

Teoría de grafos en redes neuronales

Jimara Urzola, Valentina Cárdenas, Brayan Sandoval, Jesús Esguerra, Jose Castro

14 de noviembre de 2025

Resumen

La amplia cantidad de información asociada a un usuario dentro de una red social depende fundamentalmente de su interacción con otros individuos. Este diseño demanda un sistema que permita organizar las relaciones entre usuarios dados sus distintos intereses y características. A su vez, este sistema requiere de algoritmos autónomos lo suficientemente capaces para garantizar la expansión de conexiones ya existentes a través de sugerencias personalizadas. En este proyecto, nos enfocaremos en la predicción de sugerencias de amistades a través de las redes neuronales gráficas, modelo que aprovecha la estructura de los grafos para representar las relación entre datos *no independientes e idénticamente distribuidos* (non-IID, por sus siglas en inglés), como lo son la interacción entre individuos. Obtendremos, finalmente, resultados que indicaran la probabilidad de formar nuevas conexiones entre usuarios a partir de una base de datos dada, mediante la ejecución reducida de este modelo en Python.

1. Teoría de grafos y redes neuronales

En esencia, la teoría de grafos es la rama de las matemáticas discretas que se enfoca en el análisis de los grafos, sus teoremas y propiedades. Estas estructuras son utilizadas en aplicaciones prácticas para modelar conexiones entre objetos.

Por otro lado, las redes neuronales son una herramienta y área de estudio perteneciente al aprendizaje automático (*machine learning*). Inspiradas en la estructura y el funcionamiento del cerebro humano, las redes neuronales se constituyen de capas de neuronas artificiales (nodos), conectadas entre sí por medio de aristas de un peso determinado.

En primera instancia, ambos conceptos pueden parecer incompatibles, quizá vagamente conexos. Sin embargo, el avance del aprendizaje automático como disciplina ha brindado, a lo largo de años de investigación teórica, un marco conceptual y práctico lo suficientemente riguroso para desarrollar lo que hoy se conoce como redes neuronales gráficas. Estas se definen formalmente como métodos de *deep learning* que operan sobre el dominio de los grafos [ZCH⁺20]. Su ventaja radica, precisamente, en la efectividad de estas estructuras para manejar conexiones complejas entre datos en el contexto de un espacio no-euclíadiano.

En general, se consideran los siguientes enunciados:

Definición. Un grafo G es un par ordenado $G = (V, A)$, formado por una conjunto finito, no vacío de vértices V , y una colección de pares de vértices A llamada aristas.

Definición. Un grafo $G = (V, A)$ es dirigido si los elementos de A (aristas) son pares ordenados de vértices. Esto es, $a = (x, y)$, con $x, y \in A$.

Definición. Resumidamente, un modelo neuronal está conformado por tres elementos: (1) Un conjunto de enlaces conectores. (2) Un enlace aditivo, que funciona como un combinador lineal. (3) Una función de activación limitada por la amplitud de los parámetros de salida de una neurona [Hay99]

Definición. Una red neuronal es un grafo dirigido constituido de nodos interconectados por enlaces sinápticos y de activación, y que satisface cuatro propiedades: (1) Cada neurona está representada por un conjunto de enlaces sinápticos lineales, un sesgo aplicado exteriormente y un posible enlace de activación no lineal. El sesgo está representado por un enlace sináptico conectado a una entrada fijada

en $+1$. (2) Los enlaces sinápticos de una neurona tienen el peso de sus respectivas entradas. (3) La suma del peso de las señales de entrada define el campo inducido de la neurona en cuestión. (4) El enlace de activación suprime el campo inducido de la neurona para producir una salida [Hay99].

2. Datos no-IID y datos no-euclidianos

Tradicionalmente, las redes neuronales se encuentran limitadas a manejar datos independientes y que operan en un espacio euclidiano (imágenes, series de tiempo, por ejemplo). Sus entradas son vectores que representan un punto específico del conjunto, independientes unos de otros. De este modo, se entiende que la relación entre datos no es relevante para la salida de estos tipos de modelo (Feedforward Neural Networks, Convolutional Neural Networks, etc.).

Por dicho motivo, para el manejo de información altamente dinámica e interconectada, el uso de redes neuronales gráficas es ampliamente superior. El modelo (GNN), entonces, comprende mejor la estructura del grafo al capturar información relevante entre vecindades de vértices, aprovechando las relaciones para obtener mejores representaciones.

3. Problema: Planteamiento y desarrollo

El problema planteado en cuestión es el siguiente:

Problema. Predecir nuevas amistades en una red social aplicando algoritmos de redes neuronales gráficas.

Para el desarrollo de este problema, hemos elegido la librería "PyTorch", de Python.

Hemos planteado el siguiente procedimiento para resolver:

- Generar un conjunto de datos sintéticos que represente a 20 individuos con intereses y conexiones de amistad iniciales
- Convertir los datos sintéticos generados en una representación de grafo adecuada para Redes Neuronales Gráficas usando PyTorch Geometric
- Definir e implementar una arquitectura de Red Neuronal Gráfica simple (capas GCNConv) diseñada específicamente para la tarea de predicción de enlaces
- Entrenar el modelo GNN implementado con los datos del grafo construidos
- Utilizar el modelo GNN entrenado para predecir las probabilidades de nuevas amistades en el conjunto de prueba
- Mostrar las probabilidades predichas de nuevas amistades

Acceda al [siguiente enlace](#) y verifique el proceso.

Referencias

- [Hay99] Simon S. Haykin. *Neural networks : a comprehensive foundation.* Prentice Hall, Upper Saddle River, N.J., 2 ed.. edition, 1999.
- [ZCH⁺20] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. *AI open*, 1:57–81, 2020.