International Journal of Engineering & Technology, 7 (2.27) (2018) 311-318

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Combining rule-based and bag-of-words for phase-level sentiment analysis of blog comments

Frederick F. Patacsil *

Faculty of Information Technology, Pangasinan State University Urdaneta City Campus, India
**Corresponding author E-mail: frederick_patacsil@yahoo.co.uk*

Abstract

Blogs are one of the platforms that express personal opinions, which are intended to create awareness and used as an instrument to establish trust among customers products and services or about a specific topic. A new classifying model was experimented to improve the sentiment classification of the blog comments. This technique combined Rule-Based (RB) and Bag-of-Word (BoW) model to solve major weaknesses of this Bow model in the conduct of Sentiment Analysis (SA) evaluations. The proposed technique was experimented to esti-mate the Philippine Internet customers' satisfaction related to the quality of the services provided by the ISPs in the Philippines. In addition, automatic word seeding, building of sentiment dictionary utilizing an online dictionary, n-gram, tokenization, stemming and other SA tech-niques were applied to extract useful information from the blog comment dataset. The results of the research showed that the configurations involving BoW-RB + SVM + bi-gram + Porter stemmer achieved a high classification accuracy of 88%.
Capturing the contextual meaning of the neighboring words of a given sentimental word provides a significant help to increase the classifi-cation performance of the proposed classifier method.

Keywords: *Blogs; Sentiment Analysis; Machine Learning; Satisfaction; Automatic Polarity Classifier.*

1. Introduction

A blog (a truncation of the expression weblog) is a discussion or informational site published on the World Wide Web (WWW) consisting of discrete entries ("posts") typically displayed in reverse chronological order; in some cases, the most recent post appears first. [1] Beneath the blog posts is an area for customers/readers to make their own comments about the topic of the post. Zhou & Hovy [2] viewed blog comments as "dynamic" involving frequent exchanges between various participants. These blog comments make a blog topic very interesting, interactive, and interrelated among its readers. Further, blog comments are the lifeblood of a blog and the most popular blogs have a very active community of people who voice their opinions on the posts. It is the social aspect that makes blogs a powerful component of the social web.
The Internet customer comments online review are influential factors that may affect the decision and view of other customers about the products and the services of a certain entity. Weblog comments serve as "a simple but effective way for webloggers to interact with their readership," [3] and is considered to be one of the defining sets of weblog characteristics [4]. Therefore, these sentiments are very important sources of information that should be taken into account by companies in improving their services and developing their products. However, existing research largely ignores comments by focusing only to blog posts [3], [5]. This is due to the sheer volume of online comments that makes it difficult for any human to process and extract all meaningful information. As a result, there has been a trend toward the development of systems that can automatically summarize opinions from a set of reviews and eventually display them in an easy-to-process manner. However, automated sentiment analysis of web data can be an Endeavor due to the large volumes of noisy information it contains. Consequently, previous studies have predominantly incorporated manual or semi-automated methods [6]. Manual examination of thousands of information can be an extremely tedious effort when applied across thousands of blog comment postings. With the increasing web usage, the need for automated text classification and analysis techniques has grown in recent years.
The aim of this study is to improve the classification performance of Bag-of-Word (BOW) model to classify the sentiments of blog comments. This study trained and compared the performance of the NB and SVM using Bag-of-Words and Bag-of-Words-Rule-Based (BoW-RB) automatic polarity classified dataset and tested the proposed model using manually labeled dataset. In addition, the study experimented on the different model configurations such as elimination of stopwords, stemming and the use of n-grams.

2. Research methodology

2.1. Information searching and selection

Using an Internet search engine, review blog sites that feature the "services" and "performances" of the Smart, Globe or Sun Internet providers were explored. The results of this search were manually evaluated and filtered such that blog article that contained many comments from their customers and readers were highly considered as part of the dataset. Blog articles that have more than 10 comments/replies were included as part of the dataset. Further, participants in the discussion express both agreement and disagreement with the article and show both positive and negative towards the ISP Internet providers were also considered.

Copyright © 2018 Frederick F. Patacsil. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

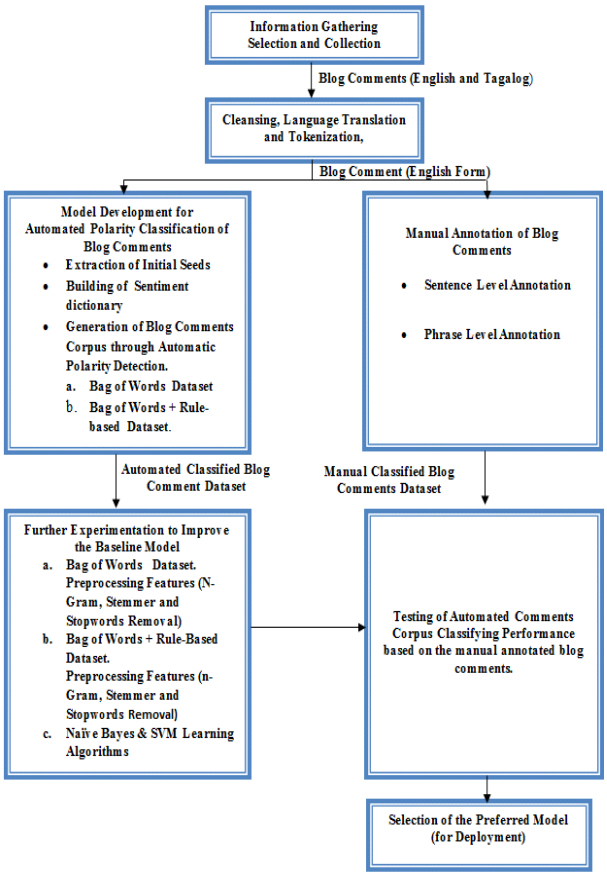


Fig. 1: Methodology Used In Building and Evaluating the Proposed Automatic Sentiment Classifier.

2.1.1. Information searching and selection

In selecting the blog articles to be considered in this study, the Google Search, a search engine, was used to review blog sites that feature the services and performances of the Smart, Globe and/or Sun Internet providers. The formulated search queries that combined the name of the three (3) ISPs with a keyword "vs". The search string was "Globe Tattoo vs. Smart Bro vs. Sun" which inhibited both the name of the ISPs and the word "vs". The results of this search were manually evaluated and filtered such as the blog articles that contained many comments from their customers and readers were highly considered as part of the dataset. Specifically, the following criteria were used in selecting the articles that were included:

- Blog articles must discuss or compare the performance or services of Philippine ISPs (Smart Globe and Sun).
- Blog articles must have at least 10 comments/replies.

The search engine results were checked sequentially following the order of relevance that the engine returned. Each accepted blog article was immediately pre-processed in order to update the total count of the blog comment sentences for the dataset.

A manual selection of blog articles continues until the desired number of blog comments reaches the desired minimum size of 14,000. The dataset included 15 blog articles that contain a total number of 1810 comments of which only 11 blog articles contain comments and the remaining 5 blog sites contain both comments and replies.

2.1.2. Data extraction

Blog comments from the selected blog sites featuring the services of the major Internet providers were extracted using a customized PHP web scraping application, which retrieved the customer comments and other important information and then stored these in a MySQL database automatically. The unnecessary parts such as the title, articles, and other contents were not included in the extraction process.

The URLs address of the blog articles was the input in the PHP application program for blog comments extraction.

2.1.3. Data cleansing

The retrieved blog comments contain several HTML tags, unnecessary characters, non-textual characters, and web codes which were automatically stripped out using a modified program in PHP. Data obtained from blog comments usually contain syntactic features, html code and entities like < > &amp which are embedded in the original comments. Thus, it was necessary to remove those contents from the data because they might affect the result of sentiment classification and were not useful for the machine learning for sentiment analysis. Hence, a PHP application module was designed and developed was used for cleaning blog comments.

2.1.4. Language translation

The collected blog comments were not all written in English. There were Tagalog and English words in some sentences and even sentences that were written in straight Tagalog. In this case, a machine translation using Google Translate was employed. Specifically, a modified application in PHP was used to automatically convert Filipino sentences into their English equivalent using the Google Translate API. However, Filipino words that were not recognized by the machine were manually corrected and converted using online Google translate and the results in English form were copied and pasted manually to the MySQL database.

2.1.5. Tokenization

Splitting up the blog comments to blog comment sentences was the next phase. For English text, white space and punctuation marks are uncomplicated and effective tokenization technique as sentence delimiters [7]. Blog sentence tokenization was useful for obtaining an array of sentences from the given blog comments. The basic and simple sentence delimiter used was (".") plus white space in dividing the sentences of a paragraph. Therefore, to handle cases correctly, the extended splitters used are (".") ("! ") ("?"). The input text was scanned and whenever it came across any of these characters, it would decide whether or not it is the end of the sentence.

2.2. Building of sentiment dictionary

Before carrying out a classification process on the dataset, there was a need to create a word dictionary to be used as a point of reference for sentiment analysis. Several research works utilized automatic seed selection using a set of words as initial seed for the dictionary. Turney's [8] and Rice & Zorn [9], built a word of seeds using two (2) human-selected seed words (the word "Poor", "Bagel" as negative and "Excellent", "Love" as positive). In another study, Zagibalov and Carroll [10] also utilized this approach and started with only a single, human-selected seed ("Good"). Nusko, Tahmasebi and Mogren [11] also used this method and employed more than two words as their initial seed. They claimed that the almost unsupervised system produces a better result.

2.2.1. Extraction of initial Seeds

This research utilized automatic identification of seed of words for the dictionary based on the dataset.

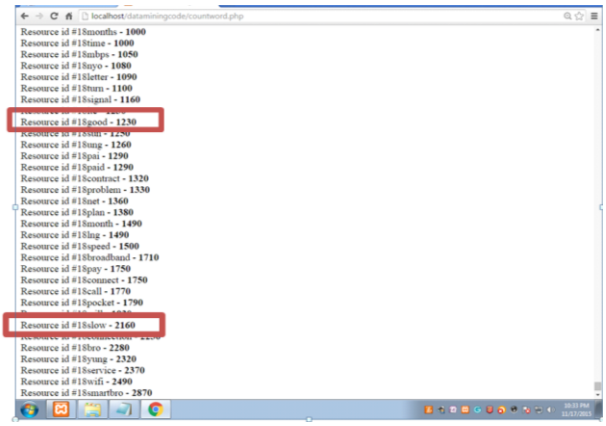
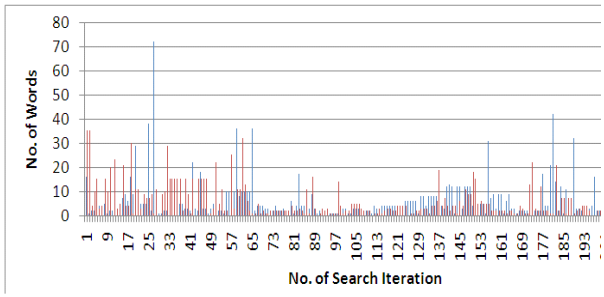


Fig. 2: Word Count Based on the Blog Comment Dataset.

The positive and negative adjectives that had the most number of occurrences as highlighted in Figure 2 were used as the initial seeds. Thus, the initial seed set is composed of the keywords "Good" and "Slow".

2.2.2. Populating sentiment words to the dictionary

Using the keywords "Good" and "Slow" as seeds, an application program was developed to automatically search for and retrieve the synonyms and antonyms of the initial seeds from an online thesaurus dictionary http://thesaurus.altervista.org/thesaurus/v1. After the first iteration, the application retrieved the first synonym word found and repeated the process of searching and retrieving of synonyms and antonyms from the online thesaurus dictionary. The process was repeated until no more new words have been added to the word collection.



'*** No. of words searched (repeated word was automatically disregarded by the application system)
Positive – 930
Negative – 1351

Fig. 3: Result Of Word Seeds Searching Using An Online Thesaurus Dictionary.

2.2.3. Building datasets

The experimentation utilized 14,517 sentences derived from 1,810 blog comments. Such data were used to identify the sentiments of the customers, whether they are satisfied or not with the Internet services provided by Globe, Smart, or Sun.

2.3. Building automated polarity classifier dataset

Two automated polarity classifiers were used to classify the polarity of the blog comment sentences are discussed below.

2.3.1. Data preprocessing

Before the 14,517 sentences were fed to the automated polarity classifier system for polarity identification and labelling, preprocessing of the sentences was required. Removing stop words (common words that have a little value in the process of identifying sentiment analysis e.g. "a", "the", etc.) and stemming (words that carry similar meanings, but in different grammatical forms such as "connect", "connects" and "connected" were combined into one word "connect") were also applied in the sentence preprocessing. In this way, the sentences can show a better representation (with stronger correlations) of these terms, and even the dataset can be reduced for achieving faster processing time.

2.3.2. Polarity identification using bag of words

The bag of words (BoW) was focused completely on the opinionated words in the blog comment sentence. These words have their own polarity value (positive 1, negative -1) in the sentiment dictionary. When these words are found in the sentences, values of the opinionated words are typically all added up and the result is a sentiment estimation of the blog comment sentence.
The identification of sentiment polarity was done using a modified PHP application specifically to estimate the sentiment of the blog comment sentence. The application counts the number of positive and negative words, then computed the total score. If the total score (positive score minus negative score) is greater than zero, it is considered as a positive sentiment and if it is smaller than zero, it is considered as a negative sentiment. In addition, if the total score of negative is equal to the score of positive the polarity is measured as neutral.

- Given a Blog Comment Sentence S
- Identify all sentimental words in S based on the sentiment dictionary
- For each sentiment word in S,
- If word is positive, increment positive_count
- Else if word is negative increment negative_count
- Return sentiment of the sentence based on the final counts
- If both are zero then NO_SENTIMENT
- Else if both are equal (and nonzero) then NEUTRAL
- Else if positive count is higher, then POSITIVE
- Else NEGATIVE

Fig. 4: Pseudo Code Bag of Words Code That Identify Sentiment Polarity of the Sentences.

However, if the presence of negative and positive words in the blog sentence is not found, it will be considered as no-sentiment. The PHP application program automatically identifies the polarity of the blog comment sentences using the bag of words and seed of words (dictionary) approach.

oh hay disgusting Smartbro ie hyper really his face hello thick fix first the service charge you before tardiness so slow connection speeds you service but charge you much complaining why so first concern costumer organize all you charge of whatever
Word of Seeds - slow tardi speed first thick
-1 -1 1 1 -1
3 (negative) – 2 (positive) = Negative (1)
Polarity of the sentence = **Negative**

Fig. 5: Sample Polarity Identification Using Bag of Words.

2.3.3. Polarity identification using a bag of words plus rule-based

The research also utilized another method to estimate the polarity of the blog comment sentences. The model used neighborhood rule and incorporated contextual information of words to enhance the performance of sentiment classifier employing bag-of-words and rule-based through the combination adjectives, adverbs, nouns, verbs, and negative terms rules.
The adverbs, adjectives, nouns and verbs are the most significant set of words that may have some effects on short and long sentences [12], [13], [14], [15], [16], and [17]. The POS Tagger identifies and explains the significance of such words in the sentences. POS tagging was used to extract adverbs, adjectives, verbs, nouns and negation terms in the sentences. Extracted phrases of the sentence which contain the combination of adjectives, adverbs, verbs, and nouns are good indicator sentiments. In addition, retrieved

sentiment words that were included in the seed of words (dictionary) were identified in the sentence and treated with rule-based analysis to ensure that the blog sentence contains only meaningful sentiment words. Rules were applied to the sentiment words and their neighboring words to ensure that the words were all used to express sentiments. Previous approaches in sentiment analysis had leveraged on semantic dependencies between words to predict sentiments [18], [19], and [20]. Wilson, Wiebe, & Hoffmann [21] observed that word patterns along with its modifiers could determine the intensity of phrase sentiments. Furthermore, this method reduces the dimensionality of feature vector and, eventually improves the efficiency of the classification algorithm [11], [12], [14], [22], [23].

2.3.3.1. Automated POS tagging

An automated POS tagging process has been performed in order to tag each term, with respect to its verbal form, to subsequently extract the adjectives, adverbs, nouns and verbs. However, the determination of their polarity depends on the number of the opinionated terms that are in the blog sentence and also found in the dictionary of positive or negative words. For the POS application, some codes here were drawn from the implementation of the Brill tagger by Mark Watson in various languages [24].

At the phrase level, the general order of sentiment computing adjectives, adverbs, verbs, and nouns combinational rules are as follows:

a) Adjective and Adjective Combination

- Unify Adjectives with Adjectives.

/NN /NN /VBZ /JJ /JJ /NN
Sun broadband has poor technical service
Poor (Adjective) - > Sentimental Word
/RB /JJ /JJ /RB /IN/PRP /NN /CC /NN
Sometimes weak signal here in our place but check
Weak (Adjective) - > Sentimental Word

- Adjective and Verb combination
- Unify Adjectives with Verb.

/IN /NN /DT /NN /VBN /DT /VBN /JJ
Within half a year experienced the promised High
/VBN /NN /JJ
Speed connection smart
High (Adjective) -> Sentimental Word

- Adjective and Noun combination
- Unify Adjective with Noun combination

/PRP /VBP /NN /IN /IN /VBN /NNS /RB /CC /NN
You are content with that speed years ahead and sign
/IN /JJ /NNP /VBN
For slow internet speed
Slow (Adjective) -> Sentimental Word

- Adjective and Adverb combination
- Unify Adjectives with Adverb

/IN /DT /VBN /JJ /NN /VBZ /VBN /RB /JJ
Since the speed, wireless broadband has been very slow
/JJR /IN /JJ /NN
Slower than dial connection
Slow (Adjective) -> Sentimental Word

- Adverb and Adverb combination
- Unify Adverbs and Adverbs

/RB /NN /RBR /IN /IN /PRP /RB /RB /RB /RB
Usually fall faster than that she even fairly fast just
/JJ /VBG /CC /NN /BVG
Normal viewing and YouTube streaming
Fast (adverb) -> Sentimental Word
F. Adverb and Noun combination.

- Unify Adverbs with Noun

/JJ /RB /NNS /IN /MD /RB /VBP /NNS /CC
Super-fast downloads like can only have Minutes and
/NN /NN
Day movie
Fast (Adverb) -> Sentimental Word
G. Verb and Noun combination.

- Unify Verb with Noun

/DT /VBN /NN /NN /RB /JJ /IN /DT /NN
The sold PLDT digetel very slow that the digital
/ VBD /RB /JJR /NN /PRP /NN
Slowed even more proof they digetel
slowed (Verb) -> Sentimental Word
H. Negative
It is a feature which presents if the opinionated clue is preceded by a negative tool. Negation is also a crucial component of natural language text, which usually flips the meaning of the sentence. This rule is used to identify whether a word is negated in the sentence.
/NN VBZ /RB /JJ
Globe is not good
Good (verb) -> Sentimental Word Negate by Not
(Positive to negative)
/IN /PRP /JJ /CC /RB /JJ /VBN /NN
Though it's slow but not slow compared sun
Slow (verb) ->Sentimental Word Negate by Not (Negative to Positive)
A set of rules was used to identify the sentence polarity by identifying neighboring relationships between words. This will also ensure that the sentiment words are used to express the sentiment in the sentences, thus, the accuracy of the end resultant polarity value of the sentences is more precise.

Given a Blog Comment Sentence S
Initializes combinational word Patterns
Identify all sentimental words in S based on the sentiment dictionary
For each sentiment word in S,

- Check if sentimental word met the combinational rules If yes Given a Blog Comment Sentence S Identify all sentimental words in S based on the sentiment dictionary For each sentiment word in S,
- If word is positive, increment positive_count
- Else if word is negative increment negative_count

Return sentiment of the sentence based on the final counts

- If both are zero then NO_SENTIMENT
- Else if both are equal (and nonzero) then NEUTRAL

Fig. 6: Pseudocode Automated Sentiment Polarity Identification Using Bag of Words+Rule-Based.

The proposed bag of words+ rule-based (BoW+RB) model integrates contextual polarity by considering the neighborhood relationship of adjectives, adverbs, verbs, and noun words generated using a PHP POS application.

However, with the sun once recv wow hrs free internet connection was slow motion you have limited sites can be browse serious
but/CC with/IN the/DT sun/NN once/RB recv/NN wow/CD hrs/NNS free/JJ internet/NNP connection/NN was/VBD slow/JJ motion/NN you/PRP have/VBP limited/JJ sites/NNS can/MD be/VB browse/VB serious/JJ 0--->slow0--->slow/JJ--->121--->limit
1--->limited/JJ--->16**Polarity is:-2**
Sun broadband has poor technical service

Fig. 7: Sample Polarity Identification Result using Automated Classification Using Bag of Words+Rule-Based.

2.3.4. Building testing dataset

2.3.4.1. Manual polarity identification using sentence level annotations

The same 14,517 sentences used in the automated sentiment classifier were also used as testing dataset for the machine learning classifiers. Four (4) groups with three (3) members were tasked to manually label the polarity of the blog comment sentences. The

formulation of the group was based on the research conducted by Bogart [25]. The results of his study suggest that one rater can rate pretty well, but three can rate better and there is not much gain after increasing to three. The members of the groups were selected using systematic random sampling.

The data were divided equally into four (4) datasets and each group was given an equal number dataset to label manually. Three (3) members of each group independently labelled same set of dataset. In order to check the rater consistency test, the results of the labelling process were tested using Kappa inter-rater reliability. Reliability of dataset is a very important component of the overall confidence in the accuracy of this research study.

Annotators were oriented and trained to apply the criteria for an event that has phenomenon (e.g. "slow internet, bad services" – Negative and "fast internet, good services"– Positive). Since the study's interest was in the explicit sentiment polarity (positive and negative), sentences that have no sentiment and are neutral were bypassed by the annotators and disregarded by the application system. Only sentences identified as positive and negative were considered as part of the testing dataset.

The annotators used a customized PHP application to label the sentiment polarity of the sentences. A screenshot of the application is shown in Figure 8.

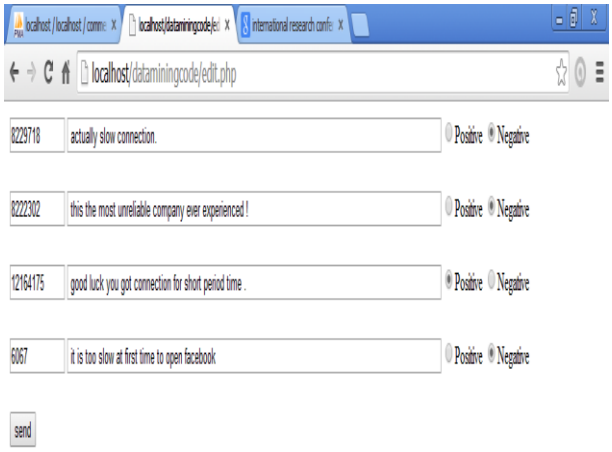


Fig. 8: Sample Screenshot of the Application Used For Manually Labeling the Polarity of the Sentences.

The kappa value was computed on a reduced dataset because only sentences that were rated either positive or negative by the annotator were included in the computation. The results reveal that group 1 has an "almost perfect agreement" with a kappa value of κ = 0.857. Groups 2 and 3 have "substantial agreement" with kappa values of κ=0.76 and κ = 0.73, respectively. κ = 0.586 is the lowest kappa value obtained by Group 4.

The average interrater reliability for the sentence polarity annotators was found to be κ = 0.74. According to the definition of the Fleiss' Kappa statistics, the accuracy of the inter-rater reliability is considered to be "Substantial agreement". By convention, a 0.60 > κ <= 0.75 is considered acceptable inter-rater reliability for applying test [26], [27], [28]. After the Fleiss' Kappa inter-rater reliability was determined, the four annotators on the group agreed on the sentiment polarity of other sentences that were labelled differently to come up with a common polarity in each sentence. Furthermore, only sentences with positive and negative polarity were included as part of the testing dataset.

2.4. Training and testing performance measurement tools

The proposed automated polarity classifier was trained to identify the polarity of unlabeled sentences and to use its own predictions to teach itself to classify unlabeled sentences using a positive and negative word of seeds, bag of words and rule-based. Training experiments were conducted to evaluate and compare the classification performance of bag of words dataset and bag of words+ rule-based dataset utilizing machine learning NB and SVM, N-gram and pre-processing features. Furthermore, the trained models were tested with their classification accuracy using a manually classified polarity dataset. The classification experimentations were carried out using 10-fold cross validation using Rapid Miner 5.3. In terms of comparing their performance, the confusion matrix that contains the precision, recall, accuracy and F-measure measurement was employed.

Table 1: Confusion Matrix for Two-Class Classifier

Predicted Class	True Negative	True Positive	Class Precision	F-Measure
Negative	*tn*	*fp*	Negative	F-Measure Negative
Positive	*fn*	*tp*	Positive	F-Measure Positive
Class Recall	Negative	Positive		

The accuracy (AC) is the proportion of the total number of predictions that were correct. It is determined using the (1):

$$AC = \frac{tn + tp}{tn + fp + fn + tp}$$

The recall (in the case of positive cases) is the proportion of positive cases that were correctly identified, as calculated using the (2):

$$Recall = \frac{tp}{tp + fp}$$

The precision (in the case of positive cases) is defined as the proportion of negative cases that were classified correctly, as calculated using the (4):

$$Precision = \frac{tp}{tp + fn}$$

The F-Measure considers both precision and recall providing a single measurement for a system avoiding having two independent measures. It is computed using the (5):

$$FM = 2\frac{(precision * recall)}{(precision + recall)}$$

The F-measure and accuracy were used as the performance indicators for evaluating the models since the models were trained and tested using unbalanced dataset. Furthermore, both of these classifying indicators have been widely used in previous researches [29], [6].

3. Results and discussions

Table 2 reveals that the dataset classified by the automated classifiers performs very well. However, BoW+R+B model outperformed BoW model in all measured areas. The results indicate that the relationship of words in the sentence provides a significant impact on the classification performance of the automated classifier. In addition, incorporating N-gram, Stemmer and machine learning also add more significant effects on the performance of the proposed automatic classifiers. In the case of N-gram features, bi-gram obtained the highest classification performance when compared to uni - gram and tri-gram in some measured areas. Further, incorporating stemming method increases the performance of bi-gram in all measured areas.

Table 2: Classification Results of the Proposed Automatic Polarity Classifiers Applying Different Learning Machine, N-gram and Pre-processing Features

N-Gram + Preprocessing Features	BoW														BoW+R-B													
	NB							SVM							NB							SVM						
	Recall		Precision		F-measure		Accuracy	Recall		Precision		F-measure		Accuracy	Recall		Precision		F-measure		Accuracy	Recall		Precision		F-measure		Accuracy
	Neg.	Pos.	Neg.	Pos.	Neg.	Pos.		Neg.	Pos.	Neg.	Pos.	Neg.	Pos.		Neg.	Pos.	Neg.	Pos.	Neg.	Pos.		Neg.	Pos.	Neg.	Pos.	Neg.	Pos.	
U-G	68%	77%	88%	49%	77%	60%	**70%**	87%	67%	87%	67%	87%	67%	**81%**	66%	90%	94%	52%	78%	66%	**73%**	85%	85%	93%	69%	89%	76%	**85%**
B-G	80%	72%	88%	59%	84%	65%	**78%**	92%	55%	84%	74%	88%	63%	**82%**	77%	85%	93%	60%	84%	71%	**80%**	89%	84%	93%	75%	91%	79%	**87%**
T-G	80%	73%	88%	60%	84%	66%	**78%**	94%	44%	81%	73%	87%	55%	**79%**	78%	86%	93%	61%	85%	71%	**80%**	90%	79%	91%	76%	91%	77%	**87%**
U-G+St	67%	78%	88%	49%	76%	60%	**70%**	86%	65%	86%	65%	86%	65%	**80%**	65%	89%	93%	50%	77%	64%	**72%**	84%	82%	92%	68%	88%	74%	**84%**
B-G+St	75%	73%	87%	54%	81%	62%	**75%**	92%	53%	83%	74%	87%	62%	**81%**	75%	88%	94%	58%	83%	70%	**78%**	88%	85%	93%	73%	90%	78%	**87%**
T-G+St	76%	73%	88%	55%	82%	63%	**75%**	94%	43%	80%	73%	86%	54%	**79%**	75%	88%	94%	58%	83%	70%	**78%**	88%	82%	92%	74%	90%	78%	**87%**
U-G+Str	66%	78%	88%	48%	76%	59%	**69%**	87%	67%	87%	68%	87%	67%	**81%**	62%	91%	94%	49%	75%	64%	**70%**	83%	85%	93%	67%	88%	75%	**84%**
B-G+Str	81%	72%	88%	60%	84%	66%	**79%**	93%	57%	84%	76%	88%	65%	**82%**	77%	87%	94%	61%	85%	71%	**80%**	89%	86%	94%	76%	91%	81%	**88%**
T-G+Str	81%	73%	88%	61%	84%	66%	**79%**	94%	45%	81%	74%	87%	56%	**80%**	78%	87%	94%	61%	85%	72%	**81%**	90%	82%	93%	77%	91%	79%	**88%**
U-G+St+Str	64%	79%	89%	47%	74%	59%	**68%**	86%	65%	86%	65%	86%	65%	**80%**	60%	90%	94%	47%	73%	62%	**69%**	83%	81%	92%	65%	87%	72%	**82%**
B-G+St+Str	81%	73%	88%	61%	84%	66%	**79%**	93%	55%	84%	75%	88%	64%	**82%**	75%	88%	94%	58%	83%	70%	**79%**	87%	85%	93%	72%	90%	78%	**86%**
T-G+St+Str	81%	73%	88%	61%	85%	66%	**79%**	94%	45%	81%	74%	87%	56%	**80%**	75%	88%	94%	59%	84%	71%	**79%**	88%	83%	93%	73%	90%	77%	**86%**

Legend:

U-G	Uni-gram		U-G+Str	Uni-gram+Stemmer	
B-G	Bi-Gram		B-G+Str	Bi-Gram+Stemmer	
T-G	Tri-Gram		T-G+Str	Tri-Gram+Stemmer	
U-G+St	Uni-gram + Stopword		U-G+St+Str	Uni-gram + Stopword+Stemmer	
B-G+St	Bi-Gram+ Stopword		B-G+St+Str	Bi-Gram+ Stopword+Stemmer	
T-G+St	Tri-Gram+ Stopword		T-G+St+Str	Tri-Gram+ Stopword+Stemmer	

However, incorporating stopwords removal and applying both stopwords and stemmer in the automated classifiers reduces the classifying performance. In the case of learning machines, SVM outclassed NB in all measured areas as revealed in the table. The best result was attained by BoW+R+B utilizing bi-gram incorporating stemmer applying SVM learning machine. On the other hand, NBs uni-gram under BoW was seen to have a negative or zero increase in classification accuracy performance both on F-measure and accuracy.

The study also reveals that BoW+RB using SVM bi-gram and tri-gram performed significantly better than its counterpart in the bag of words, as measured by classification accuracy performance. This suggests that the combination of BoW+RB and SVM bi-gram and tri-gram showed to be effective in classifying sentiments of blog comment sentences. Capturing the contextual meaning of the neighbouring words of a given sentimental word provides a significant help to increase the classification performance of the proposed classifier method. Thus, applying bi-gram and tri-gram to classify the sentiment of blog comment sentences, indeed proved to be a better feature as compared from uni-gram during the experimentation utilizing SVM learning machine.

3.1. Influence of N-gram, pre-processing features, and machine learning on the performance of the proposed classifiers

In the case of the proposed models, BoW achieved a minimal increase in F-measure and accuracy performance. One possible reason is that this model (BoW) cannot capture the effect of negation and the contextual meaning of the neighboring words in the blog sentences. Integrating the effect of negation to the BoW features plus incorporating context of adjectives, adverbs, noun, and verbs and employing neighborhood rule-based on the bag-of-words model (BoW+R+B) proved to be highly efficient.

3.2. Factors contributory to misclassifications of blog sentences

The results of the experimentation show that the proposed sentiment models did not accurately classify all the blog comment sentences. The misclassifications from negative into positive or vice versa were manually investigated and analyzed. The most apparent factors or causes for not correctly classifying the blog comment sentences were noted as follows:

1) Nonexistence of opinionated words in the dictionary. This is caused by the low frequencies of occurrence in the dictionary of sentimental words. Opinionated words are the only clues to determine the sentiment of the blog comment sentences.

E.g.
How beautiful the service provided you feel sure many customers find these behaviors! Beautiful is not in the seed of words. Pocketwifi use pretty much all the time Pretty is not in the seed of words.

2) The use of jargon and other non-dictionary words are common in blog comments. In the blog comments domain "informal words", jargon all of its own, pointed re-spelling, domain specific slang, colloquial language, and Short Messages (SMS) like those of sentences used by customers to express their sentiments. However, the proposed methods cannot establish the sentiment of sentence because these words are not part of the dictionary.

E.g.
Digital digit hell super tardiness Thickness face raises the fee from the service charge that they do not satisfy customers having
Therefore, do you feel Digi hell?

3) Negation and Sarcasm effects. Negation integration is helpful, but the proposed classifiers was not yet perfect to handle negative cases.

E.g.
Annotated as negative in the test dataset, however classified as positive during testing.
Now that does not have fast reliable internet speed
They cannot provide good services on the customers.
E.g. Annotated as positive in the test dataset, however, classified as negative during testing. Few years ago, the service was fine but now too slow. hay fast globe but the problem is that problem delayed the bill previously disconnected without he bill. However, it is slow but not slow compared sun and smart.
The word "but" and similar words here are arguably "polarity reversers" that the proposed system cannot handle efficiently.

4) Wrong grammar and spelling. Blog comment sentences, which were written by the customers, were not always grammatically correct. Misspelled words were also common. Furthermore, there were errors in the translation of sentences due to some Filipino words that have corresponding terms in English that may cause changes the polarity of the sentence limiting the efficiency of the proposed methods used.

In addition, the proposed methods were dependent in the sentiment dictionary. Misspelled opinionated words were not similar to the opinionated words stored in the dictionary, thus, they were not included in the establishment of the sentiment of the sentence.

5) Ambiguity in synonyms and antonyms of words in dictionary. Some word seeds carry polarity whose meaning cannot be resolved definitely.

E.g. "they just great service billing but they also futile" great was considered as sentiment words and counted as positive polarity.
Go to nearest smart wireless center say the best complaint. Best positive polarity but here it was used to support a negative word.
In the above examples, single words can have different meanings depending on how they were used in the sentence or in the context. One major problem that was observed during experimentation was the coverage of the sentiment dictionary. The detection and classification of blog comment sentences relied on the sentiment dictionary, hence, the minimal coverage of the sentiment words in the dictionary where the bottleneck of the proposed classifiers. Thus, building more comprehensive lists of sentiment words could be considered a possible solution to further enhance the performance.

4. Conclusions

Based on the foregoing findings, the following conclusions were drawn:

1) Two automated methods of classifying the polarity of blog comment sentences were utilized. The first method was polarity counting using sentimental dictionary. The second method is polarity identification using Part-Of-Speech (POS) tagging process and neighboring rule of adjectives, adverbs, noun, verb, and negative words.
2) Support Vector Modelling machine learning and bag of words +rule-based show a significant influence on the accuracy classification performance of the proposed automatic sentiment classifiers.
3) Bi-gram and tri-gram show a significant influence on the accuracy classification performance of the proposed automatic sentiment classifiers.
4) Stemming process and stop words removal shows no significant influence on the accuracy classification performance of the proposed automatic sentiment classifiers

5. Recommendations

The obtained experimental results are promising, and researchers are encouraged to continue to improve the model. The following are recommended:

1) The use of another machine translating tool;
2) The use of auto-spelling and auto-grammar checker;
3) Use of lexical resource for opinion mining that will assign to each of the three (3) sentiment scores: positivity, negativity, objectivity; and
4) Building more comprehensive lists of stop words could be considered to further enhance the performance of the proposed model, and adding more words in our polarity dictionary is highly recommended.

References

[1] Blood, R., 2000. Weblogs: A history and perspective. *Rebecca's pocket*, *7*(9), p.2000.

[2] Zhou, L. and Hovy, E.H., 2006, March. On the Summarization of Dynamically Introduced Information: Online Discussions and Blogs. In *AAAI Spring symposium: Computational approaches to analyzing weblogs* (p. 237).

[3] Marlow, C., 2004, May. Audience, structure and authority in the weblog community. In *International Communication Association Conference* (Vol. 27).

[4] Winer, D., 2003. What makes a weblog a weblog? *Weblogs at Harvard Law*, *23*, p.2003.

[5] Hu, M., Sun, A. and Lim, E.P., 2007, November. Comments-oriented blog summarization by sentence extraction. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management* (pp. 901-904). ACM. https://doi.org/10.1145/1321440.1321571.

[6] Zheng, R., Li, J., Chen, H. and Huang, Z., 2006. A framework for authorship identification of online messages: Writing-style features and classification techniques. *Journal of the Association for Information Science and Technology*, *57*(3), pp.378-393. HTTPS://DOI.ORG/10.1002/ASI.20316.

[7] Joshi, N.S. and Itkat, S.A., 2014. A survey on feature level sentiment analysis. *International Journal of Computer Science and Information Technologies*, *5*(4), pp.5422-5425.

[8] Turnkey, P.D., 2002, July. Thumbs up or thumbs down semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th annual meeting on association for computational linguistics* (pp. 417-424). Association for Computational Linguistics.

[9] Rice, D.R. and Zorn, C., 2013. Corpus-based dictionaries for sentiment analysis of specialized vocabularies. *Proceedings of NDATAD*, pp.98-115.

[10] Zagibalov, T. and Carroll, J., 2008, August. Automatic seed word selection for unsupervised sentiment classification of Chinese text. In *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1* (pp. 1073-1080). Association for Computational Linguistics. https://doi.org/10.3115/1599081.1599216.

[11] Nusko, B., Tahmasebi, N. and Mogren, O., 2016, July. Building a sentiment lexicon for swedish. In *Digital Humanities 2016. From Digitization to Knowledge 2016: Resources and Methods for Se-*

mantic Processing of Digital Works/Texts, Proceedings of the Workshop, July 11, 2016, Krakow, Poland(No. 126, pp. 32-37). Linköping University Electronic Press.

[12] Gamallo, P. and Garcia, M., 2014. Citius: A naive-bayes strategy for sentiment analysis on english tweets. In *Proceedings of the eighth international Workshop on Semantic Evaluation (SemEval 2014)* (pp. 171-175). https://doi.org/10.3115/v1/S14-2026.

[13] Kasthuriarachchy, B.H., De Zoysa, K. and Premaratne, H.L., 2014, December. Enhanced bag-of-words model for phrase-level sentiment analysis. In *Advances in ICT for Emerging Regions (ICTer), 2014 International Conference on* (pp. 210-214). IEEE. https://doi.org/10.1109/ICTER.2014.7083903.

[14] Benamara, F., Cesarano, C., Picariello, A., Recupero, D.R. and Subrahmanian, V.S., 2007, March. Sentiment analysis: Adjectives and adverbs are better than adjectives alone. In *ICWSM*.

[15] Patel, N.D. and Chand, C., 2014. Selecting Best Features Using Combined Approach in POS Tagging for Sentiment Analysis. *Vol. 3 Issue. [3], pg. 425-430 2014, 23*.

[16] Subrahmanian, V.S. and Reforgiato, D., 2008. AVA: Adjective-verb-adverb combinations for sentiment analysis. *IEEE Intelligent Systems*, *23*(4), pp.43-50. https://doi.org/10.1109/MIS.2008.57.

[17] An, N.T.T. and Hagiwara, M., 2014, June. Adjective-based estimation of short sentence's impression. In *KEER2014. Proceedings of the 5th Kanesi Engineering and Emotion Research; International Conference; Linköping; Sweden; June 11-13* (No. 100, pp. 1219-1234). Linköping University Electronic Press.

[18] Yu, H. and Hatzivassiloglou, V., 2003, July. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the 2003 conference on Empirical methods in natural language processing* (pp. 129-136). Association for Computational Linguistics. https://doi.org/10.3115/1119355.1119372.

[19] Roebuck, K., 2011. *Encryption: High-impact Strategies-What You Need to Know Definitions, Adoptions, Impact, Benefits, Maturity, Vendors*. Tebbo.

[20] Hatzivassiloglou, V. and McKeown, K.R., 1997, July. Predicting the semantic orientation of adjectives. In *Proceedings of the 35th annual meeting of the association for computational linguistics and eighth conference of the european chapter of the association for computational linguistics* (pp. 174-181). Association for Computational Linguistics.

[21] Wilson, T., Wiebe, J. and Hoffmann, P., 2009. Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. *Computational linguistics*, *35*(3), pp.399-433. https://doi.org/10.1162/coli.08-012-R1-06-90.

[22] Romanyshyn, M., 2013. Rule-based sentiment analysis of ukrainian reviews. *International Journal of Artificial Intelligence & Applications*, *4*(4), p.103. https://doi.org/10.5121/ijaia.2013.4410.

[23] Asghar, M.Z., Khan, A., Ahmad, S., Qasim, M. and Khan, I.A., 2017. Lexicon-enhanced sentiment analysis framework using rule-based classification scheme. *PloS one*, *12*(2), p.e0171649. https://doi.org/10.1371/journal.pone.0171649.

[24] Barber, I., PHPir, http://phpir.com/part-of-speech-tagging/. 2010.

[25] Bogartz, R.S., Interrater agreement and combining ratings. *University of Massachusetts, Amherst. http://www-unix. oit. umass. edu/~ bogartz*.

[26] Barbosa, L. and Feng, J., 2010, August. Robust sentiment detection on twitter from biased and noisy data. In *Proceedings of the 23rd international conference on computational linguistics: posters* (pp. 36-44). Association for Computational Linguistics.

[27] Klavans, J. and Kan, M.Y., 1998, August. Role of verbs in document analysis. In *Proceedings of the 17th international conference on Computational linguistics-Volume 1* (pp. 680-686). Association for Computational Linguistics.

[28] Niu, Z., Yin, Z. and Kong, X., 2012, August. Sentiment classification for microblog by machine learning. In *Computational and Information Sciences (ICCIS), 2012 Fourth International Conference on* (pp. 286-289). Ieee. https://doi.org/10.1109/ICCIS.2012.276.

[29] Wei, Z., Miao, D., Chauchat, J.H. and Zhong, C., 2008, May. Feature selection on Chinese text classification using character n-grams. In *International Conference on Rough Sets and Knowledge Technology* (pp. 500-507). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-79721-0_68.