Stock Market Prediction Using a Weighted Machine Learning Model for DowJones Industrial Average Companies

Abstract

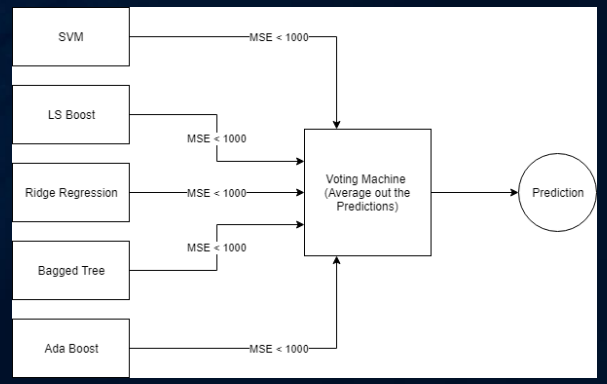
Investments into the stock market can either make one a powerful elite in a nation or a person with no value. This gamble for wealth is unpredictable as the stock market is highly volatile. Those who succeed in the stock market will be glorified for their fame and fortune. Examples of these people include Warren Buffet (CEO of Berkshire Hathaway) and Winklevoss Twins (investors of Bitcoins). Although many try to emulate these investor's investment strategies, some have begun using machine learning models to help make smart investment choices. Many of these models follow general machine learning models such as support vector machines (SVM), regressions, and trees. However, the accuracies for these models are not high enough guarantee large profits. In this research, many machine learning models will be used to predict monthly stock values for companies inside the DowJones Industrial Average. These predictions will be weighted together. The goal of this experiment is to show that weighted amounts of machine learning algorithms can improve the performance of SVMs.

Models Used

This project is inspired by the SRA (Support Vector Machine, Random Forest, Ada Boost) Voting Model from Shanghai. This model takes the predicted results the named models and do a weighted average. The results of this research showed that a weighted ensemble model has better accuracies than an SVM model. The model does fail when predicting long periods of time, so this model is best for predicting a month ahead. From doing research on other markets, Bagged Tree was a better model for the India stock market. For the New York Stock Exchange, LS Boost and ridge regression were methods with great accuracies.

Company Name	Ticker	Exchange
The 3M Company	MMM	NYSE
The American Express Company	AXP	NYSE
Apple Inc.	AAPL	NASDAQ
The Boeing Company	BA	NYSE
Caterpillar Inc.	CAT	NYSE
Chevron Corporation	CVX	NYSE
Cisco Systems, Inc.	CSCO	NASDAQ
The Coca-Cola Company	KO	NYSE
DowDuPont Inc.	DWDP	NYSE
Exxon Mobil Corporation	XOM	NYSE
The Goldman Sachs Group, Inc.	GS	NYSE
The Home Depot Inc.	HD	NYSE
International Business Machines Corporation	IBM	NYSE
Intel Corporation	INTC	NASDAQ
Johnson & Johnson	JNJ	NYSE
JPMorgan Chase & Co.	JPM	NYSE
McDonald's Corporation	MCD	NYSE
Merck & Company, Inc.	MRK	NYSE
Microsoft Corporation	MSFT	NASDAQ
Nike, Inc.	NKE	NYSE
Pfizer Inc.	PFE	NYSE
Proctor & Gamble Co.	PG	NYSE
The Travelers Companies, Inc.	TRV	NYSE
UnitedHealth Group, Inc.	UNH	NYSE
United Technologies Corporation	UTX	NYSE
Verizon Communications Inc.	VZ	NYSE
Visa Inc.	V	NYSE
Walmart Inc.	WMT	NYSE
Walgreens Boots Alliance, Inc.	WBA	NASDAQ
The Walt Disney Company	DIS	NYSE

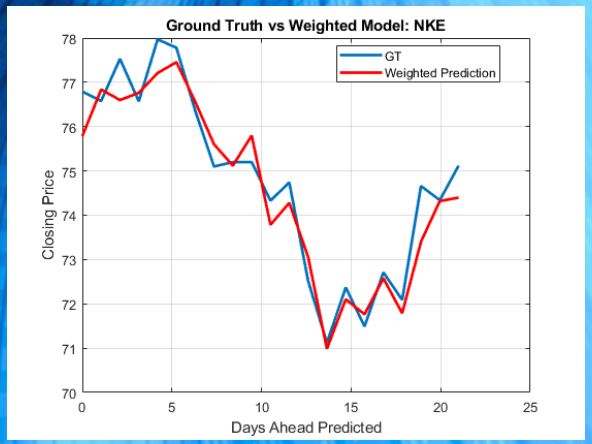
https://www.investopedia.com/terms/d/djia.asp



	tain	test	svm_mse	ridge_mse	ls_mse	bt_mse	ada_mse	weighted_r	weighted_percent error
aapl	9555	21	5239.814	4.7911	7.3848	5.575	36134.56	5.1147	1.0133
axp	11709	21	1011.564	0.80833	1.2217	1.5121	10042.75	1.0957	0.74304
ba	14308	21	46449.36	23.4131	19.9025	25.583	103141.2	20.8903	1.1311
cat	14308	21	1579.117	3.1983	9.5451	3.5991	14264.41	3.38966	1.12658
csco	7234	21	0.61826	0.1619	0.19295	0.20736	515.5325	0.19325	0.75345
cvx	12321	21	2.1963	1.3118	1.2696	1.2131	9966.361	1.2248	0.76479
dis	14308	21	0.6439	0.43951	0.46463	0.45318	10319.6	0.37096	0.43647
dwdp	11709	21	50.228	0.18355	0.48116	0.20738	1512.101	3.3446	1.13131
gs	4908	21	3218	6.1805	5.3311	6.4343	8284.759	5.7827	0.74767
hd	9360	21	1988.364	3.8388	4.0139	4.2216	31646.07	3.6708	0.94764
ibm	14308	21	0.71935	1.0289	0.70922	0.71885	7929.053	0.65653	0.56535
intc	9743	21	0.456	0.25909	0.42047	0.31059	2150.099	0.31067	0.98188
jnj	12321	21	1305.939	1.0317	0.9956	1.1869	17780.1	0.95638	0.48624
jpm	9743	21	0.8052	0.71791	0.43832	0.57102	5719.316	0.5152	0.54046
ko	14308	21	264.3557	0.11768	2.891	3.2433	1519.303	1.4674	2.2253
mcd	12321	21	435.1531	1.3684	56.997	62.1388	25676.53	27.748	2.6444
mmm	12321	21	33050.22	4.1397	4.0157	3.4883	30056.13	3.7331	0.7624
mrk	12321	21	0.33473	0.32593	0.28707	0.26173	3355.344	0.22958	0.5186
msft	8229	21	500.9215	1.1211	0.83372	0.86626	6115.662	0.83826	0.70753
nke	9563	21	0.53578	0.47407	0.49285	0.38082	554.9897	0.30938	0.60154
pfe	11709	21	0.41078	0.12467	0.06741	0.064104	1844.042	0.10893	0.63337
pg	12321	21	1.067	0.25772	0.3253	0.31612	6405.926	9.39894	0.53366
trv	9743	21	624.9604	1.7258	1.1017	1.6286	12164.89	1.4375	0.74581
unh	8582	21	20411.22	5.9988	25.9641	32.6517	69364.94	16.3455	1.1447
utx	12321	21	1.9193	1.0524	0.91133	0.77739	14802.34	0.80217	0.57549
v	2676	21	10.2753	2.3743	0.99417	1.8449	3238.512	2.1772	0.74316
vz	8811	21	0.66769	0.14876	0.16232	0.1509	909.055	0.16903	0.55395
wba	9743	21	0.56608	0.32132	0.2442	0.25947	2899.217	0.25159	0.52558
wmt	11649	21	333.7754	0.50535	0.35499	0.49133	7707.033	0.36225	0.46993
xom	14308	21	0.72208	0.48935	0.28839	0.32346	4629.731	0.30322	0.58332

Results

The result of this experiment shows that the voting model does improve the SVM models. The best experiment can be seen in Nike as it uses all 5 models. The average of all 5 models reduced the MSE. When comparing this model to the other models, the voting model was more accurate.



Barry Wu
RIT Computer Engineering
CMPE-677 Machine Intelligence
Fall 2018