Bulletin of Electrical Engineering and Informatics
Vol. 12, No. 1, February 2023, pp. 395~402
ISSN: 2302-9285, DOI: 10.11591/eei.v12i1.4228

Classifying thai news headlines using an artificial neural network

Benjamin Chanakot[1], Charun Sanrach[2]
[1]Department of Innovation Management and Business Information, Faculty of Management Technology, Rajamangala University of Technology Srivijaya, Nakhon Si Thammarat, Thailand
[2]Department of Computer Education, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

Article Info

Article history:

Received Jun 10, 2022
Revised Sep 6, 2022
Accepted Oct 12, 2022

Keywords:

Artificial neural network
Chi-square
Classification
Information gain
TFICF

ABSTRACT

This research aimed to measure the effectiveness of Thai news headlines classification using an artificial neural network (ANN). The headlines consisted of i) political news, ii) sports news, iii) economic news, and iv) crime news, 1,200 headlines in total. The distribution of headlines was measured by using chi-square, information gain, and term frequency inverse class frequency (TFICF). Threshold default value was set in relation to terms of headlines before cross-validation was employed to categorize the data to examine the efficiency of the model using a neural network algorithm in classifying the headlines. The investigation of the news headline classification efficiency revealed that the 15-fold data division using TFICF was the most accurate in classifying headlines, with the accuracy rate of 99.60% and F-measure rate of 99.05%. Moreover, it was found that when more news headlines were provided as the learning data, the news headline classification became more accurate. Likewise, appropriate threshold value determination facilitated the selection of appropriate features in the headlines and resulted in more effective and accurate classification. Hence, it can be concluded that headline classification will be more accurate if the appropriate amount of learning data exists, and appropriate threshold value was set.

Corresponding Author:

Benjamin Chanakot
Department of Innovation Management and Business Information, Faculty of Management Technology
Rajamangala University of Technology Srivijaya
109 Village No. 2, Thamyai sub-district, Thungsong District, Nakhon Si Thammarat 80110, Thailand
Email: benjamin.c@rmutsv.ac.th

1. INTRODUCTION

Advances in information technology nowadays result in greater use of various information in electronic formats including information in the news or various documents. Electronic information tends to continuously increase in volume until it is challenging to be searched or classified. Also, a large volume of documents affects the search in terms of the accuracy and speed. Previous studies have focused exclusively on investigating the classification of documents in English. Research on the classification of documents in Thai is scarcely found. More importantly, Thai language has unique characteristics when compared to others languages in that there is no space between words in the written form, so it can be ambiguous [1]. This has a negative impact on the effectiveness and accuracy of document classification.

In responding to the challenge mentioned above, the solution is to apply machine learning for text clustering or text classification. In this context, text clustering is an unsupervised learning method [2] in which documents are classified according to the content [3] where documents with similar characteristics are grouped together. In contrast, text classification refers to a supervised learning method which depends on practicing and learning [4], [5] to classify documents according to their content using word attributes as their features. In the text classification process, it is necessary to perform text mining which is the process of extracting and analyzing texts in a large database [6] to discover the pattern or feature of a text in an unstructured or structured document in natural language, [7] integrating natural language processing with machine learning to serve different purposes.

As a result, this study aims to investigate the efficiency of Thai news headlines classification through chi-square, information gain, term frequency inverse class frequency (TFICF), and threshold default value in processing the mean value of each word in each category using artificial neural network (ANN) in classifying news headlines. The findings will provide an efficient selection of news headline keywords which will result in high quality classification. In addition, it will provide a guideline for effective news headline classification.

2. METHOD

2.1. Data mining

Data mining refers to the process of data extracting from a large database [8] to look for a pattern [9] or for useful and interesting information which can be used in making prediction. The steps involved in text mining include i) data cleaning, the process of removing useless data on the database, ii) data integration, the process of data compilation, iii) data transformation, the process of transforming data suitable for data analysis, iv) data selection, the process of selecting useful data for data analysis, v) data mining, the process of using data to create a model, vi) evaluation of patterns, the process of model evaluation, and vii) knowledge presentation, the process of presenting the results obtained from the model [10], [11].

2.2. Classification

Classification refers to data classifying through a machine learning method in which a supervised learning technique where old data is used in creating a model to predict what will happen in the future [12]. To be specific, a certain amount of the data is provided for training data in creating a model [13], while the other part is used to test the efficiency of the model, known as testing data.

2.3. Steps of model development

The classification of Thai news headlines using neural network algorithms in this research involves eight steps of developing shown in Figure 1. Thai news headline classification model as follows: i) data collection, ii) data preprocessing, iii) feature selection, iv) feature weighting, v) vector space model (VSM), vi) cross-validation, vii) classification, and viii) measuring model performances.

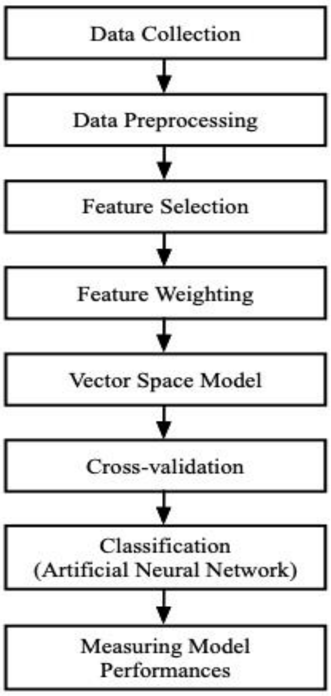

Figure 1. Steps of model development

2.4. Data collection

Thai news headlines from the following Thai news agencies were collected: i) www.thairath.co.th, ii) www.khaosod.co.th, iii) www.dailynews.co.th, and iv) www.matichon.co.th. There were 1,200 news headlines in total. They were divided into four classes: i) political news, ii) sports news, iii) economic news, and iv) crime news.

2.5. Data preprocessing

This step involves data preparation processes before performing a headline classification performance test. There were 3 sub-steps involved as follows: i) word segmentation, the process of dividing each Thai words before processing the data with natural language in the next step. Because of news written Thai, words are put together without any punctuation marks which is contrastive to English where each word is discretely separated, ii) stop word removal, the process of removing common words which are considered insignificant [14]–[16] from documents. They frequently occur in documents and are used to connect sentences or complete texts in the documents. Therefore, it is necessary to remove these words from the documents, as they are not useful in document classifying [17]. These insignificant words include: pronouns, adverbs, interjections, prepositions, conjunctions and symbols such as (!, #, +, -,*,/,=,….) [18], and iii) stemming word, the process of substituting words that have the same meaning or words with the same root with only one word. This reduces the number of redundant words in the document and help increase the efficiency of document classification [19]–[21]. Figure 2 shows the data preparation process which consists of 3 steps: i) word segmentation, the process of word cutting by comparing with words archived in a dictionary, ii) stop word removal, the process of removing words that are not important from the document by comparing it to the stop words stored on the database, and iii) word stemming, the process of replacing words with the same meaning or words with the same root with a certain word by comparing with the database.

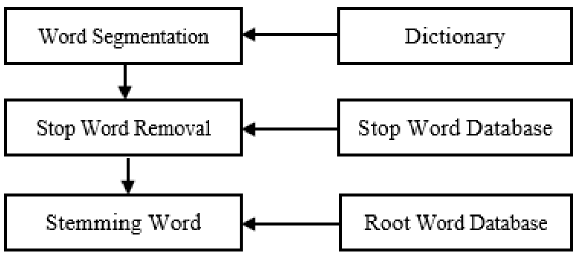

Figure 2. Step of data preprocessing

2.6. Feature selection

Feature selection involves selecting appropriate word attributes or words which are significant in feature subsets from all document attributes by eliminating duplicate, overlapping and irrelevant attributes from the document [22]. This reduces the number of features and assists in selecting the word attributes that are important in the classification of the document [23]. In this research, word attribute selection was performed using the chi-square, information gain, and the TFICF method (to calculate weight of terms).

2.7. Chi-square

Chi-square is a statistical calculation method used in examining the correlation between word features and document categories. Feature selection is processed by using frequencies of term (t) and possibility of term (t) occurring in each classification. Term characteristics were selected based on the frequency of term (t) and the probability of occurrence of term (t) in document group (C) [24] as shown in (1) [25].

$$Chi-square_{(t_k,c_i)} = \frac{N(AD-CB)^2}{(A+C)(B+D)(A+B)(C+D)} \quad (1)$$

Where,
N is the total number of documents
A is the number of documents in the c_i group in which the term t_k exits
B is the number of documents in which the term t_k appears in other groups of documents
C is the number of documents of group c_i in which the term t_k does not appear
D is the number of other groups of documents in which the term t_k does not appear

2.8. Information gain

Information gain, a method for selecting features of keywords from a group of documents, is one of the most widely used methods. The selection is performed based on entropy theory [26] by using a prediction of documents in the i category in which the term t appears and does not appears as shown in (2) [27].

$$IG(t) = -\sum P(ci)logP(ci) + P(t)\sum P(ci|t)logP(ci|t) + P(\bar{t})\sum P(ci|\bar{t})logP(ci|\bar{t}) \quad (2)$$

Where,
C is a group of documents
$P(ci)$ is the probability of a document in group i
$P(t)$ and $P(\bar{t})$ are the probability that the word t appears and does not appear in the document
$P(ci|t)$ is the probability condition for documents in group i in which the term t appear in the document
$P(ci|\bar{t})$ is the probability condition for documents in group i in which the t term does not appear in the document

2.9. TFICF

TFICF is a term weighting method to determine the correlation between a keyword and the documents of a collection, where $f_{t,d}$ is the frequency of the t term appearing in document d [28], [29] and cf is adapted from idf (term frequency inverse document frequency) [30] in which $|C|$ is the number of document categories and cf_t is the number of document categories in which the word t appears as shown in (3) [29].

$$tficf_t = f_{t,d} \times log_2(\frac{|C|}{cf_t}) \quad (3)$$

2.10. Feature weighting

Feature weighting is a process in which weight is assigned to each feature that promotes accuracy of document classification in the documents of a collection [31]. The weight of each term in the document is assigned according to its attributes [32]. As a result, in this study, the average frequency of terms in each document collection was set as a threshold value and the frequency of the term selected as a document attribute must be greater than or equal to the value in each document collection due to the fact that the term has a high frequency and is important to the document. Besides, it can represent a document and also affects the classification efficiency of the document more than terms with low frequency.

2.11. Vector space model

VSM is a mathematical model used for document classification. Each of its dimension is represented by the weight of word attributes in a document [33] in the matrix. In (4) [34] shows a VSM, in which d_j is the document and $w_{i,j}$ is the weight of t_j.

$$d_j = (t_1, w_{1,j}; t_2, w_{2,j}; \dots; t_n, w_{n,j}) \quad (4)$$

2.12. Cross-validation

In order to evaluate the effectiveness of the model, cross-validation was employed in this study to classify the data. The key concept was to divide the data into K sets, each set was the same size. Then some sets of the data were used in examining the effectiveness of the model so called the testing set, while another set was used in the model training process known as the training set. These two steps were repeated over and over until all sets of the data were processed as a training set.

2.13. Artificial neural network

ANN is a mathematical model in which a nonlinear learning element is incorporated to mimic the function of the human brain [35]. The key concept of the ANN is being progressive. In general, an ANN usually consists of three layers as show in Figure 3 which are i) the input layer, ii) the hidden layer, and iii) the output layer [36]. Every node in the same layer is connected to all nodes in the next layer [37]. As related weighted links are used in connecting between nodes, the outputs of the ANN depend on the modulation of the weight of the link. After that the data will be sent to the input layer before the output is displayed in the output layer [38]. Normally, the number of nodes in the input layer depends on the number of features of the dataset to be analyzed. Since the hidden layer affects the ability to learn about the model, the number of nodes on this layer depends on the needs of users and more nodes can be added for more accuracy. Meanwhile, the output layer is a part where the outputs of the ANN are presented [38]. The number of nodes in this layer depends on the format of the data to be classified.

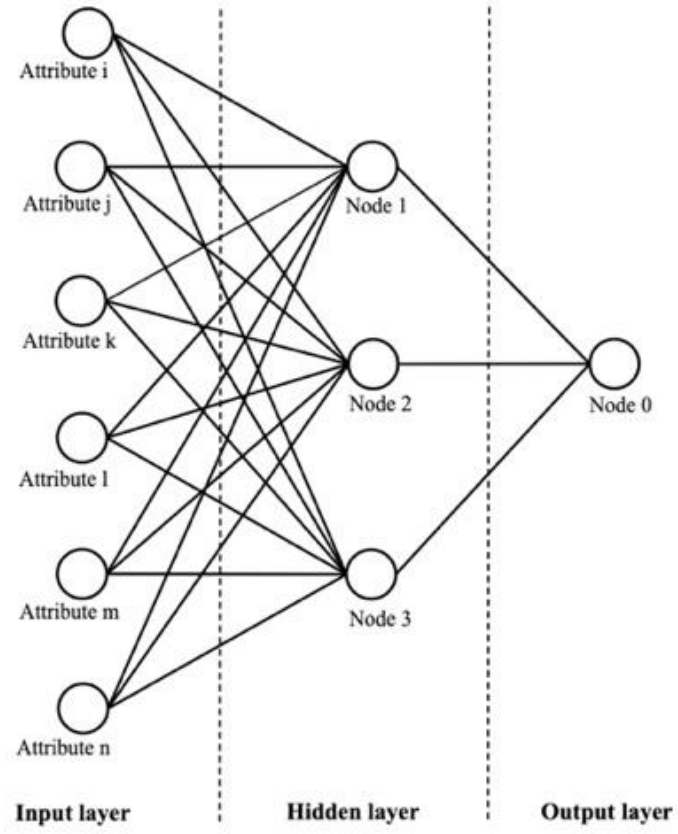

Figure 3. Network of neural network

2.14. Measuring model performances

The measurement of the model effectiveness was processed in the following areas: i) precision, ii) recall, iii) F-measure, and iv) accuracy. Each of them can be measured through the (5)-(8) [39].

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

$$F - measure = \frac{2\times Precision \times Recall}{Precision+Recall} \quad (7)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

In which TP is the number of documents correctly guessed as $Class\ C$, TN is the number of documents correctly guessed not $Class\ C$, FP is the number of documents incorrectly guessed as $Class\ C$, and FN is the number of documents incorrectly guessed as not $Class\ C$.

3. RESULTS AND DISCUSSION

According to the headline classification performance test through term weighting by the methods of information gain, chi-square, TFICF and frequency labeling of terms in each group of documents as a threshold value, the details are as follows: threshold=4 for political news, threshold=4 for sport news, threshold=3 for economic news, and threshold=5 for crime news. Then the data was divided into 3 types through cross-validation including 5-fold, 10-fold, 15-fold. During this stage, the algorithms in the neural network calculated the accuracy, precision, recall, and F-measure to measure the effectiveness of the model. The results are shown in Table 1. As shown in Table 1, the performance test of Thai news headlines classification using the ANN revealed that the TFICF method was with the highest accuracy and F-measure score, 99.60 and 99.05, respectively. This can be inferred that the TFICF weighting method is more accurate in classifying headlines than other methods.

It was also found that when classifying sports news headlines and the minimum threshold value was set, the number of words that serve as headline features became greater than that of other news categories. Consequently, the higher number of features promotes the effectiveness of the headline classification. As shown in Figure 4, when calculating the weights of terms using information gain, chi-square and TFICF methods in which 15-fold was applied in dividing the data, higher precision and F-measure scores were found regardless of the methods. Therefore, it can be concluded that the amount of the learning data affects the efficiency and accuracy of news headlines classification.

Table 1. The results of the effectiveness test of Thai news headlines classification using the ANN

Cross-validation	Feature select	Accuracy	Precision	Recall	F-measure
5-fold	Information gain	87.98	97.35	71.75	82.61
	Chi-square	89.80	81.25	74.09	77.50
	TFICF	87.90	70.62	71.06	70.84
10-fold	Information gain	95.21	96.91	88.00	92.24
	Chi-square	97.91	92.04	98.18	95.01
	TFICF	99.21	96.85	99.35	98.08
15-fold	Information gain	99.09	98.23	98.67	98.45
	Chi-square	99.01	98.86	96.66	97.75
	TFICF	99.60	99.37	98.75	99.05

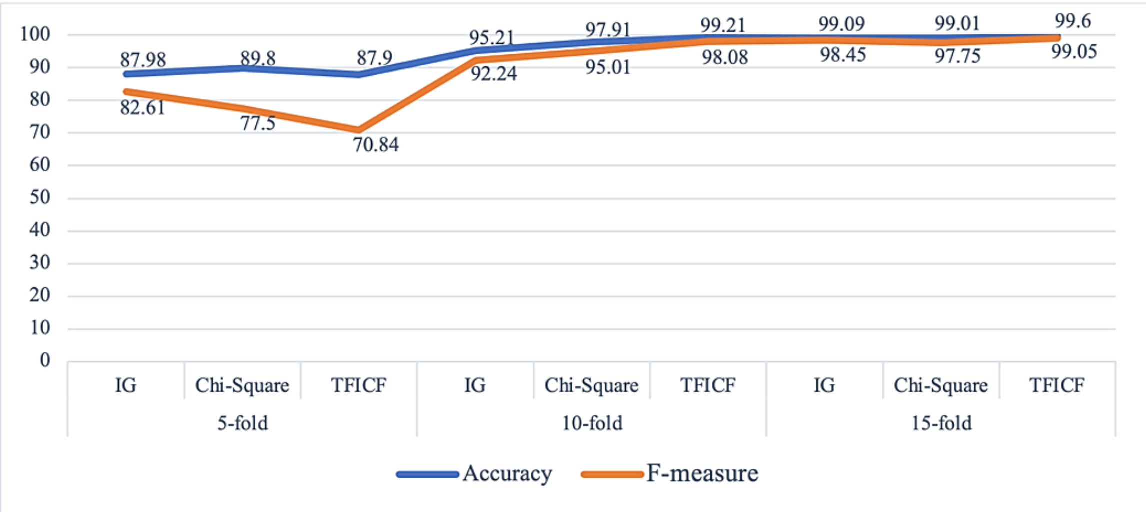

Figure 4. The efficiency of news headlines classification

4. CONCLUSION

This research aimed to test the efficiency of news headlines classification using the information gain, chi-square, TFICF methods in determining weights of terms. The cross-validation method was applied in dividing the data to evaluate the effectiveness of the model, while neural network algorithms were applied in classifying the headlines. It was found that the efficiency and accuracy of the classification depends greatly on the amount of the learning data and the numbers of terms set as the headline features. In addition, when the learning data is adequate and the terms functioning as features are able to represent the document, then news headline classification will be more effective.

ACKNOWLEDGEMENTS

We would like to thank the Faculty of Management Technology, Rajamangala University of Technology Srivijaya for their support to see this research reach completion.

REFERENCES

[1] P. Prakrankamanant and E. Chuangsuwanich, "Tokenization-based data augmentation for text classification," in *2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, Jun. 2022, pp. 1–6, doi: 10.1109/JCSSE54890.2022.9836268.

[2] J. Li, W. Liu, M. Liu, and M. Huang, "Study on Chinese Text Clustering Algorithm Based on K-mean and Evaluation Method on Effect of Clustering for Software-intensive System," *Proc. - 2020 Int. Conf. Comput. Eng. Appl. ICCEA 2020*, pp. 513–519, 2020, doi: 10.1109/ICCEA50009.2020.00114.

[3] T. Jo, "String vector based AHC for text clustering," *Int. Conf. Adv. Commun. Technol. ICACT*, pp. 673–678, 2017, doi: 10.23919/ICACT.2017.7890178.

[4] A. K. Al-Tamimi, E. Bani-Isaa, and A. Al-Alami, "Active Learning for Arabic Text Classification," *Proc. 2nd IEEE Int. Conf. Comput. Intell. Knowl. Econ. ICCIKE 2021*, pp. 123–126, 2021, doi: 10.1109/ICCIKE51210.2021.9410758.

[5] M. Sahan, V. Smidl, and R. Marik, "Active Learning for Text Classification and Fake News Detection," *Proc. - 2021 Int. Symp. Comput. Sci. Intell. Control. ISCSIC 2021*, pp. 87–94, 2021, doi: 10.1109/ISCSIC54682.2021.00027.

[6] P. K. Jayasekara and K. S. Abu, "Text Mining of Highly Cited Publications in Data Mining," *IEEE 5th Int. Symp. Emerg. Trends Technol. Libr. Inf. Serv. ETTLIS 2018*, pp. 128–130, 2018, doi: 10.1109/ETTLIS.2018.8485261.

[7] J.-C. Kim and K. Chung, "Associative feature information extraction using text mining from health big data," *Wirel. Pers. Commun.*, vol. 105, no. 2, pp. 691–707, Mar. 2019, doi: 10.1007/s11277-018-5722-5.

[8] S. Shrestha and M. Pokharel, "Educational data mining in moodle data," *Int. J. Informatics Commun. Technol.*, vol. 10, no. 1, p. 9, 2021, doi: 10.11591/ijict.v10i1.pp9-18.

[9] J. F. Andry, F. M. Silaen, H. Tannady, and K. H. Saputra, "Electronic health record to predict a heart attack used data mining with Naïve Bayes method," *Int. J. Informatics Commun. Technol.*, vol. 10, no. 3, p. 182, 2021, doi: 10.11591/ijict.v10i3.pp182-187.

[10] T. R. S. Mary and S. Sebastian, "Predicting heart ailment in patients with varying number of features using data mining techniques," *Int. J. Informatics Commun. Technol.*, vol. 8, no. 1, p. 56, 2019, doi: 10.11591/ijict.v8i1.pp56-62.

[11] C. Kaushal and D. Koundal, "Recent trends in big data using hadoop," *Int. J. Informatics Commun. Technol.*, vol. 8, no. 1, p. 39, 2019, doi: 10.11591/ijict.v8i1.pp39-49.

[12] S. Khatri, A. Arora, and A. P. Agrawal, "Supervised Machine Learning Algorithm for Credit Card Fraud Detection : A Comparison," *2020 10th Int. Conf. Cloud Comput. Data Sci. Eng.*, pp. 680–683, 2020.

[13] R. Saravanan and P. Sujatha, "A State of Art Techniques on Machine Learning Algorithms : A Perspective of Supervised Learning Approaches in Data Classification," *2018 Second Int. Conf. Intell. Comput. Control Syst.*, no. Iciccs, pp. 945–949, 2018.

[14] T. D. Roy, S. Khatun, R. Begum, and A. M. S. Chowdhury, "Vector Space Model based Topic Retrieval from Bengali Documents," *2018 Int. Conf. Innov. Sci. Eng. Technol. ICISET 2018*, no. October, pp. 60–63, 2018, doi: 10.1109/ICISET.2018.8745587.

[15] S. E. Pratama, W. Darmalaksana, D. Sa'adillah Maylawati, H. Sugilar, T. Mantoro, and M. A. Ramdhani, "Weighted inverse document frequency and vector space model for hadith search engine," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 18, no. 2, pp. 1004–1014, 2020, doi: 10.11591/ijeecs.v18.i2.pp1004-1014.

[16] M. Fikri and R. Sarno, "A comparative study of sentiment analysis using SVM and Senti Word Net," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 13, no. 3, pp. 902–909, 2019, doi: 10.11591/ijeecs.v13.i3.pp902-909.

[17] R. Janani and S. Vijayarani, "Automatic text classification using machine learning and optimization algorithms," *Soft Comput.*, vol. 25, no. 2, pp. 1129–1145, 2021, doi: 10.1007/s00500-020-05209-8.

[18] H. N. Alshaer, M. A. Otair, and L. Abualigah, "Improved ICHI square feature selection method for Arabic classifiers," *Int. J. Informatics Commun. Technol.*, vol. 9, no. 3, p. 157, 2020, doi: 10.11591/ijict.v9i3.pp157-170.

[19] A. I. Kadhim, "Survey on supervised machine learning techniques for automatic text classification," *Artif. Intell. Rev.*, vol. 52, no. 1, pp. 273–292, 2019, doi: 10.1007/s10462-018-09677-1.

[20] E. H. Muktafin and P. Kusrini, "Sentiments analysis of customer satisfaction in public services using K-nearest neighbors algorithm and natural language processing approach," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 19, no. 1, pp. 146–154, 2021, doi: 10.12928/TELKOMNIKA.V19I1.17417.

[21] C. Kaewphet and N. Wisitpongpun, "Algorithm for extracting product feature from e-commerce comment," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 22, no. 2, pp. 1199–1207, 2021, doi: 10.11591/IJEECS.V22.I2.PP1199-1207.

[22] N. K. Suchetha, A. Nikhil, and P. Hrudya, "Comparing the wrapper feature selection evaluators on twitter sentiment classification," *ICCIDS 2019 - 2nd Int. Conf. Comput. Intell. Data Sci. Proc.*, pp. 0–5, 2019, doi: 10.1109/ICCIDS.2019.8862033.

[23] Y. Chen and P. Sun, "A Chinese Text Classier Based on Strong Class Feature Selection and Bayesian Algorithm," *Proc. - 2019 Int. Conf. Intell. Transp. Big Data Smart City, ICITBS 2019*, pp. 540–543, 2019, doi: 10.1109/ICITBS.2019.00135.

[24] Y. Zhai, W. Song, X. Liu, L. Liu, and X. Zhao, "A Chi-square Statistics Based Feature Selection Method in Text Classification," *2018 IEEE 9th Int. Conf. Softw. Eng. Serv. Sci.*, pp. 160–163, 2018.

[25] S. Bahassine, A. Madani, M. Al-Sarem, and M. Kissi, "Feature selection using an improved Chi-square for Arabic text classification," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 32, no. 2, pp. 225–231, 2020, doi: 10.1016/j.jksuci.2018.05.010.

[26] C. Yin and J. Xi, "Maximum entropy model for mobile text classification in cloud computing using improved information gain algorithm," *Multimed. Tools Appl.*, vol. 76, no. 16, pp. 16875–16891, 2017, doi: 10.1007/s11042-016-3545-5.

[27] Z. Gao, Y. Xu, F. Meng, F. Qi, and Z. Lin, "Improved information gain-based feature selection for text categorization," *2014 4th Int. Conf. Wirel. Commun. Veh. Technol. Inf. Theory Aerosp. Electron. Syst. VITAE 2014 - Co-located with Glob. Wirel. Summit*, 2014, doi: 10.1109/VITAE.2014.6934421.

[28] M. Abdel Fattah, "New term weighting schemes with combination of multiple classifiers for sentiment analysis," *Neurocomputing*, vol. 167, pp. 434–442, 2015, doi: 10.1016/j.neucom.2015.04.051.

[29] J. Polpinij, "A method of non-bug report identification from bug report repository," *Artif. Life Robot.*, vol. 26, no. 3, pp. 318–328, 2021, doi: 10.1007/s10015-021-00681-3.

[30] J. P. Carvalho, H. Rosa, and F. Batista, "Detecting relevant tweets in very large tweet collections: The London Riots case study," *IEEE Int. Conf. Fuzzy Syst.*, no. 1, 2017, doi: 10.1109/FUZZ-IEEE.2017.8015635.

[31] Q. Jiang, W. Wang, X. Han, S. Zhang, X. Wang, and C. Wang, "Deep feature weighting in Naive Bayes for Chinese text classification," *Proc. 2016 4th IEEE Int. Conf. Cloud Comput. Intell. Syst. CCIS 2016*, pp. 160–164, 2016, doi: 10.1109/CCIS.2016.7790245.

[32] X. Pang and Y. Liao, "A text classification model based on training sample selection and feature weight adjustement," *Proc. - 2nd IEEE Int. Conf. Adv. Comput. Control. ICACC 2010*, vol. 3, pp. 294–297, 2010, doi: 10.1109/ICACC.2010.5486615.

[33] H. Karamollaoglu, I. A. Dogru, N. Arici, and O. Yildiz, "Content Based Sentiment Analysis for Turkish Twitter Feeds Through Vector Space Model," *1st Int. Informatics Softw. Eng. Conf. Innov. Technol. Digit. Transform. IISEC 2019 - Proc.*, pp. 0–5, 2019, doi: 10.1109/UBMYK48245.2019.8965543.

[34] Z.-Q. Xiang, Y. X. Zou, and X. Wang, “Sentiment analysis of Chinese micro-blog using vector space model,” in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific*, Dec. 2014, pp. 1–5, doi: 10.1109/APSIPA.2014.7041745.

[35] A. Eghbalzadeh, M. Javan, M. Hayati, and A. Amini, “Discharge prediction of circular and rectangular side orifices using artificial neural networks,” *KSCE J. Civ. Eng.*, vol. 20, no. 2, pp. 990–996, 2016, doi: 10.1007/s12205-015-0440-y.

[36] E. Fahimi Nezhad, G. Fallah Ghalhari, and F. Bayatani, “Forecasting Maximum Seasonal Temperature Using Artificial Neural Networks ‘Tehran Case Study,’” *Asia-Pacific J. Atmos. Sci.*, vol. 55, no. 2, pp. 145–153, 2019, doi: 10.1007/s13143-018-0051-x.

[37] S. S. Ganesh, P. Arulmozhivarman, and V. S. N. R. Tatavarti, “Prediction of PM2.5 using an ensemble of artificial neural networks and regression models,” *J. Ambient Intell. Humaniz. Comput.*, vol. 0, no. 0, pp. 1–11, 2018, doi: 10.1007/s12652-018-0801-8.

[38] P. Cachim and A. Bezuijen, “Modelling the Torque with Artificial Neural Networks on a Tunnel Boring Machine,” *KSCE J. Civ. Eng.*, vol. 23, no. 10, pp. 4529–4537, 2019, doi: 10.1007/s12205-019-0302-0.

[39] L. Jiang, F. Liu, D. Xu, and W. Zhang, “A FKSVM model based on fisher criterion for text classification,” in *2017 10th International Symposium on Computational Intelligence and Design (ISCID)*, Dec. 2017, pp. 496–499, doi: 10.1109/ISCID.2017.211.

BIOGRAPHIES OF AUTHORS

Benjamin Chanakot received Ph.D. (Computer Education) form King Mongkut’s University of Technology North Bangkok, Thailand. He has been a Head of Innovation Management and Business Information Program, Faculty of Management Technology, Rajamangala University of Technology Srivijaya. He is interested in natural language processing, gesture control, data mining, and bigdata research areas. He can be contacted via email: benjamin.c@rmutsv.ac.th.

Charun Sanrach is an Assistant Professor, Ph.D. in the Department of Computer Education, Faculty of Technical Education, King Mongkut’s University of Technology North Bangkok. He received Diplôme de Docteur (Informatique) avec la mention très honorable form INPL, Nancy, France. His research interests include computer education, information technology, apply artificial intelligence in computer education, data mining, and bigdata. He can be contacted via email: charun.s@fte.kmutnb.ac.th.