

ESTIMACIÓN DE LA INCERTIDUMBRE

Ingeniería en Inteligencia Artificial

Razonamiento bajo Incertidumbre

INTRODUCCIÓN

La estimación de la incertidumbre en problemas de inteligencia artificial sirve para **evaluar la confianza en las predicciones o decisiones** de un modelo.

Existen dos tipos de incertidumbre: epistémica y aleatoria.

1) **Espistémica**: proviene de la falta de conocimiento en el modelo. Se puede reducir con más datos o modelos más precisos.

2) **Aleatoria**: es la incertidumbre inherente en los datos mismos, debido a ruido o variabilidad intrínseca. No se puede reducir con más datos porque es una propiedad de los datos en sí.

INTRODUCCIÓN

Existen una serie de técnicas para estimar ambas incertidumbres.

- 1) **Métodos bayesianos:** permiten cuantificar la incertidumbre epistémica tratando los pesos de los modelos como distribuciones probabilísticas, en lugar de valores fijos.
- 2) **Dropout en Inferencia:** que es una forma práctica de aproximar la incertidumbre epistémica usando Dropout en la fase de inferencia.
- 3) **Ensamblado de Modelos:** entrena varios modelos de forma independiente combinando sus predicciones. La variación entre las predicciones de los diferentes modelos proporciona una estimación de la incertidumbre epistémica.

ESTUDIOS PREVIOS

Existen diferentes estudios y análisis, como los de sensibilidad y robustez, que no se diseñaron específicamente para medir la incertidumbre que ayudan a identificar áreas de incertidumbre potencial y que complementan los estudios de incertidumbre proporcionando una visión más completa de cómo un modelo puede reaccionar ante distintas condiciones.

ESTUDIOS DE ABLACIÓN

Técnica experimental que se usa para evaluar el impacto de diferentes componentes o características dentro de un modelo.

Se realiza una eliminación selectiva de uno o más elementos (como variables) y se observa cómo evoluciona el rendimiento del modelo.

Permite identificar componentes críticos e irrelevantes.

ANÁLISIS DE SENSIBILIDAD

Este tipo de análisis estudia cómo responden las salidas de un modelo a cambios en las entradas o parámetros del modelo.

Hay varias formas de hacerlo:

- 1) Análisis de sensibilidad local: examina cómo varía la salida en respuesta a pequeñas perturbaciones alrededor de un punto específico de entrada.
- 2) Análisis de sensibilidad global: evalúa cómo varía la salida a lo largo de un rango amplio de valores de entrada, permitiendo ver cómo se comporta el modelo en escenarios extremos

ANÁLISIS DE ROBUSTEZ

Los análisis de robustez son estudios donde se evalúa la capacidad del modelo para soportar variaciones o ruidos en las entradas.

Hay varias formas de hacerlo:

- 1) Pruebas de ruido o corrupción de datos: se inyecta ruido o se alteran las entradas para ver si el modelo mantiene un rendimiento adecuado.
- 2) Estudios de adversarios: se utilizan ejemplos adversos (entradas intencionalmente modificadas) para ver si el modelo hace predicciones inciertas o incorrectas.

Estos estudios son una forma de evaluar indirectamente la incertidumbre epistémica.

ESPACIO DE BÚSQUEDA

Es un conjunto de todos los estados posibles que se pueden alcanzar en el proceso de resolución de un problema. Cada estado representa una configuración única de las variables del problema.

Hay 3 tipos de espacios de búsqueda:

- 1) Espacios de Búsqueda Exhaustivos: donde se consideran todas las combinaciones posibles.
- 2) Espacios de Búsqueda con Restricciones: donde existen limitaciones sobre qué combinaciones son válidas.
- 3) Espacios de Búsqueda en Problemas Continuos: donde los estados no son discretos sino continuos, como en problemas de optimización.

ESPACIO DE BÚSQUEDA

Hay una serie de elementos a tener en cuenta en un espacio de búsqueda:

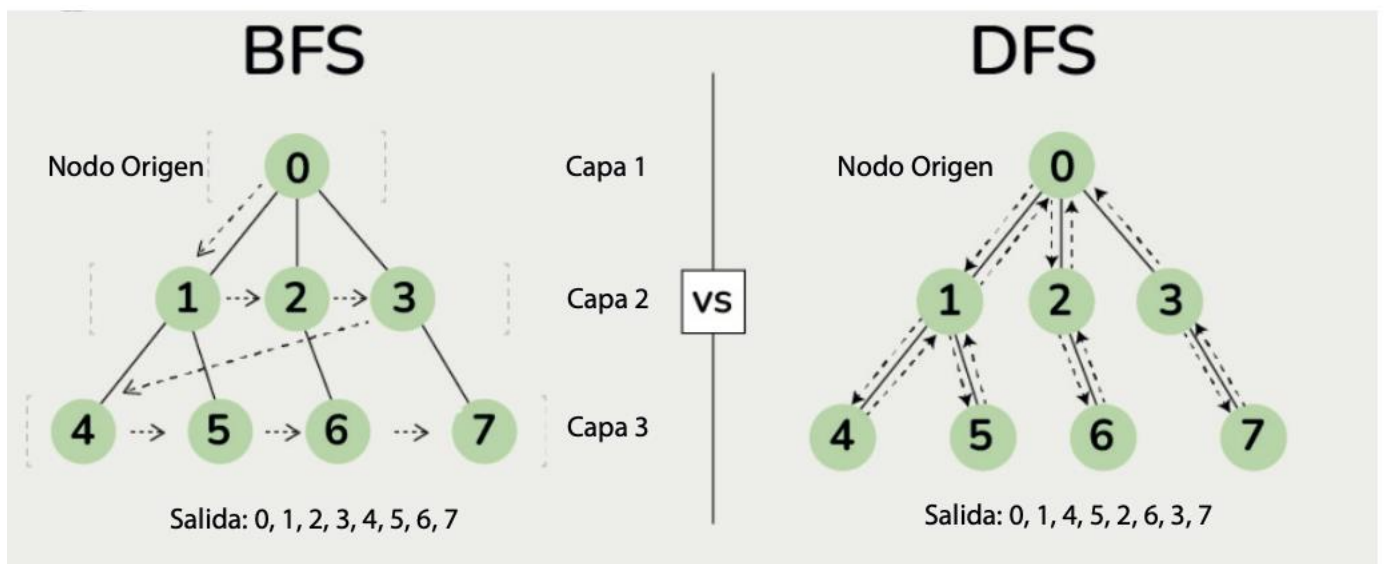
- 1) Estados: cada configuración posible del problema.
- 2) Estado inicial: el estado desde el cual se comienza la búsqueda.
- 3) Objetivo: la configuración deseada que queremos alcanzar.
- 4) Operadores: acciones que permiten movernos de un estado a otro.
- 5) Camino: secuencia de estados que conecta el estado inicial con el estado objetivo.
- 6) Costo de camino: (opcional) una medida de costo asociado con moverse entre estados, útil para problemas de optimización.

TÉCNICAS DE BÚSQUEDA

Existen diversos métodos para explorar espacios de búsqueda, divididos principalmente en técnicas de búsqueda no informada y técnicas de búsqueda informada (heurísticas).

Búsqueda no Informada: no utilizan ninguna información adicional para determinar cómo de cerca están de la meta. Exploran el espacio de búsqueda de una manera sistemática o exhaustiva.

- a) En amplitud (BFS): Busca cada nivel completo de la profundidad actual antes de avanzar a la siguiente profundidad.
- b) En profundidad (DFS): Explora cada camino hasta el final antes de retroceder y explorar otros caminos.



TÉCNICAS DE BÚSQUEDA

Existen diversos métodos para explorar espacios de búsqueda, divididos principalmente en técnicas de búsqueda no informada y técnicas de búsqueda informada (heurísticas).

Búsqueda Informada: Las técnicas de búsqueda informada utilizan una función heurística para estimar la distancia entre un estado dado y el estado objetivo. Esto permite que la búsqueda sea más eficiente, ya que se pueden priorizar los caminos que parecen más prometedores.

- a) **Búsqueda Voraz (Greedy):** utiliza únicamente una función heurística y siempre selecciona el nodo que parece estar más cerca del objetivo en cada paso. Este enfoque es rápido, pero no garantiza el camino óptimo y puede quedarse atrapado en mínimos locales.
- b) **Búsqueda en espacios continuos (algoritmos genéticos):** son técnicas de búsqueda basadas en procesos de selección natural. Son útiles en espacios de búsqueda continuos o cuando no existe una representación discreta clara.

UTILIZACIÓN DE UNA U OTRA TÉCNICA

Las técnicas como la búsqueda en amplitud y profundidad son adecuadas para problemas pequeños o sin heurística clara, mientras que la búsqueda informada, es útil para problemas con funciones heurísticas.

Los métodos evolutivos, como los algoritmos genéticos, son potentes para problemas de optimización en espacios de búsqueda continuos.

INCERTIDUMBRE EN MODELOS DE IA

La incertidumbre es fundamental en el campo de la Inteligencia Artificial y, sin embargo, es uno de los aspectos que más dificultades causa.

Hay tres fuentes principales de incertidumbre en IA:

- 1) Ruido en las observaciones: por ejemplo, errores de medición y ruido aleatorio.
- 2) Cobertura incompleta del dominio: por ejemplo, nunca se pueden observar los datos.
- 3) Modelo imperfecto del problema: por ejemplo, todos los modelos tienen errores.

La incertidumbre en modelos de IA se gestiona mediante la probabilidad:

- La probabilidad y la estadística ayudan a comprender y cuantificar el valor esperado y la variabilidad de las variables.
- La probabilidad ayuda a comprender y cuantificar la distribución y la densidad esperadas de las observaciones del dominio.
- La probabilidad ayuda a comprender y cuantificar la capacidad y la varianza esperadas en el rendimiento de nuestros modelos predictivos.