



Técnicas de Muestreo y Métodos Estadísticos

Técnicas de Muestreo



Bagging

Reduce la varianza mediante muestreo bootstrap. Evita el sobreajuste promediando predicciones.



Boosting

Aprendizaje secuencial que maneja heterocedasticidad. Fortalece modelos débiles.



Stacking

Combina múltiples modelos mediante un meta-modelo. Flexible, pero puede ser poco transparente.

Nota: Las técnicas pueden coincidir en propósito, pero difieren en sus mecanismos.



by Manoel Fernando Alonso Gadi

Tipos de Métodos



GLM / Lineales

Asumen relaciones aditivas y lineales.



Basados en Árboles

No paramétricos (Árboles de decisión, Random Forest, XGBoost, LightGBM, CatBoost). Submuestran datos.



No Lineales

Deep Learning, Redes Neuronales, RNN. Alta complejidad, identifican patrones no lineales.

Aprendizaje No Supervisado

- Análisis factorial → agrupa columnas
- Análisis de clúster → agrupa filas

Modelos de Expertos

- AHP: Experto mejora un modelo base
- Conjoint: Experto sin necesidad de modelo base
- Elección Discreta: Usuario no experto
- MaxDiff: Usuario con poco tiempo

Series Temporales

- Datos con patrones: ARIMA, ETS, Prophet
- Datos sin patrón: TBATS, LSTM

Técnicas Complementarias

Selección de Variables

Favorece la simplicidad: "menos es más"

PCA / SEM

Reduce dimensiones, enfrenta multicolinealidad y sombras poblacionales

Sectorización de Filas

Maneja la no independencia de observaciones (ej. ratings por sector)



Regularización (L1/L2)

Corrige multicolinealidad

Vectores de Soporte

Elimina valores atípicos en espacios multidimensionales

Marcado de Outliers

Detecta valores atípicos mediante variables binarias