Intelligent Automated Cognitive-Maturity Recognition System for Confidence Based E-learning

Imran Usman† and Adeeb M. Alhomoud††

†College of Computing and Informatics, Saudi Electronic University, Riyadh, Saudi Arabia
††College of Science and Theoretical Studies, Saudi Electronic University, Riyadh, Saudi Arabia

Summary
As a consequence of sudden outbreak of COVID-19 pandemic worldwide, educational institutes around the globe are forced to switch from traditional learning systems to e-learning systems. This has led to a variety of technology-driven pedagogies in e-teaching as well as e-learning. In order to take the best advantage, an appropriate understanding of the cognitive capability is of prime importance. This paper presents an intelligent cognitive maturity recognition system for confidence-based e-learning. We gather the data from actual test environment by involving a number of students and academicians to act as experts. Then a Genetic Programming based simulation and modeling is applied to generate a generalized classifier in the form of a mathematical expression. The simulation is derived towards an optimal space by carefully designed fitness function and assigning a range to each of the class labels. Experimental results validate that the proposed method yields comparative and superior results which makes it feasible to be used in real world scenarios.

Key words:
E-learning, cognitive maturity, Classification, Genetic Programming, confidence-based e-learning

1. Introduction

Electronic learning (e-learning) is an educational strategy which is based on the implication and upgradation of the Information and Communication Technology (ICT). No one can deny the influence of ICT on educational pedagogies in the current technological scenario [1], [2] especially after the recent outbreak of Covid-19 pandemic worldwide. The conventional teacher-centric educational scenario has shifted to more learner-centric one [3] and learning is becoming wide and adjustable in terms of space and time [4].

E-learning is an important and effective learner-centric educational means that can help the self-learner achieve all educational goals if it is based on adoption of confidence-based e-learning (CBL) [5]. CBeL aims to attain the desired cognitive masterly level that is enriched with great confidence and knowledge, which is pivotal to accomplish lifelong learning. CBeL is an innovative methodology that empower e-learner not only with proficiency and knowledge, but also with the assurance and tenacity in that knowledge. This confidence and assurance is a necessary precondition of a potential practical life.

The classical CBeL theory emphasizes the four possible classes pertaining to knowledge and confidence of the learner. These are as follows:

(i) Mastery: this is the highest level of both knowledge and confidence. A learner in this level is capable of applying correct knowledge with confidence.
(ii) Doubt: a state in which the e-learner possesses true knowledge, but with the low level of confidence which lets him incapable of taking bold and the right decision in practical walks of life.
(iii) Misinformation: a state in which a learner possesses wrong knowledge but with the highest level of confidence. This can be very devastating in some practical life applications. E.g., in medical, some engineering works, and other life threatening situations.
(iv) Ignorance: a state in which the learner is neither having the knowledge, nor is he confident at all about what he is learning.

These classes truly depict the outcome of an e-learning scenario. Hunt's [6] research finds a direct link between the retention of newly learnt material with people's confidence on the correctness of the learned material. While Abedi's work [7] focusses on understanding the connection between knowledge, confidence, retention and the quality of knowledge.

Bruno's research [8] concludes that knowledge without confidence is not adequate to build behavior. Chernova et al. [9] proposed a form of learning based on confidence-based autonomy (CBA), which focuses on confident execution and corrective demonstration of the

Manuscript received April 5, 2021
Manuscript revised April 20, 2021
https://doi.org/10.22937/IJCSNS.2021.21.4.27

learnt knowledge. Gozava et al. [1] have done a research on the confidence exchange using the gradient-based optimization approach.

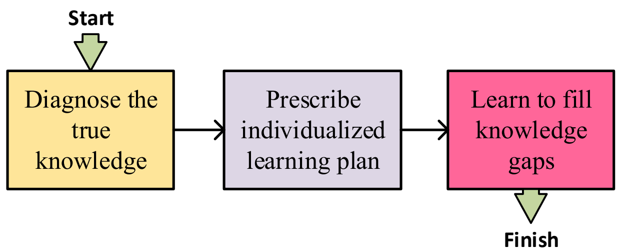

Fig. 1 Different stages of CBeL based learning paradigm

Buthipitiya et al. [10] proposed Confidence based learning ensembles (CobLE) based on assembling of classifiers by measuring their confidence function. Park et al. [11] represented a confidence based matching cost modulation scheme to improve the efficiency of different stereo matching algorithms. Erdt et al [12] highlighted the strengths and weaknesses in the technology enhanced learning Evaluation (TELE). Nevertheless, none of the above-mentioned works cover the cognitive state of the e-learner, integrating a CBeL system. In this work, we cover the learner's cognitive state on the basis of the learner's performance in a test following CBeL paradigm, where both confidence and knowledge are taken into account. This will lead towards a more effective confidence -based e-learning system.

In this work, we extend the work presented by Bhattacharya et al. [5] by incorporating an intelligent search technique that is capable to exploit the concealed reliance among different parameters pertaining to the cognitive maturity recognition, which are otherwise disregarded in conventional works in literature. We use Genetic Programming (GP) to evolve a mathematical formula that is generalized in nature and can be applied to the real-world applications. The classification accuracy and other metrics conform the suitability of the proposed technique in the results section.

The rest of this paper is organized as follows. In section 2, we present overview of the CBeL environment and describe in detail the design of the intelligent cognitive maturity recognition system for e-learner. Experimental results and analysis are presented in section 3. Finally, conclusion and future directions are presented in section 4.

2. Confidence-based E-learning Recognizer

Confidence-based e-learning environment embodies the CBeL paradigm as the most important didactic tool that facilitates the self-learner in achieving his/her learning goals. Figure 1 is presenting the overall strategy of the CBeL system representing the different stages of the system. When a sufficient amount of content has been delivered to the learner, he has to take a confidence-based test, where his performance is analyzed to find out his cognitive status. This cognitive status of the learner determines if he falls in the category of ignorance, misinformation, doubt or mastery. If the learner has yet to achieve the mastery, he is given with an appropriate guideline, such as to improve his cognitive status. This cycle continues and repeats until the learner achieves mastery.

In figure 2, the cognitive maturity matrix for an e-learner is presented depicting the class of a learner based on his knowledge and confidence. As can be seen in the figure, lower knowledge and lower confidence leads to ignorance. On the contrary, higher knowledge and higher confidence leads to the mastery level. The in-between stages are of primary concern and most dangerous in real world scenarios. A lack of knowledge and a higher level of confidence could lead to misinformation. A higher level of knowledge with a lack of confidence can lead to a doubt. Both of these former states of a learner can be very disastrous in professional environments. Hence, recognizing them and adapting the right pedagogies to overcome the shortcomings in the learning systems with the help of technology is of prime importance.

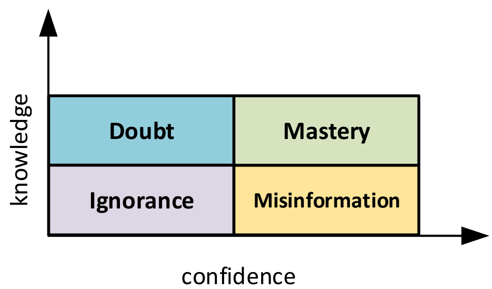

Fig. 2 Cognitive maturity matrix

2.1 CBeL test preparation

Before we are able to develop a fully functional cognitive maturity recognition system, it is essential to generate the data for the model to be employed on. To the best of our knowledge, there is no such publicly available data. Hence, we have created the data by performing the actual tests comprising a number of university level students.

We used the actual test environment to generate the data for the proposed work. A similar approach is also adopted in [5]. The test prepared was from the programming course with 50 multiple choice questions having four possible set of answers for each of the question. Along with the possible choices of answers, we also provided a scale of the confidence level of the learner with a range of values from 0-4. Where 1 being the least confident or ignorant, and 4 being the highest level of confidence. A value of 0 is reserved when the level of confidence is neutral or inconclusive by the learner.

For the purpose of labelling the actual class labels, a panel of experts comprising of academicians and psychologist is used, who actually labelled the confidence of each of the learner for every test question attempted. The experts are also given the same performa for filling the actual range of values pertaining to the confidence level against every question for a learner.

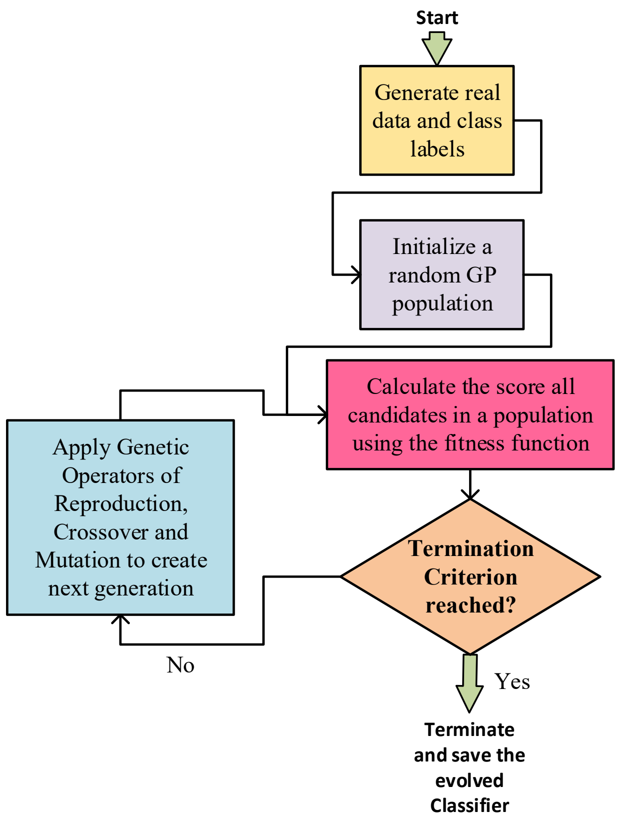

Fig. 3 General architecture of the proposed technique

2.2 Cognitive maturity recognizer modeling

Once we have the data for the actual class labels and the derived classes through the actual tests for the subjects, our goal is to construct a classifier that is capable of mapping the true confidence level of a learner. For this reason, we have utilized a Genetic Programming (GP) based modeling technique where the purpose is to construct a mathematical expression that is generic in nature and is able to predict the true class of a learner for real world applications. GP has been successfully applied in many applications in literature [13]–[16]. The general architecture of the proposed technique is displayed in figure 3. It constitutes of two phases, namely, the training phase and the testing phase. Details of these phases are described in the paragraphs that follows.

In the training phase, we use the actual class labels and the derived scores to evolve a classifier using the GP simulation. In order to evolve a suitable mathematical formula (classifier) through the GP simulation, we need to model the problem and define a set of suitable parameters. In addition, we also need to define the suitable GP Function set and the GP Terminal set. The terminal set can be further divided into constant terminal set and variable terminal set. In this work, we use the functions of mathematical Log, SIN, TAN, COSINE, exponent, addition, subtraction, multiplication, and protected division to form the GP Function Set. For the Variable Terminal Set, we use the variables α , β , γ , λ and φ , κ and τ with the values as indicated by the following equations:

$$\alpha = [-1 \times \kappa] \tag{1}$$

$$\beta = [-1 \times \tau] \tag{2}$$

$$\gamma = [1 \times \kappa] \tag{3}$$

$$\lambda = [1 \times \tau] \tag{4}$$

Where, $\kappa \in \{1, 2\}$ represents the element belonging to the set of values for the low confidence, and $\tau \in \{3, 4\}$ represent the element from the high confidence set. In equations (1) and (2), the value of -1 corresponds to an incorrect answer. Whereas, in equations (3) and (4) the value of 1 corresponds to a correct answer. For the Constant Terminal set, we use random terminals in the range of [-1,1].

Initially, the GP simulation starts with randomly initialized parameter values. A randomly initialized population of the candidate solution is created each of which forms a functional dependency with respect to the elements of the GP function set and the terminal sets as represented in the following equation.

$$\Im = (\alpha, \beta, \gamma, \lambda, \kappa, \tau, [-1,1], +, -, \times, \div, \log, \sin, \cos, \tan, \exp) \tag{5}$$

The realization of the functional dependency presented in equation (5) is in the form of a mathematical expression which is evaluated using the fitness function defined in equation (6).

$$fitness = \sum_{i=1}^{50} \theta_i \quad (6)$$

where, the value of i represent the number of questions and θ is defined as the difference between the actual class label and the evaluation of $\Im$ using the following relation:

$$\theta = \begin{cases} 1 & if\ eval(\Im) \equiv actual\_class\_range \\ 0 & otherwise \end{cases} \quad (7)$$

where, the actual class range is defined as:

$$actual\_class\_range = \begin{cases} [0, 0.2] \Rightarrow ignorant \\ [0.2, 0.4] \Rightarrow misinformation \\ [0.4, 0.6] \Rightarrow doubt \\ [0.6, 0.8] \Rightarrow mastery \\ [0.8, 1] \Rightarrow undetermined \end{cases}. \quad (8)$$

Once, all the candidate solutions in a population are scored using equation (6), we move on to the next generation by creating the candidates for the next generation through GP operators of *crossover*, *mutation*, and *reproduction*. The process continues generation by generation until we reach a stopping criterion. This could be in the form of the number of generations, or the time limit, or the desired optimal fitness value. In our experimentation we used the optimal GP value as the terminal criterion of our simulation.

Through the GP simulation, over the generations, a mathematical expression is evolved that is capable of finding the right range of values for a particular cognitive maturity of a learner. Once, generated, such an expression is saved and is then utilized in the testing module of the proposed technique, whereby a new data is used to test the accuracy of the classifier. Once adequate results are achieved in the testing phase, the best evolved expression is then used for the real-world applications. In the section that follows we present out experimental results and their analysis.

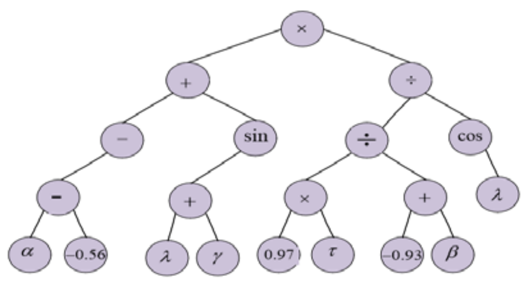

Fig. 4 An example candidate GP tree

3 Experimental Results and Analysis

The proposed technique is implemented on an AMD 7nm Ryzen 4700 series processor with 8 Gb RAM and 512 GB SSD and Radeon graphics. For the software part of the project, we have used Matlab programming suite [17]. For the implementation of the GP module we used GPLAB toolbox [18].

In order to attain the best evolved expression, several GP simulations were run and amongst them, the best result was saved. This is due to the fact that the initial population of the GP simulation is initialized randomly and repeating the simulation allows the search mechanism of GP to fully explore the solution space such as to avoid trapping in local maxima. Figure 4 demonstrates a sample GP tree which represents a candidate solution. We can see that all the Function set is at the internal node of the tree and, as the name suggests, the Terminal set constitutes the leaf nodes of the tree. It is to be noted that once the tree is properly traversed, it is denoted in the form of an equations that represents our classifier.

Table 1: GP parameter settings

Parameter	Value/s
Function Set	+, -, *, /, EXP, SIN, COS, LOG, TAN
Terminal Set	Constant terminals: [-1, 1] Variable Terminals: α, β, γ, λ and φ, κ and τ
Fitness_Function	$fitness = \sum_{i=1}^{50} \theta_i$
Generation size	50
Population size	100

Table 1 shows some of the parameter settings as done in the GPLAB toolbox in order to run the simulation. The rest of the settings in the toolbox were kept as default values. In order to measure the performance of the best evolved classifier, we use the accuracy metric defined as:

$$Ac = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

where, P is the total number of positives, N is the total number of negatives, TP, TN, FP and FN are the true positive rate, true negative rate, false positive rate and the false negative rates respectively. Further explanation of these terms can be found in [19].

Table 2 presents the performance comparison of the proposed technique with Bhattacharya et al.'s neural network based technique [5]. It can be observed that the overall accuracy of the proposed technique is a little better than the previous technique. A major contribution to this improvement is due to the fact that we have defined the range of class labels and over the generations, the GP simulation is able to associate to the appropriate class rage based on extensively finding the hidden dependencies between different attributes.

Table 2: Performance comparison of the proposed technique with [5] in terms of accuracy

Technique	Accuracy (%)
Bhattacharya et al. [5]	95.4
Proposed	96.1

In order to analyze the performance of the best evolved classifier all of the classes, we demonstrate their comparison in table 3. As can be seen from the table the variance between the accuracy measures pertaining to different classes is less and acceptable. A slight variance is always expected when it comes to a heuristic search technique. Compared to [5], the indeterminant cases in our data are also far less.

Table 3: Performance in terms of accuracy of different classes

Class	Accuracy (%)
Ignorance	97.3
Misinformation	96.1
Doubt	95.9
Mastery	96.4
Undetermined	94.8

Figure 5 demonstrates the results of the proposed technique with the previous technique in terms of the receiver operating characteristic (ROC) curve which is defined as:

$$ROC = \frac{TPR}{FPR} \tag{10}$$

where TPR denotes the true positive rate and FPR represents the false positive rate given as follows:

$$TPR = \frac{TP}{TP + FN} \tag{11}$$

$$FPR = \frac{FP}{FP + TN} \tag{12}$$

It is to be noted that squarer the curve is the better is the performance of a technique. In figure 5 it can be seen that the proposed technique is much better in terms of the ROC curve. We have used only Bhattacharya et al.'s [5] work for the comparison purposes because , to the best of our knowledge, their work is the novel work in literature on the recognition of cognitive state of an e-learner and no other such work exists in literature.

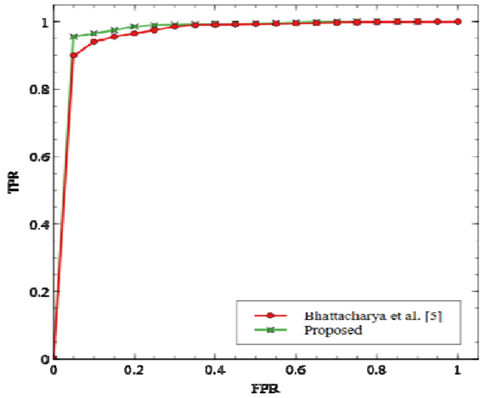

Fig. 5 Performance comparison in terms of ROC.

4. Conclusion

In this work we present an intelligent cognitive maturity recognition system for an e-learner which is very essential in confidence-based e-learning. Once the classification of the cognitive state is achieved, learning pedagogies can be tailored along with the availability of suitable ICT technologies to increase the learning capacity in the e-learning environment. Appropriate data is collected using a real examination setup with students performing the entire test. Experienced academicians are then utilized to mark the actual class labels. Once the data is collected, GP based simulation technique is applied to evolve the best classifier that is capable of exploiting the concealed dependencies between different problem attributes. An enhanced fitness function is devised with a possible range of values assigned to the classes as class labels. Experimental results confirm the suitability of the proposed approach and validate the usefulness in terms of real-world applications.

Acknowledgment

The authors highly extend their appreciation and acknowledgement to the Deanship of Scientific Research (DSR) at Saudi Electronic University, Riyadh, Saudi Arabia, for funding this work through the grant for the project titled "Intelligent automated Cognitive-maturity recognition system for confidence based e-learning".

References

[1] N. Grozavu, M. Ghassany, and Y. Bennani, "Learning confidence exchange in Collaborative Clustering," in *The 2011 International Joint Conference on Neural Networks*, Jul. 2011, pp. 872–879, doi: 10.1109/IJCNN.2011.6033313.

[2] "Curricula and the use of ICT in education: Two worlds apart? - Tondeur - 2007 - British Journal of Educational Technology - Wiley Online Library." https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.111

1/j.1467-8535.2006.00680.x (accessed Apr. 25, 2021).
[3] P. Brusilovsky and C. Peylo, "Adaptive and Intelligent Web-based Educational Systems," p. 14.
[4] A. Koohang and A. Durante, "Learners' Perceptions toward the Web-based Distance Learning Activities/Assignments Portion of an Undergraduate Hybrid Instructional Model," *Distance Learn.*, p. 9.
[5] S. Bhattacharya, S. Roy, and S. Chowdhury, "A neural network-based intelligent cognitive state recognizer for confidence-based e-learning system," *Neural Comput. Appl.*, vol. 29, no. 1, Art. no. 1, Jan. 2018, doi: 10.1007/s00521-016-2430-5.
[6] D. P. Hunt, "The concept of knowledge and how to measure it," *J. Intellect. Cap.*, vol. 4, no. 1, pp. 100–113, Jan. 2003, doi: 10.1108/14691930310455414.
[7] J. Abedi and J. Bruno, "Test-Retest Reliability of Computer Based MCW-APM Test Scoring Methods," *J. Comput.-Based Instr.*, vol. 16, no. 1, pp. 29–35, 1989.
[8] J. E. Bruno, "Using Testing to Provide Feedback to Support Instruction: A Reexamination of the Role of Assessment in Educational Organizations," in *Item Banking: Interactive Testing and Self-Assessment*, Berlin, Heidelberg, 1993, pp. 190–209, doi: 10.1007/978-3-642-58033-8_16.
[9] S. Chernova and M. Veloso, "Interactive Policy Learning through Confidence-Based Autonomy," *J. Artif. Intell. Res.*, vol. 34, pp. 1–25, Jan. 2009, doi: 10.1613/jair.2584.
[10] S. Buthpitiya, A. K. Dey, and M. Griss, "CobLE: Confidence-Based Learning Ensembles," in *2014 International Conference on Computational Science and Computational Intelligence*, Mar. 2014, vol. 1, pp. 386–391, doi: 10.1109/CSCI.2014.72.
[11] Min-Gyu Park and K.-J. Yoon, "Leveraging stereo matching with learning-based confidence measures," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 101–109, doi: 10.1109/CVPR.2015.7298605.
[12] "Evaluating Recommender Systems for Technology Enhanced Learning: A Quantitative Survey | IEEE Journals & Magazine | IEEE Xplore." https://ieeexplore.ieee.org/document/7115158 (accessed Apr. 25, 2021).
[13] I. Usman and K. A. Almejalli, "Intelligent Automated Detection of Microaneurysms in Fundus Images Using Feature-Set Tuning," *IEEE Access*, vol. 8, pp. 65187–65196, 2020, doi: 10.1109/ACCESS.2020.2985543.
[14] T. Naheed, I. Usman, T. M. Khan, A. H. Dar, and M. F. Shafique, "Intelligent reversible watermarking technique in medical images using GA and PSO," *Opt.-Int. J. Light Electron Opt.*, vol. 125, no. 11, Art. no. 11, 2014.
[15] T. Naheed, I. Usman, and A. Dar, "Lossless data hiding using optimized interpolation error expansion," in *Frontiers of Information Technology (FIT), 2011*, 2011, pp. 281–286.
[16] T. Bashir, I. Usman, A. Albesher, K. A. Almejalli, and S. S. Naqvi, "GP based smart reversible watermarking of depth image based rendering for stereoscopic images," *Multimed. Tools Appl.*, vol. 78, no. 15, Art. no. 15, Aug. 2019, doi: 10.1007/s11042-019-7399-5.
[17] "Get MATLAB." https://www.mathworks.com/products/get-matlab.html (accessed Apr. 25, 2021).
[18] "GPLAB - A Genetic Programming Toolbox for MATLAB." http://gplab.sourceforge.net/ (accessed Apr. 25, 2021).
[19] I. Usman, "Anomalous Crowd Behavior Detection in Time Varying Motion Sequences," in *2019 4th World Conference on Complex Systems (WCCS)*, Apr. 2019, pp. 1–5, doi: 10.1109/ICoCS.2019.8930795.

Imran Usman received his BE degree in Software Engineering from Foundation University, Pakistan in 2003 and MS Computer System Engineering from Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan in 2006. He joined Pakistan Institute of Engineering and Applied Sciences as a research scholar and received his PhD degree in 2010. From 2009 to 2010 he served at Iqra University Islamabad, Pakistan as an Assistant Professor in Department of Computing and Technology. From 2010 to 2012 he served as Assistant Professor and Senior In-charge Graduate Program in the Department of Electrical Engineering at COMSATS Institute of Information Technology Islamabad, Pakistan. He is presently serving as Associate Professor in College of Computing and Informatics, Saudi Electronic University, Kingdom of Saudi Arabia. His present research interests include machine learning, digital image processing, evolutionary computation and digital watermarking. Dr. Usman has a number of research papers to his credit and has supervised many BS, MS and PhD students. He is also a Senior Member of IEEE.

Adeeb Alhomoud holds a PhD in cyber security from the University of Bradford (2014) and a Master's degree from University of Essex (2005) United Kingdom. Alhomoud is technically-astute Cyber Security specialist with 15+ years of diversified experience across academic and IT sector. Significant exposure in Cyber security, Network security, Malware Attacks, Botnets, Self-healing and Offensive Security. Certified Ethical Hacker, ISO 27001 LI, with extensive knowledge of security threats, risk analysis and the development of security systems and protocols in business and corporate settings. His research interests include botnets, malware propagation, IoT, steganography Machine Learning and Artificial intelligent.