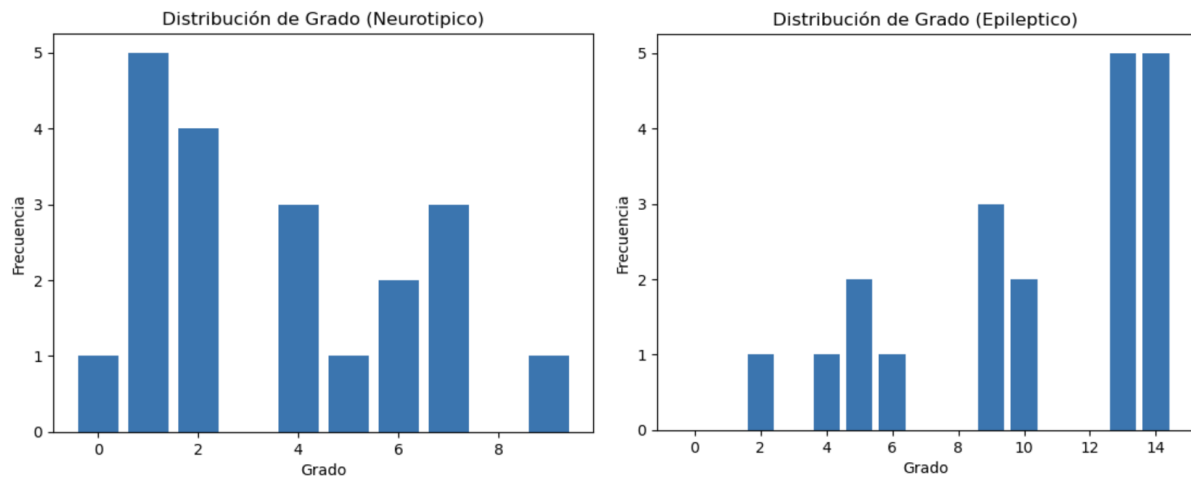
Una vez binarizada se infiere la red tomando a la matriz antes mencionada como matriz de adyacencia y se construye el grafo correspondiente.



4.3 Métricas

A partir de estos grafos es posible calcular algunas de las métricas definidas en el marco teórico.

4.3.1 Distribución de grado



4.3.2 Coeficiente de clustering

Neurotipico: 0.37

Epileptico: 0.744

4.3.3 Longitud promedio de camino más corto

Neurotipico: 2.4044117647058822 (Se eliminan los nodos sin conectar)

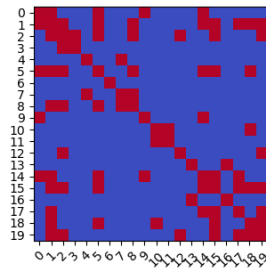
Epileptico: 1.5210526315789474

5. Resultados

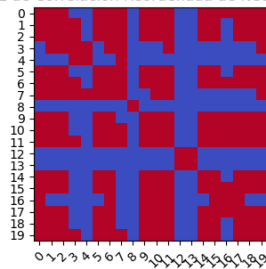
Realizando el análisis mostrado en el desarrollo en los 12 pacientes podemos obtener diferentes parámetros que nos pueden permitir caracterizar los pacientes y ver si se pueden encontrar patrones.

Matrices de correlación para umbral de 0.6

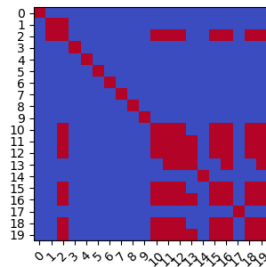
Matriz de Correlación Reordenada de Neurotípico 1



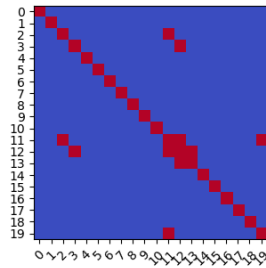
Matriz de Correlación Reordenada de Neurotípico 2



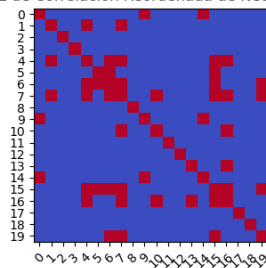
Matriz de Correlación Reordenada de Neurotípico 3



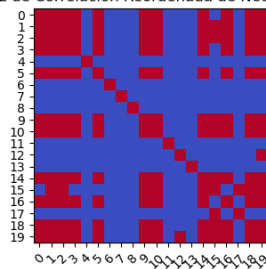
Matriz de Correlación Reordenada de Neurotípico 4



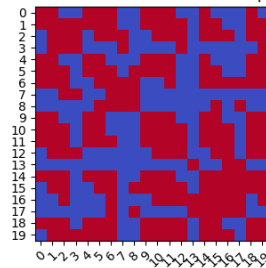
Matriz de Correlación Reordenada de Neurotípico 5



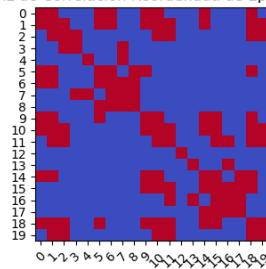
Matriz de Correlación Reordenada de Neurotípico 6



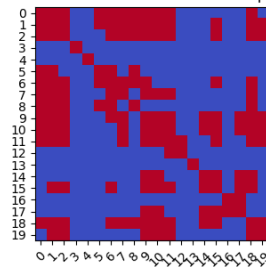
Matriz de Correlación Reordenada de Epileptico 1



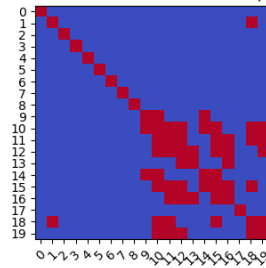
Matriz de Correlación Reordenada de Epileptico 2



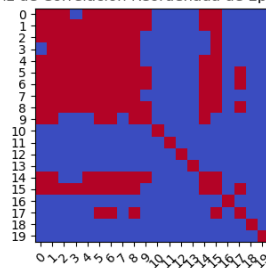
Matriz de Correlación Reordenada de Epileptico 3



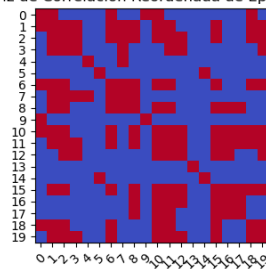
Matriz de Correlación Reordenada de Epileptico 4



Matriz de Correlación Reordenada de Epileptico 5

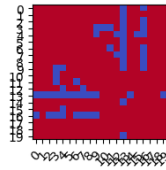


Matriz de Correlación Reordenada de Epileptico 6

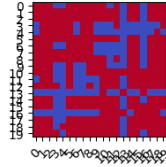


Para analizar cómo se comportan las matrices de correlación de los pacientes 0 ante diferentes umbrales (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)

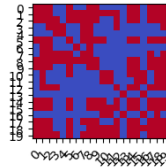
Matriz de Correlación Reordenada de Neurotípico 1



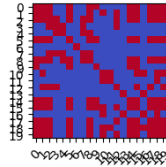
Matriz de Correlación Reordenada de Neurotípico 2



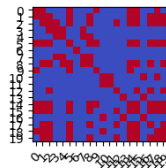
Matriz de Correlación Reordenada de Neurotípico 3



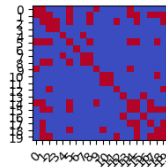
Matriz de Correlación Reordenada de Neurotípico 4



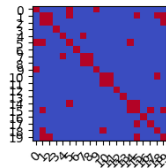
Matriz de Correlación Reordenada de Neurotípico 5



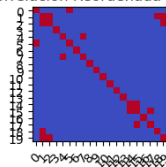
Matriz de Correlación Reordenada de Neurotípico 6



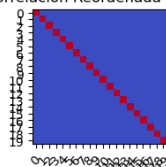
Matriz de Correlación Reordenada de Neurotípico 7



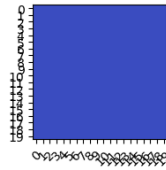
Matriz de Correlación Reordenada de Neurotípico 8



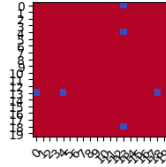
Matriz de Correlación Reordenada de Neurotípico 9



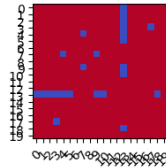
Matriz de Correlación Reordenada de Epileptico 1



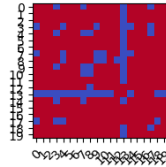
Matriz de Correlación Reordenada de Epileptico 2



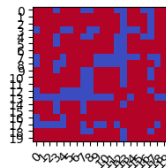
Matriz de Correlación Reordenada de Epileptico 3



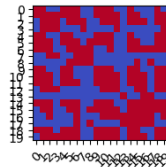
Matriz de Correlación Reordenada de Epileptico 4



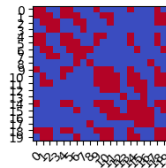
Matriz de Correlación Reordenada de Epileptico 5



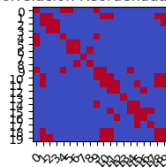
Matriz de Correlación Reordenada de Epileptico 6



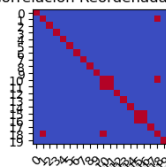
Matriz de Correlación Reordenada de Epileptico 7



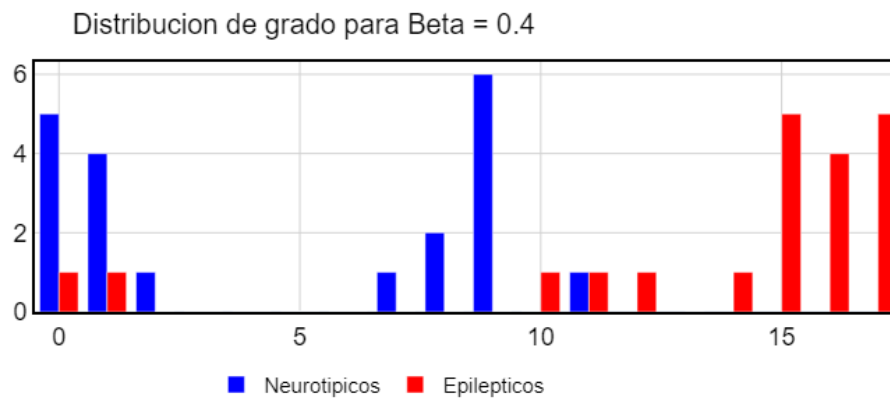
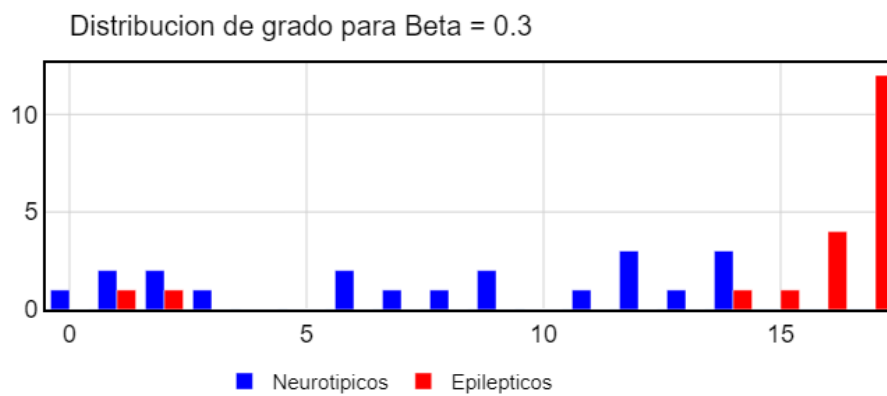
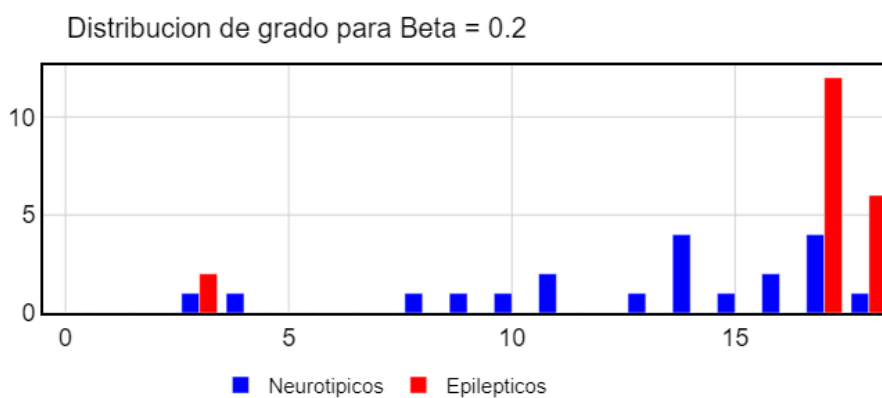
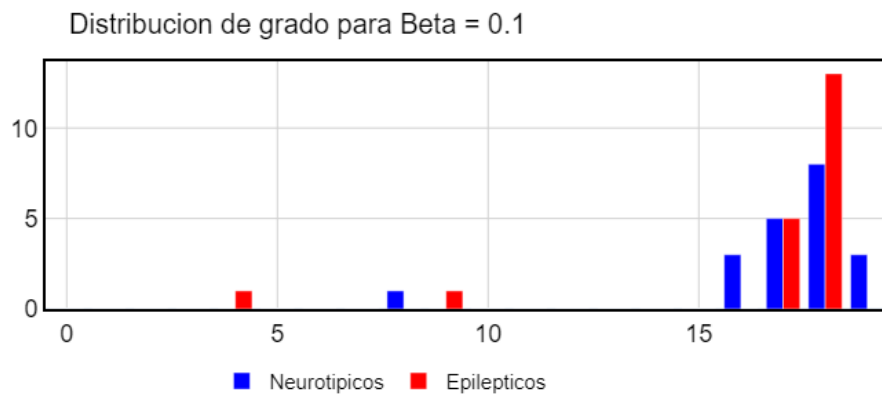
Matriz de Correlación Reordenada de Epileptico 8



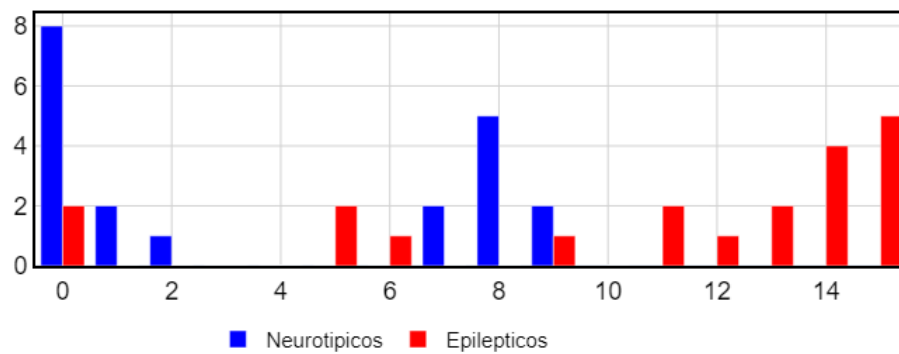
Matriz de Correlación Reordenada de Epileptico 9



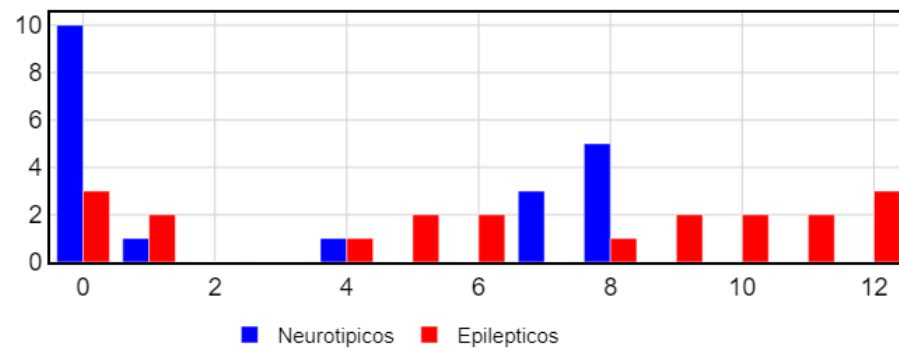
Calculando la distribución de grado de los grafos correspondientes para cada uno de los umbrales obtenemos los siguientes gráficos de barra.



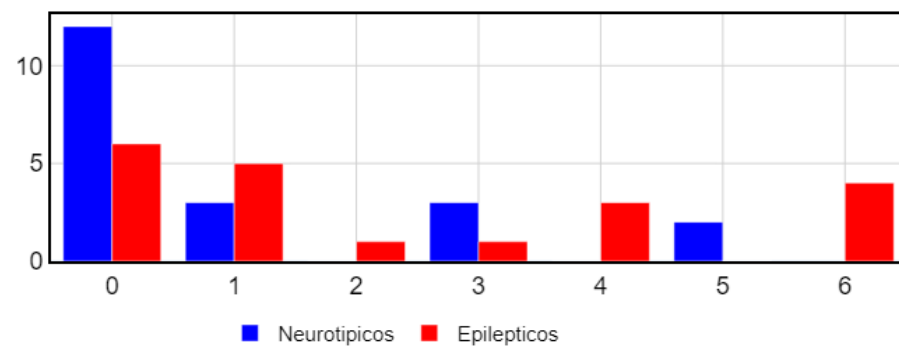
Distribucion de grado para Beta = 0.5



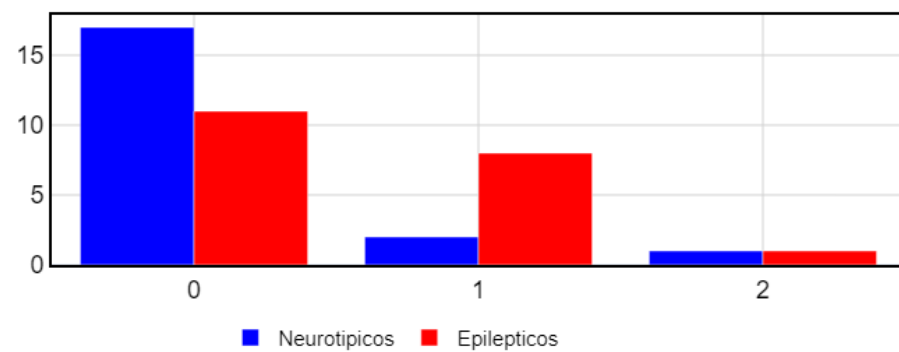
Distribucion de grado para Beta = 0.6



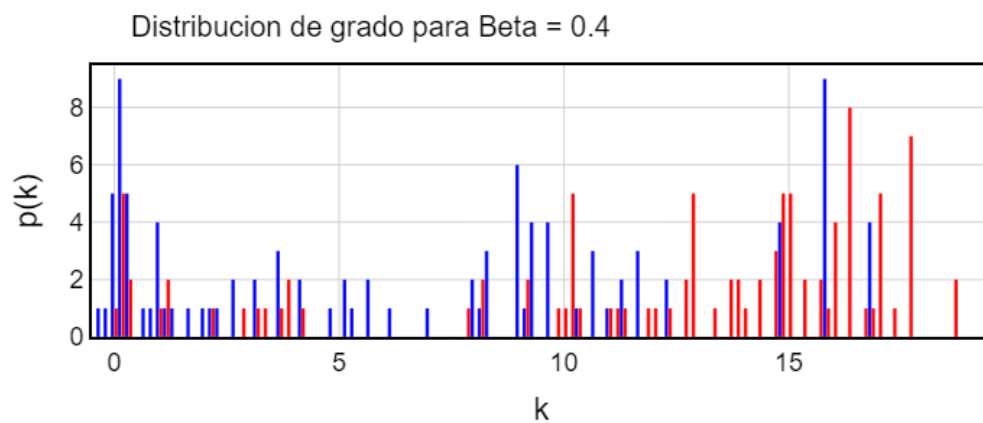
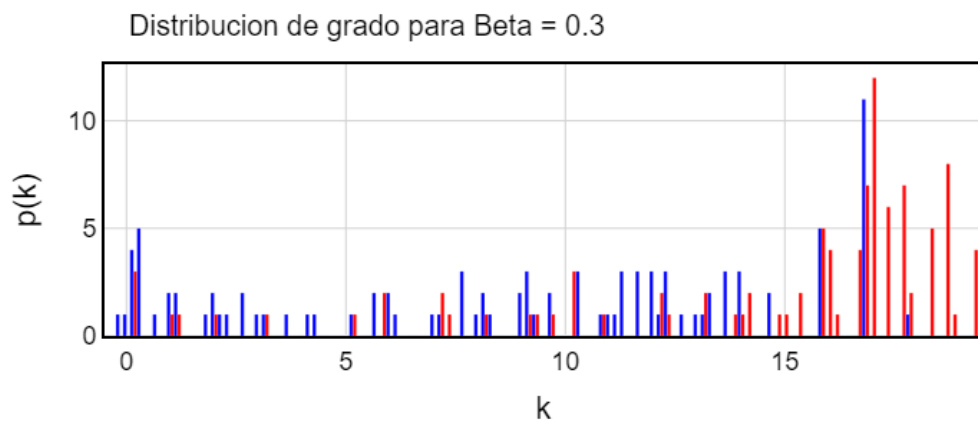
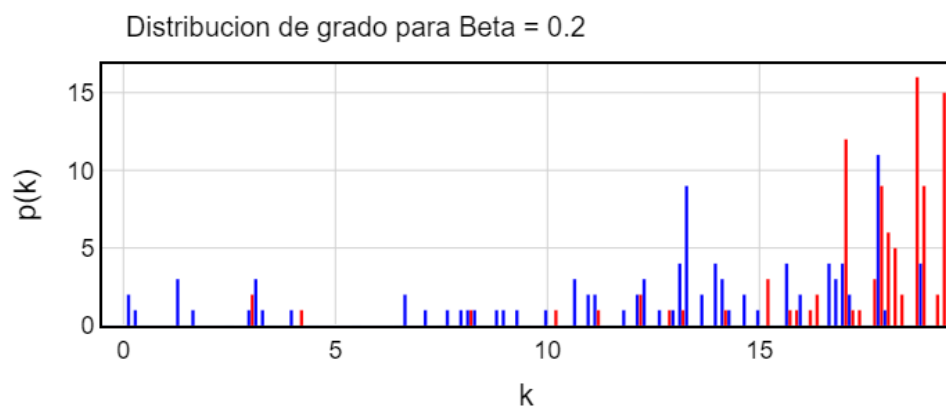
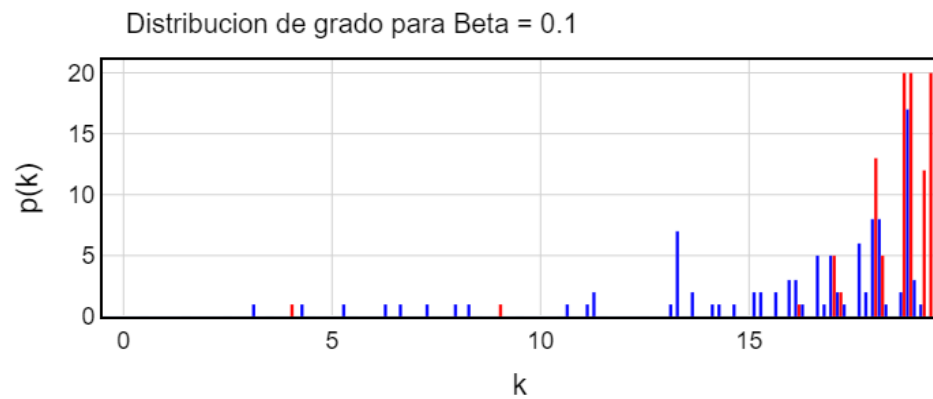
Distribucion de grado para Beta = 0.8



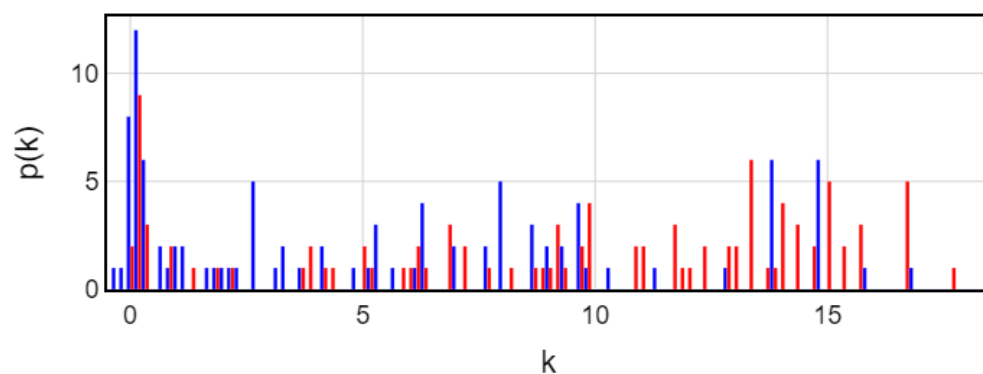
Distribucion de grado para Beta = 0.9



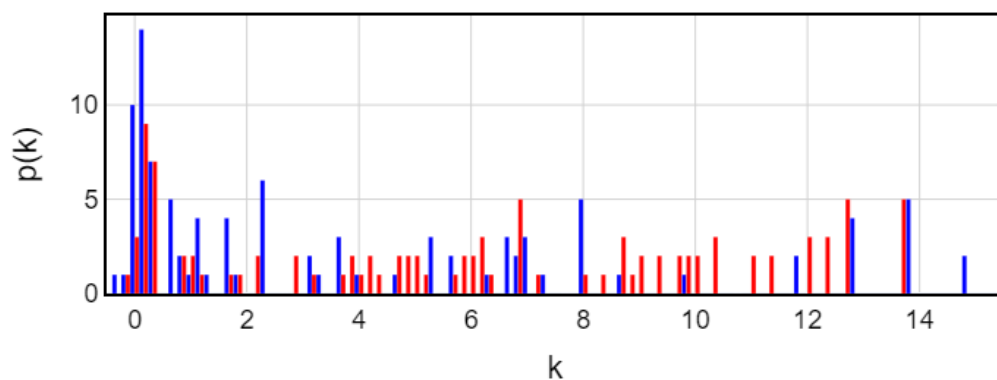
Si hacemos esto mismo para todos los pacientes y superponemos los gráficos de barra podemos ver cómo se comporta la distribución de grado para los diferentes pacientes.



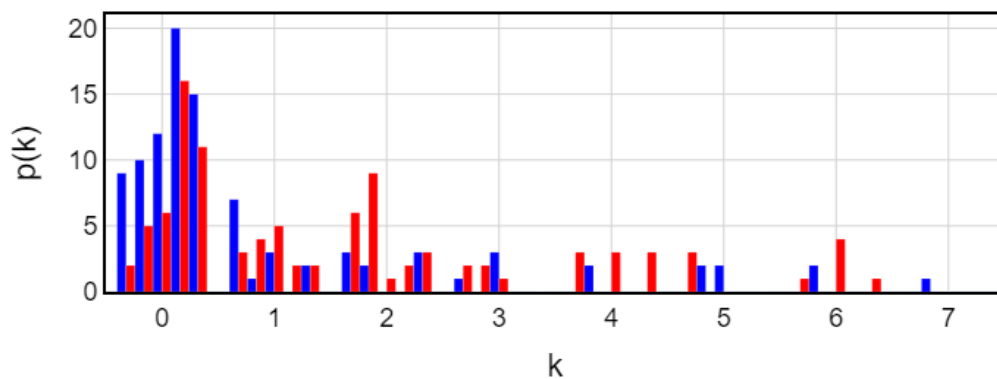
Distribucion de grado para Beta = 0.5



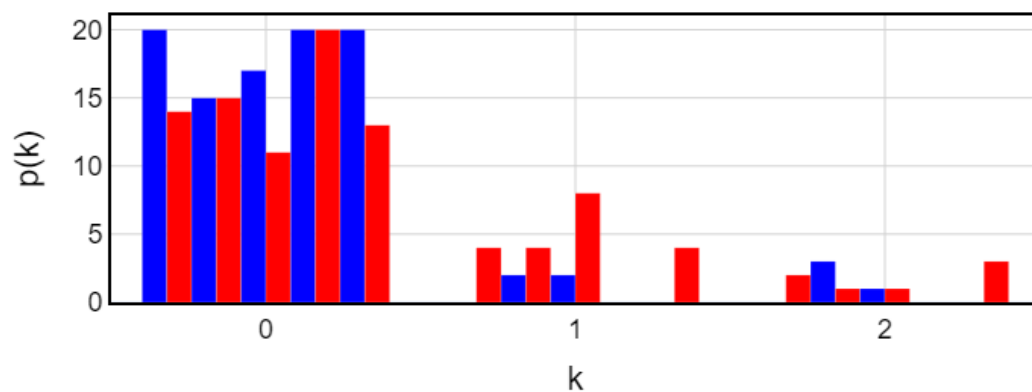
Distribucion de grado para Beta = 0.6



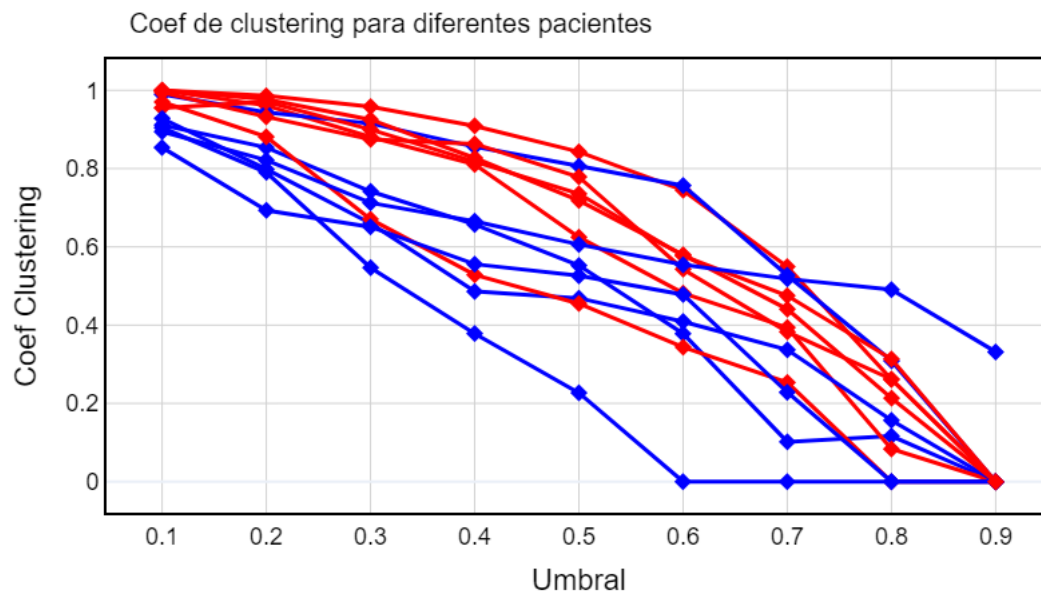
Distribucion de grado para Beta = 0.8



Distribucion de grado para Beta = 0.9



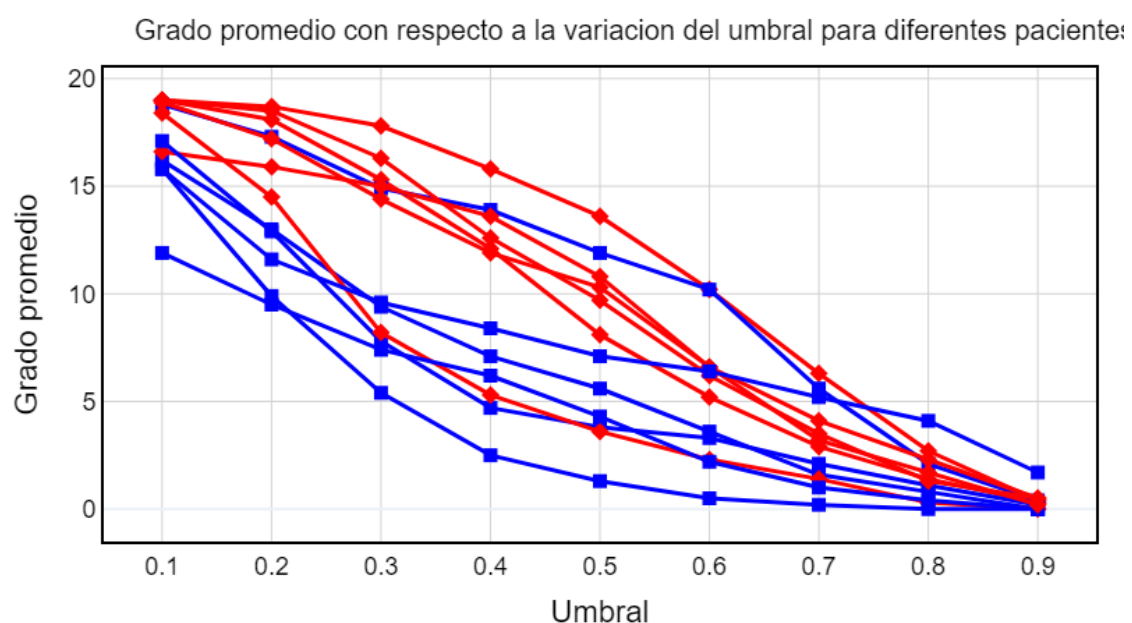
Graficando el coeficiente de clustering para los diferentes valores de umbral de todos los pacientes se obtiene el siguiente grafico:



En el gráfico podemos ver como varía el coeficiente de clustering a medida que aumenta el valor de umbral. Los diferentes trazos rojos son pacientes epilépticos y los pacientes neurotípicos en azul. que para valores bajos de umbral se ve una clara separación de los mismos. Eso en conjunto con los gráficos de distribución de grado se podrían tomar como un parámetro para elegir un valor de umbral para lograr una buena inferencia de red.

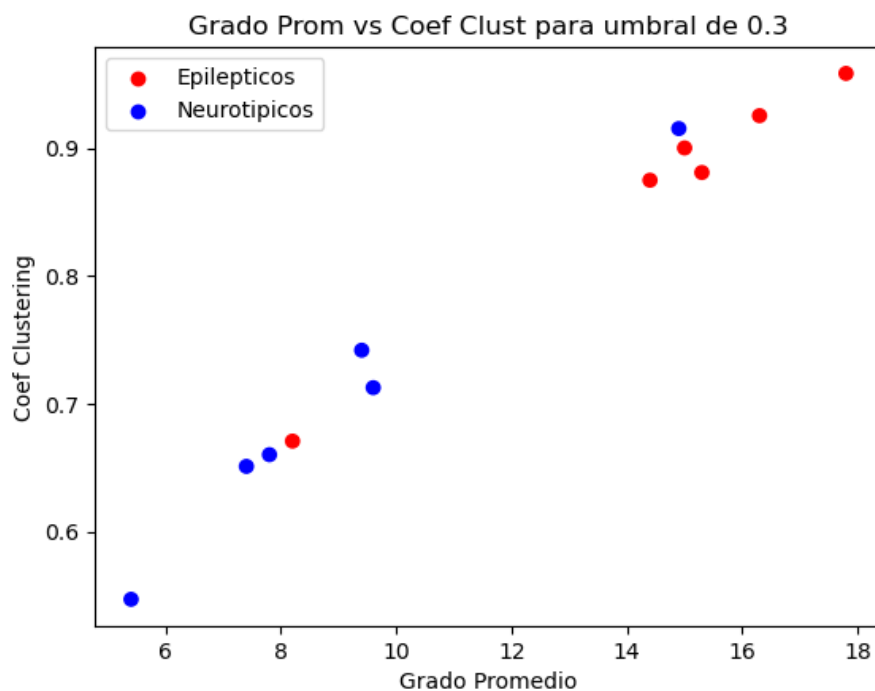
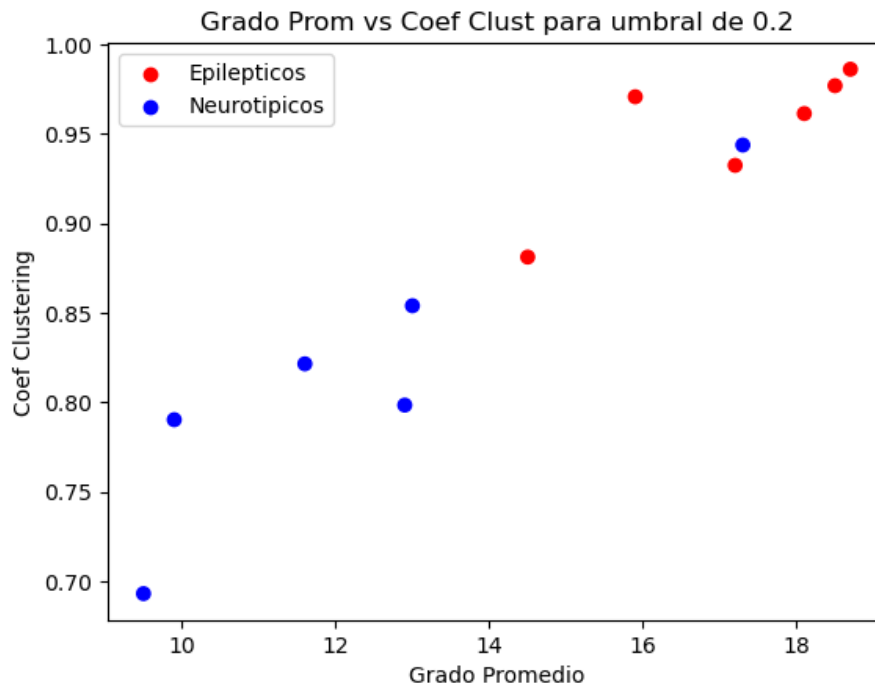
Podemos ver que para los valores de umbral 0.2 y 0.3 hay una clara separación de los coeficientes.

Por otro lado se realizó el cálculo del grado promedio de los grafos correspondientes para cada uno de los valores de umbral y se obtuvo la siguiente gráfica.



Al igual que el gráfico del coeficiente de clustering, se puede ver que para valores de umbral de 0,2 y 0,3 los valores de ambos casos presentan diferencias notables.

Si graficamos en un scatter plot el coeficiente de clustering vs el grado promedio de cada paciente para ambos valores de umbral (0.2 - 0.3) obtenemos una distribución de puntos que se separan.



6. Conclusión

Los resultados obtenidos indican que es viable utilizar una red para caracterizar a los pacientes y encontrar parámetros distintivos. A partir del análisis, se puede identificar el parámetro que mejor permita la inferencia de la red. Luego, mediante la aplicación de métricas adecuadas, es posible encontrar el valor de umbral óptimo que permita una separación efectiva de las características.

En las distribuciones de grado, podemos apreciar una distinción notable en los valores de umbral de 0.2 y 0.3 para ambos grupos. Para los pacientes epilépticos, se observa una tendencia a concentrarse en valores de grado alto, mientras que en el caso de los individuos neurotípicos, las barras se encuentran distribuidas en todos los grados. Implementando algunas técnicas que extraigan las características de las distribuciones de grado como grado acumulado, se podría obtener alguna métrica más que en conjunto con las otras nos permita la distinción de ambos casos.

Por último al observar los scatterplots de Grado promedio vs Coeficiente de clustering, se puede notar una clara separación de los pacientes epilépticos y neurotípicos. Si se dispone de más datos con una mayor cantidad de casos y se confirma que siguen el mismo patrón de separación, sería factible crear un conjunto de datos para entrenar un modelo de Machine Learning

Este modelo podría utilizarse para desarrollar un sistema que permita un apoyo en el diagnóstico de detección de casos de epilepsia. Estos avances podrían ser de gran utilidad en el campo médico.