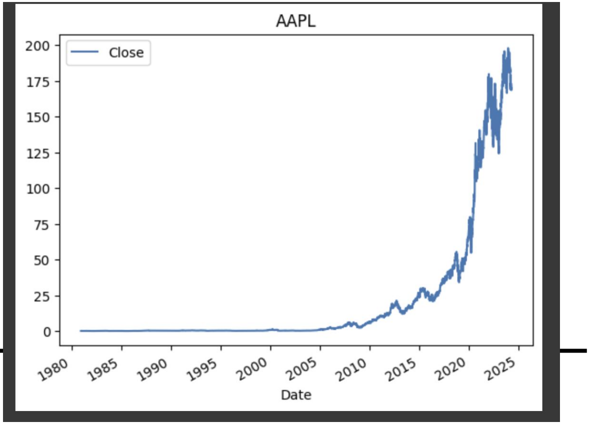
Machine Learning
Stock Advisor

Grant Harrell, Nolan Henderson, Aubrey Trumbo

Background

Predicting stock market movement using ML is a billion dollar industry that can metaphorically print money if done right.

A model that can regularly predict whether a stock will go up or down can remove lots of uncertainty normally associated with trading.



Goals

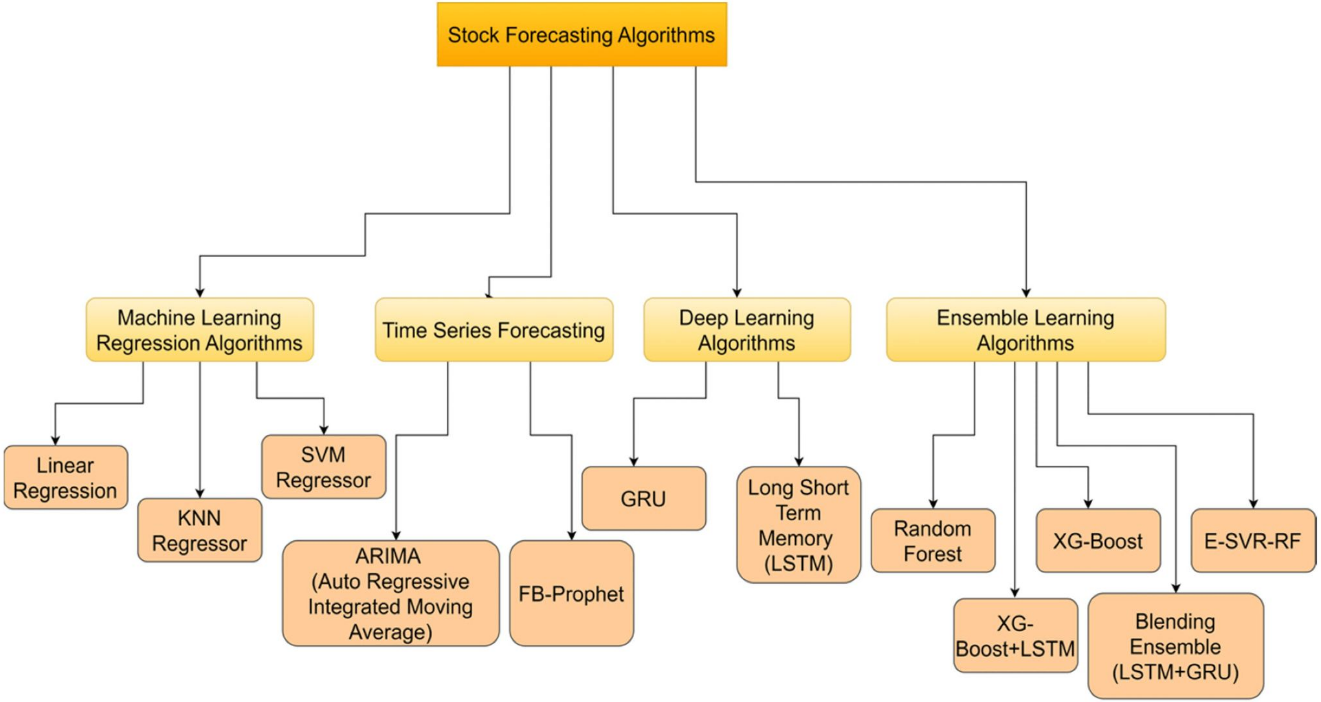
To create a model that:

- Is trained on one specific stock
- Predicts whether the stock will close higher or lower than the opening value
- Has above 50% success

If there is sufficient time remaining once these goals are met, we hope to train additional stocks with the same criteria

Industry Standard

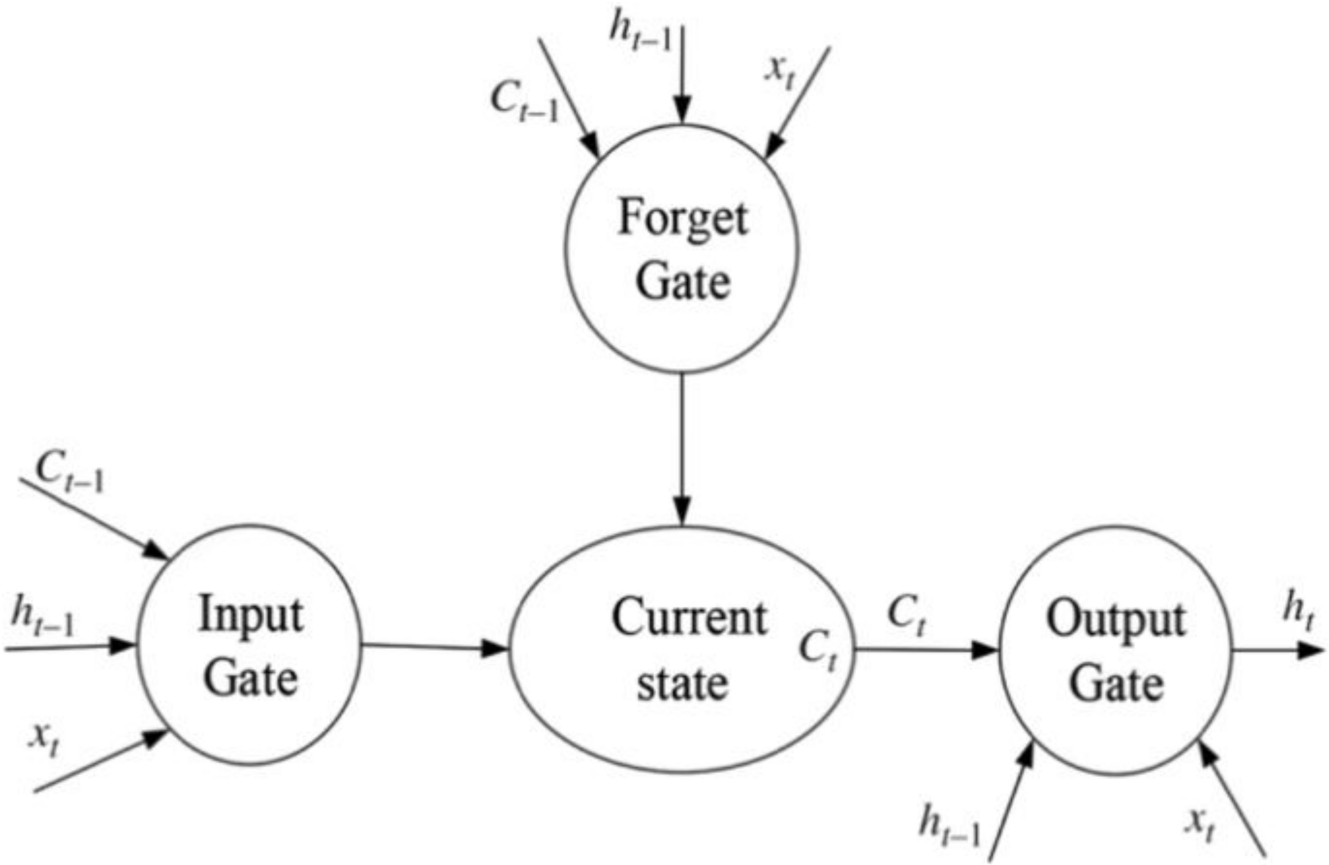
“Good Predictors” typically have an accuracy of 70% - 95%



- Stock Forecasting Algorithms
  - Machine Learning Regression Algorithms
    - Linear Regression
    - KNN Regressor
    - SVM Regressor
  - Time Series Forecasting
    - ARIMA (Auto Regressive Integrated Moving Average)
    - FB-Prophet
  - Deep Learning Algorithms
    - GRU
    - Long Short Term Memory (LSTM)
  - Ensemble Learning Algorithms
    - Random Forest
    - XG-Boost+LSTM
    - XG-Boost
    - Blending Ensemble (LSTM+GRU)
    - E-SVR-RF

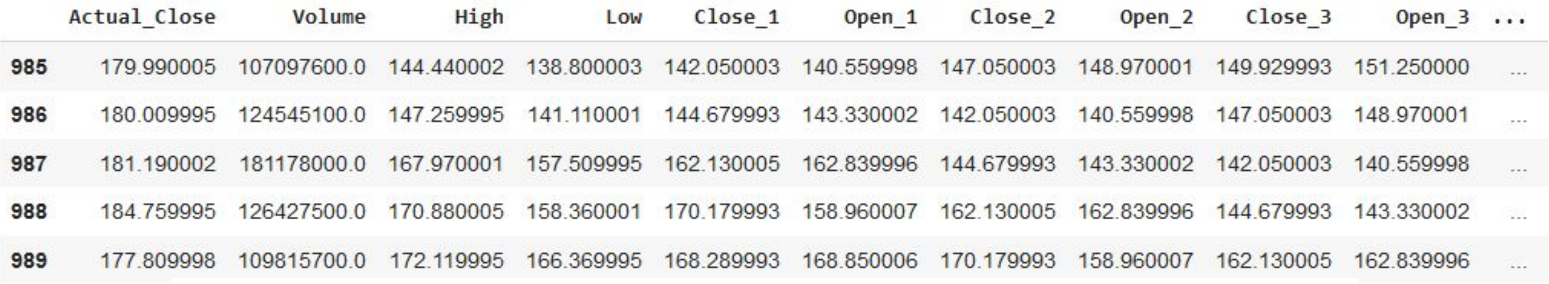
LSTM algorithms (RNN) are popular

- Around 93%

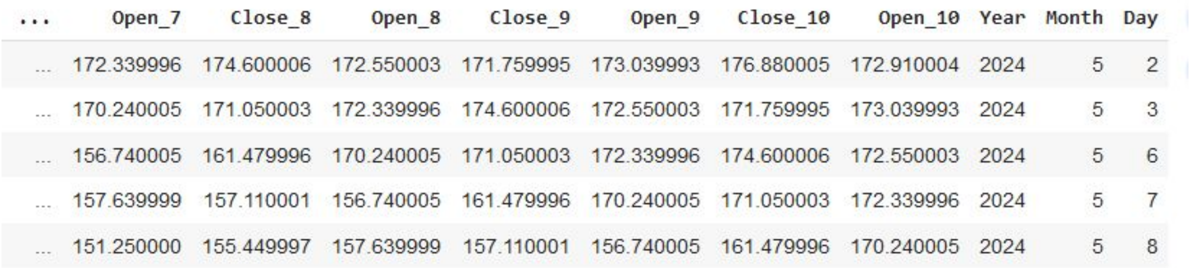


Methodology

Features - Close, Volume, Open, High, Low, Date



	Actual_Close	Volume	High	Low	Close_1	Open_1	Close_2	Open_2	Close_3	Open_3	...
985	179.990005	107097600.0	144.440002	138.800003	142.050003	140.559998	147.050003	148.970001	149.929993	151.250000	...
986	180.009995	124545100.0	147.259995	141.110001	144.679993	143.330002	142.050003	140.559998	147.050003	148.970001	...
987	181.190002	181178000.0	167.970001	157.509995	162.130005	162.839996	144.679993	143.330002	142.050003	140.559998	...
988	184.759995	126427500.0	170.880005	158.360001	170.179993	158.960007	162.130005	162.839996	144.679993	143.330002	...
989	177.809998	109815700.0	172.119995	166.369995	168.289993	168.850006	170.179993	158.960007	162.130005	162.839996	...

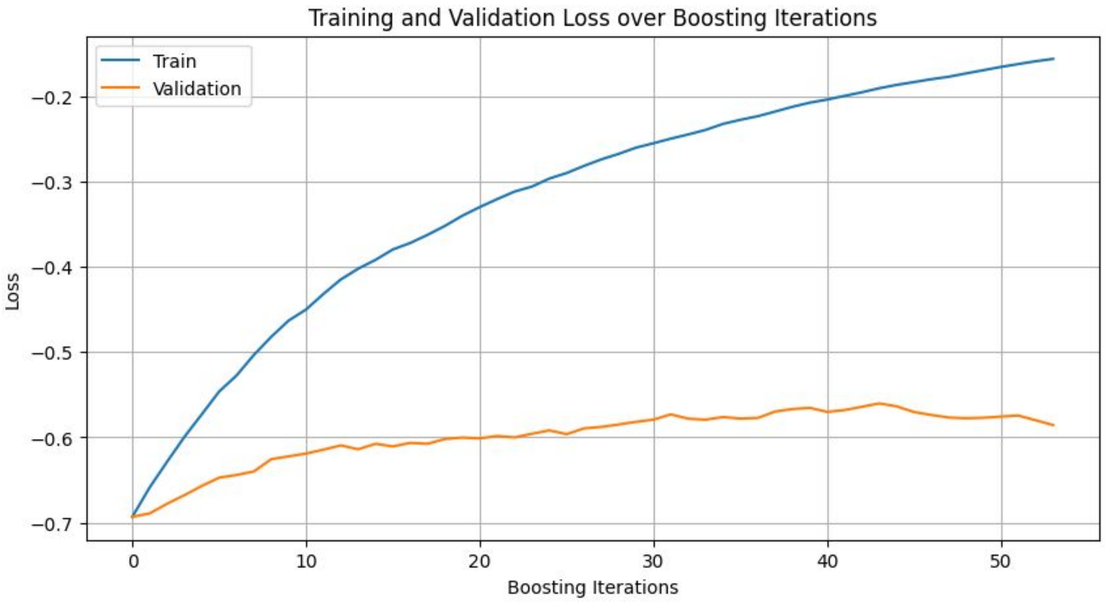


...	Open_7	Close_8	Open_8	Close_9	Open_9	Close_10	Open_10	Year	Month	Day
...	172.339996	174.600006	172.550003	171.759995	173.039993	176.880005	172.910004	2024	5	2
...	170.240005	171.050003	172.339996	174.600006	172.550003	171.759995	173.039993	2024	5	3
...	156.740005	161.479996	170.240005	171.050003	172.339996	174.600006	172.550003	2024	5	6
...	157.639999	157.110001	156.740005	161.479996	170.240005	171.050003	172.339996	2024	5	7
...	151.250000	155.449997	157.639999	157.110001	156.740005	161.479996	170.240005	2024	5	8

Methodology (cont.)

HistGradientBoosting

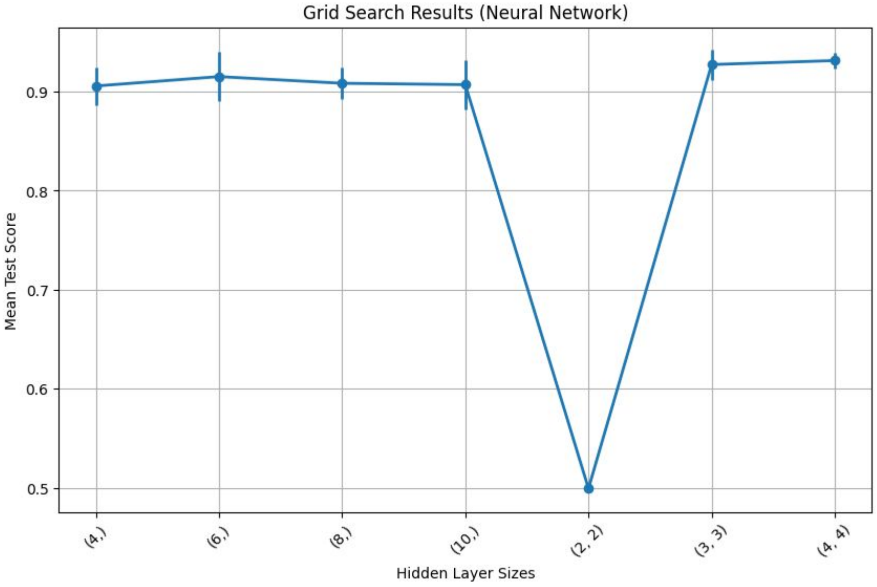
```
Best Parameters:
{'clf__learning_rate': 0.1,
'clf__max_depth': 7,
'clf__min_samples_leaf': 10}
```

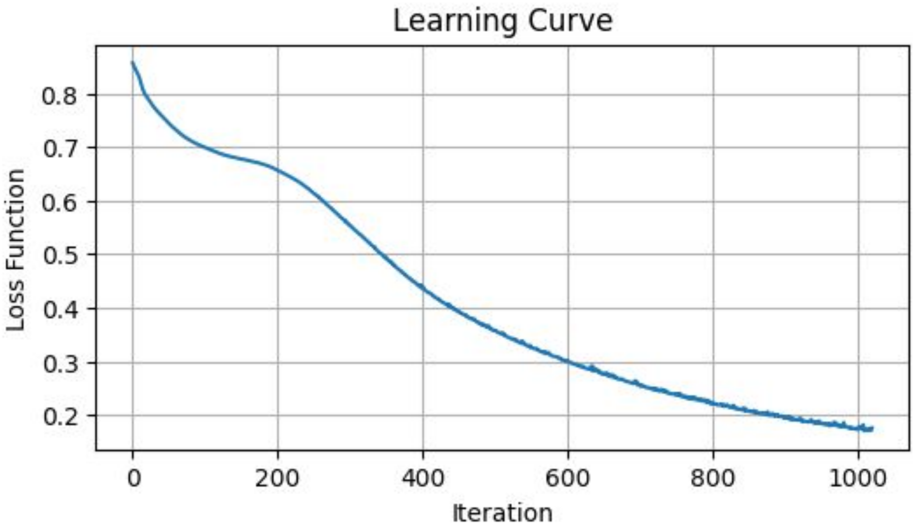


Neural Network

Methodology (cont.)

```
Best parameters:
{'clf  hidden_layer_sizes'
: (4, 4)}
```





Methodology (cont.)

AdaBoost

```
Best Parameters: {'ada__learning_rate': 1.0,
'ada__n_estimators': 100}
```

Methodology (cont.)

Bagging

```
Best Estimator: LogisticRegression

Best Parameters:
{'bagging__base_estimator__C': 10.0,
'bagging__max_features': 1.0,
'bagging__max_samples': 1.0,
'bagging__n_estimators': 50}
```

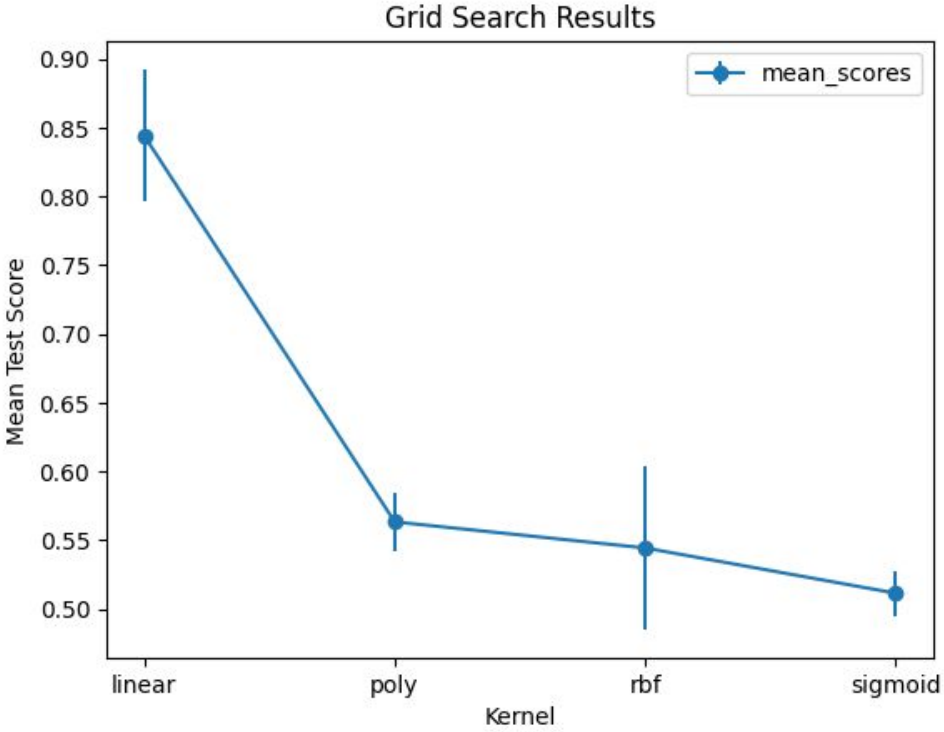
```python
# Define a list of candidate estimators to
evaluate
estimators = [
    ('RandomForest',
RandomForestClassifier()),
    ('SVM', SVC()),
    ('MLP', MLPClassifier()),
    ('KNN', KNeighborsClassifier()),
    ('LogisticRegression',
LogisticRegression())
]
```

Methodology (cont.)

SVM

```
Grid Parameters:
{'clf__kernel': ['linear',
'poly', 'rbf', 'sigmoid']}

Best parameters:
{'clf__kernel': 'linear'}
```



Methodology (cont.)

Voting

```
# Create a VotingClassifier
voting_clf = VotingClassifier(
    estimators=[
        ('gb', clf1),
        ('mlp', clf2),
        ('adaboost', clf3),
        ('bagging', clf4),
        ('svm', clf5)
    ],
    voting='hard'  # Use 'hard' voting for majority rule
)
```

Our Models

adaBoost - 57%

Gradient Boosting - 75.3%

Support Vector Machine - 84.4%

Neural Network - 93%

Bagging - 94.4%

The Real Test

How much money does the model actually make?

The Real Test

How much money does the model actually make?

$-7