REFERENCES

Agarwalla, N., Panda, D., & Modi, P. M. (2016). Deep Learning using Restricted Boltzmann Machines. *(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (3)* , 1552-1556.

Atul. (22 May, 2019). *What is Machine Learning? Machine Learning For Beginners*. Retrieved from edureka: https://www.edureka.co/blog/what-is-machine-learning/#ML

Baresa, S., Bogdan, S., & Ivanovic, Z. (2013). Strategy of Stock Valuation by Fundamental Analysis. *Special issue, UTMS Journal of Economics 4 (1)*, 45-51.

Basak, S., Kar, S., Saha, S., Khaidema, L., & Dey, S. R. (2019). Predicting the direction of stock market prices using tree-based. *The North American Journal of Economics and Finance, Vol 47*, 552-567, https://doi.org/10.1016/j.najef.2018.06.013.

BBC News. (9 April, 2020). *Coronavirus: Worst economic crisis since 1930s depression, IMF says*. Retrieved from BBC News Web site: https://www.bbc.com/news/business-52236936

Bursa Malaysia. (2020). *Bursa Malaysia announces RM151.0 million profit-after-tax-and-minority-interest for the first half of 2020 highest first half financial performance since listing in 2005*. Retrieved from Bursa Malaysia website: https://www.bursamalaysia.com/about_bursa/media_center/

Butt, S., Ramakrishnan, S., Chohan, M. A., & Punshi, S. K. (2019). Prediction of Malaysian Exchange Rate Using Microstructure Fundamental and Commodities Prices: A Machine Learning Method. *International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-2S9,*. Blue Eyes Intelligence Engineering & Science Publication.

Chen, J. (16 May, 2020). *Technical Indicator*. Retrieved from Investopedia Web site: https://www.investopedia.com/terms/t/technicalindicator.asp#:~:text=Technical%20indicators%20are%20heuristic%20or,to%20predict%20future%20price%20movements.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., . . . Li, Y. (2020). *xgboost: Extreme Gradient Boosting*. Retrieved from R package version 1.0.0.2: https://CRAN.R-project.org/package=xgboost

Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems With Applications 83* , 187–205.

Ciaburro, G., & Venkateswaran, B. (2017). *Neural Networks with R*. ISBN 9781788397872 : Packt Publishing Limited. Retrieved from https://subscription.packtpub.com/book/big_data_and_business_intelligence/9781788397872/1/ch01lvl1sec27/pros-and-cons-of-neural-networks

Dash, R., & Dash, P. K. (2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. *The Journal of Finance and Data Science 2* , 42-57.

Dey, S., Kumar, Y., Saha, S., & Basak, S. (2016). Forecasting to Classification: Predicting the direction of stock market price using Xtreme Gradient Boosting. DOI: 10.13140/RG.2.2.15294.48968.

Dongdong, L., Shuhan, Y., Meizi, L., & Yang, X. (14 April, 2019). *An Empirical Study of Machine Learning Algorithms for Stock Daily Trading Strategy*. Retrieved from Hindawi: https://doi.org/10.1155/2019/7816154

Heuer, J. (16 December, 2019). *XGBoost: Enhancement Over Gradient Boosting Machines*. Retrieved from Medium Web site: https://opendatascience.com/xgboost-enhancement-over-gradient-boosting-machines/

Huigol, P. (24 August, 2019). *Accuracy vs. F1-Score*. Retrieved from Analytics Vidhya Web site: https://medium.com/analytics-vidhya/accuracy-vs-f1-score-6258237beca2#:~:text=Accuracy%20is%20used%20when%20the,as%20in%20the%20above%20case.

Jatain, A., & Ranjan, A. (2017). A Review Study on Big Data Analysis using R Studio. *International Journal of Computer Science and Mobile Computing, Vol 6 Issue 6*, 8-13.

Kapoor, A. (25 February, 2019). *Deep Learning vs. Machine Learning: A Simple Explanation*. Retrieved from Hacker Noon: https://hackernoon.com/deep-learning-vs-machine-learning-a-simple-explanation-47405b3eef08

Karatzoglou, A., Smola, A., Hornik, K., & Zeilei, A. (2004). *kernlab -- An {S4} Package for Kernel Methods in {R}*. Retrieved from Journal of Statistical Software Vol 11, No 9, 1-20: http://www.jstatsoft.org/v11/i09/

Khoo, D., & Kaur, G. (6 May, 2017). *Rise of the Algorithms*. Retrieved from The Star: https://www.thestar.com.my/business/business-news/2017/05/06/rise-of-the-algorithms

Krauss, C., Dob, X. A., & Huck, N. (2016). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *FAU Discussion Papers in Economics, No. 03, ISSN 1867-6707*.

Kristjanpoller, W., Fadic, A., & Minutolo, M. C. (2014). Volatility forecast using hybrid Neural Network models. *Expert Systems with Applications, Vol 41, Issue 5*, 2437-2442.

Kumar, D. (14 June, 2019). *Top 4 advantages and disadvantages of Support Vector Machine or SVM*. Retrieved from Medium Web site: https://dhirajkumarblog.medium.com/top-4-advantages-and-disadvantages-of-support-vector-machine-or-svm-a3c06a2b107

Lee, Y. N. (26 August, 2020). *Retail investors with 'money to play with' help Malaysian stocks recoup nearly all losses this year*. Retrieved from CNBC Web site:

https://www.cnbc.com/2020/08/26/malaysian-retail-investors-pile-into-stocks-help-market-recoup-losses.html

Lo, A., Mamaysky, H., & Wang, J. (2000). Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementation. *The Journal of Finance, 55(4)*, 1705-1765.

L'Sousa, R., & Gupta, M. (2006). *Why MultiLayer Perceptron/Neural Network?* Retrieved from Massachusetts Institute of Technology Media Courses: https://courses.media.mit.edu/2006fall/mas622j/Projects/manu-rita-MAS_Proj/MLP.pdf

Majaski, C. (14 April, 2019). *Fundamental vs Technical Analysis: What's the Difference?* Retrieved from Investopedia Web site: https://www.investopedia.com/ask/answers/difference-between-fundamental-and-technical-analysis/

M'ng, J. C., & Aziz, A. A. (2016). Using Neural Networks to Enhance Technical Trading Rule Returns: A Case with KLCI. *Athens Journal of Business and Economics, Vol. 2, Issue 1*, 63-70, DOI: 10.30958/ajbe.2-1-5.

Monfared, S. A., & Enke, D. (2014). Volatility Forecasting using a Hybrid GJR-GARCH Neural Network Model. *Procedia Computer Science 36* , 246-253.

Nabipour, M., Nayyeri, P., Jabani, H., Mosavi, A., Salwana, E., & S., S. (2020). Deep Learning for Stock Market Prediction. *Entropy, 22*, 840; doi:10.3390/e22080840.

Nishida, K. (9 March, 2017). *Introduction to Extreme Gradient Boosting in Exploratory*. Retrieved from Medium Web site: https://blog.exploratory.io/introduction-to-extreme-gradient-boosting-in-exploratory-7bbec554ac7

Pohlert, T. (2014). *The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR)*. Retrieved from R package: https://cran.r-project.org/web/packages/PMCMR/vignettes/PMCMR.pdf

Prabhakaran, S. (2017). *Logistic Regression*. Retrieved from r-statistics.co Web site: http://r-statistics.co/Logistic-Regression-With-R.html

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. Retrieved from R Foundation for Statistical Computing: https://www.R-project.org/

Rong, X. (2014). *deepnet: deep learning toolkit in R*. Retrieved from R package version 0.2: https://CRAN.R-project.org/package=deepnet

Rout, A. R. (2 September, 2020). *Advantages and Disadvantages of Logistic Regression*. Retrieved from GeeksforGeeks Web site: https://www.geeksforgeeks.org/advantages-and-disadvantages-of-logistic-regression/

Ryll, L., & Seidens, S. (6 July, 2019). *Evaluating the Performance of Machine Learning*. Retrieved from arXiV: https://arxiv.org/abs/1906.07786

Sagir, A. M., & Sathasivam, S. (2017). The use of artificial neural network and multiple linear regressions. *MATEMATIKA, 2017, Volume 33, Number 1, 1-10*. Penerbit UTM Press.

Schwager, J. D. (2017). *A complete guide to the future market: fundamental analysis, technical analysis, trading, spreads and options Second Edition.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Shafii, N. H., Ramli, N. E., Alias, R., & Fauzi, N. F. (2019). Fuzzy Time Series and Geometric Brownian Motion in Forecasting Stock Prices in Bursa Malaysia. *Jurnal Intelek Vol 14, Issue 2* , 240-250.

Soon, G. K., On, C. K., Rayner, A., Patricia, A., & Teo, J. (2018). A CIMB Stock Price Prediction Case Study with Feedforward Neural Network and Recurrent Neural

Network. *Journal of Telecommunication, Electronic and Computer Engineering Vol. 10 No. 3-2,* 89-94.

Sze, V., Chen, Y.-H., Yang, T.-J., & Emer, J. (13 August, 2017). *Efficient Processing of Deep Neural Networks: A Tutorial and Survey* . Retrieved from arxiv.org: https://arxiv.org/pdf/1703.09039.pdf

Tan, V. (16 July, 2020). *COVID-19 lockdown stimulates Malaysia's retail investor boom*. Retrieved from CNA News Asia Web site: https://www.channelnewsasia.com/news/asia/malaysia-covid-19-lockdown-retail-investor-boom-share-trading-12894640

The Star. (2020). *Bursa to consolidate to 1,500-1,530 next week on prolonged bargain-hunting*. Retrieved from The Star Web site: https://www.thestar.com.my/aseanplus/aseanplus-news/

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., & Manzagol., P.-A. (2010). Stacked Denoising Autoencoders: Learning Useful Representations ina Deep Network with a Local Denoising Criterion. *Journal of Machine Learning Research 11*, 3371-3408.

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). *Deep Learning for Computer Vision: A Brief Review.* . Retrieved from Computational intelligence and neuroscience, 7068349.: https://doi.org/10.1155/2018/7068349

Wei, N. S. (2020). *Bursa: Retail investors need to analyse, assess companies' fundamentals before investing in market*. Retrieved from The Edge Markets Web site: https://www.theedgemarkets.com/article/bursa-retail-investors-need-analyse-assess-companies-fundamentals-investing-stock-market

Weidman, S. (2017 December, 2017). *The 4 Deep Learning Breakthroughs You Should Know About*. Retrieved from Towards Data Science Web site:

https://towardsdatascience.com/the-5-deep-learning-breakthroughs-you-should-know-about-df27674ccdf2

Yan, Y. (2016). *Mean Square Error Loss*. Retrieved from RDocumentation Web site: https://www.rdocumentation.org/packages/MLmetrics/versions/1.1.1/topics/MSE

Yao, J., Tan, C. L., & Poh, H.-L. (1999). Neural Networks for Technical Analysis: A Study on KLCI. *International Journal of Theoretical and Applied Finance ,Vol. 1, No. 2,* 221-241.

Yiing, A. T., & Thim, C. K. (2015). Prediction of Bursa Malaysia Stock Index using Autoregressive Integrated Moving Average and Artificial Neural Network.

Yong, B. X., Rahim, M. R., & Abdullah, A. S. (2017). A Stock Market Trading System using Deep Neural Network. *AsiaSim 2017, Part I, CCIS 751,* 356–364.

Zaini, B. J., Mansor, R., Yusof, N., & Sang, B. H. (2019). Classify Stock Market Movement Based on Technical Analysis Indicators Using Logistic Regression. *Journal of Advanced Research in Business and Management Studies 14, Issue 1*, 35-41, ISSN: 2462-1935.

APPENDICES

Appendix A: Features description used for ML algorithms

1. Basic symbols	Explanation	
H[i]	H[i] represents the highest price of a stock on the i day, where H indicates the highest price time series of a stock.	
L[i]	L[i] represents the lowest price of a stock on the i day, where H indicates the lowest price time series of a stock.	
C[i]	C[i] represents the closing price of a stock on the i day, where H indicates the closing price time series of a stock.	
O[i]	O[i] represents the opening price of a stock on the i day, where H indicates the opening price time series of a stock.	
V[i]	V[i] represents the volume of a stock on the i day, where H indicates the volume time series of a stock.	
SMA(x, n)	The n order simple moving average of the time series x.	
EMA(x, n)	The n order exponentially moving average of the time series x.	
1：N	1:N represents all positive integers from 1 to N.	
runSum(x, n)	runSum(x, n) indicates the rolling sum of the order n of the sequence x, for example, x=1,2,3,4,5,6,7, then runSum (x, 3) is NA, NA, 6, 9, 12, 15, 18.	
HH[i]	HH[i] represents the maximum value in the highest price sequence.	
LL[i]	LL[i] represents the minimum value in the lowest price sequence.	
runMean(x, n)	runMean(x, n) represents the rolling mean of the n order of the sequence x.	
runSD(x, n)	runSD(x, n) represents the rolling standard deviation of the n order of the sequence x.	
2. Technical indicators	Calculation method	Explanation
(1) ADX	http://www.fmlabs.com/reference/default.htm?url=DI.htm http://www.fmlabs.com/reference/default.htm?url=DX.htm http://www.fmlabs.com/reference/default.htm?url=ADX.htm	The ADX is a Welles Wilder style moving average of the Directional Movement Index (DX). The values range from 0 to 100, but rarely get above 60. To interpret the ADX, consider a high number to be a strong trend, and a low number, a weak trend.
(2) OBV	If C[i]>C[i-1]，OBV[i]=OBV[i]+V[i] If C[i]<C[i-1]，OBV[i]=OBV[i]-V[i] If C[i]=C[i-1]，OBV[i]=OBV[i]	The On-Balance Volume (OBV) is a cumulative total of the up and down volume. A series of rising peaks, or falling troughs, in the OBV indicates a strong trend. If the OBV is flat, then the market is not trending.
(3) WR	%WR[n]=100*（HH[1:n]-C[n]）/(HH[1:n]-LL[1:n])	The values range from zero to 100 and are charted on an inverted scale, that is, with zero at the top and 100 at the bottom. Values below 20 indicate an overbought condition and a sell signal is generated when it crosses

		the 20 line. Values over 80 indicate an oversold condition and a buy signal is generated when it crosses the 80 line.
（4）RSI	If C[i]>C[i-1], then up[i]=C[i]-C[i-1], dn[i]=0; If C[i]<=C[i-1], then dn[i]=C[i-1]-C[i], up[i]=0; Upave[i]=(upave*(i-1)+up)/(i); Dnave[i]=(dnave*(i-1)+dn)/(i); RSI[i]=100*upave[i]/(upave[i]+dnave[i])	The RSI is interpreted as an overbought/oversold indicator when the value is over 70/below 30. You can also look for divergence with price. If the price is making new highs/lows, and the RSI is not, it indicates a reversal.
(5) NCO	NCO[i]=C[i]-C[12]	Net Change Oscillator Index is the change of series over n periods.
(6) PO	PO=(SMA(C, 5)-SMA(C, 10))/(SMA(C, 10))	Detrended Price Oscillator Index removes the trend in prices by subtracting a moving average of the price from the price. The PO shows cycles and overbought/oversold conditions.
(7) Chaikin A/D Oscillator	AD[i]=AD[i-1]+(C[i]-L[i])+ (((C[i]-L[i])-(H[i]-C[i]))/(H[i]-L[i]+0.01))*V[i] CO=EMA（AD，3）- EMA（AD，10）	Chaikin A/D Oscillator is a stock index related to trading volume, which can be used to observe the flow of funds in the market.
(8) MOM	MOM[i]=(C[i]/C[i-9])*100	Momentum Index is the (rate of) change of a series over n periods.
(9) ROC	ROC[i]=log(C[i]/C[i-1])	The ROC indicator provides the percentage difference of a series over two observations
(10) SROC	SROC=(EMA(C, 20)/EMA(C,10))*100	Smoothed Rate of Change Index, like ROC, is used to reflect the rate of change in stock prices.
(11) Return	Ret=log(C[i]/C[i-1]) Return=runMean(Ret, 14)	Return represents means of logarithmic return rate over a n-period moving window.
(12) Sigma	Ret=log(C[i]/C[i-1]) Return=runSD(Ret, 14)	Sigma represents standard deviations of logarithmic return rate over a n-period moving window.
(13) CCI	TP[i]=（HH[1:i]+LL[i]+C[i]）/3 ATP=SMA(TP, 20) MDTP=runMean(|TP-ATP|,20) CCI=(TP-ATP)/(0.015*MDTP)	The Commodity Channel Index (CCI) attempts to identify starting and ending trends.

(14) RSV	RSV[i]=100*（C[i]-LL[(i-8):i]）/(HH[(i-8):i]-LL[(i-8):i])	The RSV index is mainly used to analyse whether the market is in an overbought or oversold state. The market is overbought when RSV is higher than 80%; The market was oversold when RSV was below 20%.
(15) Kvalue	Kvalue=EMA(RSV, 2)	The K, D, and J index also known as Stochastic Oscillator. It can be used to judge the market more quickly and intuitively and is widely used in the analysis of the short- and medium-term trend of the stock market.
(16) Dvalue	Dvalue=EMA(Kvalue，2)	
(17) Jvalue	Jvalue=3*Kvalue-2*Dvalue	
(18) MACD	Fast=EMA（C，12） Slow=EMA (C, 26) DIF=Fast-Slow MACD=EMA(DIF, 9)	The MACD signals trend changes and indicates the start of new trend direction.
(19) CAD	CAD[i] =CAD[i-1] + V[i]* CLV[i]	The Chaikin Accumulation / Distribution (CAD) line is a measure of the money flowing into or out of a security. It is similar to OBV.
(20) VOLA	EMAHL=EMA(H-L，10) VOLA[i]=(EMAHL[i]-EMAHL[i-9])/EMAHL[i-9]	Chaikin Volatility measures the rate of change of the security's trading range.
(21) Label	If log(C[i+1]/C[i])>0, Label[i]=1, else Label[i]=0	The classified label is an important sign to supervise learning algorithm.

Appendix B: Example algorithms in performing statistical testing

```
> # significant test
> setwd("~/THESIS/SCRIPT/BURSA")
>
> #Kruskal Wallis Test- to check if there is a significant diff (p-
value<0.05)
> si_func=function(alg_name,sig_name){
+   modelFile=paste("BURSA30_Performance_ML_",alg_name,".csv",sep = "")
+   sig_f=NULL
+   for(i in 1:N){
+     sig_f0=read.csv(modelFile[i],header = T,sep = ",")[,sig_name]
+     sig_f= c(sig_f,sig_f0)
+   }
+   sig_f = as.numeric(unlist(sig_f))
+   sig_name = factor(rep(1:N, each=length(sig_f0)),labels = alg_name)
+   kruskal.test(sig_f,sig_name)
+ }
>
> sig_dir=c("AR","PR","RR","F1","AUC","MSE")
> sig_perf=c("WR","ARR","ASR","MDD")
> alg= c("LR","SVM","XGB","DBN","MLP","SAE")
> strgy= c("Index","B&H","LR","SVM","XGB","DBN","MLP","SAE")
>
> si_func(alg, AR)

      Kruskal-Wallis rank sum test

data:  alg and AR
Kruskal-Wallis chi-squared = 155.41, df = 5, p-value < 2.2e-16

> #next use nemenyi test also known as post hoc test- to check the difference
between trading strategies
> sig_func=function(alg_name,sig_name){
+   modelFile=paste("BURSA30_Performance_ML_",alg_name,".csv",sep = "")
+   N=length(alg_name)
+   sig_f=NULL
+   for(i in 1:N){
+     sig_f0=read.csv(modelFile[i],header = T,sep = ",")[,sig_name]
+     sig_f= c(sig_f,sig_f0)
+   }
+   sig_f = as.numeric(unlist(sig_f))
+   sig_name = factor(rep(1:N, each=length(sig_f0)),labels = alg_name)
+   PMCMR::posthoc.kruskal.nemenyi.test(sig_f,sig_name)
+ }
> sig_func(alg, AR)

      Pairwise comparisons using Tukey and Kramer (Nemenyi) test
        with Tukey-Dist approximation for independent samples

data:  alg and AR

    LR      SVM     XGB     DBN     MLP
SVM 0.00045 -       -       -       -
XGB 0.11618 0.57501 -       -       -
DBN 0.00371 1.3e-13 1.0e-08 -       -
MLP 1.4e-05 7.8e-14 1.6e-12 0.80226 -
SAE 2.0e-05 4.0e-14 2.8e-12 0.84104 1.00000

P value adjustment method: none
Warning message:
In posthoc.kruskal.nemenyi.test.default(sig_f, sig_name) :
Ties are present, p-values are not corrected.
```