

Introducción a Machine Learning

Bias-Variance Tradeoff

MSc. Marco Sobrevilla

Objetivo



- Aprender conceptos relacionados a sub y sobre ajuste.

Agenda



- **Introducción**
- **Error por bias**
- **Error por variance**

Introducción

- Aprendizaje Supervisado
 - Algoritmo aprende un modelo a partir de data de entrenamiento
- Estimar el mejor mapeamiento (f) entre una salida «y» y una entrada «x»

Introducción

- Todo algoritmo tiene errores que están compuestos por:
 - *Bias*
 - Varianza (*Variance*)
 - Error irreducible
- Error irreducible
 - Causado por variables ocultas que no conocemos

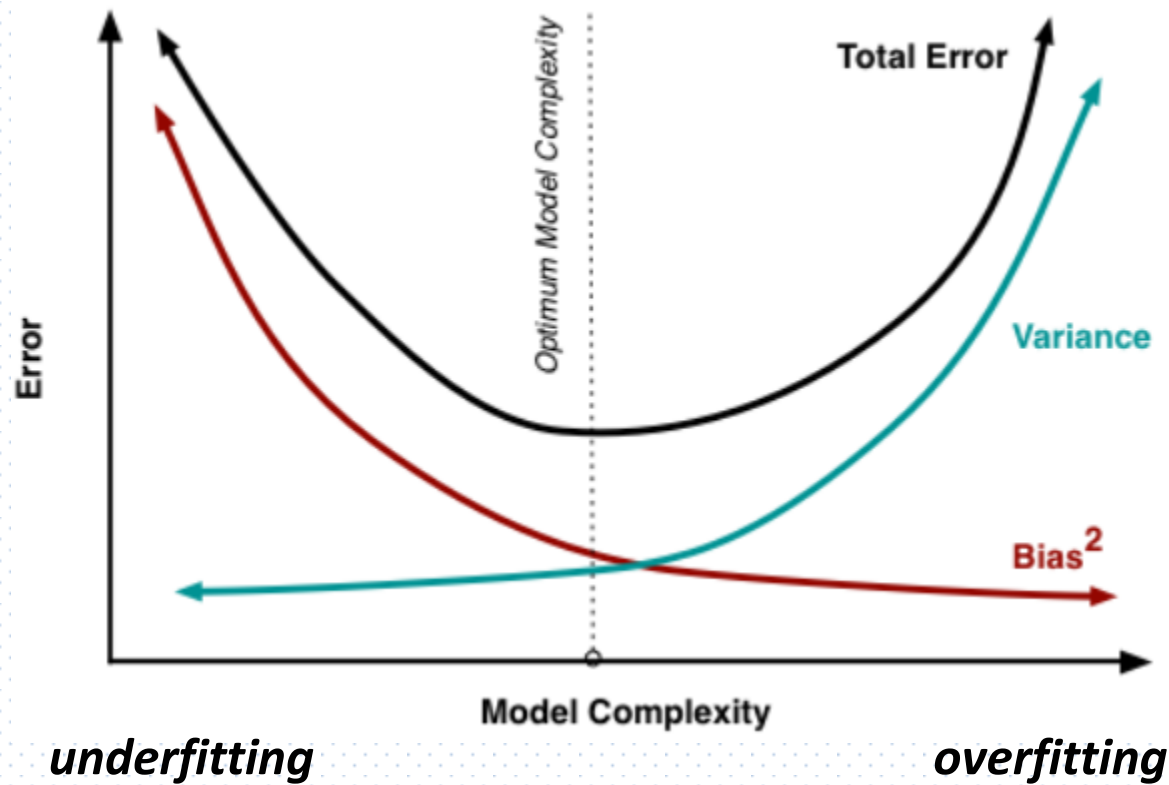
Error por Bias

- Error por *Bias*
 - Son suposiciones simplificadas hechas por un modelo para hacer que la función objetivo sea más fácil de aprender.
 - Esta relacionado a la diferencia entre lo esperado y lo obtenido

Error por Variance

- Error por *Variance*
 - Es la cantidad que la estimación de la función objetivo cambiará si se usan diferentes datos de entrenamiento.

Tradeoff



Tradeoff

- **Sub-ajuste**
 - El modelo usa muy pocas características.
 - La data de entrenamiento es insuficiente.
- **Sobre-ajuste**
 - Usar demasiadas características hace que el modelo "replique la data".
 - No podrá generalizar bien para data nueva.

Fin 😊