

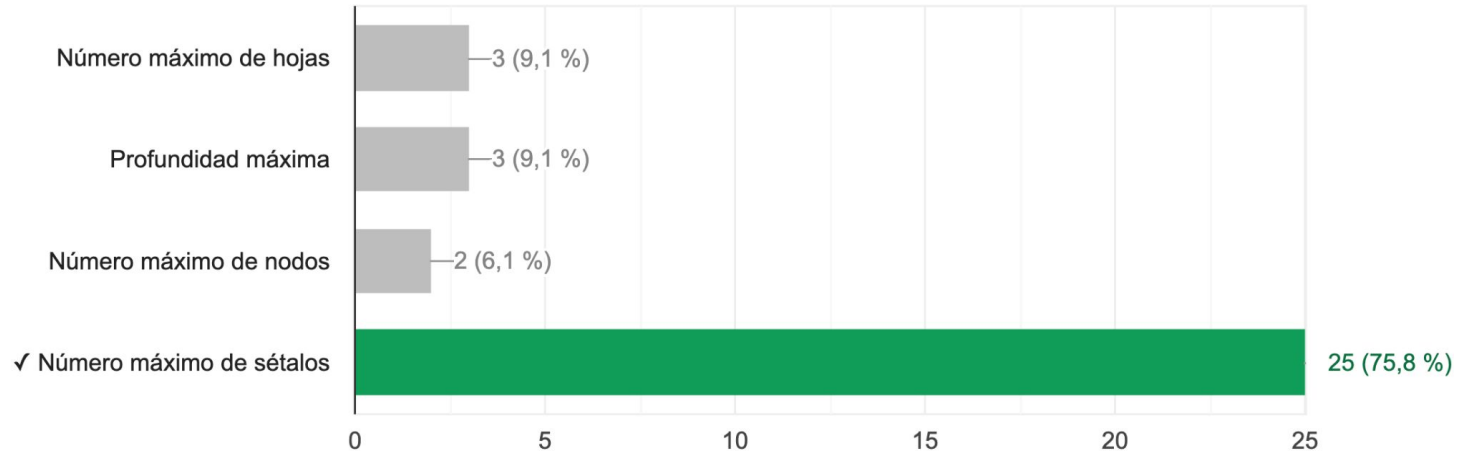


Form: Árboles



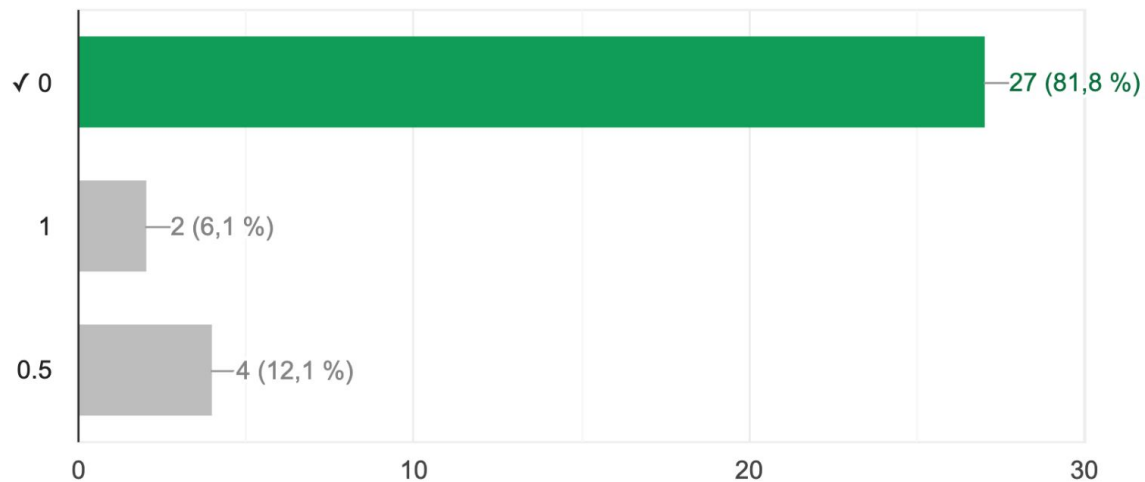
Cual de los siguientes NO es un hiper-parámetro de un árbol de decisión?

25 de 33 respuestas correctas



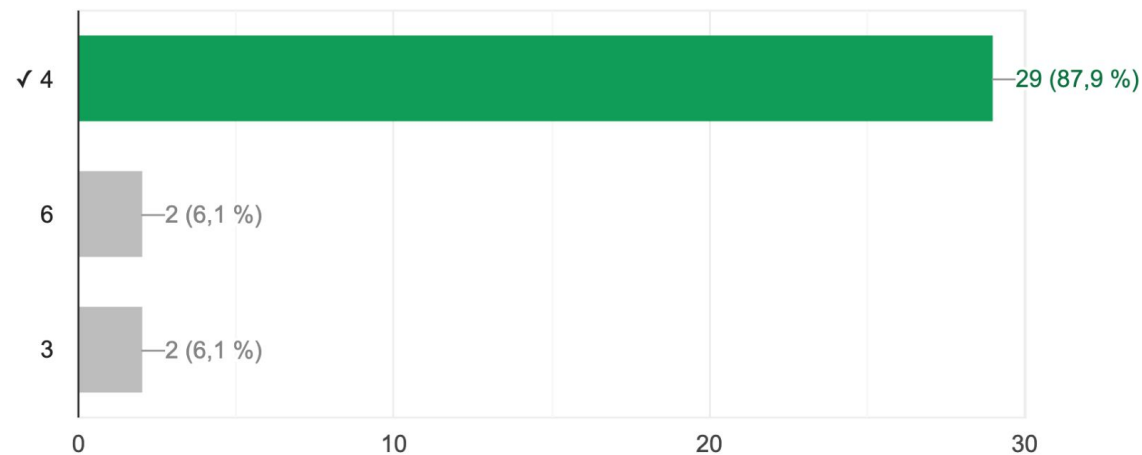
¿Cuál es la impureza de Gini de un nodo cuyas instancias pertenecen todas a una misma clase?

27 de 33 respuestas correctas



¿Cual es la profundidad (depth) del siguiente árbol?

29 de 33 respuestas correctas



IAA-2023

Clase 7: Sesgo y Varianza



UNSAM
UNIVERSIDAD
NACIONAL DE
SAN MARTÍN

Descargo de Responsabilidad

- El contenido de esta clase no entra en el parcial de la semana próxima.
- Los temas que entran en el parcial son hasta la semana pasada (inclusive).

Probabilidades

Cuando tenemos muestras en vez de distribuciones de probabilidad, podemos estimar los siguientes estadísticos como sigue:

$$\{x_i\} \quad i = 1 \dots N$$

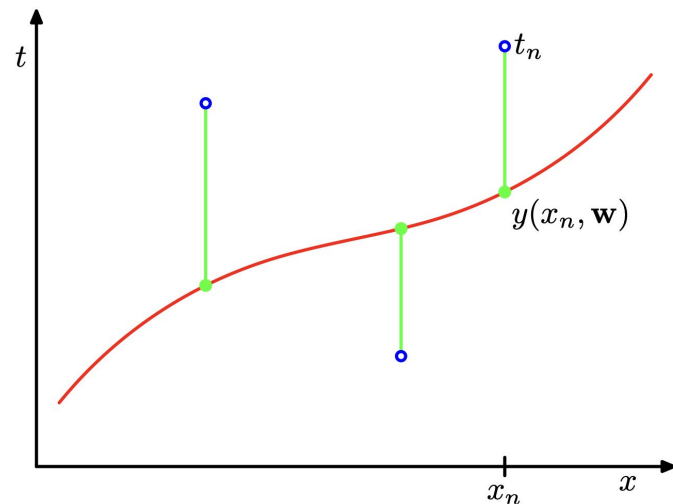
$$\mathbb{E}_f[x] \longrightarrow \bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$$

$$\text{var}_f[x] \longrightarrow \bar{S} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2$$

Conceptos Básicos: Regresor Lineal

$$y = \omega_0 + \omega_1 \cdot x$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - t_i)^2$$



Conceptos Básicos: Flexibilidad del modelo

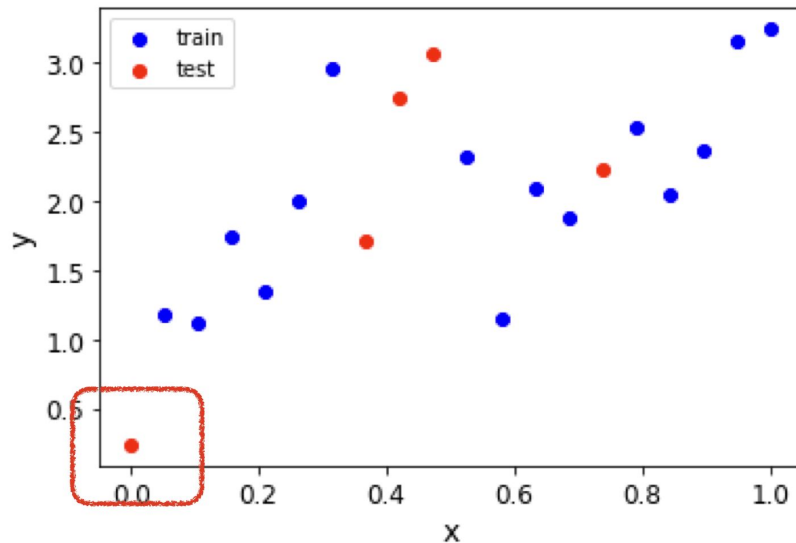
Problemas en los datos

Al hacer regresión polinómica, la performance puede depender muy fuertemente de unos pocos puntos.

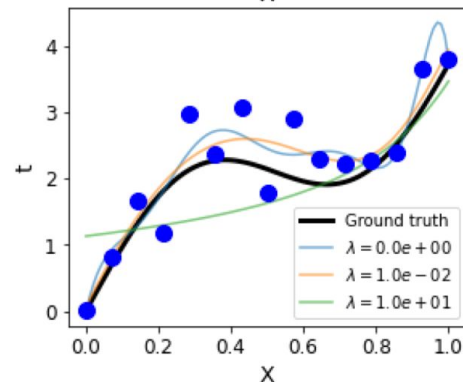
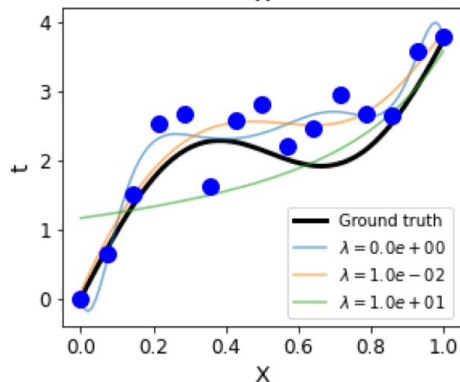
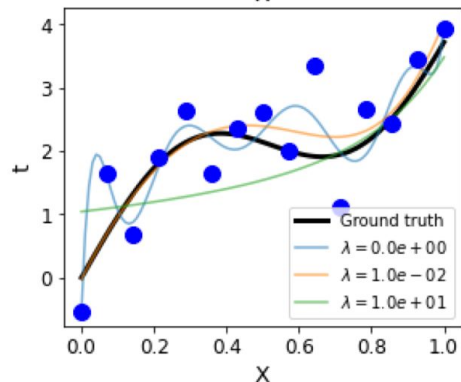
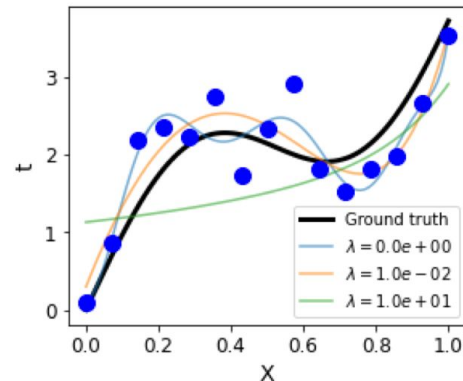
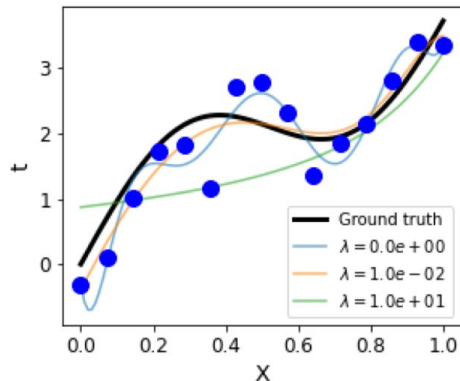
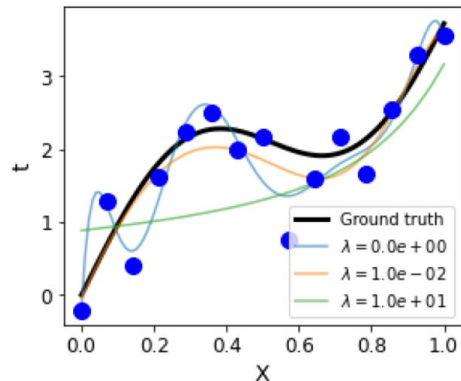
Regularización

Hemos visto que la flexibilidad del modelo (sea lo que sea eso) puede ser controlada por el hiper-parámetro de regularización λ .

- ¿Cómo podemos estimar el error esperado de un modelo, independientemente de los datos sobre los que se calcula?
- ¿Podemos sacar conclusiones basadas en un solo set de datos?
- ¿Qué podemos decir acerca de la flexibilidad del modelo?



¿Qué ocurre con diferentes sets de datos?



¿Qué querríamos tener?

- Alguna clase de **valor esperado** del error sobre diferentes sets de datos

$$\mathcal{D} = \{\mathcal{D}_1, \dots, \mathcal{D}_N\}$$

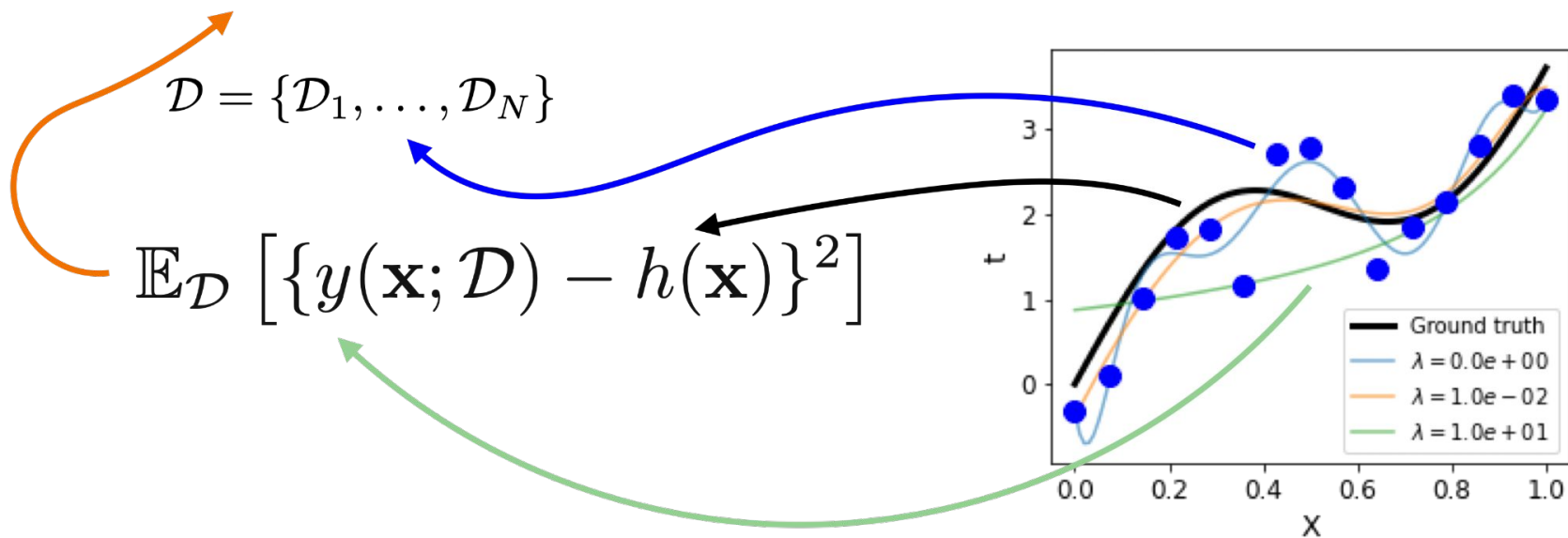
$$\begin{aligned}\mathbb{E}_{\mathcal{D}}(MSE) &= \mathbb{E}_{\mathcal{D}} \left[\{y - h(\mathbf{x})\}^2 \right] \\ &= \mathbb{E}_{\mathcal{D}} \left[\{y(\mathbf{x}; \mathcal{D}) - h(\mathbf{x})\}^2 \right] + \mathbb{E}_{\mathcal{D}} \left[\{h(\mathbf{x}) - t\}^2 \right]\end{aligned}$$

σ^2
*Error Irreducible
(o error de Bayes)*

- Independiente del modelo y
- Solo asociado al proceso que genera los datos

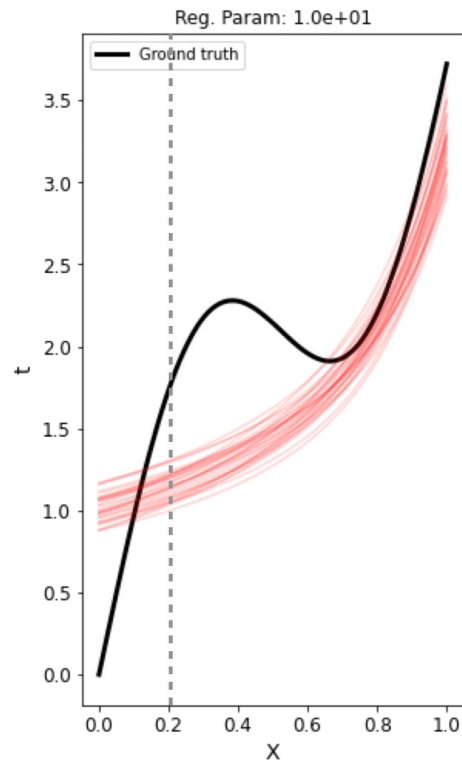
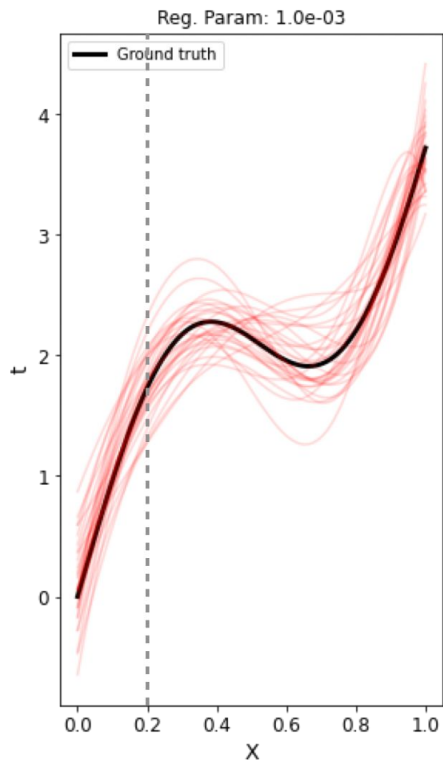
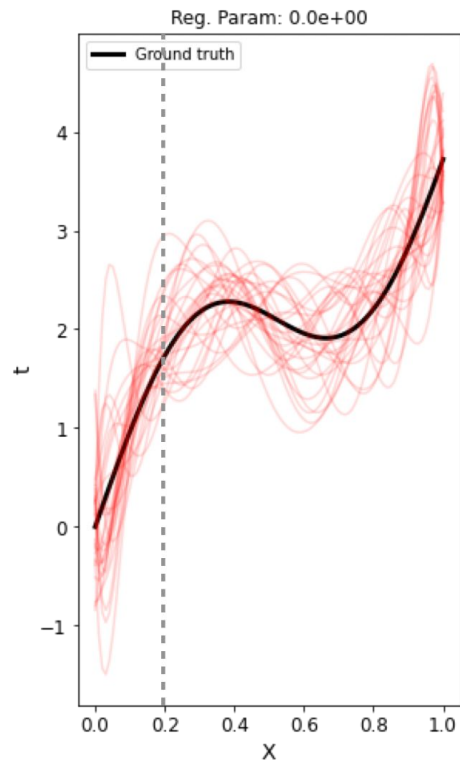
Conceptos Básicos

- Alguna clase de **valor esperado** del error sobre diferentes sets de datos

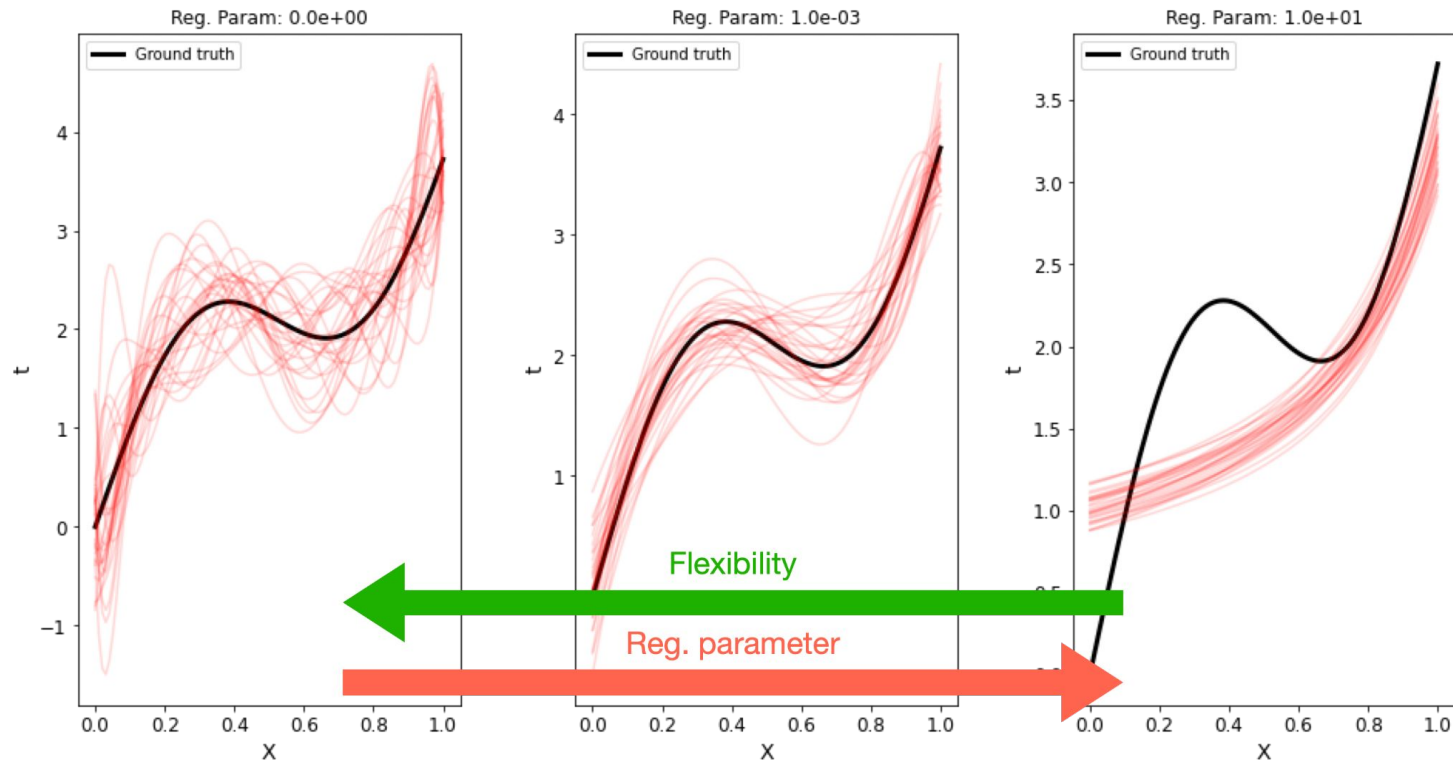


- Por simplicidad, consideremos esta cantidad para un valor *fijo* de \mathbf{x}

Conceptos Básicos



Conceptos Básicos



Conceptos Básicos

$$\mathcal{D} = \{\mathcal{D}_1, \dots, \mathcal{D}_N\}$$

$$\begin{aligned} \mathbb{E}_{\mathcal{D}} [\{y(\mathbf{x}; \mathcal{D}) - h(\mathbf{x})\}^2] \\ = \underbrace{\{\mathbb{E}_{\mathcal{D}}[y(\mathbf{x}; \mathcal{D})] - h(\mathbf{x})\}^2}_{(\text{bias})^2} + \underbrace{\mathbb{E}_{\mathcal{D}} [\{y(\mathbf{x}; \mathcal{D}) - \mathbb{E}_{\mathcal{D}}[y(\mathbf{x}; \mathcal{D})]\}^2]}_{\text{variance}} \end{aligned}$$

$$\mathbb{E}_{\mathcal{D}}(MSE) = (\text{bias})^2 + \text{variance} + \sigma^2$$

Conceptos Básicos

