Lexicon-Based and Immune System Based Learning Methods in Twitter Sentiment Analysis

Hamidah Jantan, Fatimatul Zahrah Drahman, Nazirah Alhadi, and Fatimah Mamat

Universiti Teknologi MARA (UiTM) Terengganu, Malaysia,
hamidahjtn@tganu.uitm.edu.my, {fatimahtuldh,nazirah,fatimahm}@yahoo.com

ABSTRACT

Nowadays, there are increasingly numbers of studies on seeking ways to mine Twitter for sentiment analysis. Machine learning approach such as immune system based learning methods is an alternative way for sentiment classification. This method is centered on prominent immunological theory as computation mechanisms that emulate processes in biological immune system in achieving higher probability for pattern recognition. The aim of this article attempts to study the potential of this method in text classification for sentiment analysis. This study consists of three phases; data preparation; classification model development using three selected Immune System based algorithms i.e. Negative Selection algorithm (NSA), Clonal Selection algorithm (CSA) and Immune Network algorithm (INA); and model analysis. As a result, NSA algorithm proposed slightly high accuracy in experimental phase and that would be considered as the potential classifiers for Twitter sentiment analysis. In future work, the accuracy of proposed model can be strengthened by comparative study with other heuristic based searching algorithms such as genetic algorithm, ant colony optimization, swam algorithms and etc.

Keywords: Immune system, lexicon, sentiment classification, twitter messages.

I INTRODUCTION

Micro blogging platforms such as Twitter used by many companies and media organizations to know what people think and feel about their products and services. Given the character limitation on tweets, it deals with explicit messages that contain valuable sentiments for future decision making. Sentiment analysis in natural language processing (NLP) field is ranging from document, sentence, and aspect level analysis to learning the polarity of words and phrases in text. Twitter as social media deal with short message that has limited number of characters but contains valuable sentiment for future decision making in many area. Twitter sentiment analysis commonly used in product review, stock market analysis, news articles, political debates and etc. (Alexander & Paroubek, 2010; Balahur & Steinberger, 2009; Medhat et al., 2014; Tumasjan et al., 2010; Xiaodong Li et al., 2014). In sentiment analysis, there are four main tasks i.e. sentiment identification, feature selection, classification and polarity. Sentiment classification will determine the end result of sentiment in text analysis. As an example, sentiment classification in Twitter messages focus on sentence-level analysis in order to identify the view point(s) of underlying sentiment in text such as positive, negative and neutral (Pang & Lee, 2004). In sentiment classification, there are three major classification techniques can use for sentiment analysis i.e. machine learning that focus on linguistic features, lexicon-based relies on sentiment lexicon and hybrid approaches (Choi et al., 2009; Medhat et al., 2014; Pang & Lee, 2004). Machine learning approach has been widely used for sentiment analysis in order to reduce the high-dimensional feature space feature selection in traditional approach by eliminating the noisy and irrelevant features; and their ability towards global optimization in classifier construction (Agarwal & Mittal, 2015; Medhat et al., 2014).

Immune systems based learning method is known as soft computing paradigm used in machine learning. This approach inspired by the human immune system that elicits theories which can act as an inspiration for computer-based solutions as alternative approach to solve computational problems (Azadeh et al., 2014; Hunt & Cooke, 1996; Sarafijanovic & Boudec, 2005). Artificial Immune Systems (AIS) has led to the development of numerous classification models in various areas such as pattern recognition, fault detection, computer security, data mining, engineering and computer applications (Venkatesan et al., 2013). Due to the uniqueness of self-recognition ability stimulated by the biological immune system, this technique is widely used in pattern recognition by classifying self or non-self as detectors in classification task. In this study, three selected AIS based algorithms i.e. Negative Selection algorithm (NSA), Clonal Selection algorithm (CSA) and Immune Network algorithm (INA) are used as recognition mechanism by classifying sentiment in text analysis. The rest of this paper is organized as follows: the second section discusses the related work in sentiment mining and AIS algorithms and application, the third section describes the experiment setup conducted in this study and, the fourth section discusses the results and discussions. Finally, the paper ends with the fifth section where the concluding remarks and future research directions are identified.

II RELATED WORK

A. Sentiment Analysis

Nowadays, sentiment analysis is among the key emergent technologies to navigate the huge amount of online content regarding on people opinion in product review, political view, stock market analysis, news article and etc. (Alexander & Paroubek, 2010; Kucuktunc et al., 2012; Paltoglou & Thelwall, 2012; C. Tan et al., 2011). The ability to extract opinion or sentiment from online sources can provide valuable information about people's views on various topics that are beneficial for future planning. Sentiment analysis identified the sentiment expressed in a text then analyses it to find opinions, identify the sentiments and to classify their polarity as shown in Figure 13. (Choi et al., 2009; Medhat et al., 2014; Pang & Lee, 2004). Sentiment classification is a central task that has three main classification levels i.e. document-level aims to classify an opinion document as expressing a positive or negative opinion; sentence-level focus to classify sentiment in each sentence whether subjective or objective then determine whether it is positive or negative; and aspect-level aims to classify the sentiment with respect to the specific aspects of entities.

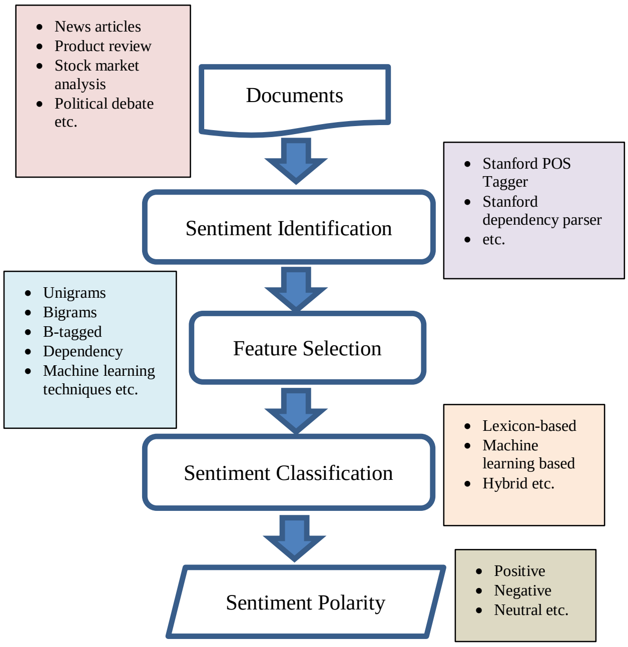


Figure 13. Fundamental of Sentiment Analysis

Sentiment classification techniques categorized into three approaches i.e. lexicon-based; machine learning; hybrid approaches (Medhat et al., 2014). Lexicon-based approach relies on sentiment lexicon which is a collection of precompiled sentiment terms that known as manual-based, dictionary-based and corpus-based collections. This approach uses statistical or semantic methods to find sentiment polarity in text analysis. There and many applications that use this approach such as Arabic Twitter corpus for subjectivity (Refaee & Rieser, 2014); Indian general election using context rules (Singhal et al., 2015); news impact on stock market (Xiaodong Li et al., 2014) and many others.

Machine learning approach uses linguistic features in supervised learning and unsupervised learning. Supervised learning deals with a large labeled training documents and unsupervised learning for unlabeled training documents for classification task. The common techniques used in supervised learning are decision tree, linear-based (Support Vector machine and Neural Network), rule-based and probabilistic classifier (Naïve Bayes, Bayesian Network and Maximum entropy). Support vector machine (SVM) is proven as high performance and robust methods as classifier in previous works (Agarwal & Mittal, 2015; Joachims, 2005; Mudinas et al., 2012; S. Tan & Zhang, 2008; Zhang et al., 2011). Besides that, Bayesian naïve methods are known as the most efficient and effective inductive learning algorithms for machine learning (Agarwal & Mittal, 2015; Dinu & Iuga, 2012; McKaughan et al., 2011). Recently, bio-inspired algorithms for classification attract attention in this area due to the ability in propose high accuracy in classification (Akhmedova et al., 2014; Puteh et al., 2013; Samsudin et al., 2013).

Hybrid approach combines both approaches where sentiment lexicons playing a key role in sentiment analysis. As example, this approach used in combining lexicon-based and learning-based methods for sentiment analysis (Akhmedova et al., 2014; Khan et al., 2015; Mudinas et al., 2012; Zhang et al., 2011). In this approach, lexicon-based approach focus on sentiment feature selection and machine learning- based approach focus on sentiment classification. Besides that, other techniques from those categories also been used in sentiment analysis such as Formal Concept Analysis (FCA) is a mathematical approach for structuring, analyzing and visualizing data, Fuzzy Formal Concept Analysis (FFCA) to conceptualize documents into a more abstract form of concepts and etc.(Medhat et al., 2014)

Twitter Sentiment Analysis

Nowadays, in the emerging social media, more and more people are expressing their likes and dislikes sentiments towards different subjects about current affairs on blogs, micro-blogs and social networking sites such as Twitter, Facebook and etc. Analyzing these expressions of short colloquial text can yield vast information about the behavior of the people that can be helpful in many. There are many studies on Twitter sentiment analysis to handle these issues especially in product review, stock market analysis, political debates and many others (Alexander & Paroubek, 2010; Balahur & Steinberger, 2009; Medhat et al., 2014; Tumasjan et al., 2010; Xiaodong Li et al., 2014). Their focus is to determine the polarity of the Twitter sentiment analysis using either lexicon-based, machine-based, hybrid or other method in sentiment classification (Kontopoulos et al., 2013; Kouloumpis et al., 2011; Singhal et al., 2015).

B. Immune Based Learing System

Bio-inspired algorithms stimulate human immune system known as Artificial Immune System (AIS). This technique inspired by the immunology immune function, principles and models to solve complex problem and implements a learning technique for natural defense mechanism that learns about foreign substances (Dasgupta, 2006; Hunt & Cooke, 1996). There are three common algorithms in AIS i.e. Negative selection algorithm (NSA), clonal selection algorithm (CSA) and artificial immune network (AiNet) algorithm.

Clonal Selection Algorithm

The clonal selection algorithm is a class of algorithms inspired by the clonal selection theory of acquired immunity that describe how B and T lymphocytes improve the reaction of antigen which is called as affinity maturation. The clonal selection theory in an immune system is used to clarify the basic reaction of the adaptive immune system to an antigenic stimulus. The theory is based on the idea that only cells capable of recognizing an antigen will proliferate (Berna & Sadan, 2011). Its main idea is the antigen can selectively react with the antibodies, which are native production and spread on the cell surface in the form of peptides. When an antigen is discovered, those antibodies that best recognize an antigen will proliferate by cloning.

Clonal selection operation has the ability to combine the global search with local search. The global optimum can be easily obtained through a series of operations including clone, mutation, and selection (Liu et al., 2009). Figure 2 shows the process flow for Clonal selection algorithm (CSA) in classification process to detect antigen(Castro & Timmis, 2003). Where P is a candidate solution, M is a memory sel, P_r is the remaining of population that indicates P= P_r+M, P_n is best individual of population, C is temporary population of clones, C* is maturated antibody population and N_d is replace d antibodies by novel ones (diversity introduction), the lower affinity cells have higher probabilities of being replaced.

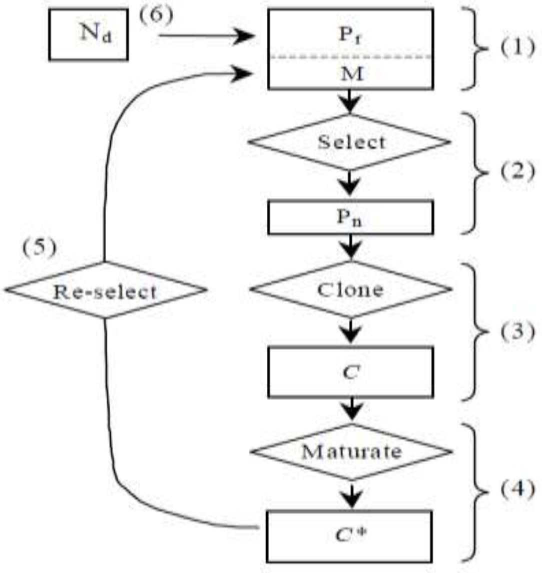


Figure 14. Clonal Selection Algorithm (CSA)

Negative selection algorithm

The Negative Selection Algorithm (NSA) is an immunology-inspired algorithm for anomaly detection application. This algorithm has been implemented with different pattern representations and various matching rules and successfully applied to a broad range of problems (Hou & Dozier, 2006).

NSA is an approach to anomaly detection using negative detectors, was originally proposed by Forest, which models the clonal deletion process in the natural immune system to prevent autoimmunity (Hou & Dozier, 2006). The NSA generates random detectors and removes the ones that detect self-patterns, which results in a collection of detectors that potentially detect non self-patterns Classification is performed by generating detectors that match none of the negative examples, and these detectors are then matched against the elements to be classified and a large number of detectors may be required for acceptable sensitivity, or finding detectors that match none of the negative examples may be difficult (Li et al., 2010). Figure 3 shows the process flow of NSA for generation and detection in classification process

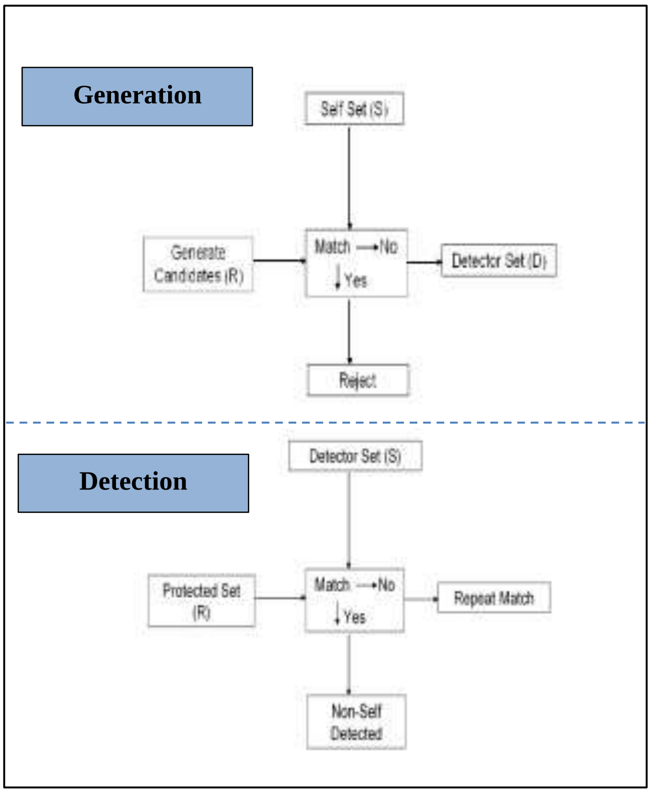


Figure 3. Fundamental of Negative Selection Algorithm

Artificial Immune Network

Artificial Immune Network (AiNet) is a bio-inspired computational model that uses ideas and concepts from the immune network theory, mainly the interactions among B cells (stimulation and suppression), and the cloning and mutation process (Fran et al., 2005; Galeano et al., 2005). The immune network theory was suggested by Jerne, as a way to describe the memory and learning capabilities exhibited by the immune system. AiNet algorithm focus on the network graph structures involved where antibody producing cells represent the nodes and the training algorithm involves growing or pruning edges between the nodes based on affinity as shown in Figure 4.

AiNet algorithm has been used to solve the computer security against computer viruses which disrupt the normal usage of the network. (Chandrasekaran & Murugappan, 2008). Besides that, another research using AiNet is used for mobile ad hoc networks to detect the node misbehavior in mobile ad hoc network using DSR (Sarafijanovic & Boudec, 2005). AiNet used for multimodal function optimization on dynamic environment to deal with the time-varying fitness function with the challenging benchmark problems in static multimodal optimization are considered to validate the new proposal (Fran et al., 2005).

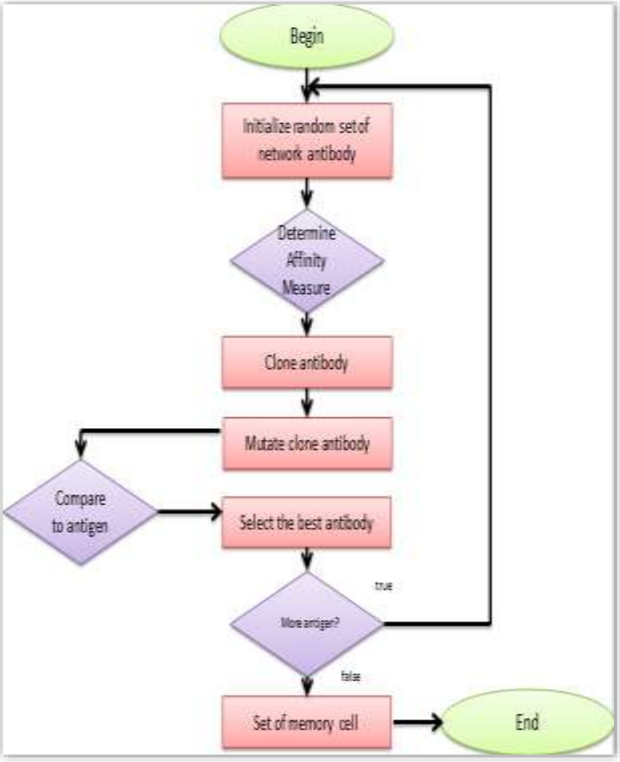


Figure 4. Artificial Immune Network Algorithm (AiNet))

III RESEARCH METHOD

Sentiment analysis in this study has four main phases using lexicon-based and immune system based approaches. Figure 5 show the process flow involved in this study.

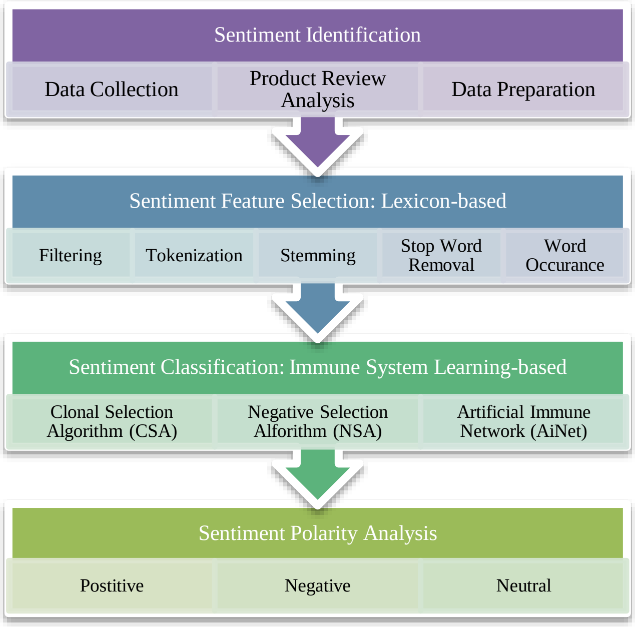
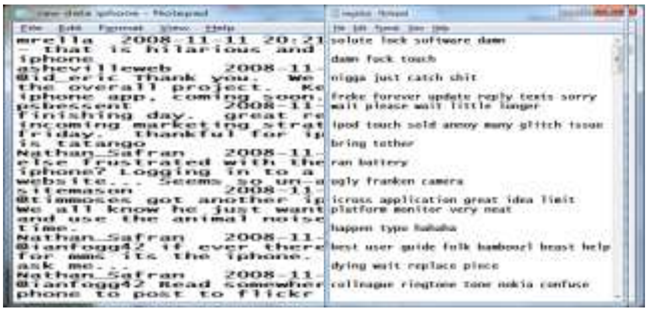

Sentiment Identification
Data Collection
Product Review Analysis
Data Preparation
Sentiment Feature Selection: Lexicon-based
Filtering
Tokenization
Stemming
Stop Word Removal
Word Occurance
Sentiment Classification: Immune System Learning-based
Clonal Selection Algorithm (CSA)
Negative Selection Alforithm (NSA)
Artificial Immune Network (AiNet)
Sentiment Polarity Analysis
Postitive
Negative
Neutral

Figure 5. Experiment Phase on Twitter Sentiment Analysis

The first phase is sentiment identification contains data collection, product review analysis and data preparation activities. Twitter dataset on hand phone product review that contain 600 Twitter messages are used for the case study. The second phase is sentiment feature selection using lexicon-based approach that contains filtering, tokenization, stemming, stop word removal and word occurrence analysis to prepare a group of word (Bag of Words (BOWs). Immune system based learning approach is used as sentiment classification method in third phase. Three common algorithms in AIS are used in this process i.e. Clonal Selection Algorithm (CSA), Negative Selection Algorithm (NSA) and Artificial Immune Network (AiNet). Dataset is divided into 480 as training dataset and 120 as testing dataset in learning process. Sentiment polarity analysis in the fourth phase is determined by the accuracy of correctly classified in testing data via proposed classification model for each algorithm.

IV RESULT ANALYSIS

In this study, the lexicon-based approach is used in sentiment feature selection process in order to prepare the collection of word (BOW) for the dataset. **Error! Reference source not found.** shows the sample of clean data by filtering, stemming and stop word removal processes in this phase. The next step is to prepare the antibody by identify the sentiment (positive, negative or neutral). The clean data than categorized according to sentiment identified from selected dictionary. As an example, in NSA algorithm training phase, a set of detectors will be generated as non-self.



```
mrella     2008-11-11 20:21
- that is hilarious and
iphone
ashevilleweb     2008-11-
@id_eric Thank you.  We
the overall project.  Ke
iphone app. coming soon.
pcbessent        2008-11-
finishing day.  great re
incoming marketing strat
friday.  thankful for ip
is tatango
Nathan_Safran    2008-11-
else frustrated with the
iphone? Logging in to a
website...  Seems so un-a
sitemason        2008-11-
@timmoses got another ip
we all know he just want
and use the animal noise
time.
Nathan_Safran    2008-11-
@ianfogg42 if ever there
for mms its the iphone.
ask me...
Nathan_Safran    2008-11-
@ianfogg42 Read somewher
phone to post to flickr
```

```
solute lock software damn

damn fuck touch

nigga just catch shit

freke forever update reply texts sorry
wait please wait little longer

ipod touch sold annoy many glitch issue

bring tother

ran battery

ugly franken camera

icross application great idea limit
platform monitor very neat

happen type hahaha

best user guide folk bamboozl beast help

dying wait replace piece

colleague ringtone tone nokia confuse
```

Figure 6. Sample of Datasets

Table 7. Sample of Detectors from NSA Algorithm

Positive Detector	Neutral Detector	Negative Detector
positive,attention,2	neutral,phones4u,2	negative,solution,1
positive,user,6	neutral,agree,2	negative,lock,1
positive,adobe,3	neutral,deal,3	negative,software,4
positive,free,8	neutral,stock,2	negative,damn,6
positive,pretty,6	neutral,guardian,1	negative,fuck,4
positive,cool,13	neutral,open,1	negative,touch,6
positive,extreme,1	neutral,process,3	negative,nigga,1
positive,slick,1	neutral,develop,9	negative,catch,1
positive,bloomberg,1	neutral,twitter,13	negative,shit,3
positive,bring,2	neutral,application,48	negative,freke,1
positive,just,16	neutral,shorten,1	negative,forever,2

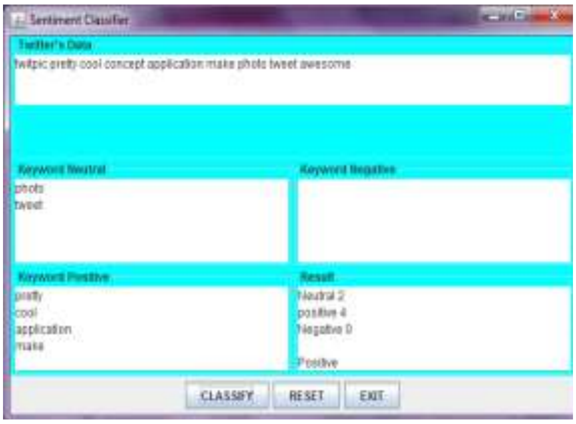
Table 7 shows the sample of detector generated by NSA algorithm for positive, neutral and negative detectors. The accuracy of proposed classification model in sentiment classification using three selected algorithms determined by the percentage of correctly classified in testing dataset.

Table 8 shows the result analysis for the selected AIS algorithms in sentiment classification phase.

Table 8. Immune System Based Learning Method Classification Analysis

AIS Algorithm	Negative Selection	Clonal Selection	Artificial Immune Network
Learning activities	Generation, detection, matching	Affinity analysis, cloning, mutation	Cloning, mutation, affinity analysis, suppress network
Classifier	Set of detector	Highest affinity list of word	List of detectors with highest affinity
Correctly Classified	76	72	67
Incorrectly Classified	44	48	53
Accuracy	63.33%	60.00%	55.83%

In this study, the accuracy for AIS selected algorithms is between 56 to 63%. The results show that the NSA algorithm has slightly higher accuracy as compared to CSA and AiNet algorithms. The accuracy is considered acceptable for the dataset and may be in future need some enhancements for better accuracy in classification process. As an example of application, the prototype system that applied selected AIS classification model for sentiment analysis is shown Figure .



twitpic pretty cool concept application make photo tweet awesome

photo
tweet

pretty
cool
application
make

Neutral 2
positive 4
Negative 0

Positive

Figure 7. Prototype System for Sentiment Analysis

V CONCLUSION

In this study, lexicon-based approach is used in sentiment feature selection process and machine learning approach that used for sentiment classification. AIS algorithm as immune system based learning is proposed method for sentiment classification. As a result, NSA algorithm produced a little bit higher accuracy compared to CSA and AiNet algorithm. In future work, the comparative study using other machine learning technique such as bio-inspired algorithm, nature-based algorithm and many other. Besides that, other techniques for data preprocessing such as back-forward stemming algorithm and porter algorithm would be considered as alternative approach in sentiment feature selection process. This would give a new direction for sentiment classification. As a conclusion, the ability to obtain new understanding of AIS algorithm in sentiment classification will lead to the imperative contribution in sentiment analysis area.

REFERENCES

Agarwal, B., & Mittal, N. (2015). Machine Learning Approach for Sentiment Analysis *Prominent Feature Extraction for Sentiment Analysis* (Vol. 2, pp. 21-45). Switzerland: Springer International Publishing

Akhmedova, S., Semenkin, E., & Sergienko, R. (2014). *Automatically Generated Classifier for Opinion Mining with Different Term Weighting Schemes*. Paper presented at the 11th International Conference on Informatics in Control, Automation & Robotics (ICINCO), Vienna, USA.

Alexander, P., & Paroubek, P. (2010). *Twitter as a Corpus for Sentiment Analysis and Opinion Mining.* Paper presented at the *Proceedings of the Seventh International Conference on Language Resources and Evaluation* Valletta, Malta.

Azadeh, A., Taghipour, M., Asadzadeh, S. M., & Abdollahi, M. (2014). Artificial immune simulation for improved forecasting of electricity consumption with random variations. *Electrical Power and Energy Systems, 55*, 205 – 224.

Balahur, A., & Steinberger, R. (2009). *Rethinking Sentiment Analysis in the News: from Theory to Practice and back.* Paper presented at the Procedding of the 1st Workshop on Opinion Mining and Sentiment Analysis.

Berna, H. U., & Sadan, K.-k. (2011). A Review of Clonal Selection Algorithm and Its Applications. *The Artificial Intelligence Review, 36*(2), 117-138.

Castro, L. N. d., & Timmis, J. I. (2003). Artificial Immune Systems as A Novel Soft Computing Paradigm. *Soft Computing 7*(8), 526-544. doi: 10.1007/s00500-002-0237-z

Chandrasekaran, S., & Murugappan, C. D. A. (2008). *An Artificial Immune Networking Using Intelligent Agents.* Paper presented at the Proceedings of the World Congress on Engineering London.

Choi, Y., Kim, Y., & Myaeng, S.-H. (2009). *Domain-specific sentiment analysis using contextual feature generation*. Paper presented at the Proceedings of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion, Hong Kong, China.

Dasgupta, D. (2006). Advances in Artificial Immune Systems. *Computational Intelligence Magazine, IEEE 1*(4), 40-49.

Dinu, L. P., & Iuga, I. (2012). *The Naive Bayes Classifier in Opinion Mining: In Search of The Best Feature Set*. Paper presented at the 3th international conference on Computational Linguistics and Intelligent Text Processing, New Delhi, India.

Fran, F. O. d., Zuben, F. J. V., & Castro, L. N. d. (2005). *An Artificial Immune Network for Multimodal Function Optimization on Dynamic Environments*. Paper presented at the *Proceedings of the 2005 conference on Genetic and evolutionary computation*, Washington DC, USA.

Galeano, J. C., Veloza-Suan, A., & Gonzlez, F. A. (2005). *A comparative analysis of artificial immune network models*. Paper presented at the *Proceedings of the 2005 conference on Genetic and evolutionary computation*, Washington DC, USA.

Hou, H., & Dozier, G. (2006). *An Evaluation of Negative Selection Algorithm with Constraint-based Detectors.* Paper presented at the Proceedings of the 44th annual Southeast regional conference, Melbourne, Florida , USA.

Hunt, J., & Cooke, D. (1996). Learning using An Artificial Immune System. *Journal or Network and Computer Applications, 19*, 189-212.

Joachims, T. (2005). Text Categorization with Suport Vector Machines: Learning with Many Relevant Features. *Machine Learning: ECML-98*, 137-142.

Khan, A. Z. H., Atique, M., & Thakare, V. M. (2015). Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis. *Special Issue of International Journal of Electronics, Communication & Soft Computing Science and Engineering, 1*, 89-75.

Kontopoulos, E., Berberidis, C., Dergiades, T., & Bassiliades, N. (2013). Ontology-based Sentiment Analysis of Twitter Posts. *Expert Systems with Applications, 40*, 4065–4074.

Kouloumpis, E., Wilson, T., & Moore, J. (2011). *Twitter Sentiment Analysis: The Good the Bad and the OMG!* Paper presented at the The Fifth International AAAI Conference on Weblogs and Social Media, Barcelona, Spain.

Kucuktunc, O., Cambazoglu, B. B., Weber, I., & Ferhatosmanoglu, H. (2012). *A Large-Scale Sentiment Analysis for Yahoo! Answers.* Paper presented at the 5th ACM International conference, New York, USA.

Li, M., kiewicz, & Textor, J. (2010). *Negative Selection Algorithms without Generating Detectors*. Paper presented at the *"Proceedings of the 12th annual conference on Genetic and evolutionary computation"*, Portland, Oregon, USA.

Liu, R., Sheng, Z., & Jiao, L. (2009). *Gene transposon based clonal selection algorithm for clustering*. Paper presented at the *"Proceedings of the 11th Annual conference on Genetic and evolutionary computation"*, Montreal, Canada.

McKaughan, D. C., Heath, Z., & McClain, J. T. (2011). *Using a text analysis and categorization tool to generate Bayesian belief networks for use in cognitive social simulation from a document corpus.* Paper presented at the Proceedings of the 2011 Military Modeling & Simulation Symposium, Boston, Massachusetts.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams Engineering Journal, 5*, 1093-1113.

Mudinas, A., Zhang, D., & Levene, M. (2012). *Combining Lexicon and Learning based Approaches for Concept-Level Sentiment Analysis.* Paper presented at the The First International Workshop on Issues of Sentiment Discovery & Opinion Mining, New York, USA.

Paltoglou, G., & Thelwall, M. (2012). Twitter, MySpace, Digg: Unsupervised Sentiment Analysis in Social Media. *ACM Transactions on Intelligent Systems and Technology, 3*(4), 1-19.

Pang, B., & Lee, L. (2004). *A Sentimental Education: Sentiment Analysis using Subjectivity Summarization Based on Minimum Cuts*. Paper presented at the Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics, Barcelona, Spain.

Puteh, M., Isa, N., Puteh, S., & Redzuan, N. A. (2013). *Sentiment Mining of Malay Newspaper (SAMNews) Using Artificial Immune System.* Paper presented at the The World Congress on Engineering London, UK.

Refaee, E., & Rieser, V. (2014). *An Arabic Twitter Corpus for Subjectivity and Sentiment Analysis*. Paper presented at the Ninth International Conference on Language Resources and Evaluation, Reykjavik, Iceland.

Samsudin, N., Puteh, M., Hamdan, A. R., & Nazri, M. Z. A. (2013). *Immune Based Feature Selection for Opinion Mining.* Paper presented at the The World Congress on Engineering, London, UK.

Sarafijanovic, & Boudec, L. (2005). An Artificial Immune System Approach with Secondary Response for Misbehavior Detection in Mobile Ad hoc Networks. *Neural Networks, IEEE Transactions on, 16*(5), 1076-1087.

Singhal, K., Agrawal, B., & Mittal, N. (2015). Modeling Indian General Elections: Sentiment Analysis of Political Twitter Data. *Information Systems Design and Intelligent Applications*, 469-477.

Tan, C., Lee, L., Tang, J., Jiang, L., Zhou, M., & Li, P. (2011). *User-Level Sentiment Analysis Incorporating Social Networks*. Paper presented at the Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, San Diego, California, USA.

Tan, S., & Zhang, J. (2008). An Empirical Study of Sentiment Analysis for Chinese Documents. *Expert Systems with Applications, 34*, 2622–2629.

Tumasjan, A., Sprenger, T., Sandner, P., & Welpe, I. (2010). *Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment.* Paper presented at the International AAAI Conference on Weblogs and Social Media, Germany.

Venkatesan, S., Baskaran, R., Chellappan, C., Vaish, A., & Dhavachelvan, P. (2013). Artificial immune system based mobile agent platform protection. *Computer Standards & Interfaces, 35*(4), 365 – 373.

Xiaodong Li, Haoran Xie, Li Chen, Jianping Wanga, & Xiaotie Deng. (2014). News Impact on Stock Price Return via Sentiment Analysis. *Knowledge-Based Systems, 69*, 14-23.

Zhang, L., Ghosh, R., Dekhil, M., Hsu, M., & Liu, B. (2011). Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis. HP Laboratories: Hewlett-Packard Development Company.