Evaluating AI Techniques for Blind Students Using Voice-Activated Personal Assistants

Tariq S Almurayziq, Gharbi Khamis Alshammari, Abdullah Alshammari , Mohammad Alsaffar and Saud Aljaloud

Department of Information and Computer Science, College of Computer Science and Engineering,
University of Ha'il, Ha'il, 81481, Saudi Arabia

Summary
The present study was based on developing an AI based model to facilitate the academic registration needs of blind students. The model was developed to enable blind students to submit academic service requests and tasks with ease. The findings from previous studies formed the basis of the study where functionality gaps from the literary research identified by blind students were utilized when the system was devised. Primary simulation data were composed based on several thousand cases. As such, the current study develops a model based on archival insight. Given that the model is theoretical, it was partially applied to help determine how efficient the associated AI tools are and determine how effective they are in real-world settings by incorporating them into the portal that institutions currently use. In this paper, we argue that voice-activated personal assistant (VAPA), text mining, bag of words, and case-based reasoning (CBR) perform better together, compared with other classifiers for analyzing and classifying the text in academic request submission through the VAPA.

Key words:
Voice-activated personal assistant, Text mining, Classification.

1. Introduction

The developmental arena of human-computer interaction cannot be undermined and it has transitioned from conventional computers to touchscreen and the use of speech recognition systems facilitating effective communication and interaction [1,2]. The progression in conversational interaction technologies has led to the development of several intelligent voice assistants including S-Voice, Siri, Google Now Cortana and Echo [1,3]. All of these programmes and interventions help to streamline various features including sending emails or text messages, setting alarms, directional searches or opening information on the web, requesting music to be played, and making purchases while making the best possible use of speech recognition technology [4,5,6]. Voice recognition technology is alternatively referred to as an intelligent virtual assistant, virtual personal assistant, mobile virtual assistant, voice-activated personal assistant or voice-activated digital assistant, all of which refer to same functionalities operating mechanisms and usage [7].

One of the key characteristics which categorises voice-activated personal assistants (VAPAs) as being the most effective interaction tool is their improvised capacity to interact, react and fulfil the requested use services [8]. Besides that, VAPA also facilitates human-like behaviour while enhancing the ability of users to evolve, learn new actions and engage in human-like behaviour with computers [1]. VAPA acts like a practical solution, especially for blind people, enabling them to perform various tasks and, hence, providing them with an alternative method of interaction by using a virtual keyboard [9].

Extensive research has been conducted to analyse the efficacy of different requests being inputted into VAPA and evaluate the issues related to the acceptance and constraints of the requests [1,2,6,10]. However, very limited research has been conducted to examine the efficacy of the VAPA system for blind students, including determining the barriers and issues they might encounter [11]. For example, Pradhan, Mehta, and Findlater (2018) conducted analysed the responsiveness of purchases undertaken by people with disabilities on Amazon Echo VAPAs. The research results declared that 38% of people with disabilities, especially those who are visually impaired, found it effective. Azenkot and Lee (2013) also emphasised the vitality of voice-based interaction as compared to text-based functionalities and acknowledged it to be a reliable input to assistive technology. The researchers further put forth the stance that VAPA presents a unique advantage which provides an interactive facility to blind individuals, including digital touch screen thermostats which would otherwise pose difficulty of access for visually impaired people [12].

1.1 Problem statement

As previously identified, a number of studies have been conducted within the domain of technology assisted personal programmes, speech recognition and the efficacy of using these programmes for the individuals concerned [5,7,10,12,13]. However, the majority of these studies have either focused on the theoretical aspect of the concepts or involved people with impairments or disabilities to explore the problems which hinder effective use of the programmes [15,16]. No previous study has considered how the existing systems can be modified for visually impaired students,

Manuscript received December 5, 2021
Manuscript revised December 20, 2021
https://doi.org/10.22937/IJCSNS.2022.22.1.10

enabling them to fulfil their academic requirements and submissions through speech control or voice based personal assistants. Therefore, this is identified as a key research gap, prompting the researchers to develop a model that students with special needs can utilise for effective academic progression.

1.2 Rationale of the study

Considering the identified problem, the present study set out to develop an AI-based model that meets the needs of blind students while increasing their capacity to submit academic services, tasks and assignments. The research was conducted with the expectation of bridging the existing gap and obstacles faced by blind students followed by urging universities, schools and other academic institutions to incorporate this system into their existing portals. In addition, the study was conducted to ensure that all students have an equal opportunity to access academic material and submit requests.

1.3 Research objectives

The present study was based on the following research objectives:

- To understand the efficacy of AI techniques for blind students.
- To ensure the successful development of a system facilitating the academic needs of blind students in terms of communication and interactions through VAPA.
- To bridge the identified research gap and contribute to the empirical literature relating to the use of technology, VAPA, artificial intelligence and their effectiveness for visually impaired students.

1.4 Research objectives

The present study was based on one primary research question which is stated as follows:

To what extent can artificially intelligent techniques and systems be used to help blind students in terms of submitting academic requests by making effective use of VAPA?

2. Related Work

There have been a number of studies emphasising the importance of VAPA and recommending its use for performing various tasks both at home and while on-the-go [2,7,17]. Abdolrahmani, Kuber and Branham (2018) conducted an empirical investigation of VAPA with a focus on blind people by interviewing 14 blind students where the participants identified the challenges they encountered regarding the responsiveness of system, input management and achieving control over the required information in terms of extracting the required information from what was presented. The research results identified the need to improve the use behaviours facilitating interaction and situational modification to ensure inclusivity, privacy management and the eradication of concerns relating to trust and functional synchronisation [1,18].

In addition, the efficacy of VAPA was also assessed within the context of determining its importance during the Covid-19 pandemic, especially within the learning environment [5]. The researchers argued that the use of advanced learning technology, when combined with the learning management system can prove vital because for various reasons. First, it provides a self-regulated learning environment while ensuring a flexible learning and development system to the students which is further reflected in their academic performance [1,7,10,16]. Second, it makes students well versed in using modern technology and enforces personal learning which increases students' satisfaction. The researchers also found a high correlation between the technological improvisation of learning management systems in terms of VAPA and student satisfaction, the intensity of which increased during the Covid-19 pandemic; however, its usability for the blind students was identified as a challenging aspect [5,19].

Arumugam and Ananthi (2020) added to the literary contributions by developing a new voice assisted system for visually impaired people that entailed making use of a gesture algorithm, taking robotisation and innovation in the field of technology to the next level. The researchers further developed a private assistant inculcating the capacity to recognise the human voice and act accordingly using exteroception devices facilitating the force feedback mechanism, data rendering and call detection. The system helped blind people through an articulation of a sound management methodology supported by instructional vibration and streamlining the process of accepting aggressive actions [6,20].

A similar project system was developed by Agrawal, Agrawal and Padiya (2020) comprising of a voice recognition mechanism for multiple choice questions in exams. The researchers identified the need to develop a dynamic system for the students, arguing that the static orientation of existing systems was inconvenient for users while not offering protection for the eyes. The new system was developed with advanced and sophisticated features with an advanced level of voice recognition which was adherent to all models until 2030 and an inbuilt capacity to ask questions, manage questionnaires and correct movements. The platform further encompassed a number of reliable features which included the recording of answers by speaking into the microphone. The key distinctive features identified in the system included its dynamic nature, the compliance of the application with android phones,

confinement to voice-based mechanisms and the ability of the system to facilitate interaction for blind individuals in terms of improving their educational progression, specifically during exams [10,21,22].

3. Methodology

This study is based on developing an artificially intelligent system catering to the needs of blind students who face challenges when submitting academic requests using any sort of technological intervention. The findings from previous studies formulated the basis of the current study whereby functionality gaps from the literary research identified by blind students were used to develop the system. a primary data was collected from the blind students. Therefore, the study relied on archival research to postulate this model.

Based on the findings of previous studies, the literary gap and research objectives, the suggested system in this research is an integrated artificially intelligent model which supports blind students in terms of processing their academic requests. The system has been developed theoretically in the form of a framework to provide the basis to its development as a prototype and the model has been examined to determine its effectiveness. A detailed description of the model is provided in the following section.

3.1 Description of the suggested framework

The proposed framework consists of four different stages, each of which facilitates a distinctive feature or functionality. The dataset which has been considered for the simulation includes common information: the name of the students, ID provided by the educational institution for recognition, gender (male or female), the age of the student, the name of the college, the type of service requested, as well as mandating the insertion of keywords to specify the nature of the requested services. All of these functionalities hold an important position in the developed model because they recognize the individuality of each student and, hence, assign them a distinctive number for future reference and database records.

There are four different stages in the framework:

- Voice recognition
- Text mining
- Classification
- Fulfilling the request

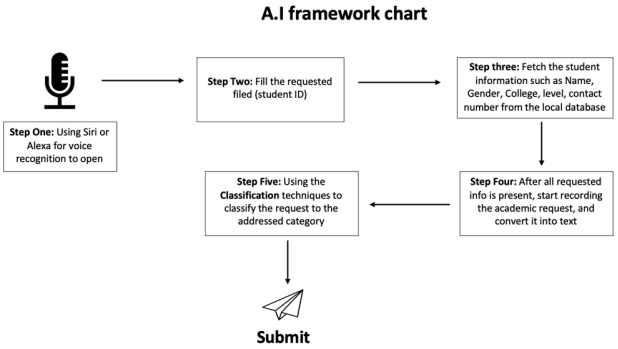

Fig. 1 Artificial intelligence support system.

3.1.1 Voice recognition

The first step in the model is voice recognition. This feature facilitates using Siri, Alexa or any such system to speak words or sentences which are further recorded into the system for processing. For example, blind students can utter a sentence requesting information about the desired services or request certain academic services as a first step to register their voice into the support system app to retrieve student information from the database by providing the student ID, allowing artificial intelligence programs and commands to proceed with the functional aspects. This first step provides the basis for the process of text mining. The currently developed system facilitates three types of academic services which include requesting information about admission to an academic institution, pursuing the registration of students onto any programme or simply obtaining information about the services being offered by the academic institution.

3.1.2 Text mining

Text mining is the second step of the framework which facilitates the processing of artificially intelligent functions or algorithms. This step is connected to the first step of voice recognition. As soon as the student records their voice or submits a query, the system starts processing it by converting it into text. At this step, ensuring the correct conversion of audio to text is a critical and challenging aspect for the programmers to develop because of different accents, use of vocabulary, proficiency of language, tone of voice and pauses breaking or forming the meaning of a sentence.

The main feature which is of great importance in this step of the framework is the extraction of keywords from the text which is generated through a voice-based system. Text mining techniques are used at this step to ensure that the accuracy of the text generated from the recording is meaningful yet brief, concise and understandable. The pre-processing procedure needs to do the following:

– Using text analysis to identify keywords using Word Cloud, the more frequently a word appears in the text, the larger the word will be. This step needs to be worked to eliminate punctuation and uninformative words and retain only meaningful words.
– The text will be converted into individual words to count how many times each word appears using the bag of words method. The effect of pre-processing is visually explored as below.

Fig. 2 Word cloud of keywords.

3.1.3 Classification

Once pre-processing has been completed, it is necessary to assess the framework classification which enables the text of the student's voice to be categorised and to determine how accurately the text complies with the categorisation keywords and information database. The system also assesses the keywords to ensure that they are relatable. Students are able to submit requests to three categories: registration, admission and an alternative academic query. Requests are processed by classification methods in the AI to produce optimal outcomes. For the model that has been proposed, KNN, logistic regression and random forest have each been assessed, revealing that performance is maximised when applying KNN with three neighbours and Euclidean similarity (1). As such, it is recommended that this approach be applied for facilitative purposes.

$$\sqrt{\Sigma_{i=0}^{n}(x_i - y_i)^2} \quad (1)$$

The simplified Euclidean distance equation is employed to establish how similar the investigated and case-base cases are. Equation 1 is applied for each individual case variable with every variable weight being initialised to 1.

3.1.4 Fulfilling the request

This is the last step of the framework which ensures the streamlining of the recorded audio using microphone, text generation through text mining techniques and classification through CBR techniques, ensuring that the correctly requested service is being provided to the user. At this step, any similarity between keywords and the final output is evaluated, completing the process of request generation, and forwarding. At this step, the most critical aspect is ensuring the extraction of the correct information for the blind users and, hence, providing alternative ways to modify the requested services if the output is not what the blind student had intended or requested. However, the inability of the system to ensure a high level of coherence between the requested service and the provided service might also question the efficacy and reliability of the system to a great extent (see Table 1).
Considering the efficacy of system for blind people, it is ascertained that the framework will certainly be useful for the visually impaired, enabling them to have easy access to all services just like other students. Besides that, the system also provides equal access to educational facilities for all students without any discrimination while eradicating the hinderances that might impede their progression to accessing academic services or making requests. Figures 4 and 5 present the framework interfaces.

Table 1: Sample of request and keyword identification

Category	Request	Keywords identified
Admission	When does admission start? How can I join the university? I want to transfer to another university.	Admission, join-university, transfer.
Registration	I would like to change my section. I want to withdraw from course number BA324. I want to request to postpone this term.	Section, withdraw, postpone.
Enquiry	Ask about CS college location? How do I contact the dean's office?	Ask or (any unclassified request).

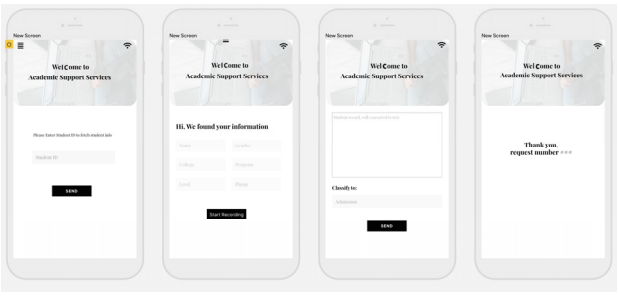

Fig. 3 Sample of request and keyword identification.

Fig. 4 Display of the reply from the institution through the app.

4. Results and evaluation

Among the most important processes involved in the current study is the identification of keywords for the academic request. The importance stems from the fact that these keywords are used to establish the appropriate category. Table 1 presents the keywords for the academic request cases based on examinations, searches and text analysis as well as the application of the Word Cloud.

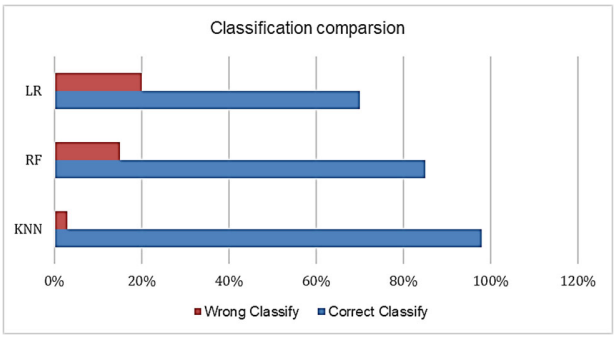

Fig. 5 Comparison of classification methods.

Figure 5 illustrates that the combination of CBR with the results of text mining is beneficial. CBR achieved accuracy of 0.991 owing to the fact that it provides rational explanations in conjunction with the retrieval stage with KNN in the CBR process. As such, combining CBR with the results of text mining provides the greatest accuracy in terms of classification and academic requests in future.

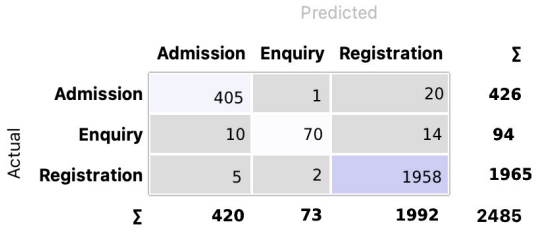

Fig. 6 Comparison of classification methods.

It is apparent from the above figures that the performance of KNN is superior to that of both logistic regression learning and the random forest regarding absolute error and accuracy. The results are based on observations of several thousand requests made in a range of academic categories. New and old cases are matched using case-based reasoning and this enables new requests to be anticipated and classified. The methodology applied in this study involves the creation of a benchmark based on previous examples, thereby enabling a sound appreciation of the different kinds of requests.

Fig. 7 KNN selection error.

It can be seen in Figure 7 that for 3 KNN, the K selection with the smallest possible error was 0.0155, whilst the largest error at 3 KNN was close to 0.02. It is possible to use the size of the dataset to determine which K number to select for CBR. Owing to the fact that there are more than 2,000 cases in the dataset being utilised in the current study, the weather status results are likely to be most reliable when utilising 3 KNN.

Model	AUC	CA	F1	Precision	Recall
kNN	0.991	0.979	0.987	0.973	1.000
Random Forest	0.980	0.970	0.980	0.970	1.000
Logistic Regression	0.970	0.967	0.960	0.957	1.000

Fig. 8 Confusion matrix applying KNN.

Several different means of evaluation have been utilised to compare the results and determine the accuracy of the outcomes. The leave-one-out (LOOCV) method is an approximate function that is trained on all but one data point and the prediction relates to this point. An estimate of the average error is produced and it is upon this that the model is evaluated. It is apparent that the LOOCV does not suggest that the predictions differ significantly in terms of how accurate they are and this may be because of the type of problem domain. For example, because weather forecasts are already reliable, it is unlikely that utilising a different method of evaluation would yield a marked improvement in accuracy. Moreover, as a result of the training and test being split at random, the results can be evaluated. No improvement is realised when splitting the dataset 90/10 or 70/30. No attempt was made to split the data 60/40 because it was deemed that this was similar to the split of the CV. Consequently, the CV is considered to offer the best available means of evaluation when predicting.

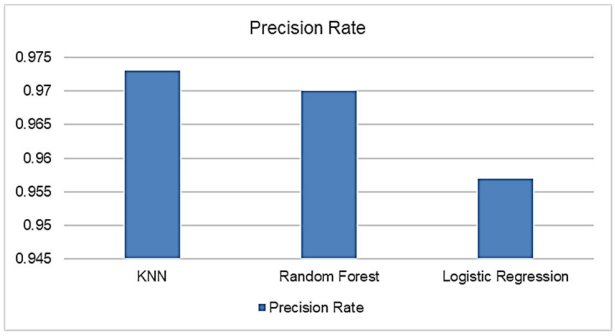

Fig. 9 Precision rate.

KNN offers greater accuracy in terms of precision as well as recall measures relative to logistic regression and random forest. More specifically, whilst the random forest performed well in terms of classification results, most of the cases classified with a high degree of accuracy were in the same category whilst additional cases were included in other categories. For this reason, text mining was no longer considered to be viable.

5. Limitations of the study

Despite its contribution to the literary research, there are several limitations associated with the current study. First, the study was confined to archival analysis in order to identify the gap and postulate the need for the development of a system for blind people. Second, a simulation of primary data was collected from the blind people about the issues and hinderances which they face when using the existing frameworks of educational services. All of these limitations formulate the basis of future prospects of the studies in the domain of technological development, the use of artificial intelligence, robotics and human computer interaction for blind people.

6. Conclusion

The present study was based on developing an AI-based model adhering to the needs of blind students while increasing their capacity to submit academic services and tasks. The research was conducted to bridge the existing gap in research and tackle the associated hinderances faced by blind students when attempting to manage their academic services. The research also sought to emphasise the need for universities, schools and other academic institutions to amalgamate this system into their existing software to ensure the facilitation of visually impaired individuals. Moreover, the research was undertaken in a way that ensured all of the students were equally likely to submit requests and access academic materials. The results confirm that good results in terms of submitting requests and classifications can be achieved by combining CBR, VAPA, bags of words and text mining. The model advocated involves a framework with five steps comprising text mining, the fulfilment of requests, voice recognition, and classification. Importantly, this approach is feasible in practical terms and, therefore, it can be recommended with the use of AI in future research. Theoretical analysis indicates that the model will enable blind students to independently make requests for academic services, albeit that further research is required to establish how well the model works in practice

Acknowledgments

I would like to thank our institute, University of Ha'il, for their unlimited support and recourses to make this work possible.
This research has been funded by research deanship at University of Ha'il – Saudi Arabia through project number (BA-2031).

References

[1] A. Abdolrahmani, R. Kuber, and S. M. Branham, "'Siri Talks at You.' An Empirical Investigation of Voice-Activated Personal Assistant (VAPA) Usage by Individuals Who Are Blind," in *Proceedings of the 20th International ACM SIGACCESS Conference*, 2018, pp. 249-258

[2] M. R. R. M Akanda, M. M. Khandake, T. Saha, J. Haque, A. Majumder et al, "Voice-Controlled Smart Assistant and Real-Time Vehicle Detection for Blind People," in *Advances in Electrical and Computer Technologies: Select Proceedings of ICAECT 2019:*

Lecture Notes in Electrical Engineering, Vol. 672. Springer: Singapore, pp. 287-297, 2019.

[3] S. A. Saparmammedovich, M. A. Al-Absi, Y. J. Koni, and H. J. Lee, "Voice Attacks to AI Voice Assistant," in *International Conference on Intelligent Human Computer Interaction*, 2020, pp. 250-261.

[4] R. Soussiel, E. Loup-Escande, n. Métayer, A. Parant, and V. Laguette, "Risks and benefits of Artificial Intelligence for humans: A literature review," *in Virtual Reality International Conference (VRIC) Proceedings*, 2020, pp. 28.

[5] M. C. Sáiz-Manzanares, R. Marticorena-Sánchez, and J. Ochoa-Orihuel, "Effectiveness of using voice assistants in learning: A study at the time of COVID-19*," International Journal of Environmental Research and Public Health*, vol. 17, no. 15, pp. 5618, 2020.

[6] S. S. L. Arumugam and D. N. Ananthi, "Voice assistants through inaudible voice commands for visually challenged people using gesture algorithm," *European Journal of Molecular and Clinical Medicine*, vol. 7, no. 4, pp. 3024-3030, 2020.

[7] M. Barata, A. G. Salman, I. Faahakhododo, and B. Kanigoro, "Android based voice assistant for blind people," *Library Hi Tech News*, vol 35, no. 6, pp. 9-11, 2018.

[8] S. S. Rautaray and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey," *Artificial Intelligence Review*, vol. 43, no. 1, pp. 1-54, 2015.

[9] A. Abdolrahmani, K. M. Storer, A. R. M. Roy, R. Kuber, and S. M. Branham, "Blind leading the sighted: drawing design insights from blind users towards more productivity-oriented voice interfaces," *ACM Transactions on Accessible Computing (TACCESS)*, vol.12, no.4, pp.1-35, 2020.

[10] S. Agrawal, M. Agrawal, and S. Padiya, "Android Application with Platform Based On Voice Recognition For Competitive Exam," *International Journal of Advanced Research is Science and Technology (IJARST),* vol. 5, no. 5, pp. 27-34, 2020.

[11] S. M. Kelly and G. Kapperman, "A second look at what high school students who are blind should know about technology," 2018.

[12] A. Pradhan, K. Mehta, and L. Findlater, "'Accessibility Came by Accident' Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities," in *Proceedings of the 2018 CHI Conference on human factors in computing systems,* 2018, pp. 1-13.

[13] S. Azenkot and N. B. Lee, "Exploring the use of speech input by blind people on mobile devices," in *Proceedings of the 15th international ACM SIGACCESS conference on computers and accessibility*, 2013, pp. 1-8.

[14] C.A. Pereira, "Digital for Life? Blind Spots of AI and its Reframing for Desirable Futures," in *IEEE 18th International Conference on Cognitive Informatics and Cognitive Computing*, 2019, pp. 323-328.

[15] J. I. Bartolome, L.C. Quero, S. Kim, M.-Y. Um, and J. Cho, "Exploring art with a voice controlled multimodal guide for blind people," in *Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction*, 2019, pp. 383-390.

[16] V. A. Kumar, P. A. Kailas, and N. John, "The Third Eye–An Assistive Technology for the Blind,"

[17] A. Sangwai, S. Deshmukh, V. Sathe, R. Agarwal, and R. Kalantri, "Eye (I) Still Know! An App for the Blind Built using Web and AI," in *An App for the Blind Built using Web and AI*, 2021.

[18] A. V. Yadav, S. S. Verma, and D. D. Singh, "Virtual assistant for blind people," *International Journal of Advanced Scientific Research and Engineering Trends*, vol. 6, no. 5, pp. 156-159, 2021.

[19] R. Z. Ul, S. Abbas, M. A. Khan, G. G. Mustafa, H. Fayyaz et al, "Understanding the Language of ISIS: An Empirical Approach to Detect Radical Content on Twitter Using Machine Learning," *Computers, Materials, and Continua*, vol. 66, no. 2, pp. 1075 - 1090, 2021.

[20] S. Real and A. Araujo, "Navigation systems for the blind and visually impaired: Past work, challenges, and open problems," *Sensors*, vol. 19, no. 15, pp. 3404, 2019.

[21] S. Rizvi, I. Sohail, M. M. Saleem, A. Irtaza, M. Zafar et al, "A smart home appliances power management system for handicapped, elder and blind people," in *4th International Conference on Computer and Information Sciences (ICCOINS)*, 2018, pp. 1-4.

[22] N. K. Sirohi, M. Bansal, and S. N. Rajan, "Recent Approaches for Text Summarization Using Machine Learning & LSTM0," *Journal on Big Data*, vol. 3, no. 1, pp. 35, 2021.

Tariq S Almurayziq master's and doctoral degrees were from the University of Brighton in the UK. at various times in his career, he has taught, full-time courses, and is currently an Assistant Professor in the Department of Computer Science and Engineering at Hail University. Interesting in developing Machine Learning solutions related to Natural Languages, and working as Data Scientist in Processing and Classifications the big data, by using

various feature reduction techniques for supervised and unsupervised expert and decision support systems. In his research area, he has published various articles in A.I to push the boundary of Artificial intelligence systems.

Gharbi Khamis Alshammari is an Assistant Professor in Computer science and Engineering College at University of Hail in Saudi Arabia. He is currently the Vice Dean of Scientific Research for Scientific Publishing. I received my BSc in Computer Science from University of Hail in 2008, the MSc degree in Computing from De Montfort University in the UK in 2011, and the PhD degree in Computer Science from Brighton University in the UK in 2019. His main research focuses on Machine Learning, Recommender Systems, cyber Security and Deep Learning.

Abdullah Alshammari is an Assistant Professor in Computer Science and Engineering College at University of Hail in Saudi Arabia. He is currently the head of Statistics and Information department in the University of Hail. In 2010, Alshammari received his BSc in Computer Science from the University of Hail. In 2014 and 2019, the University of Brighton in the UK granted him the MSc and PhD degrees consequently. Dr. Alshammari research focuses on Artificial Intelligence, Machine Learning, Recommender Systems and social media analysis by AI techniques.

Mohammad Alsaffar is an Assistant Professor in Computer science and Engineering College at University of Hail in Saudi Arabia. He is currently the vice dean of community service and continuing education deanship. Mohammad received his BSc in Computer Science from University of Hail in 2008, the MSc degree in Computing from De Montfort University in the UK in 2011, and the PhD degree in Computer Science from Brighton University in the UK in 2019. Mohammad's research focuses on User Experience, Human Computer Interaction, Usability, Interface Design and Cyber Security.

Saud Aljaloud is an Assistant Professor in Computer science and Engineering College at University of Hail in Saudi Arabia. Dr. Aljaloud received his BSc in Computer Science from University of Ha'il in 2007, then MSc degree in Software Engineering from Southampton University in the UK in 2011, and the PhD degree in Computer Science from Southampton University in 2019. His main research focus spans around Machine Learning and AI, Cloud Computing, Databases Optimizations and Blockchain technology.