Exploring Students' Feedback in Online Assessment System Using Opinion Mining Technique

Muslihah Wook, Sharmelen Vasanthan, Suzaimah Ramli, Noor Afiza Mat Razali, Nor Asiakin Hasbullah, and Norulzahrah Mohd Zainudin

***Abstract*—Opinion mining has been widely used in recent online reviews or feedback due to its ability to analyse text-based data. The use of this technique for analysing data from students' feedback needs to be addressed, since most educational institutions are focusing more on questionnaires based on the Likert-scale rather than on the open-review type. To this end, there is a lack of online assessment systems that could automatically analyse open-review questionnaires. Therefore, the main aim of this study is to analyse students' feedback in an online assessment system through the opinion mining technique, by focusing on textual form data derived from the open-review questionnaires. To achieve this aim, an opinion mining feedback system, known as the OMFeedback, was developed. The Vader Sentiment Intensity Analyser was adapted for processing students' feedback and the lexicon based approach was used for analysing the words. In addition, the OMFeedback incorporates the capitalisation of words and emoji features to enrich the capability of the system. This newly developed system could lead to new paradigms in educational institutions for enhancing students' learning process and for guiding them through their learning journey.**

***Index Terms*—Lexicon based approach, online assessment system, opinion mining, students' feedback.**

I. Introduction

Analysing students' feedback in online assessment systems is a not new concept; educational institutions use this feedback for educational insight. The purpose of obtaining such feedback is primarily to improve the teaching quality and to help the academic administration better understand students' perspective on the teaching and learning process [1]. Online assessment systems would usually consist of Likert-scale and open-review questionnaires. The Likert-scale comprises of a set of questions, such as multiple choice or scale rating, while the open-review has textual feedback questions that ask for opinions or comments. The set of questions in the Likert-scale would typically cater various dimensions, such as a lecturer's teaching style, presentation skills, time management, and course materials. Meanwhile, the textual feedback provides students with the opportunity to point out certain issues that are not directly covered in the Likert-scale questions. Balahadia *et al.* [2] asserted that most educational institutions rely on Likert-scale questions for easy analysis, compared to the textual feedback, which is hard to digest because of the high amount of text data [3], [4].

The aforementioned students' feedback would generally be available in an unstructured format, which would require extra analysing effort to gain fruitful conclusions from these feedbacks. Advanced data mining techniques that are capable of exploring the unstructured patterns inside the textual data are required. One of the most active techniques for understanding various types of students' feedback is the opinion mining technique. It is an area of research within sentiment analysis that could examine human's opinions, comments, feelings or behaviours towards various entities, such as services, products, organisations, individuals, events, issues, and highlighted topics [5]. According to Ravi and Ravi [6], opinion mining and sentiment analysis are interrelated areas of research, which extract texts or words that have been commented in numerous levels. Sivakumar and Reddy [7] reported that these texts can be classified as document level, sentence level, and word level. These different levels have text polarity that can be determined by checking the occurrences of a positive, negative or neutral word. For instance, if a positive word occurs more frequently than a negative word in a sentence, then the sentence is deemed a positive sentence.

Basically, the opinion mining technique can be divided into machine learning and lexicon based approach. The main objective of these techniques is to determine whether the written comments contain positive, negative or neutral opinions. Positive opinions represent the necessary things, negative opinions means unnecessary things, and neutral opinions are used for a better distinction between positive and negative words. Examples of positive words would include "joy", and "trust"; negative words would be "anger", and "fear"; and neutral words are "but", and "surprise" [8]. Machine learning focuses on building models with the aid of large training datasets to determine the features of text orientation [5]. Whereas, the lexicon based approach uses words contained in the sentiment dictionary (e.g., Vader Lexicon, SentiWordNet, and General Inquirer) and matches them with the collected data to determine the polarity of words [9].

Numerous machine learning techniques have been used to analyse students' feedback, such as Naïve Bayes (NB), Support Vector Machine (SVM), neural network, and k-nearest neighbour. Dhanalakshmi *et al.* [10] highlighted that NB is the most commonly used method to calculate the possibility of a given text belonging to a particular feature; SVM works best for classifying sparse text data; neural network employs multiple layers of neurons for text classification; and k-nearest neighbour uses Euclidean distance to calculate the similarity of text data. Several

Manuscript received March 29, 2020; revised June 20, 2020.

The authors are with the Department of Computer Science, Faculty of Defence Science and Technology, National Defence University of Malaysia (e-mail: muslihah@upnm.edu.my, sharmelenvasanthan@gmail.com, suzaimah@upnm.edu.my, noorafiza@upnm.edu.my, asiakin@upnm.edu.my, norulzahrah@upnm.edu.my).

studies claimed that neural network is the perfect technique for opinion mining [2], [10]–[12]. However, a study by Hutto and Gilbert [13] found that most machine learning techniques have drawbacks. First, these techniques often require large training datasets to represent various features. Second, the techniques are often computationally expensive since they need high processing time and large memory spaces. And third, the features extracted from the text are not easily interpretable and are therefore, more complex to modify, extend or generalise.

Recently, several studies revealed that the lexicon based is an optimal approach since it neither needs a large number of text data nor high computation power to produce accurate results [1], [4], [9]. This is in accordance with the results obtained by Hutto and Gilbert [13], who found that such approach can generate a high accuracy value. Previous studies on opinion mining have also shown that the capitalisation of words and the use of emoji characters may emphasise a user's intent and enhance the expressivity of the feedback [13], [14]. However, only a small number of studies have utilised these features to enrich the ability of the opinion mining technique to analyse students' feedback. Based on these reasons, this study has developed a new system to analyse students' feedback, known hereafter as the OMFeedback system. This system uses the lexicon based approach and incorporates the capitalisation of words and emoji characters that are actively used in most online-based systems.

II. Methodology

To design the OMFeedback system, this study uses sequence diagrams as shown in Fig. 1 and Fig. 2. This system consisted of students and admin as the main users. Before these users can use the system, they need to log in to the system using their username and password. Then, students need to select the name of the lecturer from a drop-down list and write their feedback regarding the teaching and learning process. The lexicon based approach, which are underpinned by Vader Sentiment Intensity Analyser (VSIA) would be used to examine and calculating the polarity of each written word in the feedback and appear the opinion results as positive, neutral and negative scores. Eventually, these feedbacks and opinion results can only be accessed by admin for further actions.

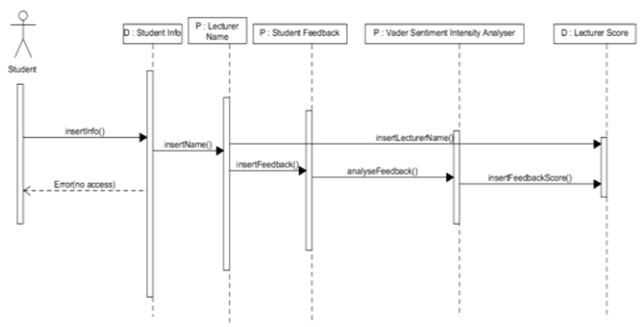


Fig. 1. Sequence diagram for student.

Fig. 3 exhibits the process of opinion mining in the OMFeedback system. This figure consists of input, process, and output phases. Initially, students' feedback is taken as inputs for the subsequent process. The VSIA will perform two sub-processes, namely, keyword identification and calculation of total scores.

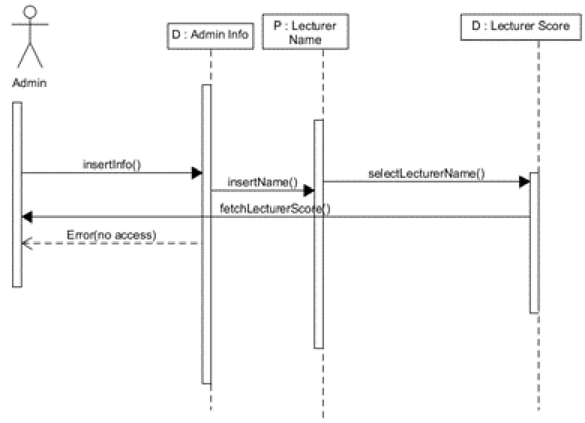


Fig. 2. Sequence diagram for admin.

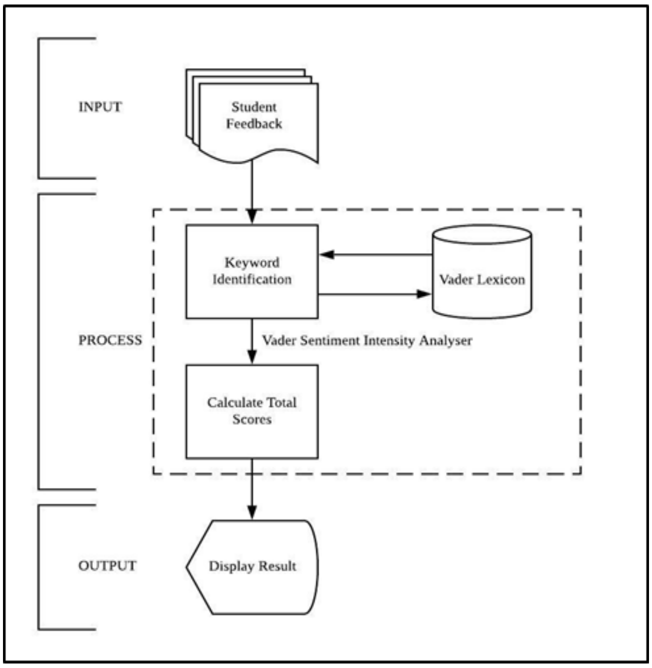


Fig. 3. The flow of the opinion mining process in OMFeedback system.

In the keyword identification process, as shown in Fig. 3, each word in the feedback are compared with the words stored in the Vader Lexicon database. This database consists of a collection of words that have fixed polarity of scores. The polarity of scores is then assigned to each word and the total scores of these words are calculated using the lexicon based approach. The OMFeedback system will display a result that contains the polarity of positive, negative, and neutral scores for all of the students' feedback. Table I exhibits some examples of polarity, words, and their scores that are stored in the Vader Lexicon database.

This study used *WxPython* and *Python 3.7* to develop the interface and the interaction of the OMFeedback system. Table II shows the algorithm of the main process in the system. First, students need to choose the name of the lecturer from a list and write down their feedback in the space provided. Then, the compiler will read these input and the VSIA will analyse the feedback text, and then assigns positive, negative, and neutral scores. The lecturer's name, the feedback text, and the score will then be stored in the

Vader Lexicon database. Finally, the sentiment scores (positive, negative, and neutral), will be displayed to the administrator in the form of pie charts. Table III shows the algorithm used by the administrator to view the results of the OMFeedback system.

TABLE I: EXAMPLE OF WORDS STORED IN THE VADER LEXICON DATABASE

Polarity of Word	Example of Words	Score
Positive	*happy*	2.7
	good	1.9
	great	3.1
	positive	2.6
	joy	2.8
Negative	*negative*	-2.7
	bad	-2.5
	awful	-2.0
	sad	-2.1
	boring	-1.3
Neutral	*but*	0
	Words that are not recorded in the Vader Lexicon database will be assigned as neutral	0

TABLE II: ALGORITHM OF A STUDENT'S TEXTUAL FEEDBACK ANALYSIS USING THE VADER SENTIMENT INTENSITY ANALYSER

Input:
Name of lecturer *lecturerName*;
Textual feedback *textualFeedback*;
Output:
Feedback score *feedbackScore*;

0. **START**
1. GET *lecturerName*;
2. GET *textualFeedback*;
3. def sentiment_analyser_scores(*textualFeedback*):
    feedbackScore = analyser.polarity_scores(*textualFeedback);*
    print("{:-<50} {}".format(*textualFeedback*, str(*feedbackScore*)));
4. Send variable *[lecturerName, courseCode, classGroup, textualFeedback, feedbackScore]* to database
5. **END**

TABLE III: ALGORITHM OF THE SENTIMENT RESULT

Input:
Name of lecturer *lecturerName;*
Output:
Feedback score *feedbackScore;*

0. **START**
1. GET *lecturerName;*
2. GET *classGroup;*
3. Search in database data that contain variable [*lecturerName, classGroup*];
4. Display sentiment result in a pie chart
5. **END**

III. EXPERIMENT AND RESULTS

This study tested the effectiveness of the OMFeedback system using feedback from third-year undergraduate students from the Department of Computer Science, National Defence University of Malaysia during the second semester for the 2018/2019 session. In this phase, 18 students were tested as a sample because they have the same lecturer. First, the students were asked to choose the name of the lecturer, class group, course name, and course code. Next, they were asked to write their opinion or comment in the feedback space provided, as shown in Fig. 4.

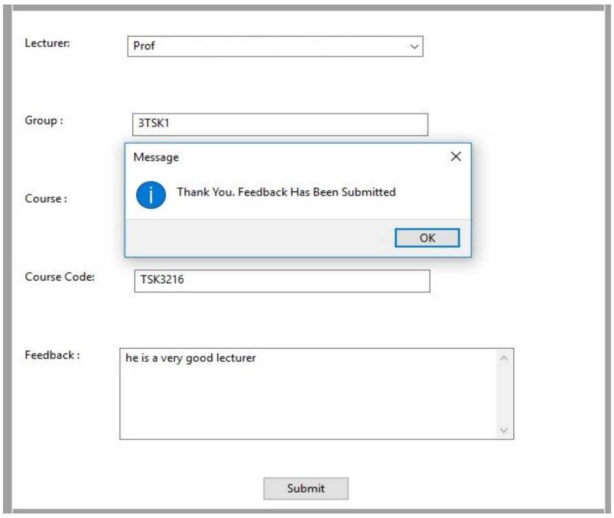
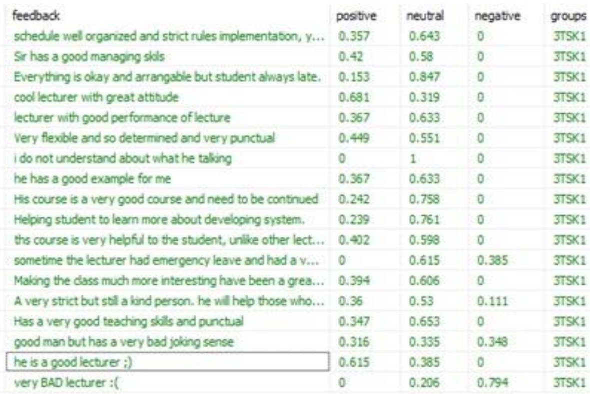


Fig. 4. The feedback form in the OMFeedback system.

Once a student has submitted the feedback form, VSIA will identify the keywords and assign the specified values obtaining from the Vader Lexicon. Then, this lexicon based approach will calculate the value of scores and display the results, as shown in Fig. 5. Based on this figure, all students' feedback has different polarity scores for being positive, negative, and neutral. The last two students have interesting scores as they use the new features of capitalised words and emoji characters to better express their feelings and opinions. For example, when a student wrote "very BAD lecturer :(", the lexicon based approach yielded a high negative score (79.4%) since both features simultaneously in a sentence.



feedback	positive	neutral	negative	groups
schedule well organized and strict rules implementation, y...	0.357	0.643	0	3TSK1
Sir has a good managing skils	0.42	0.58	0	3TSK1
Everything is okay and arrangable but student always late.	0.153	0.847	0	3TSK1
cool lecturer with great attitude	0.681	0.319	0	3TSK1
lecturer with good performance of lecture	0.367	0.633	0	3TSK1
Very flexible and so determined and very punctual	0.449	0.551	0	3TSK1
i do not understand about what he talking	0	1	0	3TSK1
he has a good example for me	0.367	0.633	0	3TSK1
His course is a very good course and need to be continued	0.242	0.758	0	3TSK1
Helping student to learn more about developing system.	0.239	0.761	0	3TSK1
ths course is very helpful to the student, unlike other lect...	0.402	0.598	0	3TSK1
sometime the lecturer had emergency leave and had a v...	0	0.615	0.385	3TSK1
Making the class much more interesting have been a grea...	0.394	0.606	0	3TSK1
A very strict but still a kind person. he will help those who...	0.36	0.53	0.111	3TSK1
Has a very good teaching skills and punctual	0.347	0.653	0	3TSK1
good man but has a very bad joking sense	0.316	0.335	0.348	3TSK1
he is a good lecturer ;)	0.615	0.385	0	3TSK1
very BAD lecturer :(	0	0.206	0.794	3TSK1

Fig. 5. List of students' textual feedback.

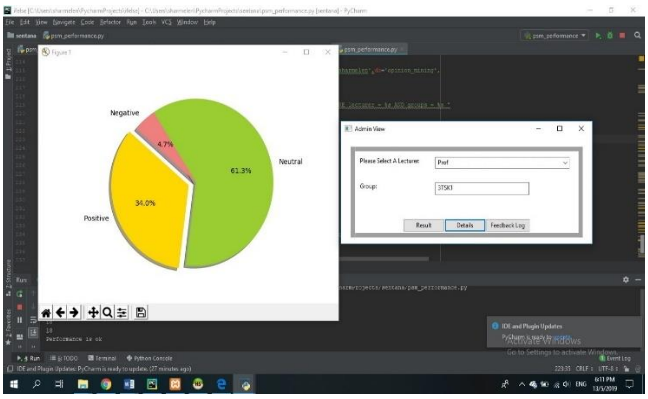


Fig. 6. Results of students' feedback towards a lecturer.

The above results can be summarised in the form of a pie chart that can offer more comprehension about the students' feedback. The pie chart in Fig. 6 shows three coloured sections. The yellow, pink, and green colours represent positive, negative, and neutral opinions, respectively. As shown in the figure, the proportion of students who made neutral opinions exceeds the sum of those who made both positive and negative opinions. In other words, most students have neutral opinions regarding their lecturer's teaching (61.3%). Meanwhile, several students expressed positive opinions (34.9%) and only a small number of students' feedback indicated negative opinions (4.7%).

IV. Conclusion

The present study was designed to explore students' feedback using opinion mining technique by integrating the lexicon based approach in the OMFeedback system. This system is capable of accepting textual opinions from students. By integrating the lexicon based approach, the OMFeefback system can automatically analyse students' opinion about their lecturer's teaching skills at the end of every semester. One of the most surprising findings to emerge from this study was the score given by the simultaneous use of capitalised words and emoji features in a sentence, which revealed a significant value. This value corroborates the ideas presented by Gkontzis *et al*. [15], who suggested that an analytical system with values that are close to 70% is considered as efficient. Thus, the 79.4% obtained from the new features of capitalised words and emoji characters in the students' feedback can be considered as efficient too.

The current study was limited in several ways. First, the OMFeedback system used a Vader Lexicon database, which has a limited number of words. The database does not store words that are related to the educational context, such as "late", "discipline", and "marks". Second, the OMFeedback system was unable to detect spelling errors and assigned a neutral score to misspelled words, which eventually led to inaccurate results. The final limitation was that the system had only focused on students' feedback that was written in English because all the words in the Vader Lexicon database are stored as English. Based on these limitations, future studies should include more words that relate to the educational context in the Vader Lexicon database. To achieve better results, it is essential to integrate a spelling and grammar checker in the OMFeedback system. Finally, future studies are recommended to import a translation programme in the OMFeedback to allow students who are not proficient in the English language the chance to use this system. Overall, educational institutions are encouraged to facilitate this new initiative to increase the effectiveness of their online assessment system.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Muslihah and Suzaimah led the research; Sharmelen developed the OMFeedback system; Nor Asiakin and Norulzahrah collected the data; Muslihah and Noor Afiza analysed the data; Muslihah wrote the paper; all the authors had agreed and approved the final version of the paper.

Acknowledgment

The authors would like to acknowledge and thank the National Defence University of Malaysia for supporting this project.

References

[1] Z. Nasim, Q. Rajput, and S. Haider, "Sentiment analysis of student feedback using machine learning and lexicon based approaches," in *Proc. 2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*, 2017, pp. 1–6.

[2] F. F. Balahadia, Ma. C. G. Fernando, and I. C. Juanatas, "Teacher's performance evaluation tool using opinion mining with sentiment analysis," in *Proc. 2016 IEEE Region 10 Symposium (TENSYMP)*, 2016, pp. 95–98.

[3] S. Khan and R. A. Khan, "Online assessments: Exploring perspectives of university students, " *Educ. Inf. Technol.*, vol. 24, no. 1, pp. 661–677. 2019.

[4] G. I. Nitin, G. Swapna, and V. Shankararaman, "Analyzing educational comments for topics and sentiments: A text analytics approach," in *Proc. 2015 IEEE Frontiers in Education Conference (FIE)*, 2015, pp. 1–9.

[5] H. C. Soong, N. B. A. Jalil, R. Kumar Ayyasamy, and R. Akbar, "The essential of sentiment analysis and opinion mining in social media: Introduction and survey of the recent approaches and techniques," in *Proc. 2019 IEEE 9th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, 2019, pp. 272–277.

[6] K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications," *Knowl.-Based Syst.*, vol. 89, pp. 14–46, 2015.

[7] M. Sivakumar and U. S. Reddy, "Aspect based sentiment analysis of students opinion using machine learning techniques" in *Proc. 2017 International Conference on Inventive Computing and Informatics (ICICI)*, 2017, pp. 726–731.

[8] S. Song, H. Kawamura, J. Uchida, and H. Saito, "Determining tourist satisfaction from travel reviews," *Inf. Technol. Tour.* 2019.

[9] K. Z. Aung and N. N. Myo, "Sentiment analysis of students' comment using lexicon based approach," in *Proc. 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, 2017, pp. 149–154.

[10] V. Dhanalakshmi, D. Bino, and A. M. Saravanan, "Opinion mining from student feedback data using supervised learning algorithms," in *Proc. 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC)*, 2016, pp. 1–5.

[11] C. Pong-Inwong and K. Kaewmak, "Improved sentiment analysis for teaching evaluation using feature selection and voting ensemble learning integration," in *Proc. 2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, 2016, pp. 1222–1225.

[12] C. W. Tseng, J. J. Chou, and Y. C. Tsai, "Text mining analysis of teaching evaluation questionnaires for the selection of outstanding teaching faculty members," *IEEE Access*, vol. 6, pp. 72870–72879, 2018.

[13] C. J. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proc. the Eighth International AAAI Conference on Weblogs and Social Media*, 2014, p. 10.

[14] M. O. Shiha and S. Ayvaz, "The effects of emoji in sentiment analysis," *Int. J. Comput. Electr. Eng.*, vol. 9, no. 1, pp. 360–369, 2017.

[15] A. F. Gkontzis, C. V. Karachristos, C. T. Panagiotakopoulos, E. C. Stavropoulos, and V. S. Verykios, "Sentiment analysis to track emotion and polarity in student fora," in *Proc. the 21st Pan-Hellenic Conference on Informatics*, 2017, pp. 1–6.

Muslihah Wook received her PhD in information science from Universiti Kebangsaan Malaysia in 2017, Master of Computer Science from Universiti Putra Malaysia in 2004, and Bachelor of Information Technology (Hons) from Universiti Utara Malaysia in 2001. Her research interests include data mining applications in various domain particularly in education, security and defence. Currently she is working as a senior lecturer at Department of

Computer Science, Faculty of Defence Science and Technology, National Defence University of Malaysia. She has become members of International Association of Computer Science and Information Technology (IACSIT) and Institute of Research Engineers and Doctors (IRED) since 2011 and 2013 respectively. Recently, she has been appointed as a technical reviewer of Education and Information Technologies —Springer's journal indexed by Scopus (Q1) and other outstanding journals as well.

Sharmelen Vasanthan was born in Teluk Intan, Perak. He received his Bachelor of Computer Science in system security from National Defence University of Malaysia in 2019. His final year project mainly focuses on sentiment analysis and artificial intelligence. He also has a military background of training under the Reserve Officer Training Unit at NDUM and was commissioned as Second Lieutenant in 2019. Currently, he is pursuing his masters of digital security in Eurecom, France.

Suzaimah Ramli was born in Temerloh Pahang. She received her PhD in electrical, electronic and system engineering from Universiti Kebangsaan Malaysia in 2011, Master of Computer Science from Universiti Putra Malaysia in 2001, and Bachelor of Information Technology (hons) from Universiti Utara Malaysia in 1997. Her research interests include image processing and artificial intelligence applications in various domain particularly in education, security and defence. Currently she is working as an associate professor at Department of Computer Science, Faculty of Defence Science and Technology, National Defence University of Malaysia. She is a member of Malaysia Board of Technologist and Informatics Intelligence Special Interest Group, UPNM. She has published and presented most of her research findings to various international conferences and articles in many international journals specifically in her research niche.

Noor Afiza Mat Razali holds a Bachelor's Degree in Computer and Information Engineering, Master of Science in Computer Science and PhD of Science in Computer Science from Japanese Universities. Afiza is a senior lecturer at Faculty of Defence Science and Technology in National Defence University of Malaysia. Afiza also appointed as a visiting lecturer and fellow at Management and Science University Malaysia, Academic Liaison Consultant for Japanese Universities at University Kuala Lumpur and Fellow at Education Malaysia Global Services (EMGS). Afiza research and expertise are in the area of cyber security, disaster management system, big data analytics, human computer interaction, artificial intelligence & robotics and blockchain technology. Afiza is also a professional technologist, a recognition given by Malaysia Board of Technologist.

Nor Asiakin Hasbullah was born in Muar, Johor. She received her PhD in Information Technology and quantitative sciences from MARA University of Technology Malaysia in 2017 and have done three months attachment in Glasgow Caledonian University of Glasgow under Privacy research in 2012. She holds a Master of Science in Information Technology from MARA University of Technology Malaysia in 2006, and bachelor of information technology (hons) from Universiti Kebangsaan in 2003. Currently she is a senior lecturer in the Department of Computer Science, National Defence University of Malaysia (UPNM). Her research interests are in the field of privacy, data protection, information security and ethics in ICT. She is a member of Malaysia Board of Technologist and Informatics Intelligence Special Interest Group, UPNM. She has published and presented most of her research findings and articles in various international conferences and international journal.

Norulzahrah Mohd Zainudin is a lecturer in the Department of Computer Science at Faculty of Defence Science and Technology, National Defence University Malaysia (UPNM). She received her MSc at Universti Putra Malaysia and joined Military Academy of Malaysia in 2002. Her main research interests are in the areas of forensic computing, online social networks and computer intelligence. She has published a number of papers in international journals and conferences. Currently she is a member of Informatics Intelligence Special Interest Group, UPNM.