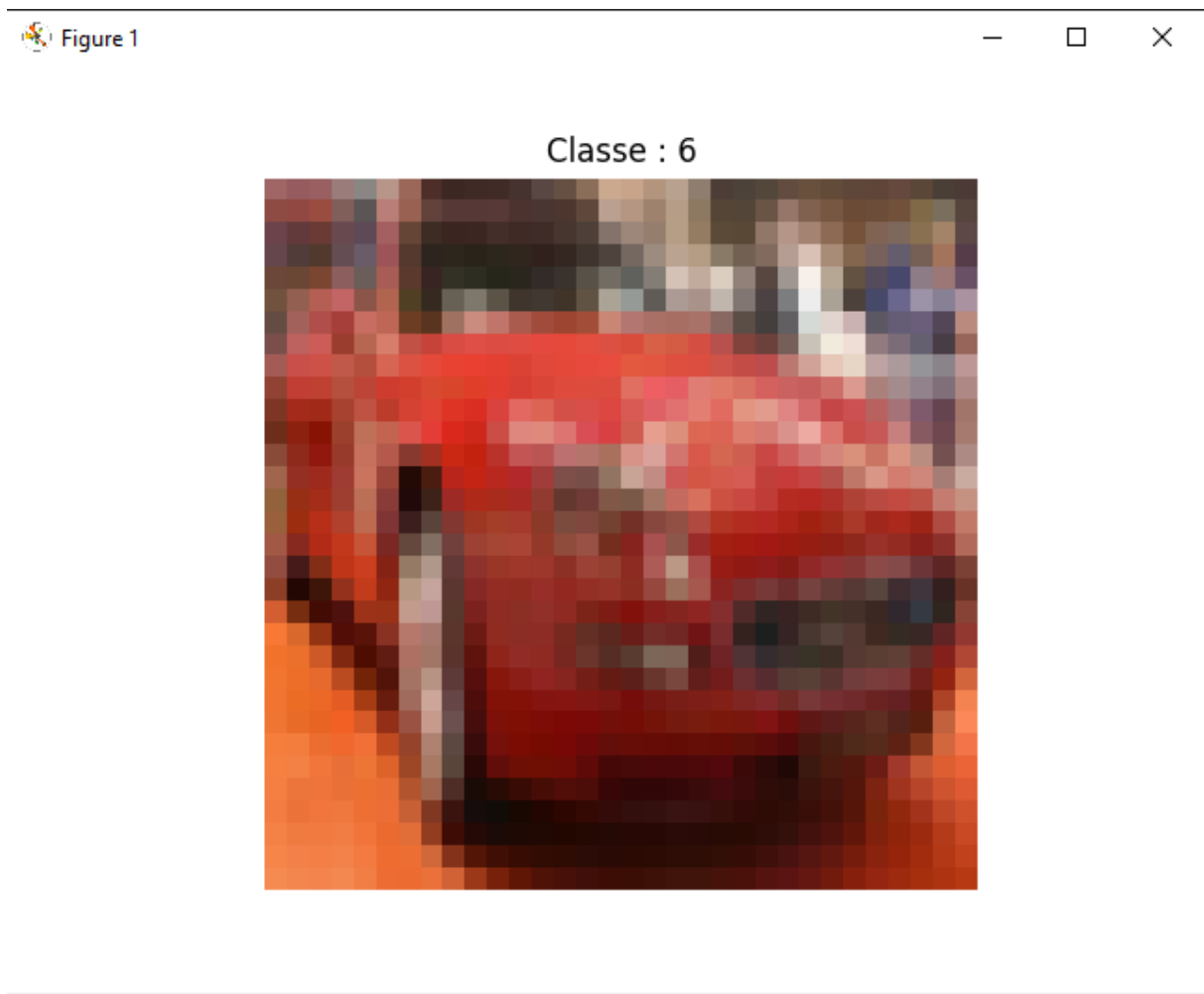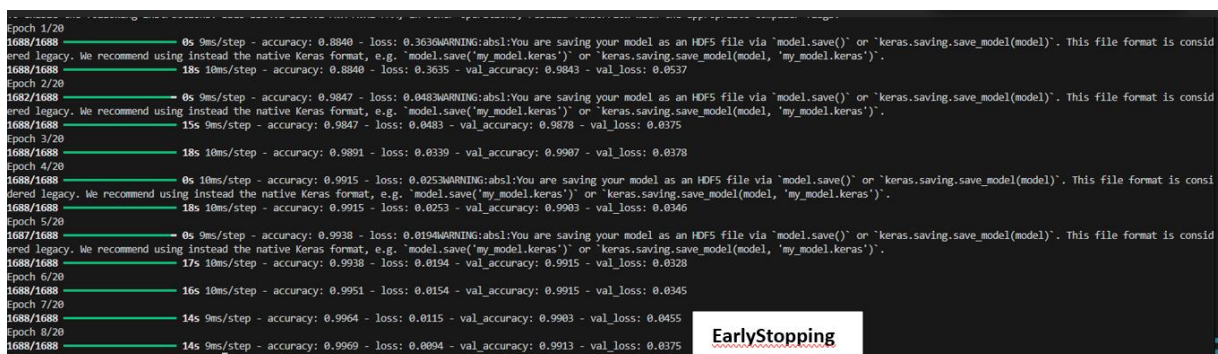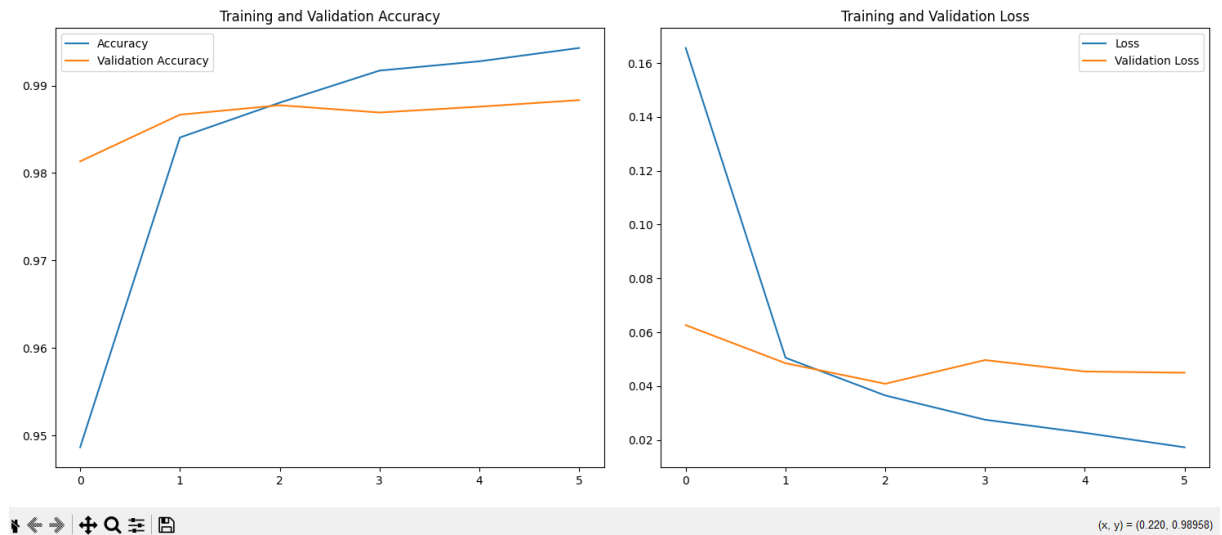Dossier data

Partie 1

Capture d'écran n°1

Capture d'écran n°2
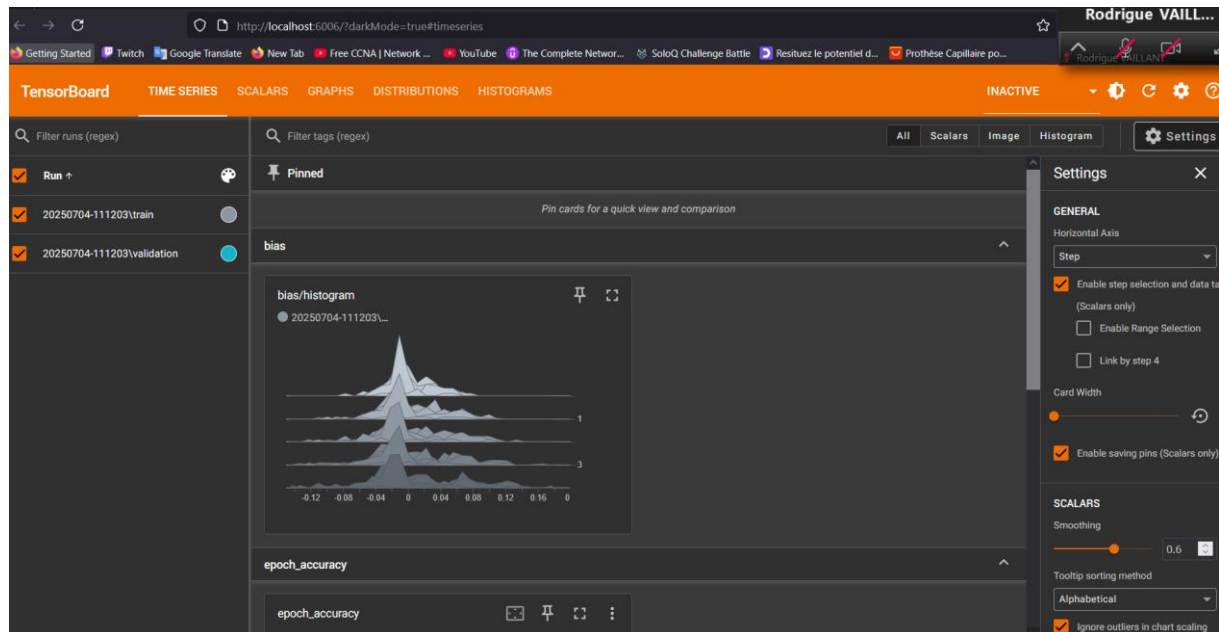
Capture d'écran n°3

Capture d'écran n°4


```
Matrice de confusion :
[[649  10  51  52  16   6   8   7 140  61]
 [ 24 659  14  24   2   5  11   2  51 208]
 [ 80   8 509 154  59  63  58  22  31  16]
 [ 18   5  56 664  45 102  56  15  20  19]
 [ 27   2 118 147 530  26  77  41  26   6]
 [  9   3  54 362  28 472  17  25  22   8]
 [  7   2  57 149  30   8 716   8  13  10]
 [ 19   3  40 127  89  89  10 586   3  34]
 [ 49  22  10  33   3   4   5   3 833  38]
 [ 17  58  10  38   4   6   8   8  57 794]]
```

Capture d'écran n°5

Capture d'écran n°6

Crash google colab & crash pc en local a cause de la ram qui n'est pas assez puissante

Partie 2

Capture d'écran n°7

```
PS C:\Users\Benjamin> & c:/Users/Benjamin/tf-env/Scripts/python.exe c:/Users/Benjamin/tf-exercice/venv/classification.py
Extrait des données (head) :
    Time       V1        V2        V3        V4        V5        V6        V7  ...       V23       V24       V25       V26       V27       V28  Amount
lass
0   0.0 -1.359807 -0.072781  2.536347  1.378155 -0.338321  0.462388  0.239599  ... -0.110474  0.066928  0.128539 -0.189115  0.133558 -0.021053  149.62
    0
1   0.0  1.191857  0.266151  0.166480  0.448154  0.060018 -0.082361 -0.078803  ...  0.101288 -0.339846  0.167170  0.125895 -0.008983  0.014724    2.69
    0
2   1.0 -1.358354 -1.340163  1.773209  0.379780 -0.503198  1.800499  0.791461  ...  0.909412 -0.689281 -0.327642 -0.139097 -0.055353 -0.059752  378.66
    0
3   1.0 -0.966272 -0.185226  1.792993 -0.863291 -0.010309  1.247203  0.237609  ... -0.190321 -1.175575  0.647376 -0.221929  0.062723  0.061458  123.50
    0
4   2.0 -1.158233  0.877737  1.548718  0.403034 -0.407193  0.095921  0.592941  ... -0.137458  0.141267 -0.206010  0.502292  0.219422  0.215153   69.99
    0

[5 rows x 31 columns]

Informations sur les données (info) :
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 284807 entries, 0 to 284806
Data columns (total 31 columns):
 #   Column  Non-Null Count   Dtype
---  ------  --------------   -----
 0   Time    284807 non-null  float64
 1   V1      284807 non-null  float64
 2   V2      284807 non-null  float64
 3   V3      284807 non-null  float64
 4   V4      284807 non-null  float64
 5   V5      284807 non-null  float64
 6   V6      284807 non-null  float64
 7   V7      284807 non-null  float64
 8   V8      284807 non-null  float64
 9   V9      284807 non-null  float64
 10  V10     284807 non-null  float64
 11  V11     284807 non-null  float64
 12  V12     284807 non-null  float64
 13  V13     284807 non-null  float64
 14  V14     284807 non-null  float64
 15  V15     284807 non-null  float64
 16  V16     284807 non-null  float64
 17  V17     284807 non-null  float64
```

```
Statistiques descriptives (describe) :
                Time            V1            V2            V3            V4  ...           V26           V27           V28          Amount          Class
count  284807.000000  2.848070e+05  2.848070e+05  2.848070e+05  2.848070e+05  ...  2.848070e+05  2.848070e+05  2.848070e+05   284807.000000  284807.000000
mean    94813.859575  1.168375e-15  3.416908e-16 -1.379537e-15  2.074095e-15  ...  1.683437e-15 -3.660091e-16 -1.227390e-16       88.349619       0.001727
std     47488.145955  1.958696e+00  1.651309e+00  1.516255e+00  1.415869e+00  ...  4.822270e-01  4.036325e-01  3.300833e-01      250.120109       0.041527
min         0.000000 -5.640751e+01 -7.271573e+01 -4.832559e+01 -5.683171e+00  ... -2.604551e+00 -2.256568e+01 -1.543008e+01        0.000000       0.000000
25%     54201.500000 -9.203734e-01 -5.985499e-01 -8.903648e-01 -8.486401e-01  ... -3.269839e-01 -7.083953e-02 -5.295979e-02        5.600000       0.000000
50%     84692.000000  1.810880e-02  6.548556e-02  1.798463e-01 -1.984653e-02  ... -5.213911e-02  1.342146e-03  1.124383e-02       22.000000       0.000000
75%    139320.500000  1.315642e+00  8.037239e-01  1.027196e+00  7.433413e-01  ...  2.409522e-01  9.104512e-02  7.827995e-02       77.165000       0.000000
max    172792.000000  2.454930e+00  2.205773e+01  9.382558e+00  1.687534e+01  ...  3.517346e+00  3.161220e+01  3.384781e+01    25691.160000       1.000000

[8 rows x 31 columns]

Répartition des classes :
Class
0    284315
1       492
Name: count, dtype: int64

PS C:\Users\Benjamin>
```

Capture d'écran n°8

```
Modèle RandomForest entraîné avec class_weight='balanced'.

Matrice de confusion :
[[56861     3]
 [   25    73]]

Rapport de classification :
              precision    recall  f1-score   support

   Non-fraude       1.00      1.00      1.00     56864
       Fraude       0.96      0.74      0.84        98

     accuracy                           1.00     56962
    macro avg       0.98      0.87      0.92     56962
 weighted avg       1.00      1.00      1.00     56962
```

Capture d'écran n°9

Matrice de confusion (heatmap)