Depression and anxiety detection through the Closed-Loop method using DASS-21

Setiyo Budiyanto*, Harry Candra Sihombing, Fajar Rahayu I. M. Department of Electrical Engineering, Faculty of Engineering, Universitas Mercu Buana, Jakarta 11650, Indonesia *Corresponding author, e-mail: sbudiyanto@mercubuana.ac.id

Abstract

The change of information and communication technology has brought many changes in daily life. The way humans interacting is changing. It is possible to express each form of communication directly and instantly. Social media has contributed data in size, diversity and capacity and quality. Based on it, the idea was to see and measure the tendency of depression and anxiety through social media using the Closed-Loop method using Facebook text mining posts. Through the stages of pre-processing including text extraction using the Naïve Bayes machine learning model for text classification, the early signs of depression and anxiety are measured using DASS-21 parameter. In total, 22,934 Facebook posts were contributed as training and learning data collected from July 2017 until July 2018. As a results, analysis and mapping of social demographics of users that are usually as a trigger of depression, and anxiety, such as grief, illness, household affairs, children education and others are available.

Keywords: Closed-Loop, DASS-21, depression and anxiety, machine learning, Naïve Bayes

Copyright © 2019 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

Social media as manifested in everyday life allows people to share thoughts and shows emotions that characterize depression and anxiety. Feelings of worthlessness, guilt, helplessness, self-hatred and other aspects. Social media has the potential as a tool for measuring and monitoring mental health depression and anxiety. Social media can help detect measure and capture the social context of sufferers of depression and anxiety in the population [1]. Characteristics of Big Data are data content quantity (Volume), data density (Velocity), data format (Variety), quality of data (Veracity). Big data can also be referred as data that tells its actual information that can be used when it needed [2]. A big data analysis conducted to ranking, finding and identifying meaningful information from large unstructured data by analyzing related details based on sequences of textual meta-data processing, identification, and time series processing [2]. The DASS-21 scale is a measurement scale that is used to facilitate grouping and categorizing social media texts into a best matched dual factor model that shows the general negative effects of specific factors of Depression, Anxiety, and Stress in psychological disorders [3]. Closed-Loop system is a control system that uses feedback which partly comes from the output signal that is fed back as an input to minimize errors and to improve accuracy on the system [4]. Social media can be used to see the condition of its users through analysis of texts using the Naïve Bayes classification algorithm as the most widely used [5] and fast classification of Naïve Bayes works better and requires only less training data to predict classes from a collection of data test when independence assumptions apply [6].

The idea is to measure depression and anxiety through text mining of Facebook, Closed-Loop method used as a learning process and Naïve Bayes machine as training model for text classification that raises several questions as follows:

- The initial signs (What) that cause depression or anxiety occur in social media users ?

- Causes (Why) of depression or anxiety can occur to social media users ?
- Identification (Who) of depression or anxiety from social media users ?
- Does (When) depression or anxiety can occur to social media users ?

- Where depression or anxiety occurs ?

- How depression or anxiety occurs, does it affect social media users ?

The approach by [5], making a decision support system to collect signals that produce important patterns that lead to child abuse through structured and unstructured data in free text using the Random Forest Model machine learning model and Support Vector Machine through text classification algorithms. In their research [6], choose a smart system approach, a data-based decision model. Intelligent systems and machine learning models through new methodologies in intelligent systems in real-world business-to-business (B2B) sales forecasting. A group of sales experts collect B2B sales case records with known results to support the task of predicting new sales opportunities, allow cognitive evaluations of models by users based on new insights. In [7] explained, depression is a general mental disorder that arises with feelings of depression, loss of interest or pleasure, decreased energy, feelings of guilt or inferiority, disturbed sleep or appetite disorders, and poor concentration. What's more, depression is often accompanied by anxiety symptoms. These problems can become chronic or recurrent and cause substantial damage and the ability of individuals to carry out their daily roles and responsibilities.

In the worst conditions, depression can cause someone to commit suicide. There are an estimated 788,000 suicide deaths worldwide. Suicide is the second leading cause of death among children aged 15-29 years globally with a global age standard suicide rate of 10.7 per 100,000 population. This tragedy has long-term effects on people who are left behind and greatly affects families, communities and countries (WHO, 2015). Meanwhile [8], in their research using the Pattern Recognition Algorithm on Big Data social media, online transactions, network sensors or mobile devices. This approach is used in the application of the Computer Vision field and image analysis with large-scale characteristics.

Depression is one of the most devastating psychiatric disorders, the main cause of disability in people in teens and productive ages. Tragically, 76% of people with moderate depression and 61% of people with major depression never get help on time [9]. Then [10] added, mental emotional disorders are the same term as psychological distress. This condition is a condition that indicates a person is experiencing psychological changes that can be experienced by everyone in certain circumstances but can recover as before. The prevalence of severe mental disorders in the Indonesian population is 1.7 per mile. Most severe mental disorders in DI Yogyakarta, Aceh, South Sulawesi, Bali and Central Java.

According to [11], worry and anxiety are two different concepts that occur in various parts of the human brain. Anxious can occur without worry, and worry without anxiety, but both tend to be inseparable conditions. anxious originating from the mind, occur in the mind, and involve the brain thinking of the prefrontal cortex, interacting with the limbic system, which controls emotions and basic human instincts. Anxiety is always portrayed physically, as symptoms, actions, and behaviors of the body that trigger brain parts to turn on the circuit of fear. Anxiety can have unconscious traits as a indications of cause and will appear as symptoms, such as abdominal pain, headache, or shortness of breath [11]. Using batch process-oriented Hadoop and MapReduce on Big Data for decision makers who adopt analysis to achieve efficient decisions in accordance with the application domain. The adoption of a mechanism by the Query Controller (QC) that is able to manage the results carried out on a number of processors each responsible for each existing cluster [12]. Based on the conditions [13], describes research on the importance of open digital collaboration as a sociotechnical system that loosely binds individuals when facing challenges in analyzing data with new datasets that reach different contexts using new computational models and analytical techniques.

Then [14], conduct research related to big data analysis, cloud computing, social networking and machine learning on computer vision face recognition and cloud computing Extreme Learning Machines techniques to perform new cloud-based Face Tagging recognition on social media and database engines face recognition is in the cloud (non-local) and operates on a large-scale image database (Big Data). The user interface communicates with a cloud-based web API containing face recognition machines and face databases consisting of face detection, extraction, matching and so on.

Research by [15], propose regression methods through Extended Markov Chain on predictions of the arrival of rain using other machine learning algorithms to predict rainfall based on predictive accuracy and with minimal correlations that exist in all climates. Research approach by [16], shows a summary of the proposed system architecture. Interactive

exploration of large-scale multi-modal data flow consists of repeated parallel assignments of MapReduce to find patterns that often appear.

In his research approach [17], use learning algorithms for Shallow and Deep Extreme Learning Machines to exploit the latest cutting-edge data technology, learning algorithms, and data-based Transient Delay Prediction System (TDPS) statistics tools for large-scale railway networks. While [18], mentions the social media as a private media has emerged as a medium to communicate opinions, products and services or even political and public events as rich resources for sentiment analysis and text mining, machine learning, statistical and computational linguistics. Through open source R applications, classification, data mining techniques such as grouping to find associations and patterns in text and in exploring and discovering new information and relationships in textual sources on Twitter data.

The approach by [19], using Cognitive Biases theory in looking an anxiety conditions and the limits of bias associated with threats in anxiety, concluding that the bias has comparable magnitude in various types of anxious populations. Three experimental paradigms have been used to study patterns related to bias affecting anxiety: emotional stroop, dot probe, and spatial emotional cuing. This solution offers a findings from existing meta-analyzes. Machine learning and text analytics have proven increasingly useful in a number of health related applications in analyzing of online data for disease epidemics as a warning signs of a variety of mental health issues such as anxiety, anorexia and depression. Personal blogs are collected from the Tumblr API, and labeled them based on whether they exhibited [20].

Semantic HMC proposed by [21] is a hierarchical multi-label classification used to automatically classify unstructured text documents according to an ontology to describe the classification model that focus on analyzing the data in Big Data sources. In his research [22], discussing depression which is a serious challenge in personal and public health. The proposed methodology uses crowdsourcing to collect a reported set of Twitter users and diagnosed with clinical depression, based on a standard psychometric instrument to establish a statistical classification that provides an estimate of the risk of depression, implemented several steps to measure individual social media behavior for one year before the onset of depression reported.

Research aims by [23] is to classify authors of tweets by comparing machine learning methods like logistic regression and naive Bayes. The Naïve Bayes classification estimates P(C) and P(X|C) of the document (X) and class (C), the the relative frequency of each target class in the training data calculated using the independence assumption that the attribute depends on the conditions determined by target class value. The act of authenticating information encountered on social media becomes very complex. The processes of this application are fetching of tweets, pre-processing, feature extraction, and developing a machine learning model for classification.

Computer network technology as a medium of communication between devices has made significant progress in terms of communication media. It is one of the fastest growing internet applications now [24]. Data traffic has increased by 131% since 2011 has encouraged some mobile operators in Europe to invest in machine-to-machine communications [25] and protocol routing can improve the performance of the network, can improve troughput and reduce handover delay [26]. The technologies that underlie the effectiveness of 3G-WiFi offloading are in constant development. An important such technology are the Cognitive Radios, which work by intelligently adjusting both signal strength and resource use [27].

2. Research Method

The purpose of this study is to overlook of opportunities in big data and machine learning through Naïve Bayes text classification using text mining social media, to detect and analyse depression and anxiety through the Closed-Loop method and DASS-21.

2.1. Initiation and Identification of Review

Combination of keywords as well as in lowercase letters without quotation marks and each word separated by a space that represent the objectives of the research, namely "Big Data", "Machine Learning", "Text Mining", "Depression" and "Anxiety" and "Closed-Loop". Number of combinations of keywords used in IEEExplore Digital Library journal searches produced 6 groups of keywords 6, 15, 20, 15, 6 and 1 combinations.

3. Results and Analysis

A number of 128 publications from the IEEEXplore obtained as main database, the next phase was to strip the results by using the stages of reviewing the literature abstract, suggesting development and conclusions and also reviewing the entire contents of the literature. The research framework was developed to group and categorize social media texts into specific factors of depression, anxiety, stress (DAS). Social media texts are collected according to identity, name, gender, age and geographic information. Check condition function that found in DASS-21 been applied to have a probability value and a condition marker. Looping is carried out until existing texts have been checked as shown in Figure 1 and Table 1.



Figure 1. The research method proposed in the study

3.1. Data Sources

Data collected from 749 Facebook accounts, an account cleansing process done by selecting social media accounts from a person name, as in total there is 654 personal accounts resulted. Random selection using Slovin minimal sample formula is applied with a margin of error of 5% or 0.05. There is 248 accounts has been selected with 109 women and 139 men in details.

3.2. Depression and Anxiety Behavior

Social media text classification based on a dataset represented on Bahasa Indonesia and English through a machine learning process using the Naïve Bayes algorithm as a training process and parameters in DASS-21 as a learning process. Based on the above proccess, 165 Facebook accounts with a tendency to depression and anxiety resulted that were in the range of years 2009-2018 as shown in Figure 2 (a) and (b). A map of depression and anxiety

trends distribution from Facebook shown in Figure 3 (a) and a detailed map of Indonesia from each condition and spreading of depression and anxiety shown in Figure 3 (b).







Figure 2. (a) Depression and Anxious posts per user, (b) Number of cleaned posts in the range of years 2018-2009 from total = 22.934 posts



Figure 3. (a) The tendency of the depression and anxious city groups, (b) geographical map of the distribution of depression (blue) anxious (yellow) and stress (purple)

3.2.1. Initial Signs (What) that Cause Depression or Anxiety?

In Table 2, the top 50 lists show the causes of depression and anxiety as expressed through the posts of Facebook. The selection of text is determined based on the highest level of probability of depression and anxiety from each post text.

Table 2. The Probability of Text Showing Early Signs of Depression and Anxiety

	Depression					Anxiety					
No.	Users	Prob.	No.	Users	Prob.	No.	Users	Prob.	No.	Users	Prob.
1	8	9.89	26	9	7.39	1	78	9.11	26	154	6.28
2	46	9.85	27	13	7.35	2	48	9.06	27	74	6.02
3	37	9.79	28	5	7.24	3	222	8.87	28	80	5.83
4	78	9.73	29	12	7.15	4	135	8.75	29	26	5.63
5	72	9.67	30	70	7.03	5	17	8.59	30	43	5.5
6	19	9.66	31	31	6.83	6	65	8.55	31	34	5.48
7	6	9.64	32	22	6.8	7	32	8.44	32	18	5.3
8	34	9.59	33	43	6.76	8	46	8.32	33	16	5.27
9	24	9.54	34	10	6.76	9	4	8.3	34	218	5.15
10	49	9.52	35	74	6.64	10	206	8	35	2	5.09
11	80	9.51	36	56	6.6	11	57	7.97	36	166	5.07
12	32	9.48	37	4	6.24	12	83	7.85	37	94	5.04
13	83	9.22	38	21	6.23	13	24	7.82	38	61	5.03
14	28	9.11	39	65	6.2	14	136	7.64	39	131	4.9
15	25	8.91	40	48	5.87	15	31	7.59	40	41	4.87
16	53	8.89	41	1	5.85	16	49	7.53	41	8	4.8
17	55	8.89	42	51	5.7	17	147	7.37	42	230	4.76
18	15	8.69	43	26	5.38	18	192	7.33	43	1	4.72
19	16	8.68	44	69	5.34	19	37	7.23	44	35	4.71
20	17	8.52	45	89	5.23	20	185	7.08	45	55	4.57
21	39	8.4	46	84	5.2	21	246	7.06	46	6	4.56
22	91	7.98	47	18	5.11	22	39	7	47	10	4.52
23	57	7.67	48	61	4.88	23	150	6.57	48	219	4.51
24	52	7.66	49	33	4.71	24	91	6.44	49		
25	2	7.56	50	66	4.68	25	72	6.34	50		

3.2.2. Causes (Why) of Depression or Anxiety can Occur to Social Media Users?

Based on the user's text post in Facebook, data cleaning such as punctuation, numbers and readable characters that 5,651 words are obtained, so that several keywords can be shown as causes of depression and anxiety as shown in Table 3 with at least words 10 times appearance.

Table 3. Things that Cause Depression and Anxiety

No.	Word	Sum	No.	Word	Sum	No.	Word	Sum	No.	Word	Sum
			-						-		
1	Tuhan	62	16	Keluarga	19	31	Nephew	14	46	Belajar	11
2	Anak	52	17	Foto	19	32	Jakarta	13	47	Allah	11
3	Selamat	49	18	Time	18	33	Brother	13	48	Film	11
4	Hidup	49	19	Year	18	34	Natal	13	49	Air	11
5	Orang	45	20	Manusia	18	35	Team	13	50	Perasaan	11
6	God	37	21	Hati	17	36	Acara	13	51	Work	11
7	Family	35	22	Home	17	37	Mama	13	52	Hasil	11
8	Gold	32	23	Salah	16	38	World	12	53	Bali	10
9	Tangan	31	24	Heart	16	39	Kuat	12	54	Mom	10
10	Day	30	25	Semoga	16	40	Beloved	12	55	Terimakasih	10
11	Takut	27	26	Kasih	15	41	Sister	12	56	School	10
12	Teman	25	27	Jalan	15	42	Morning	12	57	Sekolah	10
13	Life	22	28	Christmas	15	43	Holiday	12	58	Yesus	10
14	Bless	20	29	Enjoy	14	44	Doa	12	59	Lagu	10
15	Sehat	19	30	Friend	14	45	Women	12	60	Peace	10

3.2.3. Identification (Who) of Depression or Anxiety from Social Media Users?

Based on information of Facebook user, shown demographic data based on gender in Table 4, employment status in Table 5. Men have a higher tendency (53.0%) to experience depression and anxiety than women (46.4%) and Groups of users with has a Job

TELKOMNIKA

status have a higher tendency to experience depression and anxiety than other groups of users. The age group 41-50 years has a higher tendency to experience depression and anxiety than the age group under 30 years and above 51 years as shown in Figures 3 (a) and 3 (b). User groups with University education or equivalent have a higher tendency to experience depression and anxiety. The group of users with Married status also has a higher tendency to experience depression and anxiety. Where single status tends to be depressed and anxious which is quite high as shown in Figure 4 and Figure 5.





Figure 3. (a) Users by age group, (b) users according to birth date information; n = 165.



Figure 4. Tendency of depression and anxiety in education background



Figure 5. Tendency of depression and anxiety in married status

3.2.4. Does (when) Depression or Anxiety can Occur to Social Media Users?

More detailed description of the number of depressed and anxious tendencies of users is based on the average number per month and the average number per day in the range 2017-2018 as in Figure 6 (a) and Figure 6 (b).

3.2.5. Where Depression or Anxiety Occurs?

Based on data obtained, then the information of the top 30 cities was collected with the highest number of depression and anxiety tendencies as in Figure 7. It was found that the big cities (Jakarta, Medan, Bandung, Surabaya) showed the greatest number of depression and anxiety tendencies.





Figure 6. (a) Number of posts of anxious and, (b) depression in units of days per month



Figure 7. Trend of depression (blue) and anxiety (orange)-City

Depression and anxiety detection through the Closed-Loop method using... (Setiyo Budiyanto)

3.2.6. How Depression or Anxiety Occurs, does it Affect Social Media Users?

Based on data analysis, there are no expressions or statements from Facebook users that can be attributed that in the previous post that can show as the cause of depression or anxious.

4. Conclusion

The initial signs of depression and anxiety has been seen through Facebook text analyzing. Based on it was found that statements through social media text has relation to depression and anxiety can approach the conditions experienced by users. The source of depression and anxiety, such as grief, illnesses, household affairs, school children and others. The process of depression and anxiety can last short, medium or long time depending on conditions and social demographic background of users such as gender, age, age of marriage, education and other things. In the future, using text from Twitter, Instagram, Path and other social media sources can be used to develop an analysis of depression and anxiety using images or photos posted by users through face recognition methods using more complex algorithms and analysis to have broader results.

References

- [1] Casadei D, Serra G, Tani K. Implementation of a direct control algorithm for induction motors based on discrete space vector modulation. IEEE Transactions on Power Electronics. 2000; 15(4): 769-77.
- [2] Miah SJ, Vu HQ, Gammack J, McGrath M. A big data analytics method for tourist behaviour analysis. Information & Management. 2017; 54(6): 771-85.
- [3] Moore SA, Dowdy E, Furlong MJ. Using the Depression, Anxiety, Stress Scales–21 with US Adolescents: An Alternate Models Analysis. *Journal of Psychoeducational Assessment*. 2017; 35(6): 581-98.
- [4] Tutorial E. Closed-loop Systems. *Electron Tutor* [*Internet*]. 2014; 1. Available from: https://www.electronics-tutorials.ws/systems/closed-loop-system.html. [Accessed: July 7 2018].
- [5] Amrit C, Paauw T, Aly R, Lavric M. Identifying child abuse through text mining and machine learning. *Expert systems with applications*. 2017; 88: 402-18.
- [6] Bohanec M, Borštnar MK, Robnik-Šikonja M. Explaining machine learning models in sales predictions. *Expert Systems with Applications*. 2017; 71: 416-28.
- [7] Yasamy MM, van Ommeren M, Chisholm D, Saxena S. Depression. A global Public health Concern. WHO Department of Mental Health and Substance Abuse. World Health Organization. 2012.
- [8] Fernández A, Gómez Á, Lecumberry F, Pardo Á, Ramírez I. Pattern recognition in Latin America in the "Big Data" era. *Pattern Recognition*. 2015; 48(4): 1185-96.
- [9] Leahy RL, Holland SJ, McGinn LK. Treatment Plans and Interventions for Depression and Anxiety Disorders. Second Edi, Vol. 26. London: The Guilford Press. 2012: 17, 88, 145, 165, 556, 615.
- [10] RISKESDAS Riset Kesehatan Dasar Kementerian Kesehatan Republik Indonesia. 2013: 1-384.
- [11] Mohan N, Undeland TM, Robbins WP. Power Electronics. New York: John Wiley & Sons. 2005: 11-13 Hampton, D. Beat Depression and Anxiety by Changing Your Brain: With Simple Practices That Will Improve Your Life. CreateSpace Independent Publishing Platform. 2015: 1-138.
- [12] Kolomvatsos K, Anagnostopoulos C, Hadjiefthymiades S. An efficient time optimized scheme for progressive analytics in big data. *Big Data Research*. 2015; 2(4): 155-65.
- [13] Brunswicker S, Bertino E, Matei S. Big data for open digital innovation-a research roadmap. *Big Data Research*. 2015 Jun 30; 2(2): 53-8.
- [14] Vinay A, Shekhar VS, Rituparna J, Aggrawal T, Murthy KB, Natarajan S. Cloud based big data analytics framework for face recognition in social networks using machine learning. *Procedia Computer Science*. 2015; 50: 623-30.
- [15] Cramer S, Kampouridis M, Freitas AA, Alexandridis AK. An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives. *Expert Systems with Applications*. 2017; 85: 169-81.
- [16] Ruan G, Zhang H. Closed-loop big data analysis with visualization and scalable computing. *Big Data Research*. 2017; 8: 12-26.
- [17] Oneto L, Fumeo E, Clerico G, Canepa R, Papa F, Dambra C, Mazzino N, Anguita D. Train Delay Prediction Systems: A Big Data Analytics Perspective. *Big data research*. 2018; 11: 54-64.
- [18] Younis EM. Sentiment analysis and text mining for social media microblogs using open source tools: An empirical study. *International Journal of Computer Applications*. 2015; 112(5): 44-48.
- [19] Bar-Haim Y, Lamy D, Pergamin L, Bakermans-Kranenburg MJ, Van Ijzendoorn MH. Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological bulletin.* 2007; 133(1): 1-24.

- [20] Simms T, Ramstedt C, Rich M, Richards M, Martinez T, Giraud-Carrier C. Detecting cognitive distortions through machine learning text analytics. InHealthcare Informatics (ICHI). 2017 IEEE International Conference on 2017 Aug 23 (508-512). IEEE.
- [21] Peixoto R, Cruz C, Silva N. Adaptive learning process for the evolution of ontology-described classification model in big data context. InSAI Computing Conference (SAI), 2016: 532-540). IEEE.
- [22] De Choudhury M, Gamon M, Counts S, Horvitz E. *Predicting depression via social media*. ICWSM. 2013; 13: 1-0.
- [23] Aborisade O, Anwar M. Classification for Authorship of Tweets by Comparing Logistic Regression and Naive Bayes Classifiers. In2018 IEEE International Conference on Information Reuse and Integration (IRI) 2018: 269-276). IEEE.
- [24] Ramadhan E, Firdausi A, Budiyanto S. Design and analysis QoS VoIP using routing Border Gateway Protocol (BGP). InBroadband Communication, Wireless Sensors and Powering (BCWSP), 2017 International Conference on 2017: 1-4. IEEE.
- [25] S Budiyanto, M Asvial, D Gunawan. Performance Analysis of Genetic Zone Routing Protocol Combined with Vertical Handover Algorithm for 3G-WiFi Offload. *J. ICT Res. Appl.* 2014; 8(1): 49-63.
- [26] M Asvial, S Budiyanto, D Gunawan. An intelligent load balancing and offloading in 3G WiFi offload network using hybrid and distance vector algorithm. IEEE Symp. Wirel. Technol. Appl. ISWTA, 2014: 36-40.
- [27] S Budiyanto, M Asvial, D Gunawan. *Implementation Dedicated Sensing Receiver (DSR) in 3G-WiFi Offload*. Proceedings-2014 International Conference on Smart Green Technology in Electrical and Information Systems: Towards Greener Globe through Smart Technology, ICSGTEIS 2014: 37-42.