Formation Machine Learning

Atelier Pratique AP-ML7

ALIASE

Atelier Pratique AP-ML7 : Exercice ML7.2.1

ALIASE
IA - SERVICE & FORMATION

Objectif : Prédire l'évolution du nombre de passagers d'une compagnie aérienne.

Dataset = nombre de passagers / mois depuis janvier 1949

Month	nb_passengers
1949-01	112
1949-02	118
1949-03	132
1949-04	129
1949-05	121
1949-06	135
1949-07	148
1949-08	148
1949-09	136
1949-10	119

1) Charger le fichier *'Airline_Passengers.csv*' dans un dataframe

- dataset = pd.read_csv(. . .)

- Seule la 2ème colonne ("*nb_passengers*") nous interesse

- Convertir le dataset en array

Atelier Pratique AP-ML7 : Exercice ML7.2.1

ALIASE
IA - SERVICE & FORMATION

2) Créer le dataset avec X = nb passengers au mois courant, Y = nb passengers au mois suivant

chaque label (Y) sera égal à la valeur à t+1 (c.a.d mois suivant)

- **Créer une variable *look_back = 1***

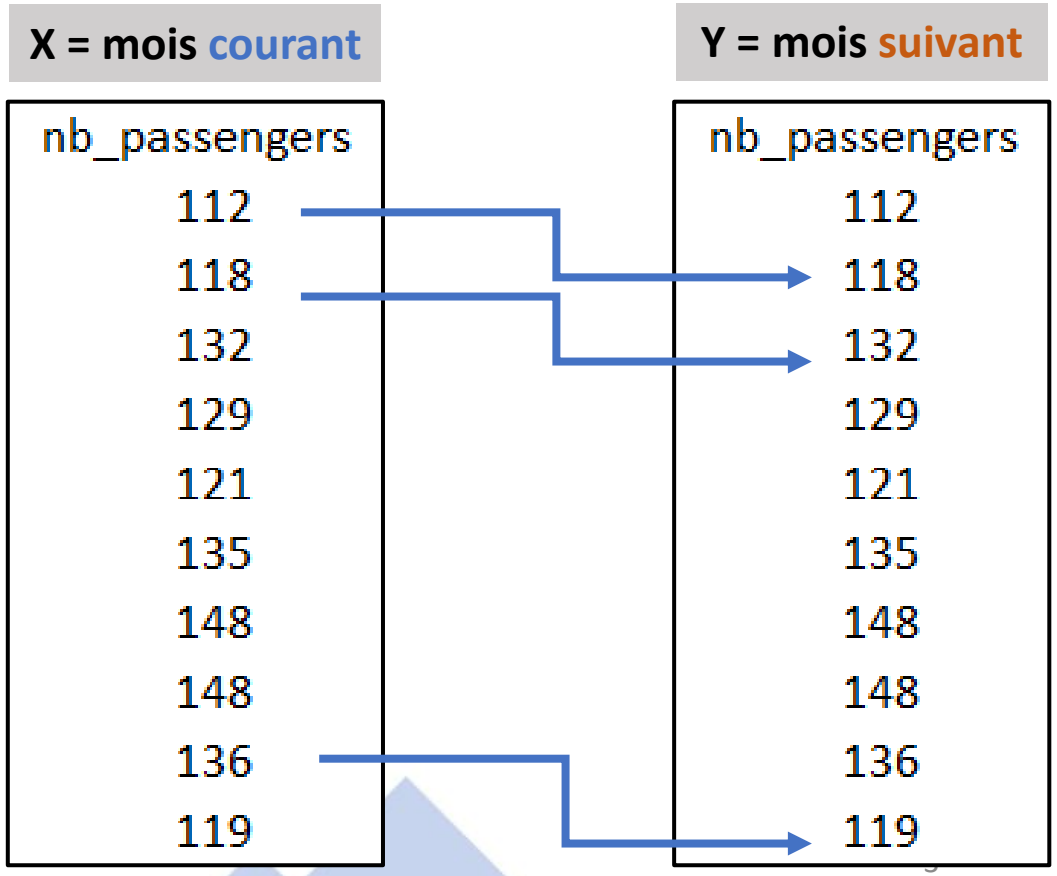
  signifie qu'on veut observer uniquement *un* mois précédent

- **Définir une fonction *create_dataset(dataset, look_back)***

  qui crée et retourne X et Y

- **Afficher X et Y avec un for … in zip(…)**

- **Convertir X et Y en tableaux de dimension 2**



Atelier Pratique AP-ML7 : Exercice ML7.2.1

ALIASE
IA - SERVICE & FORMATION

3) Normaliser X et Y avec le transformer *MinMaxScaler()*

4) Transformer X en dimension 3, car pour utiliser X dans un RNN, il faut qu'il soit de dimension 3

L'input (X) doit avoir le format *(nb samples, time steps, features)*

- **nb_samples = nb total de données dans le dataset**
- **input_length = nb de données en entrée (input), c'est aussi = à *look_back***
- **nb_features = 1 , nb de colonnes dans l'entrée**

Atelier Pratique AP-ML7 : Exercice ML7.2.1

ALIASE
IA - SERVICE & FORMATION

5) Découper X en *Training Set* (70%) et *Testing Set* (30%)

6) Créer le modèle RNN avec 3 couches LSTM

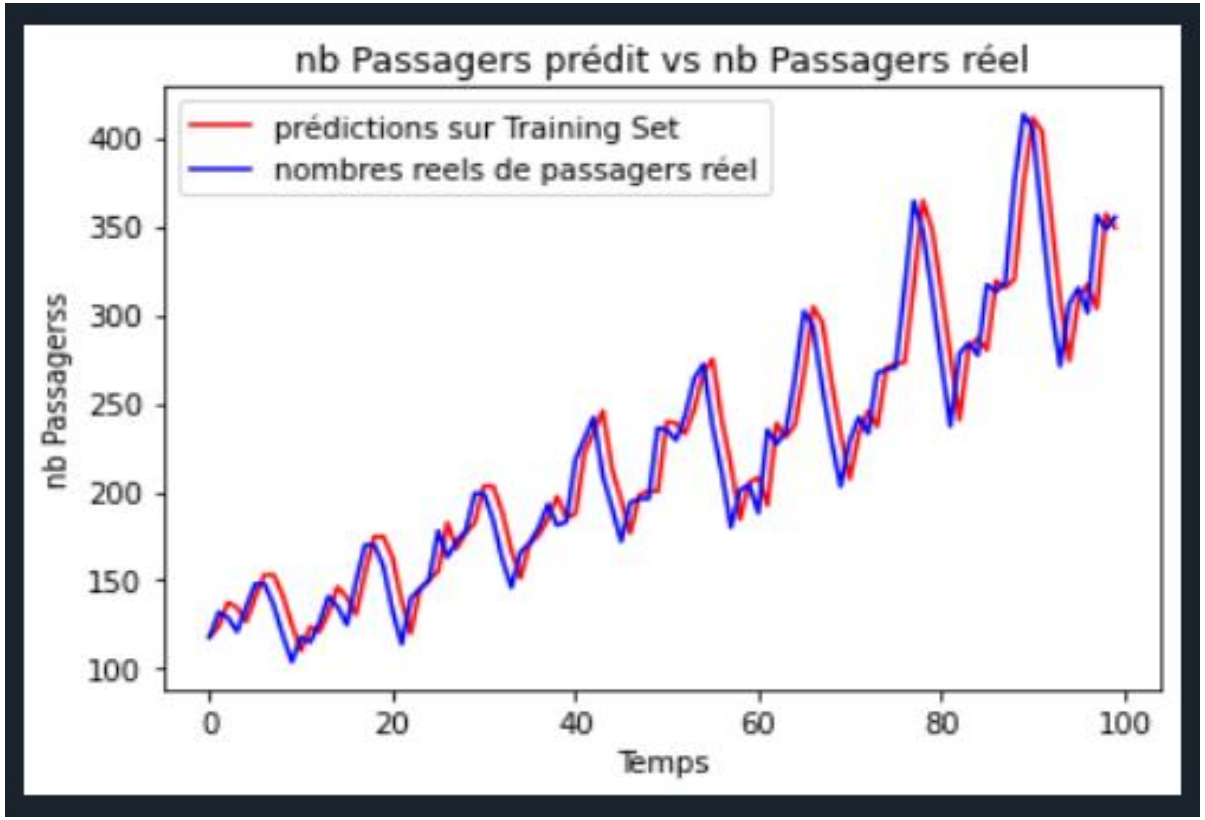
7) Compiler et entraîner le modèle

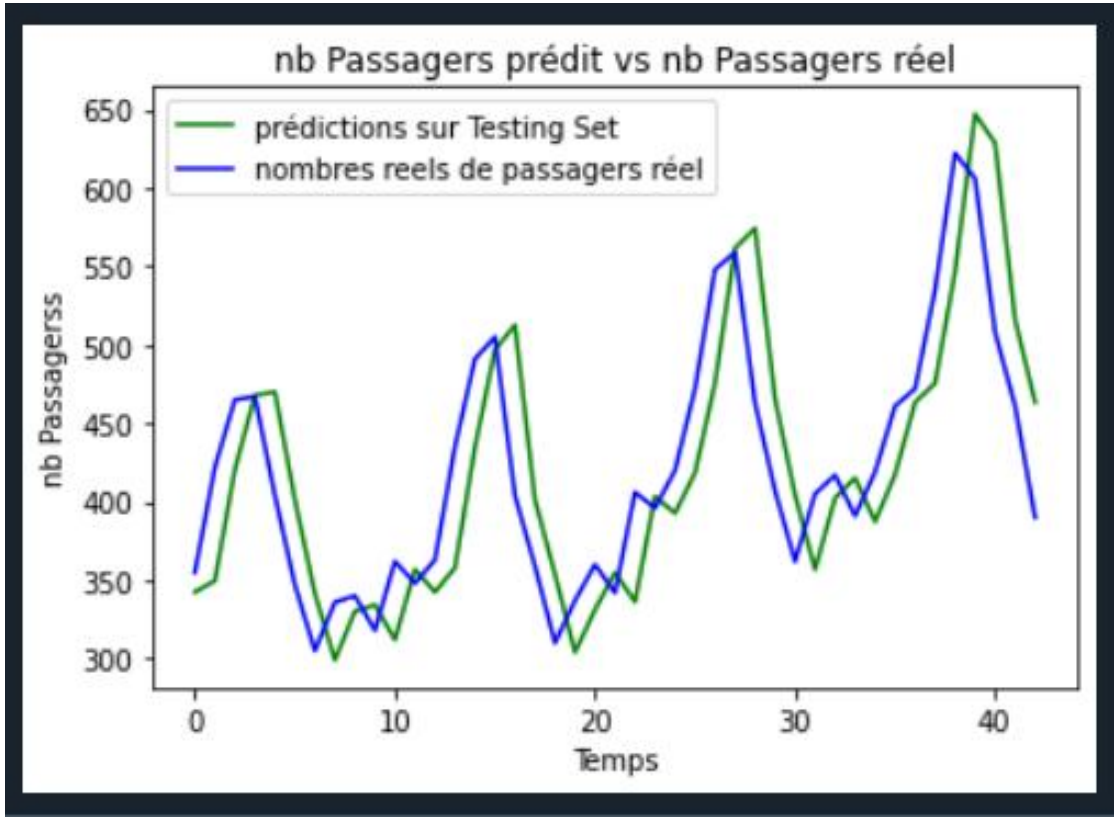
8) Faire des prédictions sur le Training Set et sur le Testing Set

Atelier Pratique AP-ML7 : Exercice ML7.2.1

ALIASE
IA - SERVICE & FORMATION

9) Afficher deux graphes : prédictions sur le Training Set puis prédictions sur le Testing Set

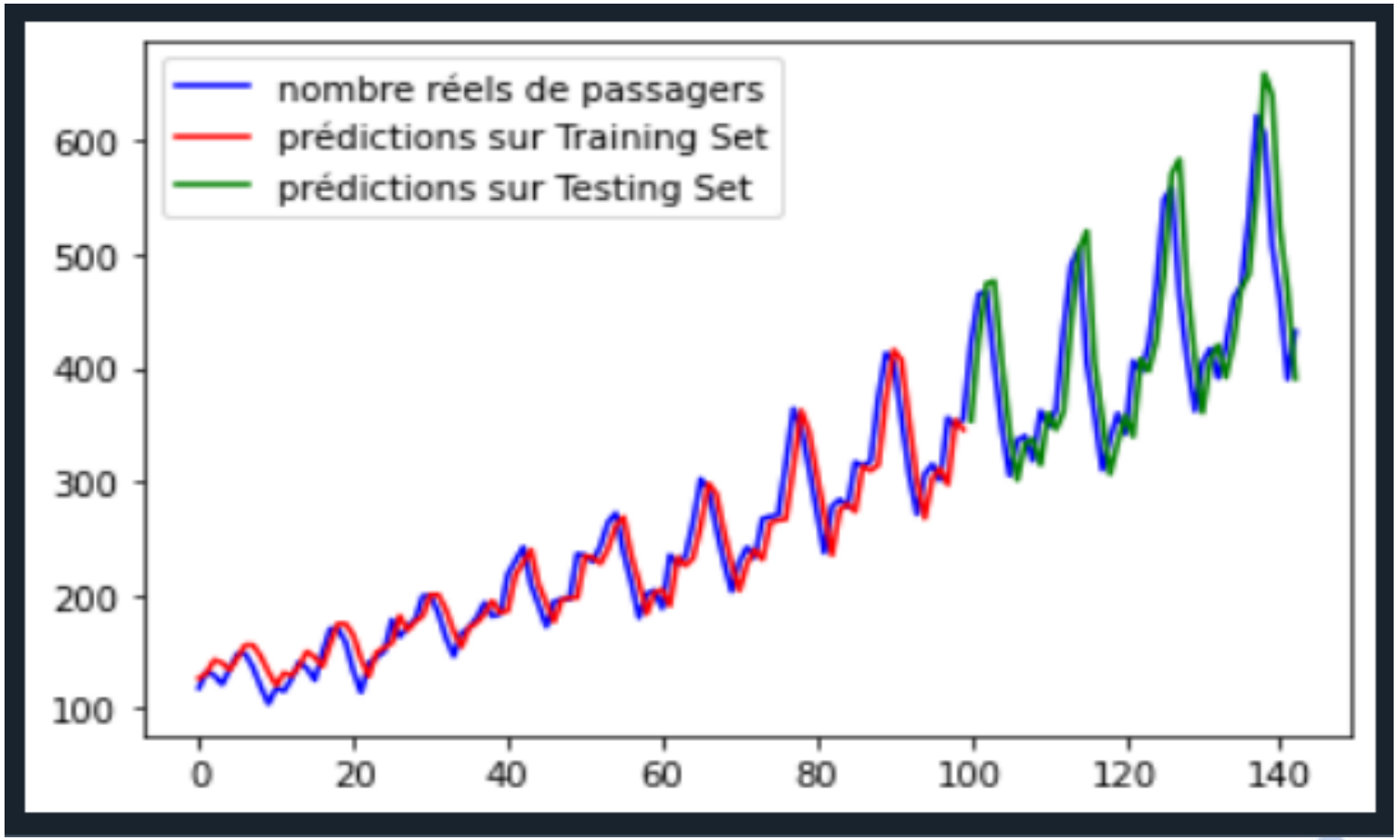




Atelier Pratique AP-ML7 : Exercice ML7.2.1

ALIAS
IA - SERVICE & FORMATION

9) Affichage graphique des prédictions sur TOUT le dataset (Training Set et Testing Set)



Solution → Voir MLExo7.2.1.py

Atelier Pratique AP-ML7 : Exercice ML7.4

ALIASE
IA - SERVICE & FORMATION

Objectif : Prédire l'évolution de la valeur d'une action en bourse

- *Entree : Liste des valeurs de l'action durant 60 jours*
- *Sortie : Prédiction de la valeur de l'action au 61ème jour*

Dataset = valeurs de l'action / jour **du 1er janvier 2011 au 31 décembre 2019**

Fichier : Bourse_dataset.xlsx , **Seule la 2ème colonne (« *Open* ») nous interesse**

Date	Open	High	Low	Close	Adj Close	Volume
1/3/2011	29.98000	30.05999947	29.87999916	30.05999947	28.47595024	1600
1/4/2011	30.30000	30.35000038	29.90999985	30.01000023	28.42859077	22200
1/5/2011	29.84000	29.89999962	29.79999924	29.82999992	28.25807381	14100
1/6/2011	29.82000	29.85000038	29.73999977	29.82999992	28.25807381	13300
1/7/2011	29.74000	29.73999977	29.67000008	29.67000008	28.10649872	1700
1/10/2011	29.47000	29.64999962	29.37000084	29.64999962	28.08755875	33200

Atelier Pratique AP-ML7 : Exercice ML7.4

ALIASE
IA - SERVICE & FORMATION

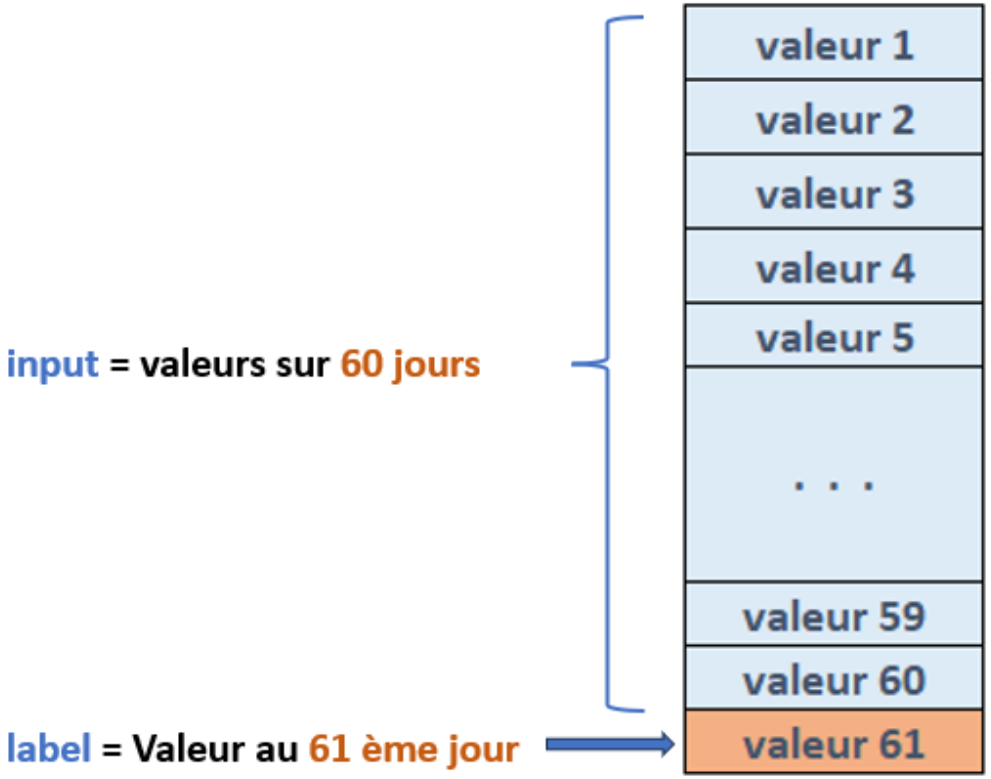
Créer le dataset avec

- **input = Liste des valeurs de l'action durant 60 jours**
- **label = valeur au 61ème jour**

Utiliser la variable look_back pour désigner

le nombre n de jours à observer avant de prédire

la valeur de l'action au (n+1)ème jour



Atelier Pratique AP-ML7 : Exercice ML7.4

ALIASE
IA - SERVICE & FORMATION

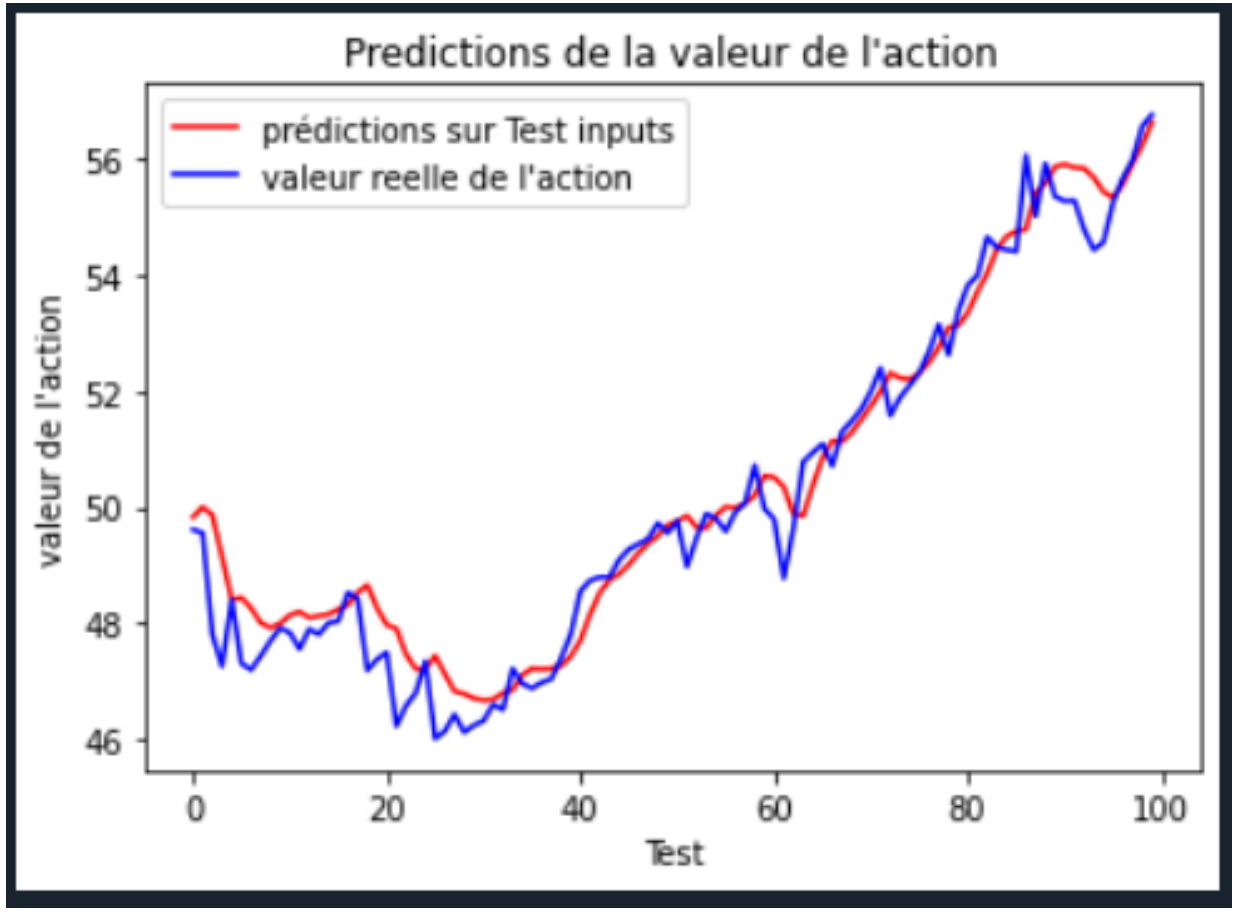
Affichage graphique des prédictions

sur les données du Testing Set

```
test_pred  = model.predict(x_test)

y_test = y_test.reshape( y_test.shape[0], 1 )

# Dénormaliser
test_pred   = scaler.inverse_transform(test_pred)
y_test      = scaler.inverse_transform(y_test)
```



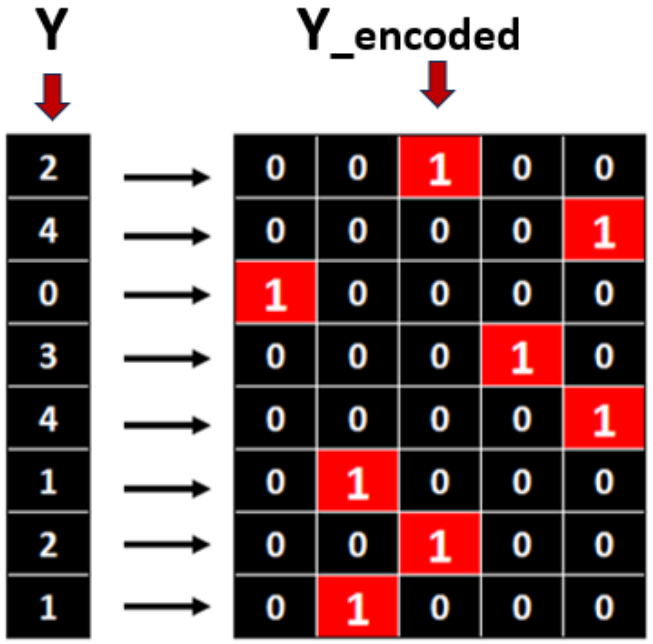
Solution → Voir MLExo7.4.py

Atelier Pratique AP-ML7 : Exercice ML7.3.7

ALIASE
IA - SERVICE & FORMATION

Refaire l'exercice 7.3.5 (Prédire le mot suivant) en utilisant le format **one-hot** des labels

Format 2 : one-hot



Chaque étiquette est représentée par un vecteur binaire où la classe correcte est représentée par 1 et les autres sont 0.

Solution → Voir MLExo7.3.7.py