

Identificación del proyecto

Nets4IML

Nombre del proyecto

Física estadística de la inferencia en redes para aprendizaje automático interpretable

Expediente numero

PID2022-142600NB-I00



Descripción del proyecto

Las redes son representaciones de sistemas cuyas partes constituyentes interactúan entre sí de formas no triviales. Se utilizan para modelizar una amplia gama de sistemas complejos, incluidos sistemas físicos, químicos, biológicos, de transporte y sociales. Dentro de este contexto, la inferencia en redes es el proceso de inferir la estructura subyacente o las propiedades subyacentes de una red a partir de datos observacionales. El objetivo de la inferencia en redes es reconstruir la red subyacente a partir de datos e identificar las características clave de la red, como la presencia de comunidades o la existencia de nodos clave, o caracterizar la estructura general de la red. La inferencia en redes es un paso importante en el análisis de sistemas complejos, ya que permite a los investigadores comprender los mecanismos subyacentes que gobiernan las interacciones entre las partes constituyentes del sistema. Aún más importante, también permite hacer predicciones específicas sobre los sistemas. En este último sentido, la inferencia en redes ha permitido que la ciencia de las redes se convierta en una ciencia predictiva, en lugar de puramente descriptiva.

Además de ayudar en problemas que son problemas de red per se, la inferencia en redes también está abriendo nuevas puertas para abordar un problema cada vez más importante en el aprendizaje automático estadístico, a saber, la falta de interpretabilidad de los modelos. De hecho, la falta de interpretabilidad es una fuente de preocupación en el aprendizaje profundo por varias razones: (i) dificultad para comprender el proceso de toma de decisiones del modelo; (ii) dificultad para identificar y corregir sesgos y errores; (iii) falta de confianza en el modelo; (iv) dificultad para cumplir con la normativa relacionada con la privacidad de datos y sesgos que exigen que los modelos sean interpretables; y (v) dificultad para usar el modelo para ciertas aplicaciones, como atención médica y finanzas, donde la interpretabilidad es crucial para garantizar la seguridad y la equidad. Aunque este problema parece, en principio, ir más allá de la ciencia de las redes, un enfoque prometedor que hemos presentado recientemente es la regresión simbólica bayesiana, que aprovecha las representaciones de expresiones matemáticas como redes, y transforma el problema de formular modelos interpretables en un problema de inferencia en redes.

El objetivo general de este proyecto es desarrollar herramientas de física estadística, basadas en representaciones de red, modelos de red e inferencia en redes, para desarrollar enfoques de aprendizaje automático estadístico que sean tanto interpretables como predictivos

Financiación

Entidad financiadora

MCIN/ AEI /10.13039/501100011033/ y por FEDER Una manera de hacer Europa

Importe

142.500 €