The Use of Sentiment Analysis in Pre-IPO Market to Predict The Stock Performance in Post-IPO Market

Author Celine Halim, Deddy P. Koesrindartoto
[1]Bandung Institute of Technology School of Business and Management

***Abstract*:** *Nowdays, social media has been used extensively as communication tools in many different areas. Through social media, user can generate information and share it with others. There are many type of social media platform such as microblogging, stockforum, web blog, etc. In this study, author will use microblogging (Twitter) and Stock Forum for predicting the stock performances in Secondary market by examining the sentiment of related stocks in Pre-IPO Market. There will be 38 stocks which doing the IPO during 2010-2014 that will be examined in this research. The method that will be used is sentiment analysis by using Linguistic Inquiry Word Count Method (LIWC). Unlike the previous research which use the Pre-IPO stock price data to predict the stock performance in secondary market, in this research the data is collected from twitter and stock forum. This research will generate the model which can be used as trading strategy on new released stocks.*

Keywords: *Trading Strategy, Pre-IPO Market, Post-IPO Maret, Twitter, Sentiment, Stock Price*

1. Introduction

In recent years, the Indonesian market had expanded aggressively. The fact is that apparently, there has been 35 companies been trying to push an IPO process to OJK in 2015, 16% higher year on year from 2014. The fact that there are 507 companies listed remains vast options for the investors and sometimes confusing. Investors usually have their own trading strategy in determining which stock is a certain BUY, one of the strategies is to seek newly listed stock via IPO. A believe (or paradigm) that newly listed stocks will always provide enormous return in short term still lingers in the market. That goes the misleading information; the fact is that buying newly listed stocks is actually very risky. A failure in choosing the right stock with good performance at that position may result a massive loss and low probability to bounce back. Therefore, this research is trying to address this issue in order to find a good proxy to predict stock performance after IPO.

On making decisions, market sentiment plays an important role in investment considerations. To investigate, nor monitor the density and diversity of market sentiment, social media is commonly used. Social media nature as a real-time vehicle to capture sentiments is perfectly suited to anticipate the volatility of the stock markets. The social media in this issue (Indonesia) includes twitter, stock-bit, klubsaham, and bei5000.Through these social media platforms, this research collect and aggregate sentiment results of a certain stocks in a specific period of time before the IPO was even conducted in order to predict the future performance of the particular stock in the secondary market.

To support this research, Fransesca Cornelli [3] stated that the grey market prices (pre IPO market) is a good predictor of company value in the secondary market. On the other hand, we can view the grey market prices as indications of investor opinions about the aftermarket prices and thus what such investor willing to pay for the shares [2] .The grey market in Indonesia is usually called as pre IPO market. This period usually is not noticed by investors. But however, there is an indication that this period can give the prediction of the particular company stock performance in the long run based on the research held by Francesca Cornelli on the European Stock Market.

So at the bottom line, this research main purpose is to identify the correlations between pre IPO sentiment analyses towards the particular stock performance in the secondary market. A slightly different approach is adopted in this research, where the researcher will collect the social media sentiment of particular stock during the Pre-IPO market; time before the stocks is officially trading in the market.

2. Literature Review

Based on previous research by Chong Oh[1], sentiment from microblogging activity have strong predictive power for future market directions .In this research, they use sentiment analysis and predictive analytic method to prove this believe. The research show that the sentiment from microblogging activity contains valuable information. Another research done by Xue Zhang[6] show that people mood turn out as one possible way to predict the stock market movement. They collected the data from twitter and analyzed it by finding the emotions keyword in every tweet. In the end, the result show that there is correlation between people's mood and stock market movement. This study specifically done based on the research by Fransesca Corneli [3]. According to Francesca Cornelli[3], investor sentiment in Pre-IPO market, in this issue is the Pre-IPO market's price, can act as an indicator for predicting the stock performance in the secondary market. Based on this research, they found that the price in secondary market is highly correlated with the price in grey market(Pre-IPO market) and concluded that there is an important information in the grey market that will determine the stock performance in the secondary market.

Using the same period of Pre-IPO time, this research will indentified the correlations between information during that time and stock performance in secondary market. The main difference lying on indicator that used for predicting stock performance in secondary market. On the previous research the indicator being used is grey market stock price, whereas in this research investor sentiment during Pre-IPO market will be used as predictive indicator for future stock performance. In this study, LIWC sentiment analysis will be used to analyze the investor sentiment data during Pre-IPO market.The research related to this method done by Tausczik and Pennebaker [4], in this research they stated that languange style convey a subtle information about emotions and other's thought.

3. Methodology

The objective of this research is to understand the relationship between sentiment in Pre-IPO time and stock performance in secondary market. Below are the stages for doing this research:

3.1 Data Collecting

This study contains two set of data, first one is historical price data that collected from www.finance.yahoo.com for 4 months after the company listed. While the second data collected from twitter and stockforum (like Klubsaham, Stockbit, and BEI 5000). The author obtained 2796 posting from twitter and stockforum. This posting specifically obtained for certain time period which is 1 month before the company listed.

3.2 Filtering the Data Set

There are some basic ways to filter the social media data. First is by using exact phrase search, exact phrase search is used to find the exact word from the data.Second is by using distance search, distance search is used to find the word on which between that word there is a distance.Third is by using fuzzy search,fuzzy search is used to find the word with the extension letter on it.

3.3 Sentiment Analysis

Sentiment analysis is a computational study of opinions,sentiment and emotions that expressed in the text [5]. In this research, the researcher will use LIWC method; where the sentiment will be calculated based on frequencies of certain single word in every sentence. In doing this, first, the researcher need to determine the keyword and weight each of the keywords: +1 for positive word, 0 for neutral and -1 for negative words. Below are the keywords that used by the researcher in this study:

TABLE I: Positive Keywords

Positive (+1)				
dilirik	oversubscribe	menggiurkan	menjanjikan	banjir peminat
dikoleksi	mantap	andalan	borong	sukses
menarik	hebat	meroket	antri	menguntungkan
perfect	keren	terbang	hajar	full cash
naik	market leader	melesat	laris	berprospek baik
untung	upsize	melonjak	laku	jangka panjang
beli	AR kanan	prospektif	laba	cuan

TABLE II: Negative Keywords

Negative (-1)				
sengketa	batal	stagnan	avoid	payah
utang	turun	cut loss	gorengan	lesu
kemahalan	melorot	rugi	kemurahan	sepi
overvalued	nyungsep	gagal	melempem	undersubscribe
risky	ambles	hancur	terpuruk	pesimis
resiko	anjlok	bocuan	suram	AR kiri
jelek	merah	amsiong	loyo	turunkan

TABLE III: Neutral Keywords

Neutral (0)		
Prediksi	Perhatikan	tahan
Update	amati	waiting
Cermati	simak	ayo

For calculating the sentiment score, in this case, the researcher will use Semantria for Excel. Semantria will not only calculating the sentiment score but also grouping each posting into positive, neutral, and negative.

3.4 Trend Analysis

To understanding when the price and sentiment starting to incline or decline, trend analysis will be conducted. First the researcher need to classify the stocks into positive stocks and negative stocks based on 4 month stock return after it listing in the market. Second, the researcher will find the moment when the price start to incline for positive stock while for the negative stock it will be identified the time when it start to decline. Third, the researcher will do the same things to the stock's sentiment; find the time when the sentiment score start to incline for positive stock and start to decline for negative stock. There are two ways to understand when the price and sentiment start to inclining or declining. First is by plotting it to graph, second is by doing the linear regression in which the sentiment and price as independent variable and time as dependent variable. In this case, the linear regression conducted to obtain β value. Positive or inclining trend shown by positive β while the negative or declining trend shown by negative β. Last step is to compare price movement and sentiment movement of each stock during the certain time and identify whether there is correlations between sentiment and price.

4. Experimental Result

4.1 Price Movement Model

All stocks prices movement will be plotted into the graph. Figure 1 shows the movement of 20 positive stocks. In this issue, positive stocks means stocks that have positive return after 4 month listing in the market. From the graph, it can be identified that the price for positive stocks have an icline trend specifically in 30 days after it listing in the market. This result also strengthen by checking the β value,which depict the trend , for these positive stocks. From 20 stocks, there are 16 stocks that have positive β value that means that 75% of these stocks following this uptrend movement during 30 days after the stocks listing in the market. In this study, the researcher also identifying the trend for 60 days and 90 days after IPO. However the inclining trend(symbolized by β) is not significantly postive, therefore 30 days period chosen as the most suitable indicator for uptrend period.

Figure 2 show the movement of 18 negative stocks. From the graph, it can be identified that the price for negative stocks have downtrend movement specifically in 30 days after listing in the market. This result also strengthen by checking the β value. From 18 stocks, there are 14 stocks that have negative β value which means that 77.77% of these negative stocks have significantly downtrend movement during 30 days after IPO. The

researcher also identifying the trend for 60 days and 90 days after IPO. However the declining trend (β) is not significantly negative. Looking through this fact, 30 days period assumed as the most suitable indicator for downtrend period.

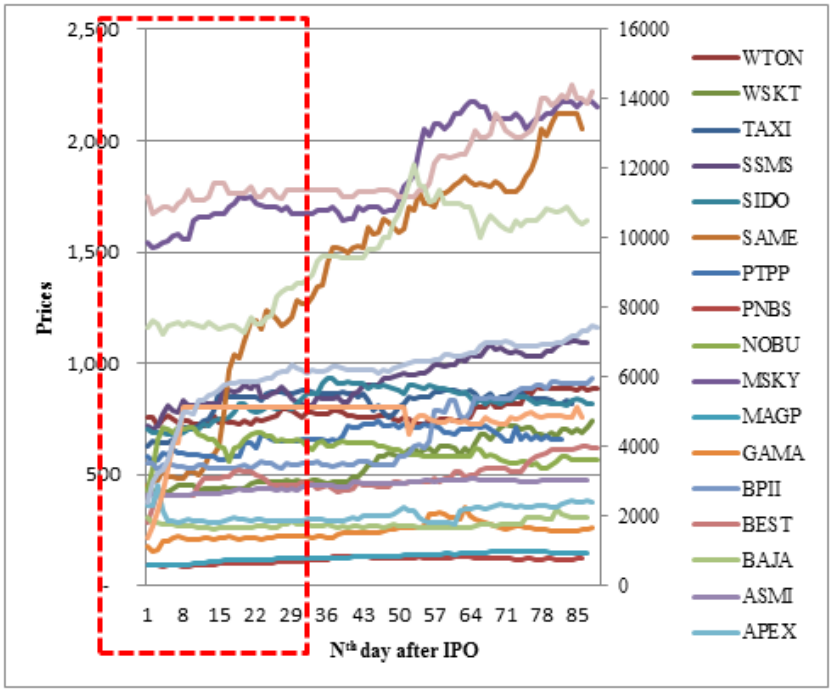

Fig. 1: Positive Stock Price Movement after IPO

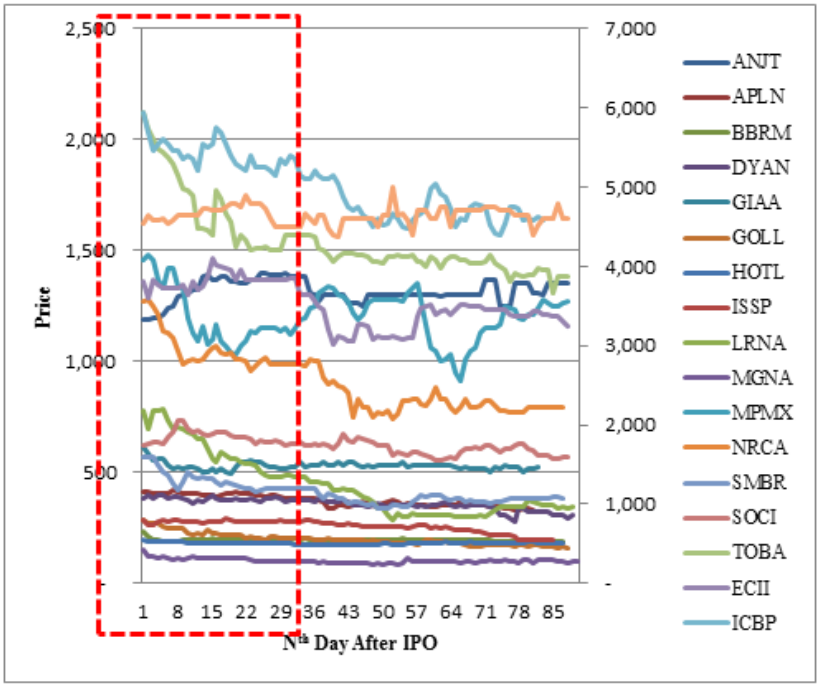

Fig. 2: Negative Stock Price Movement After IPO

4.2 Sentiment Movement Model

In order to understanding when is the crucial moment to start identifying the sentiment during Pre-IPO time, the researcher plot and integrate the sentiment score from each stocks into the graph based on its group. Figure 3 show the movement of sentiment score for positive stocks 8 weeks before the stocks listing in the market. From the graph, it can be identified overall these positive stocks have an uptrend movement,specifically during 1 weeks before IPO. This result also strengthen by checking the β value during this period. From 20 positive

stocks, there are 19 stocks that have an uptrend movement that means 95% of these positive stocks have significantly uptrend moment during 1 week before IPO.

Figure 4 show the movement of sentiment score for negative stocks 4 weeks before the stocks listing in the market. From the graph, it can be identified that overall these negative stocks have downtrend movement, specifically during 1 wee before IPO. This result also strengthen by checking the β value during this period. From 18 negative stocks, there are 17 stocks having downtrend movement which means 94% of these negative stocks have significantly downtrend moment during 1 week before IPO.

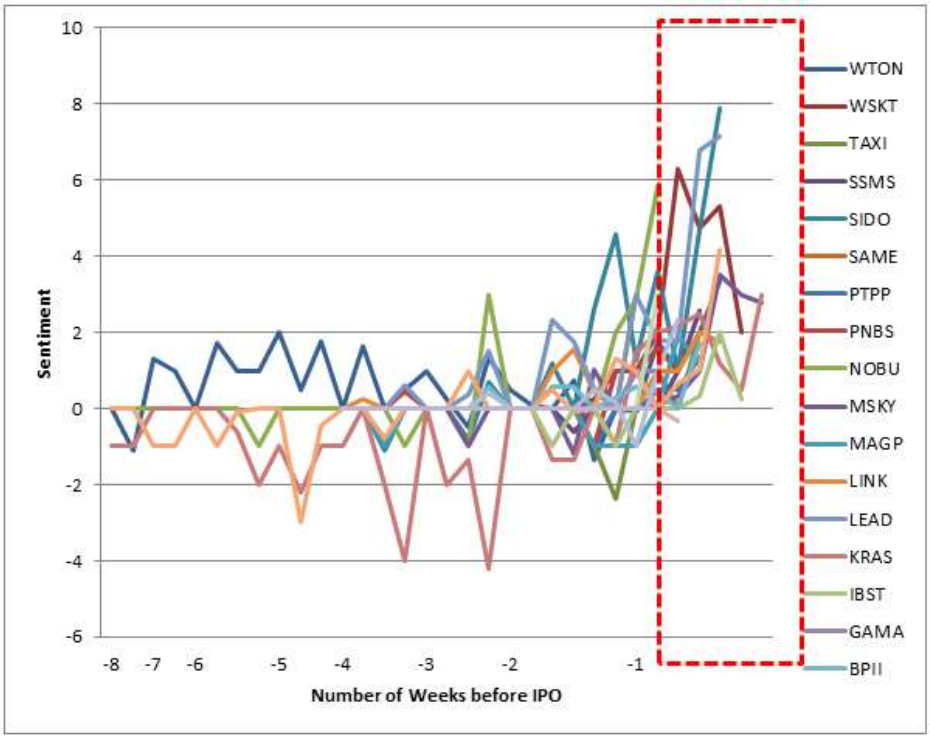

Fig. 3: Positive Stock Sentiment Score Score Before IPO

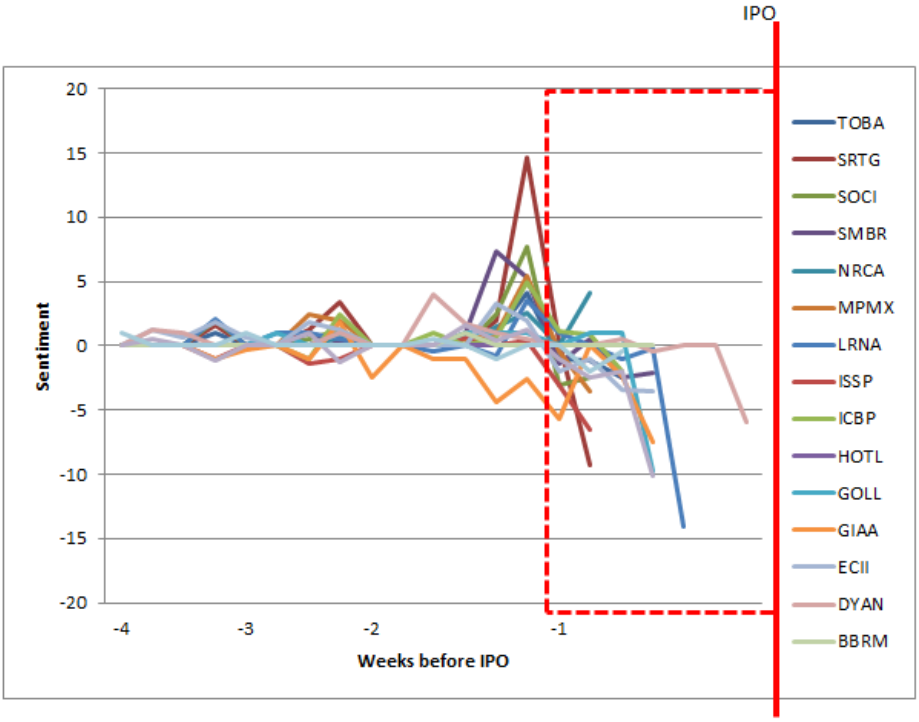

Fig. 4: Negative Stock Sentiment Score before IPO

4.3 Comparisson between sentiment trend and price trend movement

After understanding the price and sentiment movement trend for each stock categories(positive and negative), next step of the research is to identify the correlations between sentiment and price movement. Table 4 below showing the compilation result for positive stocks. Column 2 present name of stocks which exist in positive group. Column 3 present the compilation of price movement trend that already discussed in figure 1, as stated before there are 16 positive stocks which following the increasing price trend(symbolized by letter "I" in the table) during 1 month after IPO, while the remaining 4 stocks not following this trend(symbolized by letter "N" in the table). Column 4 present the compilation of sentiment movement trend which has also been discussed in figure 3 , as stated before there's 19 positive stocks which following the increasing sentiment trend(symbolized by letter "I") during 1 week before IPO, while one stock not following this trend (symbolized by letter "N"). Column 5 identifying the correlation between price and sentiment,there's assume to have correlation if and only if the previous 2 column have "I" value and in this column the exist correlation symbolized by letter "Y" and no correlation symbolized by letter "N". From column 5 , it can be concluded that there are 16 stocks (out of 20 stocks) that have correlation between sentiment and price trend movement.This result show that for positive stocks, the sentiment trend can indicate the price movement trend with 80% accuracy.

TABLE IV: Compilation Result for Positive Stocks

No	Name of Stock	Price	Sentiment	Correlation between Price and Sentiment
1	**WTON**	N	I	N
2	**WSKT**	I	I	Y
3	**TAXI**	I	I	Y
4	**SSMS**	I	I	Y
5	**SIDO**	I	I	Y
6	**SAME**	I	I	Y
7	**PTPP**	I	I	Y
8	**PNBS**	I	I	Y
9	**NOBU**	I	I	Y
10	**MSKY**	I	I	Y
11	**MAGP**	I	I	Y
12	**LINK**	I	I	Y
13	**LEAD**	I	I	Y
14	**IBST**	I	I	Y
15	**GAMA**	I	I	Y
16	**BPII**	N	I	N
17	**BIRD**	N	I	N
18	**BEST**	I	I	Y
19	**ASMI**	I	I	Y
20	**APEX**	N	N	N

TABLE V: Compilation Result for Negative Stocks

No	Name of Stock	Price	Sentiment	Correlation between Price and Sentiment
1	ANJT	N	D	N
2	APLN	D	D	Y
3	BBRM	D	D	Y
4	DYAN	D	D	Y
5	ECII	N	D	N
6	GIAA	D	D	Y
7	GOLL	D	D	Y
8	HOTL	D	N	N
9	ICBP	D	D	Y
10	ISSP	D	D	Y
11	LRNA	D	D	Y
12	MPMX	D	D	Y
13	NRCA	D	D	Y
14	SMBR	D	D	Y
15	SOCI	N	D	N
16	SRTG	N	D	N
17	TOBA	D	D	Y
18	BAJA	D	D	Y

Table 5 above showing the compilation result for negative stocks. Column 2 present name of stocks which exist in negative group. Colum 3 present the compilation movement trend that already discussed in figure 2, as stated before there are 14 negative stocks which following decreasing trend during one month after IPO (symbolized by letter "D" in the table) while the remaining 4 stocks which not following this trend(symbolized by letter "N" in the table). Column 4 present the compilation of sentiment movement trend that already

discussed in figure 4 , as stated before there's 17 negative stocks which following this decreasing sentiment trend(symbolized by by letter "D") during one week after IPO while one stocks not following this trend(symbolized by letter "N"). Column 5 identifying the correlations between price and sentiment. Same as in positive stocks, there's assumed to be a correlation between price and sentiment if and only if the previous 2 column have "D" value; the exist correlations symbolized by letter "Y" and no correlation symbolized by letter "N". From column 5, it can be concluded that there are 13 stocks (out of 18 stocks) that have correlation between sentiment and price trend movement. This result show that for negative stocks, the sentiment trend can indicate the price movement trend with 72% accuracy.

5. Conclusion

The main purpose of this research is to identify the correlations between sentiment in Pre-IPO time and stock price in secondary market both for positive stock and negative stock. The result is for positive stock, there is 80 % correlations between sentiment in Pre-IPO market and stock price in secondary market while for negative stock the correlations between sentiment and stock price is a bit lower around 72.22%.

The researcher also identifying the crucial moment to start observing the sentiment during Pre-IPO time. From the graph, it can be identified that the sentiment for positive stocks have an incline trend specifically during one week before IPO, while for negative stocks, the sentiment have an decline trend also during one week before the stocks listing in the market.

6. Acknowledgement

I would like to thank Mrs. Eneng for her advice throughout this research. I also would like to thank to my parents and all of my friends for their unlimited support, therefore I could finish this project.

7. References

[1] Chong Oh, O. R. (2011). INVESTIGATING PREDICTIVE POWER OF STOCK MICRO BLOG SENTIMENT IN FORECASTING FUTURE STOCK PRICE DIRECTIONAL MOVEMENT. Knowledge Management and Business Intelligence .

[2] Dorn, D. (2002). Does sentiment drive the retail demand for IPO? Colombia University .

[3] Francesca Cornelli, D. G. (June 2006). Investor Sentiment and Pre-IPO Markets. The Journal of Finance , 1187–1216. http://dx.doi.org/10.1111/j.1540-6261.2006.00870.x

[4] Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text. Journal of Language and Social Psychology , 24-54. http://dx.doi.org/10.1177/0261927X09351676

[5] Walaa Medhat, A. H. (2014). Sentiment Analysis Algorithms and Applications: A Survey. Ain Shams Engineering Journal , 1093-1094. http://dx.doi.org/10.1016/j.asej.2014.04.011

[6] Xue Zhang, H. F. (2011). Predicting Stock Market Indicators Through Twitter "I hope it is not as bad as I fear". Procedia-Social and Behavioural Science , 55-62. http://dx.doi.org/10.1016/j.sbspro.2011.10.562