

# Métricas de calidad de un Modelo de Predicción

MACHINE LEARNING

---

Docente: Marvin Jiménez

# Tipos de métricas según la respuesta

---

Las métricas de calidad de un modelo predictivo dependen de la naturaleza de la variable respuesta del modelo:

1. Variable respuesta **cuantitativa**
2. Variable respuesta **cualitativa**

# Caso 1: Variable respuesta **cuantitativa**

---

Dependiendo del interés, las métricas de calidad pueden ser para medir:

1. *Bondad de ajuste del modelo*
2. *Error de predicción*

# Medidas de error de predicción

---

Se dispone de un conjunto de datos (data set) con  $n$  observaciones y  $p$  variables.  
Sean:

$y_i$ : Valor de la variable respuesta de la observación  $i$ .

$\hat{y}_i$ : Predicción del valor de la variable respuesta de la observación  $i$ .

Entonces:

$e_i$ : Error de predicción de la observación  $i$ .

$$e_i = y_i - \hat{y}_i$$

Obs	Var. Resp	Pred 1	Pred 2	...	Pred $p$
1	233,4	23	1,81	.	1800.000
2	167,9	31	1,54	.	2500.000
3	198,5	18	1,72	.	3400.000
4	210,6	25	1,67	.	1350.000
5	.	.	.	.	.
6	.	.	.	.	.
7	.	.	.	.	.
8	.	.	.	.	.
9	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
$k$	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
$n - 2$	.	.	.	.	.
$n - 1$	.	.	.	.	.
$n$	.	.	.	.	.



$i$	$y$	$x_1$	$x_2$	...	$x_p$
1	233,4	23	1,81	.	1800.000
2	167,9	31	1,54	.	2500.000
3	198,5	18	1,72	.	3400.000
4	210,6	25	1,67	.	1350.000
5	.	.	.	.	.
6	.	.	.	.	.
7	.	.	.	.	.
8	.	.	.	.	.
9	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
$k$	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
$n - 2$	.	.	.	.	.
$n - 1$	.	.	.	.	.
$n$	.	.	.	.	.

$i$	$y$	$x_1$	$x_2$	...	$x_p$
1	233,4	23	1,81	.	1800.000
2	167,9	31	1,54	.	2500.000
3	198,5	18	1,72	.	3400.000
4	210,6	25	1,67	.	1350.000
5	.	.	.	.	.
6	.	.	.	.	.
7	.	.	.	.	.
8	.	.	.	.	.
9	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
$k$	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
$n - 2$	.	.	.	.	.
$n - 1$	.	.	.	.	.
$n$	.	.	.	.	.



$i$	$\hat{y}_i$
1	241,2
2	150,3
3	200
4	215,1
5	.
6	.
7	.
8	.
9	.
.	.
.	.
.	.
$k$	.
.	.
.	.
$n - 2$	.
$n - 1$	.
$n$	.

# ME (Mean Error)

---

$$ME = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)$$

# MPE (Mean Percent Error)

---

$$MPE = \frac{1}{n} \sum_{i=1}^n \frac{y_i - \hat{y}_i}{y_i}$$



# MSE (Mean Squared Error)

---

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

# RMSE (Root Mean Squared Error)

---

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

# MAE (Mean Absolute Error)

---

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

# MAPE (Mean Absolute Percent Error)

---

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{|y_i|}$$

# Bibliografía

---

- Trevor H, Robert T, Jerome F, *The Elements of Statistical Learning*, 2ª Edición, Springer.
- Gareth J, Daniela W, Trevor H, Robert T, *An Introduction to Statistical Learning with Applications in R*, 6ª Edición, Springer.
- Ethem A, *Introduction to Machine Learning*, 2ª Edición, The MIT Press Cambridge, Massachusetts.
- Max Kuhn, et al., *The caret package*, R CRAN, disponible en <http://topepo.github.io/caret/index.html>.
- Montgomery, D.; Peck, E.; Vining G.G. (2006). *Introducción al análisis de Regresión Lineal*. (3era Edición).