



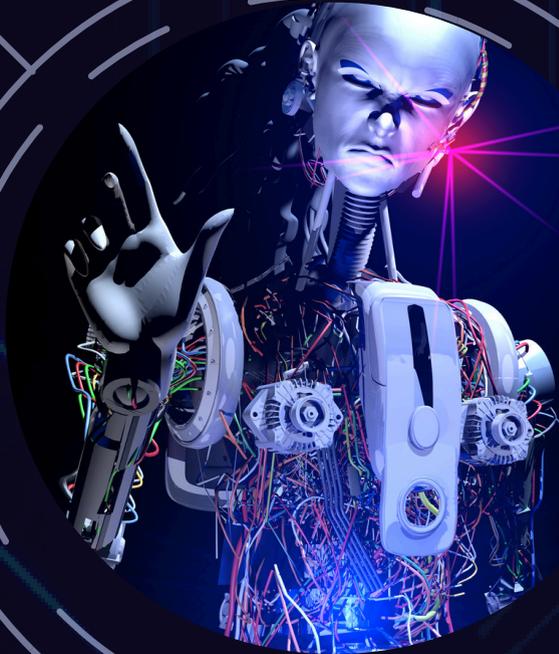
TIC



Módulo 1

Lección 2

Aprendizaje Automático
Supervisado



Contenido

5

Selección de modelos

7

Aplicaciones prácticas

6

Problemas específicos del aprendizaje supervisado

8

Avances y tendencias en aprendizaje supervisado



Haz clic sobre los títulos para navegar en cada tema.

Tema 5. Selección de modelos

La selección del modelo adecuado es una etapa crítica en el proceso de construcción de un sistema de aprendizaje supervisado. Esto ocurre por las siguientes razones:

Variedad de algoritmos	Evaluación de rendimiento	Consideraciones computacionales
Existen varios algoritmos de aprendizaje supervisado, cada uno con sus propias fortalezas y debilidades. La selección del modelo adecuado depende del tipo de problema, la naturaleza de los datos y los objetivos específicos.	Diferentes algoritmos pueden funcionar mejor para diferentes conjuntos de datos. La selección del modelo se basa en una evaluación exhaustiva del rendimiento de cada algoritmo en el contexto específico del problema.	Algunos algoritmos pueden ser más eficientes en términos computacionales que otros. La selección del modelo debe tener en cuenta la capacidad computacional disponible y los requisitos de tiempo de ejecución.

[Volver a Contenido](#)

Proceso de selección de modelos



1

División de datos: se divide el conjunto de datos en conjuntos de entrenamiento, validación y prueba. El conjunto de entrenamiento se utiliza para entrenar los modelos, el de validación para ajustar hiperparámetros y el de prueba para evaluar el rendimiento final.

2

Evaluación de modelos: se entrena cada modelo con el conjunto de entrenamiento y se evalúa su rendimiento en el conjunto de validación. Se utilizan métricas como precisión, recall, F1-score, etc, para comparar los modelos.



Proceso de selección de modelos



3

Ajuste de hiperparámetros: se ajustan los hiperparámetros del modelo para optimizar su rendimiento en el conjunto de validación. Esto puede incluir la selección de la tasa de aprendizaje, la profundidad del árbol, el número de capas en una red neuronal, etc.

4

Comparación y selección: se comparan los modelos según las métricas de rendimiento y se selecciona el modelo que mejor se ajuste al problema. También se realiza una evaluación final en el conjunto de prueba para validar la generalización del modelo.



Consideraciones en la selección de modelos

- Tipo de problema: problemas de clasificación, regresión o incluso tareas más específicas como detección de anomalías, pueden requerir enfoques de modelado diferentes.
- Características del conjunto de datos: la naturaleza de las características, la presencia de datos desbalanceados, la cantidad de datos, entre otros, influyen en la elección del modelo.
- Interpretabilidad: algunos modelos son más interpretables que otros. Dependiendo de los requisitos del problema, puede ser crucial comprender cómo toma decisiones el modelo.

La selección de modelos en el aprendizaje supervisado es un proceso iterativo y crucial. La elección adecuada impulsa el rendimiento del modelo y su capacidad para generalizar a nuevos datos; se debe tener en cuenta la naturaleza del problema y las características específicas del conjunto de datos para tomar decisiones informadas.



Tema 6. Problemas específicos del aprendizaje supervisado

Conjuntos en que se mide el rendimiento:



Rendimiento del conjunto de
entrenamiento

**Error de
entrenamiento**



Rendimiento del conjunto de
prueba

**Error de
generalización**

[Volver a Contenido](#)

Capacidad de generalización

Generalización: es la capacidad de un algoritmo de obtener un buen desempeño para entradas previamente no observadas. Error de generalización o error de prueba: se define como el valor esperado del error en una nueva entrada y se estima al medir el rendimiento en un conjunto de pruebas.

$$\frac{1}{m^{(\text{train})}} \|X^{(\text{train})}w - y^{(\text{train})}\|_2^2$$

Lo que se mide

$$\frac{1}{m^{(\text{test})}} \|X^{(\text{test})}w - y^{(\text{test})}\|_2^2$$

Lo que se quiere mejorar

¿Estrategias para abordar desafíos como el desbalance de clases, sobreajuste y subajuste?

Tema 7. Aplicaciones prácticas



Estudio de casos



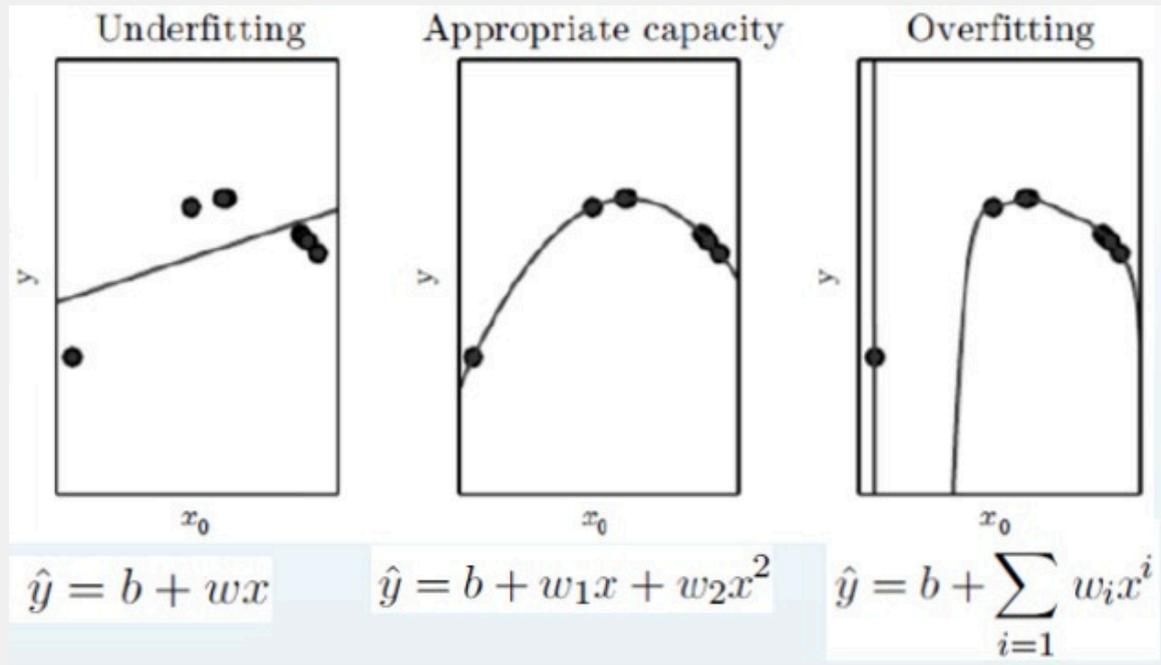
Análisis de aplicaciones prácticas del Aprendizaje Supervisado en la vida real



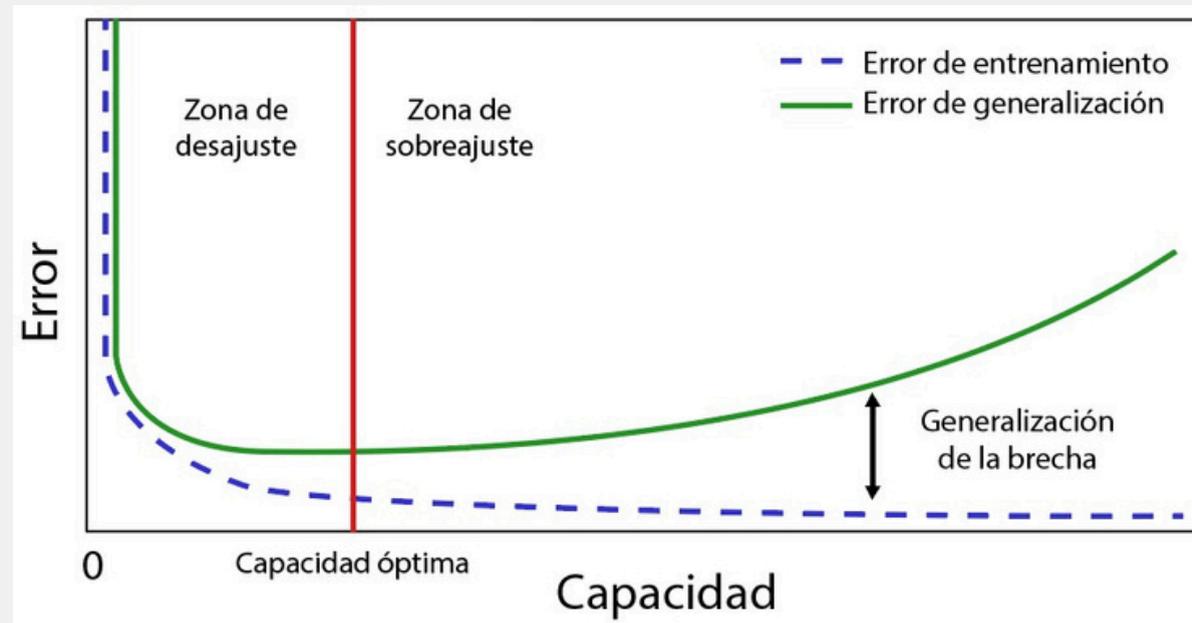
Ejemplos de cómo los modelos supervisados han abordado problemas específicos

 [Volver a Contenido](#)

Ejemplo de regresión polinomial



Capacidad vs Error



Tema 8. Avances y tendencias en aprendizaje supervisado

Modelos preentrenados: se refiere a modelos que han sido entrenados previamente en grandes conjuntos de datos y tareas específicas antes de ser ajustados para una tarea particular.

Ventajas en transferencia de conocimiento: los modelos preentrenados poseen conocimientos generales que pueden ser transferidos a nuevas tareas, acelerando el aprendizaje y mejorando el rendimiento con conjuntos de datos más pequeños.

 [Volver a Contenido](#)



- Ejemplos de Aplicación en NLP (Procesamiento del Lenguaje Natural): modelos preentrenados como BERT o GPT-3 han revolucionado el campo del procesamiento del lenguaje natural, logrando resultados destacados en tareas como comprensión del lenguaje y generación de texto.
- Técnicas de Transfer Learning: es un paradigma donde un modelo entrenado en una tarea general se adapta para tareas específicas relacionadas.
- Aplicaciones en visión por computadora: modelos entrenados en grandes conjuntos de datos para reconocimiento de imágenes, pueden transferir conocimientos a tareas específicas como detección de objetos o clasificación de imágenes médicas.



- Adaptación a dominios específicos: las técnicas de transfer learning permiten adaptar modelos a dominios específicos con menos datos de entrenamiento, facilitando su aplicación en contextos empresariales o industriales.
- Desafíos y futuras direcciones: la investigación se centra en desarrollar modelos preentrenados y técnicas de transfer learning que puedan generalizar conocimientos a una amplia gama de tareas.

Interpretabilidad y ética: a medida que los modelos preentrenados se vuelven más

- complejos, la interpretabilidad y las consideraciones éticas en el uso de grandes conjuntos de datos son temas clave a abordar.