Kharagpur Data Science Hackathon

Algorithmic Trading Model Development for BTC/USDT Crypto Market

Team : 21je0001

Aadya Dewangan
Arpita Kargaonkar
Aditya Mukherjee

Problem Statement

The problem consists of developing algorithmic trading models for the BTC/USDT cryptocurrency market, aiming to outperform benchmark returns.

The primary objective was to create trading algorithms that can generate returns while managing risk effectively in the specific BTC/USDT market.

Algorithmic Trading

A process for executing orders utilizing automated and pre-programmed trading instructions to account for variables such as price, timing, and volume.

Approach

Approach 1

Using Indices like exponential moving index**(EMA)**, Relative Strength Index**(RSI),** Bollinger bands to get an idea of the trend followed

Approach 2

Building and training an Autoregressive Integrated Moving Average**(ARIMA)** model for time series forecasting of BTC/USDT

Approach 3

Creating and instructing a model based on Long Short Term Memory**(LSTM)** for the purpose of forecasting time series data related to BTC/USDT.

Obstacles Encountered

Problem 1

During the application of indices , certain obstacles were experienced such as lagging indicators, false signals, market conditions, parameter sensitivity, etc.

Problem 2

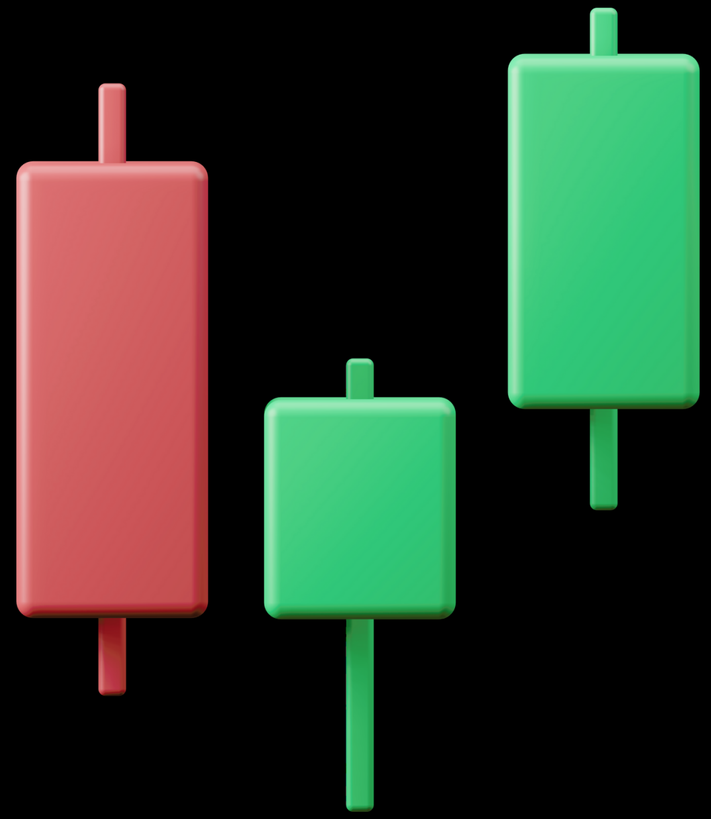
Building an ARIMA model posed us with a variety of challenges such as non-stationarity, lack of seasonality consideration, parameter estimation, overfitting, model-evaluation, etc.

Problem 3

Building an LSTM model involved challenges such as, forecasting horizons, , navigating model complexity, addressing overfitting, tuning hyperparameters, limited interpretability, etc.

Our Solution

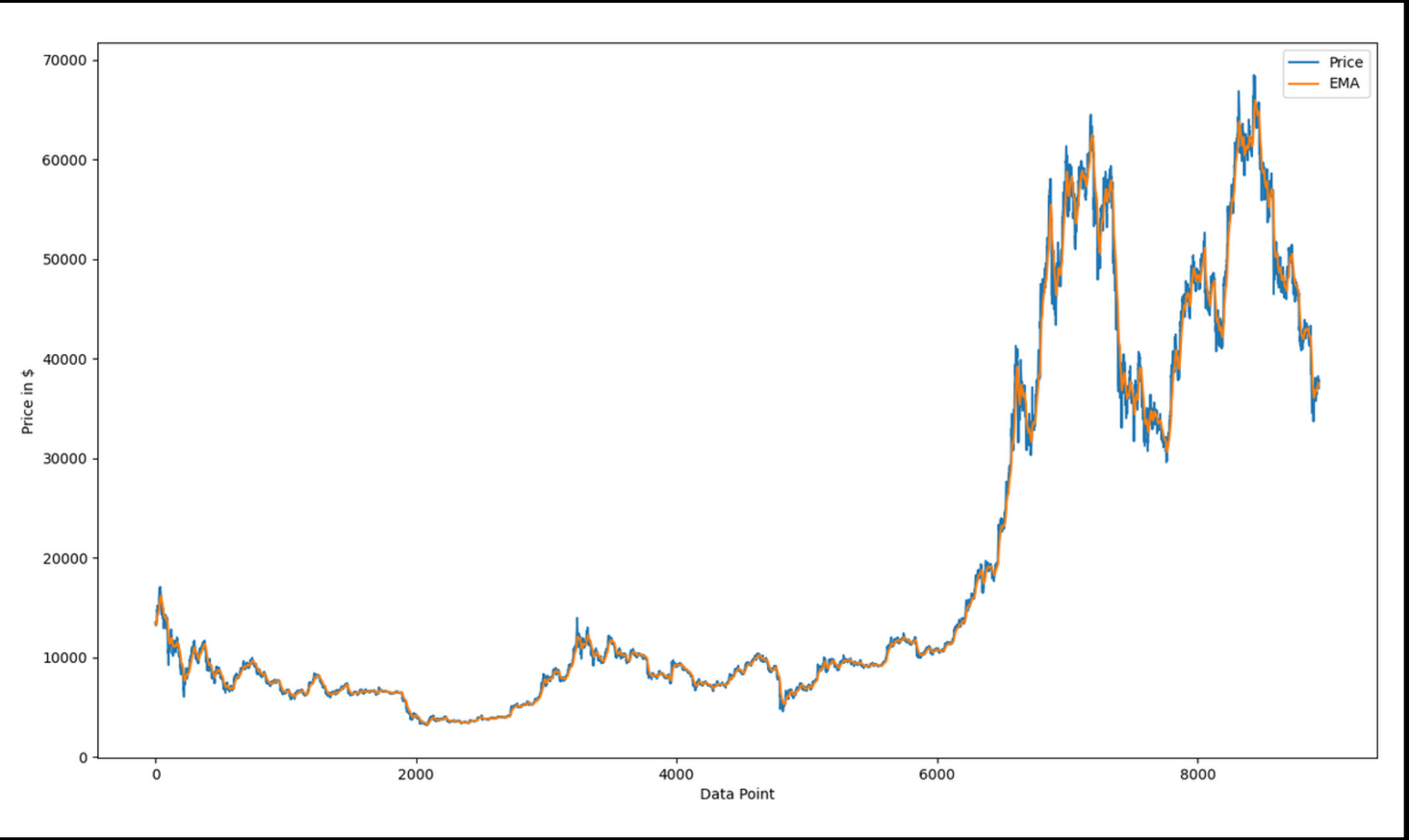
Price Action Analysis
Using LSTM Network



Technical Analysis

The process of identifying trading opportunities in price trends and patterns seen on charts.

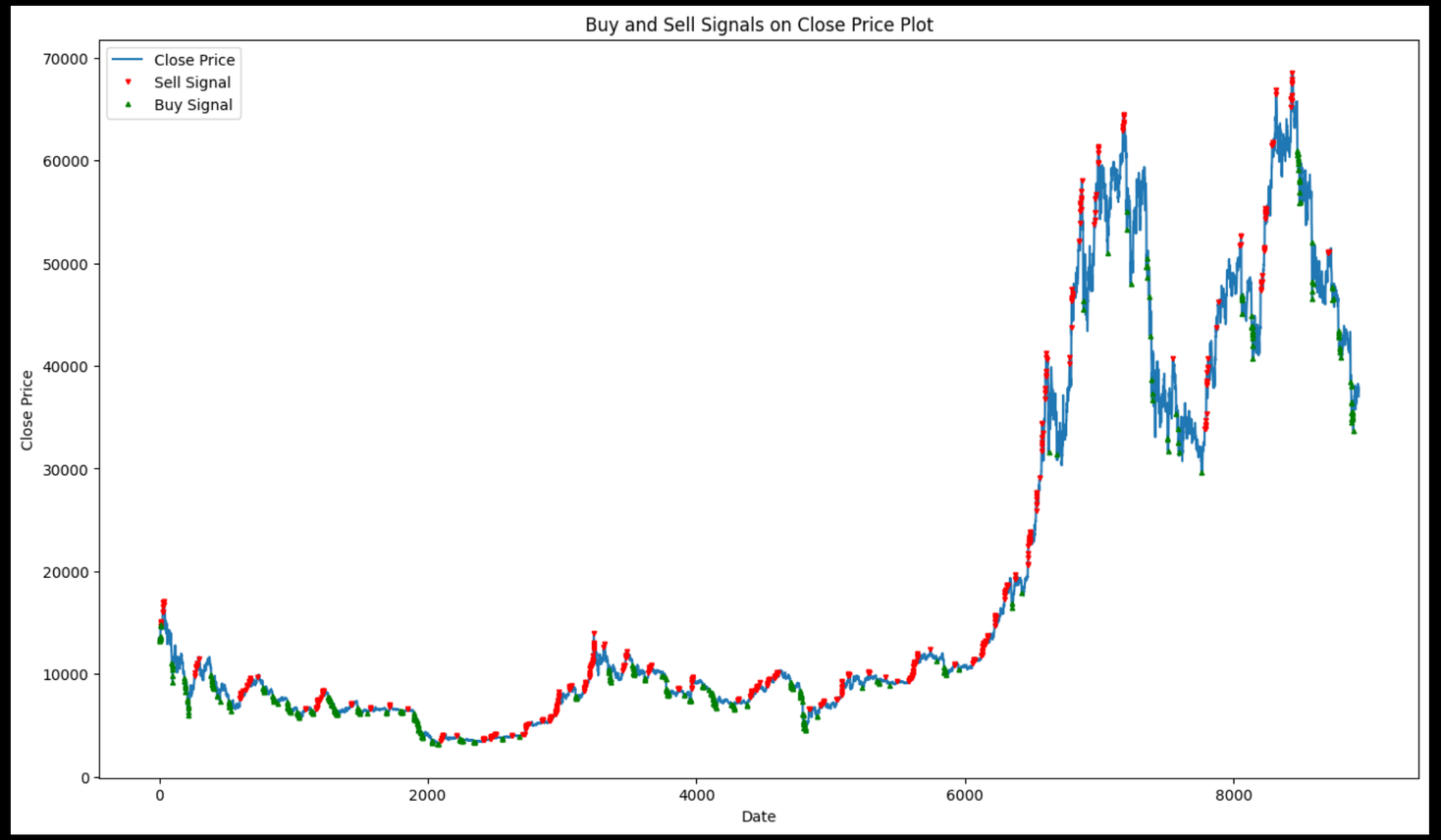
Exponential Moving Average (EMA)



EMA is a technical chart indicator that tracks the price of an investment (like a stock or commodity) over time.

In this case, Halflife is set to 7, which specifies the exponential decay factor. It determines how quickly the weights decrease exponentially.

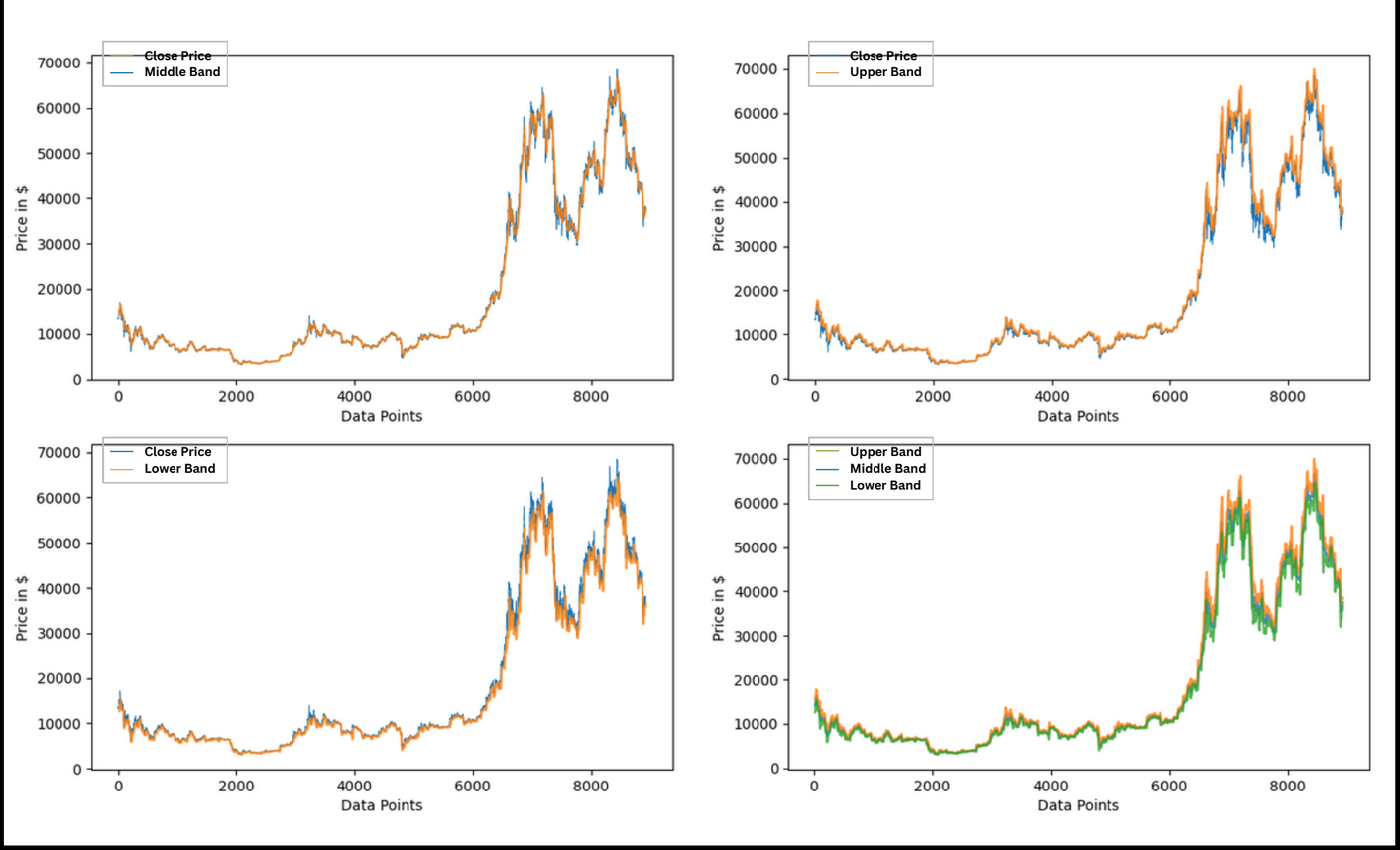
Relative Strength Index (RSI)



The Relative Strength Index (RSI), is a momentum oscillator that measures the speed and change of price movements.

Traditionally the entity is considered overbought when RSI is above 70 and oversold when below 30.

Bollinger Bands



Bollinger Bands are a technical analysis tool comprising three bands – upper, middle, and lower – plotted on a price chart. The middle band is a moving average, and the upper and lower bands are calculated based on standard deviations from the middle band.

Traders use Bollinger Bands to identify overbought or oversold conditions, potential trend reversals, and periods of price consolidation or breakout.

Time Series Analysis

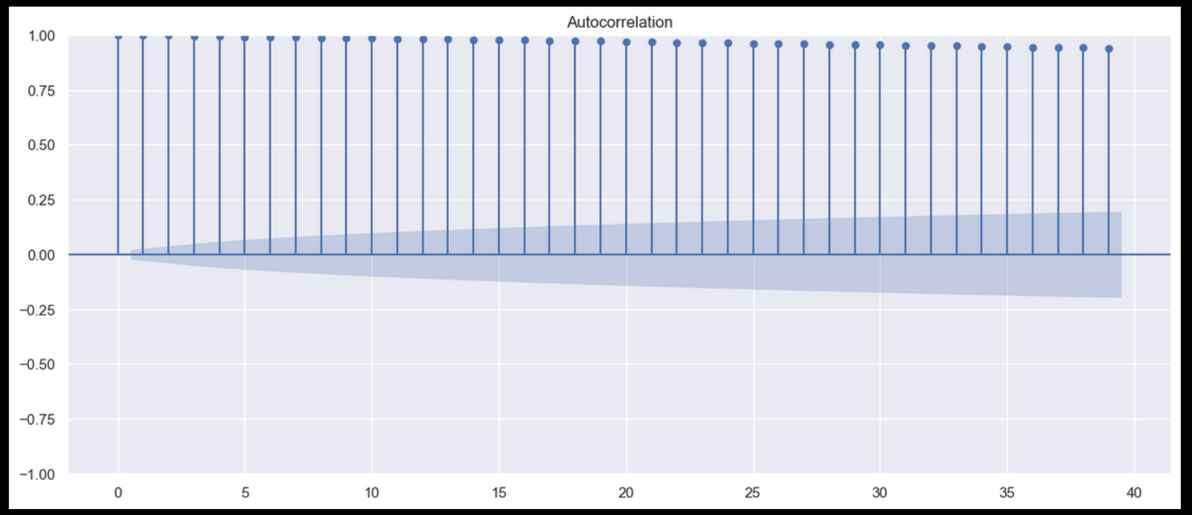
The process of inferring what has happened to a series of data points in the past and attempting to predict what will happen to it the future.

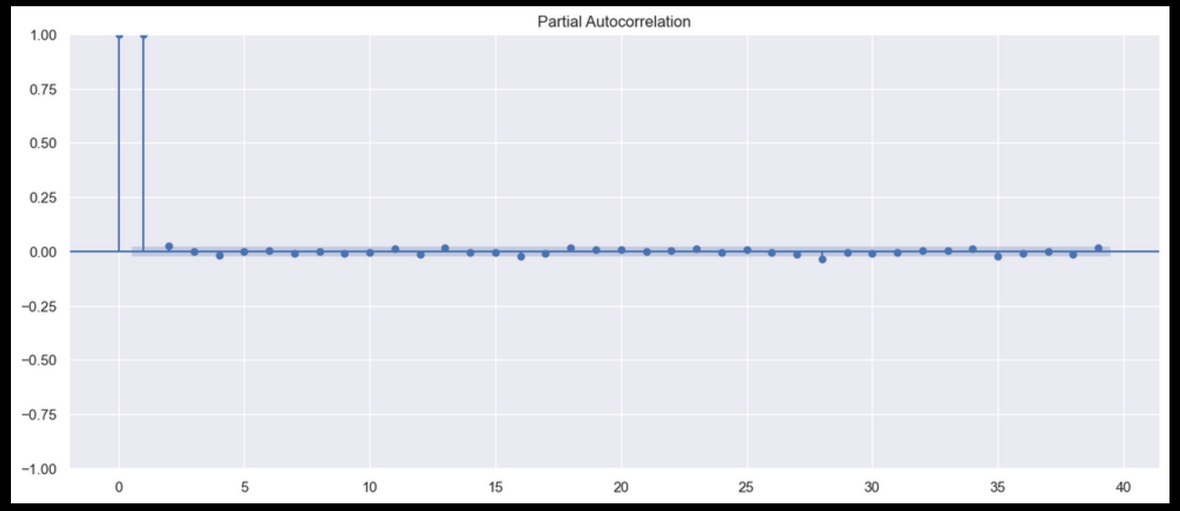
ARIMA

Application of logarithm to raw data

Logarithms are used in ARIMA models for time series analysis to stabilize variance, linearize trends, and enhance interpretability. This transformation is crucial for addressing changing variances and non-linearity in the data

ARIMA

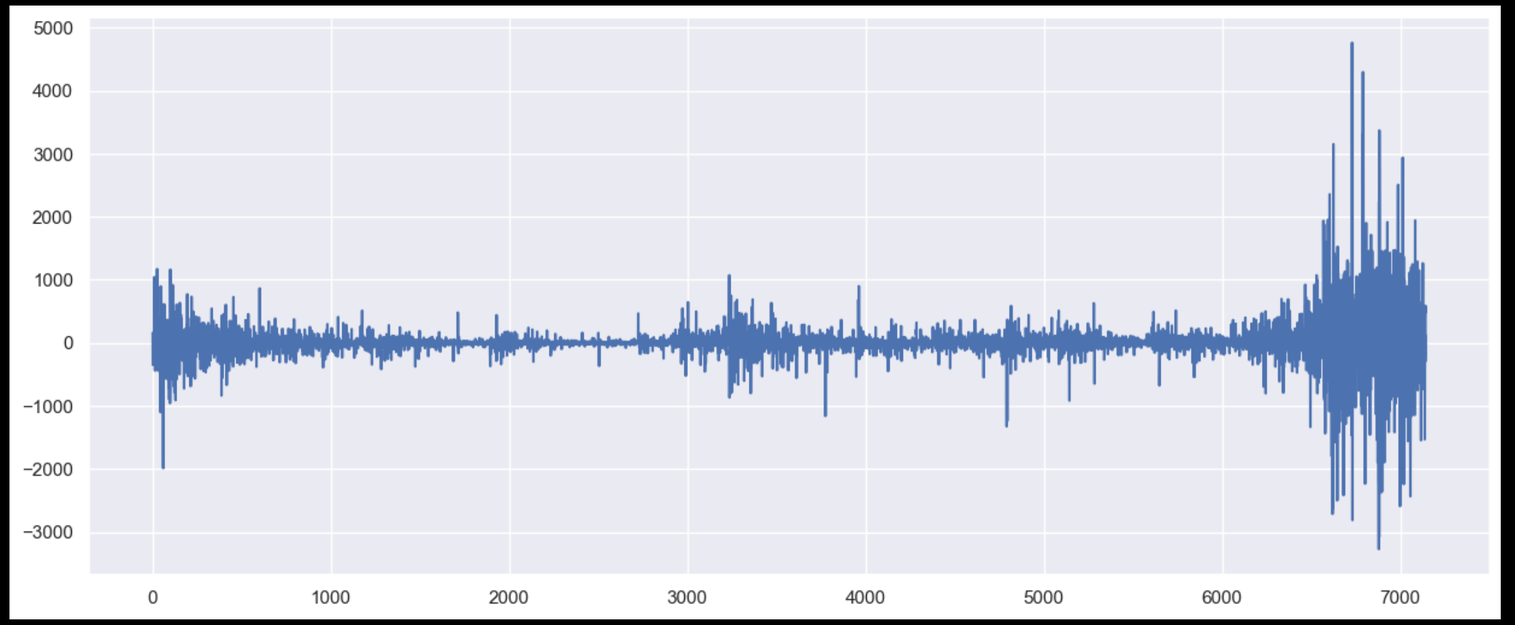




Autocorrelation and Partial Autocorrelation tests

Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) tests are vital in ARIMA modeling. ACF reveals correlation at different lags, aiding in identifying the order of the Moving Average (MA) component. PACF isolates direct relationships, assisting in determining the AutoRegressive (AR) component's order, crucial for effective ARIMA model specification.

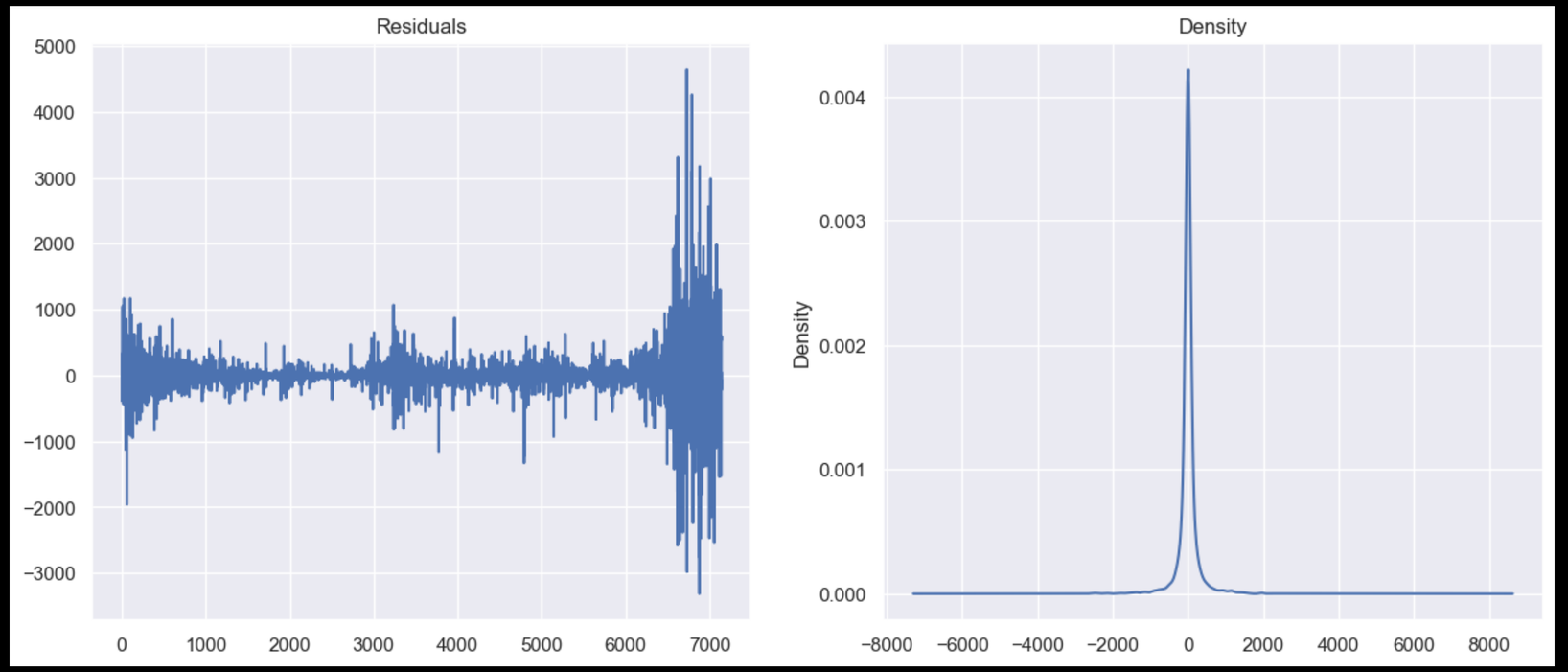
ARIMA



Application of differencing to data

Differencing is a key step in ARIMA modeling to achieve stationarity by subtracting the series' previous values from the current ones. This mitigates trends and seasonality, making the data suitable for analysis. Differencing is integral to the Integrated (I) component of ARIMA, allowing effective time series modeling.

ARIMA



Residuals and Density plots

Residuals and density plots in ARIMA models are essential for model evaluation. Examining residuals helps assess the model's accuracy, ensuring no systematic patterns remain. Density plots provide insights into the distribution of residuals, aiding in validating the assumption of normally distributed errors, a crucial aspect for reliable model performance

ARIMA

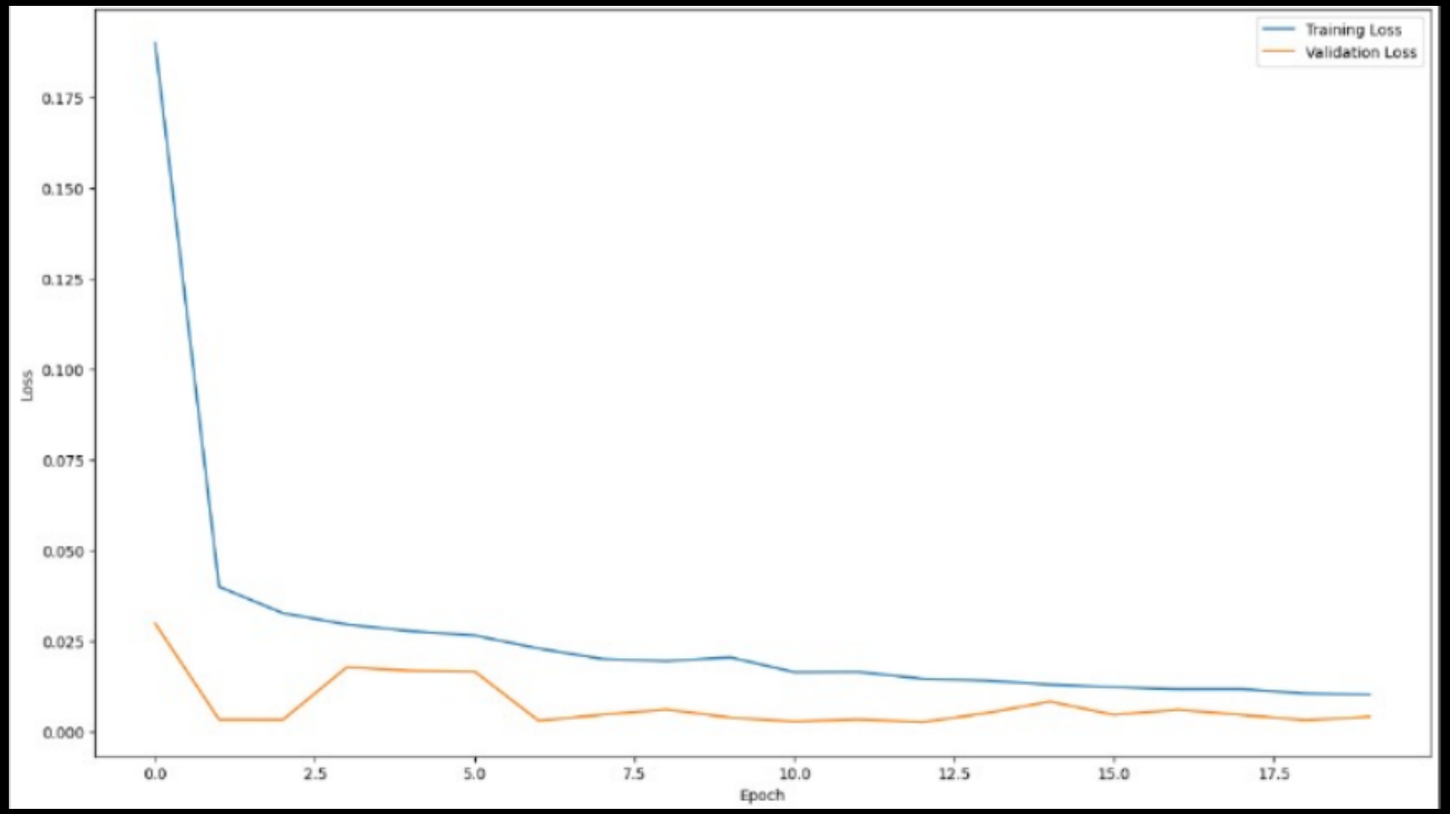
Predictions vs Orignal Values using ARIMA

An autoregressive integrated moving average, or ARIMA, is a statistical analysis model that uses time series data to either better understand the data set or to predict future trends.

The model is specified by three parameters: p (autoregressive order), d (degree of differencing), and q (moving average order). ARIMA is particularly effective for capturing linear trends and seasonality in time series data

Using Deep Learning

Bi-LSTM



Training and Validation Loss During Training

In the context of trading predictions, LSTMs can be valuable because they can effectively analyze historical price data and learn complex patterns and trends. The ability to capture long-term dependencies makes LSTMs suitable for predicting stock prices

BiLSTM, or Bidirectional Long Short-Term Memory, is an extension of the traditional LSTM architecture. The key difference is that BiLSTMs process the input sequence in both forward and backward directions, capturing intricate patterns more effectively

Bi-LSTM

R2 Score = 0.801

Predictions vs Orignal Values using LSTM

layer type	units	activation
Bi-LSTM	128	relu
Dropout	0.3	
Bi-LSTM	64	relu
Dropout	0.3	
Bi-LSTM	32	relu
Dropout	0.3	
Bi-LSTM	16	relu
Dropout	0.3	
Bi-LSTM	1	linear

LSTM Architecture

Back Testing

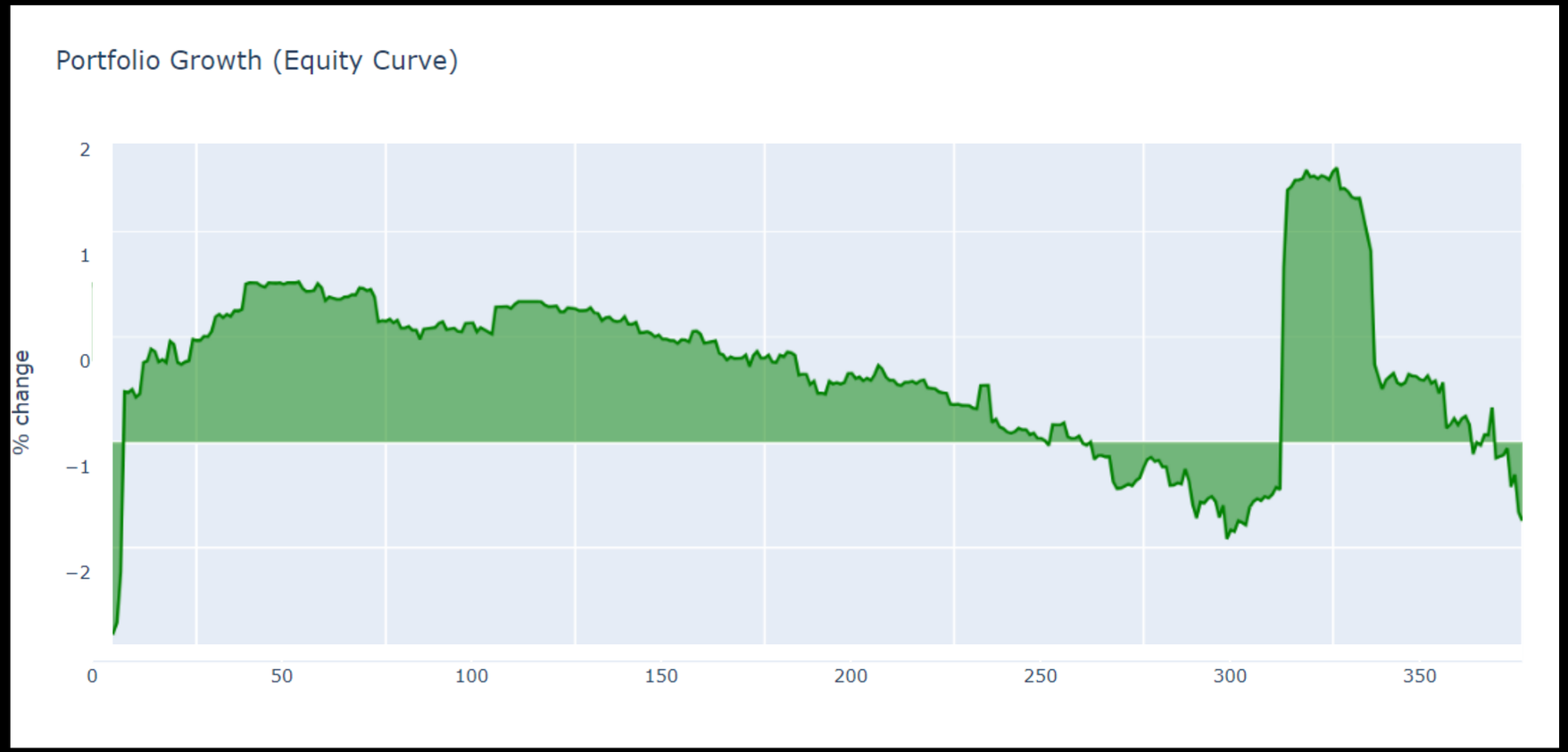
It is a method for testing how well a strategy or model would have played out using historical data.

Metrics

START	01-01-2018 05:30:00 AM
END	31-01-2022 05:30:00 AM
DURATION	1491 Days, 4h Data
INITIAL EQUITY	100,000
BUY & HOLD	97%
FINAL EQUITY	187,230
GROSS PROFIT	197094
NET PROFIT	154756
GROSS LOSS	42338
LARGEST WINNING TRADE	3190.30

LARGEST LOSING TRADE	2098.13
SHARPE RATIO	1.23
SORTINO RATIO	7.39
MAX DRAWDOWN	3.79%
AVG DRAWDOWN	2.58%
AVG WINNING TRADE	2034.49
AVG LOSING TRADE	1243.24
TOTAL CLOSED TRADES	2096

Equity Curve



The equity curve represents the cumulative performance of an investment over time.

Portfolio Volatility

Portfolio volatility is a measure of the degree of variation of a trading portfolio's returns using standard deviation..

Risk Managment

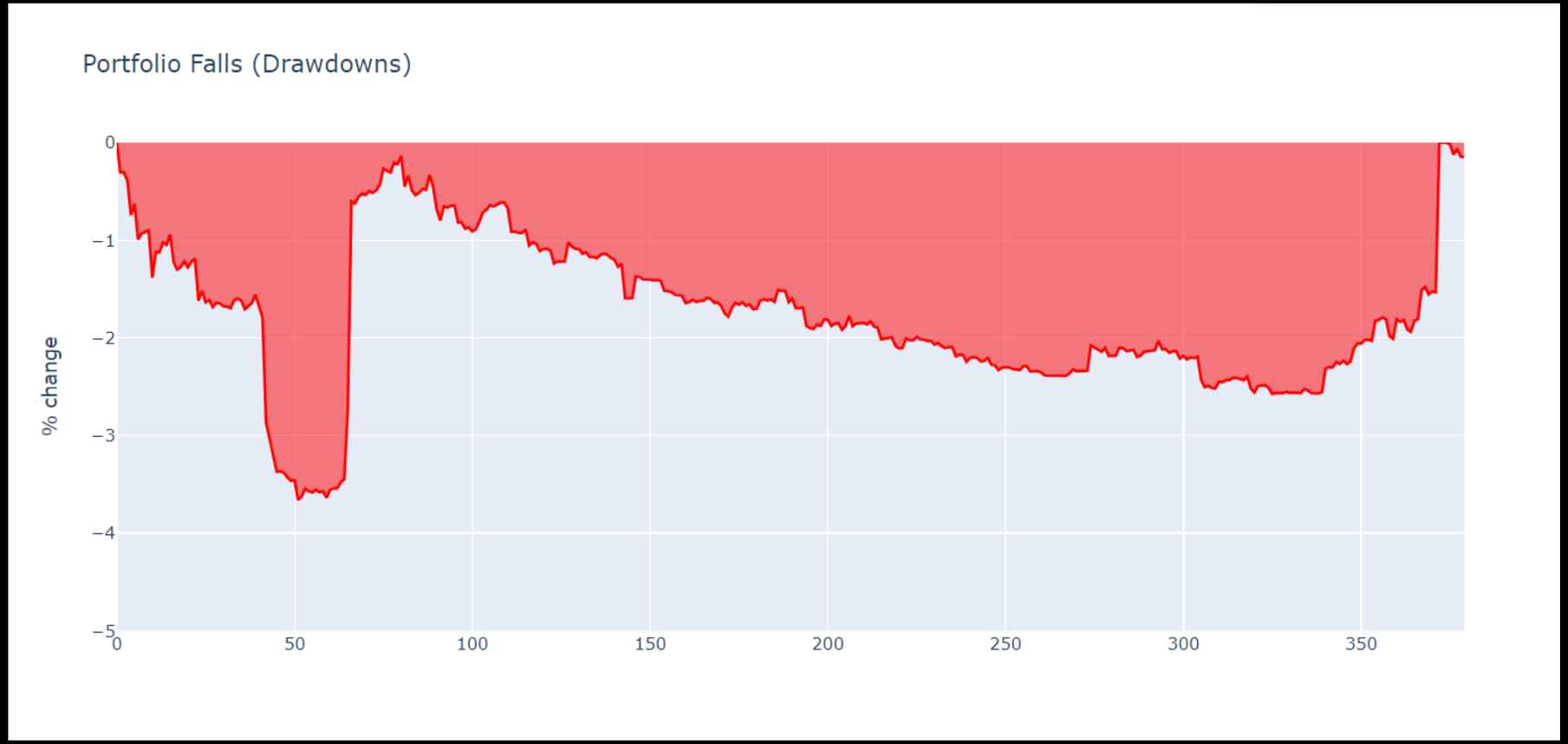
The process of measuring the size of your potential losses against the original profit potential on each new position within the financial markets

Stop Loss Strategy

A stop-loss is a risk management order placed with a broker or trading platform to sell a security automatically when its price falls to a certain predetermined level. The purpose of a stop-loss order is to limit potential losses on a trade by triggering a sell order once the asset's price reaches or falls below a specified threshold.

For our case, we have kept the max allowable percentage as 5%.

Drawdown



It measures the largest loss from a peak to a trough during a specific period

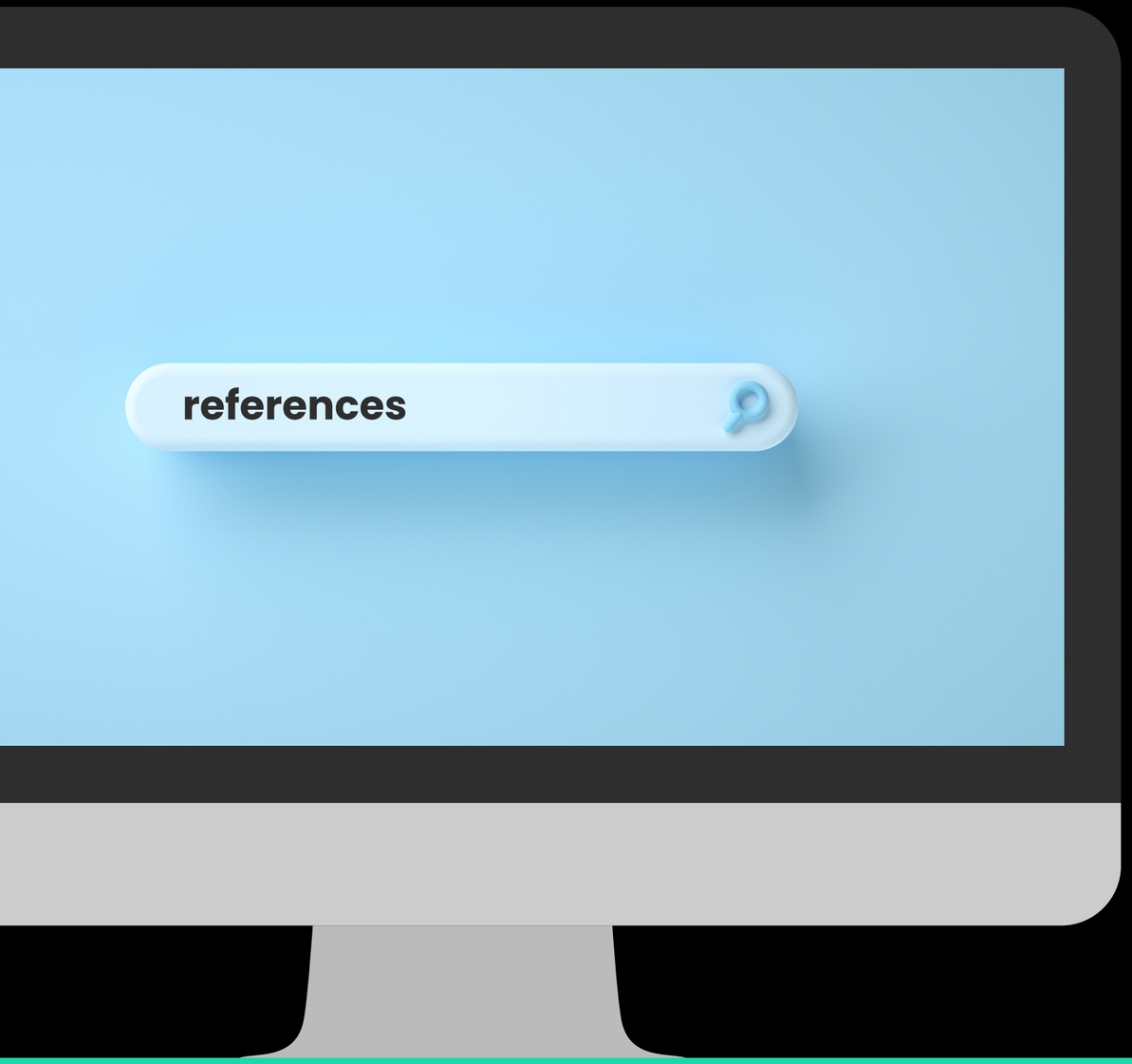
Our Team

Aadya Dewangan

Arpita Kargaonkar

Aditya Mukherjee

Thank you!



Investopedia
Medium
Youtube
Trade With Python