

# Stock Market Analysis and Prediction

Autor: Kovacs Norbi-Alexandru

Indrumator: lect. Univ. dr. Mara Macelaru



# Criptomonedele....



Razboiul curent ?  
Market Crash-ul ?  
Sau ce o sa postez



7:57 AM · Feb 4, 2021 · Twitter Web App

Cu un singur tweet, Elon a ridicat valoarea Dogecoin-ului la 37% fata de valoarea dinaintea tweet-ului.



# De ce am vrea sa prezicem pretul...

Pentru ca.... reprezinta o idee interesanta de a aplica machine learning.  
ne-ar aduce un eventual profit.



# Tehnologiile utilizate...

**Google Colab** - o unealta, gratis, in **cloud** folosita pentru a executa code python de machine learning.

**Python** - cea mai populara alegere pentru a scrie machine learning code datorita implementarii usoare.

**Setul de date** - setul de date a fost preluat de la Yahoo Finance si este public.

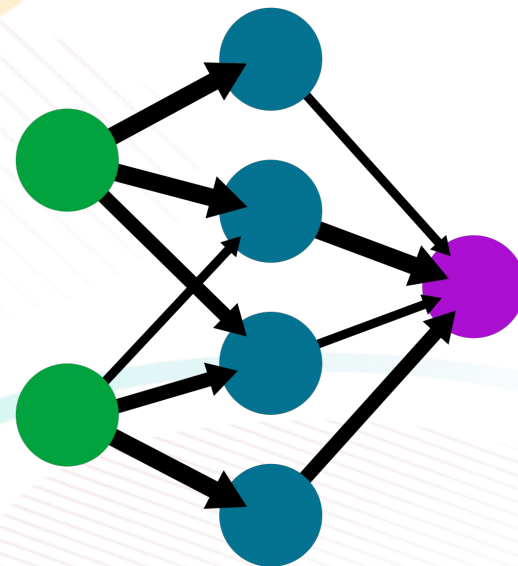


# Ce este o retea neuronală?

Este un sistem de calcul inspirat de rețelele neuronale biologice care constituie creierul animalelor.

A simple neural network

input layer    hidden layer    output layer



# Reteaua neuronală de care ne v-om folosi

Este LSTM (Long Short-Term Memory) fiind un tip de RNN (Recurrent Neural Network) care este folosită în predicția unei **time series** precum previziunile meteorologice, rezultatele alegerilor și, bineînțeles, prețul criptomonedelor. Este folosit LSTM, deoarece rezultatul pasului anterior este utilizat în pasul curent în RNN, fiind mai optim decât o implementare clasică a RNN-ului.

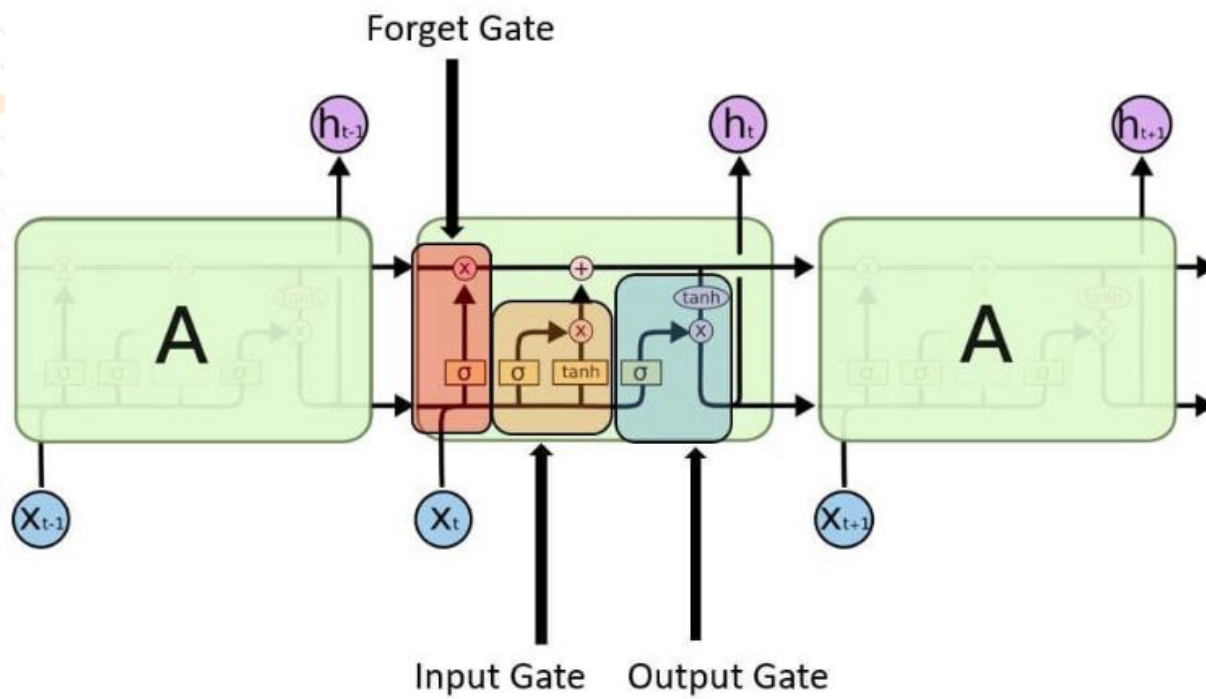
Ideea e să determinăm importanța datelor recente și a celor mai vechi și să aflăm care sunt parametrii care afectează cel mai mult prețurile din ziua “curentă” sau “următoare”.

# Structura unui LSTM

**Input gate** - Este utilizat pentru a determina care dintre valorile de intrare trebuie sa fie utilizate pentru a modifica memoria.

**Output gate** - “Blocul” de input si memoria sunt utilizate pentru a determina iesirea.

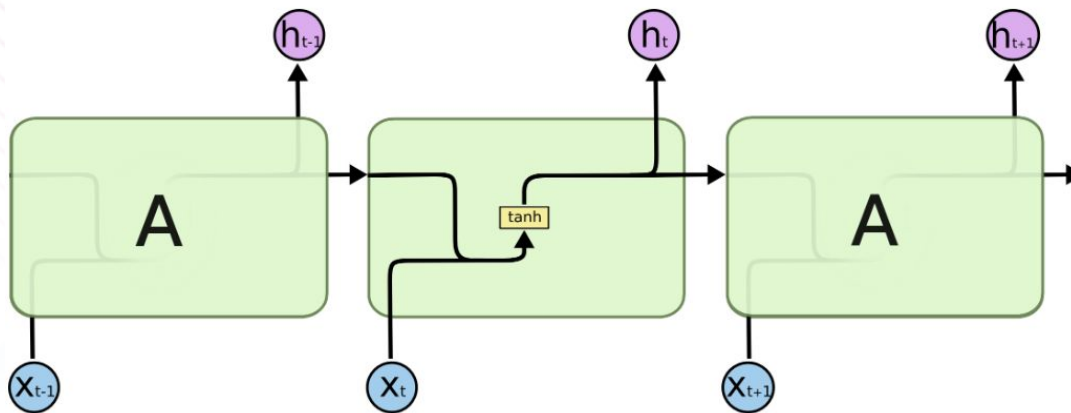
**Forget gate** - Aceasta gaseste detaliile care trebuie eliminate din bloc



# Reteaua LSTM...

O secventa de module de retele neuronale care se repeta consituie toate RNN-urile. Output-ul pasului anterior devine input-ul pasului curent, de aici denumirea de "Reccurent".

La fiecare pas modelul verifica datele de intrare curente si ce stie despre etapa anterioara.

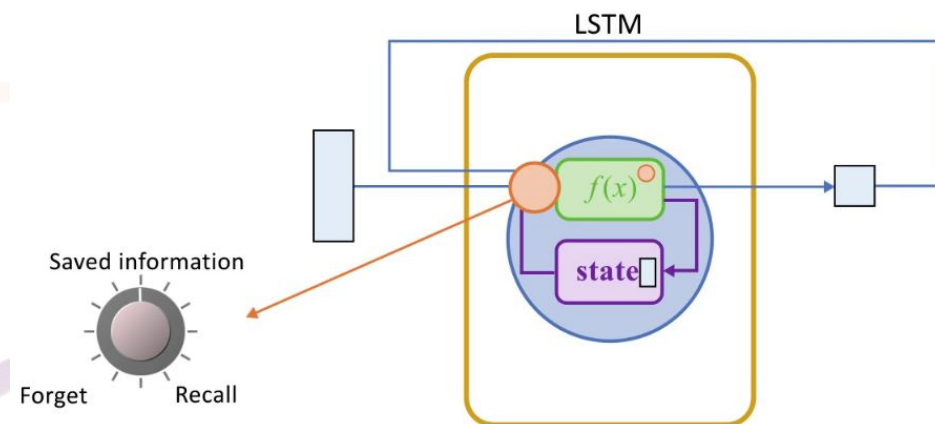




# Ciclul LSTM

Un ciclu LSTM este impartit in patru etape:

- Cu ajutorul **forget gate**-ului este identificata informatia care trebuie uitata din pasul anterior.
- Cu ajutorul **input gate**-ului se cauta informatia noua pentru actualizarea starii celului.
- Informatiile provenite din cele doua porti de mai sus sunt utilizate pentru a actualiza starea celului.
- **Output gate**-ul si operatia de “strivire” au furnizat informatiile utile;



# Partea practica

Pentru a prezice pretul token-ului bitcoin avem nevoie de urmasori pasi:

- Pretul monedei pe parcusul a mai multor ani.
- Preprocesarea datelor.
- Time series.
- Sa construim modelul LSTM.
- Sa folosim modelul pentru a prezice pretul pentru urmatoarele zile.

```
import numpy as np
import pandas as pd
import io
import matplotlib.pyplot as plt
from datetime import datetime
from keras.models import Sequential
from keras.layers import Dense, LSTM, Dropout, GRU
from keras.layers import *
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error, mean_absolute_error
from sklearn.model_selection import train_test_split
from keras.callbacks import EarlyStopping
from keras.optimizers import Adam, SGD
from google.colab import files

getData = files.upload()
```

```
[ ] df.head()
```

	Date	Open	High	Low	Close	Adj Close	Volume
0	2014-09-17	465.864014	468.174011	452.421997	457.334015	457.334015	21056800
1	2014-09-18	456.859985	456.859985	413.104004	424.440002	424.440002	34483200
2	2014-09-19	424.102997	427.834991	384.532013	394.795990	394.795990	37919700
3	2014-09-20	394.673004	423.295990	389.882996	408.903992	408.903992	36863600
4	2014-09-21	408.084991	412.425995	393.181000	398.821014	398.821014	26580100

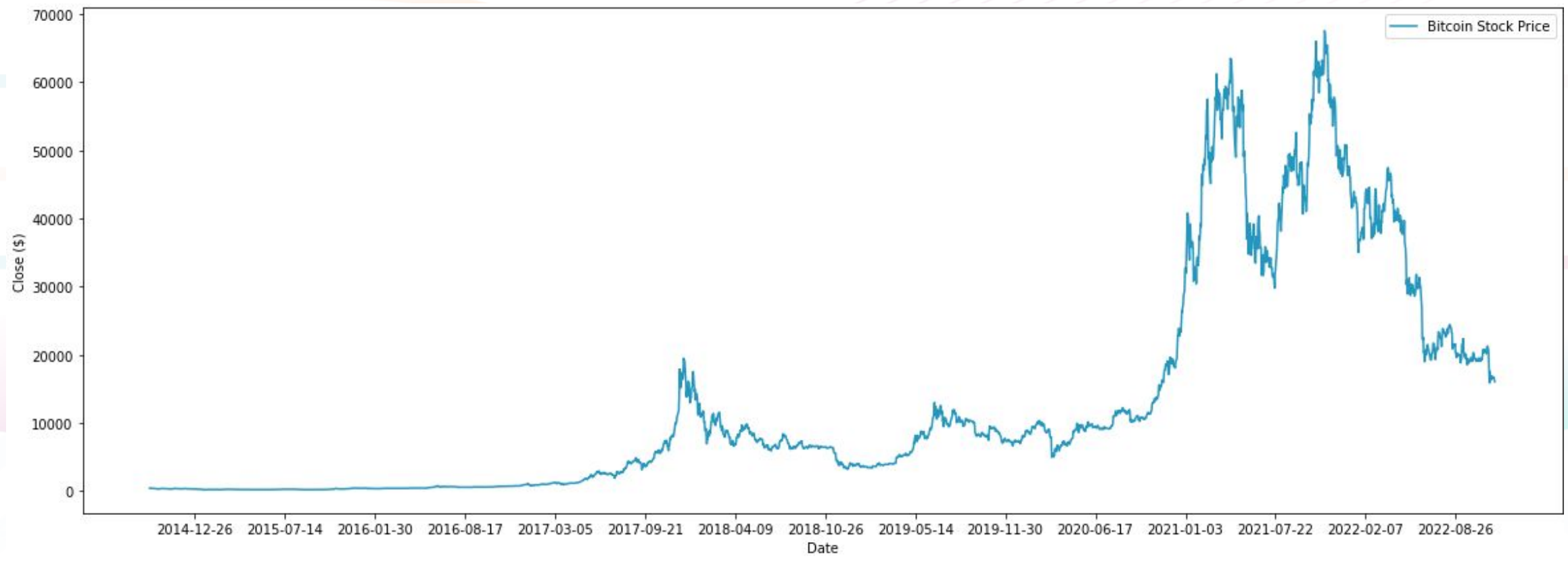
```
[ ] df.shape
```

```
(2988, 7)
```

Avem aproape 3000 de data points simbolizand zile, adica date pe parcursul a aproape 9 ani.

```
df['Close'] = df['Close'].astype(float)

plt.figure(figsize = (20, 7))
plt.plot(df['Date'].values, df['Close'].values, label = 'Bitcoin Stock Price', color = '#2596be')
plt.xticks(np.arange(100, df.shape[0], 200))
plt.xlabel('Date')
plt.ylabel('Close ($)')
plt.legend()
plt.show()
```



```
[ ] model = Sequential()

model.add(LSTM(units = 50, return_sequences = True, input_shape = (X_train.shape[1], 1)))
model.add(Dropout(0.2))

model.add(LSTM(units = 50, return_sequences = True))
model.add(Dropout(0.2))

model.add(LSTM(units = 50, return_sequences = True))
model.add(Dropout(0.2))

model.add(LSTM(units = 50))
model.add(Dropout(0.2))

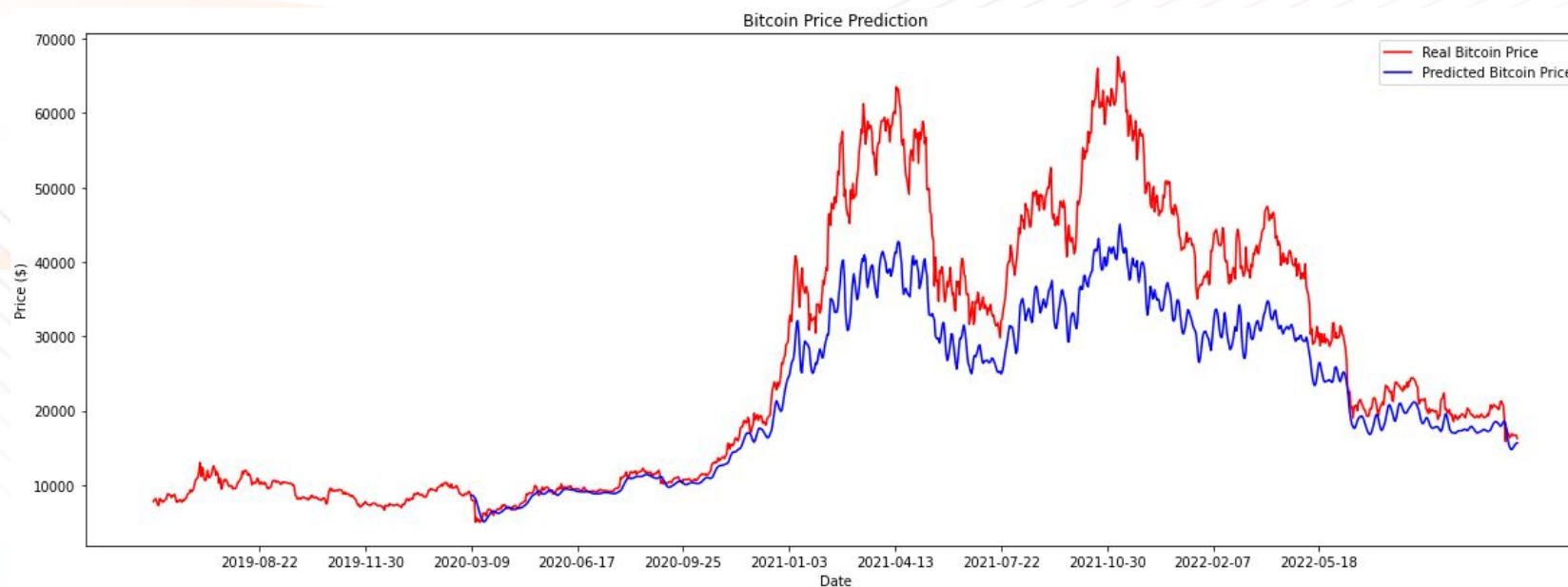
model.add(Dense(units = 1))
model.summary()
```

Initializam modelul de RNN, acesta fiind primul start in LSTM.

```
model.compile(optimizer = 'adam', loss = 'mean_squared_error')
model.fit(X_train, y_train, epochs = 100, batch_size = 128);
```

Antrenam modelul

```
[ ] plt.figure(figsize = (20,7))
plt.plot(df['Date'].values[1700:], df_volume[1700:], color = 'red', label = 'Real Bitcoin Price')
plt.plot(df['Date'][-predict.shape[0]:].values, predict, color = 'blue', label = 'Predicted Bitcoin Price')
plt.xticks(np.arange(100,df[1800:].shape[0],100))
plt.title('Bitcoin Price Prediction')
plt.xlabel('Date')
plt.ylabel('Price ($)')
plt.legend()
plt.show()
```



Dupa cum observam pretul prezis a fost de 15900\$ iar cel real este 15950\$ (21/11/2022)

# Concluzii...

Putem sa prezicem preturile unor criptomonede cu suficiente date, insa aceste preziceri niciodata nu o sa fie 100% sigure, intrucat exista si elemente care nu tin de noi.

**DISCLAIMER: IDEILE ARATATE IN ACEASTA PREZENTARE NU REPREZINTA SFATURI FINANCIARE!**

**DISCLAIMER: IDEILE ARATATE IN ACEASTA PREZENTARE NU REPREZINTA SFATURI FINANCIARE!**



Thank  
You

