HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

Stock Price Analysis and Prediction

Group 18

Tạ Ngọc Minh
ID: 20214918

Lê Ngọc Bình
ID: 20214878

Hoàng Đình Dũng
ID: 20214882

Trần Thanh Trường
ID: 20214938

Supervisor: Assoc. Prof. NGUYEN Linh Giang

Description	Net Change	% Net Change	Open	High	Low
S&P 500 (SPC)	2159.75	3.06	0.13%	2156.81	2160.58	2152.56
DJ Industrial Average (DJI)	18337.50	22.40	0.13%	18313.00	18354.10	18283.20
NYSE Financial Index (NYK)	6100.08	42.96	0.71%	6057.27	6100.45	6057.27
Biotechnology Index (BTK)	3465.00	18.40	0.53%	3433.21	3467.29	3430.07
NYSE Healthcare Index (NYPID)	13057.99	-73.20	-0.56%	13131.19	13131.19	13043.57
S&P 500 Select Sector Energy Index (IXE)	666.71	7.37	1.12%	660.13	667.45	657.02
Bank of America Corp (BAC)	14.46	0.33	2.34%	14.11	14.47	14.09
General Electric Co (GE)	31.13	0.08	0.26%	31.03	31.24	31.02
SPDR S&P 500 ETF Trust (SPY)	215.86	0.33	0.15%	215.48	215.96	215.13
E-Mini S&P 500, Sep 16 (EP)	2154.25	1.75	0.07%	2152.25	2155.50	2145.25
Crude Light (Globex), Sep 16 (CLE)	40.64	1.13	2.86%	39.70	40.68	39.19
NY Harbor ULSD, Sep 16 (HOE)	1.29	0.03	2.40%	1.27	1.29	1.26
RBOB Gasoline (Globex), Sep 16 (RBE)	1.34	0.03	2.36%	1.31	1.34	1.30
Gold (Globex), Dec 16 (GCE)	1363.10	-9.50	-0.69%	1370.50	1373.40	1360.60
DAX Index, Sep 16 (DD)	10166.00	20.00	0.20%	10129.00	10187.00	10088.00
Euro STOXX 50, Sep 16 (DSX)	2906.00	0.00	0.00%	2902.00	2919.00	2888.00
10yr US Treasury Notes (Globex), Sep 16 (TYA)	132.39	-0.16	-0.12%	132.41	132.63	132.27

5-minute Charts

S&P 500 (SPC)

NASDAQ Composite Index (NQCOMP)

Description	Last Price	Net Change	% Net Change	Open	High	Low
NASDAQ Composite Index (NQCOMP)	5146.17	8.44	0.16%	5133.24	5149.22	5128.44
NASDAQ-100 Index (NDX)	4721.87	2.66	0.06%	4716.58	4730.34	4712.10
NASDAQ Financial 100 (IXF)	3344.86	25.12	0.76%	3323.75	3345.02	3322.53
NASDAQ Biotechnology (NBI)	3099.96	16.25	0.53%	3072.36	3101.62	3069.10
Nasdaq Health Care Index (NHC)	685.81	2.75	0.40%	680.88	686.15	680.22
PHLX Oil Service Sector (OSX)	157.54	1.66	1.06%	155.81	158.48	154.98
Sirius XM Holdings Inc. (SIRI)	4.25	0.02	0.47%	4.21	4.27	4.21
Apple Inc (AAPL)	105.45	0.97	0.93%	104.85	105.84	104.77
Tesla Motors, Inc. (TSLA)	226.58	-0.62	-0.27%	227.20	229.69	225.10
Micron Technology Inc (MU)	13.56	0.15	1.12%	13.28	13.57	13.28
Powershares QQQ Trust (QQQ)	115.06	0.10	0.09%	114.88	115.27	114.80
Cisco Systems Inc (CSCO)	30.63	0.01	0.03%	30.53	30.66	30.51
Facebook, Inc. (FB)	122.59	-0.50	-0.41%	123.09	123.92	122.31
iShares 20+ Year Treasury Bond ETF (TLT)	138.21	-0.12	-0.09%	138.60	138.78	137.91
Intel Corporation (INTC)	34.20	-0.36	-1.04%	34.10	34.28	34.10
QUALCOMM Inc (QCOM)	60.76	0.16	0.26%	60.59	61.11	60.50
American Airlines Group Inc. (AAL)	33.39	-0.12	-0.36%	33.49	33.96	33.17

Middle market internals

	NYSE	NASDAQ
Total Volume	331,900	894,953
Up Volume	220,713	556,437
Down Volume	108,706	270,096
Unchanged Vol	2,479	68,420
Adv Issues	1,840	1,577
Dec Issues	1,132	1,065
Unc Issues	74	93
Tick	-15	11

CQG Top Ten Percentage Performers on the NYSE

Description	Last Price	Net Change	% Net Change	Tdy Vol 21d-Vol	Open	High	Low
Genworth Financial Inc. (GNW)	3.56	0.81	29.45%	300%	3.43	3.58	3.32
3D Systems Corporation (DDD)	14.53	2.36	19.29%	182%	13.00	14.58	12.77
American Vanguard Corp (AVD)	17.50	2.58	16.74%	227%	16.05	17.79	15.61
Meritor Inc. (MTOR)	9.23	1.31	16.39%	174%	8.36	9.33	8.23
Nationstar Mortgage Holdings, Inc. (NSM)	13.71	1.93	16.28%	159%	12.81	13.84	12.61
Comstock Resources Inc (CRK)	3.37	0.45	15.41%	232%	2.90	3.47	2.79
Walker & Dunlop, Inc. (WD)	27.04	3.29	14.14%	127%	26.05	27.04	25.80
Fitbit, Inc. (FIT)	14.91	1.75	13.30%	266%	14.11	14.96	13.92
Cheetah Mobile Inc. (CMCM)	11.92	1.34	12.67%	187%	10.50	12.05	10.47
ENPRO Industries Inc (NPO)	50.70	4.60	9.98%	207%	46.50	52.00	46.50

CQG Top Ten Percentage Performers on the NASDAQ

Description	Last Price	Net Change	% Net Change	Tdy Vol 21d-Vol	Open	High	Low
Magellan Petroleum Corp (MPET)	2.92	1.72	143.33%	13422%	1.92	5.00	1.70
China Natural Resources, Inc. (CHNR)	2.45	0.82	47.59%	8815%	1.80	3.17	1.80
Sizmek Inc (SZMK)	3.86	1.20	45.11%	3922%	3.87	3.87	3.85
Cesca Therapeutics Inc. (KOOL)	5.75	1.30	29.21%	627%	6.70	7.39	5.63
Forbes Energy Services Ltd (FES)	0.16	0.03	23.08%	161%	0.14	0.19	0.14
Stanley Furniture Co Inc (STLY)	2.98	0.53	21.63%	4173%	3.25	3.25	2.95
Neovasc Inc. (NVCN)	0.66	0.12	22.22%	315%	0.54	0.72	0.54
Ocean Power Technologies (OPTT)	8.94	1.48	19.84%	42%	7.20	9.12	7.20
Memorial Production Partners LP (MEMP)	1.66	0.27	19.42%	161%	1.40	1.69	1.40
Big 5 Sporting Goods Co (BGFV)	12.44	1.97	18.82%	339%	11.80	12.82	11.54

CQG Bottom Ten Percentage Performers on the NYSE

Description	Last Price	Net Change	% Net Change	Tdy Vol 21d-Vol	Open	High	Low
The Rubicon Project, Inc. (RUBI)	9.24	-4.43	-32.41%	1186%	9.50	9.70	9.00
Kate Spade Company (KATE)	15.92	-4.22	-20.95%	571%	16.00	16.50	15.66
ARC Document Solutions Inc (ARC)	3.39	-0.43	-11.26%	136%	3.61	3.61	3.35
Capital Senior Living Corp (CSU)	17.03	-2.05	-10.74%	324%	18.45	18.48	16.75
Stone Energy Corp (SGY)	10.06	-1.06	-9.53%	47%	8.50	10.47	8.42
Myers Industries Inc (MYE)	13.44	-1.36	-9.19%	45%	13.92	13.94	13.20
PROS Holdings, Inc. (PRO)	16.86	-1.94	-10.46%	112%	17.88	18.15	16.25
Tableau Software Inc (DATA)	51.49	-4.88	-8.71%	323%	52.87	53.50	50.27
Themon Group Holdings, Inc. (THR)	18.20	-1.40	-7.14%	212%	17.87	18.25	16.66
DRDGOLD Limited (DRD)	8.05	-0.60	-6.94%	55%	8.50	8.50	8.02

CQG Bottom Ten Percentage Performers on the NASDAQ

Description	Last Price	Net Change	% Net Change	Tdy Vol 21d-Vol	Open	High	Low
Cray Inc (CRAY)	22.04	-9.24	-29.54%	1046%	25.06	25.19	21.80
Electro Scntfc Inds Inc (ESIO)	4.85	-2.00	-29.20%	551%	5.81	5.89	4.74
The Chefs' Warehouse, Inc. (CHEF)	11.51	-4.64	-28.73%	1038%	14.00	14.39	10.78
Spherix Incorporated (SPEX)	1.52	-0.58	-27.62%	232%	1.60	1.69	1.45
Crocs Inc (CROX)	8.31	-2.72	-24.45%	605%	8.79	9.50	8.10
Sino Global Shipping America LTD (SINO)	1.79	-0.45	-20.09%	86%	2.24	2.26	1.77
Globus Maritime Limited (GLBS)	0.67	-0.15	-18.29%	65%	0.81	0.83	0.63
Cosi Inc (COSI)	0.23	-0.05	-17.86%	94%	0.26	0.26	0.22
USA Truck (USAK)	15.73	-3.62	-18.20%	83%	16.42	18.10	15.51
Odyssey Marine Exploration Inc (OMEX)	3.64	-0.75	-17.08%	469%	4.27	4.31	3.56

Rank

Google

stock prediction

Tin tức   Hình ảnh   Video   LSTM   Github   RNN   CNN-LSTM   Sách   Maps

Khoảng 271.000.000 kết quả (0,45 giây)

Google

nyse nasdaq stock prediction

Tin tức   Hình ảnh   Video   Sách   Maps   Mua sắm   Chuyến bay   Tài chính
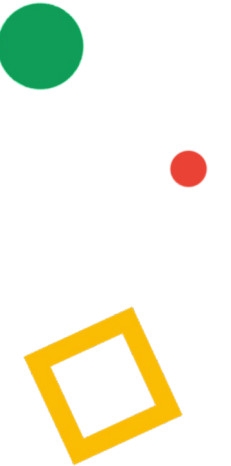
Khoảng 21.700.000 kết quả (0,40 giây)

Contents

- Data crawling and analysis

- Algorithm

- Model Tuning

- Summary and proposal of improvements

Data crawling and analysis

Data Crawling

- Data collected from TCBS Dashboard (Techcombank Securities)

- Contain data up to July 1st, 2023

- Features: Open, High, Low, Close, Volume, Trading Date, Stock Code.

		Open	High	Low	Close	Volume	TradingDate	Code
1	0	16203.0	19804.0	15753.0	18138.0	278800	2010-07-06	LCS
2	1	19399.0	19399.0	19399.0	19399.0	30700	2010-07-07	LCS
3	2	20749.0	20749.0	18453.0	20299.0	192800	2010-07-08	LCS
4	3	18904.0	18904.0	18904.0	18904.0	1000	2010-07-09	LCS
5	4	17598.0	18904.0	17598.0	17733.0	133500	2010-07-12	LCS

Data Crawling

Data Crawling

`…?ticker={}&type=stock&resolution={}&from=0&to={}`

- Here I use request library to handle the API with:

  - Ticker: Stock code

  - Resolution: Time range

  - From – to: The starting time and ending time, converted into absolute time.

Data Analysis

4 Trading fields with 2 most traded stocks:

- Constructor:
    - LCS: Licogi JSC
    - PTC: Icapital Investment JSC
- Technology:
    - VGI: Viettel Global Investment
    - ITD: Tien Phong Technology
- Consumer:
    - LSS: Lam Son Sugar
    - PLX: Vietnam National Petroleum Group
- Finance:
    - TCB: Vietnam Technological and Commercial Joint Stock Bank
    - VIG: Vietnam Financial Investment Securities Corporation

Data Analysis

Stock Price Illustration

Data Analysis

Revision of GDP growth forecast of selected countries due to the outbreak of COVID-19
Unit: percentage

Vietnam: 7.0%, 6.7%, 4.8%, 3.3%, 2.7%
United States: 2.3%, 1.9%, 0.4%, -5.9%
Japan: 0.7%, 0.5%, -1.5%, -5.2%
South Korea: 2.4%, 2.0%, 1.3%, -1.2%
China: 6.1%, 6.0%, 2.3%, 1.2%
Thailand: 3.2%, 2.4%, -4.8%, -6.7%

Legend:
- 2019
- 2020f before COVID-19
- 2020f after COVID-19 (ADB – 03 April 2020)
- 2020f after COVID-19 (IMF – 06 April 2020)
- 2020f after COVID-19 (Fitch – 08 April 2020)

Data Analysis

Vietnam's GDP growth
2023, 2024: World Bank's forecasts

World Bank, General Statistics Office
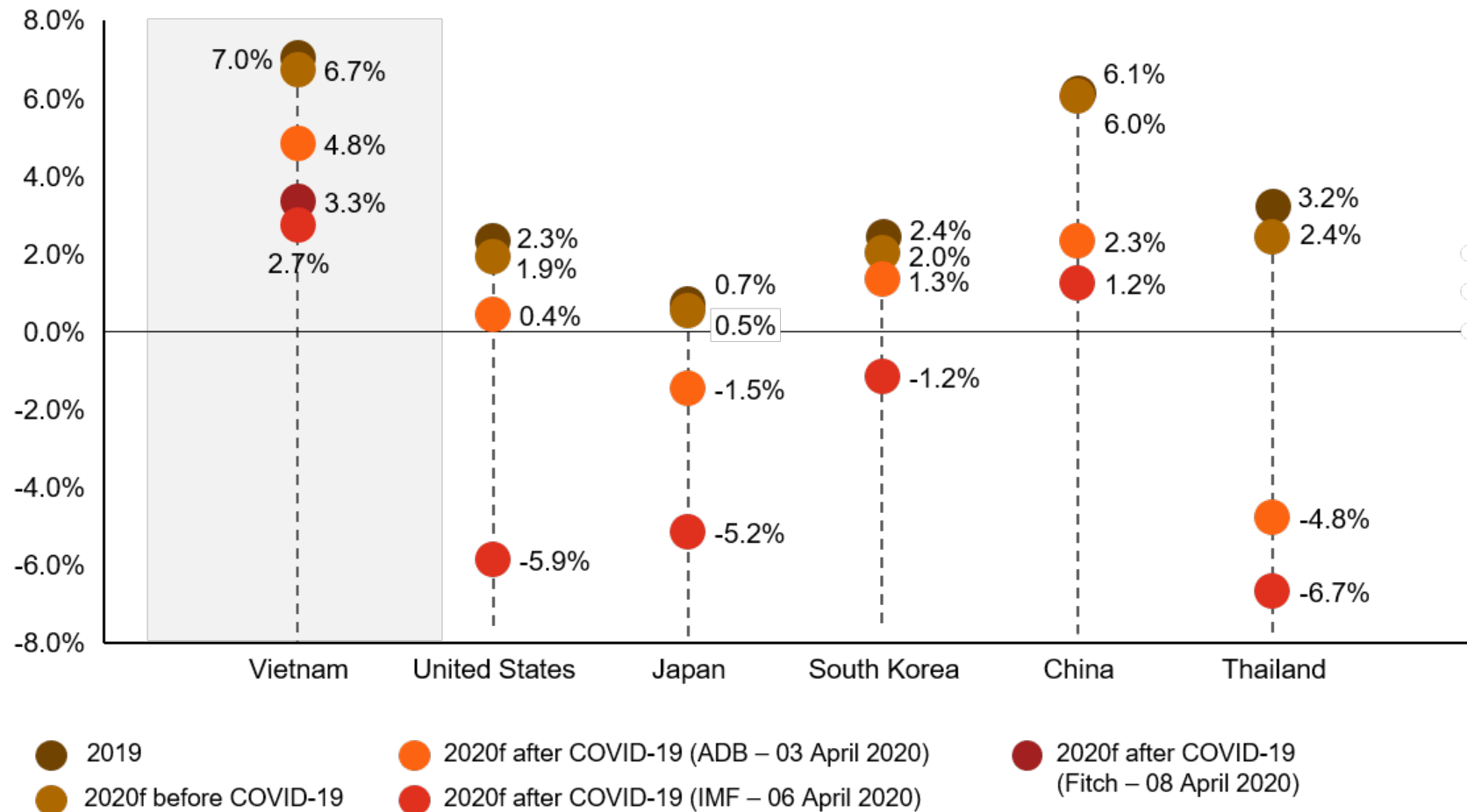
Feature Selection

- We only choose `Close` feature since the importance of all four first features are equal.

- `Volume` seems unrelated, however, we will use Volume in another way.

Correlation Matrix Of The Dataset
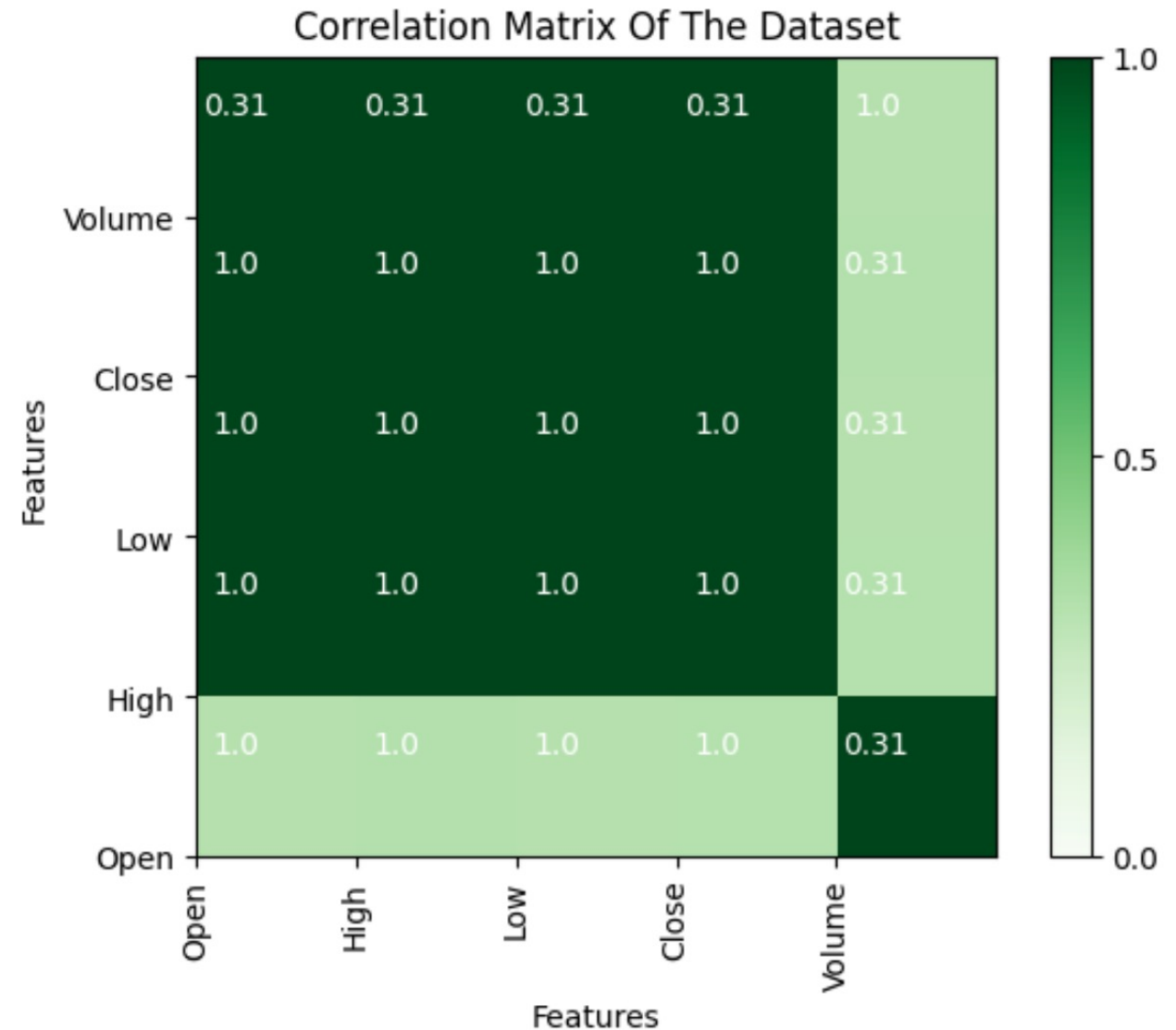
RSI Score and relation to the volume

- The relative strength index (RSI) is a momentum indicator used in technical analysis. RSI measures the speed and magnitude of a security's recent price changes to evaluate overvalued or undervalued conditions in the price of that security.
- To calculate RSI, we define upward (U) and downward (D) indicators, which are:

$$U_t = \begin{cases} P_t - P_{t-1} & \text{if } P_t > P_{t-1}, \\ 0 & \text{otherwise.} \end{cases}$$

$$D_t = \begin{cases} -P_t + P_{t-1} & \text{if } P_t < P_{t-1}, \\ 0 & \text{otherwise.} \end{cases}$$

- Then up_t and down_t are average numbers of upward moves and downward moves of closing price of past n days:

$$\text{up}_t = \frac{s}{n+1} \times U_t + \left(1 - \frac{s}{n+1}\right) \times \text{up}_{t-1}$$

$$\text{down}_t = \frac{s}{n+1} \times D_t + \left(1 - \frac{s}{n+1}\right) \times \text{down}_{t-1}$$

- Finally, we get the RSI score calculated by:

$$\text{RSI}_t = 100 - \frac{100}{1 + \frac{\text{up}_t}{\text{down}_t}}$$

RSI Score and relation to the volume

Algorithms

A. Extreme Gradient Boosting (XGBoost)

- Boosting:

    - Boosting is an ensemble modelling

    - Build a strong model from the number of weak models

- Gradient boosting:

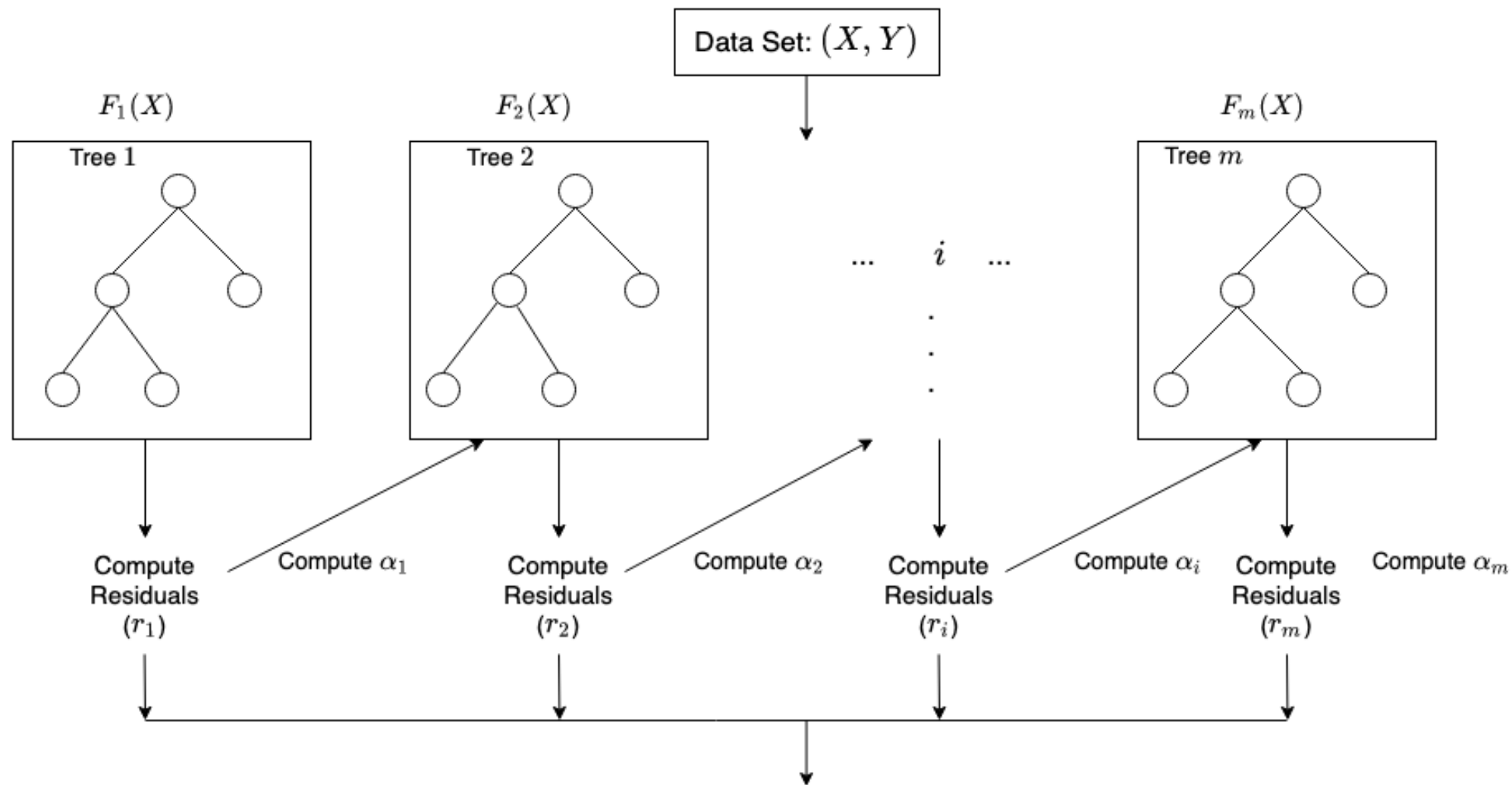   - In gradient boosting, each predictor corrects its predecessor's error

A. Extreme Gradient Boosting (XGBoost)

- Extreme Gradient Boosting:

    - An implementation of Gradient Boosted decision trees

    - Decision trees are created in sequential form

    - Weights play an important role in XGBoost

Data Set: (X, Y)

$F_1(X)$ — Tree 1

$F_2(X)$ — Tree 2

$\ldots \quad i \quad \ldots$

$F_m(X)$ — Tree m

Compute Residuals (r_1)

Compute α_1

Compute Residuals (r_2)

Compute α_2

Compute Residuals (r_i)

Compute α_i

Compute Residuals (r_m)

Compute α_m

$$F_m(X) \ = \ F_{m-1}(X) + \alpha_m h_m(X, r_{m-1}),$$

where α_i, and r_i are the regularization parameters and residuals computed with the ith tree respectfully, and h_i is a function that is trained to predict residuals, r_i using X for the ith tree. To compute α_i we use the residuals computed, r_i and compute the following: $arg \min_{\alpha} \ = \ \sum_{i=1}^{m} L(Y_i, F_{i-1}(X_i) + \alpha h_i(X_i, r_{i-1}))$ where $L(Y, F(X))$ is a differentiable loss function.

B. Support Vector Regression (SVR)

- SVR is an improved model of Support Vector Machine (SVM)

- Find an appropriate line (or hyperplane in higher dimensions) to fit the data

- Can handle high-dimension feature

B. Support Vector Regression (SVR)

The function of SVR method is: $\boxed{y = f(x) = w^T \varphi(x) + b}$

The w and b are estimated by minimizing the following optimi...

$$\min \frac{1}{2}\|w\|^2$$

subjects to $\begin{cases} y_i - w^T \varphi(x_i) - b \le \varepsilon, \\ b + w^T \varphi(x_i) - y_i \le \varepsilon + \xi_i^* \end{cases}$

Here, ξ_i and ξ_i^* are slack variables introduced to cope with training data possibly violating the condition $|f(x_i) - y_i| \le \varepsilon$.

$$\min \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^*)$$

subjects to $\begin{cases} y_i - w^T \varphi(x_i) - b \le \varepsilon, \\ b + w^T \varphi(x_i) - y_i \le \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* \ge 0, i = 1, ..., n, \end{cases}$

where C is a constant known as the penalty factor, ε is the insensitive loss parameter and the slack variables ξ_i and ξ_i^* measure the amount of difference between the estimated value and the target value beyond ε.

Decision Boundary

Decision Boundary

Hyperplane
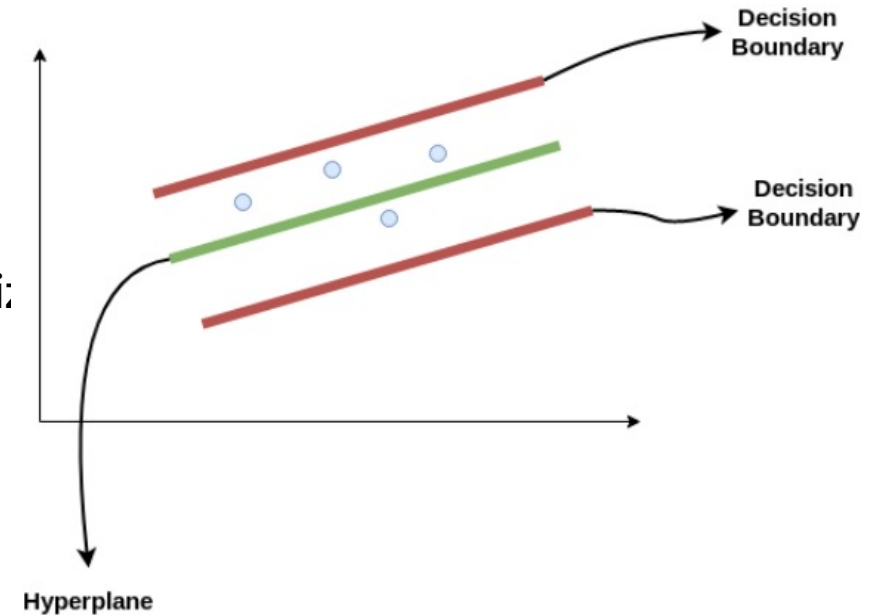
C. Auto Regressive Integrated Moving Average (ARIMA)

- ARIMA is actually a class of models that explains a given time series based on its own past values, that is, its own lags and the lagged forecast errors, so that equation can be used to forecast future values.

- An ARIMA model is characterized by 3 terms: p, d, q, where:
  - p is the order of the AR term,
  - q is the order of the MA term
  - d is the number of differencing required to make the time series stationary.

C. Auto Regressive Integrated Moving Average (ARIMA)

AutoRegressive Integrated Moving Average(ARIMA) is a time series forecasting model that incorporates autocorrelation measures to model temporal structures within the time series data to predict future values.

A pure Auto Regressive (AR only) model is one where depends only on its own lags. That is, Y_t is a function of the 'lags of Y_t'.

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + ... + \beta_p Y_{t-p} + \varepsilon_1$$

where, Y_{t-1} is the lag-1 of the series, β_1 is the coefficient of lag-1 that the model estimates and α is the intercept term, also estimated by the model.

C. Auto Regressive Integrated Moving Average (ARIMA)

A pure Moving Average (MA only) model is one where Yt depends only on the lagged forecast errors

$$Y_t = \alpha + \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + ... + \varphi_q \varepsilon_{t-q}$$

where the error terms are the errors of the autoregressive models of the respective lags. The errors Et and Et−1 are the errors from the following equations :

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + ... + \beta_t Y_0 + \varepsilon_t$$
$$Y_{t-1} = \beta_1 Y_{t-2} + \beta_2 Y_{t-3} + ... + \beta_{t-1} Y_0 + \varepsilon_{t-1}$$

An ARIMA model is one where the time series was differenced at least once to make it stationary, and you combine the AR and the MA terms. So, the equation becomes:

$$Y_t = \quad \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + ... + \beta_p Y_{t-p} + \varepsilon_t$$
$$+ \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + ... + \varphi_q \varepsilon_{t-q}$$

Model Tuning

A – Hyperparameter Tuning

- Tuning is to find the optimal parameter

- Improve predictor power of model and run speed

- Tune parameter of SVR and XGBoost model

A – Hyperparameter tuning - Choosing hyperparameters

Support Vector Regression:

- C: determines the penalty for data points that fall outside the margin or violate the regression tolerance

- Kernel: determines the type of kernel function used in SVR to transform the feature space

  - linear: used when the data can be well separated by a hyperplane.

  - rbf (Radial Basis Function): used when the data is not linearly separable in the original space.

  - Sigmoid : used when the data is not linearly separable and has similar characteristics to the sigmoid function in logistic regression.

  - γ: adjusts the influence of a training data point on other data points.

A – Hyperparameter tuning - Choosing hyperparameters

XGBoost:

- n_estimators: represents the number of individual decision trees (weak learners) to be built in the XGBoost model.

- max_depth: defines the maximum depth of each decision tree in the XGB mode

- learning_rate: controls the contribution of each tree in the ensemble

B – Cross Validation For Time Series

- Can't use Cross-Validation in this project

- Alter by Time Series Cross Validation

- The rules of splits :
  - Every test set contains unique observations
  - Observations from the training set occur before their corresponding test set

- In this project, we introduce two kind of Nested Cross Validation:
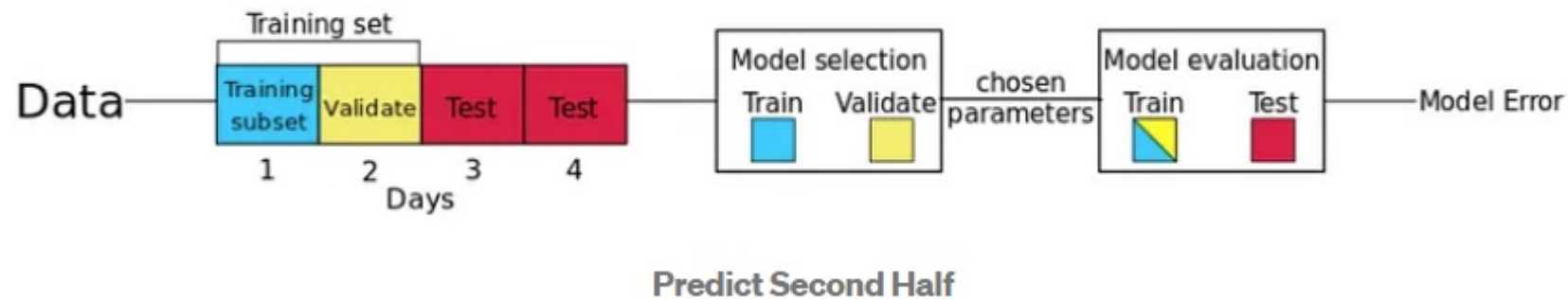 - Predict Second Half
  - Day Forward Training

B – Cross Validation For Time Series
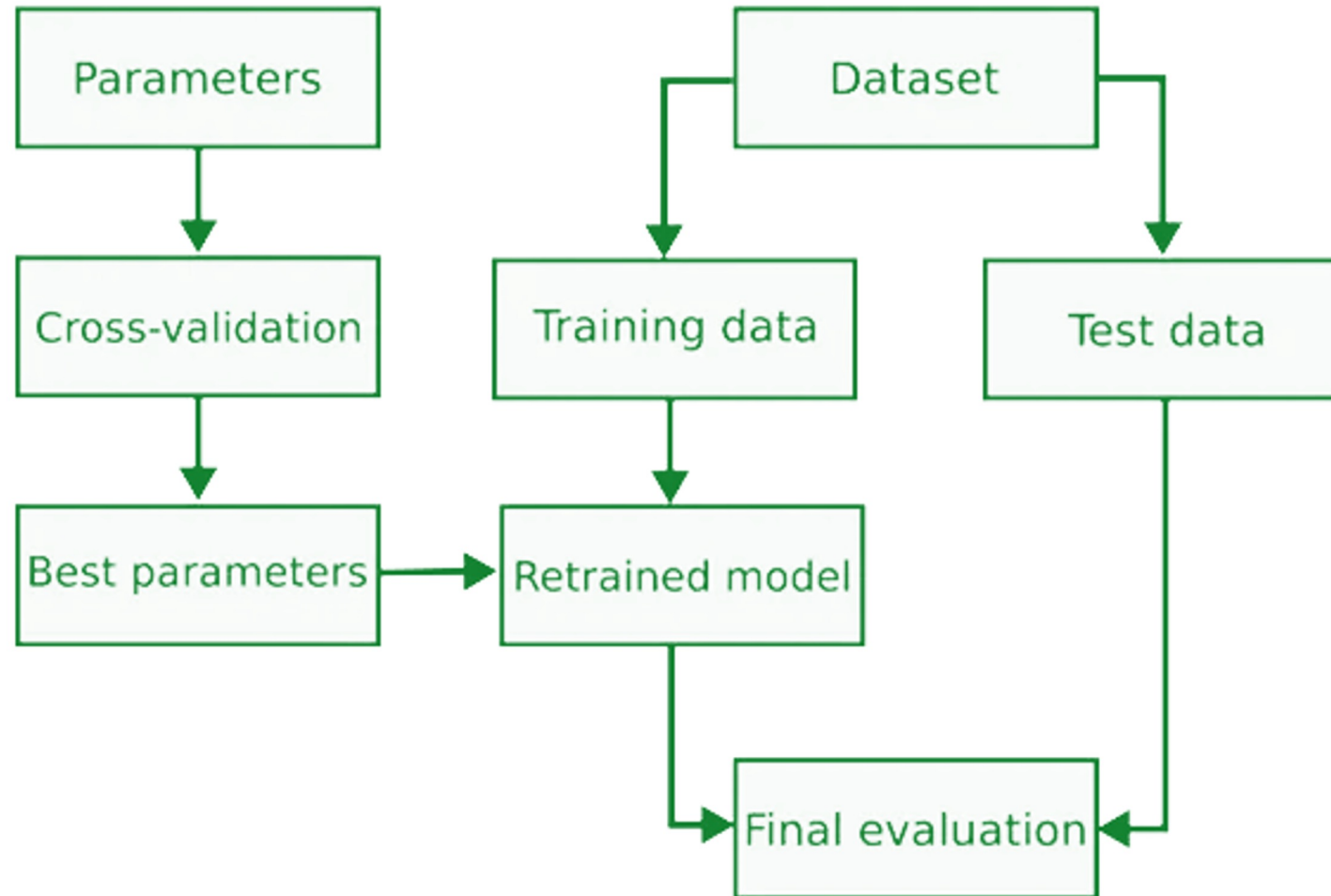
B – Cross Validation - Predict Second Half

- 50% : Training subset/ Validate
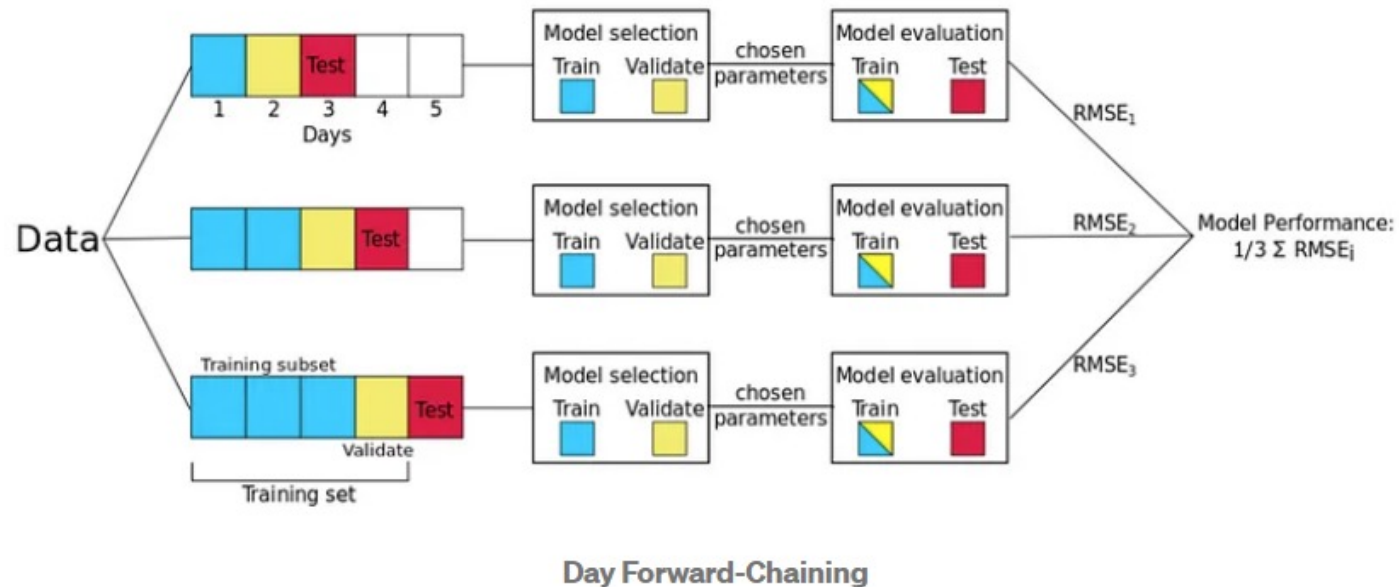- 50% : Test

Predict Second Half

- **Simply but pay attention to chronological order of the data !!!**

B – Cross Validation - Day Forward-Chaining

- Each day as the test set while assigning all previous data to the training set.

- Each split represents a distinct evaluation scenario.

Day Forward-Chaining

B – Cross Validation - Comparison and Conclusion

The Predict Second Half

- may not capture the full variability of the data or provide a robust estimate of model error

- may not effectively evaluate the model's ability to generalize to unseen future data

The Day forward-chaining

- more robust estimate of the model's error

- better assessment of its generalization capabilities.

Summary and Improvements

Summary

The number in each cells are

- Train mean squared error,
- Train R2 score,
- Test mean squared error,
- Test R2 score

respectively.

Alg.	LCS	PTC	VGI	ITD
ARIMA	0.06	0.0015	0.04	0.007
	0.94	0.998	0.96	0.94
	0.12	0.0034	0.09	0.008
	0.90	0.801	0.9	0.93
SVR	0.005	0.004	0.016	0.002
	0.995	0.995	0.983	0.997
	0.015	0.007	0.042	0.0224
	0.984	0.992	0.957	0.977
SVR & DFC	0.006	0.0044	0.014	0.0014
	0.994	0.995	0.984	0.995
	0.001	0.0043	0.018	0.006
	0.972	0.936	0.960	0.963
SVR & PSH	0.006	0.004	0.0016	0.002
	0.993	0.995	0.983	0.997
	0.020	0.005	0.019	0.008
	0.979	0.994	0.980	0.991
XGB	0.0003	0.0003	0.013	0.0002
	0.999	0.999	0.998	0.999
	0.033	0.021	0.024	0.202
	0.966	0.978	0.975	0.801
XGB & DFC	0.0015	0.0003	0.0012	0.0004
	0.998	0.999	0.998	0.998
	0.027	0.021	0.0121	0.301
	0.519	0.978	0.875	0.201
XGB & PSH	1.630	4.612	1.569	7.28
	0.999	0.999	0.999	0.999
	0.006	0.016	0.014	0.149
	0.939	0.983	0.985	0.850

Alg.	LSS	PLX	TCB	VIG
ARIMA	0.0053	0.003	0.006	0.008
	0.9947	0.996	0.992	0.93
	0.0087	0.004	0.032	0.163
	0.982	0.945	0.984	0.87
SVR	0.0049	0.0047	0.0029	0.0026
	0.995	0.995	0.997	0.9974
	0.01	0.012	0.021	0.0089
	0.99	0.988	0.978	0.991
SVR & DFC	0.006	0.0044	0.006	0.0026
	0.994	0.996	0.994	0.996
	0.005	0.006	0.005	0.004
	0.99	0.99	0.991	0.988
SVR & PSH	0.0052	0.005	0.0206	0.003
	0.9947	0.995	0.979	0.996
	0.0077	0.008	0.01	0.004
	0.992	0.992	0.99	0.995
XGB	0.13	0.12	0.08	0.06
	0.86	0.981	0.97	0.95
	0.19	0.07	0.06	0.07
	0.8	0.93	0.94	0.96
XGB & DFC	0.12	0.1	0.002	0.09
	0.88	0.9	0.98	0.91
	0.16	0.042	0.038	0.12
	0.83	0.96	0.91	0.87
XGB & PSH	0.01	0.17	0.02	0.12
	0.99	0.82	0.91	0.87
	0.072	0.21	0.03	0.18
	0.93	0.8	0.93	0.821

Proposal of Improvements

- **Enhancing Data Integration:** To improve the overall performance of the system, we recommend integrating additional relevant data sources. By incorporating financial news, social media sentiment analysis, and economic indicators, we can capture a broader range of factors that influence stock prices.

- **User-Friendly Interface and Visualization:** To ensure an intuitive and user-friendly experience, we propose developing a web-based interface that allows users to interact with the system effortlessly. The interface should provide clear visualizations of predicted stock prices, historical data, and relevant performance metrics. Additionally, incorporating features like customizable alerts and notifications will empower users to make timely and informed investment decisions.

References

1. Phan, K. C., & Zhou, J. (2014). Market efficiency in emerging stock markets: A case study of the Vietnamese stock market. IOSR Journal of Business and Management, 16(4), 61-73.
2. Ngọc, Đ. B., & Cường, N. C. (2016). Các nhân tố tác động đến dao động giá cổ phiếu của các công ty niêm yết trên thị trường chứng khoán Việt Nam. Tạp Chí Phát Triển Kinh Tế, Tạp Chí Khoa Học Kinh Tế, Tạp Chí Kinh Tế Và Tài Chính Quốc Tế, 228(6), 43-51.
3. Nguyễn, T. T. V. (2010). Cuộc khủng hoảng kinh tế thế giới năm 2008-2009 và tác động đối với kinh tế-xã hội Việt Nam: Luận văn ThS. Quan hệ quốc tế: 60 31 40 (Doctoral dissertation, Trường Đại học Khoa học Xã hội và Nhân văn).
4. Chương, P. H. (2020). Tác động của đại dịch Covid-19 đến nền kinh tế việt nam. Tạp chí Kinh tế và Phát triển, 274, 12
5. Gumparthi, S. (2017). Relative strength index for developing effective trading strategies in constructing optimal portfolio. International Journal of Applied Engineering Research, 12(19), 8926-8936.
6. Ghosh, M., & Gor, R. (2022). STOCK PRICE PREDICTION US- ING SUPPORT VECTOR REGRESSION AND K-NEAREST NEIGH-BORS: A COMPARISON. International Journal of Engineering Science Technologies, 6(4), 1–9. https://doi.org/10.29121/ijoest.v6.i4.2022.354
7. Prabhakaran, S. (2023). ARIMA Model – Complete Guide to Time Series Forecasting in Python. Machine Learning Plus. https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/
8. Shrivastava, S. (2023, May 24). Cross Validation in Time Series. Medium. https://medium.com/@soumyachess1496/cross-validation-in-time-series-566ae4981ce4

Thank you for your attention!

Don't hesitate to state any question you have.