**Commodity Price Prediction with LSTM Recurrent Neural Network**

Javier Aymachoque

Enero 2021.

Minconsulting S.A.

Data Science Departure.

# Chapter 1 Research Interest

We are in times where the economy of countries is unstable, whether due to civil wars, economic wars, pandemics, new policies, etc. That is why there is a huge need to predict emerging and stable markets in which we can trust our investment, for this is that we seek to use Machine Learning techniques to predict the price of commodities as an investment market.

In this work, the prediction of some commodities was performed using Long Short-Term Memory (LSTM) networks, a type of recurrent neural network [1], for this a Python Jupyter Notebook with the Keras library was used. The dataset used was Commodities Daily Price Dataset, which has daily closing price data for 17 commodities, where the information was extracted from Yahoo Finance webpage.

In order to determine the causal relationship between the prices of different commodities, two machine learning analyzes were carried out: the first was a Feature Importance analysis, and the second was a Dynamic Time Warping (DTW) analysis. The result of relationship between variables will help us to define the appropriate variables to consider in the price prediction of a commodity using LSTM.

Apart from the Keras library, we also used the python libraries Seaborn, Pandas, Scikit-Learn, NumPy, and dtw. In the computational aspect, a laptop with a Core I7 7700HQ processor and 16 GB of RAM was used.

# Chapter 2 Data Description

## Commodities Daily Price Dataset Description

The used dataset was the Commodities Daily Price Dataset, it is a CSV file which tells us the daily closed price from seventeen commodities since 01-01-2018 to 12-31-2020. Commodity closing price information comes from the Yahoo Finance website.[2]

The data is composed by the next 18 columns:

Table 1 Description of the Commodities Daily Price Dataset

| **Column number** | **Column name** | **Yahoo FInance Commodity Code** | **Description** |
| --- | --- | --- | --- |
| 0 | **Date** | - | Date of the day of register |
| 1 | **cocoa** | CC=F | Cocoa commodity price at closed time |
| 2 | **crude\_oil** | CL=F | Crude Oil commodity price at closed time |
| 3 | **cotton** | CT=F | Cotton commodity price at closed time |
| 4 | **gold** | GC=F | Gold commodity price at closed time |
| 5 | **lean\_hogs** | HE=F | Lean Hogs commodity price at closed time |
| 6 | **copper** | HG=F | Copper commodity price at closed time |
| 7 | **coffee** | KC=F | Coffee commodity price at closed time |
| 8 | **wheat** | KE=F | KC HRW Wheat commodity price at closed time |
| 9 | **live\_cattle** | LE=F | Live Cattle commodity price at closed time |
| 10 | **natural\_gas** | NG=F | Natural Gas commodity price at closed time |
| 11 | **platinum** | PL=F | Platinum commodity price at closed time |
| 12 | **sugar** | SB=F | Sugar commodity price at closed time |
| 13 | **silver** | SI=F | Silver commodity price at closed time |
| 14 | **corn** | ZC=F | Corn commodity price at closed time |
| 15 | **oat** | ZO=F | Oat commodity price at closed time |
| 16 | **rough\_rice** | ZR=F | Rough Rice commodity price at closed time |
| 17 | **soybean** | ZS=F | Soybean commodity price at closed time |

Below are some informative charts and tables of the Commodities Daily Price Dataset:

Table 2 Commodities Daily Price Dataset

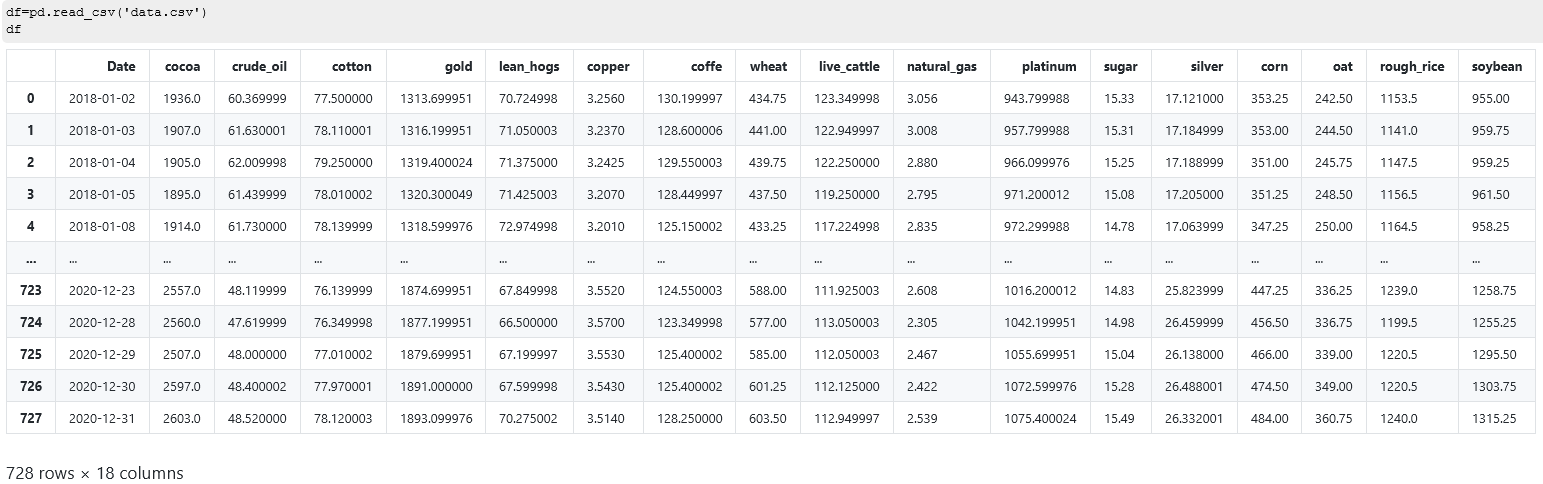




Figure 2 Historical price of the gold commodity from 01-01-2018 to 12-31-2020. Source: Yahoo Finance[3]