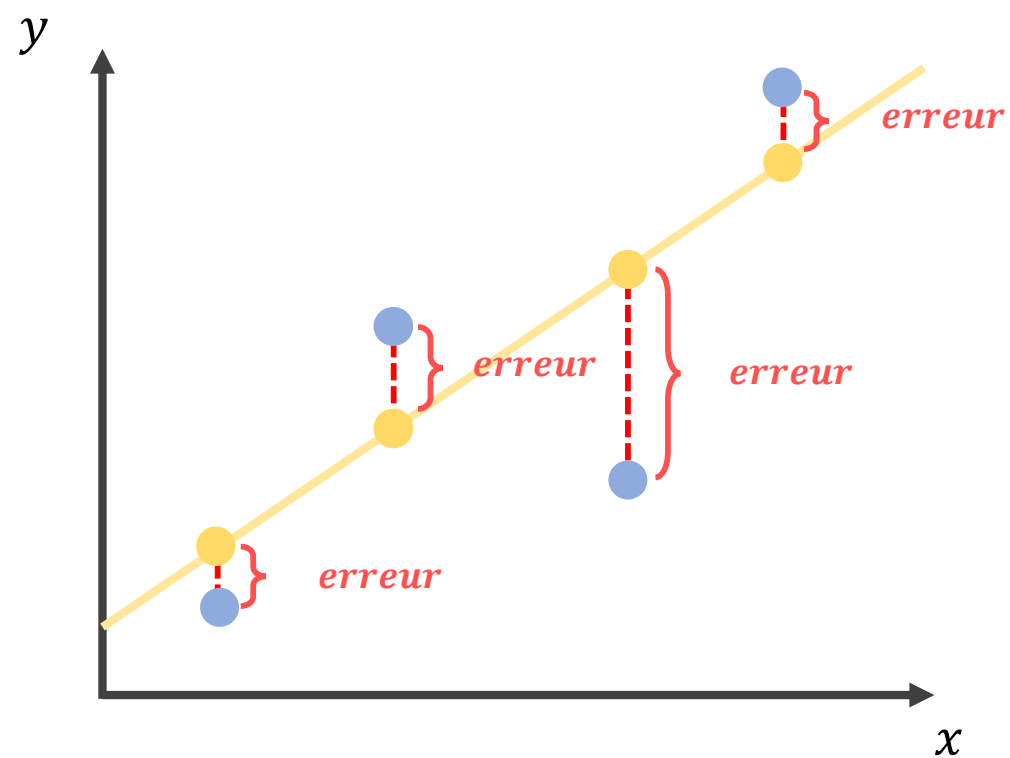


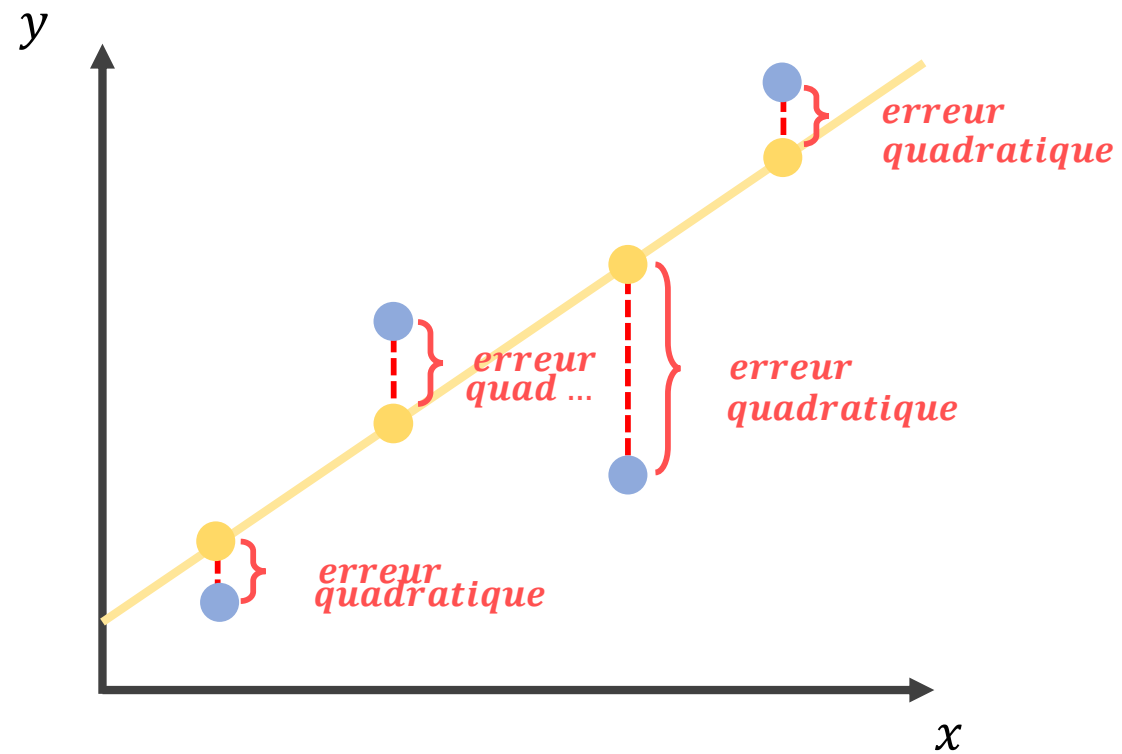
la moyenne des erreurs

$$\frac{1}{m} \sum \textit{erreur}$$



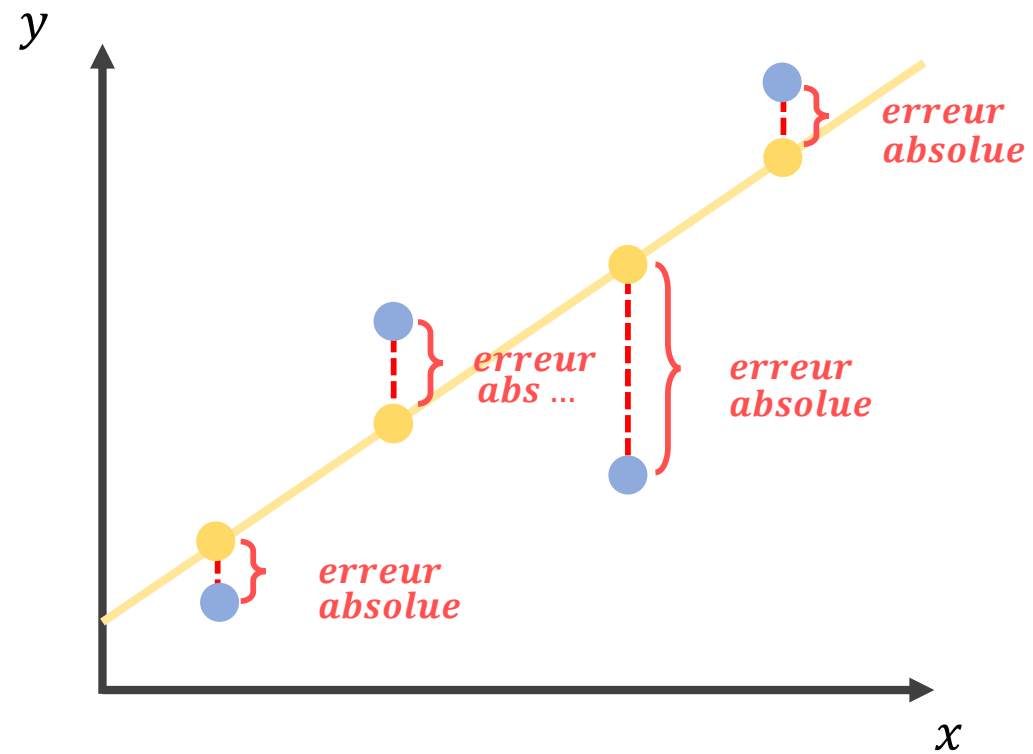
Erreur quadratique moyenne

$$MSE = \frac{1}{m} \sum (y_{vrai} - y_{pred})^2$$



Erreur absolue moyenne

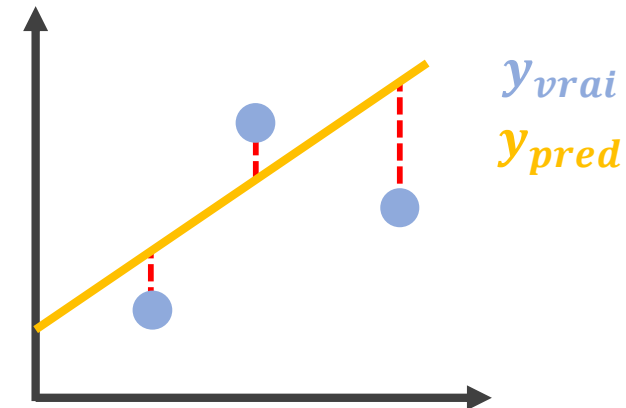
$$MAE = \frac{1}{m} \sum |y_{vrai} - y_{pred}|$$



MAE
MSE
RMSE

$$MAE = \frac{1}{m} \sum |y_{vrai} - y_{pred}|$$

$$MSE = \frac{1}{m} \sum (y_{vrai} - y_{pred})^2$$



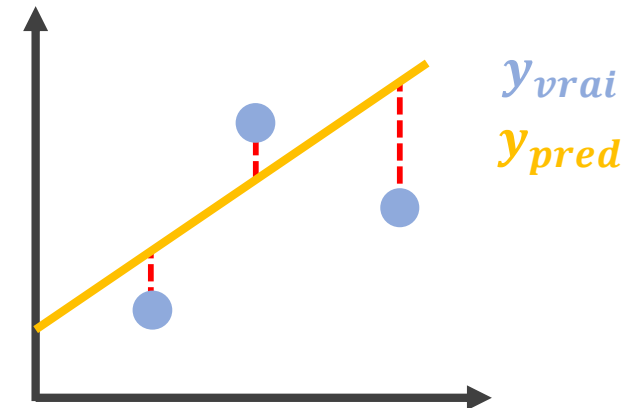
MAE
MSE
RMSE

$$MAE = \frac{1}{m} \sum |y_{vrai} - y_{pred}|$$

$$|1 - 1| = 0$$

$$MSE = \frac{1}{m} \sum (y_{vrai} - y_{pred})^2$$

$$(1 - 1)^2 = 0$$



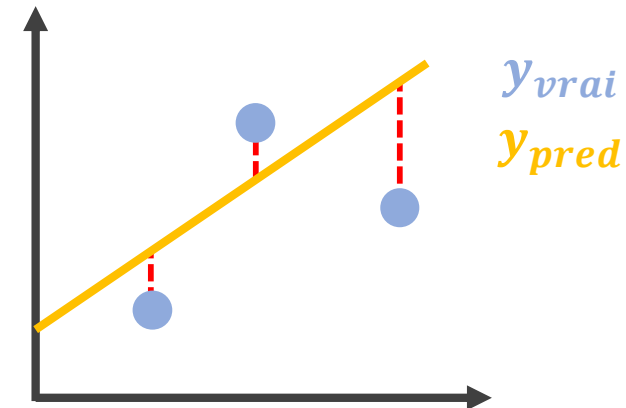
MAE
MSE
RMSE

$$MAE = \frac{1}{m} \sum |y_{vrai} - y_{pred}|$$

$$|1 - 2| = 1$$

$$MSE = \frac{1}{m} \sum (y_{vrai} - y_{pred})^2$$

$$(1 - 2)^2 = 1$$



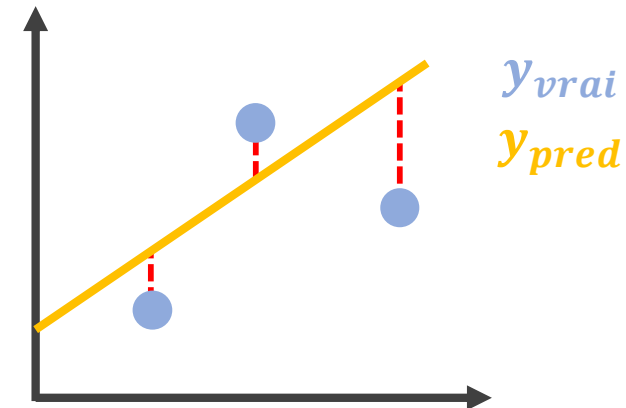
MAE
MSE
RMSE

$$MAE = \frac{1}{m} \sum |y_{vrai} - y_{pred}|$$

$$|1 - 3| = 2$$

$$MSE = \frac{1}{m} \sum (y_{vrai} - y_{pred})^2$$

$$(1 - 3)^2 = 4$$



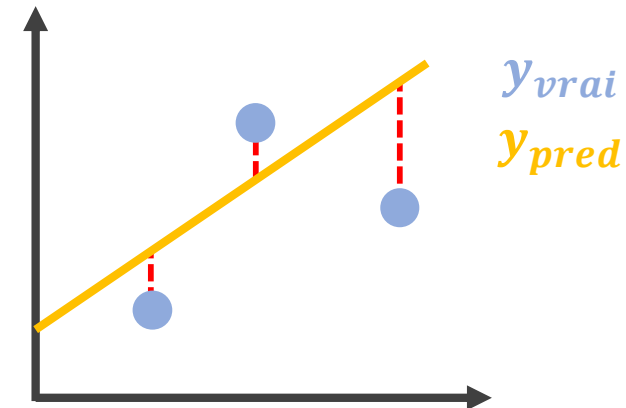
MAE
MSE
RMSE

$$MAE = \frac{1}{m} \sum |y_{vrai} - y_{pred}|$$

$$|1 - 4| = 3$$

$$MSE = \frac{1}{m} \sum (y_{vrai} - y_{pred})^2$$

$$(1 - 4)^2 = 9$$



$$\sum (erreurs)^2 \neq \left(\sum erreurs \right)^2$$

$$MSE \neq (MAE)^2$$

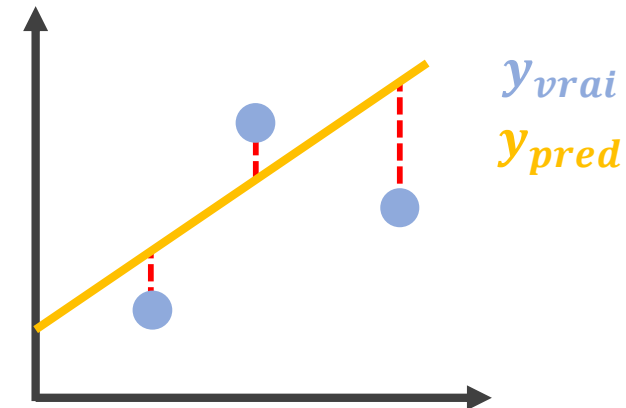
MAE
MSE
RMSE

$$MAE = \frac{1}{m} \sum |y_{vrai} - y_{pred}|$$

$$\frac{4 + 0}{2} = 2$$

$$MSE = \frac{1}{m} \sum (y_{vrai} - y_{pred})^2$$

$$\frac{4^2 + 0^2}{2} = 8$$



MAE
MSE
RMSE

$$MAE = \frac{1}{m} \sum |y_{vrai} - y_{pred}|$$

$$\frac{4 + 0}{2} = 2$$

$$MSE = \frac{1}{m} \sum (y_{vrai} - y_{pred})^2$$

$$\frac{4^2 + 0^2}{2} = 8$$

$$RMSE = \sqrt{\frac{1}{m} \sum (y_{vrai} - y_{pred})^2}$$

$$\sqrt{\frac{4^2 + 0^2}{2}} = 2.8$$

La RMSE retourne l'erreur à **son échelle initiale**, ce qui est plus compréhensible.

Mean Absolute Error

$$\frac{4 + 0}{2} = 2$$

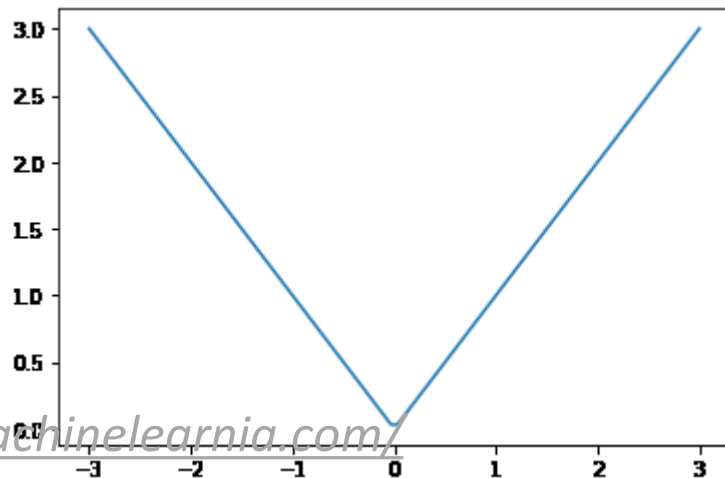
Root Mean Squared Error

$$\sqrt{\frac{4^2 + 0^2}{2}} = 2.8$$

Quand utiliser la **MSE** plutôt que la **MAE** ?

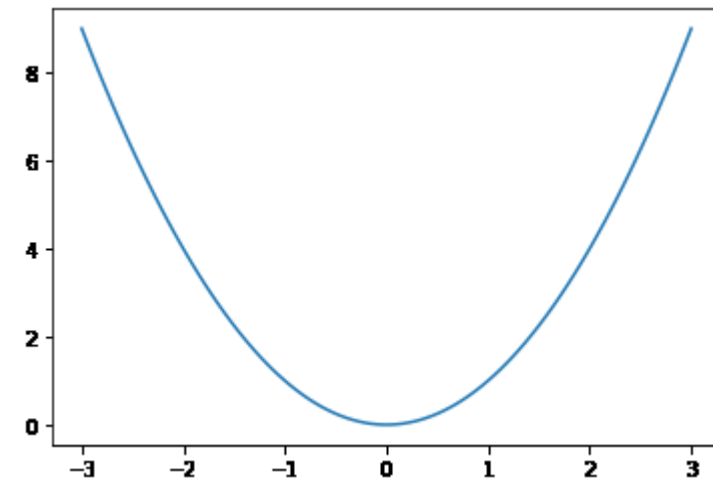
Mean Absolute Error

l'importance d'une erreur est **linéaire** avec son amplitude



Root Mean Squared Error

l'importance d'une erreur est **exponentielle** avec son amplitude





Estimateur de distance de freinage



	<i>Erreur 1</i>	<i>Erreur 2</i>	<i>MAE</i>	<i>MSE</i>
→ Modèle A	10	0	5	
Modèle B	6	5	5.5	

Estimateur de distance de freinage



			<i>Erreur 2</i>	<i>MAE</i>	<i>MSE</i>
	Modèle A		0	5	
	Modèle B		5	5.5	

Est-ce vraiment une bonne chose de choisir le modèle A ? Il peut être **très dangereux**....

Estimateur de distance de freinage




	<i>Erreur 1</i>	<i>Erreur 2</i>	<i>MAE</i>	<i>MSE</i>
Modèle A	10	0	5	7
➡ Modèle B	6	5	5.5	5.52

Estimateur de distance de freinage



Une erreur de 10 mètres, c'est **100 fois plus grave** qu'une erreur de 1 mètre.



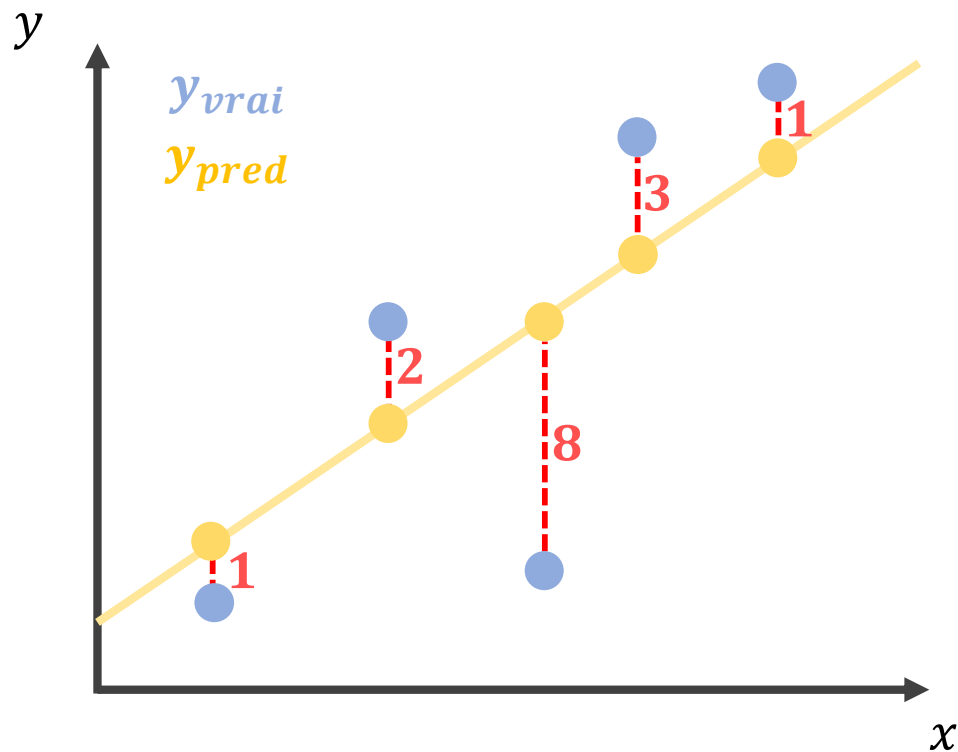
	<i>Erreur 2</i>	<i>MAE</i>	<i>MSE</i>
Modèle A	0	5	7
Modèle B	5	5.5	5.52

la MSE **pénalise beaucoup plus** les **grandes erreurs** que la MAE.

Quand utiliser la **MSE** plutôt que la **MAE** ?

MSE → vous accordez une **grande importance** aux **grandes erreurs**.

MAE → l'importance d'une erreur est **linéaire** avec son amplitude. Si le Dataset contient des valeurs **aberrantes** (*outliers*).



$erreurs = \{1, 1, 2, 3, 8\}$

↑
median

Median Absolute Error = 2

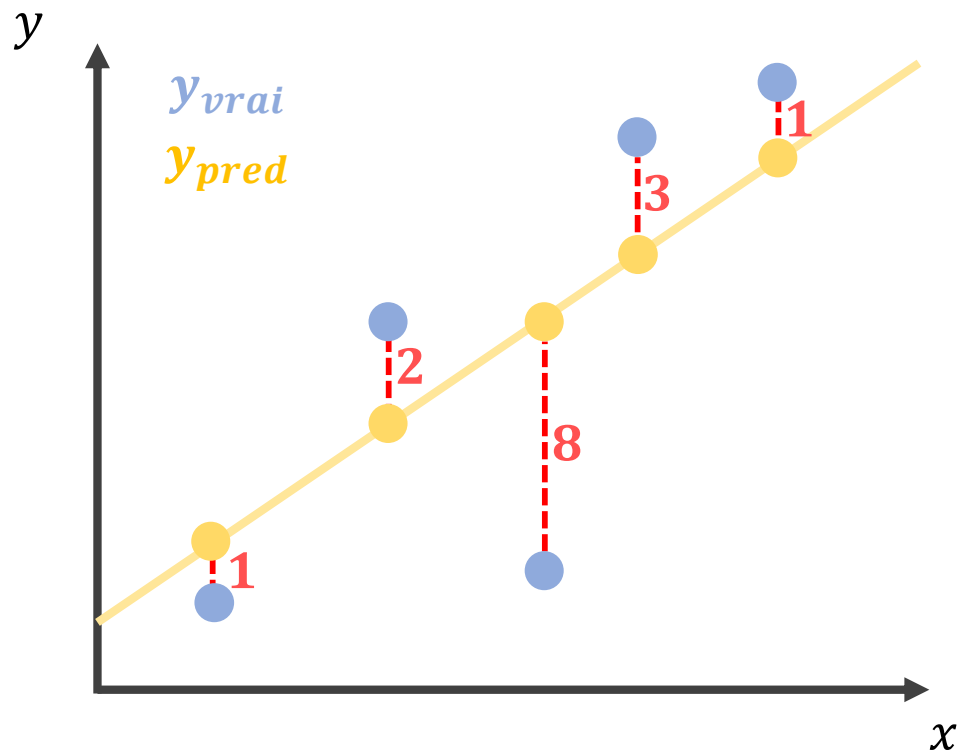
MEDIAN ABSOLUTE ERROR

Il existe aussi la *Median Absolute Error*.

Très peu sensible aux grandes erreurs.

$$MAE = \text{median} \{|y_{vrai} - y_{pred}|\}$$

Attention quand vous l'utilisez....



$erreurs = \{1, 1, 2, 3, 8\}$

↑
median

Median Absolute Error = 2

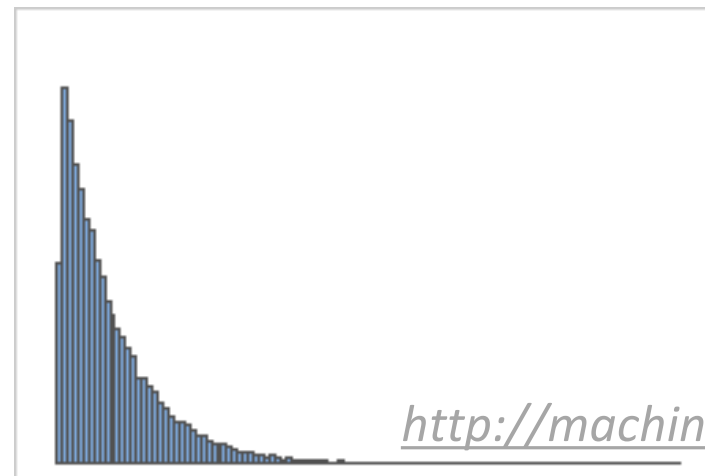
AU FINAL...

MAE, MSE, RMSE, Médiane...

Il semble que chaque mesure est des **avantages** et des **inconvénients**... Du coup, laquelle choisir ?

→ **Utilisez les toutes !**

Vous récolterez ainsi beaucoup plus **d'information** !



<http://machinelearnia.com/>

COEFFICIENT DE DETERMINATION R^2

$$R^2 = 1 - \frac{\sum (y_{vrai} - y_{pred})^2}{\sum (y_{vrai} - \overline{y_{vrai}})^2}$$

erreur

variance

Évalue la **performance** du modèle par rapport au **niveau de variation** présent dans les données.

Coefficient de détermination R2

$$R^2 = 1 - \frac{\sum (y_{vrai} - y_{pred})^2}{\sum (y_{vrai} - \overline{y_{vrai}})^2}$$

erreur quadratique

variance

moyenne des y_{vrai}

Évalue la **performance** du modèle par rapport au
niveau de variation présent dans les données.

Coefficient de détermination R2

$$\boxed{R^2_1} = 1 - \frac{\textit{erreurs}}{\textit{variance}_0}$$

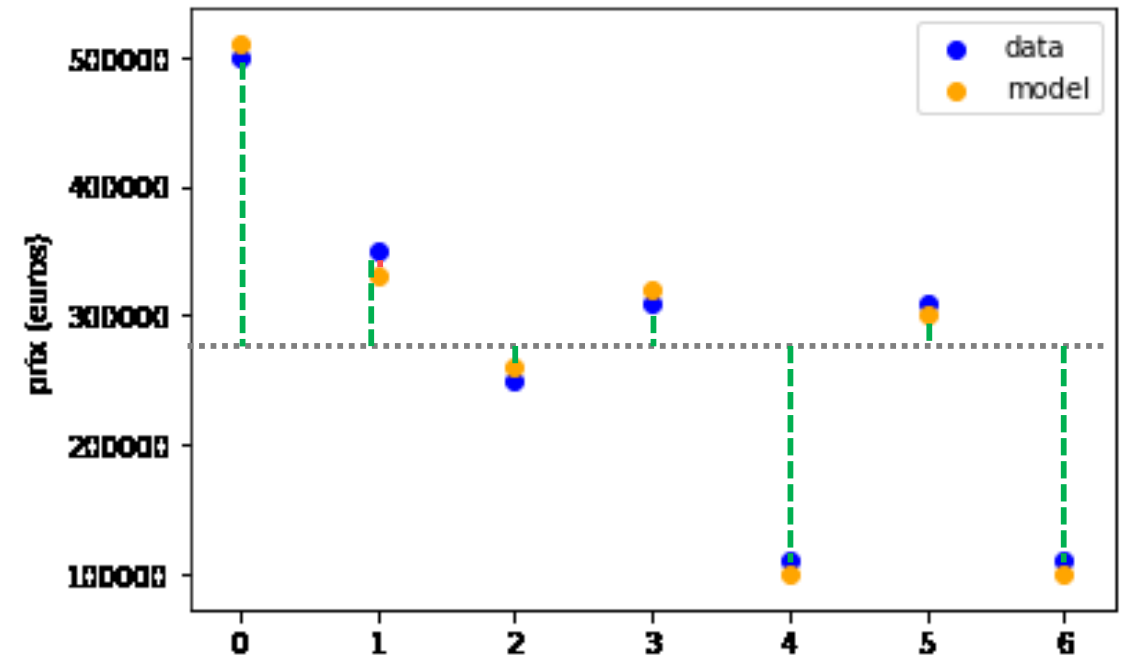
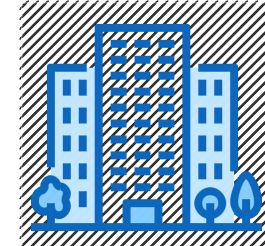
$$\textit{erreurs} \ll \textit{variance}$$

Coefficient de détermination R2

$$R^2 = 1 - \frac{\overbrace{\sum (y_{vrai} - y_{pred})^2}^{1000}}{\underbrace{\sum (y_{vrai} - \overline{y_{vrai}})^2}_{100000}}$$

$$R^2 = 1 - 0.01$$

$$R^2 = 0.99$$



Coefficient de détermination R2

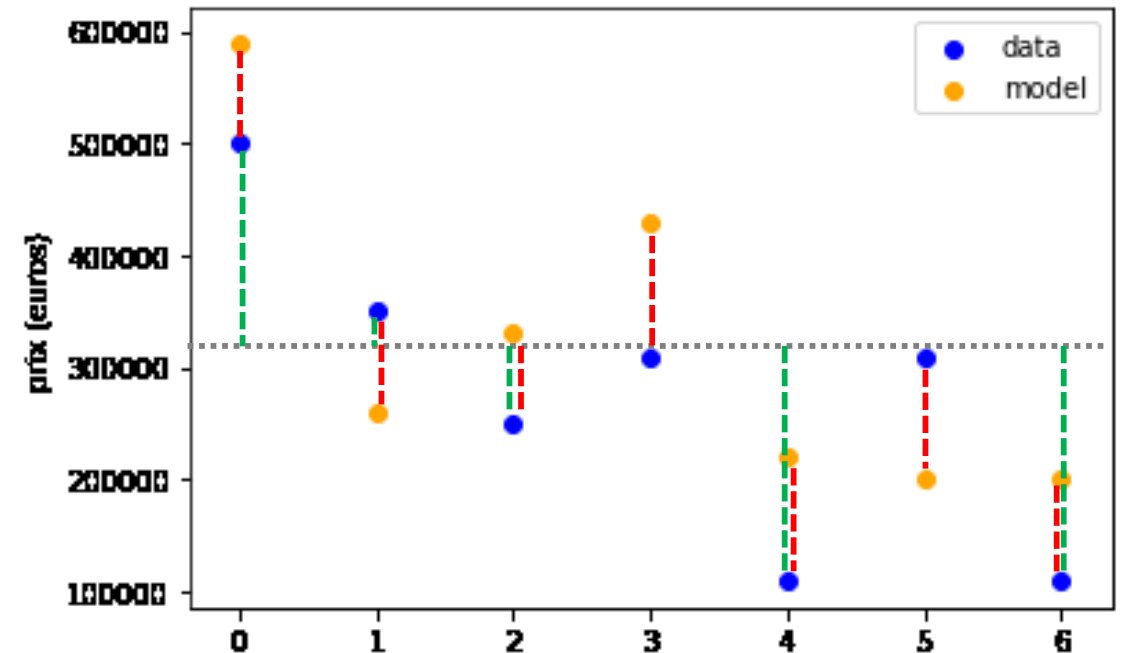
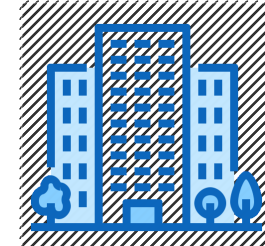
100000

$$R^2 = 1 - \frac{\sum (y_{vrai} - y_{pred})^2}{\sum (y_{vrai} - \overline{y_{vrai}})^2}$$

100000

$$R^2 = 1 - 1$$

$$R^2 = 0$$

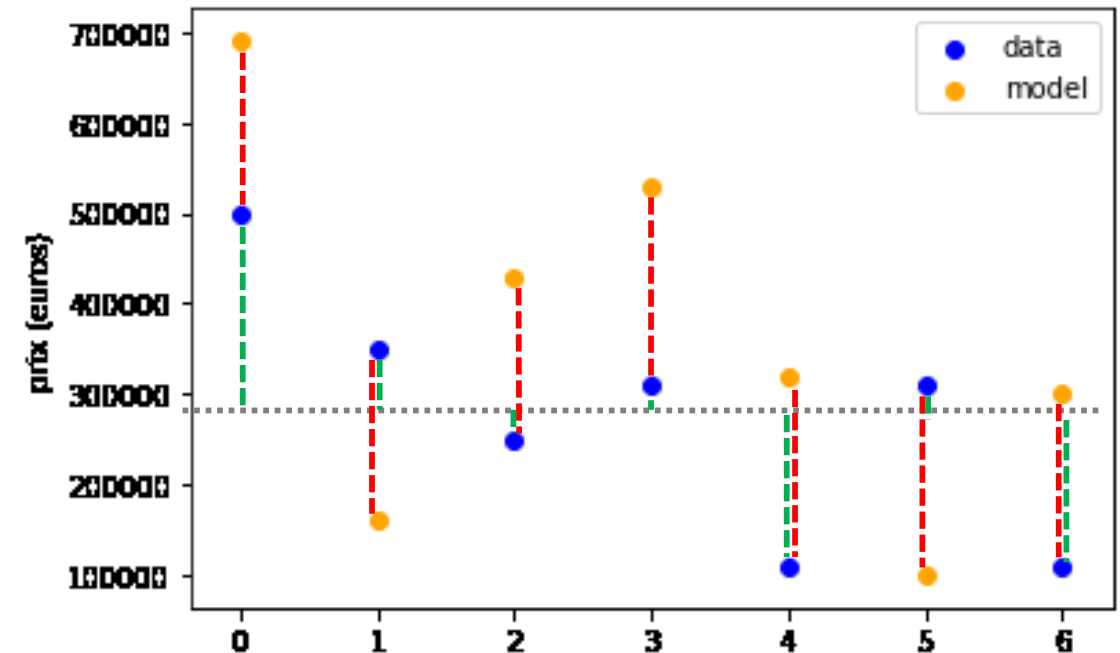
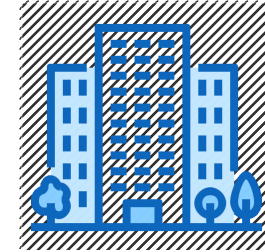


Coefficient de détermination R2

$$R^2 = 1 - \frac{\overbrace{\sum (y_{vrai} - y_{pred})^2}^{200000}}{\underbrace{\sum (y_{vrai} - \overline{y_{vrai}})^2}_{100000}}$$

$$R^2 = 1 - 2$$

$$R^2 = -1$$





Quelle formule peut sauver ce piéton ?

MAE

$$\frac{1}{m} \sum |y_{vrai} - y_{pred}|$$

MSE

$$\frac{1}{m} \sum (y_{vrai} - y_{pred})^2$$