A Predictive Analytic on Data Online Digital News using Systematic Literature Review

Razief Perucha Fauzie Afdih$^{1, 2*}$ and Zainal A. Hasibuan$^3$

$^1$Faculty of Computer Science, Universitas Indonesia  
$^2$Faculty of Mathematics and Natural Science, Universitas Syiah Kuala  
$^3$Faculty of Computer Science, Universitas Dian Nuswantoro

Email: *razief.perucha@ui.ac.id

**Abstract.** This study intends to provide an overview of the use of online digital news as a text dataset for future data analysis. Systematic literature review used as the method for collecting and analyzing the information from previous studies that used online digital news as a dataset. The result showed that the use of online digital news as a dataset can be implemented for classification and clustering. Furthermore, online digital news dataset is used to predict stock price and product price movement, to predict the approval rate for election processes, to analyze the diseases epidemiology, to detect event, classification of fake news, popularity of news in social media, and other NLP tasks. By comparing online digital news dataset versus social media dataset, it can be used to detect fake news, news popularity prediction, stock price prediction, topic detection, sentiment analysis, event detection and prediction, spam detection, trending topic prediction, and other tasks. Online digital news as a text dataset has powerful performance to be implemented in various fields such as economics, politics, health, language, and so forth.

**Keywords:** online digital news, dataset, algorithm, prediction, classification, clustering

1. **Introduction**

Text is one of the data formats that are widely used in Artificial Intelligent (AI) research. In addition to text, there are also other media formats used in AI-based research, namely audio, images, and video. The amount of data generated in text format increases with the growth of text-based applications, for example; news portals and social media portals, or micro blogs (Facebook, Twitter, etc.) An increase in the amount of data generated by news portals and social media, making it one of the largest text-based data producers known as big data.

With current technological advances, the data contained in the news portal and micro blog can be extracted for further study. Text-based data can be used to classify a topic or article whether it is fake or not [1]–[4]. The possibility of a topic being trending or popular can also be predicted using text-based data [5]–[7]. Furthermore, with the development of research in the field of NLP supported by data sourced from news articles or social media datasets, studies related to “next word prediction”, “dialogue agents”, “semantic meaning”, “text summarization”, “topic modeling” and another subfields of NLP research are very well developed [8], [9].
In this research, we conducted a literature study to explore the use of data sourced from online digital news. The contribution of this research is to give an overview of the implementation of using data source from online digital news as a dataset. In addition, we also make comparisons with data sourced from social media such as twitter or Facebook. This aims to provide a broader picture of the use of text data sourced from social media.

Research related to mapping the use of online digital news has been done before by [10]. In this study provides an overview related to the social media research topic big data analytics. Explanations related to the techniques and methods used, the application of big data social media and research opportunities in the field of big data are also discussed. Earlier research was conducted by [11]. Furthermore, in this study more specifically provides an overview related to the use of online news articles in various studies. Among them are used for topic detection and tracking, news summarization, news aggregation, event detection based on topic or categorization. This study also provide further insight related to research opportunities, especially those using online news. System architecture used, techniques and applications related to online news.

The structure of this research begins with the introduction, the related works section, the methodology section and the result and discussion section. In the methodology section we will explain the process of searching and collecting data in further detail.

2. Method

2.1 Search Strategy

In a systematic review, a well-planned search strategy is very important so that every relevant piece of work can be found in the search results. Therefore, an extensive search for research papers was conducted to try answering the proposed research questions. The search terms used in this systematic review were developed using the Kitchenham et al. (2010) steps [7]. The collection of publication data in this study is sourced from Scopus Indexer. Data collection is limited to Journal articles and conferences papers with full paper access. A search is performed for each word contained in the title, abstract and author keywords section. Metadata taken as search results are title, abstract, author keywords, year of publication, publisher and citation. The document period for the search range is from 2010 to early 2020. The keywords used in searching each database of the publication are as follows ("news articles" OR "online news" OR "news") AND ("prediction" OR "detection" OR "forecasting" OR "clustering" OR "classification").

2.2 Data Collection

The process of collecting data as shown in Figure 1. The initial process is carried out in the source database using keywords as described in the section above. Based on the search results from the initial search process, a selection of the results is obtained by displaying several criteria. In the refine search process, the journal and conference articles will be consider with full access. The results of this search are used for further processing to see an overview of the research conducted.

Figure 1. Data Collection Process
3. Results and Discussion

Based on the search results in the previous stage, it is known that the use of online news as a dataset is quite widely used in various fields. As in the economic/finance, political, journalism, health, social and other fields, including NLP. More detail will be explained in the following section.

3.1. Finance

Research that uses news articles in the financial sector is very much done. These studies are generally related to stock price movements. As research conducted by [12]–[17]. In addition, the use of financial news not only predicts stock prices, but also predicts product prices. Research conducted by [18] predict oil prices using news articles. Product price prediction was also carried out by [19] in his research to predict the movement of the price index of drug ingredients in China. Table 1 show more detail about source and number of dataset used on these research.

| Author                      | Dataset                  | Number of Dataset | Purpose                                                                 |
|-----------------------------|--------------------------|-------------------|-------------------------------------------------------------------------|
| Li et.al (2014) [12]        | FINET                    | News article      | To analyze the impact of news sentiment toward the stock price          |
|                             |                          | January 2003 –    |                                                                          |
|                             |                          | March 2008        |                                                                          |
| Shah et. Al (2019) [13]     | Moneycontrol.com         | 6 month news      | To analyzed the effects of news sentiments on the stock market          |
|                             |                          | articles          |                                                                          |
| Shynkevich et. Al (2016) [14]| PR Newswire,             | 51,435            | By using news articles to can improve the performance of financial      |
|                             | McClatchy-Tribune        |                   | forecasting and support the decision-making process                    |
|                             | Business News,           |                   |                                                                          |
|                             | Business Wire            |                   |                                                                          |
|                             | Ifeng.com                | 3,086             | To predict the trend of stock price by using the information structure   |
|                             |                          |                   | contained in daily financial news                                       |
| Matsubara et. Al (2018) [16]| Reuters and             | 292,093           | To predict daily stock price movements given by news articles           |
|                             | Bloomberg                |                   |                                                                          |
| Nam and Seong (2019) [17]   | Naver                    | 1,397,800         | To forecast stock price movement based on the financial news considering|
|                             |                          |                   | causality                                                              |
| Li et.al (2018) [18]        | Investing.com            | 6,756             | To forecast the crude oil price                                         |
| Yu and Guo (2018) [19]      | Various Chinese          | 133,457           | To predict the movements of Chinese medicinal material price indexes    |
|                             | news portal              |                   | using news                                                              |

3.2. Political

Research conducted by [20] predicts the level of acceptance of politicians. This research was conducted on the presidential election in South Korea. [21] In his research using news articles to look at political polarization and also identify partisan media. This research was conducted on news articles originating from Romanian. Other studies on analyzing online media to see patterns of media orientation were carried out by [22]. The study also predicts bias from media reporting using sentiment classifier. Earlier research conducted by [23] used news articles to look at the relationship between party politicians, the relationship between politicians and other parties as well as key political events. This research uses Portuguese news articles to see the relationship. Table 2 show more detail about source and number of dataset used on these research.
Table 2. Research using news articles as dataset in the field of politics

| Author                    | Dataset                  | Number of Dataset | Purpose                                                                 |
|---------------------------|--------------------------|-------------------|-------------------------------------------------------------------------|
| Park et.al (2018) [20]    | YTN                      | 11,476            | To Predict an approval rate of politician in election process           |
| Chiru et. Al (2017) [21]  | hotnews.ro/politic      | Period 2009 – 2013| To analyze news articles from Romanian mass media and extract opinions about political entities |
| Al-Sarraj and Lubbad (2018)[22] | Various source of news (Aljazeera, etc) | 799 | To identify patterns in the press orientation and further in the media bias towards side to another |
| Bento et.al (2014) [23]   | Portuguese newspaper Público | Period 2001 - 2010 | To visualize and explore communities of politicians arising from mentions in news articles |

3.3. Health

News articles as a dataset are also used in health research. News articles are used to monitor disease progress in an area. As research conducted by [24] to monitor the development of fever and Zika in tropical countries. Research conducted by [25] uses online news to look at news related to mammography in online media to provide insight into this. Not only related to human health, the use of news is also used to monitor the development of diseases in animals [26], [27]. The research also extracted information contained in online news to obtain specific information such as location, date, disease and origin using the principle of text mining. Table 3 show more detail about source and number of dataset used on these research.

Table 3. Research using news articles as dataset in the field of health

| Author                    | Dataset                  | Number of Dataset | Purpose                                                                 |
|---------------------------|--------------------------|-------------------|-------------------------------------------------------------------------|
| Zhang et.al (2019) [24]   | LexisNexis and Local news paper | 4,595            | To track case numbers of dengue and zika using news articles            |
| Young Lin et.al (2017) [25] | Google News              | 100               | To characterize online news coverage relating to mammography, including articles’ stance toward screening mammography |
| Valentin et.al (2018) [26] | PADI-web                 | 442               | To automatically find the similarity between different news report about epidemiological animal disease monitoring |
| Goel et.al (2019) [27]    | Google News              | -                 | To visualize and explore epidemiological data for monitoring animal disease outbreaks |

3.4. Social

In addition, online news articles are also used to detect events [28]–[30]. These research introduce event detection in stream news. Recognizing and predicting social problems in the community can also use online news articles [31]. This research extract social term in online news in order to detect social problem in the community. By using the principle of topic detection and tracking, food safety detection can be done through online news articles [32]. This research using China popular news to monitor the food safety by using topic detection in online news. Table 4 show more detail about source and number of dataset used on these research.
Table 4. Research using news articles as dataset in the field of social

| Author                  | Dataset                                | Number of Dataset | Purpose                                                        |
|-------------------------|----------------------------------------|-------------------|                                                               |
| Hu et.al (2017) [28]    | Chine news dataset (NewsMiner)         | 600,000           | To detect online event for news streams                       |
| Zhou et.al (2018) [29]  | Chinatime.com and ifeng.com            | 1,160             | Introduce new model to detect event in news streams           |
| Qian et.al (2019) [30]  | Investide.cn                           | 14,556            | To extract high-quality information about business events from massive online news headlines and leads |
| Suh (2019) [31]         | Naver.com                              | 126,402           | To identify and predict the SocialTERMs from a large number of online news articles |
| Xiao et.al (2019) [32]  | Various China popular news             | 1,255             | To discover food safety events by monitoring online news      |

3.5. Natural Language Processing (NLP)

Predicting news popularity from news [7], detecting fake news [2], [33], [34], news classification [35–37], news clustering [38], news recommendation system [39], [40] and news summarization [41], [42] are sub-fields of research that use news articles as the main dataset. In addition, by using news articles, research can also be conducted related to detecting a controversy in a news [43]. This study uses sentiment analysis and word matching to detect these issues. Table 5 show more detail about source and number of dataset used on these research.

Table 5. Research using news articles as dataset in the field of NLP

| Author                  | Dataset                                | Number of Dataset | Purpose                                                        |
|-------------------------|----------------------------------------|-------------------|                                                               |
| Stokowiec et.al (2017) [7] | Breaking News                         | 38,142            | To predict the popularity of online content using its title    |
| Sholar et.al (2017) [2]  | Kaggle, LIAR, Fake news challenge, University of Washington | 148,743           | To identify a given claim is fake or genuine using Deep Neural Networks |
| Ghosh and Shah (2018) [33] | Kaggle, PolitiFact, Emergent, LIAR, George McIntire | More than 23,358  | To detect fake news by using sentiment as a new important feature to improve the accuracy |
| Bhutani et.al (2019) [34] | Indonesian News website               | 1,309             | To classify the Indonesian News Article                       |
| Londo et.al (2019) [35]  | Sohu                                   | 8,687             | To predict hot news in the news articles                      |
| Wang and Song (2019) [36] | Indonesian News website               | 18,005            | To classify the natural disasters multi-classes problem and to classify the transportation accidents multi-classes problem |
| Lee and Ryu (2019) [38]  | Korean News articles                  | 20,929            | To examine users’ online news commenting behavior by analyzing substantial online news articles and comments |
| Ren, Long and Xu (2019) [39] | GF Securities                         | 64,881            | Financial news recommendation                                 |
| Hu et.al (2020) [40]    | Adressa (Norwegian news)              | approximately 20 million page | To personalize news recommendation for user from online news articles |
| Mishra and Gayen (2018) [41] | CNN/Dailymail                         | More than 250,000 | To summarize news articles                                   |
3.6. Field of implementation using social media as a dataset

The use of social media (Twitter) in predicting, classifying and classifying, has been done in several studies. Research conducted by [45] conducted fake news detection. The research compared the results using twenty-three different supervised artificial intelligence algorithms. Research on the prediction of the popularity of a tweet, conducted by [46]. Using social media, it can also be used to predict stock prices [47]. In addition, by using social media, event detection [48], spam detection [49] and trending topic prediction [50]. More details can be seen in the following table 6.

| No of Dataset | Algorithm/Method | Author |
|---------------|------------------|--------|
| 46,600        | twenty-three different supervised artificial intelligence algorithms | Ozbay and Alatas (2019) [45] |
| 10^4 (original) and 10^6 (re-tweet) | exponential regression prediction model | Wu and Shen (2015) [46] |
| Tweet period 2016 - 2017 | SVM, Naïve Bayes, Neural Network, TF-IDF | Coelho et.al (2019) [47] |
| 10GB          | SVM non-parametric model | Peng et.al (2016) [51] |
| 980           | Naïve Bayes Maximum Entropy Support Vector Machine | Lunando and Purwarianti (2013) [52] |
| Not mention   | SVM Neural Networks | Shareen et.al (2018) [48] |
| Not mention   | KNN, Best First Decision Tree | Jeyapriyanga et.al (2019) [49] |
| 111,387       | Document Pivot BN-Gram | Indra et.al (2018) [50] |

4. Conclusion

Online digital news is one source of data that can be used to perform various machine learning tasks. For example the classification of news, the prediction of the price of a product or stock, see the user's view of an issue, monitor the development of a disease, monitor criminal actions and other topic modeling tasks. In addition, by using social media datasets, we can perform any of machine learning task. Our future research is to explore the possibility of using online digital news as text dataset, to predict the changes of topics in the online digital news, spatial as well as temporal.
References

[1] Q. Zeng, Q. Zhou, and S. Xu, “Neural Stance Detectors for Fake News Challenge,” Stanford, pp. 1–9.
[2] J. M. Sholar, S. Chopra, and S. Jain, 2017. “Towards Automatic Identification of Fake News: Headline-Article Stance Detection with LSTM Attention Models,” IOP Conf. Series: Materials Science and Engineering, pp. 1–15.
[3] K. Miller and A. Oswalt, 2017. “Fake News Headline Classification using Neural Networks with Attention.
[4] C. Aymanns, J. Foerster, and C.-P. Georg, 2017. “Fake News in Social Networks,” SSRN.
[5] X. Guan, Q. Peng, Y. Li, and Z. Zhu, 2017. “Hierarchical neural network for online news popularity prediction,” Proc. - 2017 Chinese Autom. Congr. CAC 2017, pp. 3005–3009.
[6] T. Uddin, M. J. A. Patwary, T. Ahsan, and M. S. Alam, 2017. “Predicting the popularity of online news from content metadata,” 2016 Int. Conf. Innov. Sci. Eng. Technol. ICISSET 2016.
[7] W. Stokowiec, T. Trzciński, K. Wolk, K. Marasek, and P. Rokita, 2017. “Shallow reading with deep learning: Predicting popularity of online content using only its title,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 10352, pp. 136–145.
[8] T. Bansal, D. Belanger, and A. McCallum, 2016. “Ask the GRU: Multi-Task Learning for Deep Text Recommendations,” pp. 107–114.
[9] J. Li, W. Monroe, A. Ritter, M. Galley, J. Gao, and D. Jurafsky, 2016. “Deep Reinforcement Learning for Dialogue Generation, pp. 4.
[10] N. A. Ghani, S. Hamid, I. A. Targio Hashem, and E. Ahmed, 2018. “Social media big data analytics: A survey,” Comput. Human Behav.
[11] W. M. S. Yafooz, S. Z. Z. Abidin, and N. Omar, 2011. “Challenges and issues on online news management,” Proc. - 2011 IEEE Int. Conf. Control Syst. Comput. Eng. ICCSCE 2011, pp. 482–487.
[12] X. Li, H. Xie, L. Chen, J. Wang, and X. Deng, 2014. “News impact on stock price return via sentiment analysis,” Knowledge-Based Syst. 69(1) pp. 14–23.
[13] D. Shah, H. Isah, and F. Zulkernine, 2019. “Predicting the Effects of News Sentiments on the Stock Market,” Proc. - 2018 IEEE Int. Conf. Big Data, Big Data 2018, pp. 4705–4708.
[14] Y. Shynkevich, T. M. McGinnity, S. A. Coleman, and A. Belatreche, 2016. “Forecasting movements of health-care stock prices based on different categories of news articles using multiple kernel learning,” Decis. Support Syst. 85, pp. 74–83.
[15] W. Long, L. Song, and Y. Tian, 2019. “A new graphic kernel method of stock price trend prediction based on financial news semantic and structural similarity,” Expert Syst. Appl. 118, pp. 411–424.
[16] T. Matsubara, R. Akita, and K. Uehara, 2018. “Stock price prediction by deep neural generative model of news articles,” IEICE Trans. Inf. Syst. E101D(4), pp. 901–908.
[17] K. H. Nam and N. Y. Seong, 2019. “Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market,” Decis. Support Syst. 117, pp. 100–112.
[18] X. Li, W. Shang, and S. Wang, 2018. “Text-based crude oil price forecasting: A deep learning approach,” Int. J. Forecast. 35(4), pp. 1548–1560.
[19] M. Yu and C. Guo, 2018. “Using news to predict Chinese medicinal material price index movements,” Ind. Manag. Data Syst. 118(5), pp. 998–1017.
[20] J. Park, Y. Na, and I. C. Moon, 2017. “Text augmented automatic statistician for predicting approval rates of politicians,” 2017 IEEE Int. Conf. Syst. Man, Cybern. SMC 2017, pp. 954–959.
[21] C. G. Chiru, T. Dimcica, and S. Caciandone, 2017. “Expression of Political Opinions in Press,” Proc. - 2017 21st Int. Conf. Control Syst. Comput. CSCS 2017, pp. 182–189.
[22] W. F. Al-Sarraj and H. M. Lubbad, 2018. “Bias Detection of Palestinian/Israeli Conflict in Western Media: A Sentiment Analysis Experimental Study,” Proc. - 2018 Int. Conf. Promis.
Electron. Technol. ICPET 2018, pp. 98–103.

[23] C. Bento, D. Goncalves, and B. Martins, 2014. “Visualizing the evolution of groups of politicians mentioned in the news,” Proc. Int. Conf. Inf. Vis., pp. 366–367.

[24] Y. Zhang, M. Ibaraki, and F. W. Schwartz, 2019. “Disease Surveillance Using Online News: Dengue and Zika in Tropical Countries,” J. Biomed. Inform. 102.

[25] L. L. Young Lin and A. B. Rosenkrantz, 2017. “The U.S. Online News Coverage of Mammography Based on a Google News Search,” Acad. Radiol. 24(1), pp. 1612–1615.

[26] S. Valentin, R. Lancelot, and M. Roche, 2018. “How to combine spatio-temporal and thematic features in online news for enhanced animal disease surveillance?,” Procedia Comput. Sci. 126, pp. 490–497.

[27] R. Goel et al., 2019. “EpidNews: Extracting, exploring and annotating news for monitoring animal diseases,” J. Comput. Lang. 56, pp. 100936.

[28] L. Hu, B. Zhang, L. Hou, and J. Li, 2017. “Adaptive online event detection in news streams,” Knowledge-Based Syst. 138, pp. 105–112.

[29] P. Zhou, Z. Cao, B. Wu, C. Wu, and S. Yu, “EDM-JBW: A novel event detection model based on JS-IDForder and Bikmeans with word embedding for news streams,” J. Comput. Sci. 28, pp. 336–342.

[30] Y. Qian, X. Deng, Q. Ye, B. Ma, and H. Yuan, “On detecting business event from the headlines and leads of massive online news articles,” Inf. Process. Manag. 56(6), pp. 102086.

[31] J. H. Suh, 2019. “SocialTERM-Extractor: Identifying and Predicting Social-Problem-Specific Key Noun Terms from a Large Number of Online News Articles Using Text Mining and Machine Learning Techniques,” Sustainability. 11(1).

[32] K. Xiao, C. Wang, Q. Zhang, and Z. Qian, 2019. “Food safety event detection based on multi-feature fusion,” Symmetry (Basel). 11(10) pp. 1–12.

[33] S. Ghosh and C. Shah, 2018. “Towards automatic fake news classification,” Proc. Assoc. Inf. Sci. Technol. 55(1), pp. 805–807.

[34] B. Bhutani, N. Rastogi, P. Sehgal, and A. Purwar, 2019. “Fake News Detection Using Sentiment Analysis,” 2019 12th Int. Conf. Contemp. Comput. IC3 2019, pp. 1–5.

[35] G. L. Y. Londo, D. H. Kartawijaya, H. T. Ivariyani, Y. S. Purnomo WP, M. R. Aryasuta P, and Di. Ariyandi, 2019. “A Study of Text Classification for Indonesian News Article,” 2019 Int. Conf. Artif. Intell. Inf. Technol., pp. 205–208.

[36] Z. Wang and B. Song, 2019. “Research on hot news classification algorithm based on deep learning,” Proc. 2019 IEEE 3rd Inf. Technol. Networking, Electron. Autom. Control Conf. ITNEC 2019, no. Itnec, pp. 2376–2380.

[37] T. F. Abidin, N. R. Sari, A. Z. Ramadhan, Irvanizam, and R. P. F. Afidh, 2018. “Performance evaluation of n-grams ratio features in solving multi-classes classification problems,” Proc. 2018 10th Int. Conf. Inf. Technol. Electr. Eng. Smart Technol. Better Soc. ICITEE 2018, pp. 307–311.

[38] S. Y. Lee and M. H. Ryu, 2019. “Exploring characteristics of online news comments and commenters with machine learning approaches,” Telemat. Informatics. 43, pp. 101249.

[39] J. Ren, J. Long, and Z. Xu, 2019. “Financial news recommendation based on graph embeddings,” Decis. Support Syst. 125, pp. 113115.

[40] L. Hu, C. Li, C. Shi, C. Yang, and C. Shao, 2020. “Graph neural news recommendation with long-term and short-term interest modeling,” Inf. Process. Manag. 57(2), pp. 10214.

[41] R. Mishra and T. Gayen, 2018. “Automatic Lossless-Summarization of News Articles with Abstract Meaning Representation,” Procedia Comput. Sci. 135, pp. 178–185.

[42] R. Y. Rumagit, N. Setiyawati, and D. H. Bangkalang, 2019. “Comparison of graph-based and term weighting method for automatic summarization of online news,” Procedia Comput. Sci. 157, pp. 663–672.

[43] A. Sriteja, P. Pandey, and V. Pudi, 2017. “Controversy detection using reactions on social media,” IEEE Int. Conf. Data Min. Work. ICDMW. 2017, pp. 884–889.
[44] G. Margagliotti, T. Bollé, and Q. Rossy, 2019. “Worldwide analysis of crimes by the traces of their online media coverage: The case of jewellery store robberies,” Digit. Investig. 31.

[45] F. A. Ozbay and B. Alatas, “Fake news detection within online social media using supervised artificial intelligence algorithms, 2019.” Phys. A Stat. Mech. its Appl. 540, pp. 123174.

[46] B. Wu and H. Shen, 2015. “Analyzing and predicting news popularity on Twitter,” Int. J. Inf. Manage. 35(6), pp. 702–711.

[47] J. Coelho, D. D’Almeida, S. Coyne, N. Gilkerson, K. Mills, and P. Madiraju, 2019. “Social media and forecasting stock price change,” Proc. - Int. Comput. Softw. Appl. Conf., 2, pp. 195–200.

[48] N. Shahreen, M. Subhani, and M. Mahfuzur Rahman, 2018. “Suicidal Trend Analysis of Twitter Using Machine Learning and Neural Network,” 2018 Int. Conf. Bangla Speech Lang. Process. ICBSLP 2018, pp. 1–5.

[49] S. Jeyapriyanga, B. Mahalakshmi, and C. Anuradha, 2019. “Spam detection in twitter using machine learning algorithms,” Int. J. Eng. Adv. Technol. 8(6) Special Issue 2, pp. 174–178.

[50] Indra, E. Winarko, and R. Pulungan, 2018 “Trending topics detection of Indonesian tweets using BN-grams and Doc-p,” J. King Saud Univ. - Comput. Inf. Sci. 31(2), pp. 266–274.

[51] B. Peng, J. Li, J. Chen, X. Han, and R. Xu, 2015. “Trending Sentiment-Topic Detection on Twitter, 3406, pp. 66–77.

[52] E. Lunando and A. Purwarianti, 2013. “Indonesian social media sentiment analysis with sarcasm detection,” 2013 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACSIS 2013, pp. 195–198.