Using Algo

July 8, 2020

1 Examples of how to use the build_algo.py script

[44]:
```
import sys
sys.path.insert(1, '../scripts/')

from build_algo import DataAnlysis, Modelling
import pandas as pd
from sklearn.metrics import confusion_matrix
from sklearn.metrics import f1_score
```

[6]:
```
STOCK_NAME = "BTC-USD"
```

1.1 Data Analysis

[8]:
```
da = DataAnlysis(STOCK_NAME)
```

[9]:
```
da.create_laggedFeatures('Close', 1)
```

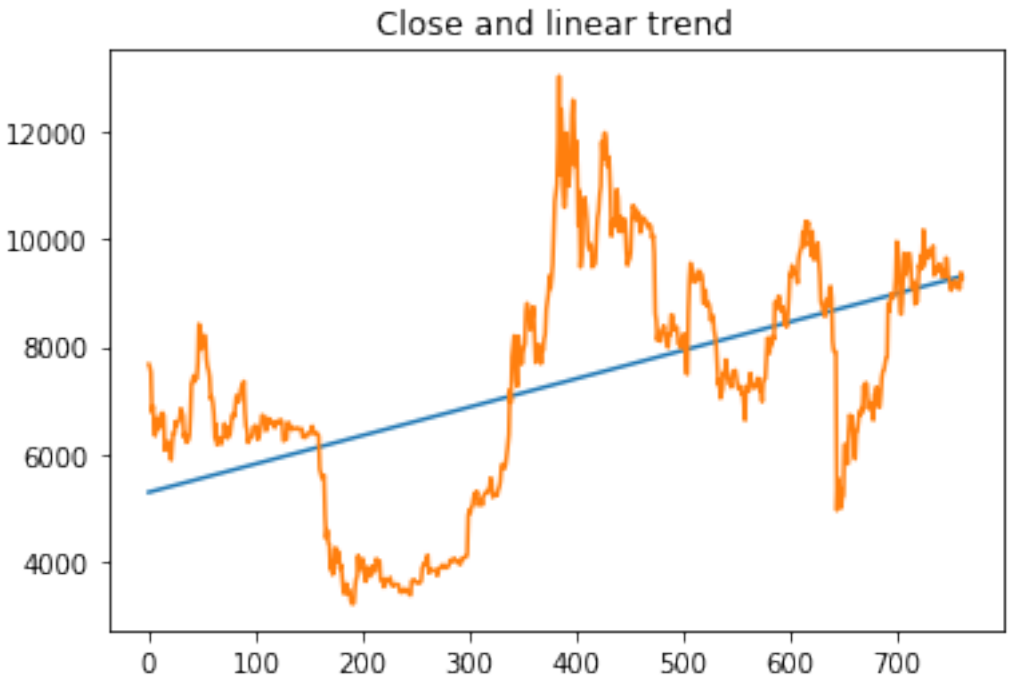
[9]:

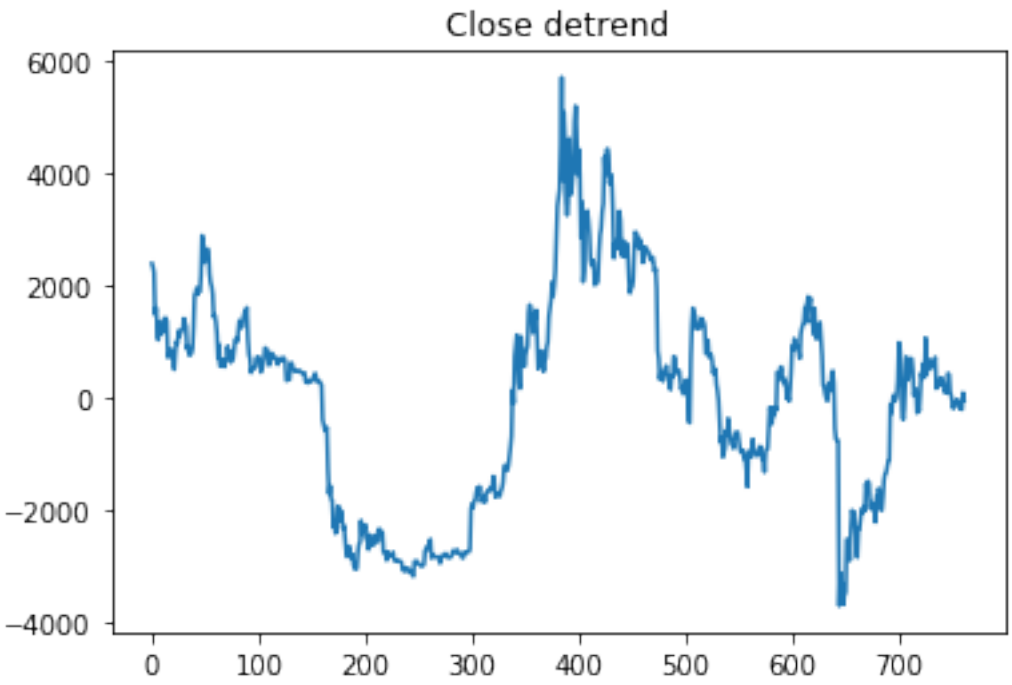
	Date	Open	High	Low	Close	Volume	Dividends	Stock Splits	Close_lagged
0	2018-06-07	7650.82	7741.27	7650.82	7678.24	4485799936	0	0	NaN
1	2018-06-08	7685.14	7698.19	7558.40	7624.92	4227579904	0	0	7678.24
2	2018-06-09	7632.52	7683.58	7531.98	7531.98	3845220096	0	0	7624.92
3	2018-06-10	7499.55	7499.55	6709.07	6786.02	5804839936	0	0	7531.98
4	2018-06-11	6799.29	6910.18	6706.63	6906.92	4745269760	0	0	6786.02
..	...	...	...	...	...	...	...	...	...
757	2020-07-03	9124.84	9202.34	9058.79	9087.30	13078970999	0	0	
758	2020-07-04	9084.23	9183.30	9053.63	9132.49	12290528515	0	0	
759	2020-07-05	9126.09	9162.18	8977.02	9073.94	12903406143	0	0	
760	2020-07-06	9072.85	9375.47	9058.66	9375.47	17889263252	0	0	
761	2020-07-07	9323.49	9359.54	9216.21	9249.49	14392401920	0	0	

..	...	...
757	0	9123.41
758	0	9087.30
759	0	9132.49
760	0	9073.94
761	0	9375.47

[762 rows x 9 columns]

[10]:
```
da.detrend('Close')
```





[10]: array([2.37191244e+03, 2.31333237e+03, 2.21513230e+03, 1.46391223e+03,
1.57955216e+03, 1.24973209e+03, 1.01201202e+03, 1.33220195e+03,
1.10817188e+03, 1.19649182e+03, 1.14034175e+03, 1.37063168e+03,
1.40049161e+03, 1.40184154e+03, 1.34977147e+03, 6.98461399e+02,
7.71991330e+02, 7.77481261e+02, 8.48171191e+02, 6.87401122e+02,
7.45601053e+02, 4.86650983e+02, 7.96250914e+02, 9.76690845e+02,
9.53250775e+02, 1.17635071e+03, 1.08650064e+03, 1.14920057e+03,
1.18553050e+03, 1.21463043e+03, 1.39280036e+03, 1.30449029e+03,
1.26710022e+03, 8.50040151e+02, 9.09540082e+02, 7.38380013e+02,
7.42359943e+02, 7.75169874e+02, 8.53429805e+02, 1.23027974e+03,
1.80430967e+03, 1.84878960e+03, 1.93960953e+03, 1.82161946e+03,
1.88151939e+03, 1.87545932e+03, 2.16281925e+03, 2.87071918e+03,
2.62257911e+03, 2.38750904e+03, 2.59567897e+03, 2.61755890e+03,
2.63860883e+03, 2.59536877e+03, 2.19006870e+03, 2.02927863e+03,
1.96625856e+03, 1.82823849e+03, 1.42143842e+03, 1.45180835e+03,
1.32986828e+03, 1.12592821e+03, 6.73348141e+02, 9.30518072e+02,
5.41738003e+02, 6.47497933e+02, 6.69197864e+02, 6.38817795e+02,
5.35697725e+02, 6.39247656e+02, 6.60197587e+02, 9.00837517e+02,
7.38707448e+02, 8.15757379e+02, 6.12957309e+02, 7.87927240e+02,
6.70617171e+02, 8.23527101e+02, 1.00334703e+03, 1.04131696e+03,
9.80126893e+02, 1.15224682e+03, 1.35862675e+03, 1.30424669e+03,
1.23005662e+03, 1.28414655e+03, 1.43455648e+03, 1.50876641e+03,
1.49084634e+03, 1.58718627e+03, 1.01309620e+03, 7.44176131e+02,
6.76816062e+02, 4.30465992e+02, 5.00085923e+02, 5.23665854e+02,
5.09905784e+02, 5.35245715e+02, 6.95495646e+02, 6.85635576e+02,

7.10865507e+02, 6.79585438e+02, 4.38345368e+02, 5.23185299e+02,
5.45165230e+02, 6.61035160e+02, 8.71055091e+02, 8.52825022e+02,
8.36214953e+02, 7.15734883e+02, 5.61534814e+02, 6.04804745e+02,
7.81294675e+02, 7.43414606e+02, 6.95984537e+02, 7.14324467e+02,
6.73124398e+02, 6.34344329e+02, 5.75574259e+02, 6.44414190e+02,
6.84944121e+02, 6.45514051e+02, 6.54893982e+02, 6.98913913e+02,
6.84063843e+02, 6.21693774e+02, 2.87143705e+02, 3.00223635e+02,
3.06373566e+02, 3.06053497e+02, 6.06403427e+02, 6.00713358e+02,
5.43773289e+02, 4.70793220e+02, 4.54233150e+02, 4.72753081e+02,
4.60653012e+02, 4.60202942e+02, 4.43522873e+02, 4.58362804e+02,
4.33552734e+02, 4.26752665e+02, 4.27122596e+02, 4.27872526e+02,
2.68852457e+02, 2.65232388e+02, 2.43312318e+02, 2.98222249e+02,
3.03622180e+02, 2.71182110e+02, 2.80792041e+02, 3.19061972e+02,
3.55151902e+02, 4.19021833e+02, 3.37341764e+02, 2.63981694e+02,
2.82321625e+02, 2.79111556e+02, 2.33851487e+02, 2.16811417e+02,
-4.09588652e+02, -5.05168721e+02, -5.82908791e+02, -6.09388860e+02,
-5.45438929e+02, -1.30274900e+03, -1.72762907e+03, -1.58258914e+03,
-1.82407921e+03, -1.84816928e+03, -2.31977935e+03, -2.19582941e+03,
-2.43192948e+03, -2.39559955e+03, -1.96415962e+03, -1.94798969e+03,
-2.21482976e+03, -2.02268983e+03, -2.10273990e+03, -2.35374997e+03,
-2.29625004e+03, -2.50441011e+03, -2.74256018e+03, -2.84898025e+03,
-2.79807032e+03, -2.66521039e+03, -2.78204045e+03, -2.86537052e+03,
-2.80827059e+03, -2.98680066e+03, -3.06326073e+03, -3.07424080e+03,
-3.06342087e+03, -2.77566094e+03, -2.63072101e+03, -2.58609108e+03,
-2.20286115e+03, -2.44602122e+03, -2.33364129e+03, -2.35410136e+03,
-2.27974142e+03, -2.54811149e+03, -2.51156156e+03, -2.71929163e+03,
-2.45546170e+03, -2.56423177e+03, -2.52395184e+03, -2.65246191e+03,
-2.55690198e+03, -2.46227205e+03, -2.57420212e+03, -2.55848219e+03,
-2.57627226e+03, -2.35009233e+03, -2.40673240e+03, -2.40639246e+03,
-2.40720253e+03, -2.76884260e+03, -2.76565267e+03, -2.79698274e+03,
-2.91059281e+03, -2.76275288e+03, -2.84338295e+03, -2.82431302e+03,
-2.80602309e+03, -2.83200316e+03, -2.76653323e+03, -2.89935330e+03,
-2.92959337e+03, -2.90630344e+03, -2.93102350e+03, -2.92053357e+03,
-2.92689364e+03, -2.92946371e+03, -2.95321378e+03, -3.07199385e+03,
-3.09958392e+03, -3.06678399e+03, -3.10043406e+03, -3.07553413e+03,
-3.04768420e+03, -3.10999427e+03, -3.12011434e+03, -3.11816441e+03,
-3.17601447e+03, -3.19557454e+03, -2.93352461e+03, -2.93436468e+03,
-2.92063475e+03, -2.96765482e+03, -2.96781489e+03, -2.99453496e+03,
-3.01498503e+03, -3.01631510e+03, -3.01259517e+03, -2.97380524e+03,
-2.73719531e+03, -2.71107538e+03, -2.66360545e+03, -2.71456551e+03,
-2.66841558e+03, -2.53667565e+03, -2.87403572e+03, -2.80702579e+03,
-2.84062586e+03, -2.84919593e+03, -2.85071600e+03, -2.85118607e+03,
-2.85160614e+03, -2.87410621e+03, -2.96498628e+03, -2.83542635e+03,
-2.83312642e+03, -2.83084649e+03, -2.84645655e+03, -2.78953662e+03,
-2.80650669e+03, -2.85813676e+03, -2.85946683e+03, -2.86716690e+03,
-2.85477697e+03, -2.82349704e+03, -2.74093711e+03, -2.76969718e+03,
-2.76767725e+03, -2.73425732e+03, -2.72322739e+03, -2.78663746e+03,

-2.79725752e+03, -2.79065759e+03, -2.80957766e+03, -2.87393773e+03,
-2.85718780e+03, -2.76045787e+03, -2.78367794e+03, -2.75967801e+03,
-2.75664808e+03, -2.76316815e+03, -2.71564822e+03, -1.99920829e+03,
-1.91132836e+03, -1.96680843e+03, -1.85818850e+03, -1.84030856e+03,
-1.70648863e+03, -1.62087870e+03, -1.71094877e+03, -1.59661884e+03,
-1.86193891e+03, -1.84214898e+03, -1.84035905e+03, -1.77448912e+03,
-1.88035919e+03, -1.71716926e+03, -1.70604933e+03, -1.66485940e+03,
-1.66469947e+03, -1.63587954e+03, -1.66449960e+03, -1.58491967e+03,
-1.41718974e+03, -1.52993981e+03, -1.78954988e+03, -1.72597995e+03,
-1.74230002e+03, -1.73071009e+03, -1.77376016e+03, -1.67564023e+03,
-1.62893030e+03, -1.53161037e+03, -1.27386044e+03, -1.21624051e+03,
-1.25696058e+03, -1.31112064e+03, -1.23369071e+03, -1.08599078e+03,
-8.99180852e+02, -7.00120922e+02, 1.20539009e+02, -1.17121060e+02,
7.20168870e+02, 8.94408801e+02, 1.09989873e+03, 7.74378662e+02,
2.28108593e+02, 1.50158524e+02, 1.07137845e+03, 8.46738385e+02,
8.26498316e+02, 5.37978247e+02, 7.34498177e+02, 8.34758108e+02,
8.94668039e+02, 1.51008797e+03, 1.63738790e+03, 1.54630783e+03,
1.48057776e+03, 1.13529769e+03, 1.38506762e+03, 1.36932755e+03,
1.54300748e+03, 1.00378741e+03, 4.97297345e+02, 6.08497276e+02,
6.01027207e+02, 8.17697137e+02, 7.22617068e+02, 4.51306999e+02,
7.58296929e+02, 6.80416860e+02, 8.93306791e+02, 9.73106722e+02,
1.43075665e+03, 1.57004658e+03, 1.72089651e+03, 2.04149644e+03,
1.79764637e+03, 1.98414631e+03, 2.23252624e+03, 2.84466617e+03,
3.39653610e+03, 3.54495603e+03, 3.69542596e+03, 4.46998589e+03,
5.69003582e+03, 3.85135575e+03, 5.07061568e+03, 4.61739561e+03,
3.46992554e+03, 3.23063547e+03, 3.44392540e+03, 4.59825534e+03,
3.84716527e+03, 3.60492520e+03, 3.82975513e+03, 4.06679506e+03,
4.89664499e+03, 5.17923492e+03, 4.75667485e+03, 3.95356478e+03,
4.40563471e+03, 3.97676464e+03, 2.83518457e+03, 3.46895450e+03,
2.04624443e+03, 2.25714436e+03, 3.22456430e+03, 3.08355423e+03,
3.31470416e+03, 3.14141409e+03, 2.88015402e+03, 2.43255395e+03,
2.33845388e+03, 2.43310381e+03, 2.38630374e+03, 1.98842367e+03,
2.05834360e+03, 2.01937353e+03, 2.10238346e+03, 2.57533339e+03,
2.88411332e+03, 2.99735326e+03, 3.29565319e+03, 3.43884312e+03,
4.26905305e+03, 3.93631298e+03, 4.39485291e+03, 4.41403284e+03,
4.30530277e+03, 3.79112270e+03, 3.95542263e+03, 3.80920256e+03,
3.31715249e+03, 2.46776242e+03, 2.72235235e+03, 2.77988229e+03,
2.63202222e+03, 2.74083215e+03, 3.30581208e+03, 3.14773201e+03,
2.51729194e+03, 2.50504187e+03, 2.77668180e+03, 2.52342173e+03,
2.49672166e+03, 2.72376159e+03, 2.53318152e+03, 2.09684145e+03,
1.84736138e+03, 1.93007131e+03, 1.95730125e+03, 2.07935118e+03,
2.66288111e+03, 2.93440104e+03, 2.90009097e+03, 2.87587090e+03,
2.64838083e+03, 2.80707076e+03, 2.72584069e+03, 2.61427062e+03,
2.39002055e+03, 2.44715048e+03, 2.67365041e+03, 2.61881034e+03,
2.61105027e+03, 2.59545021e+03, 2.51927014e+03, 2.47849007e+03,
2.43021000e+03, 2.49311993e+03, 2.40307986e+03, 2.23589979e+03,
2.28130972e+03, 1.93497965e+03, 8.20969582e+02, 6.82129512e+02,

3.08849443e+02, 4.36469374e+02, 4.25279304e+02, 2.78289235e+02,
4.62709166e+02, 5.06859096e+02, 5.51359027e+02, 4.13048958e+02,
3.53738888e+02, 2.94038819e+02, 1.25438750e+02, 3.77638680e+02,
3.55538611e+02, 7.17238542e+02, 7.02708473e+02, 4.32738403e+02,
4.42278334e+02, 4.21468265e+02, 4.69888195e+02, 2.95308126e+02,
1.32208057e+02, 1.83327987e+02, 4.73679180e+01, 5.74578487e+01,
2.85717779e+02, 3.02097710e+02, 1.31317641e+02, -4.37472429e+02,
-4.63912498e+02, 6.98037433e+02, 1.27704736e+03, 1.57852729e+03,
1.27770722e+03, 1.44398716e+03, 1.21676709e+03, 1.20535702e+03,
1.26161695e+03, 1.31997688e+03, 1.22534681e+03, 1.39734674e+03,
1.32200667e+03, 1.33509660e+03, 1.23651653e+03, 7.68576462e+02,
7.72016393e+02, 1.00870632e+03, 7.05706254e+02, 7.58316185e+02,
7.45656116e+02, 6.40226046e+02, 4.18865977e+02, 4.72375908e+02,
4.94335838e+02, 2.20385769e+02, 1.11985700e+02, -7.21543696e+01,
-4.61934439e+02, -8.13364508e+02, -7.17404578e+02, -1.07254465e+03,
-9.79594716e+02, -9.12614785e+02, -6.04584855e+02, -6.78394924e+02,
-3.85524993e+02, -5.82395063e+02, -7.32995132e+02, -8.40555201e+02,
-8.47655271e+02, -9.21035340e+02, -7.30015409e+02, -6.36585479e+02,
-6.32605548e+02, -6.29755617e+02, -7.98465687e+02, -9.26505756e+02,
-9.92455825e+02, -9.72015895e+02, -9.50725964e+02, -1.10099603e+03,
-1.07862610e+03, -1.30370617e+03, -1.60092624e+03, -9.69906311e+02,
-1.04912638e+03, -1.03840645e+03, -1.07132652e+03, -7.56156588e+02,
-9.17376657e+02, -9.55736726e+02, -1.00836680e+03, -1.04981687e+03,
-1.00395693e+03, -9.81317004e+02, -8.81917073e+02, -1.01682714e+03,
-1.12148721e+03, -1.12017728e+03, -1.34013735e+03, -9.85987420e+02,
-9.25467489e+02, -9.30067558e+02, -5.77427628e+02, -1.88217697e+02,
-2.77307766e+02, -4.83357836e+02, -2.01137905e+02, -3.35407974e+02,
-1.85718043e+02, -2.39278113e+02, 4.39031818e+02, 4.13021749e+02,
3.24541679e+02, 5.24531610e+02, 5.33041541e+02, 2.91221471e+02,
2.37351402e+02, 3.20341333e+02, 2.50071263e+02, -2.95488060e+01,
4.10112468e+00, -7.87389446e+01, 1.44980986e+02, 4.52710917e+02,
8.96220847e+02, 8.49000778e+02, 1.03610071e+03, 8.72380639e+02,
9.09470570e+02, 8.55700501e+02, 7.99590431e+02, 6.81770362e+02,
1.10897029e+03, 1.22009022e+03, 1.28097015e+03, 1.34489008e+03,
1.59118002e+03, 1.32585995e+03, 1.67222988e+03, 1.78477981e+03,
1.66784974e+03, 1.76032967e+03, 1.33236960e+03, 1.37211953e+03,
1.12256946e+03, 1.56916939e+03, 1.05529932e+03, 1.02512925e+03,
1.09782918e+03, 1.06930911e+03, 1.32538905e+03, 1.04577898e+03,
7.32058906e+02, 2.05608837e+02, 1.64318768e+02, 4.70286985e+01,
-3.11813708e+01, -7.35014401e+01, 2.28458491e+02, 1.41318421e+02,
1.03518352e+02, 4.21768283e+02, 4.60298213e+02, 2.42438144e+02,
-5.64651925e+02, -7.54391995e+02, -7.73562064e+02, -7.77122133e+02,
-3.72302220e+03, -3.13536227e+03, -3.50396234e+03, -3.31728241e+03,
-3.70037248e+03, -3.49448255e+03, -3.48693262e+03, -2.53944269e+03,
-2.53711276e+03, -2.55608283e+03, -2.91616290e+03, -2.33536297e+03,
-2.02213303e+03, -2.08113310e+03, -2.05101317e+03, -2.30291324e+03,
-2.53578331e+03, -2.86119338e+03, -2.35865345e+03, -2.35511352e+03,

```
       -2.19223359e+03, -2.01065366e+03, -2.07614373e+03, -1.94726380e+03,
       -2.02892387e+03, -1.55353394e+03, -1.65416400e+03, -1.50173407e+03,
       -1.53900414e+03, -1.98086421e+03, -1.99253428e+03, -1.88578435e+03,
       -2.01709442e+03, -2.02496449e+03, -2.23054456e+03, -1.76111463e+03,
       -1.78699470e+03, -1.63076477e+03, -1.70427484e+03, -2.01699491e+03,
       -2.02389498e+03, -1.79226504e+03, -1.48501511e+03, -1.36909518e+03,
       -1.35531525e+03, -1.25064532e+03, -1.14017539e+03, -1.13397546e+03,
       -1.45255530e+02, -2.93005599e+02, -9.20456686e+01,  2.65242621e+01,
       -6.98658072e+01, -5.99458766e+01,  2.52140541e+01,  2.85643985e+02,
        9.63143915e+02,  8.49033846e+02,  5.95003777e+02, -2.47726292e+02,
       -4.07616362e+02, -2.10196431e+02,  2.50053500e+02,  7.08523430e+02,
        2.97743361e+02,  3.41293292e+02,  6.29763222e+02,  6.80343153e+02,
        6.77543084e+02,  4.66223014e+02,  1.97429450e+01,  1.15302876e+02,
        1.36752806e+02, -2.87427263e+02, -1.76127332e+02, -2.53267402e+02,
        8.74425291e+01,  4.26912460e+02,  3.35022390e+02,  5.91052321e+02,
        3.46442252e+02,  1.04739218e+03,  4.04662113e+02,  5.26322044e+02,
        6.64981975e+02,  5.24611905e+02,  5.07501836e+02,  6.07411767e+02,
        6.14791697e+02,  6.33741628e+02,  7.02871559e+02,  1.49301489e+02,
        3.03101420e+02,  2.92281351e+02,  1.98531281e+02,  2.57181212e+02,
        3.39241143e+02,  2.76211073e+02,  2.02541004e+02,  7.34609347e+01,
        1.12520865e+02,  7.85507961e+01,  4.18380727e+02,  3.94060657e+02,
        7.27505882e+01,  1.86905188e+01, -8.84595505e+01, -2.11249620e+02,
       -1.18319689e+02, -7.63097584e+01, -1.34429828e+02, -4.93498971e+01,
       -1.59529966e+02, -2.00900036e+02, -1.60970105e+02, -2.24780174e+02,
        7.14897563e+01, -5.97503130e+01])
```

[11]:
```
da.ma_smoothing(7)
```

[11]:

	Date	Open	High	Low	Close	Volume	Dividends	Stock Splits	Close_lagged	ma_7
0	2018-06-07	7650.82	7741.27	7650.82	7678.24	4485799936	0	0	NaN	NaN
1	2018-06-08	7685.14	7698.19	7558.40	7624.92	4227579904	0	0	7678.24	NaN
2	2018-06-09	7632.52	7683.58	7531.98	7531.98	3845220096	0	0	7624.92	NaN
3	2018-06-10	7499.55	7499.55	6709.07	6786.02	5804839936	0	0	7531.98	7065.762857
4	2018-06-11	6799.29	6910.18	6706.63	6906.92	4745269760	0	0	6786.02	6922.492857
..	...	...	...	...	...	...	...	...	...	...
757	2020-07-03	9124.84	9202.34	9058.79	9087.30	13078970999	0	0		
758	2020-07-04	9084.23	9183.30	9053.63	9132.49	12290528515	0	0		
759	2020-07-05	9126.09	9162.18	8977.02	9073.94	12903406143	0	0		
760	2020-07-06	9072.85	9375.47	9058.66	9375.47	17889263252	0	0		
761	2020-07-07	9323.49	9359.54	9216.21	9249.49	14392401920	0	0		

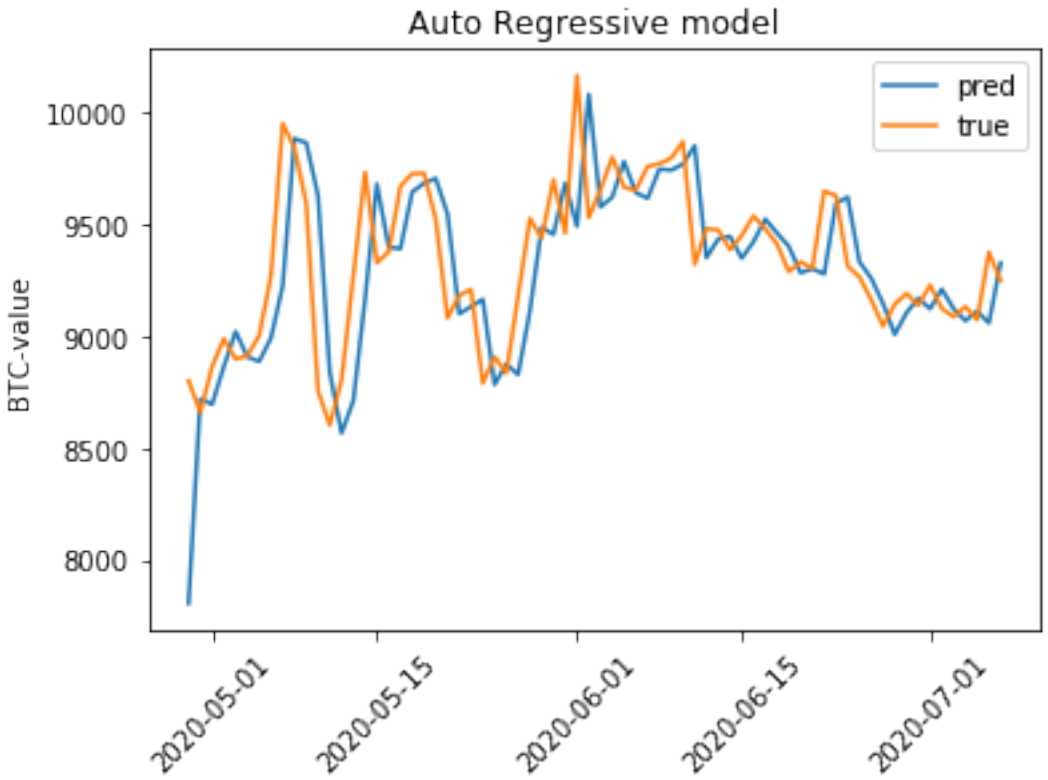
757	0	9123.41	9165.561429
758	0	9087.30	9181.490000
759	0	9132.49	NaN
760	0	9073.94	NaN
761	0	9375.47	NaN

[762 rows x 10 columns]

1.2 Modelling

[12]:
```
m = Modelling(da.df)
```

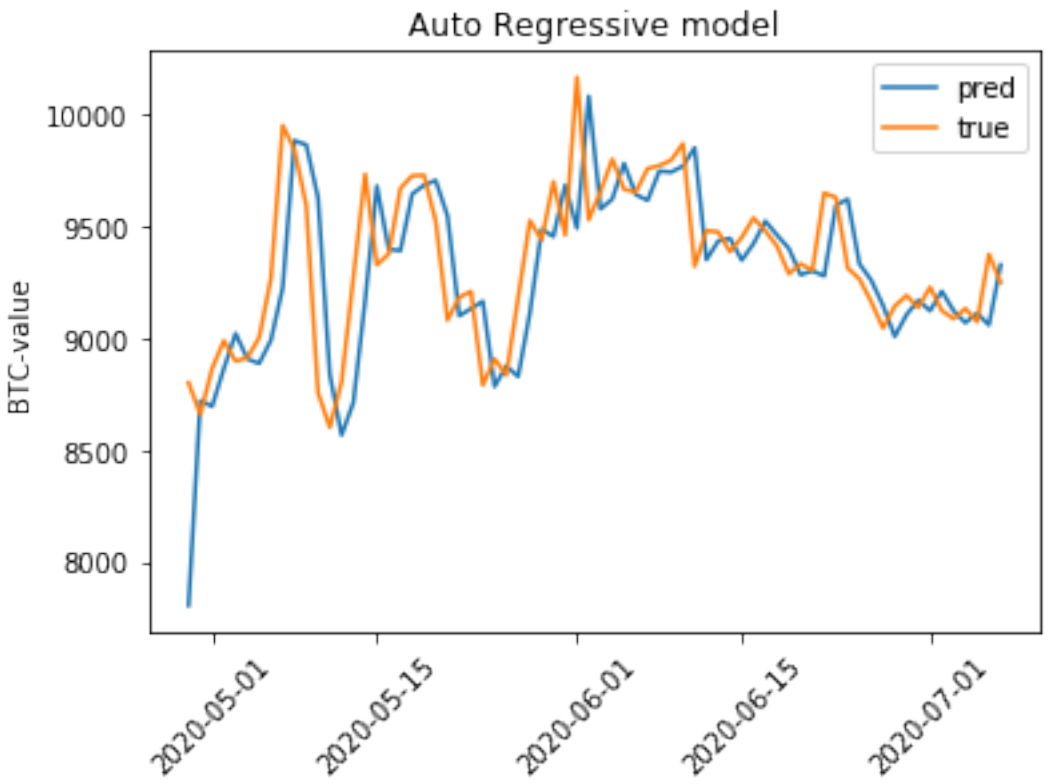
[13]:
```
m.AutoRegressive()
```



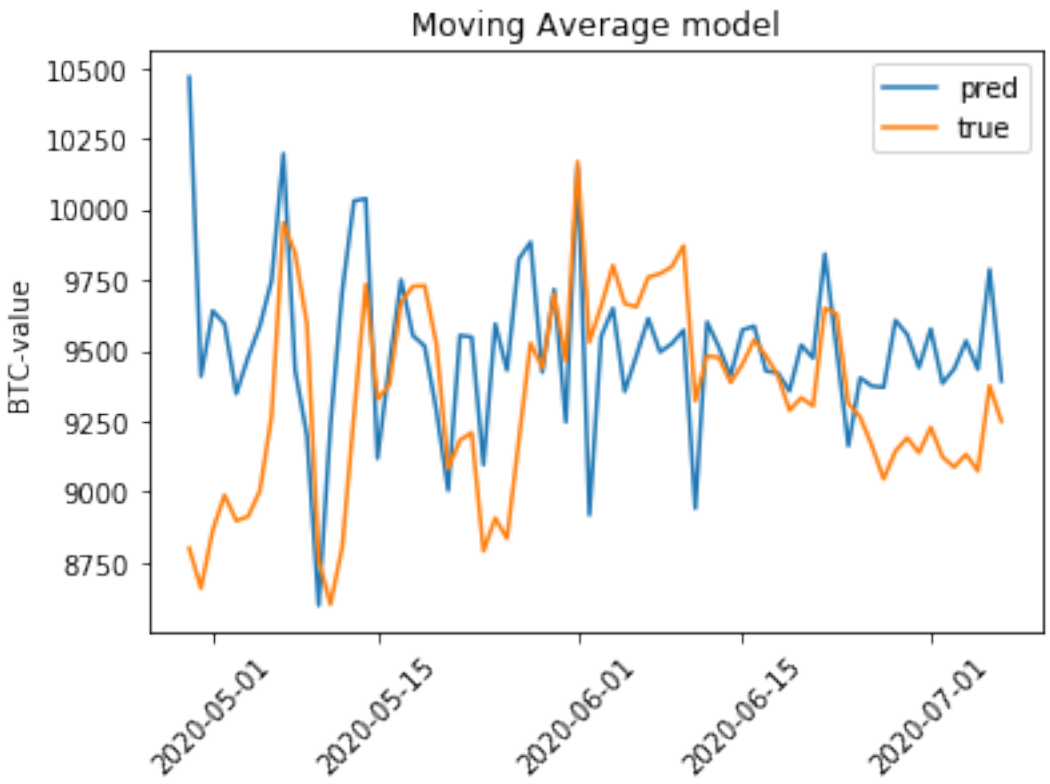
RMSE of Auto Regressive: 290.24

[13]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)

[14]:
```
m.MovingAverage()
```



RMSE of Auto Regressive: 290.24

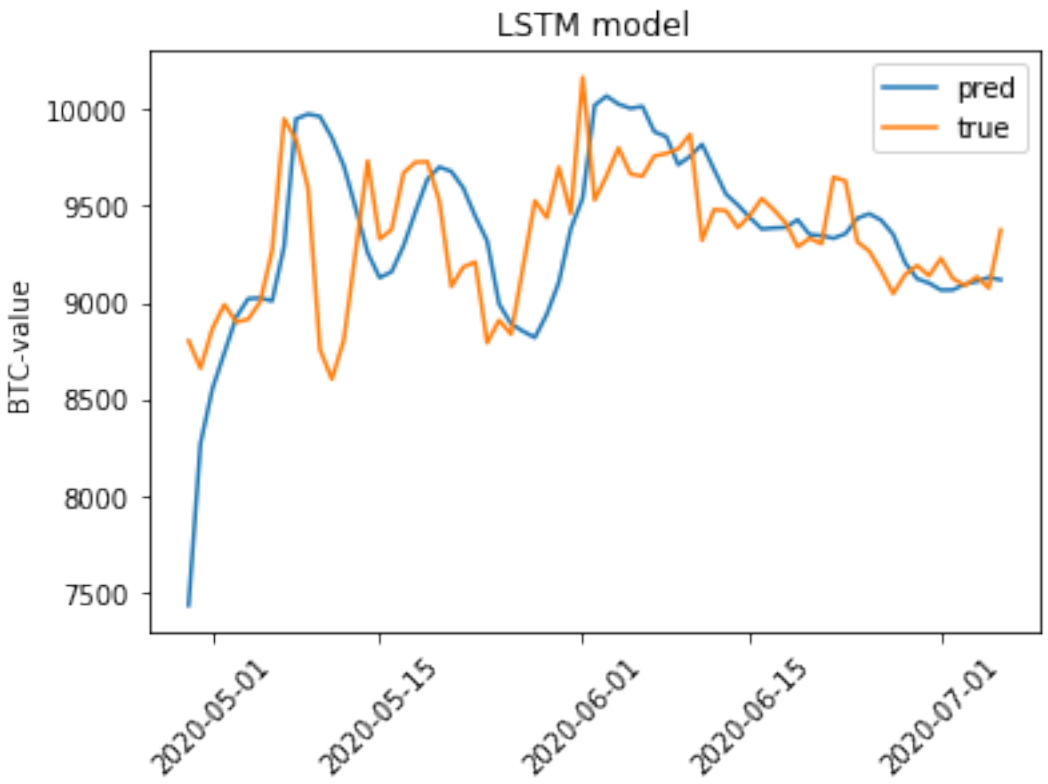


RMSE of Auto Regressive: 416.67

[14]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)

[15]:
```
y_pred = m.LSTM_model()
```

Epoch 1 completed!
Epoch 2 completed!
Epoch 3 completed!
Epoch 4 completed!
Epoch 5 completed!



RMSE of LSTM: 407.16

[20]:
```
df = da.df.copy()
```

[25]:
```
TEST_SIZE = int(len(df)*(9/10))
df_test = df[TEST_SIZE:]
y_true = df_test[7:len(df_test)-1]['Close'].values
table = pd.DataFrame({'y_pred': y_pred, 'y_true': y_true})
table['y_true_lagged'] = table['y_true'].shift(1)
table = table[1:]
table['dif_true'] = table['y_true'] - table['y_true_lagged']
table['dif_pred'] = table['y_pred'] - table['y_true_lagged']
table['sign_true'] = table['dif_true'].apply(lambda x: 1 if x>=0 else 0)
table['sign_pred'] = table['dif_pred'].apply(lambda x: 1 if x>=0 else 0)
```

[48]:
```
table.head()
```

[48]:

	y_pred	y_true	y_true_lagged	dif_true	dif_pred	sign_true	\
1	8268.144531	8658.55	8801.04	-142.49	-532.895469	0	
2	8552.465820	8864.77	8658.55	206.22	-106.084180	1	
3	8738.349609	8988.60	8864.77	123.83	-126.420391	1	
4	8924.604492	8897.47	8988.60	-91.13	-63.995508	0	
5	9019.111328	8912.65	8897.47	15.18	121.641328	1	

	sign_pred
1	0
2	0
3	0
4	0
5	1

[40]:
```
confusion_matrix(table['sign_true'], table['sign_pred'])
```

[40]:
```
array([[18, 14],
       [14, 22]])
```

[47]:
```
print('F1-score: {}'.format(round(f1_score(table['sign_true'], table['sign_pred']),2 )))
```

F1-score: 0.61