Prédire l’avenir avec Python

PONDERING MY ORB RN

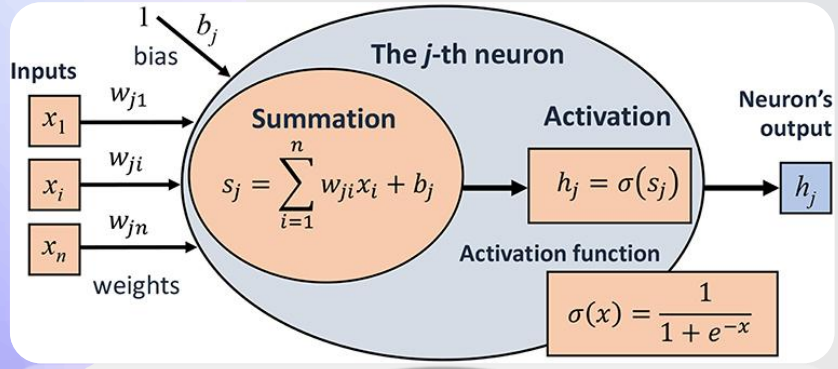
Sommaire

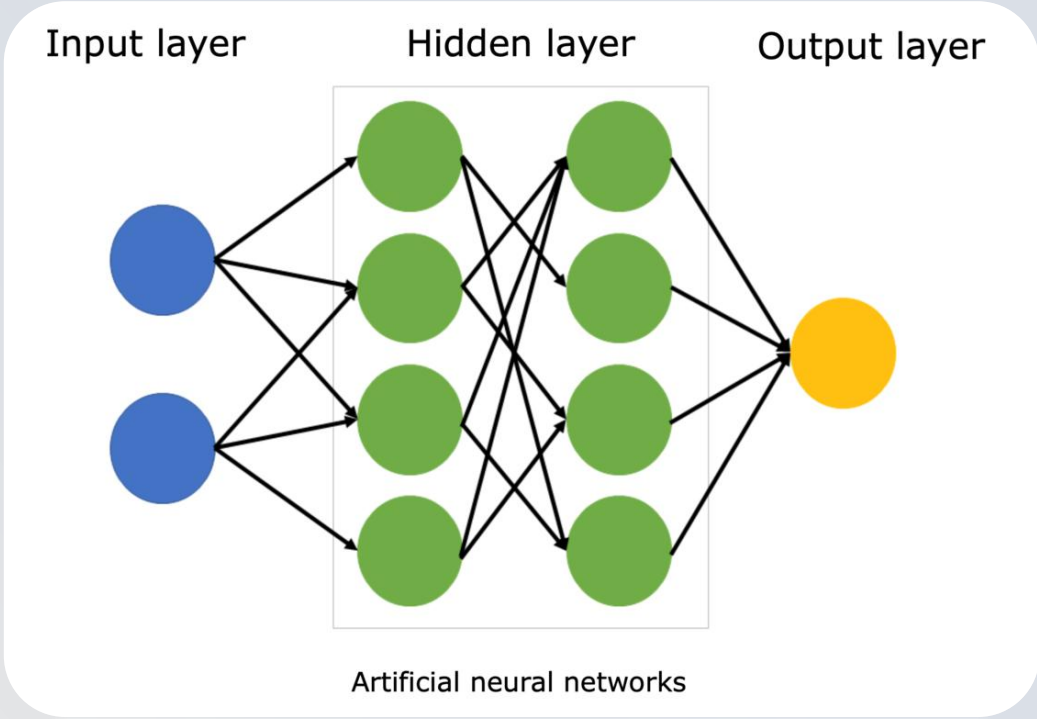
RÉSEAU DE NEURONES ???
LSTM ET TERMES TECHNIQUES
COMMENT ÇA MARCHE
GRIND LA DATA
ENTRAÎNER LE MODÈLE

Un cerveau dans votre PC

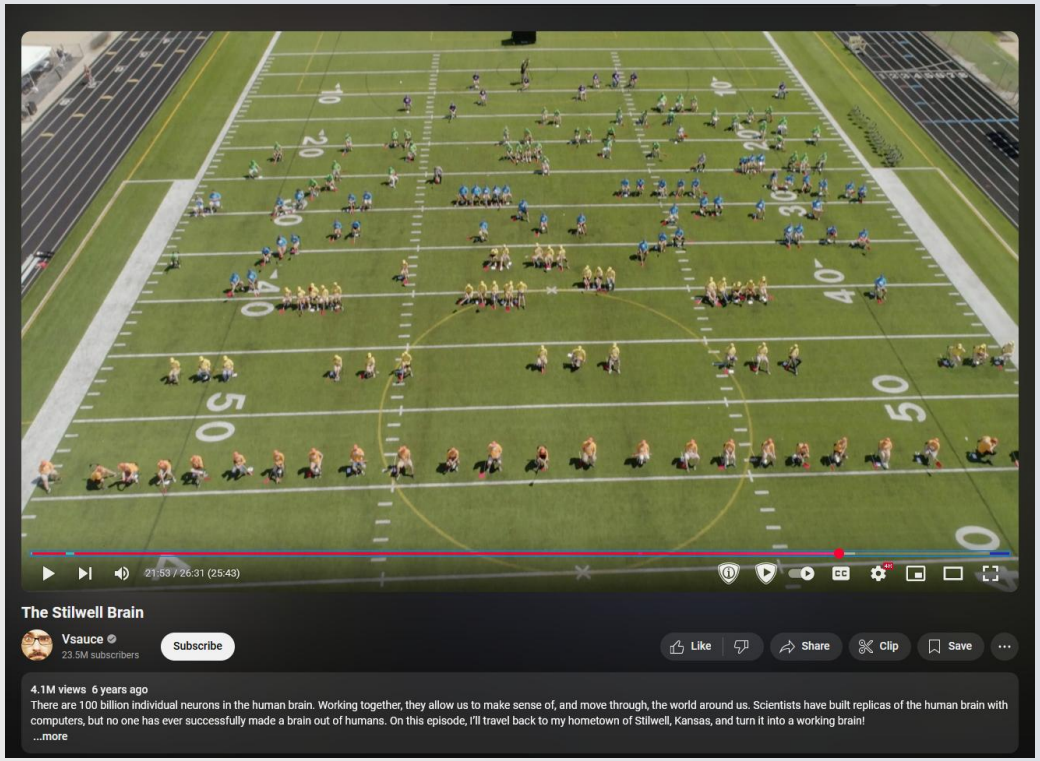
C’EST QUOI UN RÉSEAU DE NEURONES

C’est très simple (presque)





Une très bonne vidéo pour approfondir



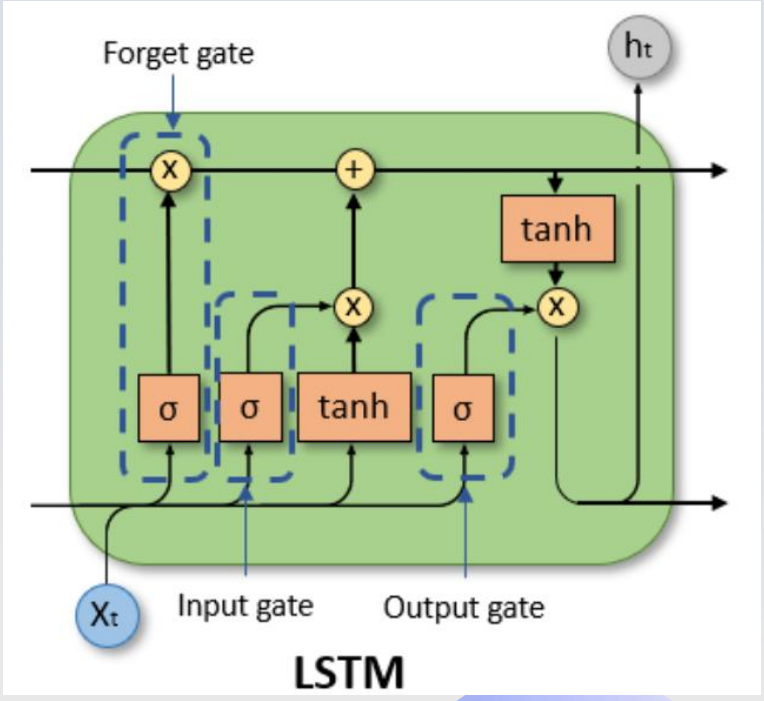
Vsauce le goat

Les réseaux LSTM

LONG SHORT TERM MEMORY

Une cellule particulière

En plus des entrées et sorties standard, une cellule LSTM ajoute des mécanismes de mémoire, notamment l'apprentissage et l'oubli



Mais comment ça marche ?

TRÈS BONNE QUESTION !

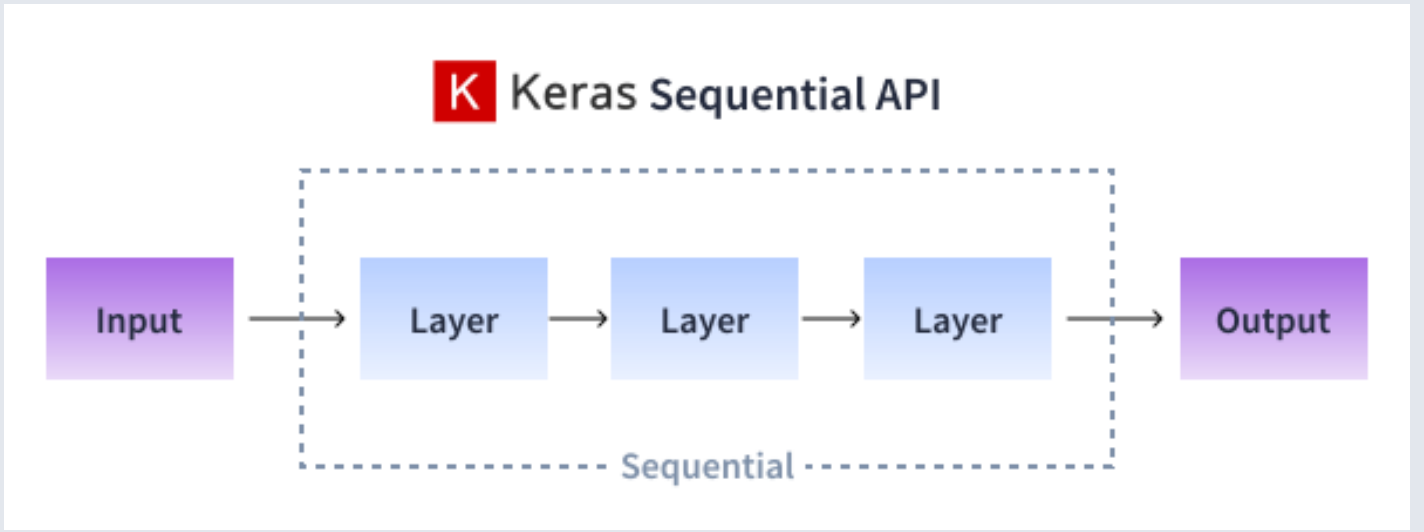
La magie Tensorflow

- Développé par Google depuis le début des années 2010
- Abstractions de haut niveau pour du machine learning
- Possibilité de “fabriquer” ses modèles brique par brique

Fabriquer des modèles ?

- Les deux couches LSTM traitent la séquence de données pour en extraire progressivement l'essentiel.
- La première LSTM gère la séquence complète et renvoie un ensemble de vecteurs, la seconde condense ensuite cela en un unique vecteur de taille 32.
- Enfin, la couche Dense transforme ce vecteur en un seul nombre (sortie du modèle).

```
model = tf.keras.Sequential([
    tf.keras.layers.LSTM(64, return_sequences=True, input_shape=(window_length, len(features))),
    tf.keras.layers.LSTM(32),
    tf.keras.layers.Dense(1)
])
```

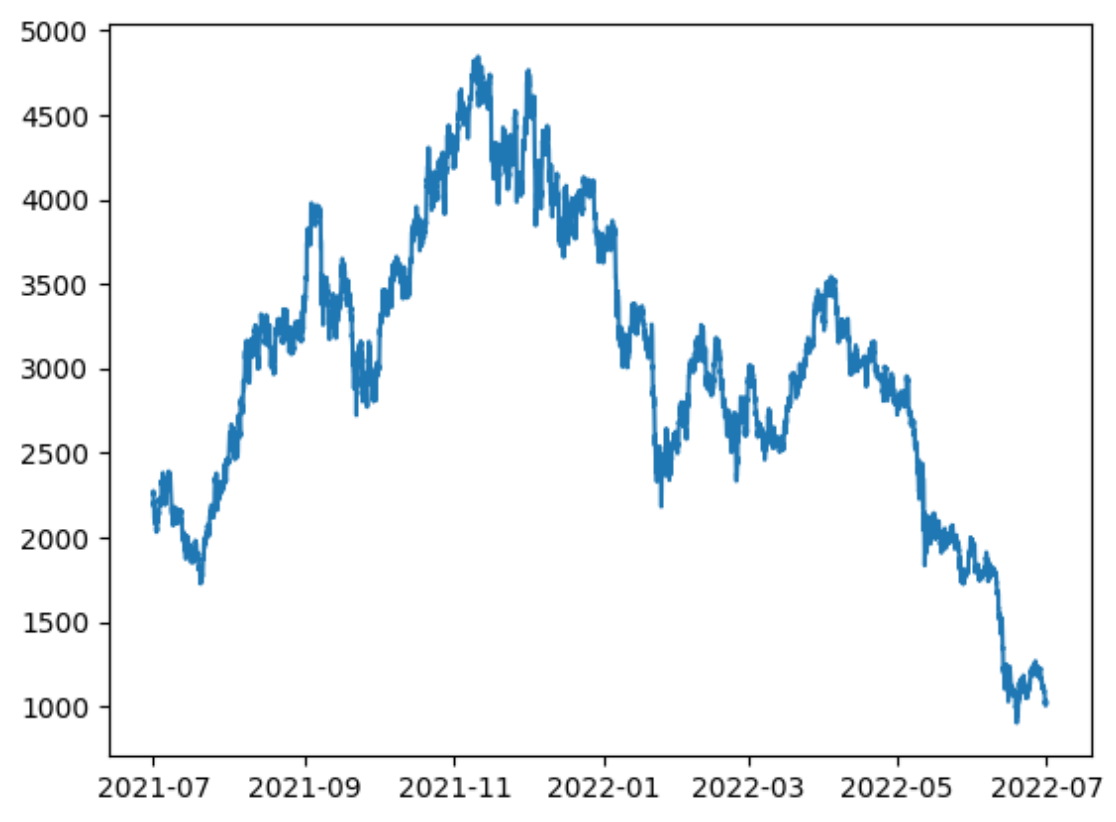


Les données

(important)

Ethereum Juillet 2021 – Juillet 2022

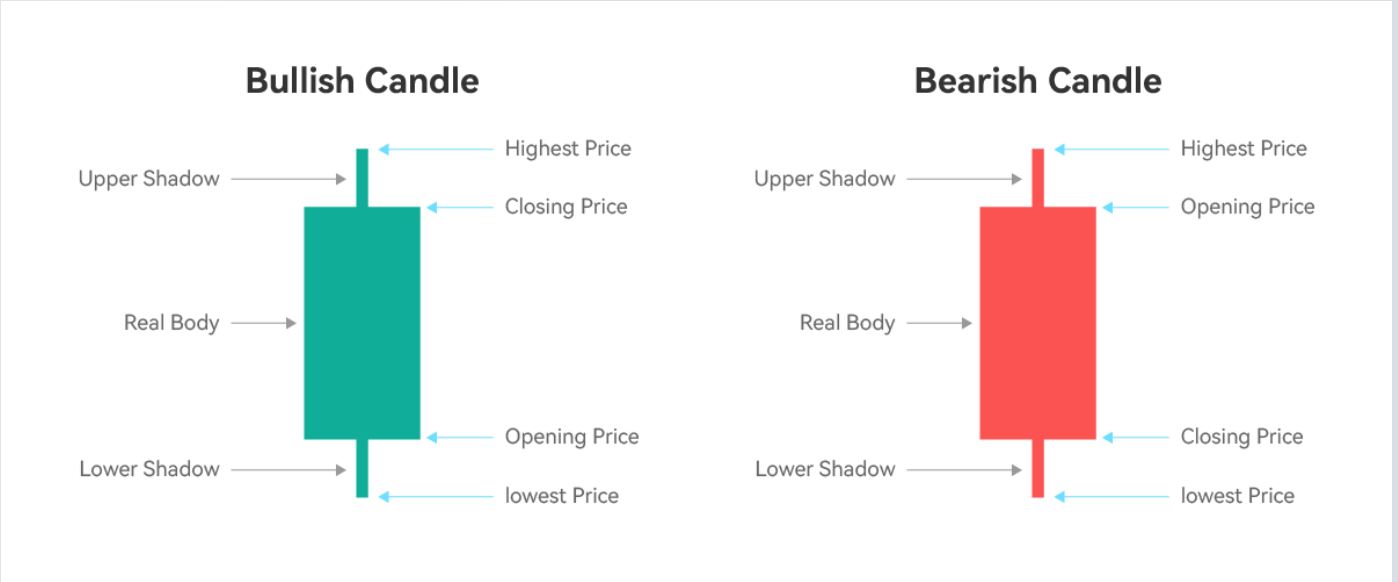
Pas zinzin pour la performance financière, mais parfait pour nous



Le trading pour les nuls

Bullish = ça monte

Bearish = tu perds tes sous mon pote



Pratique pour les humains, moins pour le code

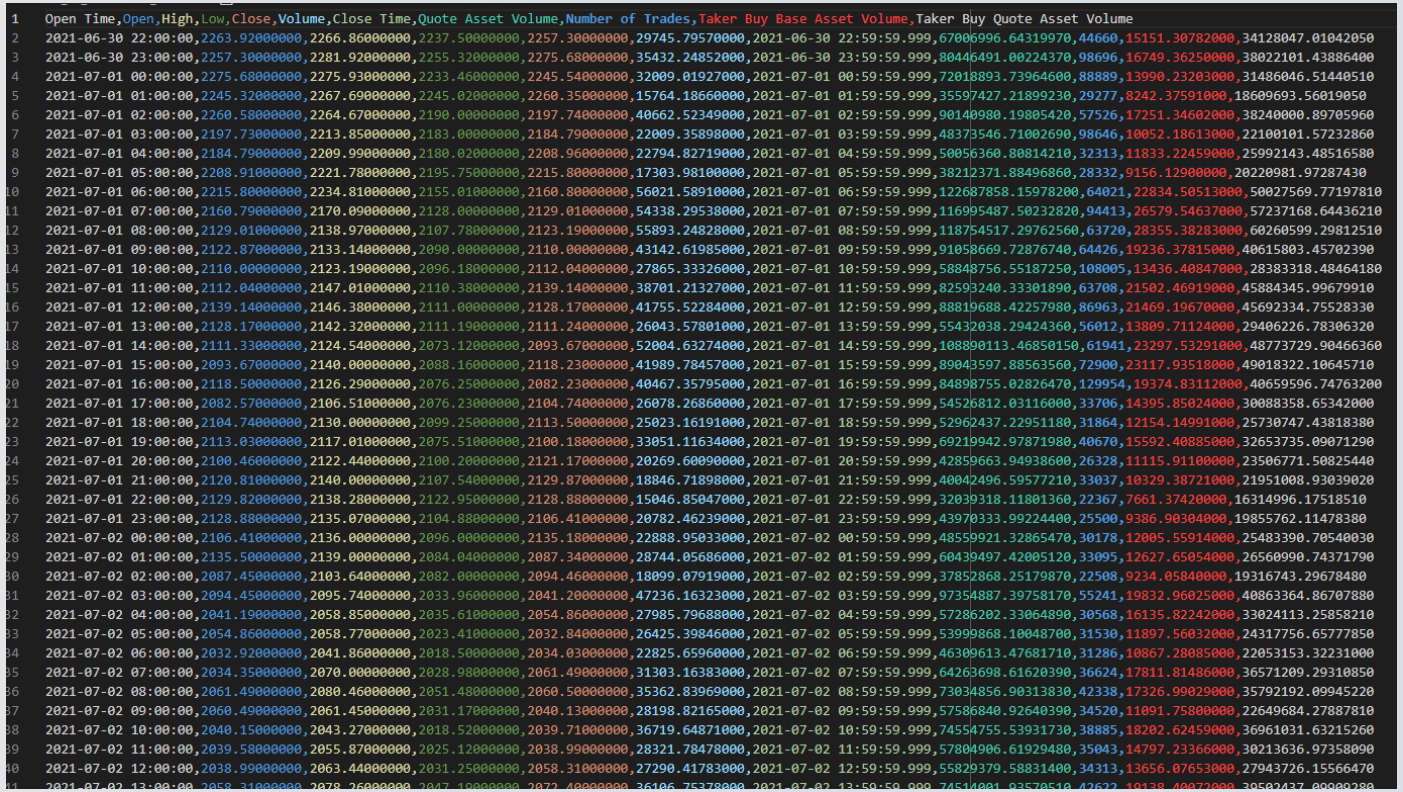
Si seulement on pouvait trouver un format facilement lisible par une machine…

Fichiers CSV sortis droit de chez Binance

Mêmes composants que les bougies Kline :

- Open
- High
- Low
- Close

Et même un peu plus de données utiles...

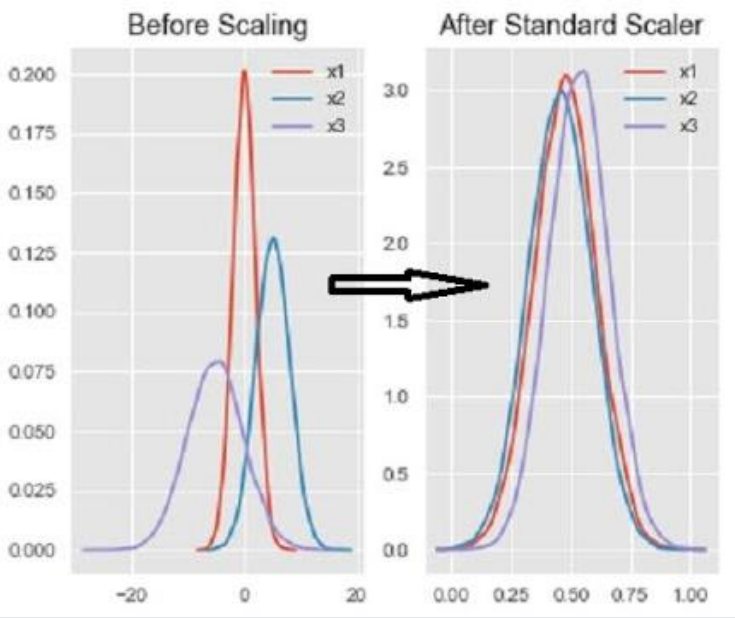


```
Open Time,Open,High,Low,Close,Volume,Close Time,Quote Asset Volume,Number of Trades,Taker Buy Base Asset Volume,Taker Buy Quote Asset Volume
2021-06-30 22:00:00,2263.92000000,2266.86000000,2237.50000000,2257.30000000,29745.79570000,2021-06-30 22:59:59.999,67006996.64319970,44660,15151.30782000,34128047.01042050
2021-06-30 23:00:00,2257.30000000,2281.92000000,2255.32000000,2275.68000000,35432.24852000,2021-06-30 23:59:59.999,80446491.00224370,98696,16749.36250000,38022101.43886400
2021-07-01 00:00:00,2275.68000000,2275.93000000,2233.46000000,2245.54000000,32009.01927000,2021-07-01 00:59:59.999,72018893.73964600,88889,13990.23203000,31486046.51440510
2021-07-01 01:00:00,2245.32000000,2267.69000000,2245.02000000,2260.35000000,15764.18660000,2021-07-01 01:59:59.999,35597427.21899230,29277,8242.37591000,18609693.56019050
2021-07-01 02:00:00,2260.58000000,2264.67000000,2190.00000000,2197.74000000,40662.52349000,2021-07-01 02:59:59.999,90140980.19805420,57526,17251.34602000,38240000.89705960
2021-07-01 03:00:00,2197.73000000,2213.85000000,2183.00000000,2184.79000000,22009.35898000,2021-07-01 03:59:59.999,48373546.71002690,98646,10052.18613000,22100101.57232860
2021-07-01 04:00:00,2184.79000000,2209.99000000,2180.02000000,2208.96000000,22794.82719000,2021-07-01 04:59:59.999,50056360.80814210,32313,11833.22459000,25992143.48516580
2021-07-01 05:00:00,2208.91000000,2221.78000000,2195.75000000,2215.80000000,17303.98100000,2021-07-01 05:59:59.999,38212371.88496860,28332,9156.12900000,20220981.97287430
2021-07-01 06:00:00,2215.80000000,2234.81000000,2155.01000000,2160.80000000,56021.58910000,2021-07-01 06:59:59.999,122687858.15978200,64021,22834.50513000,50027569.77197810
2021-07-01 07:00:00,2160.79000000,2170.09000000,2128.00000000,2129.01000000,54338.29538000,2021-07-01 07:59:59.999,116995487.50232820,94413,26579.54637000,57237168.64436210
2021-07-01 08:00:00,2129.01000000,2138.97000000,2107.78000000,2123.19000000,55893.24828000,2021-07-01 08:59:59.999,118754517.29762560,63720,28355.38283000,60260599.29812510
2021-07-01 09:00:00,2122.87000000,2133.14000000,2090.00000000,2110.00000000,43142.61985000,2021-07-01 09:59:59.999,91058669.72876740,64426,19236.37815000,40615803.45702390
2021-07-01 10:00:00,2110.00000000,2123.19000000,2096.18000000,2112.04000000,27865.33326000,2021-07-01 10:59:59.999,58848756.55187250,108005,13436.40847000,28383318.48464180
2021-07-01 11:00:00,2112.04000000,2147.01000000,2110.38000000,2139.14000000,38701.21327000,2021-07-01 11:59:59.999,82593240.33301890,63708,21502.46919000,45884345.99679910
2021-07-01 12:00:00,2139.14000000,2146.38000000,2111.00000000,2128.17000000,41755.52284000,2021-07-01 12:59:59.999,88819688.42257980,86963,21469.19670000,45692334.75528330
2021-07-01 13:00:00,2128.17000000,2142.32000000,2111.19000000,2111.24000000,26043.57801000,2021-07-01 13:59:59.999,55432038.29424360,56012,13809.71124000,29406226.78306320
2021-07-01 14:00:00,2111.33000000,2124.54000000,2073.12000000,2093.67000000,52004.63274000,2021-07-01 14:59:59.999,108890113.46850150,61941,23297.53291000,48773729.90466360
2021-07-01 15:00:00,2093.67000000,2140.00000000,2088.16000000,2118.23000000,41989.78457000,2021-07-01 15:59:59.999,89043597.88563560,72900,23117.93518000,49018322.10645710
2021-07-01 16:00:00,2118.50000000,2126.29000000,2076.25000000,2082.23000000,40467.35795000,2021-07-01 16:59:59.999,84898755.02826470,129954,19374.83112000,40659596.74763200
2021-07-01 17:00:00,2082.57000000,2106.51000000,2076.23000000,2104.74000000,26078.26860000,2021-07-01 17:59:59.999,54526812.03116000,33706,14395.85024000,30088358.65342000
2021-07-01 18:00:00,2104.74000000,2130.00000000,2099.25000000,2113.50000000,25023.16191000,2021-07-01 18:59:59.999,52962437.22951180,31864,12154.14991000,25730747.43818380
2021-07-01 19:00:00,2113.03000000,2117.01000000,2075.51000000,2100.18000000,33051.11634000,2021-07-01 19:59:59.999,69219942.97871980,40670,15592.40885000,32653735.09071290
2021-07-01 20:00:00,2100.46000000,2122.44000000,2100.20000000,2121.17000000,20269.60090000,2021-07-01 20:59:59.999,42859663.94938600,26328,11115.91100000,23506771.50825440
2021-07-01 21:00:00,2120.81000000,2140.00000000,2107.54000000,2129.87000000,18846.71898000,2021-07-01 21:59:59.999,40042496.59577210,33037,10329.38721000,21951008.93039020
2021-07-01 22:00:00,2129.82000000,2138.28000000,2122.95000000,2128.88000000,15046.85047000,2021-07-01 22:59:59.999,32039318.11801360,22367,7661.37420000,16314996.17518510
2021-07-01 23:00:00,2128.88000000,2135.07000000,2104.88000000,2106.41000000,20782.46239000,2021-07-01 23:59:59.999,43970333.99224400,25500,9386.90304000,19855762.11478380
2021-07-02 00:00:00,2106.41000000,2136.00000000,2096.00000000,2135.18000000,22888.95033000,2021-07-02 00:59:59.999,48559921.32865470,30178,12005.55914000,25483390.70540030
2021-07-02 01:00:00,2135.50000000,2139.00000000,2084.04000000,2087.34000000,28744.05686000,2021-07-02 01:59:59.999,60439497.42005120,33095,12627.65054000,26560990.74371790
2021-07-02 02:00:00,2087.45000000,2103.64000000,2082.00000000,2094.46000000,18099.07919000,2021-07-02 02:59:59.999,37852868.25179870,22508,9234.05840000,19316743.29678480
2021-07-02 03:00:00,2094.45000000,2095.74000000,2033.96000000,2041.20000000,47236.16323000,2021-07-02 03:59:59.999,97354887.39758170,55241,19832.96025000,40863364.86707880
2021-07-02 04:00:00,2041.19000000,2058.85000000,2035.61000000,2054.86000000,27985.79688000,2021-07-02 04:59:59.999,57286202.33064890,30568,16135.82242000,33024113.25858210
2021-07-02 05:00:00,2054.86000000,2058.77000000,2023.41000000,2032.84000000,26425.39846000,2021-07-02 05:59:59.999,53999868.10048700,31530,11897.56032000,24317756.65777850
2021-07-02 06:00:00,2032.92000000,2041.86000000,2018.50000000,2034.03000000,22825.65960000,2021-07-02 06:59:59.999,46309613.47681710,31286,10867.28085000,22053153.32231000
2021-07-02 07:00:00,2034.35000000,2070.00000000,2028.98000000,2061.49000000,31303.16383000,2021-07-02 07:59:59.999,64263698.61620390,36624,17811.81486000,36571209.29310850
2021-07-02 08:00:00,2061.49000000,2080.46000000,2051.48000000,2060.50000000,35362.83969000,2021-07-02 08:59:59.999,73034856.90313830,42338,17326.99029000,35792192.09945220
2021-07-02 09:00:00,2060.49000000,2061.45000000,2031.17000000,2040.13000000,28198.82165000,2021-07-02 09:59:59.999,57586840.92640390,34520,11091.75800000,22649684.27887810
2021-07-02 10:00:00,2040.15000000,2043.27000000,2018.52000000,2039.71000000,36719.64871000,2021-07-02 10:59:59.999,74554755.53931730,38885,18202.62459000,36961031.63215260
2021-07-02 11:00:00,2039.58000000,2055.87000000,2025.12000000,2038.99000000,28321.78478000,2021-07-02 11:59:59.999,57804906.61929480,35043,14797.23366000,30213636.97358090
2021-07-02 12:00:00,2038.99000000,2063.44000000,2031.25000000,2058.31000000,27290.41783000,2021-07-02 12:59:59.999,55829379.58831400,34313,13656.07653000,27943726.15566470
2021-07-02 13:00:00,2058.31000000,2078.26000000,2047.19000000,2072.40000000,36106.75378000,2021-07-02 13:59:59.999,74514001.93570510,42622,18138.48073000,37502437.09909280
```

Pandas (pd pour les intimes)

```
data = pd.read_csv('ethusdt.csv')
```

Le scaling ?



- Mise à l'échelle des données
- Réduction de l'importance des "outliers"
- Facilité d'interprétation pour notre modèle

Création des time series

```
train_dataset = tf.keras.preprocessing.timeseries_dataset_from_array(
    data=train_data,
    targets=train_targets,
    sequence_length=window_length,
    sequence_stride=1,
    shuffle=True,
    batch_size=batch_size,
)

test_dataset = tf.keras.preprocessing.timeseries_dataset_from_array(
    data=test_data,
    targets=test_targets,
    sequence_length=window_length,
    sequence_stride=1,
    shuffle=False,
    batch_size=batch_size,
)
```

- Utilisation des builtins tensorflow
- Création de “fenêtres” de données
- Association des “fenêtres” avec le résultat correspondant

Et maintenant...

FAITES CHAUFFER LES CARTES GRAPHIQUES !

github.com/TEK-STUFF/workshop-lstm

TEK-STUFF/
workshop-lstm

Comment utiliser python, tensorflow et les time series pour prédire l'avenir. Ce workshop va vous permettre de découvrir les bases...

1 Contributor 0 Issues 0 Stars 0 Forks