

Graph Neural Networks

Jose Luis Rodríguez, Gonzalo Martinez y Alexandre Muñoz

November 2023

1 Graph Neural Networks

Las redes neuronales de gráficos (GNN) son una clase de métodos de aprendizaje profundo diseñados para realizar inferencias sobre datos descritos por gráficos.

Las Graph Neural Networks (GNN) aplican la potencia predictiva del deep learning a estructuras de datos ricas que representan los objetos y sus relaciones como puntos conectados por líneas en un gráfico.

En las GNN, los puntos de datos se denominan nodos, que están vinculados por líneas (llamadas bordes) con elementos expresados matemáticamente para que los algoritmos de machine learning puedan realizar predicciones útiles al nivel de nodos, edges o graphs enteros.

A diferencia de las redes neuronales convencionales que operan en datos tabulares o secuenciales, las GNNs se utilizan para aprender representaciones y realizar tareas de predicción en datos con estructura de grafo, como redes sociales, sistemas de recomendación, análisis de lenguaje natural y más.

Una descripción detallada de las GNNs:

- **Grafos como estructura de datos:** En un grafo, los datos se representan como nodos (vértices) y relaciones (aristas) entre estos nodos. Cada nodo puede contener información asociada (atributos), y las aristas representan relaciones entre nodos. Los grafos son una forma poderosa de representar datos que muestran interconexiones y dependencias entre elementos.
- **Objetivo de las GNNs:** El objetivo principal de las GNNs es aprender representaciones efectivas de los nodos en un grafo para tareas como clasificación, regresión, agrupación, recomendación, entre otras. La idea es capturar la información de los nodos vecinos para tomar decisiones sobre un nodo dado.
- **Operaciones en GNNs:** Las GNNs aplican operaciones iterativas a cada nodo y sus vecinos. Estas operaciones se realizan en múltiples capas (como en las redes neuronales convencionales). En cada capa, un nodo actualiza su representación considerando su propia información y la de sus vecinos.

- **Mensaje Passing:** La técnica central en GNNs es el "mensaje passing" o paso de mensajes. Cada nodo "envía" un mensaje a sus vecinos, y esos mensajes se agregan para actualizar la representación del nodo. Los mensajes pueden codificar información sobre los atributos del nodo, sus relaciones con los vecinos y otros detalles del grafo.
- **Agregación de Información:** Después de recibir mensajes de los vecinos, los nodos agregan esta información en una representación actualizada. Las GNNs utilizan diferentes estrategias de agregación, como la suma, promedio o redes neuronales para combinar mensajes.
- **Capas en GNNs:** Las GNNs constan de múltiples capas, donde las operaciones de mensaje passing se repiten. Cada capa captura información a diferentes distancias en el grafo. Esto permite que las GNNs comprendan relaciones locales y globales en el grafo.
- **Aprendizaje supervisado y no supervisado:** Las GNNs se pueden utilizar tanto en tareas de aprendizaje supervisado, donde se tiene información etiquetada, como en tareas de aprendizaje no supervisado, donde el objetivo es aprender patrones y estructuras en el grafo.
- **Variaciones de GNNs:** A lo largo de los años, se han desarrollado numerosas variaciones de GNNs, como Graph Convolutional Networks (GCNs), GraphSAGE, Gated Graph Neural Networks (GGNNs), Graph Attention Networks (GAT), Graph Isomorphism Networks (GIN), y más. Cada variación utiliza diferentes estrategias para actualizar las representaciones de los nodos.
- **Aplicaciones de GNNs:** Las GNNs se aplican en una amplia gama de campos, como recomendación de productos, análisis de redes sociales, química computacional, detección de fraude, análisis de lenguaje natural, procesamiento de imágenes basado en gráficos y más.
- **Desafíos:** Aunque las GNNs son poderosas, presentan desafíos, como la generalización en grafos con diferentes estructuras, la escalabilidad en grafos masivos y la interpretación de las representaciones aprendidas.

En resumen, las Graph Neural Networks son un avance significativo en el procesamiento de datos estructurados en forma de grafos, permitiendo la aplicación de técnicas de aprendizaje automático en una variedad de dominios. Están diseñadas para capturar relaciones y dependencias en grafos, y su flexibilidad las hace adecuadas para una amplia gama de aplicaciones.

2 Redes Convolucionales Gráficas

Los GCN se introdujeron por primera vez en "Redes espectrales y redes profundas conectadas localmente en gráficos" (Bruna et al, 2014), como un método para aplicar redes neuronales a datos estructurados en gráficos.

El GCN más simple tiene sólo tres operadores diferentes:

- Convolución gráfica
- Capa lineal
- Activación no lineal

Las operaciones suelen realizarse en este orden. Juntos, forman una capa de red. Podemos combinar una o más capas para formar un GCN completo.

Desarrollando un poco mas, las redes neuronales son un subconjunto del machine learning y el núcleo de los algoritmos de aprendizaje profundo. Están compuestas por capas de nodos, que incluyen una capa de entrada, una o más capas ocultas y una capa de salida. Cada nodo está conectado a otro y tiene un peso y un umbral asociados. Si la salida de un nodo individual está por encima del valor de umbral especificado, ese nodo se activa y envía datos a la siguiente capa de la red. De lo contrario, no se pasa ningún dato a la siguiente capa de la red.

2.1 ¿Cómo funcionan las redes neuronales convolucionales?

Las redes neuronales convolucionales se distinguen de otras redes neuronales por su rendimiento superior con entradas de imagen, voz o señales de audio. Se componen de tres tipos principales de capas:

Capa convolucional

Capa de agrupación

Capa totalmente conectada

La capa convolucional es la primera capa de una red convolucional. Si bien las capas convolucionales pueden ir seguidas de otras capas convolucionales o de capas de agrupación, la capa final es la capa totalmente conectada. Con cada capa, la CNN aumenta en complejidad, identificando partes cada vez más grandes de la imagen. Las primeras capas se centran en características simples, como colores y bordes. A medida que los datos de la imagen avanzan a través de las capas, la CNN comienza a reconocer elementos o formas más grandes hasta que finalmente identifica el objeto esperado.

Capa convolucional

La capa convolucional es el bloque de creación principal de una CNN, y es donde se realizan la mayoría de los cálculos. Requiere algunos componentes, como datos de entrada, un filtro y un mapa de características. Ejemplo, si la entrada es una imagen en color compuesta por una matriz de píxeles en 3D. Esto significa que la entrada tendrá tres dimensiones: altura, anchura y profundidad, que corresponden a la composición RGB en una imagen. También

hay un detector de características, conocido como kernel o filtro, que se mueve por los campos receptivos de la imagen para comprobar si la característica está presente. Este proceso se denomina convolución.

El detector de características es una matriz bidimensional (2D) de pesos que representa una parte de la imagen. A continuación, el filtro se aplica a un área de la imagen y se calcula un producto escalar entre los píxeles de entrada y el filtro. Este producto escalar se introduce en una matriz de salida. Después, el filtro se desplaza un poco y repite el proceso hasta que el kernel haya recorrido toda la imagen. El resultado final de la serie de productos escalares de la entrada y el filtro se conoce como mapa de características, mapa de activación o característica convolucionada.

Después de cada operación de convolución, la CNN aplica una transformación de unidad lineal rectificada (ReLU) al mapa de características, introduciendo la no linealidad en el modelo.

Como se ha mencionado anteriormente, la capa de convolución inicial puede ir seguida de otra capa de convolución. Cuando esto sucede, la estructura de la CNN puede volverse jerárquica, ya que las capas siguientes pueden ver los píxeles en los campos receptivos de las capas anteriores. Por ejemplo, si queremos determinar si una imagen contiene una bicicleta. Podemos considerar la bicicleta una suma de varias partes. Se compone de un cuadro, un manillar, ruedas, pedales, etc. Cada parte de la bicicleta forma un patrón de nivel inferior en la red neuronal, y la combinación de todas las partes representa un patrón de nivel superior, lo que crea una jerarquía de características dentro de la CNN.

Capa de agrupación

La agrupación de capas, también conocida como submuestreo, permite reducir la dimensión mediante la reducción del número de parámetros de la entrada. De manera similar a la capa convolucional, la operación de agrupación barre toda la entrada con un filtro, pero la diferencia es que este filtro no tiene ningún peso. En su lugar, el kernel aplica una función de agregación a los valores dentro del campo receptivo y llena así la matriz de salida. Hay dos tipos principales de agrupación:

- Agrupación máxima: a medida que el filtro recorre la entrada, selecciona el píxel con el valor más alto para enviarlo a la matriz de salida. Este enfoque suele utilizarse más que la agrupación media.
- Agrupación media: conforme el filtro avanza por la entrada, calcula el valor promedio dentro del campo receptivo para enviarlo a la matriz de salida.

Aunque se pierde mucha información en la capa de agrupación, esta tiene una serie de beneficios para la CNN. Ayuda a reducir la complejidad, mejora la eficiencia y limita el riesgo de sobreajuste.

Capa totalmente conectada

El nombre de la capa totalmente conectada describe con precisión la capa en sí. Como se ha mencionado anteriormente, los valores de píxel de la imagen de entrada no están conectados directamente con la capa de salida en las capas

parcialmente conectadas. Sin embargo, en la capa totalmente conectada, cada nodo de la capa de salida sí está conectado directamente a un nodo de la capa anterior.

Esta capa realiza la tarea de clasificación basándose en las características extraídas de las capas anteriores y sus diferentes filtros. Las capas convolucionales y de agrupación suelen utilizar funciones ReLU, mientras que las capas totalmente conectadas generalmente usan una función de activación softmax para clasificar adecuadamente las entradas y generar una probabilidad entre 0 y 1.

2.2 Ventajas de las GCN

Una de las ventajas clave de los GCN es su capacidad para aprender tanto de las características de los nodos como de la estructura del gráfico. Esto les permite capturar la rica información presente en los datos gráficos, lo que a menudo no es posible con los métodos tradicionales de aprendizaje automático. Por ejemplo, en una red social, un GCN puede aprender a identificar usuarios influyentes considerando tanto sus atributos (p. ej., número de seguidores) como su posición en la red (p. ej., centralidad). Esto puede conducir a predicciones más precisas y mejores conocimientos sobre la dinámica subyacente de la red.

Otro aspecto importante de las GCN es su capacidad para manejar gráficos a gran escala de manera eficiente. Los métodos tradicionales basados en gráficos a menudo sufren problemas de escalabilidad, ya que requieren el cálculo de relaciones por pares entre todos los nodos del gráfico. Los GCN, por otro lado, utilizan un enfoque localizado, donde cada nodo solo necesita agregar información de sus vecinos inmediatos. Esto reduce significativamente la complejidad computacional del modelo, lo que hace factible aplicar GCN a gráficos a gran escala con millones o incluso miles de millones de nodos.

2.3 Aplicaciones de las GCN

Las aplicaciones potenciales de las GCN son amplias y abarcan varios dominios. En el campo del análisis de redes sociales, los GCN se pueden utilizar para predecir el comportamiento de los usuarios, detectar comunidades e identificar usuarios influyentes. En bioinformática, los GCN pueden ayudar a analizar las redes de interacción proteína-proteína, predecir las funciones de los genes e identificar posibles objetivos farmacológicos. En el ámbito del transporte, los GCN se pueden emplear para optimizar el flujo de tráfico, predecir la congestión y modelar la propagación de enfermedades. Hay gran cantidad de posibilidades que las GCN pueden desbloquear en el mundo del análisis de datos.

A pesar de su inmenso potencial, las GCN siguen siendo una tecnología relativamente nueva y existen varios desafíos y preguntas abiertas que deben abordarse. Uno de los principales desafíos es el desarrollo de algoritmos de entrenamiento eficientes para GCN, ya que los métodos actuales pueden ser lentos y requieren mucha memoria. Otro desafío es el diseño de arquitecturas GCN más expresivas y flexibles que puedan manejar diferentes tipos de datos y tareas de gráficos. Finalmente, existe la necesidad de una mejor comprensión

de las propiedades teóricas de los GCN, como su poder expresivo, capacidades de generalización y robustez al ruido.

En conclusión, las redes convolucionales de IA y gráficos son muy prometedoras para revolucionar el campo del análisis de datos y el aprendizaje automático. Al capturar de manera efectiva los patrones y estructuras complejos presentes en los datos gráficos, los GCN pueden desbloquear información valiosa y permitir predicciones más precisas en una amplia gama de aplicaciones. A medida que avanza la investigación en esta área, podemos esperar ver modelos GCN aún más potentes y versátiles que ampliarán aún más los límites de lo que es posible con la IA y el análisis de datos.

3 Generalidad y Conclusión

Relevancia y Aplicabilidad de GNN y GCN en Otros Contextos y Problemas

Las Graph Neural Networks (GNN) y, en particular, las Redes Convolucionales en Grafos (GCN) han demostrado ser herramientas poderosas en una amplia variedad de contextos y problemas. Su capacidad para modelar la estructura y las relaciones en datos de tipo grafo las hace aplicables en diversas disciplinas. Aquí se destacan algunas reflexiones sobre su relevancia y aplicabilidad:

1. **Aplicaciones en Ciencia de Datos y Redes Sociales:** GNN y GCN han encontrado aplicaciones destacadas en ciencia de datos, especialmente en el análisis de redes sociales. Pueden ser utilizadas para descubrir patrones de conexión, identificar comunidades y prever la propagación de información en redes complejas.
2. **Biología y Química Computacional:** En biología y química, donde las interacciones moleculares pueden representarse como grafos, GNN y GCN se han utilizado para predecir propiedades químicas, descubrir nuevos compuestos y entender la estructura de proteínas.
3. **Recomendación y Filtrado Colaborativo:** En sistemas de recomendación, GNN se ha aplicado con éxito para modelar las relaciones entre usuarios y productos. Pueden capturar mejor la información contextual y las preferencias individuales, lo que mejora la precisión de las recomendaciones.
4. **Infraestructura y Logística:** En el ámbito de la infraestructura y logística, GNN puede ser utilizado para optimizar rutas de transporte, planificación de redes eléctricas y gestión de cadenas de suministro, aprovechando la capacidad de modelar conexiones complejas.

Conclusión General sobre el Potencial y Limitaciones de Estas Arquitecturas

Las Graph Neural Networks, y específicamente las Redes Convolucionales en Grafos, representan un avance significativo en el procesamiento de datos estructurados en forma de grafo. Su capacidad para incorporar información de vecindario y patrones de relación ha impulsado avances sustanciales en una variedad de campos.

Potencial

- **Flexibilidad en la Representación:** GNN y GCN ofrecen una representación flexible que se adapta a la estructura de los datos de grafo, permitiendo modelar relaciones no lineales y aprender representaciones más significativas.
- **Aplicabilidad a Diversos Dominios:** La capacidad de aplicar estas arquitecturas a problemas en distintos dominios, desde biología hasta ciencia de datos y recomendación, resalta su versatilidad.
- **Mejora en la Interpretación de Datos Complejos:** En comparación con enfoques tradicionales, GNN y GCN mejoran la interpretación de datos complejos mediante la captura de información contextual y relaciones más ricas.

Limitaciones

- **Requerimientos Computacionales:** El entrenamiento de GNN y GCN puede ser intensivo en términos computacionales, especialmente para grafos grandes, lo que puede limitar su aplicabilidad en entornos con recursos limitados.
- **Interpretabilidad:** La interpretabilidad de los modelos GNN sigue siendo un desafío, ya que entender cómo toman decisiones en entornos complejos puede ser complicado.
- **Necesidad de Datos Etiquetados:** En algunos casos, GNN puede requerir grandes cantidades de datos etiquetados para entrenarse de manera efectiva, y la disponibilidad de esos datos puede ser un desafío.

En resumen, GNN y GCN son herramientas prometedoras con un potencial sustancial en diversos campos. Aunque presentan desafíos, su capacidad para modelar datos de grafo complejos sugiere un futuro emocionante para estas arquitecturas en la resolución de problemas del mundo real.