ABSTRACT

Stock price manipulation is an essential issue for both developed and emerging stock markets. It has been recognized as harmful practices in global share markets. We investigated the characteristics of the stock price manipulation. Two manipulation techniques were studied: pump and dump, spoof trading. Due to the new technology, machine learning algorithms now supporting to detecting stock price manipulation.

In this paper, we use Indonesia Stock Exchange intraday data, and we investigated stocks involved in the Unusual Market Activity (UMA) Announcements during 2008-2017, is there any connection between the UMA and the stock price manipulation. We use Neural network model to detect stock price manipulation for each manipulation techniques, after that we want to detect price manipulation, so we took randomly 30 stocks and input it into 2008-2017 model and from the output we got in the first year of the UMA policy the announcement can detect manipulation but in the next following years the effectiveness is decreasing. From this study, we discover that UMA can be proxies to detect stock price manipulation, but for the long term the announcement can't be a proxy anymore because of technology and science are continuously evolving, the Indonesia Stock Exchange should look for better consideration and evaluation in assessing and publishing UMA announcements

Keywords: Stock Price Manipulation, Unusual Market Activity, Machine Learning, Pump and Dump, Spoof Trading, Neural Network

1. Introduction

Today, the stock market is an significant and active component of the financial market. By means of a fundamental or technical analysis, investors want to produce the biggest possible profit on the market. The potential for stock markets (both developed and emerging) to be manipulated is a significant challenge to trade regulation and market efficiency.

The manipulation of the stock market is a kind of deception that generates an artificial stock price. There are many techniques to create price manipulation, such as scam-making in-stock discussion or misinformation and news. When there are individuals in the business who know the important news about the business and then take advantage of it to avoid losses or profits by buying or selling their own shares. Manipulator purchases enough of a particular stock, commodity, or other asset to regulate supply and price it.

Weak market structure and regulation make some investors want to benefit from tendentially significant illegal intervention, which is likely to occur in emerging markets such as Indonesia. Most emerging capital markets function as joint ventures so that stockbrokers own those markets. It is primarily managed by brokers, i.e. most of the board of exchange directors, including the chairperson, is brokers. In addition, trading on the stock exchange can only be carried out by licensed brokers. A stockbroker, as a business player and also as a shareholder, has a conflict of interest. The Indonesian Capital Market Act 1998 stated that only stockbrokers are permitted to own Indonesian capital market shares. In order to enhance governance and market integrity, the global trend shows that global capital markets have changed their market ownership structure. External owners make the policy technique more transparent in the capital market.

Starting April 2008, Indonesia Stock Exchange announce Unusual Market Activity , this say that there is a unusual price (increasing or decreasing) from a stock that possibility disrupt the market, Indonesia Stock Exchange said that the announcement does not necesserely indicate there is a manipulation on that stock. After being announced, IDX keep alert from that stock if there any suspicious movement or not, if there still, the IDX will suspend the stock. So in this research we want to know if the UMA announcement can detect stock price manipulation, this research can be used by investor to know is it safe to invest when the stock being announced by IDX, and it can be benchmark for IDX for knowing if the stock being manipulated or not. Felisca (2013) said that the UMA effect will occur 6 days after the Indonesia Stock Exchange announce the UMA . Hanafi (2010) said that there is a chance a stock being manipulated using information based manipulation.

This research discusses an approach to detect trade-based price manipulations in stock market using Unusual Market Activities (UMA) announcements in Indonesia Stock Exchange (it happen when the stocks increased or decreased the price at least 25% in some day transactions). In this research, we use two manipulations: pump-and-dump and spoof trading. Pump-and-dump is purchasing stock, making the price higher, and then selling in a short time to others. Spoof trading is an action to send high-volume passive orders to trick others to sell the stock at that cost. They cancel their passive orders after the manipulators obtain advantages from the fake price. These activities enable manipulators to sell their stock at a higher than usual cost. In some nations, these activities are illegal and should be brought under control. The more information we have will improve manipulation detection efficiency. We depend on the price information sent to the market by buyers and sellers.

2. Previous Research

ANN has been introduced in a number of fields. You can discover a survey of their applications in finance in Fadlalla and Lin (2001). Several neural network studies have been performed in financial markets. Collard (1990) stated that his commodity-training neural network model would have resulted in significant profits over other trading strategies. Kamijo and Tanigawa (1990) used a neural network to map Tokyo Stock Exchange information. They found that the model's results would defeat the "purchasing and keeping" policy. Finally, a neural model was developed, using a number of economic indicators, to predict the percentage change in the S&P 500 five days ahead (Fishman et al., 1991). The writers argue that the model used the same indices to provide more precise predictions than the supposed field specialists. Neural networks have been efficiently trained to determine whether to approve loan applications (Gallant, 1988). Neural networks have also been shown to be better prepared to predict mortgage applicants ' solvency than mortgage writers (Collins et al., 1988). Another region where neural networks have been efficiently introduced predicts the rating of corporate bonds and attempts to predict their profitability (see Dutta and Shakhar, 1988; and Surkan and Singleton, 1990). Neural networks have outperformed regression analysis and other mathematical modeling instruments to predict bond rating and profitability. The primary conclusion reached was that the neural networks offered a more general structure for linking the company's economic data to the corresponding bond rating. Fraud prevention is another area for the use of neural networks in company. Chase Manhattan Bank has effectively used neural networks to deal with loan card fraud (Rochester, 1990), with neural network models outperforming traditional regression methods. Furthermore, neural networks were used to validate bank signatures (see Francett, 1989; and Mighell, 1989). These networks have acknowledged far stronger falsehoods than any human expert. Leangarun (2016) predict price manipulation using pump-dump technique and spoof trading techniques with neural network model, he compared level 1 data and level 2 data, the result pump-dump level 1 data and level 2 data were close, that showing the effectiveness of the pump-dump model, but in spoof trading they can't compare it because of limitness of information in level 1 data.

3. Manipulation Strategies

Manipulation of the market is nothing new and has existed for as long as there have been markets. Most techniques depend on changing market perceptions, such as trade volumes, demand and supply in order books, and other important determinants. However, today's increasingly complicated computerized markets enable for fresh and innovative techniques of market manipulation. Due to the rapid speed of innovation and sometimes the overwhelming complexity of trading activities, regulation has lagged behind. In this paper we use two techniques : pump and dump technique and spoof trading technique

3.1 Pump and Dump

Traditionally, fraudsters are implementing a pump and dump system that attempts to inflate a security's price by making false or misleading claims about its worth, generally through cold calling or distributing false press releases. These traders, who had accumulated a big holding in the safety before the "pump" stage began, liquidate their position as prices rise, giving birth to the "dump" stage, in which the security generally ends up erasing its prior profits, inflicting huge losses on the "pump" phase's late joiners. More lately, high-frequency traders have often embraced pump and dump approaches, creating algorithms that temporarily drive up a safety price, only to reverse their stance quickly and capitalize on false momentum at the cost of other traders.

The most frequent victims of pump and dump manipulations are micro and small-cap stocks, which due to their small trading volumes tend to be the easiest to manipulate. Moreover, it is harder for most investors to assess the real value of these companies since they are relatively unknown to the public and not covered by research analysts, and thus information concerning their fundamental value is hard to find. The algorithm will provide in this paper. Lengarun (2016) said that there are two condition to detect pump and dump, first when dump period and the second one is when pump period.

Step 1: In the dumping period, the number of cancellation orders and price matching orders were verified. For the first condition, we classified the event as a dumping action when the number of the withdrawal of buy orders is more than threshold1(25%) of the average volume of buy orders during that period.

$$volume = \begin{cases} 1; V_{buy}^{cancel}(t) > (threshold1)(E\left[V_{buy}^{matched}(t)\right]) \\ 0; otherwise \end{cases} \quad (1)$$

For the second condition, the difference between the high price of sell orders and the low price of sell orders is more than threshold2(25%).

$$dump = \begin{cases} 1; \frac{P_{sell}^{max}(t) - P_{sell}^{min}(t)}{P_{sell}^{max}(t)} > threshold2 \\ 0; otherwise \end{cases} \quad (2)$$

Then, we treated that the dumping action occurs when both condition 1 and condition two are met.

$$dump = volume \cap dump \quad (3)$$

When dump activities (step 1) are detected, the condition in step 2 is tested. If dump activities are not detected, we identify it as normal activities or non-manipulated of a stock.

Step 2: After the dumping points are detected**,** we test whether the price was pumping up before. We treated the event as a pumping activity when the highest bid price increases more than threshold3(25%) compared to the lowest bid price at that point.

$$pump = \begin{cases} 1; \frac{P_{bid}^{max}(t-1) - P_{bid}^{min}(t)}{P_{bid}^{max}(t-1)} > threshold3 \\ 0; otherwise \end{cases} \quad (4)$$

$$pump\ and\ dump = pump \cap dump \quad (5)$$

When the conditions of both step 1 and step 2 are satisfied, we treated them as pump and dump points.

$V_{buy}^{cancel}(t)$: cancellation volume of buy orders at time t

$E\left[V_{buy}^{matched}(t)\right]$: average volume of buy orders that have been matched at time t

$P_{sell}^{max}(t)$: the highest price of sell orders at time t

$P_{sell}^{min}(t)$: the lowest price of sell orders at time t

$P_{bid}^{max}(t)$: the highest bid price at time t

$P_{bid}^{min}(t)$: the lowest bid price at time t

3.3.2 Spoofing

Spoofing is a tactic by which one places orders to restrict and remove them before executing them. By spoofing order limits, perpetrators hope to distort market demand and supply perceptions of other traders. For example, with the intention of being cancelled before it is executed, a large bid limit order could be placed. The spoofer would then seek to benefit from rising prices as other people would see in the market structure as a result of false optimism. Due to the large quantity of spoof orders, other investors generally believe this price comes from consensus and is sensible.

More contentious was the act of layering that shares many similarities with outright spoofing, but differs in that orders are put uniformly across prices with the aim of reserving a priority for early execution at each specified price level. If at that price point the person has no trade to execute, the orders will be removed. Although the act of layering is more benign in nature, it also distorts market demand and perception of supply.

From leangrun (2016) there is three condition to detect spoof trading: cancellation order has its price close to the current bid or ask price, cancellation volume is high enough, and there is high volume for the last buy order. For the first condition, we treated the event as a spoof trading action when the cancellation sell orders have a price $P_{sell}^{cancel}(t)$ within 0.25% (threshold_4) of the current ask price $P_{buy}^{matched}(t)$.

$$spoofcond1 = \begin{cases} 1; \left|\frac{P_{sell}^{cancel}(t) - P_{buy}^{matched}(t)}{P_{sell}^{cancel}(t)}\right| < threshold4 \\ 0; otherwise \end{cases} \quad (6)$$

For the second condition, the amount of the cancellation sell orders $V_{sell}^{cancel}(t)$ is more than threshold_5, where threshold_5 is five times of the summation of matching orders $\sum_n^{t-1} V^{matched}(n)$ since the starting point.

$$spoofcond2 = \begin{cases} 1; V_{sell}^{cancel}(t) > (threshold5)(\sum_1^t V^{matched}(t)) \\ 0; otherwise \end{cases} \quad (7)$$

the third condition, the amount of matching buy orders $V_{buy}^{matched}(t)$ is more than 25% (threshold_6) of the summation of matching orders $\sum_n^{t-1} V^{matched}(n)$ since the starting point.

$$spoofcond3 = \begin{cases} 1; V_{buy}^{matched}(t) > (threshold6)(\sum_1^t V^{matched}(t)) \\ 0; otherwise \end{cases} \quad (8)$$

So the spoof trading will occur when all spoof condition above satisfied.

$$spoof\ trading = spoofcond1 \cap spoofcond2 \cap spoofcond3. \quad (9)$$

$V_{sell}^{cancel}(t)$: cancellation volume of sell orders at time t

$V^{matched}(t)$: volume of buy and sell orders that have been matched at time t

$V_{buy}^{matched}(t)$: volume of buy orders that have been matched at time t

$P_{sell}^{cancel}(t)$: the price of sell orders that have been cancelled at time t

$P_{buy}^{matched}(t)$: the price of the last buy orders that have been matched at time t

4. Model Building

4.1 Neural Network Flow Chart

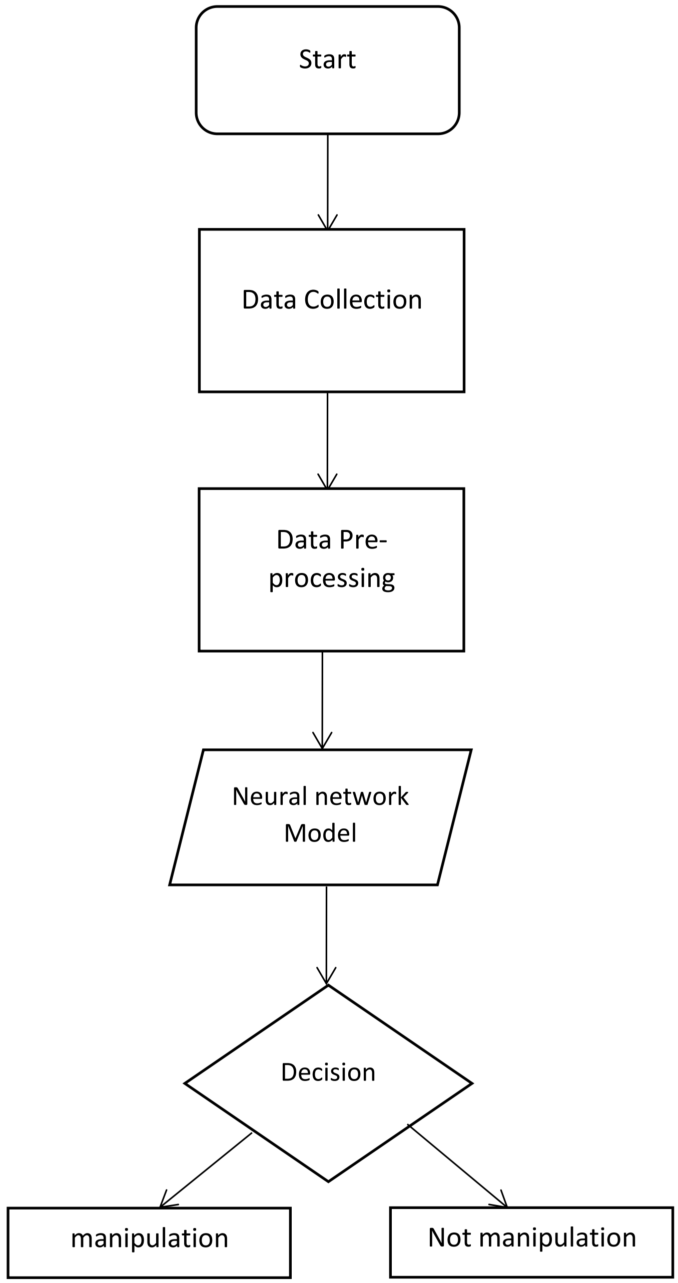

Figure 4.1 Model Flowchart

4.2 Data Collection

Starting in April 2008, the Indonesian Stock Exchange (IDX) has occasionally issued Unusual Market Activity (UMA) announcements. In this research, we used 2008-2017 intraday stock UMA. In 2008, there were 49 UMA announcements issued by the IDX. The IDX issued positive UMA more often than negative UMA, i.e., 38 and 11, respectively. In this research, we only use the positive UMA, in table 4.1, we can see the positive UMA in 2008-2017. Because of the pump and dump technique, this technique need a pumping price to execute the action, so we need a positive UMA, which is the price increasing. And the negative UMA in next following years is decreasing, so the action is not many as the positive one.

Year	UMA
2008	38
2009	77
2010	45
2011	68
2012	80
2013	71
2014	90
2015	59
2016	118
2017	113

Table 4.1 2008-2017 Positive UMA

We can see from table 4.1 that every year stock that being announced by UMA increasing. We use 38 positive UMA announcements in 2008 to build our base model. The intraday trading data get from IDX.

4.3 Data Pre-processing

After we collected the raw data, we transformed the data using equation 1,2,4,6,7,8 (pump & dump ,spoofing) into the variable , which used as inputs, and the result from the calculation will be the output . The model make an output value in the probability that stock price will be manipulated or not[0,1]. Table 4.2 is how to measure the variable for each manipulation.

Pump and Dump Variable	**How to Measure**
Pump Condition	$\frac{P_{bid}^{max}(t-1) - P_{bid}^{min}(t)}{P_{bid}^{max}(t-1)}$
Dump Condition	$\frac{P_{sell}^{max}(t) - P_{sell}^{min}(t)}{P_{sell}^{max}(t)}$
Volume Condition	$\frac{V_{buy}^{cancel}(t)}{(E\left[V_{buy}^{matched}(t)\right]}$
Spoof Trading Variable	
Condition 1	$\left\lvert\frac{P_{sell}^{cancel}(t) - P_{buy}^{matched}(t)}{P_{sell}^{cancel}(t)}\right\rvert$
Condition 2	$\frac{V_{sell}^{cancel}(t)}{(\sum_1^t V^{matched}(t)}$
Condition 3	$\frac{V_{buy}^{matched}(t)}{(\sum_1^t V^{matched}(t)}$

Table 4.2 Equation of Variable From Each Manipulation Techniques

4.4 Neural Network Models

After we get all data and input them as variables, we went to detect which stock manipulated and what techniques that manipulator use in Indonesia, we use Neural Networks in Matlab, we used neural network because to their ability to "learn" from the data, their nonparametric nature (i.e., no rigid assumptions), and their ability to generalize. Neural computing refers to a pattern recognition methodology for machine learning. Many business applications using neural network for pattern recognition, forecasting, prediction, and classification.

A neural network, called neurons, is a bio-inspired system with several single processing components. A joint mechanism, consisting of a set of allocated weights, connects the neurons to each other. While the developers of this technique used many biological terms to explain the inner workings of the modeling process of the neural network, it has a simple mathematical basis. Equation 10 is example a linear model of mathematics:

$$y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \qquad (10)$$

Where y is the calculated output and X_1, X_2, and X_3 are input attributes. α is the intercept and β_1, β_2, and β_3 are the coefficients for the input attributes X_1, X_2, and X_3, respectively. . We can represent this simple linear model in a topological form as shown in Figure 3.3

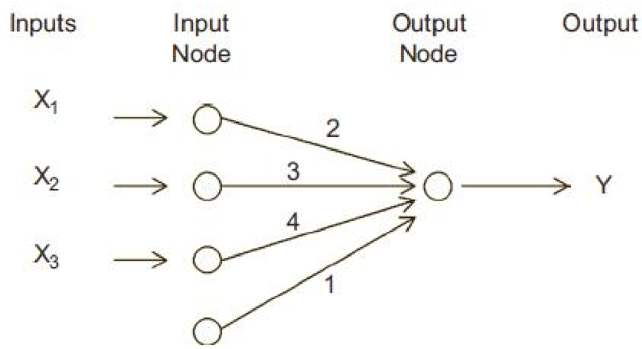

Figure 4.2 Model Topology

The topology is also a simple neural artificial network (ANN). The ANN is a biological nervous system-inspired computational and mathematical model. Some of the terms used in an ANN are therefore borrowed from biological counterparts. Typically, an artificial neural network is used to model nonlinear, complex relationships between variables of input and output. In addition to the input and output layers called hidden layers, this is created possible by the presence of more than one layer in topology. A hidden layer includes a node layer that links input from prior layers and an activation function is applied. A more complicated mixture of input values now calculates the yield, as shown in Figure 3.4

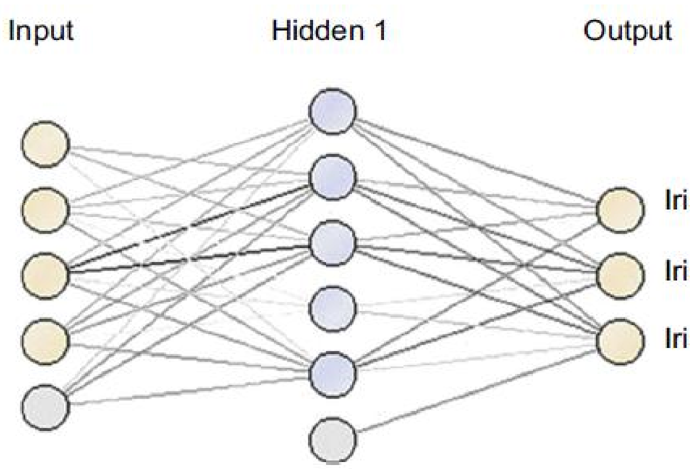

Figure 4.3Topology of Neural Network Model

The information will be processed if the structure of a neural network determined. Now we show the major concepts of neural network processing.

Inputs : each input corresponds to a single attribute. For example, if the problem is to decide on pass or not pass of an exam, the attributes are study or not, duration of the study, have the book or not,etc. In this research we using pump and dump condition , and spoofing condition. So there are three node for each manipulation, for pump and dump there are pump condition, dump condition, and volume condition. For spoof trading there are spoofcondition1,spoofcondition2,spoof condition3.

Outputs : the outputs of a network contain the solution to a problem. For example, in the case of an exam, the outputs can be pass or not. The output that ANN generate will be binomial condition, such as 1 for yes and 0 for no. The purpose of the network is to compute the values of the output. The output needed because several networks use two outputs: one for yes and the other for no. Usually we have to round it(near to 1 or near to 0). In this research there are one node (manipulated or not manipulated)

Connection Weights: connection weights are the key components of this processing. They demonstrate how many connections from layer to layer transferring information. Weights are essential because they maintain the data patterns that have been learned. The network is learning by repeating weight changes.

Summation Function : the summation function computes the weighted sums of all the input elements entering each processing element. A summation function multiplies each input value by its weight and totals the values for a weighted sum Y.

Transformation function: the function of summation calculates the neuron's inner stimulation or amount of activation. Based on this point, the neuron may or may not produce an output. The link between the internal activation level and the output can be linear or nonlinear. The link reflects one of several types of transformation (transfer) characteristics. The conversion (transfer) function combines (i.e. adds) the inputs from other neurons / sources into a neuron and then produces an output based on the choice of the transfer function. We use activation function tansig in this study, it is an S-shaped transfer function in the range of-1 to 1. This activation transforms the level of output from 0 to 1.

4.5 Decision

There are three datasets to build the model in machine learning: training dataset, test dataset, and validation dataset . Dataset training is the data sample used to fit the model. The real dataset we use to train the model (in the event of Neural Network weights and biases). From this data, the model sees and learns.

Validation Dataset is the data sample used to provide an unbiased model fit assessment on the training dataset while tuning hyperparameters of the model. The assessment becomes more biased as skill is integrated into the setup of the model on the validation dataset.To evaluate a given model, the validation set is used, but this is for frequent evaluation. These data are used by machine learning engineers to fine-tune the hyperparameters of the model.

Test Dataset: The data sample used to provide the training dataset with an unbiased evaluation of a final model fit. The Test dataset offers the model's gold standard. Only when a model is fully trained (using train and validation sets) is it used. Generally speaking, the test set is used to assess competing models. The validation set is often used as the test set, but it's not excellent practice. Generally speaking, the test set is well cured. It includes thoroughly sampled information spanning the different classes facing the model when used in the actual globe.

The iteration stops when any of these conditions occur:

- The maximum number of epochs (repetitions) is reached.
- The maximum amount of time is exceeded.
- Performance is minimized to the goal.
- The performance gradient falls below min_grad.

We use cross entropy in this study to calculate our efficiency because we want to drive output node values to either 1.0 or 0.0 depending on the target values during back-propagation training. If we use MSE, the weight adjustment factor (the gradient) contains a term of (output) * (1 – output). As the calculated output approaches either 0.0 or 1.0, the value of (output) * (1 – output) becomes lower and lower. As the adjustment factor gets smaller and smaller, the change in weights gets smaller and smaller and training can stall out.

But if we use cross-entropy error, the (output) * (1 – output) term goes away. So, the weight changes don't get smaller and smaller and so training isn't s likely to stall out. In this research we want the performance small as possible so we can create a model with a good performance.

5. Result and Analysis

5.1 Pump and Dump Techniques

First, take six days pre UMA announcement and post UMA announcement because, in this range of days, the UMA effect will occur (Felisca 2013). then from the raw data, we choose variable that in (example for one stock see in table 5.1)

day	match bid	min sell	max bid	bid start	max sell	volume cancel	average match buy
h-7	90	90	90	90	105	72500	500
h-6	93	92	93	90	92	0	12500
h-5	0	100	0	90	100	0	0
h-4	92	90	92	90	92	2500	7307,692308
h-3	90	0	90	90	90	0	14166,66667
h-2	0	0	0	90	0	0	0
h-1	90	90	90	90	90	0	50000
0	90	89	90	90	90	0	16666,66667
h+1	92	90	92	90	98	100000	34750
h+2	95	94	95	90	104	0	19600
h+3	101	94	101	90	101	0	74500
h+4	105	105	105	90	130	35000	6375
h+5	105	104	105	90	125	0	20333,33333
h+6	110	105	110	90	110	0	25250

Table 5.1 ABBA stock in one trading day

Then, we input the variable into equation 1,2,4 and the result for one stock in table 5.2, because in 2008 there are 38 positive UMA, so there are 608 data for 2008.

Day	dump	Volume	pump
h-6	0,086956522	0,0752	0,032258065
h-5	0,1	0	0
h-4	1	0,1504	0,02173913
h-3	1	0,1504	0
Day	dump	Volume	pump
h-2	0	0,1598	0
h-1	0,011111111	0,1504	0
0	0	0,1034	0
h+1	0	0,188	0,02173913
h+2	0	0,0658	0,052631579
h+3	0	0,1504	0,108910891
h+4	0,2	0,0846	0,142857143
h+5	0,16	0,1692	0,142857143
h+6	0,227272727	0,0658	0,181818182

Table 5.2 ABBA pre-processing data

Then we calculate the output layer with equation 5 and 9, and in table 4.5 is the example from the data using ABBA data. We can see that in table 4.5 there are 5 data that categorized pump and dump.

pump condition	volume condition	dump condition	pump and dump
1	1	1	1
0	1	0	0
1	1	1	1
1	1	0	0
1	1	1	1
0	0	0	0

0	1	0	0
0	0	0	0
1	0	0	0
1	0	0	0
1	0	0	0
1	1	1	1
1	1	1	1

Table 5.3 ABBA Output Layer

After we get all the data we input this into the matlab software, and we get the iteration process.

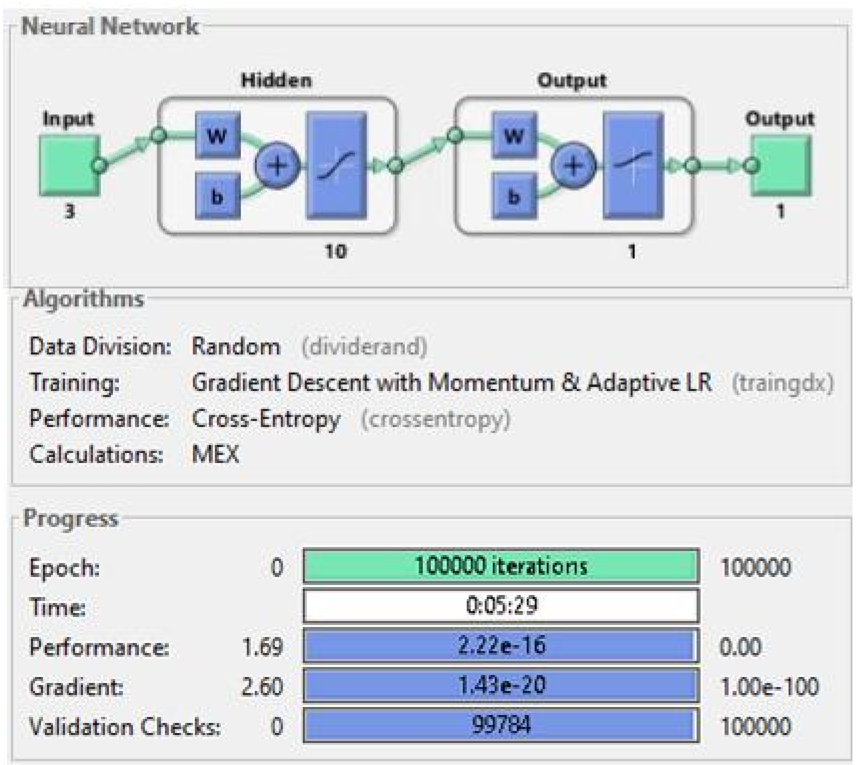

Figure 5.1 Iteration Process for Pump and Dump Technique

We use Gradient Descent with Momentum & Adaptive LR because this training function is always updates the weight and bias values. First we input the three condition that state in table 3.2, after that in the hidden layer we use 10 node, we use 10 node by trial and error. Then the weight and bias values updating until the performance is good enough, if the performance good enough the output from the model will come. We can see from figure 5.1 the model need 100000 iterations and five minutes to complete the iteration, the model is good because of the performance of the model is 2.22e-16 (0 is the best), from chapter III we said that in machine learning there are 3 datasets to build the model: Training dataset, Validation dataset, and Test data set. So first we look at the training dataset. We use 70% for the data for training, 15% for Validation, and 15% for test.

So the equation for 2008 data is

$$-0,9795+\tanh\left(w_1 * x_{2008}\right)$$

Where

w_1 is weight value for output2008.

After that we do it for 2009 until 2017 with same procedure above, in this case the model will learn from the data which one is manipulated and which one is not, and the model will change every year, for 2009-2017 the equation are :

$$\text{Output2009}=-3.4239+\tanh\left(w_2 x_{2009}\right)$$
$$\text{Output2010}=-2,239+\tanh\left(w_3 x_{2010}\right)$$
$$\text{Output2011}=-1,4412+\tanh\left(w_4 x_{2011}\right)$$
$$\text{Output2012}=-2,7834+\tanh\left(w_5 x_{2012}\right)$$
$$\text{Output2013}=-0,7283+\tanh\left(w_6 x_{2013}\right)$$
$$\text{Output2014}=-3,2782+\tanh\left(w_7 x_{2014}\right)$$
$$\text{Output2015}=-2,3821+\tanh\left(w_8 x_{2015}\right)$$
$$\text{Output2016}=-1,7343+\tanh\left(w_9 x_{2016}\right)$$
$$\text{Output2017}=-0,3247+\tanh\left(w_{10} x_{2017}\right)$$

Where :

w_2 is weight value for output2009

w_3 is weight value for output2010

w_4 is weight value for output2011

w_5 is weight value for output2012

w_6 is weight value for output2013

w_7 is weight value for output2014

w_8 is weight value for output2015

w_9 is weight value for output2016

w_{10} is weight value for output2017

$w_1 w_2 w_3 w_4 w_5 w_6 w_7 w_8 w_9 w_{10}$ are weight values for each model. This value will change if we input the new data, so each model will always learn which one is manipulated and which one is not manipulated.

After that, we want to know if the UMA announcement can detect price manipulation, so we try with 38 stock randomly that in UMA announcement 2008-2017 and then we input the data to their year output model, and the result in table 5.4

Year	Manipulated	not manipulated
2008	33	5
2009	31	7
2010	28	10
2011	28	10
2012	26	12
2013	25	13
2014	23	15
2015	23	15
2016	22	16

2017	21	17

Table 5.4 Output Decision from Randomly Stock Each Year

We can see from the table that every year, the effectiveness of UMA to detect price manipulation is declining. Figure 5.2 has two meaning; first, the model discourage the manipulator from doing price manipulation.

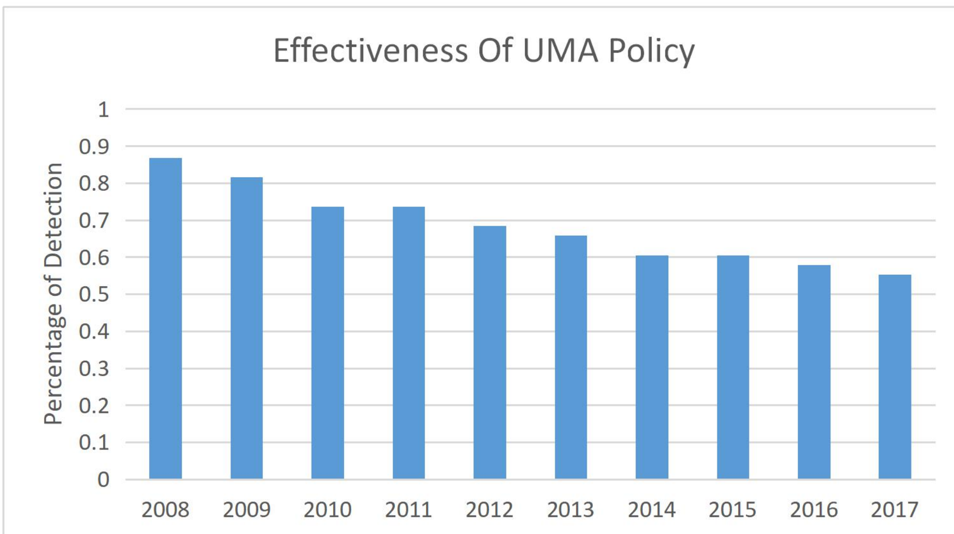

Figure 5.2 Effectiveness of UMA Detection

Second the manipulator search for a new technique to do their action, so the graph is declining. We do not know which one the meaning of this model because on this model, we only use pump and dump technique. Science and technology are expanding, and people get smart, there is some reason that UMA policy cannot detect price manipulation well.

5.2 Spoofing Techniques

Because of spoofing technique allows the execution of large trade orders in a very short time, so we only look at the day when UMA announced for each stock. We used 30 minutes trading in 2008 stock market data for one-day trading because Indonesia stock market is not liquid as America stock market. then we choose the data that we want to compute for the equation 6,7,8 in chapter 3 some example in table 5.5

time	cancel sell	last buy matched	cancel volume sell	volume buy matched	volume buy and sell
10.00	0	93	0	30000	30000
10.30	0	0	0	0	0
11.00	0	0	0	0	0
11.30	0	0	0	0	0
12.00	0	0	0	0	0
13.30	0	90	0	25000	28000
14.00	94	0	15000	0	3000
14.30	98	0	15000	0	2500
15.00	0	92	0	12000	12000
15.30	92	0	57500	0	0
16.00	130	0	25000	0	0

Table 5.5 ABBA in one trading days

After we get all the data we transform to equation in table 4.2 , and some result in table 5.6

time	condition1	condition2	condition3
10.00	#DIV/0!	0	1
10.30	#DIV/0!	0	0
11.00	#DIV/0!	0	0
11.30	#DIV/0!	0	0
12.00	#DIV/0!	0	0
13.30	#DIV/0!	0	0,431034483
14.00	0,042553191	0,245901639	0,049180328
14.30	0,06122449	0,236220472	0,039370079
15.00	#DIV/0!	0	0,158940397

15.30	0,02173913	0,761589404	0
16.00	0,284615385	0,331125828	0

Table 5.6 ABBA pre-processing data

From chapter III, we said that in machine learning, there are three datasets to build the model: Training dataset, Validation dataset, and Test data set. So first we look at the training dataset. We use 70% for the data for training, 15% for Validation, and 15% for test.

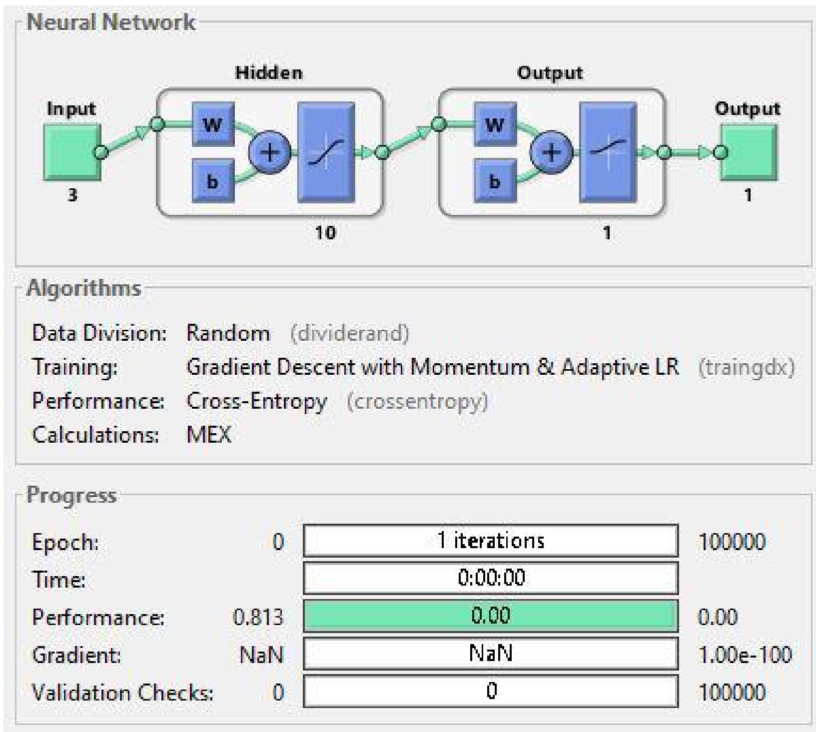

Figure 5.3 Iteration Process for Spoofing Technique

We use Gradient Descent with Momentum & Adaptive LR because this training function is always updates the weight and bias values. First we input the three condition that state in table 4.2, after that in the hidden layer we use 10 node, we use 10 node by trial and error. Then the weight and bias values updating until the performance is good enough, if the performance good enough the output from the model will come (manipulated or not). We can see from figure 5.3 that the algorithm can't build the model because of in condition1 variable there is undefined (divided by zero) variable, so the iteration won't work.

Because of there is no output from the model, we cant use it to detect price manipulation, it because there are some condition1 variable is undefined (the calculation divided by 0) so the algorithm can't generate the model. In America spoof trading occur in less than five minutes, and we use based on 30 minutes trading in Indonesia and we can't build the mode. So we can't use spoof technique to detect stock manipulation because spoofing need fast trading to execute their action.

6. Conclusions

This research develops the model that predicts the probability of stock price being manipulated using Intra-day data trading. We create the model based on market data and corporate action, and we used two techniques manipulation there are pump and dump technique, and spoofing. From this research, we used level 2 data to build the model.

From the result, It is found in the first year of UMA announcements (2008) can detect price manipulation but in the next following years the effectiveness of the UMA announcements is declining, the output has two meanings; first, the model discourage the manipulator from doing price manipulation. Second, the manipulator search for a new technique to do their action. So the IDX should change their threshold (25% price increased or decreased). Unfortunately for the spoofing model show that the model not enough to predict the stock being manipulated or not. So in Indonesia, manipulator using more likely to use pump and dump technique to manipulate the price of the stock. The spoofing technique cant applied in Indonesia because in Indonesia is not a liquid market like the American market and this manipulation need a execute fast like 3-4 seconds after the manipulator put the stock in the market.

REFERENCES

Allen F. and D. Gale. 1992. Stock-Price Manipulation. The Review of Financial Studies 5(3): pp 503-529.

Allen, F., and Gorton, G. 1992. Stock price manipulation, market microstructure and asymmetric information. European Economic. Review 36: pp. 624–630.

Banz, R. W. 1981. The relationship between return and market value of common stock. Journal of Financial Economics 9 (1) (March): 3-18

Barclays, M., and B. W. Jerold. 1993. Stealth trading and volatility: Which trades move prices. Journal of Financial Economics34 (3) (December): 281-305.

Camerer, C. 1998. Can asset markets be manipulated? A field experiment with racetrack betting. Journal of Political Economy 106 (3) (June): 457-482.

Cheuk, M-Y., K. F. Dennis, and W. S. Raymond. 2006. Insider trading in Hong Kong: Some stylized facts. Pacific-Basin Finance Journal 14 (1): 73-90.

Collard, J.E. 1990. "Commodity Trading with a Neural Net." Neural Network News, Vol. 2, No. 10.

Collins, E., S. Ghosh, C.L. and Scofield. 1988. "An Application of a Multiple Neural Network Learning System to Emulation of Mortgage Underwriting Judgments," IEEE International Conference on Neural Networks.

Dissanaike G. and Kim H. L. 2015. Detecting and Quantifying Insider Trading and Stock Manipulation in Asian Markets. Asian Economic Papers 14:3

Dutta, S., and S. Shakhar. 1988. "BondRating: A Non-Conservative Application of Neural Networks" Proceedings of the IEEE International Conference on Neural Networks, San Diego

Fadlalla, A., and C. Lin. 2001. "An Analysis of the Applications of Neural Networks in Finance." Interfaces, Vol. 31, No. 4.

Fishman, M., D. Barr, and W. Loick.1991. “Using Neural Networks in Market Analysis,” Technical Analysis of Stocks and Commodities

Felisca, O.S.., Koesrindartoto, D. P. 2013. An Alternative Approach in Determining Unusual Market Activity (UMA) in Indonesia Srock Exchange.(Magister Science Management), Institut Teknologi Bandung

Francett, B. 1989. “Neural Nets Arrive.” Computer Decisions.

Gallant, S. 1988. “Connectionist Expert Systems,” Communications of the ACM, Vol. 31, No. 2

Guresen E., Kayakutlu G., Daim T. U. Using artificial neural network models in stock market index prediction.Expert Systems with Applications. 2011;38(8):10389–10397.

H. Öğüt, M. Mete Doğanay, and R. Aktaş.2009. Detecting stock-price manipulation in an emerging market: The case of Turkey. Expert System 36(9): pp. 11944–11949.

Hanafi, M.M. 2010 Unusual Market Activity Announcements A Study of Price Manipulation on the Indonesian Stock Exchange.Gadjah Mada International Journal of Business Vol. 12, No. 2, pp. 159–187

Imisiker, S., and B.K.O. Tas. 2013. “Which Firms Are More Prone to Stock Market Manipulation?” Emerging Markets Review 16: 119–130.

Jarrow, R. 1992. Market manipulation, bubbles, corners and short squeezes. Journal of Financial and Quantitative Analysis 27 (3) (September): 311-336.

Jiang, G., P. G. Mahoney, and J. Mei. 2005. Market manipulation: A comprehensive study of stock pools. Journal of Financial Economics 77 (1) (July): 147-170.

Kamijo, K., and T. Tanigawa. 1990. “Stock Price Pattern Recognition: A Recurrent Neural Network Approach,” International Joint Conference on Neural Networks, San Diego

Khwaja, A. I., and A. Mian. 2005. Unchecked intermediaries: Price manipulation in an emerging stock market. Journal of Financial Economics78 (1).

K. Golmohammadi, O. R. Zaiane, D. Díaz.2014. Detecting stock market manipulation using supervised learning algorithms.in : Data Science and Advanced Analytics(DSAA), 2014 International Conference on, IEEE, pp 435-441

Mighell, D. 1989. "Back-Propagation and Its Application to Handwritten Signature Verification," in D. S. Touretsky (ed.), Advances in Neural Information Processing Systems. San Mateo, CA: Kaufman.

Rochester, J. (ed.). 1990. "New Business Uses for Neurocomputing." I/S Analyzer.

Surkan, A., and J. Singleton. 1990. "Neural Networks for Bond Rating Improved by Multiple Hidden Layers." Proceedings of the IEEE International Conference on Neural Networks, Vol. 2.

Tang, Z., C. de Almieda, and P. Fishwick. 1991. "TimeSeries Forecasting Using Neural Networks vs. Box Jenkins Methodology." Simulation, Vol. 57, No. 5.

T. Leangarun, P. Tangamchit, and S. Thajchayapong. 2016. Stock price manipulation detection based on mathematical model. International Journal of Trade, Economics and Finance, Vol. 7, No. 3