Introduction toexplain the purpose of the review and the motivation of the research topic

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