



ENSEMBLES DE ÁRBOLES

Seguimos con bagging

- Otro modelo de bagging más extremo que Random Forest es Extratrees: Extremely Randomized Trees.
- Es similar en construcción a RF pero busca mayor robustez frente al sobreajuste
- Para ello introduce más aleatoriedad en la selección de las características para el muestreo y en los umbrales para la división de los nodos



Formulación del Modelo

- Selección Aleatoria de Características: En cada nodo y para cada árbol en el conjunto, se selecciona un subconjunto aleatorio de características. Esto se hace para introducir más variabilidad en el modelo.
- Selección Aleatoria de Umbrales: A diferencia de Random Forest, donde se buscan umbrales óptimos para la división, en Extra Trees, los umbrales se seleccionan completamente al azar dentro del rango de valores de las características.
- Construcción de Árboles: Se construyen múltiples árboles de decisión utilizando las características y umbrales seleccionados aleatoriamente. Cada árbol se entrena en un subconjunto aleatorio de los datos de entrenamiento.
- Votación Ponderada: Durante la predicción, la clase final de una instancia se determina mediante la votación ponderada de todos los árboles en el conjunto.



Ventajas

- **Robustez Frente al Sobreajuste:** La selección extrema y aleatoria de características y umbrales hace que Extra Trees sea más robusto frente al sobreajuste en comparación con algunos otros algoritmos de árboles de decisión.
- **Menos Sensible a Datos Ruidosos:** La naturaleza aleatoria de Extra Trees lo hace menos sensible a datos ruidosos o atípicos, mejorando su capacidad para generalizar a nuevos datos.
- **Menos Necesidad de Ajuste de Hiperparámetros:** Extra Trees tiende a requerir menos ajuste manual de hiperparámetros en comparación con algunos otros algoritmos de ensemble.
- **Mayor Diversidad en los Árboles:** La aleatoriedad en la selección de características y umbrales conduce a una mayor diversidad entre los árboles en el conjunto, lo que puede mejorar la calidad del ensemble.
- **Facilidad de Implementación:** En general, Extra Trees es fácil de implementar y se puede utilizar como un sustituto de Random Forest con menos necesidad de ajuste.



Desventajas

- Menor Interpretabilidad: La selección extrema y aleatoria de características y umbrales puede hacer que el modelo resultante sea menos interpretable en comparación con árboles de decisión más simples.
- Mayor Costo Computacional: La aleatoriedad adicional en la construcción de árboles puede aumentar el costo computacional de Extra Trees en comparación con árboles de decisión estándar.
- Menor Control en la Variabilidad: Aunque la variabilidad puede ser una ventaja, en algunos casos, la falta de control preciso sobre la selección aleatoria de características y umbrales puede ser percibida como una desventaja.
- Posible Pérdida de Interpretación: La aleatoriedad puede hacer que sea más difícil interpretar y explicar cómo se toman decisiones específicas en el modelo, lo que puede ser importante en ciertos contextos.



- ***PREGUNTAS?***
- ***GRACIAS POR SU ATENCIÓN!!!***

- Fuente:

<https://www.linkedin.com/pulse/extra-trees-m%C3%A1s-all%C3%A1-de-la-frontera-en-arbolpedia-mora-caballero-dkyie/>

