

# Social Media Analysis of BPS Data availability in Economics using Decision Tree Method

Pramana Yhoga Chandra Kusuma  
Department of Multimedia and Network  
Engineering, Institute of Technology  
Sepuluh Nopember,  
Surabaya, Indonesia  
pramana14@mhs.ee.its.ac.id

Surya Sumpeno  
Department of Multimedia and Network  
Engineering, Institute of Technology  
Sepuluh Nopember,  
Surabaya, Indonesia  
surya@ee.its.ac.id

Adhi Dharma Wibawa  
Department of Multimedia and Network  
Engineering, Institute of Technology  
Sepuluh Nopember,  
Surabaya, Indonesia  
adhiosa@te.its.ac.id

**Abstract** - Badan Pusat Statistik (BPS) is one of the government institution with the duty to provide accurate and reliable data for the government and society uses. Currently BPS does not have basic tool that can be used to evaluate usage of their data by people or government. The aim of this study is to analyze the usage of BPS data, especially in economics by people or society via social media “Twitter”, mapping them and comparing the result with the availability of the current data in BPS so that some suggestions for upgrading and advancing BPS in providing data to the public can be made. Some comments, especially about economics, from the public via Twitter were mined and classified by Decision tree method. Four categories data were produced: (1) Not in the field of economic, (2) in the field but not using BPS data, (3) in the field that may be using BPS data and (4) in the field that using BPS data. Tweets harvesting were done during two months period January 2016 til February 2016 by using Twitter Archiver. About 77.854 tweets that indicates to economic topic were collected. 2.360 data were used for training data. The result showed that 49.425 tweets were classified as economic topic. The majority of class was obtained from the tweets of economics but not using BPS data. 12 major keywords were obtained during this study. These keywords were generated using unigram, bigram, and trigram. In conclusion, we found that 3 primary keywords about economics were needed to be concerned by BPS: The economic policy package, Palm Oil, and Creative Economy.

**Keywords:** BPS, Badan Pusat Statistik, Social Media Mapping, public opinion based on Twitter, Social media analysis

## I. INTRODUCTION

BPS is one of the government institutions which its duty is to provide statistical data. BPS has various kind of products data such as inflation, economic growth, export, import, poverty, etc. As a government institution, BPS is active in the implementation of bureaucratic reform gathering all public important data for public consumption. One of its implementation is by arranging Quick Wins program. The program aims is to ensure that BPS products can be reached and used by public as wide as possible. One important rule of BPS in society is giving data services such as in a public library, email services, website, and social media [5].

Nowadays, social media has been growing rapidly in Indonesia. This could be trigger by mobile devices technology and cellular network development and public penetration [13, p. iv]. According to Marius [13, p. 20] internet users in Indonesia are 88,1 million users in 2014, these reach 34,9% of the population. Marius [13, p. 31] also showed that 87,4% of those internet users are using social media. According to those facts, media social analysis is important to describe what are the public think and what are their opinion regarding to many issues in the society.

This study is analyzing and mapping the social media “Twitter” user’s comments on the availability of BPS data especially in economics issues that were used by the public. The main purpose of this study is measuring the BPS data availability and suggesting some aspects that can be used for extending and advancing the BPS products so that its performance in providing data to the public can be optimized.

## II. METHODS

Research methodology of this study in general is described in the flowchart in figure 1:

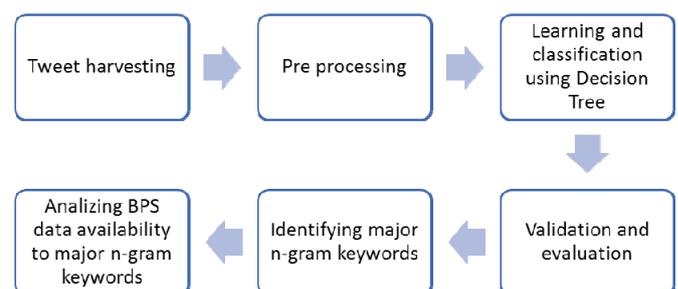


Fig. 1. Research methodology.

Tweets data harvesting was done during two month period: January to February 2016 using Twitter Archiver tool. For gathering appropriate tweets, we set Twitter Archiver to

collect in Indonesian language only. Ekonomi was used for the initial keyword so tweets that contain a word “ekonomi” were selected and stored. The process resulted 77.854 data tweets. The examples of Twitter Archiver result is shown in table I.

TABLE I. EXAMPLES OF HARVESTED TWEETS USING TWEET ARCHIVER

Date	Screen Name	Full Name	Tweet Text
01/11/2016	@amin khariman	Ben	Ffs. Stop with "Pakar IT pakar ekonomi" jokes δÿ™„δÿ”«
01/11/2016	@Pajak_News	Pajak News	https://t.co/fccUGCAZNw: Penerimaan Pajak 2015 Bertambah Rp 5,24 Triliun https://t.co/I9h3iKjqG #BeritaPajak
01/11/2016	@Bari katulHikmah	Barikat ul Hikmah	Leo Messi naik pesawat kelas ekonomi δÿ± https://t.co/I6HGz6w3Ro
01/11/2016	@fajar_milenia	fajar purnama	R&D diarahkan untuk mendapatkan benih unggul kelapa sawit dari sisi produktivitas maupun penggunaan bahan pendukung https://t.co/IDd443G4GN
01/11/2016	@Hakanochi	Mario suarez	RT @Beritasatu: Dorong Pertumbuhan, RI Bisa Andalkan Ekspor Barang Konsumsi https://t.co/P5PXRmxioE
01/11/2016	@GeniusDigger	Hahnz River Fandub l	Facebook Kembangkan Aplikasi Messenger untuk Mac?: Facebook Kembangkan Aplikasi Messenger untuk Mac? https://t.co/uWEelDa6Jl

A number of 2.360 tweets were randomly classified into training data. Two persons were involved in classifying the data in order to avoid subjectivity of the result. If there is some ambiguity in the tweet classification, two BPS officers must decide which class of these tweets belongs to. Table II showed the example of tweets which have been manually classified. The tweets were cleaned from nonusable content. Some character/strings were removed such as texts that mention (@), retweet (RT), hastag (#), numbers, web links, and other unused character. Case folding was also used for getting uniformed characters, since upper and lower case of character would deliver different words. Cleaning process was done in Microsoft Excel by applying regular expression code.

TABLE II. EXAMPLES OF MANUAL CLASSIFICATION RESULT

Tweet Text	Class
penyusunan master plan pembangunan ekonomi kabupaten gunungkidul	3
daya saing ekspor minyak sawit indonesia merosot	4
harga minyak usd barel	2
gara gara kegawatan ekonomi ayam goreng inti angin dijual pasaran	1
jelang jeda rupiah poin level rupiah	2
pertumbuhan ekonomi persen potensi	4

After cleaning process, some words that don't have significant value were removed which called feature selection process. Stopword removal and stemming were used in feature selection process. Stopword was done using Microsoft Excel

and VBA code. Stopword removal works by comparing tweets in the first column with stopword list in the second column. List of stopword used here is from F. Z Tala [6]. The list contains 758 stopwords in the Indonesian language. The examples of stopword removal result is in table III.

TABLE III. EXAMPLES OF STOPWORD REMOVAL PROCESS

Tweet Text	Result
smg cepat kau hancur rusia jahat krisis ekonomi warga rusia antri makanan gratis	smg cepat hancur rusia jahat krisis ekonomi warga rusia antri makanan gratis
ini kronologis dugaan pemalsuan flight approval airfast	kronologis dugaan pemalsuan flight approval airfast
tuntutan jihad msykt negara jihad bangunkan ekonomi politik dan alam sekitar yang sihat keadilan kebebasan terpimpin	tuntutan jihad msykt negara jihad bangunkan ekonomi politik alam sihat keadilan kebebasan terpimpin
jangan bikin hastag tentang prayjkt semacamnya ntar jadi obrolan dunia investor berkurang rupiah bisa ribu eko	bikin hastag prayjkt semacamnya ntar obrolan dunia investor berkurang rupiah ribu eko

Stopword-free tweets were stemmed in order to find the basic word. This process uses Sastrawi library for PHP. This library was installed on localhost server and some codes were created to call it. The result of stemming was stored in MySQL database. Codes for stemming is mentioned in figure 2.

```

$query = "SELECT cleaned_stop_word, Ran_2 FROM `tweet77854`";
$result_query = $mysqliconn->query($query);

$stmtmerFactory = new \Sastrawi\Stemmer\StemmerFactory();
$stmtmer = $stmtmerFactory->createStemmer();

if ($result_query->num_rows > 0) {
    while ($row = $result_query->fetch_assoc()) {
        echo $row['Ran_2'].". ". $row['cleaned_stop_word']." =====> ";

        $sentence = $row['cleaned_stop_word'];
        $output = $stmtmer->stem($sentence);

        echo $output . "====>";

        $query_stemmed = "UPDATE tweet77854 SET stemmed = '". $output.'" WHERE Ran_2 = $row['Ran_2']";
        $result_query_stemmed = $mysqliconn->query($query_stemmed);
    }
}

```

Fig. 2. Code for Stemming Tweets

From total tweets, 2.360 tweets were taken to be used for building a model using Decision Tree method in Weka free application. StringToWordVector filter was applied with TF-IDF and minTermFreq 3. Training process showed 70,55% accuracy of the model. Detail summary of Confusion Matrix is shown in figure 3.

```

=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      1665          70.5508 %
Incorrectly Classified Instances    695          29.4492 %
Kappa statistic                    0.5809
Mean absolute error                 0.1969
Root mean squared error             0.3351
Relative absolute error             55.6124 %
Root relative squared error         79.6286 %
Total Number of Instances          2360

=== Confusion Matrix ===
 a  b  c  d  <-- classified as
530 231 38 24 | a = 1
130 660 39  8 | b = 2
 59  77 151 28 | c = 3
 13  16  32 324 | d = 4

```

Fig. 3. Confusion Matrix of Decision Tree

Decision tree model that resulted by Weka is not easily to be read. For increasing readability of tree model FreeMind program was used. This program can be used to visualize tree model. Visualization of tree model from FreeMind was showed by figure 4.

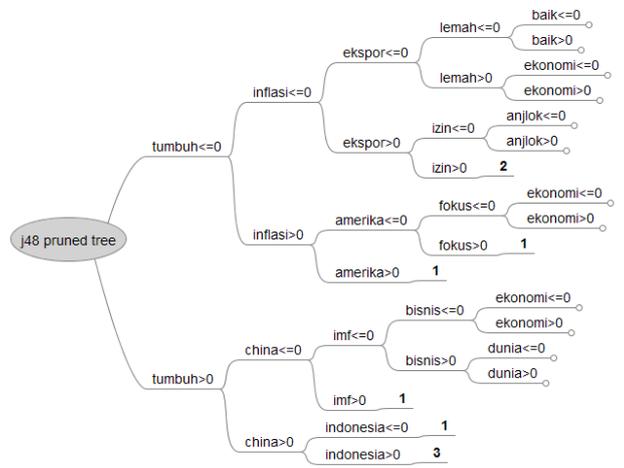


Fig. 4. Visualization of Tree Model

### III. RESULT

The result of the classification is shown in table IV.

TABLE IV. TWEETS CLASSIFICATION RESULT

Class Number	Class	Number of tweets	Percentage (without class 1)
1	The topic is not about economic field	28.429	
2	The topic of economic field that is not using BPS data	41.822	84,62
3	The topic of economy that maybe using BPS data	3.267	6,61
4	The topic of economy that using BPS data	4.336	8,77
Total		77.854	100,00

From class number 2 that had 84,62% Tweets, we identified major keywords from its frequency of occurrence. We are seeking candidates for the primary keywords. These keywords will be proposed to BPS for enhancing their

dissemination of data which are related to that. To get this primary keyword candidate, we count the frequency of words in class number 2. We count unigram, bigram, and trigram that appear most frequently.

TABLE V. UNIGRAM FREQUENCY IN THE CLASS NUMBER 2

Unigram	Frequency	Unigram	Frequency
Ekonomi	9.841	Jokowi	2.278
Indonesia	7.690	Sawit	2.093
Rupiah	3.912	Usaha	2.085
Perintah	3.474	Bangun	1.767
Harga	2.857	Bank	1.662
Industri	2.315	Paket	1.622

Unigram frequency showed some top keywords such as ekonomi, indonesia, rupiah, perintah, harga, etc (see Table V). According to the result, it shows 12 keywords that most appeared on the tweets about economy which were not related to BPS data. The major unigram keywords are too general to be connected with BPS data. For example, ekonomi (economy), Indonesia, rupiah, perintah (instruction), harga (price), and industri (industry) are the topic that too general, so we found it difficult to determine their relationship with the BPS data.

TABLE VI. BIGRAM FREQUENCY IN THE CLASS NUMBER 2

Bigram	Frequency	Bigram	Frequency
Ekonomi indonesia	1.218	Jusuf kalla	643
Kereta cepat	1.036	Ekonomi kreatif	624
Paket bijak	928	Paket ekonomi	616
Bijak ekonomi	900	Ekonomi digital	600
Bank indonesia	895	Harga minyak	545
Triliun rupiah	814	Ekonomi asean	506
Kelapa sawit	692	Masyarakat ekonomi	503

TABLE VII. TRIGRAM FREQUENCY IN THE CLASS NUMBER 2

Trigram	Frequency	Trigram	Frequency
Paket bijak ekonomi	837	Bijak ekonomi jilid	177
Masyarakat ekonomi	422	Rupiah dolar	167

asean		amerika	
Proyek kereta cepat	387	Kereta cepat jakarta	166
Bank indonesia rate	294	Industri kelapa sawit	164
Cepat jakarta bandung	247	Indeks harga saham	160
Badan ekonomi kreatif	243	Nilai tukar rupiah	160
Ekonomi digital indonesia	183	Paket ekonomi jokowi	160

Bigram and trigram frequency (see table VI and VII) show a more specific major keywords to be identified. Those tables each contains 14 major keywords. We combine those bigram and trigram to find relationship between them to minimize the number of major keyword candidates. We do it by identifying same unigram between bigram and trigram. If a keyword has same unigram with other keywords unigram, we find their meaning. If they have similar meaning then we consider them to equal. Table VIII shows 12 major n-gram keywords that generated from combination of 28 total keywords. We use 12 of these n-gram keywords as the main keyword candidates from the class number 2.

TABLE VIII. RELATIONSHIP BETWEEN BIGRAM AND TRIGRAM

N-gram keywords	Relation to another n-gram
Ekonomi indonesia	Has no relation to other n-gram keywords
Kereta cepat	Has relation to : Proyek kereta cepat, kereta cepat jakarta, cepat jakarta bandung
Paket bijak	Has relation to : bijak ekonomi, paket ekonomi paket bijak ekonomi, bijak ekonomi jilid, paket ekonomi jokowi
Bank indonesia	Has relation to : Bank indonesia rate
Triliun rupiah	Has relation to : Rupiah dolar amerika, nilai tukar rupiah
Kelapa sawit	Has relation to : Industri kelapa sawit
Jusuf kalla	Has no relation to other keywords
Ekonomi kreatif	Has relation to : Badan ekonomi kreatif
Ekonomi	Has relation to :

digital	ekonomi digital indonesia
Harga minyak	Has no relation to other keywords
Ekonomi asean	Has relation to : masyarakat ekonomi masyarakat ekonomi asean
Indeks harga saham	Has no relation to other keywords

Those main keyword candidates need to be connected with BPS data. Table IX shows types of data produced by BPS. We get those types of data from the official website of BPS (Statistics Indonesia).

TABLE IX. BPS DATA TYPE IN ECONOMIC AND TRADE TOPICS

Type of Data	Type of Data
Business and consumer tendency index	Input output
Communication	Labour wages
Construction	Large and medium manufacturing
Consumer Prices Indices	Micro and small manufacturing
Energy	Producer price indices
Farmer terms of trade	Public finance
Flow of funds	Retail
Foreign trade	Small-scale and micro establishment
Gross domestic product (expenditure)	Social accounting matrix
Gross domestic product (industrial origin)	Tourism
Gross regional domestic product (expenditure)	Transportation
Gross regional domestic product (industrial origin)	Wholesale price indices

TABLE X. KEYWORDS STATUS

Main Keyword Candidates	Too common issues	Temporary issues	Can be supported by BPS data	Supported by other data provider
Ekonomi indonesia	Yes			
Kereta cepat	No	Yes		
Paket bijak ekonomi	No	No	Yes	
Bank indonesia	No	Yes		
Triliun rupiah	Yes			
Kelapa sawit	No	No	Yes	
Jusuf kalla	Yes			
Ekonomi	No	No	Yes	

Ekonomi digital	No	Yes		
Harga minyak	No	Yes		
Masyarakat ekonomi asean	No	Yes		
Indeks harga saham	No	No	No	Yes

TABLE XI. RELATION BETWEEN N-GRAM KEYWORDS AND BPS DATA

N-gram Keyword	Relation to BPS data
Paket bijak ekonomi	Can be related to : Export, import, inflation, gross domestic product
Kelapa sawit	Can be related to : Farmers exchange rate, export, industries, palm oil statistic
Ekonomi kreatif	Can be related to : Communication
Indeks harga saham	Not supported by any BPS data

#### IV. DISCUSSION AND CONCLUSION

Table IV shows that the BPS data were used in 8,77% of tweets on the economic field (class number 4). Currently, we do not have a benchmark to measure this percentage. We cannot determine whether it's high or low. Class number 3 contains tweets that may be related to BPS data. This class may contain data from BPS or other data providers. This class consists of 6,61% of economic tweets. We do not discuss further those classes since we do not have a benchmark.

Table IV also shows that 84,62% of tweets in economic topic were not mentioning BPS as the source. We speculate that this indicates three possibilities. The first is tweets do not need data. Twitter users can say something about economic topic without a data as basis. For example they can say opinion, think, or idea about government's policy without data. Second is, Twitter users didn't know if BPS had data that can support their opinion/comment. This is due to the lack of data dissemination that has been done by BPS. Even though BPS has done some data disseminations using some ways such as printed publication in library, soft copy in email service, downloadable content in website, and also social media sharing.

Third is, Twitter users didn't want to use BPS data. This could be caused by the fact that BPS data are not relevant to the topic they discuss. This irrelevancy could be triggered by the timing of BPS data release. For example BPS' foreign exchange data is released once a month, however most foreign exchange businessmen need this data every day. Besides that, they can get foreign exchange data from other source such as foreign exchange market.

We are using table VIII until XI to analyze the relationship between main keyword candidates and BPS data. Based on 12 main keywords candidates in table VIII, we indicate that ekonomi indonesia, triliun rupiah and jusuf kalla keywords are too common to be connected with BPS data. These keywords can be related to many topics of BPS data. We exclude it out of main keywords candidates. Kereta cepat keyword can be interpreted as the fast train. This keyword became popular because there was controversial news about during tweet harvesting time [15]. We call it a temporary effect of news media. It won't be a long-lasting topic so we exclude it from main keywords candidates.

Paket bijak ekonomi keyword can be interpreted as economic policy package. This keyword should be supported by export, import, economic growth, and labor data from BPS. However, tweets do not mention any of those BPS data. We found that actually those data were used to compose economic policy package. However, in dissemination process those data were not mentioned directly. This keyword will be more long-lasting because the president of Republic Indonesia always updates this policy [2]-[4]. We included this keyword as main candidates because the role of the BPS data was needed.

Bank Indonesia can be interpreted as Indonesian Central Bank. It is one of the government agencies that have the purpose of achieving stability in the rupiah. To achieve its objectives Indonesian banks need inflation data and the exchange rate. Both of data are provided by BPS every month. However, In January and February 2016 Indonesian bank rate declined in a row from 7,50% to 7,00% [1]. We indicated that the decline of Indonesian bank rate is the cause of the popularity of this keyword. So we exclude it out of the main keywords candidates.

Kelapa sawit keyword can be interpreted as palm oil. This keyword related to exchange rate of the farmer, export, industry, and palm oil statistic data from BPS. This is one of major keyword that represents agricultural products. There is no special case which related to kelapa sawit during tweets harvesting period. We include this keyword as main keyword candidates.

Ekonomi kreatif keyword can be interpreted as the creative economy. This keyword is a new topic in indonesia. This topic is introduced by president of Republic Indonesia since 2014. It is implemented by creating Bureau of Creative Economy under Minister of Tourism structure. The main pillar of creative economy is information and creativity. BPS can support the development of a creative economy by providing data on the communication field.

Ekonomi digital keyword can be interpreted as the digital economy. In February 2016 the president of the Republic of Indonesia visited Silicon Valley in United States. He came to talk about e-commerce with the chairman of some world famous IT company i.e. Google, Facebook, and Twitter [12]. We indicate this keyword's popularity can be caused by those visit. We have to follow this keyword in a longer time for ensuring whether this popularity is permanent or temporary. So we exclude it out of the main keyword candidates right

now. We also presume that digital economy keyword may have a relationship with the creative economy keyword.

Harga minyak keyword can be interpreted as oil price. We indicated that this keyword became popular because of oil price trend during tweet harvesting. At January until February 2016 there was significant decrease in oil price in the world. This decrease reaches US\$ 26 per barrel which is the lowest price in 13 years [8]. We assume that this popularity is temporary effect that depends on oil price. So we exclude it from main keyword candidates.

Masyarakat ekonomi asean keyword can be interpreted as ASEAN Economic Community (AEC). This topic can be supported by labor, market and industry data from BPS. AEC started in January 2016, this is same period of tweets harvesting started [17]. For now we exclude it out of the main keyword candidates.

Indeks harga saham can be interpreted as stock price index. There are no special events at price index during tweets harvesting period [14]. BPS don't have data regarding stock price index. However, BPS don't need to enter to this topic since data of stock price index have been provided well by Indonesia Stock Exchange [7]. So we exclude it out of the main keywords candidates.

From above analysis of main keyword candidates, we categorize them into 2 categories. First category is unused keyword candidate that contains (1) ekonomi indonesia, (2) kereta cepat, (3) bank indonesia, (4) triliun rupiah, (5) jufus kalla, (6) ekonomi digital, (7) harga minyak, (8) masyarakat ekonomi asean, and (9) indeks harga saham keywords. We do not use these candidates because they are too common issues, only temporary issues, and already provided well by other data provider.

Second category is main keywords that need to be concerned by BPS for increasing its data availability coverage and quality. This category contains : (1) economic policy package, (2) palm oil, and (3) creative ekonomi. In economic policy package BPS has provided data to support it such as export, import, economic growth, and labour. BPS don't have to worry about this topic because actually they have provided data for arrangement of economic policy package. In palm oil topic, three possibilities mentioned at the beginning of discussion can be used. So, BPS need to evaluate dissemination, period, and variety of data. In creative economy topic, BPS don't have any specific data that can be used to support this topic. The only type of data that can be used to support this main keyword is communication field data. So BPS need more evaluation to support this main topic, especially in variety of data.

## V. REFERENCES

- [1] Bank of Indonesia. (2016, June 5). *BI Rate (Based on Decision of Board Meeting)* [online]. Available: <http://www.bi.go.id/en/moneter/bi-rate/data/Default.aspx>.
- [2] Bappenas. (2016, June 4). *Paket Ekonomi IX: Pemerataan Infrastruktur Ketenagalistrikan dan Stabilisasi Harga Daging Hingga ke Desa* [online]. Available: <http://www.bappenas.go.id/berita-dan-siaran-pers/paket-ekonomi-ix-pemerataan-infrastruktur-ketenagalistrikan-dan-stabilisasi-harga-daging-hingga-ke-des/>
- [3] Bappenas. (2016, June 4). *Paket Kebijakan Ekonomi X: Memperlonggar Investasi Sekaligus Meningkatkan Perlindungan Bagi UMKM* [online]. Available: <http://www.bappenas.go.id/berita-dan-siaran-pers/paket-kebijakan-ekonomi-x-memperlonggar-investasi-sekaligus-meningkatkan-perlindungan-bagi-umkm/>
- [4] Bappenas. (2016, June 5). *Paket Kebijakan Ekonomi VIII: Satu Peta Nasional, Kilang Minyak dan Bea Masuk Suku Cadang Pesawat* [online]. Available: <http://www.bappenas.go.id/berita-dan-siaran-pers/paket-kebijakan-ekonomi-viii-satu-peta-nasional-kilang-minyak-dan-bea-masuk-suku-cadang-pesawat/>
- [5] BPS, "Road Map Reformasi Birokrasi Badan Pusat Statistik," Central Bureau of Statistics, p. 74, 2011.
- [6] F.Z. Tala, "A Study of Stemming Effect on Information Retrieval in Bahasa Indonesia," M.S. Thesis, University of Amsterdam, 2003.
- [7] Indonesian Stock Exchange (2016, June 5). *Stock Trading Snapshot* [online]. Available: <http://www.idx.co.id/en-us/home/publication/statistic.aspx>.
- [8] J.C. Liputra. (2016, June 3). *WTI Crude Oil Nymex* [online]. Available: <http://hargaminyak.net/trend-grafik/WTI-Crude-Oil-Nymex/>
- [9] M. Khoiron, "Analisa Microblogging untuk Menentukan Prioritas Kebijakan Publik Berdasarkan Opini Masyarakat Menggunakan Algoritma Naive Bayes dan Analytical Hierarchy Process," M.S. Thesis, Institute of Technology Sepuluh Nopember, 2016.
- [10] Nazief, and Adriani, "Confix Stripping : Approach to Stemming Algorithm for Bahasa Indonesia," Technical Report, Faculty of Computer Science, University of Indonesia, Depok, 1996.
- [11] N.W. Saraswati, "Text Mining dengan Metode Naive Bayes Classifier dan Support Vector Machines untuk Sentiment Analysis," M.S. Thesis, University of Udayana, 2011.
- [12] Public Relation of Setkab of Republic Indonesia. (2016, June 3). *Kunjungan Presiden Jokowi ke Silicon Valley, Menkominfo: Kita Akan Kembangkan Ekonomi Berbasis Digital* [online]. Available: <http://setkab.go.id/kunjungan-presiden-jokowi-ke-silicon-valley-menkominfo-kita-akan-kembangkan-ekonomi-berbasis-digital/>
- [13] P. Marius, S. Anggoro, "Profil Pengguna Internet Indonesia 2014," APJII, Jakarta, 2015.
- [14] R and D Division of Indonesia Stock Exchange. (2016, June 4). *IDX Statistics 1st Quarter 2016* [online]. Available: [http://www.idx.co.id/Portals/0/StaticData/Publication/Statistic/Quarterly/20160509\\_IDX\\_1st\\_Quarter\\_2016.pdf](http://www.idx.co.id/Portals/0/StaticData/Publication/Statistic/Quarterly/20160509_IDX_1st_Quarter_2016.pdf)
- [15] R. Hardjanti. (2016, June 5). *Kereta Cepat Jakarta Bandung, Gebrakan yang Menuai Polemik* [online]. Available: <http://economy.okezone.com/read/2016/01/21/320/1293670/kereta-cepat-jakarta-bandung-gebrakan-yang-menuai-polemik>
- [16] Statistics Indonesia. (2016, June 5). *Statistics Indonesia* [online]. Available: <http://bps.go.id>
- [17] The ASEAN Secretariat. (2016, June 5). *ASEAN Economic Community Blueprint 2025* [online]. Available: <http://www.asean.org/storage/images/2015/November/aec-page/AEC-Blueprint-2025-FINAL.pdf>
- [18] T.K. Das, D.P. Acharjya, and M.R. Patra, "Opinion Mining about a Product by Analyzing Public Tweets in Twitter" International Conference on Computer Communication and Informatics (ICCCI), pp. 1-4, 2014.