**Cryptocurrency trend Analysis And Forecasting**

Abstract

Cryptocurrency is a new sort of asset that has emerged as a result of the advancement of financial technology and it has created a big opportunity for researches. Cryptocurrency price forecasting is difficult due to price volatility and dynamism. Around the world, there are hundreds of cryptocurrencies that are used. The future of the cryptocurrency market looks promising with opportunities in the peer-to-peer payment, remittance, e-commerce and retail, and media & entertainment industries. The global cryptocurrency market is expected to grow with a CAGR of 32% from 2019 to 2024. The major growth drivers for this market are transparency and immutability of the distributed ledger technology, growing remittance in developing countries, fluctuating monetary regulations, and a significant increase in venture capital investments.. Using the ARIMA model, we made a short-term forecast (from 5 to 30 days) for the cryptocurrencies. We found that the proposed approach was more accurate than the ARIMA-ARFIMA models in forecasting cryptocurrencies time series both in the periods of slow rising (falling) and in the periods of transition dynamics (change of trend).

Introduction

Cryptocurrency is a virtual or digital currency used in financial systems. It is secured by cryptography that makes it impossible to be counterfeited or double-spent. Furthermore, it is not issued from a central authority or central banks, and it is decentralized virtual currencies that can be converted via cryptographic procedures and this make it distinguishable from traditional currencies. The other feature is that it is created by technology called blockchain , which is an extremely complex, and aims to storing data that makes it difficult or impossible to alter, hack, or defraud the system. Bitcoin has begun to carve out a niche for itself, which may either help cryptocurrencies to gain widespread acceptance or be the major cause of their demise. Cryptocurrencies are still in their infancy, and it is difficult to predict whether they will ever be widely used in global markets or not . The most prominent cryptocurrency, Bitcoin, was established in 2009 and for more than two years was the sole Blockchain-based cryptocurrency. Today, however, there are over 5000 cryptocurrencies and 5.8 million active users in the cryptocurrency industry . Because of its intrinsic nature of mixing encryption technology with monetary units, Bitcoin has recently gotten a lot of attention in the disciplines of economics, cryptography, and computer science .

The rapid development of digital currencies during the last decade is one of the most controversial and ambiguous innovations in the modern global economy. Significant fluctuations in the exchange rate of cryptocurrencies and their high volatility, as well as the lack of legal regulation of their transactions in most countries resulted in significant risks associated with investment into crypto assets. This has led to heated discussions about their place and role in the modern economy.

Problem statement

To analyse and visualize trends in past crypto prices and forecast future prices using ARIMA

Technical

* Software

Python

Python is a [high-level](https://en.wikipedia.org/wiki/High-level_programming_language), [general-purpose programming language](https://en.wikipedia.org/wiki/General-purpose_programming_language). Its design philosophy emphasizes [code readability](https://en.wikipedia.org/wiki/Code_readability) with the use of [significant indentation](https://en.wikipedia.org/wiki/Off-side_rule). Its [language constructs](https://en.wikipedia.org/wiki/Language_construct) and [object-oriented](https://en.wikipedia.org/wiki/Object-oriented_programming) approach aim to help [programmers](https://en.wikipedia.org/wiki/Programmers) write clear, logical code for small- and large-scale projects.

* Dataset

The yfinance is one of the famous modules in Python, which is used to collect online data, and with it, we can collect the financial data of Yahoo. With the help of the yfinance module, we retrieve and collect the company's financial information (such as financial ratios, etc.) as well as the histories of marketing data by using its functions. But, before we start learning more about this module and its implementation as well as applications, we have to install the yfinance module in our system (as it is not a built-in module in Python). Once the installation process is done, we will proceed to the implementation part of the yfinance module.

[<https://finance.yahoo.com/>](https://finance.yahoo.com/)

* Library

1. Streamlit

Streamlit is an open source app framework in Python language. It helps us create web apps for data science and machine learning in a short time. It is compatible with major Python libraries such as scikit-learn, Keras, PyTorch, SymPy(latex), NumPy, pandas, Matplotlib etc. With Streamlit, no callbacks are needed since widgets are treated as variables. Data caching simplifies and speeds up computation pipelines.

1. Matplotlib

Matplotlib is a [plotting](https://en.wikipedia.org/wiki/Plotter) [library](https://en.wikipedia.org/wiki/Library_(computer_science)) for the [Python](https://en.wikipedia.org/wiki/Python_(programming_language)) programming language and its numerical mathematics extension [NumPy](https://en.wikipedia.org/wiki/NumPy). It provides an [object-oriented](https://en.wikipedia.org/wiki/Object-oriented_programming) [API](https://en.wikipedia.org/wiki/API) for embedding plots into applications using general-purpose [GUI toolkits](https://en.wikipedia.org/wiki/GUI_toolkit) like [Tkinter](https://en.wikipedia.org/wiki/Tkinter" \o "Tkinter), [wxPython](https://en.wikipedia.org/wiki/WxPython" \o "WxPython), [Qt](https://en.wikipedia.org/wiki/Qt_(software)), or [GTK](https://en.wikipedia.org/wiki/GTK). There is also a [procedural](https://en.wikipedia.org/wiki/Procedural_programming) "pylab" interface based on a [state machine](https://en.wikipedia.org/wiki/State_machine) (like [OpenGL](https://en.wikipedia.org/wiki/OpenGL)), designed to closely resemble that of [MATLAB](https://en.wikipedia.org/wiki/MATLAB), though its use is discouraged.[[3]](https://en.wikipedia.org/wiki/Matplotlib#cite_note-3) [SciPy](https://en.wikipedia.org/wiki/SciPy) makes use of Matplotlib.

1. Pandas

pandas is a [software library](https://en.wikipedia.org/wiki/Software_library) written for the [Python programming language](https://en.wikipedia.org/wiki/Python_(programming_language)) for data manipulation and [analysis](https://en.wikipedia.org/wiki/Data_analysis). In particular, it offers [data structures](https://en.wikipedia.org/wiki/Data_structure) and operations for manipulating numerical tables and [time series](https://en.wikipedia.org/wiki/Time_series).

1. Seaborn

Seaborn is a Python data visualization library based on [matplotlib](https://matplotlib.org/). It provides a high-level interface for drawing attractive and informative statistical graphics.

1. Beautiful Soup

BeautifulSoup is a [Python](https://en.wikipedia.org/wiki/Python_(programming_language)) package for parsing [HTML](https://en.wikipedia.org/wiki/HTML) and [XML](https://en.wikipedia.org/wiki/XML) documents (including having malformed markup, i.e. non-closed tags, so named after [tag soup](https://en.wikipedia.org/wiki/Tag_soup)). It creates a parse tree for parsed pages that can be used to extract data from HTML,[[3]](https://en.wikipedia.org/wiki/Beautiful_Soup_(HTML_parser)#cite_note-3) which is useful for [web scraping](https://en.wikipedia.org/wiki/Web_scraping).

Algorithm

ARIMA

ARIMA is an acronym that stands for AutoRegressive Integrated Moving Average. It is a class of model that captures a suite of different standard temporal structures in time series data.

ARIMA is an acronym that stands for AutoRegressive Integrated Moving Average. It is a generalization of the simpler AutoRegressive Moving Average and adds the notion of integration.

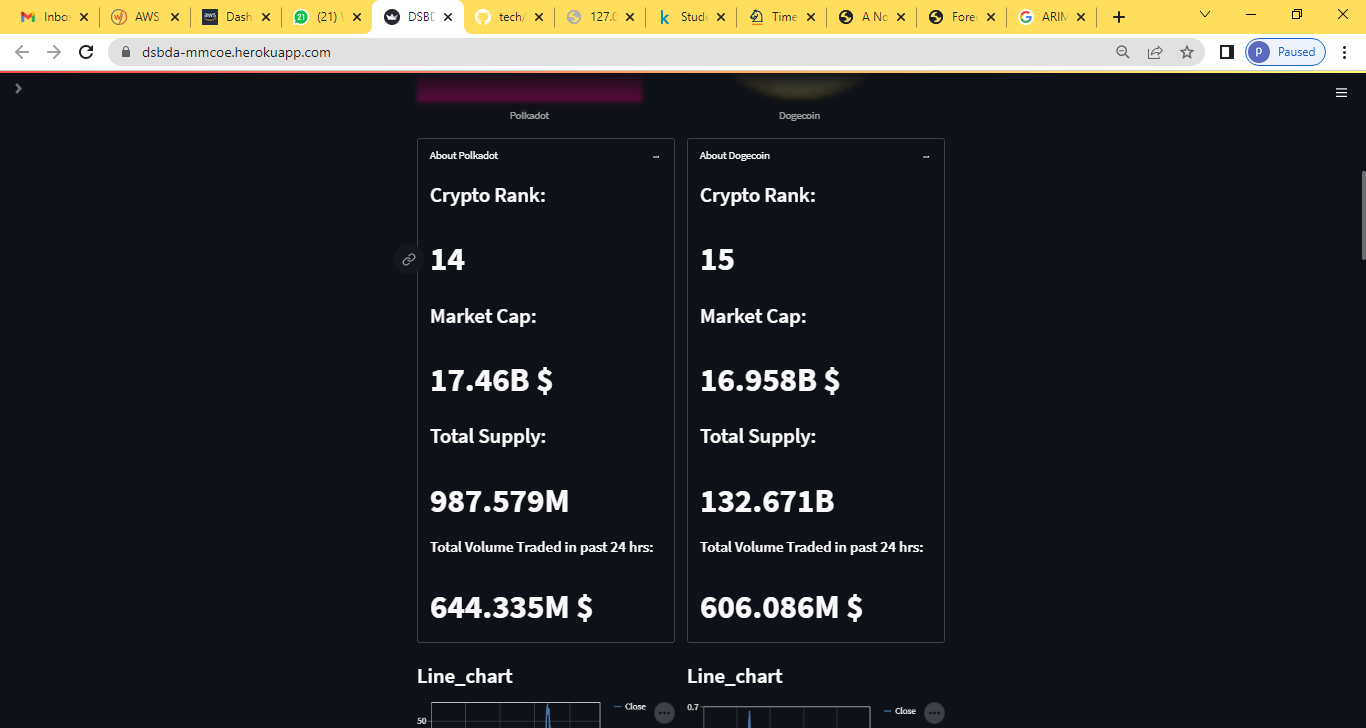
This acronym is descriptive, capturing the key aspects of the model itself. Briefly, they are:

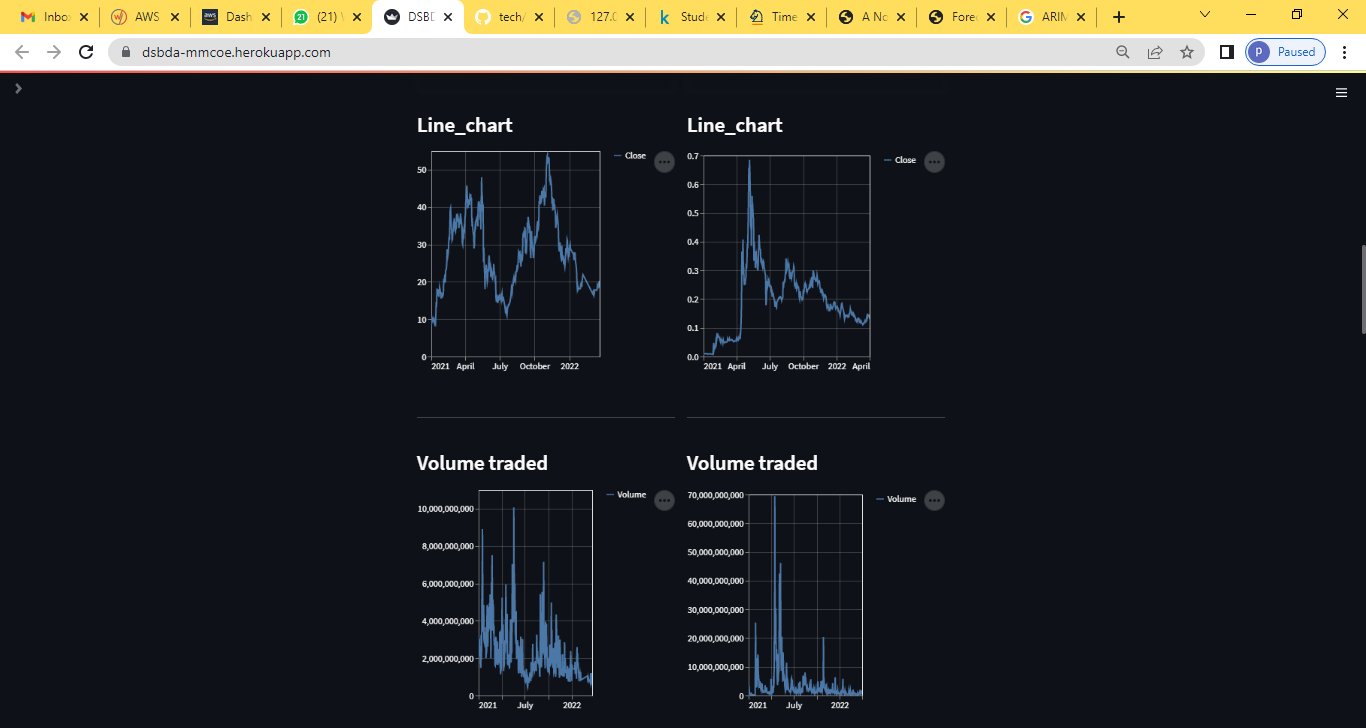
* AR: *Autoregression*. A model that uses the dependent relationship between an observation and some number of lagged observations.
* I: *Integrated*. The use of differencing of raw observations (e.g. subtracting an observation from an observation at the previous time step) in order to make the time series stationary.
* MA: *Moving Average*. A model that uses the dependency between an observation and a residual error from a moving average model applied to lagged observations.

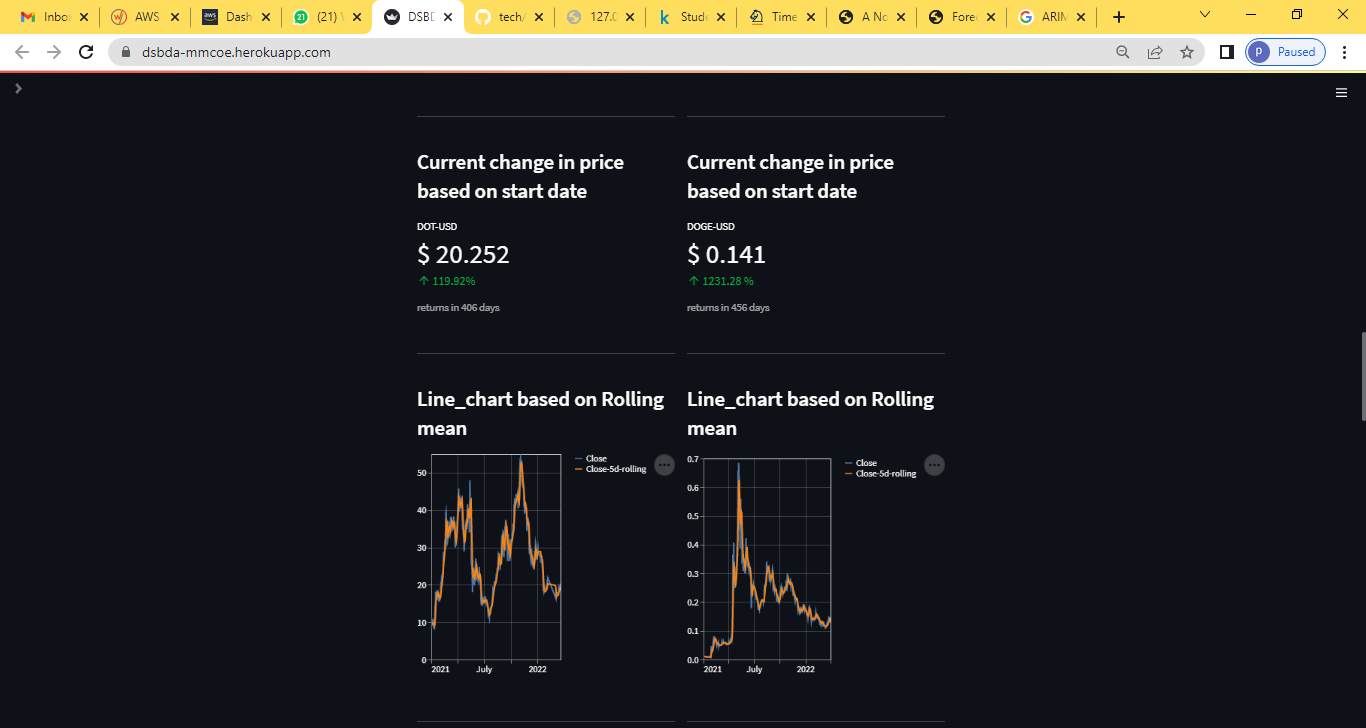
The parameters of the ARIMA model are defined as follows:

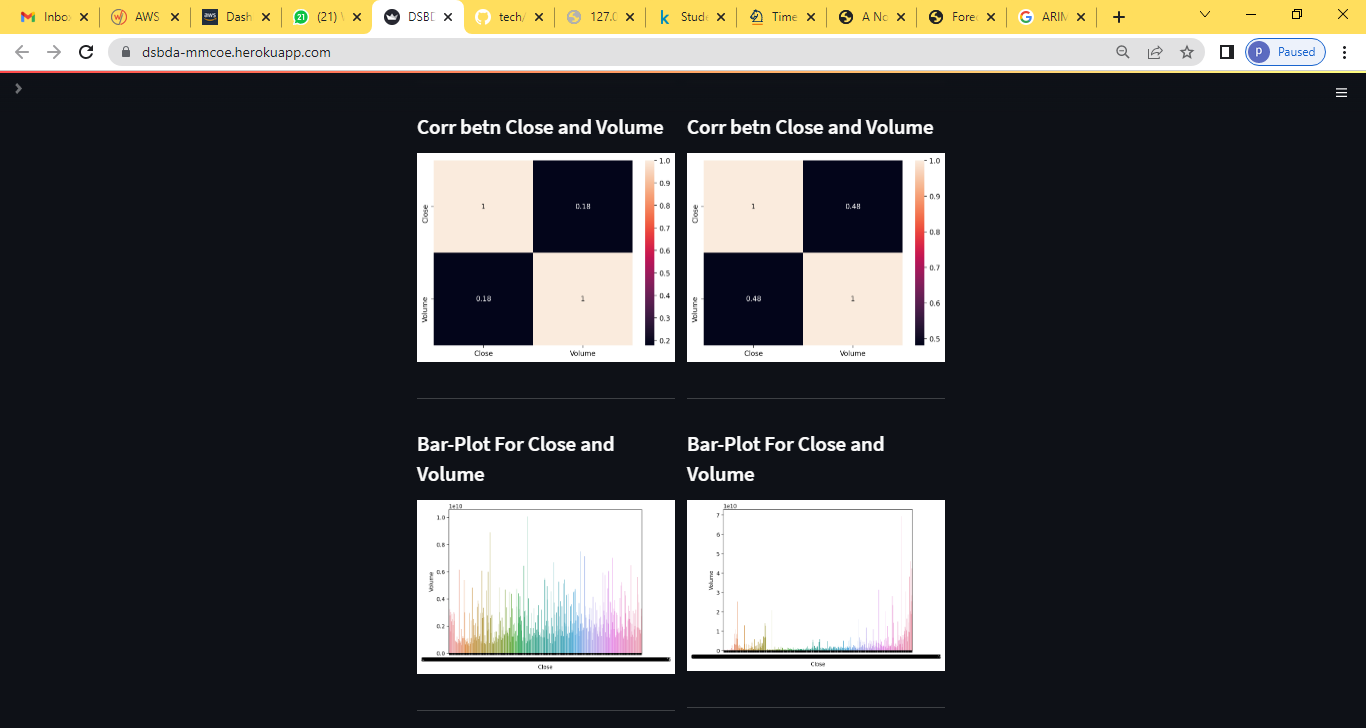
* p: The number of lag observations included in the model, also called the lag order.
* d: The number of times that the raw observations are differenced, also called the degree of differencing.
* q: The size of the moving average window, also called the order of moving average.

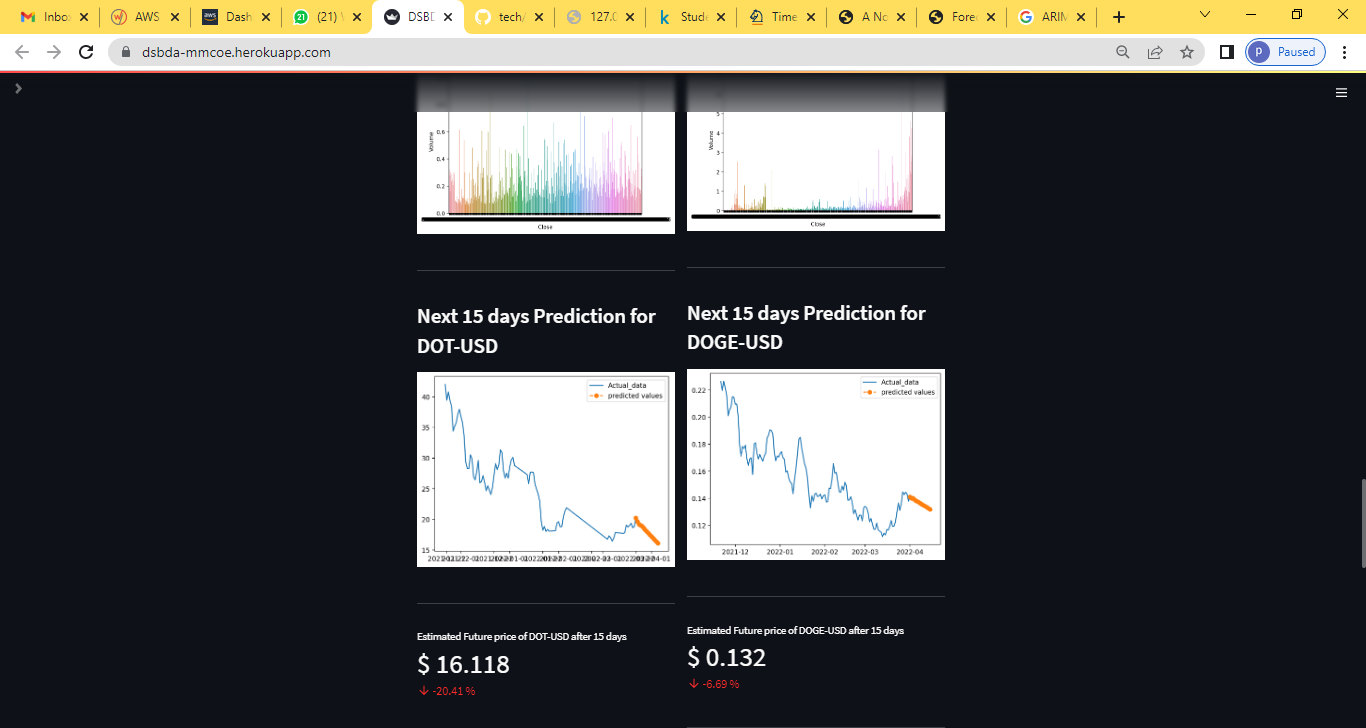
Result / Visualizations

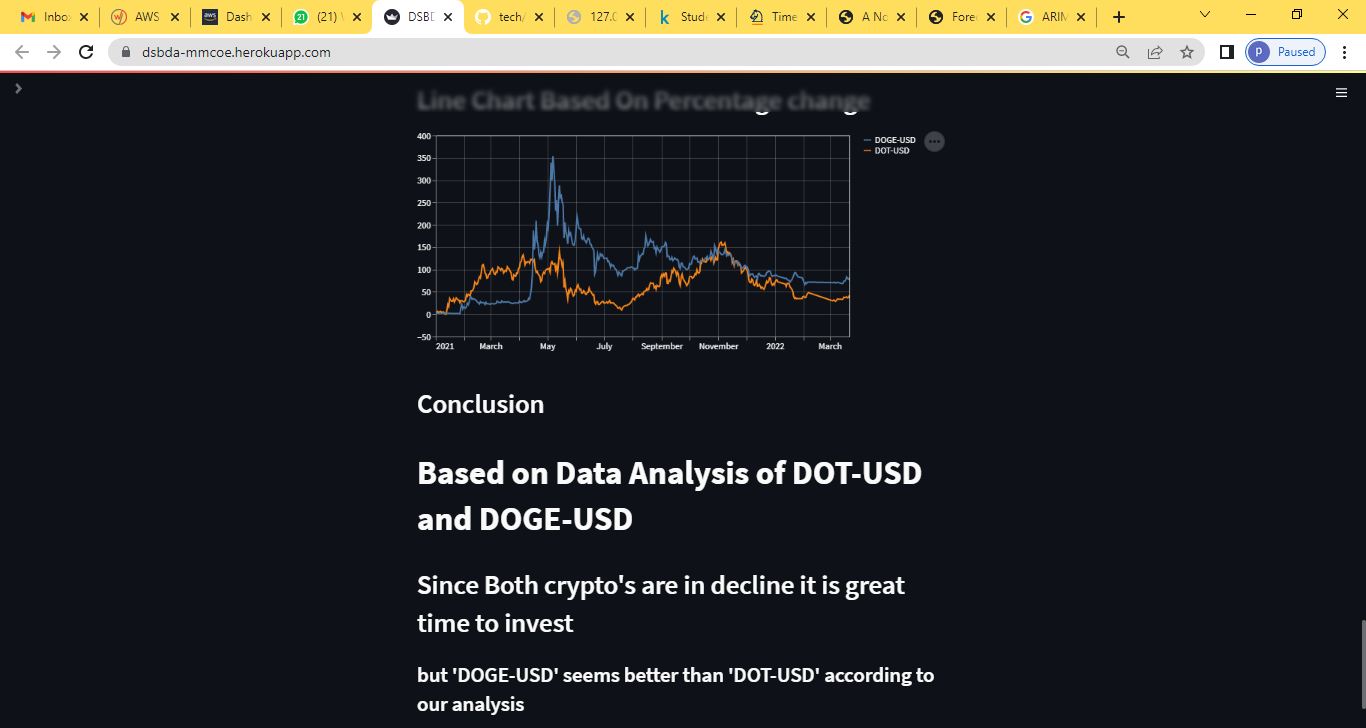












References

1] A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms [<https://bit.ly/3LlaYfa>](https://bit.ly/3LlaYfa)