International Journal of Economics and Financial Issues

ISSN: 2146-4138

available at http: www.econjournals.com

International Journal of Economics and Financial Issues, 2019, 9(4), 1-10.

An Analysis on Investment Performance of Machine Learning: An Empirical Examination on Taiwan Stock Market

Chia-Cheng Chen[1], Yi-Sheng Liu[2]*, Ting-Hsin Hsu[3]

[1]Department of Finance, Ling Tung University of Science and Technology, Taiwan, [2]Department of Finance, National Yunlin University of Science and Technology, Taiwan, [3]Department of Finance, National Taichung University of Science and Technology, Taiwan. *Email: goever.tw@gmail.com

Received: 13 April 2019 **Accepted:** 10 June 2019 **DOI:** https://doi.org/10.32479/ijefi.8129

ABSTRACT

This study aims to explore the prediction of Taiwan stock price movement and conduct an analysis of its investment performance. Based on Taiwan stock market index, the study compares four machine learning models: artificial neural networks (ANN), support vector machines (SVM), random forest and Naïve-Bayes. With a performance evaluation of Taiwan stock market index historical data spanning from 2014 to 2018, we find: (1) By overall performance measures, machine learning models outperform benchmark market index. (2) By risk-adjusted measures, the empirical results suggest that ANN generates the best performance, followed by SVM and random forest, and Naïve-Bayes coming in last.

Keywords: Naive-Bayes, Artificial Neural Networks, Support Vector Machines, Random Forest, Machine Learning, Investment Performance
JEL Classifications: C11, C53, C63, G11

1. INTRODUCTION

In recent years, with the advance of hardware devices, the computing capacity of computers has also made a great leap forward, rendering machine learning and artificial intelligence (AI) mainstream. Despite the fact that there have been a number of empirical researches conducted on machines learning to predict share prices movement (Bisoi et al., 2019; Henrique et al., 2019; Hsu et al., 2009; Huang and Liu, 2019; Kara et al., 2011; Karaatli et al., 2005; Khan et al., 2016; Lee et al., 2019; Leung et al., 2000; Long et al., 2019; Hiransha et al., 2018; Patel et al., 2015; Zhang et al., 2019), the attention has been paid majorly to the prediction efficacy of machine learning rather and little on the aspects of performance and risk measurement. Therefore, this study aims to fill the gap and delve into the financial evaluation of machine learning application in Taiwan stock market.

Stock market price movement prediction has to confront the strongest rejection from the academic paradigm of efficient market hypothesis states that prices of stocks are informationally efficient which means that it is impossible to predict stock prices based on the trading data (Malkiel and Fama, 1970). However, more recent results show that, if the information obtained from stock prices is pre-processed efficiently and appropriate algorithms are applied then trend of stock or stock price index may be predictable (Patel et al., 2015). The new discovery can greatly benefit market practitioners because accurate predictions of movement of stock price indexes are very important for developing effective market trading strategies (Leung et al., 2000).

The core objective of the research is to input the results of ten technical analysis indicators into artificial neural networks (ANN), support vector machines (SVM), Random Forest and Naive-Bayes models to predict stock price movement and evaluate investment performance and risk measurement. In the circumstance, the machine learning models buy stocks when predicting a rise and short stocks when predicting a decline in prices. Based on Taiwan Stock Exchange (TWSE) Index from 2014 to 2018, this research compares the investment performance among the machine learning models.

The remainder of this paper are organized as follows. Section 2 provides a brief overview of the theoretical literature. Section 3 describes the research data. Section 4 provides the prediction models and risk-adjusted measures used in this study. Section 5 reports the empirical results from the comparative analysis. Finally, Section 6 contains the concluding remarks.

2. LITERATURE REVIEW

For a long time, it was believed that changes in the prices of stocks is not forecastable. Predicting returns in the stock market is usually posed as a forecasting problem where prices are predicted. Intrinsic volatility in the stock market across the globe makes the task of prediction challenging. Stock prediction and selection has long been identified as an important but challenging topic in the research area of financial market analysis (Becker et al., 2007; Huang et al., 2011). In this section, we focus the review of previous studies on ANN, SVM, Random Forest and Naive-Bayes applied to stock market prediction and investment performance.

To explore the future features of stock markets, various forecasting algorithms have been employed, of which, computational intelligence (CI) (or AI) has become increasingly dominant due to its powerful learning capability and high prediction accuracy. Typical CI techniques in stock market prediction (for stock prices, stock returns, market indexes, etc.) are ANNs (Du, 2018; Kim and Shin, 2007; Qiu et al., 2016; Xi et al., 2014) and SVMs (Kazem et al., 2013; Li et al., 2014; Li et al., 2015; Yang et al., 2019).

ANN and SVM have been demonstrated to provide promising results in predict the stock price return (Avcı, 2007; Huang and Liu, 2019; Kara et al., 2011; Karaatli et al., 2005; Olson and Mossman, 2003; Patel et al., 2015; Yang et al., 2019). Cao et al., (2005) uses ANN to predict stock price movement (i.e., price returns) for firms traded on the Shanghai stock exchange. They compare the predictive power using linear models from financial forecasting literature to the predictive power of the univariate and multivariate neural network models. Their results show that neural networks outperform the linear models compared. Hassan et al., (2007) propose and implement a fusion model by combining the Hidden Markov Model (HMM), ANN and genetic algorithms (GA) to forecast financial market behavior. Using ANN, the daily stock prices are transformed to independent sets of values that become input to HMM. Forecasts are obtained for a number of securities in the IT sector and are compared with a conventional forecast method.

Wang et al. (2016) is developed and combined a hybrid v- support vector regression (SVR) model with principal component analysis and brain storm optimization for stock price index forecasting. Numerical results indicate that the developed hybrid model is not only simple but also able to satisfactorily approximate the actual CSI300stock price index, and it can be an effective tool in stock market mining and analysis. Fenghua et al. (2014), using the singular spectrum analysis (SSA), decomposes the stock price into terms of the trend, the market fluctuation, and the noise with different economic features over different time horizons, and then introduce these features into the SVM to make price predictions. The empirical evidence shows that, compared with the SVM without these price features, the combination predictive methods- the EEMD-SVM and the SSA-SVM, which combine the price features into the SVMs perform better, with the best prediction to the SSA-SVM.

Yang et al. (2019) predict stock market price with a forecasting model based on chaotic mapping, firefly algorithm, and SVR. Compared with GA-based SVR (SVR-GA), chaotic GA-based SVR (SVR-CGA), firefly-based SVR (SVR-FA), ANNs and adaptive neuro-fuzzy inference systems, the proposed model performs best based on two error measures, namely mean squared error and mean absolute percent error. Hsu et al. (2009) employs a two-stage architecture for better stock price prediction. Specifically, the self-organizing map is first used to decompose the whole input space into regions where data points with similar statistical distributions are grouped together, so as to contain and capture the non-stationary property of financial series. After decomposing heterogeneous data points into several homogenous regions, SVR is applied to forecast financial indices. The proposed technique is empirically tested using stock price series from seven major financial markets. The results show that the performance of stock price prediction can be significantly enhanced by using the two-stage architecture in comparison with a single SVR model.

Ciner (2019) show that when the random forest method, which accounts for both linear and nonlinear dynamics, is used for regression, industry returns indeed contain significant out of sample forecasting power for the market index return. Basak et al., (2018) develop an experimental framework for the classification problem which predicts whether stock prices will increase or decrease with respect to the price prevailing n days earlier. Two algorithms, random forests, and gradient boosted decision trees facilitate this connection by using ensembles of decision trees. Gupta et al., (2018) use quantile random forests to study the predictive value of various consumption-based and income-based inequality measures across the quantiles of the conditional distribution of stock returns. Results suggest that the inequality measures have predictive value for stock returns in sample, but do not systematically predict stock returns out of sample.

Khan et al. (2016) employ several algorithms in stock prediction such as SVM, ANN, linear discriminant analysis, linear regression, K-NN, and Naïve Bayesian classifier to approach the subject of predictability with greater accuracy. Chatzis et al., (2018) leverage the merits of a series of techniques including classification trees, SVM, random forests, neural networks, extreme gradient boosting, and deep neural networks and find significant evidence of interdependence and cross-contagion effects among stock, bond and currency markets.

3. RESEARCH DATA

The data used in this paper all comes from the Taiwan database of the Taiwan Economic Journal (TEJ). We collect 1229 TWSE Index samples from the TEJ over January 2014-December 2018 period. In our study, we adopt a 5-year horizon of historical data, as suggested by Jobson and Korkie (1981), in order to strike an appropriate balance between rapidly-changing market conditions and statistical confidence. These data form our entire data set. Percentage wise increase and decrease cases of each year in the entire data set are shown in Table 1.

There are some technical indicators through which one can predict the future movement of stocks. Here in this study, total ten technical indicators as employed in Huang and Liu (2019) are used. These indicators are shown in Table 2.

Table 1: The number of increase and decrease cases percentage in each year in the entire data set of TWSE

Year	Increase	%	Decrease	%	Total
2014	136	55	112	45	248
2015	119	49	125	51	244
2016	139	57	105	43	244
2017	140	57	106	43	246
2018	127	51	120	49	247
Total	661	54	568	46	1229

TWSE: Taiwan stock exchange

Table 2: Selected technical indicators and their formulas

Name of indicators	Formulas
SMA (10-day)	$\frac{C_t + C_{t-1} + \cdots + C_{t-9}}{10}$
WMA (10-day)	$\frac{((n)\times C_t + (n-1)\times C_{t-1} + \cdots + C_{t-9})}{(n+(n-1)+\cdots+1)}$
Momentum	$C_t - C_{t-n}$
Stochastic K%	$\frac{C_t - LL_{t-n}}{HH_{t-n} - LL_{t-n}} \times 100$
Stochastic D%	$\frac{\sum_{i=0}^{n-1} K_{t-i}\%}{n}$
Relative Strength Index	$100 - \frac{100}{1 + (\sum_{i=0}^{n-1} Up_{t-i}/n)/(\sum_{i=0}^{n-1} Dw_{t-i}/n)}$
MACD	$MACD(n)_{t-1} + \frac{2}{n+1} \times (DIFF_t - MACD(n)_{t-1})$
Larry William's R%	$\frac{H_n - C_t}{H_n - L_n} \times 100$
A/D Oscillator	$\frac{H_t - C_{t-1}}{H_t - L_t}$
Commodity Channel Index	$\frac{M_t - SM_t}{0.015 D_t}$

C_t is the closing price, L_t is the low price, H_t is the high price at time t, $DIFF_t = EMA(12)_t - EMA(26)_t$, EMA is exponential moving average, $EMA(k)_t = EMA(k)_{t-1} + \alpha \times (C_t - EMA(k)_{t-1})$, α is a smoothing factor, $\alpha = \frac{2}{k+1}$, k is time period of k day exponential moving average, LL_t and HH_t mean lowest low and highest high in the last t days, respectively. $M_t = \frac{H_t + L_t + C_t}{3}$, $SM_t = \frac{\left(\sum_{i=1}^{n} M_{t-i+1}\right)}{n}$, $D_t = \frac{\left(\sum_{i=1}^{n} |M_{t-i+1} - SM_t|\right)}{n}$, UP_t means upward price change while DW_t is the downward price change at time t.

In the research, we input the results of ten technical analysis indicators into ANN, SVM, random forest and Naive-Bayes models to predict stock price movement. In the circumstance which the transaction costs are calculated, the machine learning models buy stocks when predicting a rise and short stocks when predicting a decline in prices. Based on TWSE Index, the research compares the investment efficiency between the machine learning models.

4. PREDICTION MODELS AND RISK-ADJUSTED MEASURES

4.1. Prediction Models

4.1.1. ANN

The ANN are non-linear models that make use of a structure capable to represent arbitrary complex non-linear processes that relate the inputs and outputs of any system (Chatzis et al., 2018; Chen et al., 2003; Hassan et al., 2007; Henrique et al., 2019; Huang and Liu, 2019; Kara et al., 2011; Khan et al., 2016; Leung et al., 2000; Olson and Mossman, 2003; Patel et al., 2015). ANN represents one widely used soft computing technique for stock market forecasting. ANN has demonstrated capability in financial modeling and prediction (Huang and Liu, 2019; Kara et al., 2011; Leung et al., 2000; Olson and Mossman, 2003; Patel et al., 2015). In this study, a three-layered feedforward ANN model was structured to predict stock price index movement. This ANN model consists of an input layer, a hidden layer and an output layer, each of which is connected to the other. The ANN architecture is defined by the way in which the neurons are interconnected. The network is fed with a set of input-output pairs and is trained to reproduce the output. The number of neurons (hn) in the hidden layer, value of learning rate (lr), momentum constant (mc) and number of iterations (ep) are ANN model parameters that must be efficiently determined. Inputs for the network were ten technical indicators which were represented by ten neurons in the input layer. The architecture of the three-layered feedforward ANN is illustrated in Figure 1. The optimal parameters and investment performance of ANN prediction model is summarized in Table 3.

4.1.2. SVM

In machine learning, SVM are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. SVM emerged from research in statistical learning theory on how to regulate generalization and find an optimal tradeoff between structural complexity and empirical risk. SVMs classify points by assigning them to one of two disjoint half spaces, either in the pattern space or in a higher-dimensional feature space. One of the most popular SVM classifiers is the "maximum margin" one, which aims to minimize an upper bound on the generalization error through maximizing the margin between two disjoint half planes (Burges, 1998; Cortes and Vapnik, 1995; Patel et al., 2015). A SVM is a discriminative classifier formally defined by a separating hyperplane. In other words, given labeled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes new examples. In two dimensional space this hyperplane is a line dividing a plane in two parts where in each class lay in either side (Bhatia and Madaan, 2018). The main idea of support vector

Figure 1: Architecture of artificial neural networks

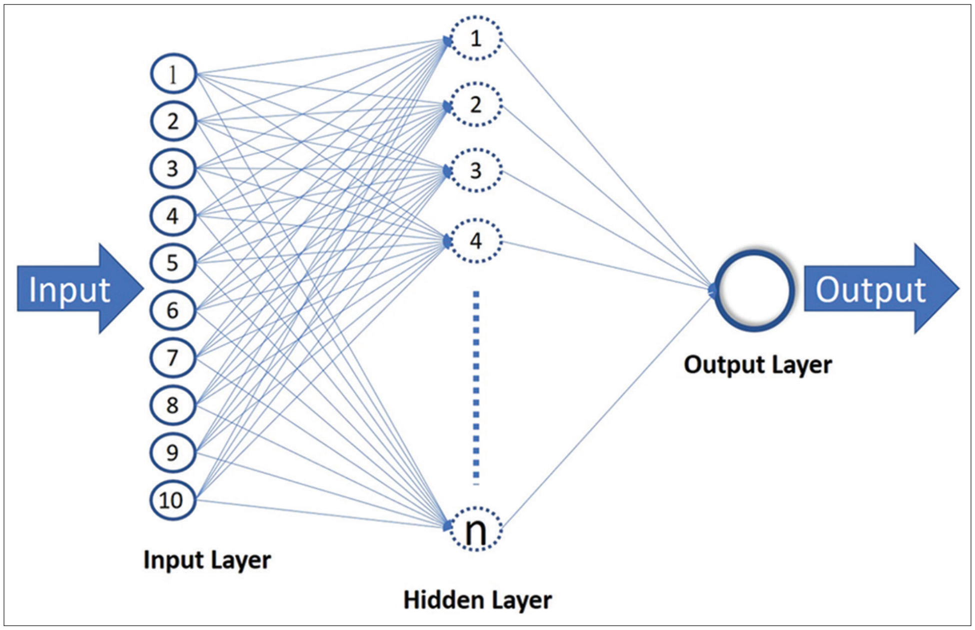


Table 3: Optimal parameters and annual returns of four prediction model with benchmark

Parameters	Prediction models	2014	2015	2016	2017	2018	5 years
None	TWSE (Benchmark)	8.7566	−12.108	14.1587	14.767	−10.218	15.3566
n=4000;ep=10; mc=0.5	ANN	8.3469	39.9949	17.5216	4.3377	40.9534	111.1545
k=radial basis; c=5;g=0.5	SVM	7.235	54.9908	32.289	4.5193	8.5091	107.5432
ct=entropy; md=4 nt=20	Random Forest	22.3404	14.4165	13.5843	18.7226	29.2117	98.2755
None	Naïve-Bayes	2.7612	11.4268	2.7631	−4.4958	6.6701	19.1254

ANN is artificial neural network; SVM is the support vector machines; TREE is the Random Forest; GNB is the Naïve-Bayes; TWSE is the Taiwan stock exchange index, it's also benchmark in the study

machine is to construct a hyperplane as the decision surface such that the margin of separation between positive and negative examples is maximized (Xu et al., 2009). The equation of the hyperplane can be given as:

$$\omega^T + b = 0 \tag{1}$$

The margin is width is 2/||ω|| and the learning problem is equivalent to unconstrained optimization problem over ω.

$$min\omega^2 + C\sum_{i}^{N}\max(0,1-y_i f(x_i)) \tag{2}$$

SVM are highly effective in high dimensional spaces but under perform when target classes (for classification problems) are overlapping i.e. kernel functions need to be used.

Choice of kernel function, degree of kernel function (d) in case of polynomial kernel, gamma in kernel function (g) in case of radial basis kernel and regularization constant c are the parameters of SVM. The optimal parameters and investment performance of SVM prediction model is summarized in Table 3.

4.1.3. Random forest

Random Forest is an ensemble, data-miner which uses "deep" (unpruned) decision trees as base learners. It is a modification of applying bagging to multiple classification and regression trees, and averaging the predictions of the approximately uncorrelated trees to yield the final estimate. Random Forest model was unable to show any clear patterns in the data through variable importance plots and did not show any significant improvement in performance in comparison to generalized linear models (Bhatia and Madaan, 2018). Decision tree learning is one of the most popular techniques for classification. Its classification accuracy is comparable with other classification methods, and it is very efficient.

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classification. It uses decision tree as the base learner of the ensemble. The idea of ensemble learning is that a single classifier is not sufficient for determining class of test data. Reason being, based on sample data, classifier is not able to distinguish between noise and pattern. So it performs sampling with replacement such that given n trees to be learnt are based on these data set samples. After creation of n trees, when testing data is used, the decision which majority of trees come up with is considered as the final output. This also avoids problem of over-fitting.

Choice of criterion function to measure the quality of a split. Supported criteria are "gini" for the Gini impurity and "entropy" for the information gain. Number of trees in the ensemble "ntrees" and the maximum depth of the tree are considered as the parameter of random forest. The optimal parameters and investment performance of random forests prediction model is summarized in Table 3.

4.1.4. Naïve-Bayes

The Naive Bayes classifier technique is based on the so-called Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. In machine learning, Naive Bayes is a simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. There is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: All naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable. For some types of probability models, naive Bayes classifiers can be trained very efficiently in a supervised learning setting.

In many practical applications, parameter estimation for naive Bayes models uses the method of maximum likelihood; in other words, one can work with the naive Bayes model without accepting Bayesian probability or using any Bayesian methods. Despite its simplicity, Naive Bayes can often outperform more sophisticated classification methods. For example, under specific assumptions, it can be demonstrated that many neural networks and curve-fitting algorithms output the maximum posteriori hypothesis, as does the naive Bayesian classifier. The investment performance of Naive-Bayes prediction model is summarized in Table 3.

4.2. Risk-Adjusted Measures

4.2.1. Sharpe ratio

The Sharpe ratio is used to help investors understand the return of an investment compared to its risk. The ratio is the average return earned in excess of the risk-free rate per unit of volatility or total risk. Subtracting the risk-free rate from the mean return allows an investor to better isolate the profits associated with risk-taking activities. Generally, the greater the value of the Sharpe ratio, the more attractive the risk-adjusted return. In particular, a negative Sharpe ratio indicates a situation of "anti-skill," since the performance of the riskless asset is clearly superior. The Sharpe ratio is defined in Equation (3).

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p} \tag{3}$$

where:
R_p = return of portfolio
R_f = risk-free rate
σ_p = standard deviation of the portfolio's excess return.

4.2.2. Treynor ratio

The Treynor ratio, also known as the reward-to-volatility ratio, is a performance metric for determining how much excess return was generated for each unit of risk taken on by a stock. Excess return in this sense refers to the return earned above the return that could have been earned in a risk-free investment. Risk in the Treynor ratio refers to systematic risk as measured by a stock's beta. Beta measures the tendency of a stock's return to change in response to changes in return for the overall market. The higher the Treynor ratio, the better the performance of the stock under analysis. The Treynor ratio is estimated using Equation (4).

$$\text{Treynor Ratio} = \frac{R_p - R_f}{\beta_p} \tag{4}$$

where:
R_p = return of portfolio
R_f = risk-free rate
β_p = beta of the portfolio.

4.2.3. Information ratio

The information ratio measures the risk-adjusted returns of a financial asset or stock relative to a certain benchmark. This ratio aims to show excess returns relative to the benchmark, as well as the consistency in generating the excess returns. The consistency of generating excess returns is measured by the tracking error.

The information ratio and the Sharpe ratio are similar in a way. Both ratios determine the risk-adjusted returns of a security or stock. However, the information ratio measures the risk-adjusted returns relative to a certain benchmark while the Sharpe ratio compares the risk-adjusted returns to the risk-free rate. The Information ratio is estimated using Equation (5).

$$\text{Information Ratio} = \frac{R_p - R_b}{\delta_{pb}} \tag{5}$$

where:
R_p = return of portfolio
R_b = return of benchmark
δ_{pb} = standard deviation of difference between portfolio and benchmark returns.

4.2.4. Jensen's alpha

The Jensen's Alpha is an absolute measure of performance. It was developed by American economist Michael Jensen in 1968 (Jensen, 1968). It is given by the annualized return of the stock, deducted the yield of an investment without risk, minus the return of the benchmark multiplied by the stock's beta during the same period. The Jensen's Alpha gives the excess return obtained when deviating from the benchmark (Jensen, 1972).

The magnitude of the Jensen's Alpha depends on two key variables: the return of the benchmark and the beta. This indicator represents the part of the mean return of the stock that cannot be explained by the systematic risk exposure to market variations.

As it is an absolute measure, it does not reflect completely the risk of the stock. It is then generally easier for a more risky stock to exhibit a greater Jensen's Alpha than for a less risky stock. It should be then applied on homogenous class of assets. Moreover, the validity of this measure depends crucially on the hypothesis that the beta of the stock is stationary. The validity of this hypothesis has to be tested before focusing on the value of this indicator (Grinblatt and Titman, 1987, 1989, 1992).

4.2.5. Modigliani ratio

The Modigliani risk ratio, often called M2, measures the return provided by an investment in the context of the risk involved. It was developed by Franco Modigliani and Leah Modigliani in the year 1997.

Modigliani and Modigliani (1997) believed that an ordinary investor would find it easier to understand the Modigliani measure compared to Sharpe ratio. The reason behind this was that their measure is expressed in percentage points. It shows how well the investor is rewarded for taking a certain amount of risk, relative to the benchmark and the risk free rate.

In general, the riskier an investment is, the less inclined investors will be to put their money into it. So riskier investments have to offer a higher potential return that is, deliver a greater profit if the investment succeeds. In simple words, it measures the returns of an investment index or stock for the amount of risk taken relative to some benchmark index.

5. EXPERIMENTAL RESULTS AND ANALYSIS

The research empirically examines the financial performance of machine learning through performance measures, such as Jensen's Alpha, Sharpe ratio, Treynor ratio, information ratio and Modigliani ratio. Our experimental results are based on data retrieved from Taiwan Stock Market Index (from January, 2014 to December, 2018). The empirical results are presented firstly by descriptive statistics, followed by an annual evaluation analysis, and concluded with overall performance comparison among machine learning models.

5.1. Descriptive Statistics

The analysis of the overall performance of ANN, SVM, Random Forest and Naive-Bayes models is undertaken base on the benchmarks indices, the TWSE index. The descriptive statistics of the daily returns for the benchmark indices, and for the four machine learning models are reported in Table 4. Figure 2 is the daily return chart of machine learning models and TWSE from 2014 to 2018.

There is a 7% daily price limit in Taiwan stock market and thus, both the descriptive statistics in Table 4 and the daily return chart in Figure 2 show data of <7% of the movement in stock prices. As presented in Table 4, our proposed machine learning models generate significantly higher mean returns. In terms of average daily return, ANN and SVM generate 7 times higher returns than TWSE (benchmark); with random forest of 6.4 and Naïve-Bayes of 1.25 times. As for the figures of maximum, minimum, and standard deviation of returns, machine learning models exhibit higher max, lower min, and less variation than benchmark index. Among the machine learning models, the results indicate that the order of investment performance excellence can be put down as follows: ANN, SVM, Random Forest and Naïve-Bayes.

The 5-year excess returns of machine learning models during the sample period are reported in Table 5 and their graphic depiction is referred to Figure 3. For entire sample period, the excess returns of ANN, SVM, Random Forest, and Naïve-Bayes are respectively of 95.80%, 92.19%, 82.92%, and 3.77%. By annual data, our machine learning models are particularly impressive during the bear markets. Specifically, in 2015 with market return of −12.11%, ANN, SVM, Random Forest, and Naïve-Bayes are respectively of 52%, 67%, 27%, and 24%. Likewise, in 2018 with market return of −10.22%, ANN, SVM, Random Forest, and Naïve-Bayes are respectively of 51%, 19%, 39%, and 17%. Overall, our empirical results suggest machine learning is promising of higher cumulative returns in TWSE application.

Table 4: Descriptive statistics of daily returns

Indicator	ANN	SVM	TREE	GNB	TWSE (benchmark)
Max.	6.3125	6.3125	6.3125	4.9538	3.5801
Min.	−4.9538	−4.9538	−3.5801	−6.3125	−6.3125
Standard deviation	0.8205	0.8209	0.8216	0.8254	0.8267
Mean	0.0905	0.0875	0.0800	0.0156	0.0125
Median	0.0858	0.0857	0.0796	0.0302	0.0581

ANN is artificial neural network; SVM is the support vector machines; TREE is the Random Forest; GNB is the Naïve-Bayes; TWSE is the Taiwan Stock Exchange Index, it's also benchmark in the study

Figure 2: Daily return of machine learning with Taiwan Stock Exchange benchmark investment performance

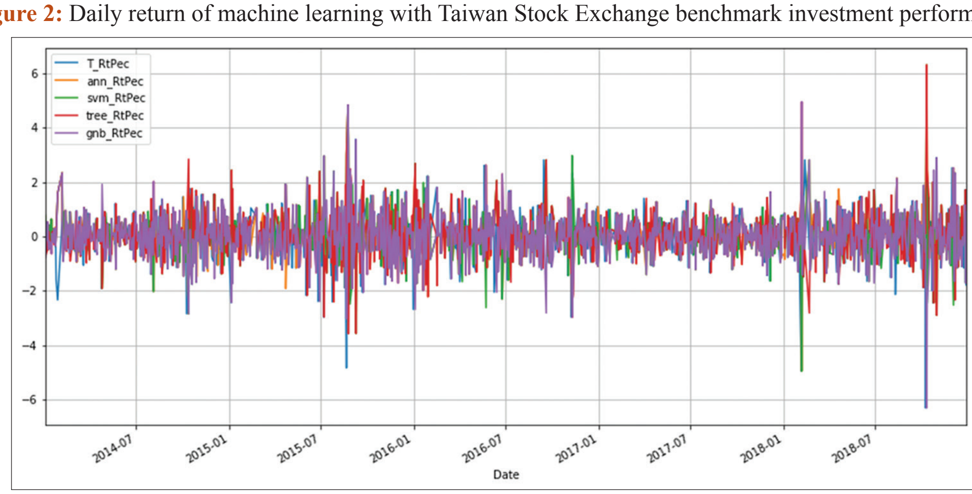
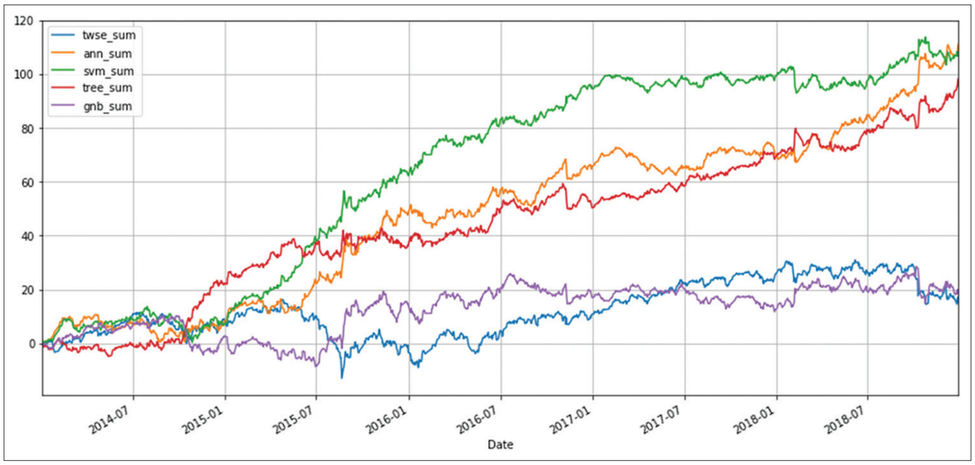


Table 5: Excess return with the TWSE benchmark

Model	2014	2015	2016	2017	2018	5 years
TWSE (Benchmark)	8.7566	−12.108	14.1587	14.767	−10.218	15.3566
ANN	8.3469	39.9949	17.5216	4.3377	40.9534	111.1545
Excess return	−0.4097	52.1029	3.3629	−10.4293	51.1711	95.7979
SVM	7.235	54.9908	32.289	4.5193	8.5091	107.5432
Excess return	−1.5216	67.0988	18.1303	−10.2477	18.7268	92.1866
TREE	22.3404	14.4165	13.5843	18.7226	29.2117	98.2755
Excess return	13.5838	26.5245	−0.5744	3.9556	39.4294	82.9189
GNB	2.7612	11.4268	2.7631	−4.4958	6.6701	19.1254
Excess return	−5.9954	23.5348	−11.3956	−19.2628	16.8878	3.7688

ANN is artificial neural network; SVM is the support vector machines; TREE is the Random Forest; GNB is the Naïve-Bayes; TWSE is the Taiwan Stock Exchange Index, it's also benchmark in the study

Figure 3: Cumulative return of machine learning with Taiwan Stock Exchange benchmark investment performance



5.2. Investment Performance Evaluation by Years

The Sharpe ratio uses the capital market line as a benchmark to measure the depth and breadth of performance, with a higher Sharpe ratio being better than a lower Sharpe ratio. In particular, a negative Sharpe ratio indicates a situation of "anti-skill," since the performance of the riskless asset is clearly superior.

According to Table 6, it is evident from the empirical results presented in proposed machine learning models outperforms the benchmark. More specifically, the Sharpe ratios of TWSE in 2015 and 2018 produce negative values, indicating no investment worth. The Sharpe ratios of ANN, SVM and random forest are positive for all 5 years and outperform that of TWSE. In short, the Sharpe ratio of ANN, SVM, Random Forest and Naïve-Bayes all exceed that of TWSE benchmark index.

Although originally referred to by Treynor and Black (1973) as the "appraisal ratio," the information ratio is the ratio of relative returns to relative risk, and whilst the sharpe ratio examines the returns relative to a riskless asset, the information ratio is based upon returns relative to a risky benchmark.

A comparison of machine learning models' performance during these 5 years manifests that ANN, SVM and random forest have higher information ratios whereas Naïve-Bayes have the lowest. In Tables 5 and 7, the excess returns of ANN and SVM in 2014 and 2017 produce negative values; therefore, their information ratios are accordingly lower, with the values respectively of −0.0547 and −0.0575. Naïve-Bayes exhibits the worst investment performance in 2017 with its excess return reaching −19.2628% and its information ratio coming to −0.1098, the lowest among the four models.

Tracking errors are calculated as the relative standard deviation of returns between a stock and a benchmark. A tracking error is a useful performance measure relative to a benchmark since it is measured in units of asset returns. The comparative empirical tracking errors of the machine learning models with respect to the benchmark indices are reported in Table 8.

Concretely speaking, we found that the graver the machine learning models' tracking errors are, the better the investment performances turn out. For instance, the excess returns of ANN in 2015 and 2018 are 52.1029% and 51.1711%. Their tracking errors are 1.4248 and 1.3876. The excess returns in 2014 and 2017 are −0.4097% and −10.4293% and their tracking errors are 0.9896 and 0.7739. The results justifies the value of active management by machine learning compared to passive index tracking.

5.3. Overall Performance Comparison

The Jensen's Alpha provides quite a robust measure of the abnormal returns that are generated by the stock as compared to a passive combination of the risk-free asset and a market index with exactly the same risk characteristics as the stock.

Table 10 shows that the Jensen's Alphas of the machine learning models are positive, transpiring that their investment performances are better than that of benchmark index. Among the machine learning models, the Alphas of ANN, SVM and random forest are far higher than that of Naïve-Bayes, generating respectively the values of 0.0909, 0.0869, 0.0798 and 0.0153. Therefore, the sequence of investment performance excellence can be put down as follows: ANN, SVM, random forest and Naïve-Bayes.

Table 6: Sharpe index

Year	TWSE	ANN	SVM	TREE	GNB
2014	0.0458	0.0433	0.0368	0.1264	0.0105
2015	−0.0529	0.1627	0.2281	0.0554	0.0431
2016	0.0670	0.0823	0.1566	0.0628	0.0094
2017	0.1073	0.0274	0.0288	0.1385	−0.0398
2018	−0.0450	0.1688	0.0321	0.1186	0.0245
5 years	0.0109	0.1061	0.1025	0.0932	0.0147

ANN: Artificial neural networks, SVM: Support vector machines, TWSE: Taiwan stock exchange

Table 7: Information ratio

Year	ANN	SVM	TREE	GNB
2014	−0.0016	−0.0061	0.0579	−0.0240
2015	0.1498	0.2004	0.0774	0.0752
2016	0.0116	0.0680	−0.0019	−0.0390
2017	−0.0547	−0.0575	0.0205	−0.1098
2018	0.1492	0.0572	0.1204	0.0478
5 years	0.0659	0.0662	0.0585	0.0026

ANN: Artificial neural networks, SVM: Support vector machines

Table 8: Tracking error

Year	ANN	SVM	TREE	GNB
2014	0.9896	1.0013	0.9486	1.0096
2015	1.4248	1.3718	1.4027	1.2822
2016	1.1829	1.0923	1.1804	1.1954
2017	0.7739	0.7239	0.7813	0.7129
2018	1.3876	1.3241	1.3255	1.4295
5 years	1.1822	1.1324	1.1524	1.1546

ANN: Artificial neural networks, SVM: Support vector machines

Table 9: Treynor ratio

Year	ANN	SVM	TREE	GNB
2014	−1.3638	−0.6566	2.8000	−0.1448
2015	−7.3928	5.5199	2.6292	0.2352
2016	−2.3004	1.4344	−2.1031	−0.1748
2017	−0.4733	0.3860	−1.5553	−0.3879
2018	−11.4458	0.3611	1.4658	−0.3947
5 years	−3.4136	1.8832	4.3296	0.6589

ANN: Artificial neural networks, SVM: Support vector machines

The beta is a measurement of its volatility of returns relative to the entire market. It is used as a measure of risk and is an integral part of the capital asset pricing model. An index with a higher beta has greater risk and also greater expected returns.

In Table 10, we can see that ANN generates the greatest excess return, which reaches 95.7979%, and yet, its beta is −0.0255 and the resulting Treynor ratio is −3.4136. The reason behind its beta's negative value is that machine learning models' primary function is to predict stock price movement and buy stocks when prices rise and short stocks when prices fall. It aims to benefit both from rising and falling, and thus, its nature resembles active management funds rather than tracking index. . It's Treynor ratio also exhibits negative value because risk in the Treynor ratio refers to systematic risk as measured by stock's beta. Such instances can be discerned in Table 9, where its data shows that ANN's Treynor ratio produces negative values all five years; excess return of Random Forest reaches 82.9189% but its Treynor ratio generates negative values in 2016 and 2017. In particular, when markets are most down and corrected predicted by ANN, the returns of ANN strategy will inverse with market returns and generate negative covariance between ANN and market index. In addition to ANN whose beta turns out to be negative, the betas of SVM, Random Forest and Naive-Bayes fall between zero and one, manifesting that although they are less volatile than the market, their investment performances proves to be more outstanding than that of benchmark index.

Modigliani and Modigliani (1997) propose M2 performance measure by using return per unit of total risk as measured with the standard deviation. The Modigliani ratio measures the returns of the stock, adjusted for the risk of the stock relative to that of some benchmark. To calculate the M2 ratio, we first calculate the Sharpe ratio and then multiply it by the annualized standard deviation of a chosen benchmark. We then add the risk-free rate to the derived value to give M2 ratio.

In terms of Sharpe ratio, Treynor ratio, Information ratio and Modigliani ratio, the four prediction models all excels benchmark index. Among them, ANN, SVM and random forest outperform Naive-Bayes.

In brief, conclusions can be elicited from results above that, firstly, the four machine learning models all exhibit better investment performance than TWSE in terms of higher excess returns or less beta volatility. Among them, the excess returns of ANN, SVM and Random Forest are 6-7 times greater than that of TWSE. For risk-adjusted performance measures, such as Shape ratio, Jensen's

Table 10: Index of risk-adjusted return based on volatility

Indicator	ANN	SVM	TREE	GNB	TWSE
5Y return	111.15	107.54	98.28	19.13	15.36
Excess return	95.7979	92.1866	82.9189	3.7688	
Sharpe ratio	0.1061	0.1025	0.0932	0.0147	0.0109
Jensen's alpha	0.0909	0.0869	0.0798	0.0153	
Beta	−0.0255	0.0446	0.0176	0.0184	
Treynor ratio	−3.4136	1.8832	4.3296	0.6589	
Information ratio	0.0659	0.0662	0.0585	0.0026	
Modigliani ratio (M2)	0.0898	0.0869	0.0795	0.0155	0.0125

ANN: Artificial neural networks, SVM: Support vector machines, TWSE: Taiwan stock exchange alpha, information ratio, and Modigliani ratio, machine learning models surpass benchmark index. Moreover, among the four machine learning model, ANN generates the best performance, followed by SVM and Random Forest, and Naïve-Bayes coming in last.

6. CONCLUSIONS AND REMARKS

The task focused in this paper is an analysis on investment performance of machine learning models. Investment performance of four models namely ANN, SVM, random forest and Naive-Bayes is compared based on 5 years (2014–2018) of historical data of Taiwan Stock Market (TWSE) Index.

The daily price limit of Taiwan stock market falls within 7% of the price range, and therefore, the data of daily return does not transcend the fluctuation limit of 7%. In terms of average daily return, ANN and SVM generate 7 times more return than TWSE (benchmark); Random Forest produces 6.4 more while Naive-Bayes exhibits 1.25 more return, validating that the machine learning models all outperform TWSE by average daily return.

Experiments with investment performance show that ANN, SVM and Random Forest exhibit higher performances with 95.79%, 92.19% and 82.92% cumulative excess returns. ANN and SVM generate the greatest cumulative return of 111% and 107%; random forest produces 98% cumulative return. It transpires that the cumulative return and excess return of ANN, SVM and Random Forest all exceed that of TWSE benchmark index. Nevertheless, Naïve-Bayes does not vary prominently from it.

The Jensen's Alphas of all four machine learning models produce positive values, indicating that their investment performances surpass that of benchmark index. Among them, the Alphas of ANN, SVM, and random forest outperform that of naïve-Bayes by a noticeable margin, producing the values respectively of 0.0909, 0.0869, 0.0798 and 0.0153. Therefore, the sequence of investment performance excellence can be put down as follows: ANN, SVM, random forest and Naïve-Bayes.

The foremost object of machine learning models is to predict stock price movement and buy stocks when prices rise and short stocks when prices fall, hoping to obtain profits both from rising and falling. Therefore, it bears a resemblance to active management funds rather than tracking index, which accounts for their beta's negative values and lucrative performance. With regard to Sharpe ratio, Treynor ratio, Information ratio and Modigliani ratio, the four prediction models all excels benchmark index.

To sum up, machine learning models exceed benchmark index in investment performance. Among the machine learning models, ANN and SVM excel the others, with Random Forest ranking third and Naïve-Bayes coming in last.

REFERENCES

Avcı, E. (2007), Forecasting daily and sessional returns of the ISE-100 index with neural network models. Journal of Dogus University, 8(2), 128-142.

Basak, S., Kar, S., Saha, S., Khaidem, L., Dey, S.R. (2018), Predicting the direction of stock market prices using tree-based classifiers. The North American Journal of Economics and Finance, 47, 552-567.

Becker, Y.L., Fei, P., Lester, A.M. (2007), Stock Selection: An Innovative Application of Genetic Programming Methodology. In: Genetic Programming Theory and Practice IV. Berlin: Springer. p315-334.

Bhatia, M., Madaan, A. (2018), Stock Portfolio Performance by Weighted Stock Selection. Abstract: The Aim of this Project.

Bisoi, R., Dash, P.K., Parida, A.K. (2019), Hybrid variational mode decomposition and evolutionary robust kernel extreme learning machine for stock price and movement prediction on daily basis. Applied Soft Computing, 74, 652-678.

Burges, C.J. (1998), A tutorial on support vector machines for pattern recognition. Data Mining Knowledge Discovery, 2(2), 121-167.

Cao, Q., Leggio, K.B., Schniederjans, M.J. (2005), A comparison between fama and French's model and artificial neural networks in predicting the Chinese stock market. Computers and Operations Research, 32(10), 2499-2512.

Chatzis, S.P., Siakoulis, V., Petropoulos, A., Stavroulakis, E., Vlachogiannakis, N. (2018), Forecasting stock market crisis events using deep and statistical machine learning techniques. Expert Systems with Applications, 112, 353-371.

Chen, A.S., Leung, M.T., Daouk, H. (2003), Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. Computers and Operations Research, 30(6), 901-923.

Ciner, C. (2019), Do industry returns predict the stock market? A reprise using the random forest. The Quarterly Review of Economics and Finance, 72, 152-158.

Cortes, C., Vapnik, V. (1995), Support-vector networks. Machine learning, 20(3), 273-297.

Du, Y. (2018), Application and Analysis of Forecasting Stock Price Index Based on Combination of ARIMA Model and BP Neural Network. Paper Presented at the 2018 Chinese Control and Decision Conference (CCDC).

Fenghua, W.E.N., Jihong, X., Zhifang, H.E., Xu, G. (2014), Stock price prediction based on SSA and SVM. Procedia Computer Science, 31, 625-631.

Grinblatt, M., Titman, S. (1987), The relation between mean-variance efficiency and arbitrage pricing. Journal of Business, 60, 97-112.

Grinblatt, M., Titman, S. (1989), Portfolio performance evaluation: Old issues and new insights. The Review of Financial Studies, 2(3), 393-421.

Grinblatt, M., Titman, S. (1992), The persistence of mutual fund performance. The Journal of Finance, 47(5), 1977-1984.

Gupta, R., Pierdzioch, C., Vivian, A.J., Wohar, M.E. (2018), The predictive value of inequality measures for stock returns: An analysis of long-span UK data using quantile random forests. Finance Research Letters. https://doi.org/10.1016/j.frl.2018.08.013.

Hassan, M.R., Nath, B., Kirley, M. (2007), A fusion model of HMM, ANN and GA for stock market forecasting. Expert Systems with Applications, 33(1), 171-180.

Henrique, B.M., Sobreiro, V.A., Kimura, H. (2019), Literature review: Machine learning techniques applied to financial market prediction. Expert Systems with Applications, 124, 226-251.

Hiransha, M., Gopalakrishnan, E.A., Menon, V.K., Soman, K.P. (2018), NSE stock market prediction using deep-learning models. Procedia Computer Science, 132, 1351-1362.

Hsu, S.H., Hsieh, J.J.P., Chih, T.C., Hsu, K.C. (2009), A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression. Expert Systems with Applications, 36(4), 7947-7951.

Huang, C.F., Hsieh, T.N., Chang, B.R., Chang, C.H. (2011), A Comparative

Study of Stock Scoring Using Regression and Genetic-based Linear Models. Paper Presented at the 2011 IEEE International Conference on Granular Computing.

Huang, C.S., Liu, Y.S. (2019), Machine learning on stock price movement forecast: The sample of the taiwan stock exchange. International Journal of Economics and Financial Issues, 9(2), 189-201.

Jensen, M.C. (1968), The performance of mutual funds in the period 1945-1964. The Journal of Finance, 23(2), 389-416.

Jensen, M.C. (1972), Optimal utilization of market forecasts and the evaluation of investment performance. In: Mathematical Methods in Investment and Finance. Amsterdam: North-Holland Publishing Company.

Jobson, J.D., Korkie, R.M. (1981), Putting Markowitz theory to work. The Journal of Portfolio Management, 7(4), 70-74.

Kara, Y., Boyacioglu, M.A., Baykan, Ö.K. (2011), Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul stock exchange. Expert Systems with Applications, 38(5), 5311-5319.

Karaatli, M., Gungor, I., Demir, Y., Kalayci, S. (2005), Estimating stock market movements with neural network approach. Journal of Balikesir University, 2(1), 22-48.

Kazem, A., Sharifi, E., Hussain, F.K., Saberi, M., Hussain, O.K. (2013), Support vector regression with chaos-based firefly algorithm for stock market price forecasting. Applied Soft Computing, 13(2), 947-958.

Khan, W., Ghazanfar, M., Asam, M., Iqbal, A., Ahmad, S., Khan, J.A. (2016), Predicting trend in stock market exchange using machine learning classifiers. Science International, 28(2), 1363-1367.

Kim, H.J., Shin, K.S. (2007), A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets. Applied Soft Computing, 7(2), 569-576.

Lee, T.K., Cho, J.H., Kwon, D.S., Sohn, S.Y. (2019), Global stock market investment strategies based on financial network indicators using machine learning techniques. Expert Systems with Applications, 117, 228-242.

Leung, M.T., Daouk, H., Chen, A.S. (2000), Forecasting stock indices: A comparison of classification and level estimation models. International Journal of Forecasting, 16(2), 173-190.

Li, X., Xie, H., Chen, L., Wang, J., Deng, X. (2014), News impact on stock price return via sentiment analysis. Knowledge-Based Systems, 69, 14-23.

Li, X., Xie, H., Song, Y., Zhu, S., Li, Q., Wang, F.L. (2015), Does summarization help stock prediction? A news impact analysis. IEEE Intelligent Systems, 30(3), 26-34.

Long, W., Lu, Z., Cui, L. (2019), Deep learning-based feature engineering for stock price movement prediction. Knowledge-Based Systems, 164, 163-173.

Malkiel, B.G., Fama, E.F. (1970), Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383-417.

Modigliani, F., Modigliani, L. (1997), Risk-adjusted performance. Journal of Portfolio Management, 23(2), 45.

Olson, D., Mossman, C. (2003), Neural network forecasts of Canadian stock returns using accounting ratios. International Journal of Forecasting, 19(3), 453-465.

Patel, J., Shah, S., Thakkar, P., Kotecha, K. (2015), Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Systems with Applications, 42(1), 259-268.

Qiu, M., Song, Y., Akagi, F. (2016), Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market. Chaos, Solitons and Fractals, 85, 1-7.

Treynor, J.L., Black, F. (1973), How to use security analysis to improve portfolio selection. The Journal of Business, 46(1), 66-86.

Wang, J., Hou, R., Wang, C., Shen, L. (2016), Improved Support vector regression model based on variable selection and brain storm optimization for stock price forecasting. Applied Soft Computing, 49, 164-178.

Xi, L., Muzhou, H., Lee, M.H., Li, J., Wei, D., Hai, H., Wu, Y. (2014), A new constructive neural network method for noise processing and its application on stock market prediction. Applied Soft Computing, 15, 57-66.

Xu, X., Zhou, C., Wang, Z. (2009), Credit scoring algorithm based on link analysis ranking with support vector machine. Expert Systems with Applications, 36(2, Part 2), 2625-2632.

Yang, F., Chen, Z., Li, J., Tang, L. (2019), A novel hybrid stock selection method with stock prediction. Applied Soft Computing, 80, 820-831.

Zhang, K., Zhong, G., Dong, J., Wang, S., Wang, Y. (2019), Stock market prediction based on generative adversarial network. Procedia Computer Science, 147, 400-406.