Novelty Journals

ISSN 2394-7314

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 9, Issue 2, pp: (44-51), Month: May - August 2022, Available at: www.noveltyjournals.com

Prediction of Psychological Disorders among Palestinian Children Using Data Mining Techniques

Naji Shukri Alzaza

Department of Software Engineering
University of Palestine, Gaza, Palestine

DOI: https://doi.org/10.5281/zenodo.6967463

Published Date: 05-August-2022

Abstract: **The psychological disorders have increased especially during the COVID-19 pandemic. Such mental disorders increased with the continuity of the military escalation, wars, and violence in Palestine. Such an unhealthy environment caused a third of the Palestinian population to need psychological health services. Furthermore, children have more suffering from such psychological disorders and need much effort to make recover their normal mental health. This study aims to investigate the prediction of psychological disorders among Palestinian children and identify the proper machine learning classifier using data Mining techniques. The dataset was collected from handwriting clinical psychologist papers under the supervision of a psychologist. 133 instances were distributed on 54 mental attributes including the class attribute. The study focuses on eight simple and complex psychological disorders. However, eight classification models of machine learning were utilized. Results show that all models (AdaBoost, Neural Network, Logistic Regression, Random Forest, SVM, Tree, KNN, and Naive Bayes) have an excellent AUC (more than 0.9). Furthermore, five models that are AdaBoost, Neural Network, Logistic Regression, Random Forest, and SVM have an acceptable accuracy between 94% to 98.5%.**

Keywords: **Psychological Disorder Classification, Data Mining, Mental Illness, Mental Disorder Prediction.**

I. INTRODUCTION

In the last decade, statistics indicate that psychological disorders have become more prevalent in recent years as a result of wars, natural disasters, or pandemics. According to World Health Organization (WHO) in 2019, there was one in every eight people in the world living with a mental disorder. Furthermore, during the COVID-19 pandemic, such numbers of psychological disorders have increased significantly in the year 2020 by 26% for anxiety and 28% for major depressive disorders (WHO, 2022). In Palestine, around a third of the population is in need of psychological health services (Brady, 2020; Marie, SaadAdeen, & Battat, 2020). Furthermore, Palestinian have the highest rate of mental health disorders in the Middle East (Brady, 2020).

A psychological disorder is a synonym of a mental disorder and is sometimes called a mental illness (Cherry, 2022; American Psychiatric Association, 2013; Sumathi & Poorna, 2016). However, it involves significant disturbances in an individual's thinking, cognition, emotional regulation, or behavior.

WHO reported that while effective prevention and treatment options are available; unfortunately, the most people do not have an effective care (WHO, 2022). Diagnosing a psychological disorder is a process, and it could include the evaluation of mental health disorders as well as physical conditions that could be related to the symptoms being experienced. Statistics show that people have more than one psychological disorder at the same time (Espié, et al., 2009), so the diagnosis should address all the individual symptoms.

One of the fundamental evaluation steps is the inspection and inquiring about the medical history and the symptoms. However, diagnosing children with psychological disorders is more problematic than diagnosing adults. Hence, their parents should be involved in the diagnosis sessions to answer some questions. (Sumathi & Poorna, 2016).

Several studies reported that children are suffering from wars and violence (Baird, Panlilio, Seager, Smith, & Wydick, 2022; Espié, et al., 2009; Brady, 2020; Marie, SaadAdeen, & Battat, 2020). Furthermore, such effects on the mental health of children are often expressed through severe and chronic reactive psychological syndromes, including post-traumatic stress disorder, anxiety, and other behavioral problems (Espié, et al., 2009). However, after military escalation especially in Gaza Strip, Palestinian children suffering from such psychological disorders and need much international efforts to make recover their normal mental health (Espié, et al., 2009; Baird, Panlilio, Seager, Smith, & Wydick, 2022; WHO, 2022). Medical Aid for Palestinians (MAP) estimates that after 2009, 30% of adults reported symptoms of post-traumatic stress disorder. This percentage increased to 54% after 2014. The children's reported symptoms included flashbacks, nightmares, distressing thoughts, trouble sleeping, and avoidant behavior (Medical Aid for Palestinians, 2020). However, MAP declared that 75% of children and adults have greater self-confidence after treatments via life skills sessions.

Data Mining has been widely used in disease prediction to increase the accuracy of the rate of detection and reduce the tests that could be conducted (Sumathi & Poorna, 2016; Baird, Panlilio, Seager, Smith, & Wydick, 2022). Furthermore, applying data Mining techniques can be of great help to the clinical decision, diagnosis prediction, and improve the patient's quality of life (Alonso, et al., 2018).

Six psychological disorders were identified the most cases in Palestine among children (Espié, et al., 2009; Marie, SaadAdeen, & Battat, 2020) which are Post-Traumatic Stress Disorder (PTSD), Specific Learning Disability (SLD), Intellectual Disability (ID), Autism Spectrum Disorder (ASD), Epilepsy, and Attention Deficit Hyperactivity Disorder (ADHD). Each disorder is defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2013) as follow:

PTSD refers to an anxiety disorder that develops in relation to an event which creates psychological trauma in response to actual or threatened death, serious injury, or sexual violation.

SLD refers to a type of Neurodevelopmental Disorder that impedes the ability to learn or use specific academic skills such as reading, writing, or arithmetic.

ID refers to a disorder that starts during the developmental period. It consists of certain intellectual deficits and challenges handling aspects of daily life like school, work, home, social life, health, among other things.

ASD refers to a range of neurodevelopmental disorders that includes autism and related conditions.

Epilepsy refers to a central nervous system disorder in which brain activity becomes abnormal, causing seizures or periods of unusual behavior, sensations and sometimes loss of awareness.

ADHD is a disorder with its onset in early childhood, and is characterized by symptoms of hyperactivity, inattention, and impulsivity that interfere with daily and occupational functioning.

II. DATA MINING PSYCHOLOGICAL DISORDERS' PREDICTION

Data mining is one of the Artificial Intelligent (AI) services constructed as a process that incorporates two elements which are database and machine learning. However, data mining needs machine learning techniques and algorithms, while machine learning doesn't necessarily need data mining.

In earlier years, in 1980s the expert systems technique was utilized for psychological disorders' diagnoses (Sumathi & Poorna, 2016). The expert systems were used to manage the interview then used data mining for classification of the disorders (Miller, Pople, & Myers, 1982; Yap & Clarke, 1996). Luxton (2014) was investigated the implications of Artificial Intelligence (AI) on the psychological healthcare. That study focuses on the current and future applications in four activities comprises clinical training, treatment, psychological assessment, and clinical decision making. Furthermore, Alonso, et al. (2018) presented a review of the techniques and algorithms of data mining in mental health that existing research works in the literature between 2008-2018.

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 9, Issue 2, pp: (44-51), Month: May - August 2022, Available at: www.noveltyjournals.com

Several studies utilized different classification models on the psychological disorders datasets and different mental disorders types, the most widely models used (Lahmiri, Dawson, & Shmuel, 2018) such as Support Vector Machines (SVM) (Das & Ari, 2014; Gandomkar & Bahrami, 2014), Artificial Neural Networks (ANN) and SVM (Gil & Johnson, 2009), Decision Trees (DT), Naive Bayes (NB), and Multilayer Perceptron (MLP) (Kambhampati, Singh, Manikandan, & Ramkumar, 2015). However, normally each machine learning and data mining model has deferent classification scores (Pirooznia, Seifuddin, Judy, & Mahon, 2011).

However, the proper model chosen of the specific dataset is one of the most critical issues in data mining. In practice, the comparing the utilized models could figure out which is better based on its accuracy (Kambhampati, Singh, Manikandan, & Ramkumar, 2015; Gil & Johnson, 2009; Alzaza, 2021).

III. METHODOLOGY

The dataset for predicting psychological disorders was collected from a clinical psychologist. It has 133 instances distributed on 54 attributes including the class attribute. Unfortunately, there are no computerized records for such psychological diagnoses, especially for children. However, the dataset was taken from handwriting papers with the supervision of a psychologist.

Table 1 includes all attributes which formulate a question that should be answered by the parent of the child. Each question (attribute) represents a symptom that affects one or more expected psychological disorders. The class attribute represents one of the six psychological disorders identified as the most cases in Palestine.

Table 1: dataset attributes

#	**Attribute (Question)**
1	Does he have a difficulty in social interaction
2	Does he have difficulty learning skills
3	Does he have difficulty eye contact
4	Does he suffer from delayed pronunciation
5	Does he have repetitive patterns of movement and speech
6	Does he have a habit of sticking to a routine
7	Does he have an overabundance of sensory aspects
8	Have symptoms lasted more than 6 months
9	Fail to pay attention to details
10	Is it difficult for him to perform activities
11	Does he fail to follow instructions
12	Does he not listen to the hadith
13	Does he dislike jobs that require mental effort
14	Do you lose essential items
15	Is he distracted by an external stimulus
16	Does he suffer from excessive forgetfulness
17	Is he writhing in his chair
18	Does he run and climb a lot
19	Does he suffer from impulsivity when playing
20	Does he have a lot of movement
21	Does he talk excessively
22	Is it difficult for him to wait for his responses
23	Does he interrupt others when speaking
24	Is it difficult for him to read the words
25	Is it difficult for him to understand
26	Is it difficult for him to express it
27	Are the numbers difficult to know

28	Is mathematical thinking difficult
29	Does he suffer from cramps
30	Has the disorder persisted for more than a month
31	Was he exposed to a traumatic event
32	Is he repeatedly exposed to the traumatic event
33	Does he have painful memories of the event traumatic
34	Does he have dreams about the traumatic event repeatedly
35	Does he reenact the traumatic event
36	Does he suffer from psychological depression
37	Does he avoid memories of the traumatic event
38	Does he forget an important aspect of the traumatic event
39	Does he have negative beliefs about himself
40	Does he have tantrums
41	Does he lazy to participate in important activities
42	Does he suffer from recklessness
43	Does he have watchful sums
44	Does he have concentration problems
45	Does he have a sleep disorder
46	Does he have a feeling of detachment from himself
47	Is he suffering from loss of consciousness
48	Does he suffer from a lack of response after losing consciousness
49	Does he suffer from staring after unconsciousness
50	Does he suffer from fear and anxiety
51	Is he out of control of the bowel and bladder
52	Does he suffer from headache and fatigue
53	Does he suffer from confusion
54	Does he suffer from a mood disorder

As shown in Table 2, the dataset has nine classes (categories). Six is the main mental disorders are SLD, PTSD, ID, ASD, Epilepsy and ADHD. Two represent a complex case which include more than one disorder class such as (ID and ADHD) and (ADHD and ASD). However, that last class represent the normal case of the children that were diagnosed.

Table 2: The nine classes of the dataset

#	**Class Name**
1	Specific Learning Disability (SLD)
2	Posttraumatic stress disorder (PTSD)
3	Intellectual Disability (ID)
4	Autism Spectrum Disorder (ASD)
5	Epilepsy
6	Attention Deficit Hyperactivity Disorder (ADHD)
7	ID and ADHD
8	ADHD and ASD
9	Normal (No Psychological Disorder)

Data mining process can be divided to two major parts that are data preprocessing and data mining. Data preprocess includes data cleaning, data transformation, and data reduction. Furthermore, data mining usefulness is very sensitive to preprocess steps in terms of its accuracy and consistency. Such steps are valuable in this study since it collected data from paper based which could include some missing data or human errors during data entry process to computer.

Novelty Journals

ISSN 2394-7314

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 9, Issue 2, pp: (44-51), Month: May - August 2022, Available at: www.noveltyjournals.com

- **Data Cleaning:** This step involves the removal of noisy or incomplete data from the original dataset. All data has been double checked and screened by psychologists to be sure that dataset has no outliers or inconsistencies.

- **Data Transformation:** This step includes data mapping and code generation process. However, all attributes (questions) have yes/no formats, so, data were coded to 1 and 0 respectively.

- **Data Reduction:** This step has a vital role since it is applied to obtain relevant data for analysis from the original dataset. The collected dataset has 54 attributes which take long time to be answered by parents. Principal Component Analysis (PCA) was conducted to check ability for summarizing the attributes. Results indicates that 44 attributes can explains 99% of the dataset variance, see figure 1.

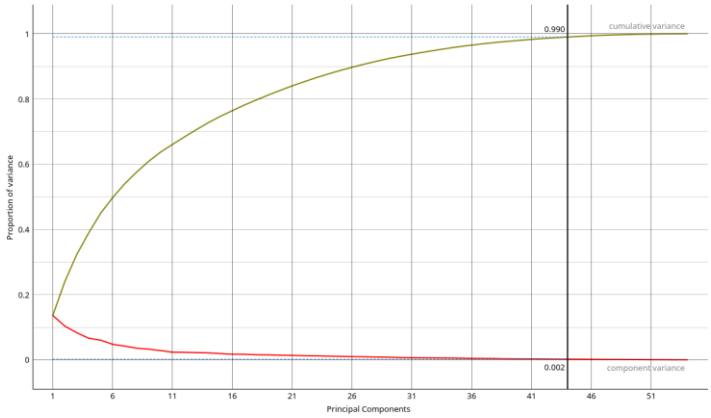

Figure 1: Results of PCA for data reduction

Data mining part of the data process will be presented in the next section.

IV. EVALUATION OF DATA MINING ALGORITHMS

Evaluation and comparing several classification algorithms provide the most accurate algorithm that could proper for such dataset. Fortunately, there are several algorithms of data mining that can be applied on such dataset. Orange Data Mining was used as a data mining tools. Orange is an open source data analysis, machine learning, and visualization tool using Python scripting (OrangeDataMining, 2016). Eight classification algorithms have been evaluated and tested on the psychological disorders dataset that collected from the Palestinian children (see Figure 2).

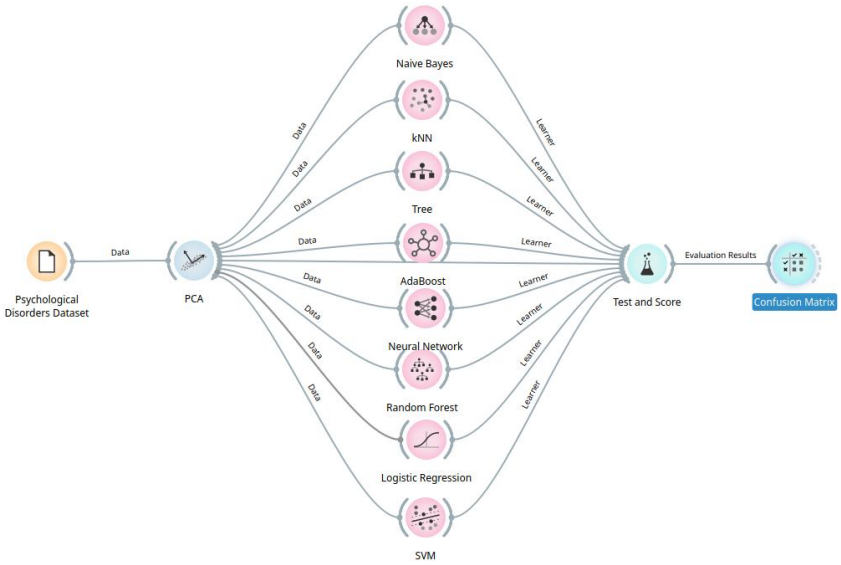

Figure 2: Data mining with all classification models

Novelty Journals

ISSN 2394-7314

International Journal of Novel Research in Computer Science and Software Engineering
Vol. 9, Issue 2, pp: (44-51), Month: May - August 2022, Available at: www.noveltyjournals.com

As shown in Table 3, the Area Under the Curve (AUC) of all classification models are more than 0.9 which considered excellent. Regarding the Classification Accuracy (CA), result shows that AdaBoost, Neural Network, Logistic Regression, Random Forest, and SVM were the best Prediction data mining models with CA percentages: 0.985, 0.985, 0.962, 0.940, 0.940 respectively.

Table 3: Test and scores of classifications

Classification Model	AUC	CA	Precision
AdaBoost	**1.000**	**0.985**	**0.987**
Neural Network	**1.000**	**0.985**	**0.986**
Logistic Regression	**0.998**	**0.962**	**0.968**
Random Forest	**0.997**	**0.940**	**0.929**
SVM	**0.981**	**0.940**	**0.923**
Tree	0.988	0.880	0.871
KNN	0.980	0.842	0.830
Naive Bayes	0.981	0.744	0.925

However, the data mining results show that the dataset was accurate and consistent which is reflected in the percentage of all classification algorithms.

For more practical three of the most predictable models were utilized in some real cases of children. Figure 3 shows the testing results of the real cases based on the testing models that are shown in Figure 4.

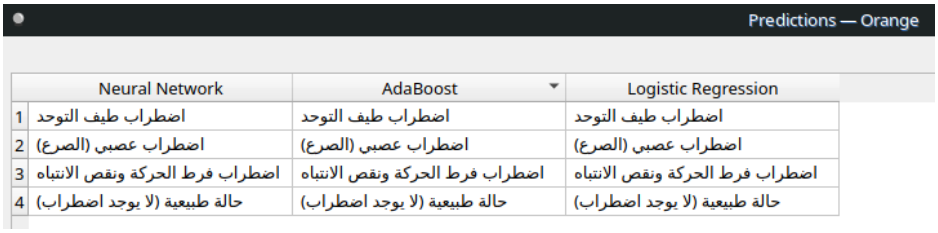

Figure 3: Testing results on real cases

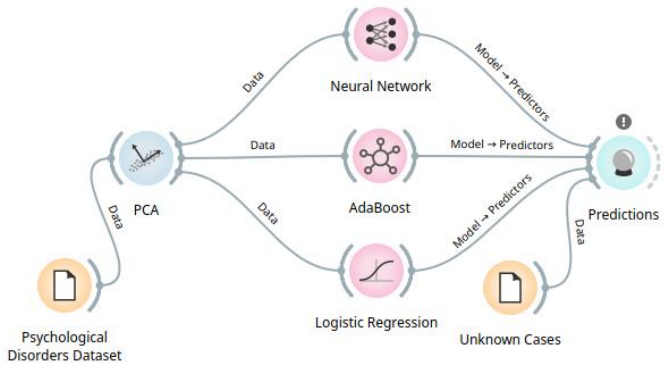

Figure 4: Testing model of classification models

V. CONCLUSION

Prediction in early stages is one of the most factors that affected on treatment process and can make it much effectively and efficiently. Several data mining models are available to make such prediction more accurate and precise. However, this study utilized eight classification models. Results show that all models (AdaBoost, Neural Network, Logistic Regression, Random Forest, SVM, Tree, KNN, and Naive Bayes) have an excellent AUC, i.e. more than 0.9. Furthermore, five models (AdaBoost, Neural Network, Logistic Regression, Random Forest, and SVM) have an acceptable accuracy between 94% to 98.5%.

Indeed, this study provides a dataset and algorithms ready to integrate with software applications such as mobile, websites, or desktop that can enable parents to make personal check-ups on the psychological disorders of their children.

ACKNOWLEDGEMENT

I gratefully thank Shaimaa H. Rostom, Rana O. Alnajjar, and Aya M. Elagha for their helping to collect and digitize the dataset. This cooperation was motivated author to go on with this study, since there was no computerized recorded data available for Palestinian children psychological disorders.

REFERENCES

[1] Alonso, S. G., Torre-Díez, I. d., Hamrioui, S., López-Coronado, M., Barreno, D. C., Nozaleda, L. M., & Franco, M. (2018). Data Mining Algorithms and Techniques in Mental Health: A Systematic Review. Journal of medical systems, 42(9). doi:10.1007/s10916-018-1018-2

[2] Alzaza, N. S. (2021). Author Prediction in Text Mining of the Opinion Articles in Arabic Newspapers. Journal of Electronics and Communication Engineering (IOSR-JECE), 16(2), S. 1-5. doi:10.9790/2834-1602020105

[3] American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Washington, DC, USA: American Psychiatric Association. doi:https://doi.org/10.1176/appi.books.9780890425596

[4] Baird, S., Panlilio, R., Seager, J., Smith, S., & Wydick, B. (2022). Identifying psychological trauma among Syrian refugee children for early intervention: Analyzing digitized drawings using machine learning. Journal of Development Economics, 156. doi:10.1016/j.jdeveco.2022.102822

[5] Brady, C. (5. December 2020). Tackling Health Mental in Palestine. Abgerufen am 4. April 2022 von The Borgen Project: https://borgenproject.org/mental-health-in-palestine/?nowprocket=1#:~:text=Research that has been done, had emotional or behavioral disorders

[6] Cherry, K. (10. May 2022). Theories: A List of Psychological Disorders. Abgerufen am 19. July 2022 von https://www.verywellmind.com/a-list-of-psychological-disorders-2794776

[7] Das, M. K., & Ari, S. (2014). Patient-specific ECG beat classification technique. Healthcare Technology Letters, 1(3), S. 98-103. doi:10.1049/htl.2014.0072

[8] Espié, E., Gaboulaud, V., Baubet, T., Casas, G., Mouchenik, Y., Yun, O., . . . Moro, M. R. (2009). Trauma-related psychological disorders among Palestinian children and adults in Gaza and West Bank, 2005-2008. International journal of mental health systems, 3(1). doi:10.1186/1752-4458-3-21

[9] Gandomkar, Z., & Bahrami, F. (2014). Method to classify elderly subjects as fallers and non-fallers based on gait energy image. Healthcare Technology Letters, 1(3), S. 110-114. doi:10.1049/htl.2014.0065

[10] Gil, D., & Johnson, M. (2009). Diagnosing Parkinson by using Artificial Neural Networks and Support Vector Machines. Global Journal of Computer Science and Technology, 9(4), S. 63-71.

[11] Kambhampati, S. S., Singh, V., Manikandan, M. S., & Ramkumar, B. (2015). Unified framework for triaxial accelerometer-based fall event detection and classification using cumulants and hierarchical decision tree classifier. Healthcare Technologies in Falls: Risk Assessment, Prediction and Detection, 2(4), S. 101-107. doi:10.1049/htl.2015.0018

[12] Lahmiri, S., Dawson, D. A., & Shmuel, A. (2018). Performance of machine learning methods in diagnosing Parkinson's disease based on dysphonia measures. Biomed Eng Lett, 1, S. 29-39. doi:10.1007/s13534-017-0051-2

[13] Luxton, D. D. (2014). Artificial intelligence in psychological practice: Current and future applications and implications. Professional Psychology: Research and Practice, 45(5), 332–339. doi:doi.org/10.1037/a0034559

[14] Marie, M., SaadAdeen, S., & Battat, M. (2020). Anxiety disorders and PTSD in Palestine: a literature review. BMC Psychiatry, 502(20). doi:10.1186/s12888-020-02911-7

[15] Medical Aid for Palestinians. (2020). Mental Health and Psychosocial Support. Abgerufen am 5. April 2022 von Medical Aid for Palestinians: https://www.map.org.uk/what-we-do/mental-health-and-psychosocial-support

[16] Miller, R. A., Pople, H. E., & Myers, J. D. (1982). Internist-1, an experimental computer-based diagnostic consultant for general internal medicine. New England Journal of Medicine, 307(8), 468-476. doi:10.1056/NEJM19820819 3070803

[17] OrangeDataMining. (2016). Orange: Documentation. Abgerufen am 9. May 2022 von https://orangedatamining.com/ docs/

[18] Pirooznia, M., Seifuddin, F., Judy, J., & Mahon, P. B. (2011). Data Mining Approaches for Genome-Wide Association of Mood Disorders. Psychiatric Genetics, 22(2), S. 55-61. doi:10.1097/YPG.0b013e32834dc40d

[19] Sumathi, M. R., & Poorna, B. (2016). Prediction of Mental Health Problems Among Children Using Machine Learning Techniques. International Journal of Advanced Computer Science and Applications (IJACSA)(1), S. 552-557.

[20] WHO. (8. June 2022). Mental disorders: Key Facts. Abgerufen am 19. July 2022 von https://www.who.int/news-room/fact-sheets/detail/mental-disorders

[21] Yap, R. H., & Clarke, D. M. (1996). An expert system for psychiatric diagnosis using the DSM-III-R, DSM-IV and ICD-10 classifications. proceedings of the Annual Symposium on Computer Application in Medical Care (AMIA), (S. 229-233).

Dr. Naji Shukri Alzaza *obtained his BSc degree in Computer/Mathematics in 2000 from the Islamic University of Gaza (IUG), Palestine, and MSc and PhD degrees in Information Technology (IT) in 2007 from the University Utara Malaysia (UUM), Malaysia. He is an Oracle Certified Professional since 2004. He got a Golden Medal from Malaysia Technology Expo (MTE2010) and a Silver Medal from MTE2008. His research interests include mobile learning, mobile commerce, virtual reality, data mining, and information security. Author of several books in IT, community and cultural issues. A political and social community analyst. In 2013, he was the Dean of the Faculty of Information Technology, University of Palestine (UP). In 2014, he was the Dean of the Community Service and Continuing Education at UP. Currently, he is an Associate Professor of Mobile Technology.*