Ideas:

Gamestop bubble, is there any way to predict this?

Can we predict when a particular stock will reach its peak?

Which stock is the most predictable?

Which stock should who buy?

When to buy, when to sell?

Machine Learning focused:

compare different ML techniques on the predictions of stock market trends

Introduction to explain the purpose of the review and the motivation of the research topic

Find the correlation of variables effecting the forecast error/stockmarket

Thinking of a research topic

1. Extension of a current model?
2. Correction/modification of a current model?
3. Something new from something old?
4. Same case, but different methodology and ML method to prove/disprove the prediction
5. Study of a new topic built up on several old ones
6. Evaluation and corrections on methodologies?’
7. Evaluation of a different case of the same topic using same/different/similar methodologies
8. Consider the gaps of the research🡪 fill in the gaps of a current research
9. Find correlations of variables which are not discussed

Learning about research methods, the different steps and narrowing down

See how and what data needs to be collected

Learning algorithms

Learning the different challenges and different factors we need to include

Learning the limitations of different models

Shed insight of what future work has to be done

Learning the structure and streamline of the report: Intro🡪 Methodology 🡪 Results 🡪 Conclusion 🡪 different data that needs to be shown

Body paragraphs to compare the literature (e.g. their main insights, arguments and conclusions)

Typically organised by themes or dates/periods

Themes could be split by topic questions or methodologies

* What problem is the author trying to solve?

This literature tries to find a function which demonstrates the likelihood of market fluctuation, the ups and downs. It is still not clear that machine learning algorithms are extracting information beyond that contained in autocorrelation patterns.

This paper reexamines the direction-of-change predictability in weak-form tests.

They used ML algorithm named Adaboost a general method for improving the performance of any learning algorithm:

Steps:

1. Start with weights
2. Compute error/loss function (), which is a function of an indicator function and transform its output so the weights normalizes to 1 using an activation function c\_m

the lack of autocorrelation in stock returns does not permit Adaboost to discover a function that discriminates between upwards and downwards movements better than random

simple random classifiers (i.e., cointoss classifiers) are able to explain the apparent predictability in such periods.

* What are the key concepts in the model/method? And how were they derived (i.e. the assumptions and theory behind them)?

Declare stock price movements {y}

Yt = \phi + et

ε is the noise component and is the “output” or “response” variable y y ∈C

positive equity premiums were codified with 1’s. Hence, we consider here a two-class case, i.e., {0,1}

using Adaboost

using 2-node tree-based models/stumps as base learners

Tree-based techniques involve partitioning the explanatory variables space into a set of rectangles and then fit a simple model to each one. A tree-based model tries to find the split that maximizes the decrement in a loss function in order to make a tree grow. This is done iteratively until a certain amount of observations is reached or no further decrements in the loss function are found. The main problem is to find the region Rj, which we need to use heuristic methods, such as top-down recursive partitioning, which starts with a single region covering the entire space of all joint input values.

Replaced loss function with the Gini index

* What was innovative about the approach used in this study?
* What are the results and conclusions of the study? Is it justified by the evidence? Or any there any bias, errors, or inconsistencies?

Using data from S&P 500 daily closing prices from August 7, 1962 to December 31, 2004

Data set divided into 4 non-overlapping sets 🡪 each set into two sub-samples

First sub sample for training 🡪 second for testing

ASSUME future stock price movements {y} may be related to past returns

Results:

Table 1:

1. Error rate🡪 the total number of misclassified observations divided by the total number of observations
2. the bias 🡪 systematic loss incurred by the function
3. unbiased variance (denoted by Vu) 🡪 evaluates the extent to which the estimated function deviates from the correct predictions
4. biased variance (denoted as Vb) 🡪 assesses the extent to which the estimated function deviates from the incorrect predictions

First two data sets: bias plays a significant role in its contribution to the error rate. In other words, the systematic loss incurred by the functions is higher than the total error rate Adaboost’s error has a positive relationship with the total number of iterations this later result indicates that Adaboost rapidly over-fits the data.

Last two data sets: bias is lower than the error rate 🡪 only occurs when the loss incurred by function’s fluctuations around the central tendency in response to different samples has a direct effect on error. (error decrease has the number of iterations increases)

Simulated 1000 coin-toss classifiers for each data set to analyze the extent the results can be explained by randomness, only 1 and 0 possible, with 0.5 probability each

Table 2:

randomness can explain up to 46 percent, approximately, of out-of-sample errors 🡪 classifiers achieving higher out-of-sample error rates can be considered as random

only in the first two data sets was Adaboost able to obtain lower out-ofsample error rates

Answering the question of what factors that affect Adaboost’s ability to discriminate between stock price movements 🡪 gauging traditional benchmarks 🡪 can evaluate whether or not simple linear models are able to explain Adaboost’s predictability 🡪 estimated a simple first-order autoregressive model for each period 🡪 hence going to table 3

Table 3:

Displays the same accuracy measures as Table 1

the autoregressive models are able to obtain in-sample predictability but fail to detect out-of-sample predictability in the last two data sets 🡪 The disappearance of the predictability cohererent with other literature results

Conclusion:

implemented a classifier induction approach to analyze the sample evidence on return predictability

General results:

1. periods characterized by high first-order serial correlation in stock returns allow both in-sample and out-of-sample direction-of-change predictability🡪 using Adaboost to find a stable function which discriminates, better than randomly made decisions, between upward and downward movements
2. , Adaboost does over-fit 🡪 Functions induced in periods characterized by the lack of autocorrelation in stock returns are able to obtain in-sample predictability but fail to detect out-of-sample predictability.

Adaboost’s out-of-sample performance decreases as more iterations are run

examined different Adaboost specifications, such as using 4- and 8-node tree-based models instead of stumps, and achieved faster over-fitting.

Natural extensions:

1. machine learning algorithms can be used to examine large price change predictability 🡪 can also be modified to study predictability of large absolute price movements, which are useful for option trading strategies.
2. machine learning algorithms are sufficiently flexible to examine the performance of nested models 🡪 one can induce classifiers for small-cap indices using small-cap’s or large-cap’s lags, and evaluate the lead-lag effect in terms of movement predictability
3. machine learning algorithms can be used to identify risk exposures 🡪 codify costly lower-tail outcomes and search for “inputs” or “explanatory” variables that help a machine learning algorithm discriminate between the costly lower-tail outcomes and the remainder of outcomes

* How does this paper compare to the other papers on this topic? Does it confirm or challenge them?
* What are the key insights and arguments of the paper?
* What are the strengths and weaknesses of the study?

Ref: P. Rodriguez and S. Sosvilla Rivero, "Using Machine Learning Algorithms to Find Patterns in Stock Prices", *SSRN Electronic Journal*, 2006. Available: 10.2139/ssrn.893141 [Accessed 5 February 2021].

* What problem is the author trying to solve?

Forecasting is challenging application for the growing of information technology 🡪 Stock price prediction is a method by which the stock market data at a specific time in the past is processed to forecast the stock price in a future term

Time series prediction using neural networks: This study applies a hybrid method of Genetic Algorithm (GA) and Artificial Neural Network (ANN) technique to develop a method for predicting stock price and time series 🡪 At first, the GA approach optimizes the weights with the SSE cost function such that desired values are calculated from a sigmoidal function. In continuation, ANN control the answers, and GA outputs decrease the iterations of ANN.

Sudden changes in price of stock cause deficient accuracy when method use ANN+BP approaches together. In the beginning the GA is used for initializing the weights, and in this mechanism at the end, proposed method can work with any price of stock market (like Apple, Pepsi, IBM and etc.)

analysis suggests that the GA and ANN can increase the accuracy in fewer iterations. The analysis is conducted on the 200-day main index, as well as on five companies listed on the NASDAQ

Applying proposed method to the Apple stocks dataset, based on a hybrid model of GA and Back Propagation (BP) algorithms, the proposed method reaches to 99.99 % improvement in SSE and 90.66 % in time improvement, in comparison to traditional methods.

From previous research: grouped under two categories: statistical and artificial intelligence model

* for good forecasting accuracy, a large amount of information should be considered
* Stock indices usually contain four factors, which include Open, High, Low, and Close
* The higher accuracy has the results with minimum error, and can predict long time periods based upon past data
* Gas, ANNs, fuzzy logic methods were successful in stock price forecasting
* Feed forward neural network or multi-layer perceptron (MLP) predicts in two steps: training the network, and forecasting the future data. Training set includes majority volume of data, though 10 % to 25 % of the data set, uses for the test step
* ANN: can be used in combination with another algorithm for better and more accurate results, eg with support vector machine (SVM), Decision Tree (DT); or combine two DT models 🡪 results show ANN + DT is more accurate than DT
* [80% data is used in training;

the accuracy of ANN and DT combined model is higher than two combined DT models. In this research, BP and Delta-Rules are used to establish the ANN model and DT is used to predict the selling or buying of the stock. The output in this method is 1 or 0 (sigmoid function), which means stock price will rise or fall, respectively]

* Predicting stock price with noisy data based on back propagation neural network (BPNN) method 🡪 decomposes the data set of several layers and then it uses the BP model on every layer for prediction 🡪 Each BP network has weighted connections between nodes (also known as neuron) 🡪 these weights change until the mean square error (MSE) is to be minimized, though each node can only solve a linear problem
* forecasting based on heuristic search and genetic algorithm/genetic programming (GA&GP) (problem space is large and not well known): many advantage points over custom models, such as vector auto regression (VAR) model in forecasting
* ,financial forecasting, trading strategies, trading system development, volatility modeling and important problems in the finance domain can be solved easily.
* , GA is used to improve the learning and reduce complexity in the problem space. Experts select 12 features including momentum, rate of change (roc), based on prior research
* financial problems have a strong relation to time 🡪 financial time-series data are important to predict the data of a future period 🡪 GA can be singly/combined with other methods to search problem domains
* SVM 🡪 linear model to classify non-linear data 🡪 SVMs transfer data from non- linear space to linear space with a specific function and in the new space they can classify data with linear methods
* Multi-Layer Perceptron (MLP) + BO for training and classification; compared with probabilistic neural network PNN for showing performance
* Prediction of stock price one day ahead: used ANN for forecasting and fuzzy method for analyzing the predicted values
* at first, the data were processed and transformed from real-world data to a new dataset vector, following which the processed data were used by ANN input and the forecasted outputs were then analyzed by a fuzzy system
* used ANN and autoregressive integrated moving average (ARIMA) 🡪 combined method can improve the accuracy in forecasting to a greater extent than when these methods are used separately. ARIMA is a linear model, and the result value of this work is a linear function that obtained from past data
* research on the benchmark of ensemble approaches: Random Forest, Ada-boost and Kernel Factory, in comparison to the single classifier models such as Neural Networks, Logistic Regression, Support Vector Machines and K-Nearest Neighbors
* the combination models always produced much better results than single algorithm application
* What are the key concepts in the model/method? And how were they derived (i.e. the assumptions and theory behind them)?
* assumed that the price of tomorrow stock with a satisfying accuracy can be predicted based on past prices of stock
* What was innovative about the approach used in this study?
* This work is proposed for modeling a technique based on opening and closing prices.
* Time series data is used: None of the time series predictive methods are used
* simple ANN+BP is used (no recurrent networks or others).
* the used methods (ANN with BP and GA) are trained on input and output (open and close price) data of time series.
* method provides a global pattern between data. Using past data, an attempt has been made to create a model for prediction. Time series prices are used as an observation of history of stock market.
* uses Genetic algorithm (GA) to find weights of ANN algorithm, and then implements the ANN to set errors on a fixed number with minimum iteration 🡪 to reduce time taken, as it consumes much time and many iterations are required to train an ANN [when new data is being added to a set, the process faces a delay, and causes a repeat]
* Data in population are called chromosomes. These chromosomes are capable of potential answers that with crossover or mutation, they can create the best response.
* back propagation (BP) to train ANN 🡪 It is a common method for training a neural network and it uses delta rule with iterations to find the best weights between the units (intricately explained). In the proposed method, the time consumption is reduced and a highly accurate forecasting is reached.
* BP minimizes the sum square error SSE (cost function by gradient descent)
* Sigmoid function to normalize between 0 and 1
* this method uses GA for initializing the weights that are used in ANN method. Moreover, the fitness function of GA method is a part of BP rule for calculating the SSE.
* What are the results and conclusions of the study? Is it justified by the evidence? Or any there any bias, errors, or inconsistencies?
* Results: result of this paper shows that ensemble algorithms that are included in the benchmark are ranked as the top approaches (Random Forest, Ada-boost and Kernel Factory, in comparison to the single classifier models such as Neural Networks, Logistic Regression, Support Vector Machines and K-Nearest Neighbors)
* assumption is correct with a high accuracy (assumed that the price of tomorrow stock with a satisfying accuracy can be predicted based on past prices of stock)
* high accuracy in less time
* initially, the GA for preliminary of ANN weights is used. With the implementation of GA, more accurate weight in a short time is found, and then with the use of ANN the SSE on a specific value is fixed. ANN with BP algorithm can minimize SSE in each iteration
* a very good improvement in terms of accuracy which is 99.42 % in SSE and 88.75 % reduction in time consumption

Future work: it is suggested to combine the proposed method with other methods like SVM or DT approaches or combined model of these approaches.

* How does this paper compare to the other papers on this topic? Does it confirm or challenge them?
* What are the key insights and arguments of the paper?
* What are the strengths and weaknesses of the study?

Ref: O. M. E. Ebadati and M. T. Mortazavi, "AN EFFICIENT HYBRID MACHINE LEARNING METHOD FOR TIME SERIES STOCK MARKET FORECASTING: INTERNATIONAL JOURNAL ON NEURAL AND MASS - PARALLEL COMPUTING AND INFORMATION SYSTEMS," *Neural Network World,*vol. 28, *(1),*pp. 41-55, 2018. Available: https://www-proquest-com.libproxy.ucl.ac.uk/scholarly-journals/efficient-hybrid-machine-learning-method-time/docview/2012893389/se-2?accountid=14511. DOI: http://dx.doi.org.libproxy.ucl.ac.uk/10.14311/NNW.2018.28.003.

* What problem is the author trying to solve?
* implement an orthogonality test of the rationality of aggregate(forming a class of) stock market forecasts
* Main finding: given our set of predictor variables, the rational expectations hypothesis (REH) cannot be rejected for short-term forecasts and that there is evidence against the REH for longer term forecasts
* What are the key concepts in the model/method? And how were they derived (i.e. the assumptions and theory behind them)?
* ML METHOD: BRT (boosted regression tree)algorithm endogenously selects the predictor variables used to proxy the information set of forecasters so as to maximize the predictive power for the forecast error (measuring the forecast error)🡪 also accounts for a potential non-linear dependence of the forecast error on the predictor variables and for interdependencies between the predictor variables
* Main idea: replace the systematic part of a conventional regression model, et + 1 = Xt𝛽, with a more general model with a systematic part of the format et + 1 = T(Xt, R), where T denotes a regression tree and R denotes the parameters of the regression tree
* Regression trees: captures even complex non-linearities in the potential link between forecast errors and thevariablesina forecaster's information set; on non-linearities in the formation of stock market expectations.

1. Non – linearities arises when the strength of the correlation between variables in the information set, It, is different for small and for large forecast errors
2. RTs provide a natural platform for investigating potential interaction effects between the variables in a forecaster's information set 🡪 the strength of the correlation between forecast errors and, say, the dividend yield depends on the value assumed by some other variable like the T-bill rate
3. they allow the relative importance of the variables in a forecaster's information set, It, to be assessed and that they are insensitive to the inclusion of irrelevant variables 🡪 potentially important for studying the orthogonality property of stock market forecasts because a researcher cannot directly observe the information set 🡪 to proxy the information available to a forecaster as close as possible, a researcher can use various variables known from earlier literature to predict stock returns. Some of these variables, however,arelikely to be redundant
4. are robust to outliers in the data. This property is useful given the financial market jitters of the recent past
5. they have the drawback that their specific structure makes them high-variance predictors, but overcame by BRT alg by using boosting techniques to additively combine several regression trees to form a low-variance predictor

* Orthogonality test:

1. e represented by a boosted model with a systematic part of the format et+1 = ∑ TT(Xt, RT)
2. individual trees are not estimated in a single step but rather in an iterative stagewise process 🡪 iterative process🡪 combines an ensemble of simple base learners (i.e., the individual trees) in an additive way to build a potentially complicated function known as a strong learner
3. resulting boosted ensemble of trees is then used to model the forecast errors

* Previously used to do: applications of regression trees in monetary and financial economics // on the determinants of financial crises // used by economists to model exchange rates & forecast output // model stock market volatility, and Ng (2014) and Döpke, Fritsche, and Pierdzioch (2017), who use BRT techniques to forecast recessions
* For re-examine the orthogonality property of forecast errors implied by aggregate stock market forecasts as represented by a consensus forecast implied by the stock market forecasts from the Livingston Survey.
* Gradient-descent least-squares boosting aims at minimizing the squared error loss function, L = (y − G(x))2, by using a stagewise algorithm to approximate the unknown function, G(x) 🡪 G(x) captures how the realizations of the predictor variables map into predictions of the response variable. The unknown function, G(x), consists of an additive combination of regression tree // G(x) is a strong learner, individual regression trees are base learners
* learning rate, 0 < 𝜁 ≤ 1 to curb the influence of individual regression trees on the approximation of the function G(x) and, thereby, to strengthen the robustness of the BRT algorithm
* If possible to forecast the forecast error🡪 rational expectation hypothesis (REH) should be rejected // if impossible to forecast the error 🡪 REH should be accepted
* Data: semi-annual Livingston Survey maintained by the Federal Reserve Bank of Philadelphia (2014) is particularly useful in this regard because it contains forecasts of the S&P 500 stock market index for various forecast horizons
* Measuring the forecast error:

1. stock market forecasts are correlated with various variables in the forecasters' information set, where evidence against orthogonality is stronger at a forecast horizon of 12 months than for 6-month-ahead forecast
2. forecast errors are correlated with information available to forecasters at the time a forecast was made
3. against the REH, where orthogonality tests suggest that forecast errors are correlated with various macroeconomic variables
4. report that stock market forecasts are unbiased and informationally efficient
5. stock market forecasts from the Livingston Survey are not efficient and that forecasts do not outperform naive benchmark models
6. t forecasts do not outperform the historical mean or small-scale regression models using the dividend yield or the T-bill rate as predictors

* What was innovative about the approach used in this study?
* What are the results and conclusions of the study? Is it justified by the evidence? Or any there any bias, errors, or inconsistencies?

RESULTS:

* REH cannot be rejected for short-term forecasts. We find evidence against the REH for longer term forecasts. Results for three different groups of forecasters corroborate our main findings
* Documented the relative importance of various predictor variables and marginal effects that capture how the forecast errors depend on the predictor variables
* findings for forward forecast errors, which are computed by combining forecasts for different forecast horizons
* machine-learning technique known as boosted regression techniques can be a useful modeling platform for studying the rationality of survey forecasts
* , the REH cannot be rejected for aggregate short-term stock market forecasts. This finding is consistent with results reported in earlier research (Brown & Maital, 1981).
* there is evidence against the REH for longer term aggregate stock market forecasts.
* evidence against the REH when we study forward forecast errors
* results are conditional on the predictor variables we have used to model the forecasters' information set and that we have analysed aggregate stock market forecast
* aggregate stock market forecasts may also cloud important differences across individual forecasters. Such differences may arise in terms of the rationality of forecasts, but also in terms of, for example, (rational) forecaster anti-herding (Pierdzioch & Rülke, 2012) and the shape of the loss function 🡪 accounting for these differences: reported three different groups of forecasters:

1. Group 1 consists of forecasts from academics and government economists.
2. Group 2 consists of forecasts from forecasters from the following four industries: commercial banking, consulting, insurance company, and investment banking.
3. Group 3 consists of forecasts from forecasters belonging to one of the following sectors: industry trade group, labour, and nonfinancial sector.

* LIMITATIONS AND FUTURE WORK
* study whether well-known predictor variables have predictive value for the sign of the forecast error implied by stock market forecasts 🡪 predictor variables can be found that help to predict the sign of the forecast error then it may be possible to build a simple trading strategy that exploits this predictive value 🡪 well known that the statistical value of predictions is not necessarily closely correlated with their economic value and that directional accuracy typically is a better metric of the economic value of foreasts than other statistical criteria
* Adapting the BRT algorithm such that it can be used to study the sign of the forecast error implied by stock market forecasts is straightforward 🡪 all a researcher has to do is to replace the regression trees that we have studied in our empirical research with classification trees 🡪 a classification-tree-based approach that sheds light on the predictability of the sign of the forecast error would open up the possibility to apply the BRT algorithm to tackle many more interesting research questions in empirical finance
* How does this paper compare to the other papers on this topic? Does it confirm or challenge them?
* What are the key insights and arguments of the paper?
* What are the strengths and weaknesses of the study?

Ref: C. Pierdzioch and M. Risse, "A machine-learning analysis of the rationality of aggregate stock market forecasts", *International Journal of Finance & Economics*, vol. 23, no. 4, pp. 642-654, 2018. Available: 10.1002/ijfe.1641 [Accessed 5 February 2021].

* What problem is the author trying to solve?

betting on the direction of stock markets is regarded as a high-risk strategy because there are too many external factors affecting it 🡪 this strategy could significantly influence portfolio returns with only a slight change in proportion of allocation of assets owing to high volatility 🡪 establishing a strategy for assigning weights to stocks is crucial for portfolio returns in asset management fields

author wants to design an optimal investment strategy (a regional allocation strategy), and it is important to accurately predict the market by understanding its characteristics

(need to consider the relative attractiveness of specific regions in terms of investment returns 🡪 portfolio managers adjust the proportions over regions based on a pre-assigned benchmark regional weight, as it is less risky than stock market directional strategy, it is used more frequently -- > regional allocation strategy contributes as much as stock market directional strategy to global portfolio returns)

Prev studies: on market valuation based on price or financial statement information; focused on bubble signals to capture market collapses; apply machine learning techniques as they have shown relative success in predicting financial time series; use of price technical parameters and compared the performances of several machine le models for stock prediction; applied a genetic algorithm for selecting machine learning model input variables. In addition, multiple machine learning models have been combined for constructing forex portfolio trading strategies; applied a genetic algorithm for selecting machine learning model input variables. In addition, multiple machine learning models have been combined for constructing forex portfolio trading strategies; study has been published which uses online data sources including google trends, Wiki and financial news with an ensemble learning model; Volatility of markets is also considered an important factor to describe market environments; there has been research on forecasting market volatility with machine learning

Characterisitic of for global investment: stock markets of countries across the world are closely linked 🡪 need to examine the global linkage effect, which can be obtained from the complex connection among various markets

By constructing several banking systems with various key parameters of the structure of the financial system, the authors showed that concentrated banking systems tend to be prone to systemic risk

suggested a model to measure systemic risk and estimated the risk contagion with financial connectivity.

o describe complex linkage structures over global regions or markets more effectively. The phenomenon of financial connectivity including risk contagion and market collapses has been explained

AIM: to present a practical methodology to design an optimal portfolio strategy of the global stock market with financial network indicators 🡪 coz there have been only a few attempts that construct investment strategies through these network indicators: expand to practical investment strategy and effects in portfolios 🡪 attempts to utilize them in the most important and difficult global financial market outlook and portfolio strategy

purpose of this study was to examine the effectiveness and enhancement of a financial network based on global stock indices of 10 countries in building a global stock portfolio strategy

Results: shows the impacts of the financial network on financial market among market turn-moil periods 🡪 proved it with more objective and quantitative method by data simulation, came up with 1) stock market prediction strategy (forecasting total stock market direction) and 2) regional allocation strategy (forecasting the relative direction of developed market (DM)/emerging market (EM))

global stock market network indicators are effective for forecasting market direction (up and down) during the market turmoil period as expected. Moreover, mid-term investments with short-term volatility showed better performance than other strategies in terms of stock market prediction

better performance obtained over one week is because the financial network measure does not immediately affect the market rise or fall but affects the implicit market risk, eventually influencing the market direction in the mid-term periods. The short-term volatility shows better performance among the other volatility periods as short-term is a more sensitive measure than others.

network indicators can be performance enhancers for the regional allocation strategy, especially during market volatility period

information of network in the regional allocation level is stably effective similar to that in the total market direction level over the prediction and volatility periods. The network measures are not used in markets for regional allocation yet, but we observed that they are good indicators for regional direction forecasting and allocation strategies.

usefulness of the network indicators can be confirmed

Special: first to demonstrate the applicability of network indicators to global investment. 🡪 suggests steps for the implementation of the strategy and shows the effect of network indicators

Method to develop the strategy: propose to construct volatility networks of the global stock market based on a simple pair-wise correlation and system-wide connectedness of national representative financial indices. 🡪 extending the study of an analysis on the US & Europe banking networks, we broadly construct a system-wide volatility network based on global stock indices 🡪 then examine the effect and usefulness of network indicators in terms of the global investment strategy

Machine learning approaches: logistic regression (LR), support vector machine (SVM), random forest (RF) and Deep Neural Network (DNN)

Predicting the directions of the total global stock market indices and regional relative attractiveness using simple index prices and network indicators 🡪 conduct sensitivity analyses by changing the forecasting and volatility periods to determine which period is the most appropriate for volatility construction to predict market directions

lastly recommend a suitable investment strategy and provide guidance regarding global portfolio management strategies

differences in the performance results among the three machine learning methodologies employed🡪 the effectiveness of SVM, a nonparametric approach, was higher than that of LR or RF. It can be concluded that the methodology of SVM is more appropriate for forecasting the market through a non-linear relationship among variables.

* What are the key concepts in the model/method? And how were they derived (i.e. the assumptions and theory behind them)?

1. Collect stock indices data of 10 individual countries and construct two global volatility networks by using their Pearson correlation and VAR (**Vector autoregression** (**VAR**)) model
2. , a time-series dataset including stock indices and connectedness indicators is converted to z-score based on historical time-series data of 52 weeks as they are observed in different scales over countries and observation periods
3. , a time-series dataset including stock indices and connectedness indicators is converted to z-score based on historical time-series data of 52 weeks as they are observed in different scales over countries and observation periods
4. evaluate the performance of our proposed approach and propose global investment portfolios based on the global investment strategy, which constitutes different stock volatility and prediction periods

Resources/data used:

* 10 countries selected selected by considering the Morgan Stanley Capital International (MSCI) Index countries and the regional representation. Five of the 10 countries were classified as developed market (DM) and the remaining countries were classified as emerging market (EM) based on the MSCI world index market classification
* data were collected over a period of 22 years from January 1995 to December 2016, and all data were sourced from Thompson Reuters DataStream Database

Defining the financial volatility network: **Volatility** is a statistical measure of the dispersion of returns for a given security or market index. In most cases, the higher the **volatility**, the riskier the security. <https://www.investopedia.com/terms/v/volatility.asp#:~:text=Volatility%20is%20a%20statistical%20measure,same%20security%20or%20market%20index>.

* The volatility of a stock index is defined as the standard deviation of the daily log returns (each period for 1,3,8,12 weeks)
* 1. we construct a correlation-based network by using an edge with the volatility correlation between the stock indices of 10 countries 🡪 we construct a financial volatility network from the variance decomposition matrix of the VAR model
* 3-lag VAR model and H-step was used for 10 days (2 weeks)
* explanatory power of future forecast error variance of individual variables derived using variance decomposition indicates the degree of effectiveness of stock index of each country in the financial network. The countries have in and out constituents for effectiveness. The net values of effectiveness are the edges between countries 🡪 constituted a VAR network

To identify the structural changes in finial networks:

* we compare three periods (1996: stable period, 1997/2008: turmoil period)
* In Pearson: he Pearson correlation network shows that the inter-state volatility network was strongly built during the emerging market crisis and the 2008 global financial crisis
* In VAR: model, the soaring influence of Asian countries during the emerging market crisis was depicted and the US influence surged during the 2008 financial crisis. The VAR network provided further detailed information as it is a directed method

ML techniques:

* LR:
* forecasting model to predict the direction of movements of individual stocks and stock price indices
* a multivariate analysis model, is an appropriate model for predicting the presence of a characteristic with a set of predictor variables
* LR is useful when the dependent variable is binary or a multinomial categorical variable, it has commonly been used in the field of investments for the prediction of directions or classification of companies.
* RF: ensemble algorithms: classification learning technique developed by Breiman (2001)
* extension of traditional decision tree techniques, and has a meta-learning form of decision trees instead of one
* composed of multiple decision trees, which are randomly selected variables. The combined prediction trees are expected to increase the accuracy and stability of the model performances as compared with those of a single classification model
* the larger the size of the forest (the number of trees), the more the convergence of the generalization error to a specific value, and thus, the over-fitting can be avoided. RF uses randomly extracted data from the total training dataset and is not significantly affected by noise or outliers
* RF has been presented as the best predictive classification algorithm for imbalanced data sets according to Brown and Mues (2012). In this study, we use the Gini index to calculate the criterion of decision trees
* SVM: learning processes for determining decision boundaries should maximize the shortest distance (margin) to the boundary in order to determine the best classifier in the given data.
* well suited to many classification problems owing to the minimization of over-fitting, high accuracy, easy modification, and ability to handle high-dimensional data (Cortes & Vapnik, 1995
* requires a long time because it has a large computational complexity depending on the kernel function. In this study, we selected a radial basis kernel function
* , the performance of the reference model, which only uses the stock index Z-score of each country as inputs without connectedness measures, is compared with that of the suggested model, which uses connectedness measures and stock index Z-score
* attempted to recommend the most suitable model for portfolio strategy by analyzing the change in forecasting period and volatility depending on the observation period and the machine learning model.

Investment framework: (1) stock market prediction strategy (investing 1/10 each to 10 countries based on global stock market directional forecasting), (2) regional allocation strategy (investing 50% each to DM/EM regions as long/short position by comparing the attractiveness between DM/EM country stock indices)

* What was innovative about the approach used in this study?
* What are the results and conclusions of the study? Is it justified by the evidence? Or any there any bias, errors, or inconsistencies?

LIMITATIONS: For network analysis in the field of finance, symptoms of market crisis and structural changes could be closely described. However, practical usefulness studies are scant in this regard

1. we have used only a connectedness measure but we must develop other network indicators for capturing diverse market dynamics.
2. the three machine learning techniques are not sufficient for measuring the effects of numerous latent factors in complex financial markets 🡪 new methodology such as model improvement or deep learning will be considered in the future
3. forecasting indicators such as price and network indicators have only been used whereas other indicators that are already considered to be important in the market such as valuation and economic indicators have not been used

ML: in investment simulations, average performances are not good enough, with low returns for some parameter combinations. It is necessary to improve parameter selection techniques to obtain more robust simulation results

* How does this paper compare to the other papers on this topic? Does it confirm or challenge them?
* What are the key insights and arguments of the paper?
* What are the strengths and weaknesses of the study?

Ref: T. Lee, J. Cho, D. Kwon and S. Sohn, "Global stock market investment strategies based on financial network indicators using machine learning techniques", *Expert Systems with Applications*, vol. 117, pp. 228-242, 2019. Available: 10.1016/j.eswa.2018.09.005 [Accessed 5 February 2021].

Conclusion to summarise the main agreements and disagreements, any gaps or areas for further work, and your overall viewpoint on the topic