



Apply classification of Gradient Boosting to identify the trend of Cryptocurrency

Hung Lam, Phung Ly



GRADIENT BOOSTING MODEL

- A method standing out for its prediction speed and accuracy, particularly with large and complex datasets
- From Kaggle competitions to machine learning solutions for business, this algorithm has produced the best results.
- Gradient boost algorithm helps us minimize bias error of the model



Why preprocessing data importance need to do first?

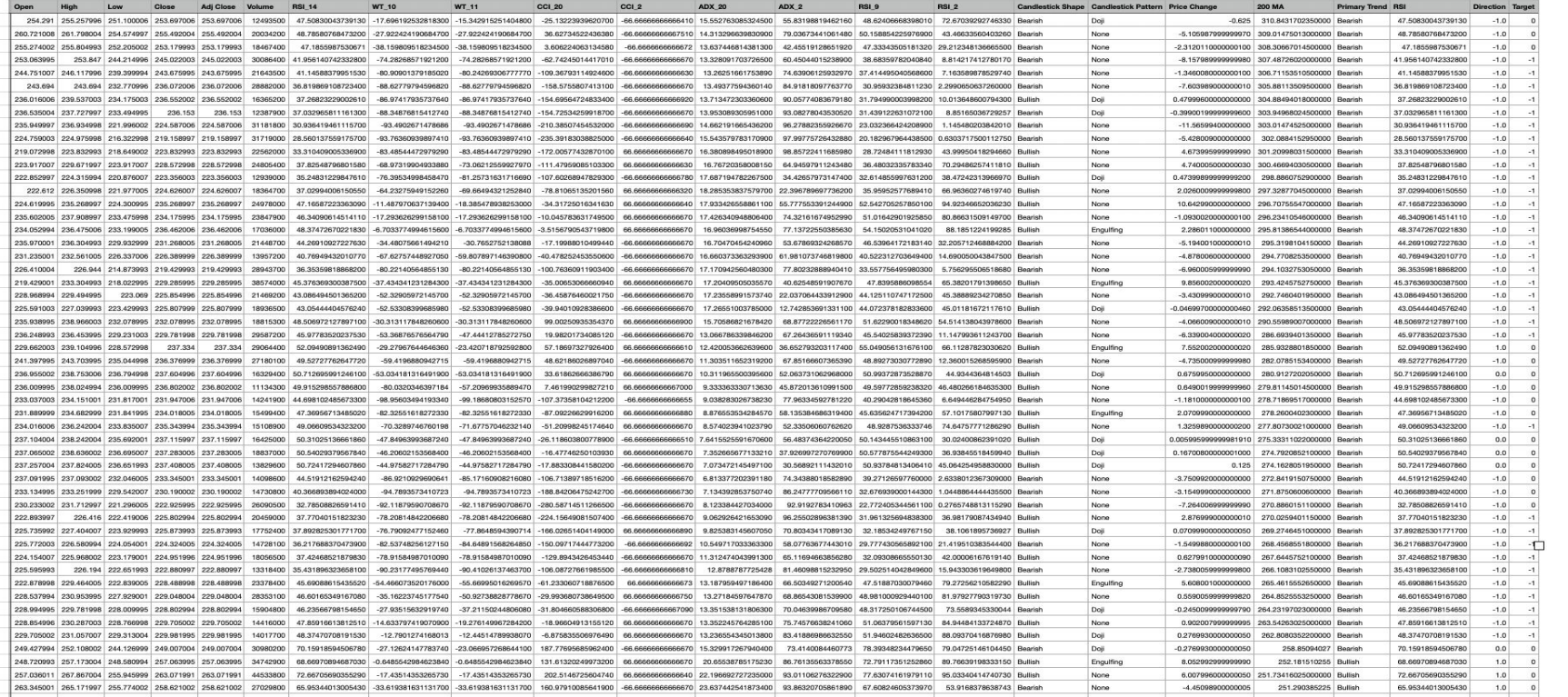
- Convert raw data into suitable form
- Store raw data in the file -> refine or preprocessing data before passing to the machine learning model for the training purpose so that the machine learning can understand data and give better performance



Dataset Example

btc_usd

Date	Open	High	Low	Close	Adj Close	Volume
2014-09-17	465.864014	468.174011	452.421997	457.334015	457.334015	21056800
2014-09-18	456.859985	456.859985	413.104004	424.440002	424.440002	34483200
2014-09-19	424.102997	427.834991	384.532013	394.795990	394.795990	37919700
2014-09-20	394.673004	423.295990	389.882996	408.903992	408.903992	36863600
2014-09-21	408.084991	412.425995	393.181000	398.821014	398.821014	26580100
2014-09-22	399.100006	406.915985	397.130005	402.152008	402.152008	24127600
2014-09-23	402.092010	441.557007	396.196991	435.790985	435.790985	45099500
2014-09-24	435.751007	436.112000	421.131989	423.204987	423.204987	30627700
2014-09-25	423.156006	423.519989	409.467987	411.574005	411.574005	26814400
2014-09-26	411.428986	414.937988	400.009003	404.424988	404.424988	21460800
2014-09-27	403.556000	406.622986	397.372009	399.519989	399.519989	15029300
2014-09-28	399.471008	401.016998	374.332001	377.181000	377.181000	23613300
2014-09-29	376.928009	385.210999	372.239990	375.467010	375.467010	32497700





Imagine we have a dataset that contains information about the price of a cryptocurrency, in this case, Bitcoin. We want to predict whether the price will go up, down, or stay the same. To do that, we need to analyze different features in the dataset.

First, we load the dataset and select the features that we think are important for prediction, such as the closing price, candlestick shape, candlestick pattern, 200-day moving average, primary trend, RSI (Relative Strength Index), and direction.

Since some of these features are categorical (e.g., candlestick shape and pattern), we encode them into numerical values so that our model can understand them.

Next, we scale the numerical features (closing price, 200-day moving average, and RSI) using a process called standardization. This ensures that all features are on a similar scale, making it easier for the model to learn.

We then split the data into training and testing sets. The training set is used to train our model, while the testing set is used to evaluate its performance.

Now comes the exciting part. We create a Gradient Boosting model, which is a machine learning algorithm known for its ability to make accurate predictions. We want to find the best configuration for this model, so we use a technique called grid search to explore different combinations of hyperparameters (settings) and select the ones that give us the best results.

Once we have the best model, we make predictions on the testing set and even predict the target for tomorrow based on the features of the last day in the dataset.

We evaluate the performance of our model by comparing the predicted values with the true target values in the testing set. This gives us metrics like precision, recall, and F1-score, which tell us how well our model is performing.

Finally, we save our best model and export the testing set with predictions and true target values to a CSV file for further analysis if needed.

In summary, this script shows how we preprocess and analyze data using machine learning techniques to predict the future price movement of a cryptocurrency.

MODEL RESULT

70%

View

Zoom

Sheet 1

Add Category

Pivot Table

Insert

Table

Chart

Text

Shape

Media

Comment

Collaborate

Format

Organize

Close	200 MA	RSI	Direction	Candlestick Shape, Bearish	Candlestick Shape, Bullish	Candlestick Pattern, Doji	Candlestick Pattern, Engulfing	Candlestick Pattern, Hammer	Candlestick Pattern, None	Primary Trend, Bearish	Primary Trend, Bullish	Prediction	True Target
0.4622695935978400	0.4064551137334400	2.184895958347900	1.0	1	0	1	0	0	0	0	0	1	1
-0.0284688170211940	-0.1702712168817870	-0.2819413232175500	1.0	0	1	0	1	0	1	0	1	1	1
0.16424411960963000	0.4117062713796880	-0.14026720032030600	1.0	0	1	0	1	0	0	1	0	0	0
-0.3711886135400270	-0.562325425833110	0.086654646466020	1.0	0	1	0	0	1	0	1	1	1	1
-0.28844633829309000	0.27441013341444500	-0.891116050321960	1.0	0	0	1	0	1	0	0	1	1	0
0.261816436110620	2.116501797907300	-0.172311168991870	1.0	0	1	0	1	0	1	0	1	1	1
1.2378565401673600	1.87692181582560	-0.5227593270065160	-1.0	1	0	1	0	1	0	1	0	-1	-1
2.1698497621927300	2.139036021228460	0.4472143002277700	1.0	1	0	1	0	0	0	0	1	1	1
-0.4906055651041400	-0.3647597122508920	-1.187416644937760	-1.0	0	1	0	0	0	1	0	0	-1	-1
0.813043039830300	0.502129234812290	-0.74987617737258	0.0	1	0	1	0	1	0	0	1	0	0
-0.63929111343228	-0.7620479788112340	-0.02014160209484200	-1.0	0	1	0	1	0	1	0	1	0	1
0.2941371885357100	0.8504853361014630	-0.767473525125450	-1.0	0	0	1	0	0	1	1	0	-1	-1
-0.862667109577980	-0.8496210747200720	-1.139311242785680	-1.0	0	1	0	0	1	1	0	1	-1	-1
-0.2852516665187260	-0.3746715070734680	-0.10188513565525900	0.0	0	1	0	0	0	0	0	1	0	0
-0.7654732602818870	-0.800374282681960	1.456529246502960	1.0	0	1	1	0	0	0	0	1	1	1
-0.861845023119020	-0.850217665472350	0.7738471265552060	1.0	0	1	0	1	0	1	1	1	1	1
-0.8622918144827410	-0.8818225569823470	-0.9940150048787030	-1.0	1	0	0	0	0	1	1	0	0	0
1.499845886816800	1.7663820753048000	-1.274745616232020	-1.0	0	1	0	1	0	0	1	0	-1	-1
1.531049040653300	2.343208661151780	-0.625078006932180	-1.0	0	1	0	0	1	1	1	0	0	-1
0.304061622758190	1.587077585130300	-1.759595658485420	-1.0	1	0	1	0	0	0	0	0	-1	-1
-0.707360520234586	-0.761705853932584	0.2337788542120630	1.0	0	1	0	0	0	0	1	1	1	1
-0.8115643600524600	-0.286931548481423	-1.728648789824960	-1.0	1	0	1	1	0	1	1	0	-1	-1
-0.2978738111094630	-0.31668686039142100	-0.0603678207130260	1.0	0	1	0	0	1	0	1	1	1	1
-0.505091919293010	-0.368542156276160	0.9031281531991270	1.0	0	1	0	1	0	0	0	1	1	1
-0.841284139139190	-0.852185142478940	-0.397103812905140	0.0	0	1	1	0	0	0	0	1	0	0
0.5752629672588200	0.419594167931213	-0.6746264150251870	0.0	0	1	0	1	0	1	0	1	0	0
-0.65366196931890	-0.4838128996720180	-1.042682632429690	-1.0	1	0	0	0	1	1	1	0	-1	-1
-0.850628530512110	-0.8423257503826870	-0.6748301856054510	0.0	1	0	1	0	0	0	0	1	1	1
-0.148003128334000	-0.2884182349108	0.7184181930275730	1.0	0	1	0	0	1	0	0	1	1	1
-0.813269183832780	-0.825731641761960	-1.569237110350260	-1.0	0	1	0	0	1	0	1	1	-1	-1
-0.371321038616930	-0.3056243584844300	-1.4572137066944400	-1.0	1	0	1	0	0	1	0	1	-1	-1
0.583641853151000	-0.04350275119681220	1.168226612007100	1.0	0	1	0	0	0	0	0	1	1	1
-0.8623719135687630	-0.8499310576313970	-0.8147764786625110	-1.0	0	1	1	0	0	1	0	0	-1	-1
-0.840694848386310	-0.847851078001500	1.158047348727360	1.0	0	1	0	0	1	0	1	1	1	1
2.50754857717084	2.12425006296780	-0.6782682743814420	1.0	1	0	1	0	1	1	0	1	1	1
2.48632636078772	2.066147396320380	0.79125448042009	1.0	1	0	1	0	0	0	0	1	1	1
-0.1568225872848720	-0.20663106212417200	0.7915671261856340	1.0	0	1	0	0	1	0	1	1	1	1
0.609090587022270	0.418425334290000	0.3772390895918800	1.0	0	1	0	0	1	0	1	1	0	0
-0.813096402277740	-0.826270189501062	0.2334627252293060	1.0	0	1	0	0	1	0	1	1	1	1
-0.80487477905110	-0.816268790543202	1.214813146682220	1.0	0	1	1	0	0	1	1	1	1	1
-0.4554145682751300	-0.340002914769910	-0.371077298102080	-1.0	1	0	1	0	0	1	0	0	0	0
-0.0448350344051780	-0.2044889993985200	1.3279945159666100	0.0	1	0	0	0	1	0	1	0	1	1
0.398335228183910	1.503346826068900	-1.43324833615060	-1.0	1	0	1	0	0	1	1	0	-1	-1
-0.7043342874698940	-0.705469741418410	0.109807099202860	1.0	0	1	0	0	1	0	0	1	1	1
0.957461132769890760	-1.75151697872150	1.2347123186356200	1.0	0	1	0	0	1	0	1	0	1	1
0.3027665334802000	0.7970566676451680	-0.5672019717563100	-1.0	0	1	0	0	1	0	1	0	-1	-1
-0.2660154177727200	-0.28691778155680500	0.2640262672957380	1.0	0	1	0	0	1	0	1	0	1	1
-0.6549342171354400	-0.449374089391190	-1.579148108620703	-1.0	0	1	0	0	1	0	1	0	-1	-1
-0.8474755050509740	-0.862974959789030	1.797390545428140	-1.0	1	0	1	0	0	1	1	0	0	0
2.9936378978664700	2.8032684544372860	0.6412276337153296	1.0	0	1	0	0	1	0	0	1	1	1
-0.478890682932380	-0.358841821208332	-0.5003688291037130	-1.0	1	0	0	0	1	1	0	0	-1	0
-0.8187113480991970	-0.8185929203365300	-0.10607762353977100	1.0	0	1	0	0	1	0	0	1	1	1
1.884043648309300	2.123278709884600	-0.75389761858050	-1.0	1	0	1	0	0	1	1	0	0	0
-0.84167810904229	-0.836281484971950	1.470723205399800	-1.0	0	1	0	0	1	0	1	1	1	1
-0.859688692310900	-0.8602217599635960	1.0291824853720220	1.0	0	1	0	0	1	0	1	1	1	1
-0.1288511704813800	-0.24124929514907400	0.4832299824582890	1.0	0	1	0	0	1	0	0	1	1	1
0.3322053309878820	0.08676158415008010	-0.2771783967138130	1.0	0	1	0	0	1	0	0	1	1	1
-0.8391275667150480	-0.8318467621991850	0.058769835248892	1.0	0	1	0	0	1	0	1	1	1	1
0.86209614120043	0.86209614120043	-0.8171096405811718	-1.0	1	0	1	0	0	1	0	0	-1	0
0.8405937272412720	0.8405937272412720	0.3019527272727200	1.0	0	1	0	0	1	0	1	1	1	1

Sheet Name

Sheet 1

Background

Duplicate Sheet

Delete Sheet



THANK YOU!!!