## PM\_RNN2

## November 30, 2017

```
In [1]: import numpy as np
        import pandas as pd
        import os
        import matplotlib.pyplot as plt
In [2]: os.listdir('data')
Out[2]: ['Env_QH.csv',
         'micro_sud3.pkl',
         'AllPM_QH.csv',
         'micro_sud3_normalized.pkl',
         'AllNO2_QH.csv']
In [3]: df = pd.read_pickle('data/micro_sud3_normalized.pkl')
        df = df.reset_index()
        def split_dataframe(dataframe, percent):
            nb_rows = int(np.floor(percent * len(dataframe)))
            return dataframe[:nb_rows], dataframe[nb_rows:]
        def dataframe_to_xy(df):
            return (np.array(df[['PM_6182', 'PM_6179', 'PM_617B', 'PM25_6182', 'PM25_6179',\
                                 'PM25_617B', 'temp', 'rh',\
                                 'tgrad', 'pressure', 'pluvio']]),\
                    np.array(df['PM_ref']))
        def dataframe_to_xy(df, look_back):
            i = look_back
            while True:
                sequence = df.iloc[i - look_back:i]
                yield np.array(sequence[['PM_6182', 'PM_6179', 'PM_617B', 'PM25_6182', 'PM25_6
                                 'PM25_617B', 'temp', 'rh',\
                                 'tgrad', 'pressure', 'pluvio']]).reshape(look_back, 1, 11) , :
                i += 1
                if i == len(df):
                    i = look_back
        def dataframe_to_xy_test(df, look_back):
            X_test, y_test = [], []
            i = look_back
```

```
while i < len(df):
                sequence = df.iloc[i - look_back:i]
                X_test.append(np.array(sequence[['PM_6182', 'PM_6179', 'PM_617B', 'PM25_6182',
                                 'PM25_617B', 'temp', 'rh',\
                                 'tgrad', 'pressure', 'pluvio']]).reshape(look_back, 1, 11))
                y_test.append(np.array(sequence['PM_ref']))
                if i == len(df):
                    break
            return np.array(X_test), np.array(y_test)
        df_train, df_test = split_dataframe(df, 0.5)
        df_valid, df_test = split_dataframe(df_test, 0.5)
       X_train = dataframe_to_xy(df_train, 24)
        X_valid = dataframe_to_xy(df_valid, 24)
        X_test, y_test = dataframe_to_xy_test(df_test, 24)
In [4]: a = next(X_train)
In [5]: from keras.layers import SimpleRNN, Dense, LSTM, GRU
        from keras.models import Sequential
        from keras.callbacks import EarlyStopping
        def simple_rnn_model(nb_units, input_dim, loss='mean_squared_error', optimizer='adam')
           model = Sequential()
            model.add(SimpleRNN(nb_units, input_shape=input_dim))#input_dim=input_dim[1], inpu
           model.add(Dense(1, kernel_initializer='normal'))
           model.compile(loss=loss, optimizer=optimizer)
           model.summary()
            return model
        def lstm_model(nb_units, input_dim, loss='mean_squared_error', optimizer='adam'):
           model = Sequential()
           model.add(LSTM(nb_units, input_shape=input_dim))#input_dim=input_dim[1], input_len
           model.add(Dense(1, kernel_initializer='normal'))
           model.compile(loss=loss, optimizer=optimizer)
           model.summary()
            return model
        def gru_model(nb_units, input_dim, loss='mean_squared_error', optimizer='adam'):
           model = Sequential()
            model.add(GRU(nb_units, input_shape=(input_dim)))#input_dim=input_dim[1], input_le
           model.add(Dense(1, kernel_initializer='normal'))
           model.compile(loss=loss, optimizer=optimizer)
           model.summary()
            return model
```

```
Layer (type)
      Output Shape
           Param #
______
simple_rnn_1 (SimpleRNN)
      (None, 16)
            448
    (None, 1)
dense_1 (Dense)
           17
______
Total params: 465
Trainable params: 465
Non-trainable params: 0
Epoch 1/100
Epoch 2/100
Epoch 3/100
Epoch 4/100
Epoch 5/100
Epoch 6/100
Epoch 7/100
Epoch 8/100
Epoch 9/100
Epoch 10/100
Epoch 11/100
Epoch 12/100
Epoch 13/100
Epoch 14/100
Epoch 15/100
Epoch 16/100
```

```
Epoch 17/100
Epoch 18/100
Epoch 19/100
Epoch 20/100
Epoch 21/100
Epoch 22/100
Epoch 23/100
Epoch 24/100
Epoch 25/100
Epoch 26/100
Epoch 27/100
Epoch 28/100
Epoch 29/100
Epoch 30/100
Epoch 31/100
Epoch 32/100
Epoch 33/100
Epoch 34/100
Epoch 35/100
Epoch 36/100
Epoch 37/100
Epoch 38/100
Epoch 39/100
Epoch 40/100
```

```
Epoch 41/100
Epoch 42/100
Epoch 43/100
Epoch 44/100
Epoch 45/100
Epoch 46/100
Epoch 47/100
Epoch 48/100
Epoch 49/100
Epoch 50/100
Epoch 51/100
Epoch 52/100
Epoch 53/100
Epoch 54/100
Epoch 55/100
Epoch 56/100
Epoch 57/100
Epoch 58/100
Epoch 59/100
Epoch 60/100
Epoch 61/100
Epoch 62/100
Epoch 63/100
Epoch 64/100
```

```
Epoch 65/100
Epoch 66/100
Epoch 67/100
Epoch 68/100
Epoch 69/100
Epoch 70/100
Epoch 71/100
Epoch 72/100
Epoch 73/100
Epoch 74/100
Epoch 75/100
Epoch 76/100
Epoch 77/100
Epoch 78/100
Epoch 79/100
Epoch 80/100
Epoch 81/100
Epoch 82/100
Epoch 83/100
Epoch 84/100
Epoch 85/100
Epoch 86/100
Epoch 87/100
Epoch 88/100
```

```
Epoch 89/100
Epoch 90/100
Epoch 91/100
Epoch 92/100
Epoch 93/100
Epoch 94/100
Epoch 95/100
Epoch 96/100
Epoch 97/100
Epoch 98/100
Epoch 99/100
Epoch 100/100
In [7]: df = pd.read_pickle('data/micro_sud3_normalized.pkl')
   df = df.reset_index()
   def dataframe_to_xy_no_generator(df, look_back):
     X, y = [], []
     i = look_back
     while i < len(df):
      sequence = df.iloc[i - look_back:i]
      X.append(np.array(sequence[['PM_6182', 'PM_6179', 'PM_617B', 'PM25_6182', 'PM2
              'PM25_617B', 'temp', 'rh',\
              'tgrad', 'pressure', 'pluvio']]).reshape(look_back, 1, 11))
      y.append(np.array(df.iloc[i]['PM_ref']))
      i += 1
      if i == len(df):
        break
     X = np.array(X)
     y = np.array(y)
     X = X.reshape((X.shape[0], X.shape[1], X.shape[3]))
     return X, y
In [8]: X_train, y_train = dataframe_to_xy_no_generator(df_train, 24)
```

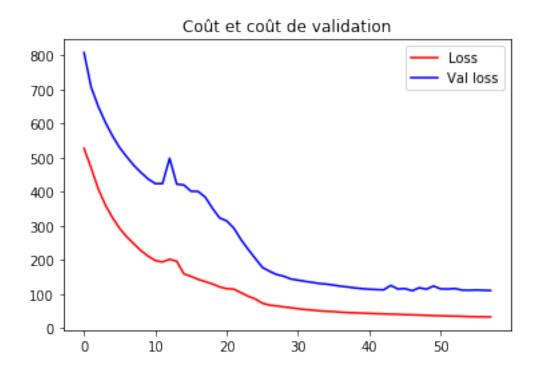
```
X_valid, y_valid = dataframe_to_xy_no_generator(df_valid, 24)
      X_test, y_test = dataframe_to_xy_no_generator(df_test, 24)
In [9]: from keras.layers import SimpleRNN, Dense, LSTM, GRU
      from keras.models import Sequential
      from keras.callbacks import EarlyStopping
      def simple_rnn_model(nb_units, input_dim, loss='mean_squared_error', optimizer='adam')
          model = Sequential()
          model.add(SimpleRNN(nb_units, input_shape=input_dim))#input_dim=input_dim[1], inpu
          model.add(Dense(1, kernel_initializer='normal'))
          model.compile(loss=loss, optimizer=optimizer)
          model.summary()
          return model
      def lstm_model(nb_units, input_dim, loss='mean_squared_error', optimizer='adam'):
          model = Sequential()
          model.add(LSTM(nb_units, input_shape=input_dim))#input_dim=input_dim[1], input_len
          model.add(Dense(1, kernel_initializer='normal'))
          model.compile(loss=loss, optimizer=optimizer)
          model.summary()
          return model
      def gru_model(nb_units, input_dim, loss='mean_squared_error', optimizer='adam'):
          model = Sequential()
          model.add(GRU(nb_units, input_shape=(input_dim)))#input_dim=input_dim[1], input_le
          model.add(Dense(1, kernel_initializer='normal'))
          model.compile(loss=loss, optimizer=optimizer)
          model.summary()
          return model
In [10]: model = simple_rnn_model(16, X_train.shape[1:])
Layer (type) Output Shape Param #
______
simple_rnn_2 (SimpleRNN) (None, 16)
-----
dense_2 (Dense) (None, 1)
                                             17
______
Total params: 465
Trainable params: 465
Non-trainable params: 0
______
In [11]: early_stopping = EarlyStopping(monitor='val_loss', verbose=1, mode='auto', patience=1
```

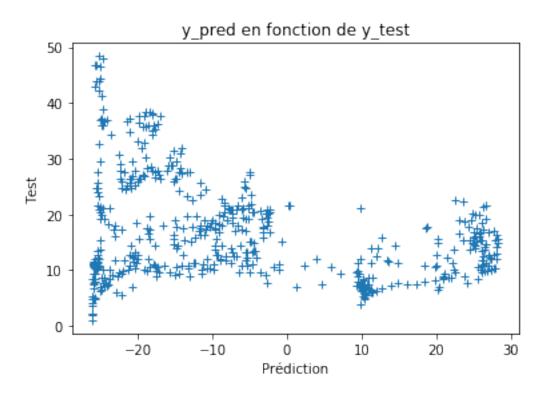
history = model.fit(X\_train, y\_train, batch\_size=16, epochs=5000, validation\_data=(X\_

```
Train on 1102 samples, validate on 539 samples
Epoch 1/5000
Epoch 2/5000
Epoch 3/5000
Epoch 4/5000
Epoch 5/5000
Epoch 6/5000
Epoch 7/5000
Epoch 8/5000
Epoch 9/5000
Epoch 10/5000
Epoch 11/5000
Epoch 12/5000
Epoch 13/5000
Epoch 14/5000
Epoch 15/5000
Epoch 16/5000
Epoch 17/5000
Epoch 18/5000
Epoch 19/5000
Epoch 20/5000
Epoch 21/5000
Epoch 22/5000
Epoch 23/5000
Epoch 24/5000
```

```
Epoch 25/5000
Epoch 26/5000
Epoch 27/5000
Epoch 28/5000
Epoch 29/5000
Epoch 30/5000
Epoch 31/5000
Epoch 32/5000
Epoch 33/5000
Epoch 34/5000
Epoch 35/5000
Epoch 36/5000
Epoch 37/5000
Epoch 38/5000
Epoch 39/5000
Epoch 40/5000
Epoch 41/5000
Epoch 42/5000
Epoch 43/5000
Epoch 44/5000
Epoch 45/5000
Epoch 46/5000
Epoch 47/5000
Epoch 48/5000
```

```
Epoch 49/5000
Epoch 50/5000
Epoch 51/5000
Epoch 52/5000
Epoch 53/5000
Epoch 54/5000
Epoch 55/5000
Epoch 56/5000
Epoch 57/5000
Epoch 58/5000
Epoch 00057: early stopping
In [12]: y_pred = model.predict(X_test)
   plt.title('Coût et coût de validation')
   line1,=plt.plot(history.history['loss'], label="Loss", linestyle='-', color='r')
   line2,=plt.plot(history.history['val_loss'], label="Val_loss", linestyle='-', color='
   first_legend = plt.legend(handles=[line1, line2], loc=1)
   plt.show()
   plt.title('y_pred en fonction de y_test')
   plt.plot(y_pred[:], y_test[:], '+')
   plt.ylabel('Test')
   plt.xlabel('Prédiction')
   plt.show()
```





In [15]: model = lstm\_model(16, X\_train.shape[1:])

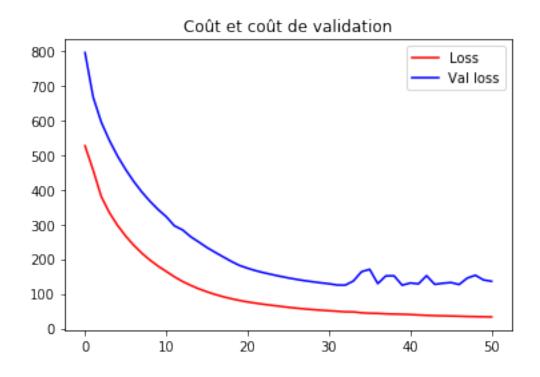
Total params: 1,809 Trainable params: 1,809 Non-trainable params: 0

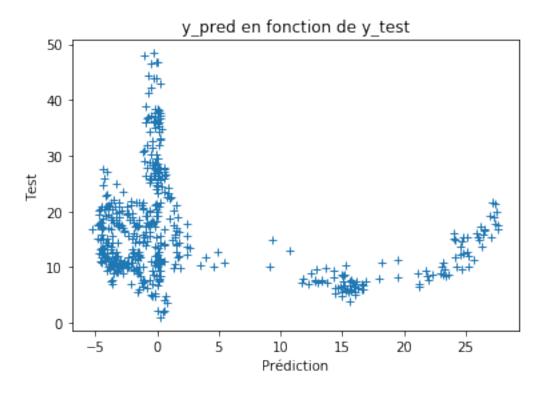
\_\_\_\_\_

```
Train on 1102 samples, validate on 539 samples
Epoch 1/5000
Epoch 2/5000
Epoch 3/5000
Epoch 4/5000
Epoch 5/5000
Epoch 6/5000
Epoch 7/5000
Epoch 8/5000
Epoch 9/5000
Epoch 10/5000
Epoch 11/5000
Epoch 12/5000
Epoch 13/5000
Epoch 14/5000
Epoch 15/5000
Epoch 16/5000
```

```
Epoch 17/5000
Epoch 18/5000
Epoch 19/5000
Epoch 20/5000
Epoch 21/5000
Epoch 22/5000
Epoch 23/5000
Epoch 24/5000
Epoch 25/5000
Epoch 26/5000
Epoch 27/5000
Epoch 28/5000
Epoch 29/5000
Epoch 30/5000
Epoch 31/5000
Epoch 32/5000
Epoch 33/5000
Epoch 34/5000
Epoch 35/5000
Epoch 36/5000
Epoch 37/5000
Epoch 38/5000
Epoch 39/5000
Epoch 40/5000
```

```
Epoch 41/5000
Epoch 42/5000
Epoch 43/5000
Epoch 44/5000
Epoch 45/5000
Epoch 46/5000
Epoch 47/5000
Epoch 48/5000
Epoch 49/5000
Epoch 50/5000
Epoch 51/5000
Epoch 00050: early stopping
In [17]: y_pred = model.predict(X_test)
   plt.title('Coût et coût de validation')
   line1,=plt.plot(history.history['loss'], label="Loss", linestyle='-', color='r')
   line2,=plt.plot(history.history['val_loss'], label="Val_loss", linestyle='-', color='
   first_legend = plt.legend(handles=[line1, line2], loc=1)
   plt.show()
   plt.title('y_pred en fonction de y_test')
   plt.plot(y_pred[:], y_test[:], '+')
   plt.ylabel('Test')
   plt.xlabel('Prédiction')
   plt.show()
```





In [18]: model = gru\_model(16, X\_train.shape[1:])

```
Layer (type) Output Shape Param #

gru_1 (GRU) (None, 16) 1344

dense_5 (Dense) (None, 1) 17

Total params: 1,361
```

Total params: 1,361 Trainable params: 1,361 Non-trainable params: 0

\_\_\_\_\_

```
Train on 1102 samples, validate on 539 samples
Epoch 1/5000
Epoch 2/5000
Epoch 3/5000
Epoch 4/5000
Epoch 5/5000
Epoch 6/5000
Epoch 7/5000
Epoch 8/5000
Epoch 9/5000
Epoch 10/5000
Epoch 11/5000
Epoch 12/5000
Epoch 13/5000
Epoch 14/5000
Epoch 15/5000
Epoch 16/5000
```

```
Epoch 17/5000
Epoch 18/5000
Epoch 19/5000
Epoch 20/5000
Epoch 21/5000
Epoch 22/5000
Epoch 23/5000
Epoch 24/5000
Epoch 25/5000
Epoch 26/5000
Epoch 27/5000
Epoch 28/5000
Epoch 29/5000
Epoch 30/5000
Epoch 31/5000
Epoch 32/5000
Epoch 33/5000
Epoch 34/5000
Epoch 35/5000
Epoch 36/5000
Epoch 37/5000
Epoch 38/5000
Epoch 39/5000
Epoch 40/5000
```

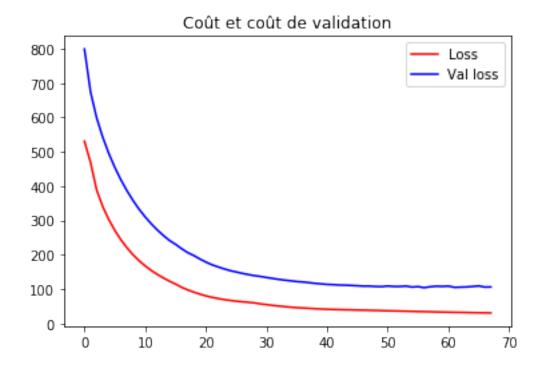
```
Epoch 41/5000
Epoch 42/5000
Epoch 43/5000
Epoch 44/5000
Epoch 45/5000
Epoch 46/5000
Epoch 47/5000
Epoch 48/5000
Epoch 49/5000
Epoch 50/5000
Epoch 51/5000
Epoch 52/5000
Epoch 53/5000
Epoch 54/5000
Epoch 55/5000
Epoch 56/5000
Epoch 57/5000
Epoch 58/5000
Epoch 59/5000
Epoch 60/5000
Epoch 61/5000
Epoch 62/5000
Epoch 63/5000
Epoch 64/5000
```

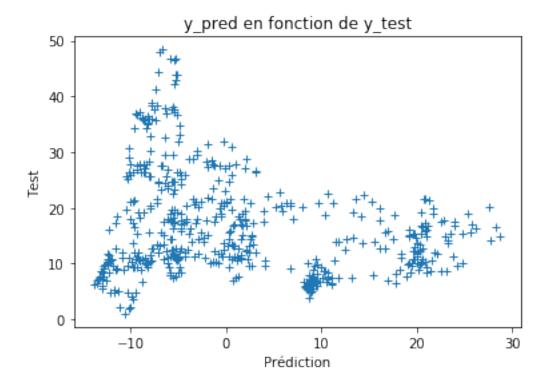
```
In [20]: y_pred = model.predict(X_test)
    plt.title('Coût et coût de validation')
    line1,=plt.plot(history.history['loss'], label="Loss", linestyle='-', color='r')
    line2,=plt.plot(history.history['val_loss'], label="Val loss", linestyle='-', color='rifirst_legend = plt.legend(handles=[line1, line2], loc=1)

plt.show()

plt.title('y_pred en fonction de y_test')

plt.plot(y_pred[:], y_test[:], '+')
    plt.ylabel('Test')
    plt.xlabel('Prédiction')
    plt.show()
```





In []: