



Full length article

Ina-BWR: Indonesian bigram word rule for multi-label student complaints

Tora Fahrudin ^{a,b,*}, Joko Lianto Buliali ^a, Chastine Fatichah ^a^a Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia^b School of Applied Science, Telkom University, Bandung 40257, Indonesia

ARTICLE INFO

Article history:

Received 23 November 2018

Revised 6 March 2019

Accepted 17 March 2019

Available online 26 March 2019

Keywords:

Multi-label

Student complaints

Bag of word

Indonesian bigram word rule

Opinion identification rules

ABSTRACT

Handling multi-label student complaints is one of interesting research topics. One of techniques used for handling multi-label student complaints is Bag of Word (BoW) method. In this research bigram word rule and preprocess are proposed to increase the accuracy of multi-label classification results. To show the effectiveness of the proposed method, data from Telkom University student data and additional relevant data by using hashtag are used as testing data. We develop Indonesian Bigram Word Rule for Multi-label Student Complaints (Ina-BWR) to identify multi-label student problems based on Bigram Word Rule. Ina-BWR consists of three processes such as preprocessing informal text, identifying complaint and object from text. Additional preprocessing techniques are conducted to formalize the text such as parsing a hashtag, correcting affixes word, correcting a conjunction word, parsing suffix people pronoun and correcting typo words. Indonesian bigram word rule is adopted from opinion identification rules with 3 additional corpuses (-)NN, (-)JJ and (-)VB to identify student complaints. To identify complaints, four label corpuses have been created manually. The experimental results show that Ina-BWR can increase Personal, Subject and Relation label accuracies. The best accuracy for four labels is obtained when Ina-BWR is combined with BoW method.

© 2019 Production and hosting by Elsevier B.V. on behalf of Faculty of Computers and Information, Cairo University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The number of smartphone users is growing rapidly in the world. In line with that, the internet accessibility is increasing year to year. Both contribute to the growing of social media users. Based on January 2017 data from Hootsuite, the world smartphone penetration is about 66%, the world internet penetration is about 50% and the world social media active user penetration is about 37% [1]. Based on 2018 social media trend [2], the first ranks in social media applications is Facebook which has 2.07 billion active users, and then followed by Instagram (800 million), Twitter (330 mil-

lion), LinkedIn (500 million), Pinterest (200 million), and also Snap-Chat (178 million). According to ministries of communication and information of Indonesia web news, in 2013, the number of Indonesian Twitter user is 19.5 million [3]. Twitter is chosen as data source in this research for reasons such as: tweets are generally public and concise (140–280 characters), tweets can be acquired by using API [4], Twitter is easily operated by students [5] and is also used in many research such as sentiment analysis [6], trending topic detection in Indonesian language [7], and some Indonesian local government offices use Twitter for acquiring public complaints in Indonesian language (such as Bandung local government offices [8]).

For students, Twitter is a media for sharing experiences, sharing emotions and seeking social support. Chen [4] used data from Twitter to understand issues and problems in their educational experiences. The results of the research can be used by education institutions to detect students who have problems in their studies. In academic failure detection topics, using social media as a data source is relatively new. Traditionally, demographic, academic, social network, or combination of those data are used to detect academic failure. To the best of our knowledge, content extraction from English tweet posting to understand student problems is so

* Corresponding author.

E-mail addresses: tora15@mhs.if.its.ac.id, torafahrudin@telkomuniversity.ac.id (T. Fahrudin), joko@cs.its.ac.id (J.L. Buliali), chastine@if.its.ac.id (C. Fatichah).

Peer review under responsibility of Faculty of Computers and Information, Cairo University.



<https://doi.org/10.1016/j.eij.2019.03.001>

1110-8665/© 2019 Production and hosting by Elsevier B.V. on behalf of Faculty of Computers and Information, Cairo University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

far only done by [4], which use five prominent labels in student problems: heavy study load, lack of social engagement, negative emotions, sleep problems and diversity issues.

Zhang [9] shows 7 factors which are related to student problems: Personal, Family, Peer-Related, Subject/content, Institutional Agent, Institutional, and Social. Four factors were chosen from seven factors for prominent labels: Personal, Relation, Subject and Institutional Agent. Personal label is a label which contains some personal complaints such as sick, confused, etc. Relation label is combination between family and peer-related label which contains some relation complaints such as longing for family, having problems with friends, etc. Subject label is a label which contains heavy study load, heavy of material course, etc. Lastly, Institutional Agent label is a label which contains some complaints about lecturer, faculty, advisor, etc. In this research we use these four labels as category labels. Those labels already cover five labels in [4].

Bag of Word method which is used in [4] has some weakness such as missing word order, neglecting grammar, etc [10]. We propose Indonesian Bigram Word Rule for Multi-label Student Complaints (Ina-BWR), a semantic method for identifying Indonesian multi-label student complaints to overcome missing word order and neglecting grammar. Ina-BWR method consists three processes: preprocess an informal text, identify complaint and identify object.

Preprocessing text data is needed because text data in social media usually informal and unstructured [4,11]. Some studies have been done in preprocessing Indonesian social media data and result some preprocess such as removing punctuation mark, converting number to letter, removing letter repetition [11], removing hashtag, converting slang language to formal language [12]. However, on multi-label student complaints, some existing preprocess are needed to be modified such as: not removing hashtag but parsing a hashtag, correcting a separate word, correcting a conjunction word, parsing suffix people pronoun and auto-correct word to formalize a sentences and make the text are ready for complaint and object identification.

Identifying a complaint sentence is done by using bigram word rule. Bigram word rule is developed by adopting bigram word rule for identifying Indonesian opinion which has been conducted by [13,14]. Part of Speech (POS) tagger is used to determine word by word tag label. Some modifications are conducted and additional rules are added to determine a new bigram word rule for multi-label student complaints.

Identifying a complaint object is done by using matching an object word with four corpuses: Personal, Relation, Subject and Institutional Agent. The four corpuses are mainly taken from IndoWordList¹ with some additional data from other resources to help object complaint identification process. IndoWordList is a dictionary which is used by Microsoft word to spell check Indonesian grammar automatically. Indonesia Natural Language Processing Toolkit (InaNLP) [15] is used to formalize sentences and to give POS tagger for each word and to lemmatize an input word.

The motivations for this research are following: to produce four Indonesian objects corpuses which related to multi-label student complaints and to produce Indonesian pattern rule for student complaints identification which can support another research in the future specially in Indonesian language. There are two main contributions of this paper. First, adding some new preprocess method for multi-label student complaints problems. Second, introducing Ina-BWR, an Indonesian bigram word rule methods to identify complaint and multi-label objects complaint.

2. Materials

2.1. Data set

In this research, two data sets are used: primary data set from Telkom University student data in the first semester and secondary data set from hashtag data. The data acquisition process is done by using Twitter API [16]. There are 151 primary students data who enter university in 2014/2015 and 2015/2016 who gave legal access license. The first semester student tweets were chosen because that semester is first transition period of students from senior high school to university, which is usually not easy [17], and senior high school had not adequately prepared the students for university in terms of the number of tasks, stress, difficulty working and understand the instructions given at the University [18]. Secondary data consists of fifty-five hashtag data, such as: #algorithm (algorithm), #penat (tired), #anakteknik (engineering student), #banyaktugas (a lot of work), #banyakpikiran (a lot of things in mind), #bosan (bored), #bingung (confused), #emosi (emotion), #malas (lazy), #mengantuk (sleepy), etc. From 4428 tweets (from 151 students) we obtained 625 relevants students tweets. Moreover, from 4736 hashtag tweets we obtained 426 relevants tweets. Therefore the total of relevants tweets is 1051.

From 625 relevant students tweets, personal label is the biggest complaint (67%), followed by subject, relation and institutional agent label (15%, 15%, 3% respectively), as shown in Fig. 1. Domination of personal problems in the first semester are personal problems (71.62%), followed by student tired and sick (18.02%), lack of sleep (7.21%) and inability to manage finance (1.35%).

For subject label, the domination of complaints in the first semester are: a lot of tasks (34%), material course difficulties (16%), full of schedules (10%) and other course activities such as a lot of laboratory work, a lot of tasks in student orientation, unpreparedness of midterm exam and final exam, and other issues (40%).

For relation label, the domination of complaints in the first semester is about: love relationship (36.73%), difficulty in making new friend relation (30.61%), homesickness (20.41%), and other issues (12.24%). For institutional agent label, the domination of complaints in the first semester is about: delay or absence of lecturer (70%), and other issues (30%).

2.2. Twitter

Twitter is one of microblogging in which people around the world publish their feelings. Twitter is the third biggest social media application around the world [2]. Twitter allows users to share their feeling in 140 up to 280 characters. Twitter offers some features such as mention, hashtag, retweet, emoticon and etc. Mention ('@') is a symbol to connect to another user in our tweet. Hashtag ('#') is a symbol to associate our tweet to any topic that can be traced by people around the world. Retweets ('RT') is a symbol to broadcast other user tweet to our followers [19].

Research using Twitter as data source has been conducted by many researchers. That research consists of sentiment analysis, emotion/mood recognition, opinion mining, topic modeling, trending topic, social network, complaint mining and mining multi-label student problems. Examples of research on sentiment analysis such as sentiment classification in Indonesian Twitter on Jakarta governor election [12], sentiment classification in Indonesian Twitter with four classes: positive, negative, neutral and question [11], and implicit sentiment expression by using similar contextual semantics and sentiment [6].

Another topics in Twitter is for emotion/mood recognition such as automatic mood classification of Indonesian Tweets using grammatical rule [20] and positive negative sentiment and seven

¹ <http://indodic.com/SpellCheckInstall.html1>

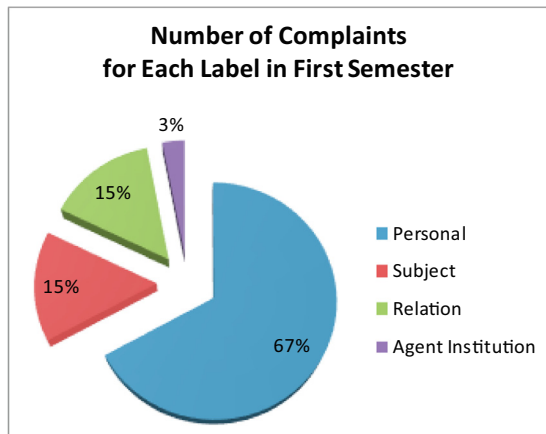


Fig. 1. Student complaint percentage for each label in the first semester.

emotion classification by using hierarchical classification [21]. One of opinion mining in Twitter has been conducted by [22] to get topic model by using intention with verb and topic with non-verb (V-BTM: Verb-Biterm Topic Model). Indra [7] compared two methods: document pivot and BN-grams for detecting Indonesian tweets trending topics. Zhao [23] analyzed Twitter data to get better understand about analytics research weeks in the Australian Department of Immigration and Citizenship (DIAC). Chen [4] used student tweets data for mining multi-label student problems in English language. To the best of our knowledge, there is no research about multi-label student problems in Indonesian language. There are some differences between English and Indonesian language such as syntax, phonemes, pronunciation, meaning, grammar, and tenses [24]. In Indonesian tweets, many informal features which appear in text such as abbreviations, interjections, foreign words, blending of Indonesian and foreign words, emoji, and etc [25]. For that reasons, many kinds of issues or challenges in multi-label student problems in Indonesian language would be explored to get more precise approach.

2.3. Educational data mining (EDM)

EDM is a research topic which implements data mining methods in educational environment. The main objective of EDM is to find out descriptive pattern that characterize learners behaviours and achievements, domain knowledge content, assessments, educational functionalities and application [26]. One of the main objectives of EDM is to predict student performance. Some research has been conducted to predict student performance with many kinds of data sources such as demographic data [27,28], academic data [29–31], demographic and academic data [32–34], demographic; academic and social network data [35] and e-learning data [36]. To the best of our knowledge, another data which has been explored to support student performance prediction model is multi-label student problems [4]. They captured student expressions in social media which related to learning experiences. Fahrudin [27] used multi-label student problems as a part of the main frame for the first year academic failure detection.

2.4. Bag of word method to solve multi-label student problems

Chen [4] developed a workflow to integrate both qualitative analysis and large-scale data mining to solve multi-label student problems using Bag of Word method at Purdue University. They added hashtag data #engineeringproblem to get additional related

data. Five prominent categories of student problems are defined: heavy study load, lack of social engagement, negative emotion, sleep problems, and diversity issues. They used multi-label classification method because one tweet can contain more than one category (y_1, y_2, \dots, y_n) with $n \geq 1$.

Multi-label classification in [4] used problem transformation methods as one of the multi-label learning method (MLL) [37]. Binary Relevance is used as problem transformation approach. The Binary Relevance classifies each data for each label class separately, and then combines into one multi-label prediction result. The Binary Relevance assumes each label class independently. According to [4], Naïve Bayes achieves best accuracy result compared to SVM and M3L in multi-label student problems using Bag of Word method.

Bag of Word method which is used in [4] has some weakness such as missing word order, neglecting grammar, neglecting number, etc. Table 1 gives two examples of error in the BoW method for multi-label student complaints. Therefore, semantic method using bigram word rule and four corpuses are proposed to correct that error.

2.5. Opinion mining

Opinion mining is a technique for extracting, classifying, understanding, and assessing the opinions expressed in various user-generated content [38]. The objectives in opinion mining is to classify text as an opinion or fact [13]. Research in Indonesia opinion mining has been conducted by [13,14]. There are 3 subprocesses in that research: document subjectivity, opinion orientation and target detection. Document subjectivity is sub process to identify a sentence is opinion or not. Opinion orientation is sub process which aims to give opinion orientation (positive or negative). Target detection is a part of opinion mining process for recognizing the target object of the opinion. One of the results of [13] and [14] is a list of bigram rule for document subjectivity sub process. In this research nineteen bigram rules from [13] as in Table 2 are used.

2.6. InaNLP

InaNLP is a toolkit for natural language processing both on formal and informal Indonesian language [15]. InaNLP has nine modules for text processing such as sentence splitter, tokenization, word formalization, morphologically analyzer (stemmer), POS tagger, phrase chunker, named entity tagger, syntactic parser, and semantic analyzer. Several modules were built using rule based approach and others using statistical based approach. InaNLP modules can be accessed independently. The accuracies of InaNLP are 93.41% for name entity (NE) tagger and 96.5% for POS Tagger [15].

Table 1

Two examples of error in BoW method for multi-label student complaints.

Tweets examples	Multi-label results
saya bersyukur, ada pretest yg sengaja ngga mereka ikuti. mereka malas. kesempatan besar untuk bisa tahu kualitas dan motivasi mereka	BoW: Personal & Relation
(I am grateful; there is a pre-exam that they do not intentionally follow. They are lazy , a great opportunity to know their quality and motivation)	Semantic: Relation
kamu tau gak rasanya jadi aku sekarang ini pusing mikir kamu	BoW: Personal & Relation
(You know, currently I feel so dizzy thinking of you)	Semantic: Relation

Table 2
Rule for opinion detection.

No	Rule	Examples
1	RB JJ	sangat buruk (very bad)
2	RB VB	semoga berjalan (hopefully work)
3	NN JJ	buku bagus (good book)
4	NN VB	perkataannya menjengkelkan (his words are annoying)
5	JJ VB	cepat memahami (quickly understand)
6	CK JJ	bagus atau baik (nice or good)
7	JJ BB	sama bagus (equally good)
8	VB VB	membikin pusing (make a headache)
9	JJ RB	indah sekali (very beautiful)
10	VB JJ	membikin bingung (making confuse)
11	NEG JJ	tidak mudah (not as easy)
12	NEG VB	tidak mengerti (do not understand)
13	PRP VBI	Saya menyukai (i like)
14	PRP VBT	kita suka (we like)
15	VBT NN	Memiliki kedekatan (have closeness)
16	MD VBT	Perlu mengambil referensi (need to take reference)
17	MD VBI	Perlu dikembangkan (need to be developed)
18	UH VBP	Tolong dicat (please painted)
19	JJ VBP	Mudah diterima (easily accepted)

There are some studies have been conducted using InaNLP such as: A tweets classification for complaints messages [15], Automatic multi-label classification for Indonesian News Articles [39], Indonesian Essay Grading Module using Natural Language Processing [40], Integrated Social Media Knowledge Capture Model in Medical Domain of Indonesia [41], Dominant Emotion Recognition in Short Story Using Keyword Spotting Technique and Learning Based Method [42], Aspect Based Sentiment Analysis for Review Rating Prediction [43].

In this research, word formalization, POS tagger and stemmer modules are used. Word formalization module is used to formalize informal sentences to formal sentences. POS Tagger is used to tag a label for each word in a tweet. Stemmer is used to lemmatize a word and provides its lemma and affixes of the word.

3. Proposed method

The proposed method aims to improve Bag of Word multi-label classification method which was used in [4] in two contributions: transforming informal conversation into formal conversation and improving bag of word method in multi-label student complaint tweets. To the best of our knowledge, this BoW method is state of the art for feature representation in multilabel classification [44] and very effective for document classification [45]. In the first contribution, informal conversation patterns in our data are analyzed and additional preprocessing steps to solve those informal conversation pattern problems are added. In the second contribution, Ina-BWR uses semantic approaches to solve incorrect word order and grammatical error in bag of word method.

3.1. Additional text preprocessing in Ina-BWR

In text mining, preprocess is a necessary step before a text is processed by information retrieval method. In Indonesian language, preprocessing informal text to formal text can reduce Out Of Vocabulary condition, so it can improve classification accuracy by 3–15% [11]. Table 3 shows examples of words before and after preprocessing methods were conducted.

Parsing hashtag is an additional process to parse concatenated word in hashtag, such as: #banyaktugas into “banyak tugas” (a lot of tasks), #banyaktgsberat into “banyak tugas berat” (a lot of heavy tasks), #fluberat into “flu berat” (severe flu), etc. Parsing process is done by using IndoWordList and formalizationDict dictionary. An algorithm has been developed to parse a hashtag word.

Table 3
Examples of informal student complaint tweets before and after preprocessing methods were conducted.

Input	Process	Output
Derita seorang murid #banyaktgs (an suffering student #manytasks)	Parsing hashtag	Derita seorang murid banyak tugas (an suffering student many tasks)
hari yang di penuhi quiz (a day fulfills with a quiz)	Correcting affixes word	hari yang dipenuhi quiz (a day fulfills with a quiz)
Mata dan tugas tidak bekerjasama (eye and tasks are not cooperated)	Correcting word conjunction	Mata tidak bekerjasama , tugas tidak bekerjasama (eye is not cooperated , tasks is not cooperated)
Aku mulai tersadar bahwa aku membutuhkannya (I start realized that I need it)	Removing Stop word	Aku tersadar aku mebutuhkannya (I realized I need it)
Kamu pergi meninggalkanku (you leave me)	Parsing suffix on people's pronoun	Kamu pergi meninggalkan aku (you leave me)
6 menit lagi desen gak masuk , libur?? (6 minute again lecturer not coming , holiday??)	Correcting typo words	6 menit lagi dosen gak masuk , libur?? (6 minute again lecturer not coming , holiday??)

This algorithm tries to separate a word in hashtag into appropriate words in dictionary by using consecutive character comparison. The separation process is done if appropriate word which has same word length and same order of character is found. Fig. 2 shows the examples of parsing process for #banyaktgs.

Limitation in the number of characters in a tweet may force user to abbreviate some words such as in affix word, examples: “d pndh” (should be “dipindah”). Furthermore, the mistake may happen in using affix word in Indonesian languages, such as: “di kerjakan” (should be “dikerjakan”). That mistake may cause Out of Rule condition because word can not be found in dictionary so it can not match our complaint bigram word rules. The word concatenation process for separate connecting word is done by using IndoWordList dictionary assistance. If these separate words are found, dictionary is used and the concatenation process is conducted.

Conjunction word is a word which connect two separate words. One of it's functions is additive function, which is connecting two separate words which have same position. One examples of additive conjunction is “and”, i.e., “mata dan tugas tidak bekerjasama (eye and task are not cooperated)”. There are two objects inside sentence, eye and task. Therefore, preprocess is needed to separate two objects in one sentence into two separate sentences, so identification process in Ina-BWR will find two complaints with two objects.

In the student complaints identification processes, one of objects which need to be identified is a people's pronoun, such as: dia (him) and mereka (them). Aku (i), saya (me), diriku (myself) are the examples of personal corpuses. Meanwhile, dirimu (yourself), dirinya (him), kamu (you) are the examples of relation corpuses. So, people's pronoun is needed to be separated to identify object based on corpuses. InaNLP stemmer is used to separate suffix on people's pronoun of a word. Table 4 shows illustration about how to separate suffix on people's pronoun.

The auto-correction process is performed on a word which is not exists in dictionary list. The smallest levensthein distance is used to select the best appropriate candidate. Levensthein is an algorithm which compare two string: source string (ss) and destination string (dt), and then calculates deletion, insertions or substitution step which replaces (ss) to (dt) [46].



Fig. 2. Parsing hashtag illustration.

Table 4 Separating suffix on people's pronoun illustration.

Possesive pronoun suffix	Suffix people pronoun	Examples
-ku	Aku (me)	pergi meninggalkanku >>> pergi meninggalkan aku (leave me)
-mu	Kamu (you)	bosan menunggumu >>> bosan menunggu kamu (boring wait for you)

3.2. Bigram word concept

Bigram is one of n-gram concepts which is used in most of NLP domain [47]. N-gram has two terminologies: n-gram character and n-gram word. N-gram word is contiguous sequences of n words, meanwhile N-gram character is contiguous sequences of n character [48]. In this research, bigram word is used. This is an example of bigram word which consists of five words “Dompot makin menipis haduhhhhh Pusing (run out of money, i feel dizzy)”. There are (n-1) combinations of bigram word, such as:{dompot, makin}, {makin, menipis}, {menipis, haduhhhhh}, {haduhhhhh, pusing}. Complaint identification process and object identification process are conducted by using those combinations.

3.3. Ina-BWR

Ina-BWR consists of three process: preprocessing informal text, identifying complaint and identifying object. Fig. 3 shows Ina-BWR block diagram.

Identifying complaint process is performed by adopting Hamzah bigram rules [13] for document subjectivity process in Indonesian opinion mining. Some stop words in Ina-NLP are preserved such as: sangat (very), banyak (many), selalu (always), boleh (can), tidak (not). Those words are needed in Ina-BWR (as shown in Table 5).

Identifying opinion is performed by using POS tagger result from InaNLP and some additional corpuses, such as: verb (943 words), noun (1386 words) and adjective (167 words) which have negative meanings. Those corpuses were selected and taken man-

ually from IndoWordList dictionary and the Indonesian Medical Council for additional list of diseases. Table 6 shows list examples of our corpuses.

Complaint identification process is started by tagging word of each tweet. After that, bigram tokenization processed is employed. Furthermore, those bigram tokens are paired with complaint bigram word rules. The Final results of this process are a lists and index positions of complaints bigram word rules.

After complaints bigram word rules and index positions are founded, object and its location are identified. There are three locations of object complaints: inside bigram word rule, in front of bigram word rule, and behind bigram word rule. To help identify the object, four object complaints corpuses are manually arranged from IndoWordList dictionary (Personal (47 words), Subject (3480 words), Relation (93 words) and Institutional Agent (2120 words)). For subject corpuses, additional corpuses data from subject names and its abbreviations are added. While some lists of lecturer code are added to institutional agent as additional corpuses data. Table 7 shows examples of complaints matching process. Table 8 shows examples of each label corpuses: Personal, Subject, Relation and Institutional Agent. Table 9 shows illustration of object identification process in Ina-BWR.

Object identification process is conducted by matching object using four corpuses (Fig. 4). There are three patterns for object locations based on POS tagger results which can be seen in Table 10.

4. Experiment and results

In this section, experiment scenarios to observe Ina-BWR method and BoW method are conducted. For BoW method, two observations scenarios are observed: without additional preprocessing and with Ina-BWR additional preprocessing (Fig. 5). For each scenarios, Unigram and Bigram Bag of Word are applied. According to [4], BoW method in multi-label using binary relevance transformation approach gives best performance in Naïve Bayes, so Naïve Bayes is chosen as BoW base classifier in this experiment. To investigate the performance of Ina-BWR and BoW method, 10-fold cross validation is employed. Last, combination of Ina-BWR and BoW method also examined.

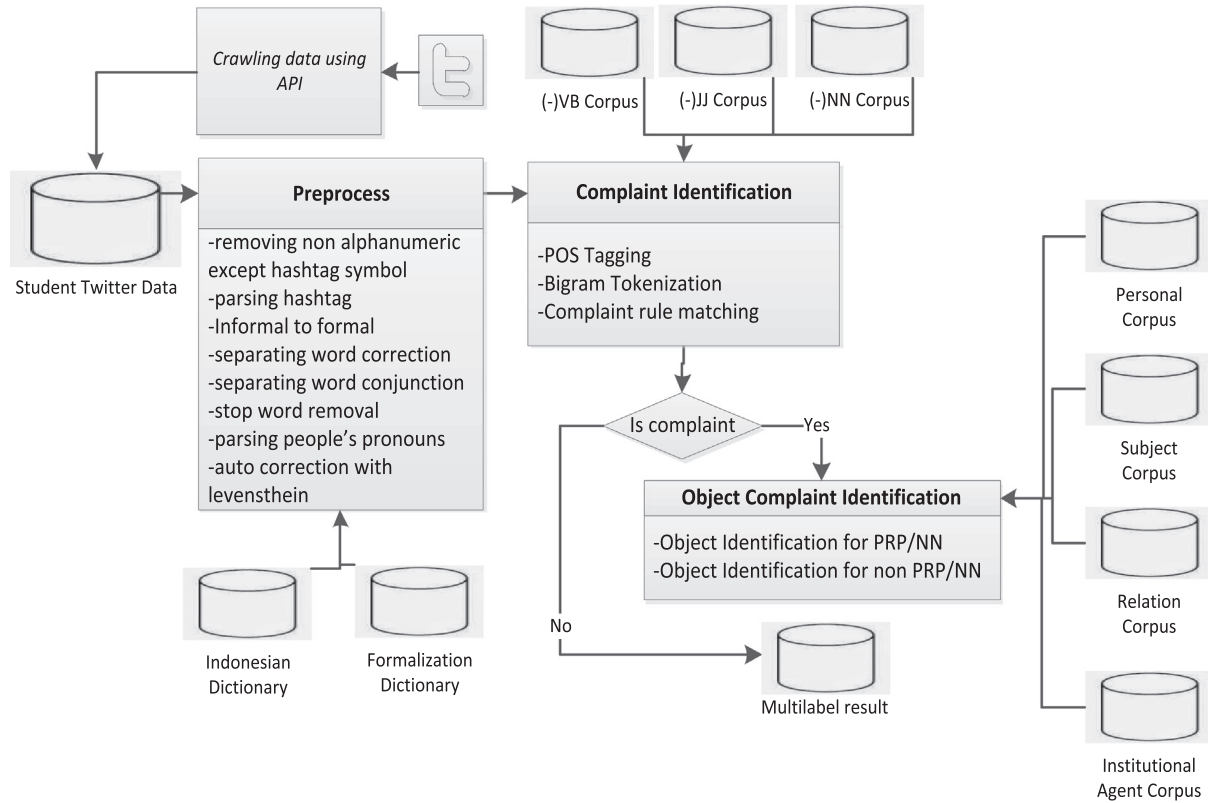


Fig. 3. Ina-BWR Block Diagram.

Table 5
Bigram word rules modification for complaint.

No	Rule	Modification Rule	Examples in Opinion	Examples in Complaints
1	RB JJ	RB (-)JJ	sangat bagus (very good)	sangat buruk (very bad)
2	RB VBT	RB (-)VBT	semoga berjalan (hopefully work)	selalu mengalah (always succumbs)
3	NN JJ	NN (-)JJ (-)JJ NN JJ (-)NN (-)NN JJ	bukunya bagus (his book is good) bagus bukunya (good book) - -	selalu mengalah (the course is difficult) sulit matakuliahnya (difficult course) banyak hambatan (many obstacles) hambatan banyak (many obstacles)
4	NN VB	NN (-)VB	perkataannya menyenangkan (he said pleasantly)	pelajarannya membosankan (the lessons are boring)
5	JJ VB	(-)JJ VB JJ (-)VB	mudah dipahami (easy to understand) -	sulit dipahami (difficult to understand) mudah mengumpat (easy to curse)
6	VB VB	VB (-)VB	membuat begadang (make stay up)	membuat begadang (make stay up)
7	JJ RB	(-)JJ RB	indah sekali (very beautiful)	jelek sekali (really ugly)
8	VB JJ	VB (-)JJ	membikin bingung (make confused)	membikin bingung (make confused)
9	JJ RB	(-)JJ RB	indah sekali (very beautiful)	jelek sekali (really ugly)
10	VB JJ	VB (-)JJ	membikin bingung (make confused)	membikin bingung (make confused)
11	NEG JJ	NEG JJ	tidak mudah (not easy)	tidak mudah (not easy)
12	NEG VB	NEG VB	tidak mengerti (don't understand)	tidak mengerti (don't understand)
13	PRP VBI	PRP (-)VBI	kita suka (we like)	kita benci (we hate)
14	PRP VBT	PRP (-)VBT	saya menyukai (i like)	saya membenci (i hate)
15	VBT NN	(-)VBT NN VBT (-)NN	memiliki kepekaan (have sensitivity) -	membenci pelajaran (hates lessons) membuat kegaduhan (make rowdy)
16	MD VBI	MD (-)VBI	perlu dikembangkan (need to be developed)	boleh berkilah (may argue)
17	-	PRP (-)JJ	-	mereka malas (they are lazy)

In multi-label, evaluation measures can be divided into two kinds: example-based measures and label-based measures [49]. Accuracy in example-based measures is calculated on each tweet and then averaged over all tweets. While in label-based measures, accuracy is calculated for each label and then averaged over all labels. The contingency table for each category label (a) can be seen in Table 11.

4.1. Example-based evaluation measures

For one tweet (t) which has actual label C and prediction label D, the accuracy of n tweets is D divided by the number of labels in the union of C ∪ D. Precision accuracy is obtained from correctly predicted number of label divided by the total number of D. Recall accuracy is obtained from correctly predicted number of label

Table 6
List examples of our corpuses.

(-) VB	(-) NN	(-) JJ
ancam (intimidate)	alergi (allergy)	aneh (strange)
begadang (stay up)	amnesia (amnesia)	arogan (arrogant)
benci (hate)	beban (load)	bau (smell)
bentak (snapped)	cemooh (scorn)	bencong (sissy)
bentrok (clashing)	demam (fever)	bengkak (swollen)
berbohong (lie)	halangan (obstacles)	beringas (violent)
.....
tidur (sleep)	ujian (exam)	sulit (difficult)

Table 7
Examples of complaint bigram word rules matching process.

Sentences after text preprocessing	Bigram word list rule	Index position of bigram world rule
Lagi musim sakit aku harap cukup aku sakit (sick season, i hope no one got sick except me)	NN (-)JJ = musim sakit (-)JJ NN = sakit aku NN (-)JJ = aku sakit	1 2 6
Lama mengerjakan tugas depan laptop lelah mata aku (doing task in front of the laptop for a long time,my eyes was tired)	(-)JJ VBT = lama mengerjakan (-)JJ NN = lelah mata	0 5
Lari lari kecil dari asrama sampai ruangan dosen tidak ada (jogging from student dormitory to lecturer room but lecturer does not exist)	NEG VBI = tidak ada	8

Table 8
Examples of Personal, Subject, Relation and Institutional Agent corpuses.

Personal	Subject	Relation	Institutional Agent
aku (me)	absensi (presence)	abah (father)	aab (aab)
badan (body)	akuntansi (accounting)	abang (brother)	aac (aac)
bibir (lip)	agama (religion)	adek (sister)	admisi (admission)
buku (book)	alin (linear algebra)	adik (sister)	akademik (academic)
dada (chest)	alpro (programming algorithm)	adinda (sister)	almamater (alma mater)
diri (self)	hafalan (recitation)	dia (you)	asisten (assistant)
mata (eye)	jadwal (schedule)	dirimu (yourself)	dosen (lecturer)
tangan (hand)	tugas (task)	suami (husband)	rector (rector)
tenggorokan (throat)	uts (midterm exam)	teman (friend)	senat (senate)
tulang (bone)	uas (final exam)	wanita (woman)	wakil rector (vice rector)

divided by the total number of C. So, for total T tweets, which derive from $t_1, t_2, t_3, \dots, t_n$, the accuracy, precision, recall and f-measure are calculated using Eqs. (1)(4) respectively.

Table 9
Examples of object identification process.

Sentences after text preprocessing	Object identification	Labelization
lagi musim sakit aku harap cukup aku sakit (sick season, i hope no one got sick except me)	1, object = - 2, object = aku 6, object = aku	Personal = 1 Subject = 0 Relasi = 0 Agent Institution = 0
lama mengerjakan tugas depan laptop lelah mata aku lelah (doing task in front of the laptop for a long time, my eyes was tired)	0, object = tugas 5, object = mata	Personal = 1 Subject = 1 Relasi = 0 Agent Institution = 0
lari lari kecil dari asrama sampai ruangan dosen tidak ada bangk*k (jogging from student dormitory to lecturer room but lecturer does not exist)	8, object = dosen	Personal = 0 Subject = 0 Relasi = 0 Agent Institution = 1

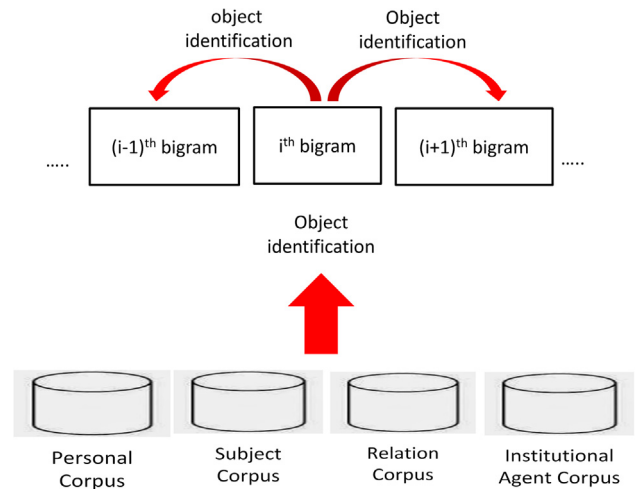


Fig. 4. Identification object process.

$$Accuracy(a) = \frac{1}{n} \sum_{i=1}^n \frac{C_i \cap D_i}{C_i \cap D_i} \quad (1)$$

$$Precision(p) = \frac{1}{n} \sum_{i=1}^n \frac{C_i \cap D_i}{D_i} \quad (2)$$

$$Recall(r) = \frac{1}{n} \sum_{i=1}^n \frac{C_i \cap D_i}{C_i} \quad (3)$$

$$Fmeasure(Fm) = \frac{1}{n} \sum_{i=1}^n \frac{2 \cdot p_i \cdot r_i}{p_i + r_i} \quad (4)$$

4.2. Label-based evaluation measures

In label-based, performance is calculated for each label. So, using matrix in Table 11, the accuracy, performance, recall and f-measure is calculated for each label (l) from total m label using Eqs. (5)(8) respectively.

$$Accuracy(a_l) = \frac{TP_l + TN_l}{TP_l + FN_l + FP_l + TN_l} \quad (5)$$

$$Precision(p_l) = \frac{TP_l}{TP_l + FP_l} \quad (6)$$

Table 10
Location of object based on POS Tagger.

POS Tagger	Bigram Identification	Examples / Bigram Rule	Object
PRP / NN	In the i^{th} bigram observation	mereka malas (PRP (-)JJ) (they are lazy) kita benci (PRP (-)VBI) (we hate) saya membenci (PRP (-)VBT) (I hate) matakuliahnya sulit (NN (-)JJ) (that course is difficult) sulit matakuliahnya ((-)JJ NN) (difficult course) pelajarannya membosankan (NN (-)VB) (the lessons are boring)	mereka kita saya matakuliah matakuliah pelajaran
All	In front of the i^{th} bigram observation	dosen jarang muncul ((-)NN VBT) (lecturer appear infrequently) mata dan tugas tidak bekerja sama (NEG VBI) (eye and task not cooperated) kuliah makin berat jelang pra-thesis (RB (-)JJ) (course getting heavier toward pre-thesis)	dosen mata & tugas kuliah
All	Behind the i^{th} bigram observation	gagal lagi aku ((-)JJ RB) (I failed again) mudah mengumpat dia sekarang (JJ (-)VBT) (he is easily curse now) panas banget badanku ((-)JJ NN) (my body get fever)	aku dia badan

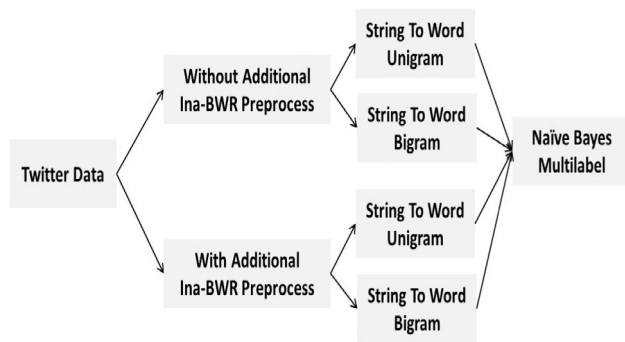


Fig. 5. Two Bag of Word experiments scenario.

Table 11
Contingency label.

	True a	True not a
Predicted a	TP_a	FP_a
Predicted not a	FN_a	TN_a

$$Recall(r_i) = \frac{TP_i}{TP_i + FN_i} \quad (7)$$

$$Fmeasure(Fm_i) = \frac{2 \cdot p_i \cdot r_i}{p_i + r_i} = \frac{2 \cdot TP_i}{2 \cdot TP_i + FP_i + FN_i} \quad (8)$$

For m label, there are two kinds of f-measured: micro-averaged Fmeasure and macro-averaged Fmeasure. Micro averaged works well on large categories, otherwise macro average works well on small categories [4].

$$Micro_{avg}Fm = \frac{2 \cdot \sum_{j=1}^m TP_j}{2 \cdot \sum_{j=1}^m TP_j + \sum_{j=1}^m FP_j + \sum_{j=1}^m FN_j} \quad (9)$$

$$Macro_{avg}Fm = \frac{1}{m} \sum_{j=1}^m \frac{2 \cdot TP_j}{2 \cdot TP_j + FP_j + FN_j} \quad (10)$$

5. Results

The performance results of each scenario can be seen in Tables 12–14. The best results for each evaluation measures from those tables are highlighted in bold typeface. For each evaluation measures, the result of Ina-BWR and Best BoW combination is also evaluated.

Table 12 shows evaluation results of examples based evaluation. Additional preprocessing of Ina-BWR can improve Unigram BoW accuracies for the accuracy, precision, recall and f-measure by up to 6.27%, 6.53%, 6.79% and 6.67% respectively. Additional preprocessing of Ina-BWR also improve accuracies of Bigram BoW method for the accuracy, precision, recall and f-measure by up to 2.4%, 1.85%, 2.28% and 2.07% respectively. While Ina-BWR compared to the best BoW method (Unigram BoW with additional preprocessing of Ina-BWR scenario) gives precision, recall and f-measure improvement by up to 10.47%, 2.84% and 6.3% respectively. Combination of Ina-BWR and Unigram BoW with Ina-BWR additional preprocessing compared to the best BoW method increase the accuracies for the precision, recall and f-measure by up to 10.54%, 2.89% and 6.35% respectively. Both Bigram BoW scenarios (with and without Ina-BWR additional preprocessing) give lower accuracies compared to Unigram BoW.

Table 13 presents average evaluation results from four labels on label based evaluation measures. Ina-BWR additional preprocessing can improve average Unigram BoW average evaluation results of accuracy, precision, recall, f-measure micro and f-measure macro by up to 3.3%, 10.31%, 5.51%, 6.01% and 8.82% respectively. For Bigram BoW scenario, Ina-BWR additional preprocessing can improve average of accuracy, precision, recall, f-measure micro and f-measure macro by up to 1.01%, 25.06%, 3.39%, 2.24% and 6.41% respectively. While Ina-BWR compared to the best BoW method shows average improvement of accuracy, precision and f-measure micro by up to 1.7%, 6.18% and 0.8% respectively. Combination of Ina-BWR and Unigram BoW with Ina-BWR additional preprocessing compared to the best BoW method (Unigram BoW with Ina-BWR additional preprocessing scenario) also increase the average of accuracy, precision, f-measure micro and f-measure macro by up to 2.18%, 7.4%, 2.01% and 0.95% respectively.

Table 14 shows accuracies for each label. The improvement of each label in Unigram BoW method with Ina-BWR additional pre-

Table 12
Example based evaluation results.

Scenario	Accuracy	Precision	Recall	F-Measure
BoW Unigram without preprocess	0.6146	0.5672	0.5420	0.5542
BoW Bigram without preprocess	0.5636	0.5489	0.5484	0.5486
BoW Unigram with preprocess	0.6773	0.6325	0.6099	0.6209
BoW Bigram with preprocess	0.5881	0.5674	0.5712	0.5693
Ina-BWR	0.6597	0.7372	0.6383	0.6842
Ina-BWR + Best BoW	0.6655	0.7379	0.6388	0.6844

Table 13
Label based evaluation results.

Scenario	Accuracy avg	Precision avg	Recall avg	F-M micro avg	F-M macro avg
BoW Unigram without preprocess	0.8165	0.5837	0.6781	0.6570	0.6145
BoW Bigram without preprocess	0.7850	0.2723	0.2792	0.5769	0.2328
BoW Unigram with preprocess	0.8495	0.6868	0.7332	0.7171	0.7027
BoW Bigram with preprocess	0.7951	0.5229	0.3131	0.5993	0.2969
Ina-BWR	0.8670	0.7486	0.6034	0.7251	0.6539
Ina-BWR + Best BoW	0.8713	0.7608	0.6839	0.7372	0.7122

Table 14
Accuracy for each label.

Scenario	Personal	Subject	Relation	Institutional Agent
BoW Unigram without preprocess	69.05%	82.98%	80.54%	94.03%
BoW Bigram without preprocess	59.37%	77.65%	81.73%	95.26%
BoW Unigram with preprocess	72.52%	84.10%	86.07%	97.07%
BoW Bigram with preprocess	60.96%	79.63%	82.10%	95.35%
Ina-BWR	75.29%	89.41%	86.75%	95.92%
Ina-BWR + Best BoW	75.29%	89.41%	86.75%	97.07%

Table 15
Out of Vocabulary words.

Scenario	OOV words
BoW Unigram without preprocess	252
BoW Bigram without preprocess	748
BoW Unigram with preprocess	128
BoW Bigram with preprocess	596
Ina-BWR	25
Ina-BWR + Best BoW	128

processing for personal, subject, relation and institutional agent are 3.47%, 1.12%, 5.53% and 3.04% respectively. For Bigram BoW scenario, Ina-BWR additional preprocessing can improves accura-

cies of personal, subject, relation and institutional agent by up to 1.59%, 1.98%, 0.37% and 0.09% respectively. While Ina-BWR compared to the best BoW method (Unigram BoW with Ina-BWR additional preprocessing scenario) shows improvement of accuracies for personal, subject and relation by up to 2.77%, 5.31% and 0.68% respectively. Combination of Ina-BWR and Unigram BoW with Ina-BWR additional preprocessing gives the best accuracy for all labels.

Table 15 shows the number of out of vocabulary (OOV) words for both BoW and Ina-BWR methods. For BoW method, the additional Ina-BWR preprocessing can decrease the OOV words up to 49.20% for BoW Unigram scenario and 20.32% for BoW Bigram scenario. The OOV number of BoW Bigram is larger than BoW Unigram, so it can causes the accuracies of BoW Bigram are lower than BoW Uni-

Table 16
Error examples of comparison methods.

Method	Examples error in predicting tweets	Reasons of errors
BoW without preprocess	#curhatanmahasiswaabanyaktugas #tugasmenumpuk	Can't parse complaints text inside hashtag
	(#sharestudentfeelinglotoftasks #alotoftasks) Nunggu dosen #kampus #dosentelat (wait lecturer #campus #latelecturer) tugasnya nupuk (the task is stack)	Can't parse complaintstext inside hashtag typo error (nupuk should be numpuk)
BoW with preprocess	#Pusing pikiran ... gak butuh teman (#headache mind ... does not need friends) aqu jenuh dan #lelah (I am saturated and #tired) bingung, orang dibantuin tapi tidak ada usaha, capek (confuse person is helped but no effort tired)	Wrong label detected (predicted label: personal) Wrong label detected (predicted label: personal, relation) Wrong label detected (predicted label: personal)
	Ina-BWR gile aje, psikotes 5 jam (it's crazy, psychology test 5 h) Habiskan masa mudamu dengan ... ngerjain tugas (spend your youth with ... doing task) nambah tugas lagi yang lain, rekreasi gagal (other tasks increase, recreation failed)	opinion aspects in number implicit complaints object is located far away from bigram word

gram in all scenarios (with or without additional preprocess). InaBWR gives lowest OOV words because its corpus derived from indowordlist dictionary directly and not depend on training data such as BoW method. Besides word normalization, other strategies are needed to be conducted on BoW method to improve accuracies such as synonym matching and replace OOV words by the best similar word using word similarity methods.

6. Discussion

Generally, Ina-BWR method is able to recognize multi-label student complaints in Indonesian language better than BoW method because of two folds (as shown in Table 16). First, Ina-BWR additional preprocessing can parse and correct informal words so the OOV will be decreased and accuracy will be increased. Second, simple complaint identification and object complaint identification process by using bigram word rule can improve accuracy. It can be seen that Ina-BWR performance results better than BoW in almost all accuracies metric except in example based accuracy, average recall on label based accuracy and institutional agent accuracy. Furthermore, combination of Ina-BWR method with the best BoW method (Unigram BoW with Ina-BWR additional preprocessing scenario) can improve Institutional Agent accuracy.

Some mistakes of BoW without additional preprocessing method such as can't parse complaints bigram word rule and typo error can be handled by adding additional preprocessing method. This additional preprocessing method can increase accuracy improvement by up to 3.47%, 1.12%, 5.53% and 3.04% for personal, subject, relation and institutional agent respectively. Two of the greatest accuracy improvements are for relation and personal labels. That because, in BoW without additional preprocessing method gives more false positive prediction in relation and personal label, while in the BoW with additional preprocessing method that false prediction is corrected. Decreasing accuracy on institutional agent label in BoW with additional preprocessing method is caused by error in false positive institutional agent prediction after parsing word in hashtag in BoW method. But that can be handled by using combination of Ina-BWR and the best BoW method (Unigram BoW with Ina-BWR additional preprocessing scenario).

The approach in Ina-BWR method is generally appropriate for other alphabetic languages for some reasons:

- Rules for complaint detection such as RB (-)] = sangat buruk = very bad (english) = sehr schlecht (german) = totemo warui (japan) are applicable for any alphabetic languages.
- Some preprocessed techniques such as parsing hashtag, correcting word conjunction, removing stop words and correcting typo words are universal techniques which can be used in other languages.
- Lists of corpus for complaint identifications and object identification can be adjusted to other languages.
- Object identification using index position of bigram is a common technique, therefore can be applied directly in other languages.

There are some limitations of Ina-BWR method which needed to be conducted on future works:

- It can not handle implicit opinion. Example: “gile aje, **psikotes 5 jam** (it's crazy, **psychology test 5 h**)”. In that sentence, the speaker regard 5 h is too much for doing psychology test.
- It can not handle object (gram word) that are separated by one or more words. Example: “nambah **tugas** lagi yang lain,

rekreasi gagal (other **tasks** increase, **recreation failed**)”. Words that are typed in bold are bigram word complaints and the object of complaints is located in fourth bigram before bigram word complaints.

7. Conclusion

The transition from high school to university gives a special experience for student in the first semester. Many complaints on student tweets in first semester show that the students need a social media support, so they can survive in new environment. In this research, four categories are conducted based on seven academic success factors: Personal, Subject, Relation and Institutional Agent. A new semantic methods, Ina-BWR has been proposed to improve BoW method on two contributions: 1) adding additional text preprocessing such as parsing a hashtag, correcting a separate word, correcting a conjunction word, parsing suffix people pronoun and auto-correcting word 2) adopting bigram word rule in Indonesian opinion identification and manual corpus assistance. Generally, two contributions in Ina-BWR can improve accuracy in all accuracy metric except on institutional agent. But this drawback can be solved by combining Ina-BWR and the best BoW method (Unigram BoW with Ina-BWR additional preprocessing scenario) which.

Limitation points of Ina-BWR in Section 6 can be other issues which are left to be considered in the future works:

- Improving Ina-BWR which can handled some limitations as mentioned in Section 5.
- Integrating our Ina-BWR result with other educational data to predict first year evaluation study.
- Using name entity recognition as alternative for object identification process in Ina-BWR.

Declarations of interest

None.

References

- [1] Kemp S. Digital snapshot: internet and social media use in 2017;2017:1. . <https://www.techinasia.com/talk/digital-snapshot-internet-social-media-2017> (accessed March 27, 2018).
- [2] Lua A. Social media trends report: 10 key insights into the present and future of social media. Buffer Soc 2018;2018:1. . <https://blog.bufferapp.com/social-media-trends-2018> (accessed March 27, 2018).
- [3] kominfo. Kominfo: Pengguna Internet di Indonesia 63 Juta Orang. Web Kominfo 2013. https://kominfo.go.id/index.php/content/detail/3415/Kominfo+%3A+Pengguna+Internet+di+Indonesia+63+Juta+Orang/0/berita_satker (accessed May 23, 2018).
- [4] Chen X, Vorvoreanu M, Madhavan K. Mining social media data for understanding students' learning experiences. IEEE Trans Learn Technol 2014;7:246–59. doi: <https://doi.org/10.1109/TLT.2013.2296520>.
- [5] Denker KJ, Manning J, Heuett KB, Summers ME. Twitter in the classroom: modeling online communication attitudes and student motivations to connect. Comput Human Behav 2018;79:1–8. doi: <https://doi.org/10.1016/j.chb.2017.09.037>.
- [6] Saif H, He Y, Fernandez M, Alani H. Semantic patterns for sentiment analysis of twitter. Proc. 13th Int. Semant. Web Conf. - Part II, 8797. p. 324–40. doi: https://doi.org/10.1007/978-3-319-11915-1_21.
- [7] Indra Winarko E, Pulungan R. Trending topics detection of Indonesian tweets using BN-grams and Doc-p. J King Saud Univ – Comput. Inf Sci 2018;324–9. doi: [10.1016/j.jksuci.2018.01.005](https://doi.org/10.1016/j.jksuci.2018.01.005).
- [8] Anggareska D, Purwarianti A. Information extraction of public complaints on Twitter text for bandung government. Proc. 2014 Int. Conf. Data Softw. Eng. ICODSE 2014. doi: <https://doi.org/10.1109/ICODSE.2014.7062658>.
- [9] Zhang A, Aasheim CL. Academic success factors: an IT student perspective. J Inf Technol Educ Res 2011;10:309–31.
- [10] El-din DM. Enhancement bag-of-words model for solving the challenges of sentiment analysis. Int J Adv Comput Sci Appl 2016;7:244–52.
- [11] Naradhupa AR, Purwarianti A. Sentiment classification for Indonesian message in social media. Proc. 2011 Int. Conf. Electr. Eng. Informatics. p. 1–4. doi: <https://doi.org/10.1109/ICCCSN.2012.6215730>.

- [12] Fatyanosa TN, Bachtiar FA. Classification method comparison on Indonesian social media sentiment analysis. 2017 Int. Conf. Sustain. Inf. Eng. Technol., p. 310–5. doi: <https://doi.org/10.1109/SIET.2017.8304154>.
- [13] Hamzah A, Widyastuti N. Document subjectivity and target detection in opinion mining using HMM POS-Tagger. Proc. 2015 Int. Conf. Inf. Commun. Technol. Syst. ICTS 2015. p. 83–7. doi: <https://doi.org/10.1109/ICTS.2015.7379876>.
- [14] Rozi IF, Pramono SH, Dahlan EA. Implementasi Opinion Mining (Analisis Sentimen) untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi. Electr Power. Electron Commun Control Informatics Semin 2012;6:37–43.
- [15] Purwarianti A, Andhika A, Wicaksono AF, Afif I, Ferdian F. InaNLP: indonesia natural language processing toolkit case study: complaint tweet classification. Int Conf Adv Informatics Concepts, Theory Appl 2016;2016:5–9.
- [16] Twitter API n.d.
- [17] Brinkworth R, Ben Mccann AE, Matthews AEC, Nordstro AEK. First year expectations and experiences: student and teacher perspectives. High Educ 2009;58:157–73. doi: <https://doi.org/10.1007/s10734-008-9188-3>.
- [18] Away EAH, Kelloway KR. First Year at University: perceptions and experiences of students. Can J High Educ 1987;XVII:47–63.
- [19] Mosley RC. Social media analytics: data mining applied to insurance Twitter posts. Casualty Actuar Soc E-Forum 2012;2:1–36.
- [20] Alva WVE. Automatic mood classification of Indonesian tweets using linguistic approach. 2013 Int. Conf. Inf. Technol. Electr. Eng., IEEE; 2013. p. 41–6.
- [21] Esmin Jr Ahmed AA, De Oliveira Jr Roberto L, Matwin S. Hierarchical classification approach to emotion recognition in twitter. 2012 11th Int Conf Mach Learn Appl. p. 381–5. doi: <https://doi.org/10.1109/ICMLA.2012.195>.
- [22] Lu T, Hou S, Chen Z, Cui L, Zhang L. An intention-topic model based on verbs clustering and short texts topic mining. Proc. – 15th IEEE Int. Conf. Comput. Inf. Technol. CIT 2015, 14th IEEE Int. Conf. Ubiquitous Comput. Commun. IUCC 2015, 13th IEEE Int. Conf. Dependable, Auton. Se. p. 837–42. doi: <https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.124>.
- [23] Zhao Y. Analysing twitter data with text mining and social network analysis. 11th Australas Data Min Conf (AusDM 2013). p. 41–7.
- [24] Sholihah U. The differences between English and Indonesian language. Magistra 2014;47–51.
- [25] Le TA, Moeljadi D, Yasuhide M, Ohkuma T. Sentiment analysis for low resource languages: a study on informal Indonesian tweets. 12th Work. Asian Lang. Resour., p. 123–31.
- [26] Pena Ayala A. Educational data mining applications and trends. London: Springer; 2013.
- [27] Fahrudin T, Buliali JL, Faticah C. Predictive Modeling of the First Year Evaluation Based on Demographics Data: Case Study Students of Telkom University, Indonesia. 2016 Int. Conf. Data Softw. Eng., Bali: IEEE; 2016. doi:10.1109/ICODSE.2016.7936158.
- [28] Maharani K, Adji TB, Setiawan NA, Hidayah I. Comparison Analysis of Data Mining Methodology and Student Performance Improvement Influence Factors in Small Data Set. 2015 Int. Conf. Sci. Inf. Technol. Comp., 2015, p. 169–74.
- [29] Baron J, Lauria EJ. Mining Sakai to Measure Student Performance: Opportunities and Challenges in Academic Analytics. Proc. Enterp. Comput. Community Conf., 2011, p. 1–16.
- [30] Li KF, Rusk D, Song F. Predicting Student Academic Performance. 2013 Seventh Int. Conf. Complex, Intelligent, Softw. Intensive Syst., 2013, p. 27–33. doi:10.1109/CISIS.2013.15.
- [31] Wang H. A comparative study of feature selection techniques for classify student performance. 2015 7th Int. Conf. Inf. Technol. Electr. Eng., p. 425–9. doi: <https://doi.org/10.1109/ICMLA.2010.27>.
- [32] Goga M, Kuyoro S, Goga N. A recommender for improving the student academic performance. Procedia – Soc Behav Sci 2015;180:1481–8. doi: <https://doi.org/10.1016/j.sbspro.2015.02.296>.
- [33] Mohamed A, Husain W, Rashid A. The Third Information Systems International Conference A Review on Predicting Student's Performance using Data Mining Techniques. Procedia - Procedia Comput. Sci., vol. 72, Elsevier Masson SAS; 2015, p. 414–22. doi:10.1016/j.procs.2015.12.157.
- [34] Paradigms SD. Educational data mining: a case study for predicting dropout-prone students. Int J Knowl Eng Soft Data Paradig 2009;1:101–11. doi: <https://doi.org/10.1504/IJKESDP.2009.022718>.
- [35] Bayer J, Bydzovská H, Géryk J, Obsvac T, Popelínský L. Predicting drop-out from social behaviour of students. Proc. 5th Int. Conf. Educ. Data Min. 2012:103–9.
- [36] Romero C, Espejo PG, Zafra A, Romero JR, Ventura S. Web usage mining for predicting final marks of students that use Moodle courses. vol. 21. Wiley; 2013. doi:10.1002/cae.20456
- [37] Gibaja E, Ventura S. Multi-label learning: a review of the state of the art and ongoing research. Wiley Interdiscip Rev Data Min Knowl Discov 2014;4:411–44. doi: <https://doi.org/10.1002/widm.1139>.
- [38] Abbasi A. AI and opinion mining, part 2. IEEE Intell Syst 2010;25:72–9. doi: <https://doi.org/10.1109/MIS.2010.94>.
- [39] Rahmawati D, Khodra ML. Automatic multilabel classification for Indonesian news articles. Adv. Informatics Concepts, Theory Appl., Chonburi: IEEE 2015:1–6. doi: <https://doi.org/10.1109/ICAICTA.2015.7335382>.
- [40] Ajitono T, Widyani Y. Indonesian essay grading module using Natural Language Processing. 2016 Int. Conf. Data Softw. Eng., 2016, p. 1–5. doi:10.1109/ICODSE.2016.7936117.
- [41] Yodihartomo F. Integrated social media knowledge capture model in medical domain of Indonesia. 2017 Int. Conf. Sustain. Inf. Eng. Technol., p. 379–84.
- [42] Amelia W, Maulidevi NU. Dominant emotion recognition in short story using keyword spotting technique and learning-based method. 4th IGNITE Conf. 2016 Int. Conf. Adv. Informatics Concepts, Theory Appl. ICAICTA 2016. doi: <https://doi.org/10.1109/ICAICTA.2016.7803131>.
- [43] Gojali S, Khodra ML. Aspect based sentiment analysis for review rating prediction. 4th IGNITE Conf. 2016 Int. Conf. Adv. Informatics Concepts, Theory Appl. ICAICTA 2016. doi: <https://doi.org/10.1109/ICAICTA.2016.7803110>.
- [44] Liu SM, Chen J. A multi-label classification based approach for sentiment classification. Expert Syst Appl 2015;42:1083–93. doi: <https://doi.org/10.1016/j.eswa.2014.08.036>.
- [45] Enríquez F, Troyano JA, López-solaz T. An approach to the use of word embeddings in an opinion classification task. Expert Syst Appl 2016;66:1–6. doi: <https://doi.org/10.1016/j.eswa.2016.09.005>.
- [46] Vladimir Levenstein. Binary Codes Capable of Correcting Deletions, Insertions, and Reversals. Soviet Physics Doklady; 1966.
- [47] Sun S, Luo C, Chen J. A review of natural language processing techniques for opinion mining systems. Inf Fusion 2017;36:10–25. doi: <https://doi.org/10.1016/j.inffus.2016.10.004>.
- [48] Kanaris I, Kanaris K, Houvardas I, Stamatatos E. Words vs. character n-grams for anti-spam filtering. Int J Artif Intell Tools 2007;16:1–20. doi: <https://doi.org/10.1142/S0218213007003692>.
- [49] Tsoumakas G, Katakis I, Vlahavas I. Mining Multi-label Data. Data Min Knowl Discov Handb 2010:667–85. doi: https://doi.org/10.1007/978-0-387-09823-4_34.