An Arabic Intelligent Diagnosis Assistant for Psychologists using Deep Learning

Asmaa Alayed, Manar Alrabie, Sarah Aldumaiji, Ghaida Allhyani, Sahar Siyam, and Reem Qaid
College of Computer and Information Systems,
Umm Al-Qura University,
Makkah, Saudi Arabia

***Abstract*—Mental illnesses have increased in recent years, especially after Covid-19 pandemic. In Saudi Arabia, the number of psychiatric clinics is small compared to the population density. As a result, psychologists encounter a variety of difficulties at work. The main goal of the current research is to develop a system that assists psychologists in the diagnosis process, which will be based on the DSM-5 (Diagnosis and Statistical Manual of Mental Disorders). The work on this research started with collecting the requirements and identifying users' needs. In this matter, several interviews have been conducted with Saudi Psychologist and then a questionnaire was developed and distributed to psychologists in Saudi Arabia. Following an analysis of the needs and requirements, the system was designed. A deep learning technique was applied during the diagnosing process to address the issues mentioned by psychologists. Additionally, the proposed system helps psychologists by quickly calculating the results of psychological tests. The system was built as a website. The Convolutional Neural Network (CNN) algorithm was used with 96% accuracy to automatically predict the appropriate diagnosis and suggest the most suitable psychological test for the patient to take. System testing and usability testing were also conducted by involving patients and Saudi psychologists to test the usability of the system and the accuracy of the CNN model. The results indicate that the diagnosis prediction was accurate, and that each activity was completed faster. This demonstrated the model's high degree of accuracy and the system's interfaces' clarity. Additionally, psychologists' comments were encouraging and positive.**

***Keywords*—Mental health; psychologist; mental illness diagnosis; psychological test; deep learning; CNN algorithm**

I. INTRODUCTION

Health care is an important aspect of human life, such as eating and drinking. Mental disorders are one of the most important health problems that are beginning to increase recently. According to the latest Saudi National Mental Health Survey report published by King Salman Center for Disability Research [1], two out of five young Saudi nationals have been diagnosed with mental disorders at a certain time in their lives. Around 80% of Saudi nationals have been diagnosed with severe mental disorders. However, they do not seek to receive any type of mental health treatment. Moreover, the number of mental health clinics in Saudi Arabia does not exceed 100 [1], which is a very small number compared to the population density in the country.

As noticed in [1], the number of patients is large compared to the number of clinics which would put more pressure on psychologists. Limited number of patients would be seen daily due to the time each session takes. Sometimes, psychologists need more time to diagnose the patient's condition, either due to the lack of experience or because the patient complains of rare or similar symptoms that require careful consideration to be diagnosed appropriately. Additionally, the psych-diagnostic may differ from one psychologist to another due to the difference in their references and method of diagnosis. The psychologist's human nature would also affect the final diagnosis.

During an interview with an experienced Saudi psychologist [2], he pointed out that there is a problem faced by psychologists, which is most people in our society do not prefer to visit a psychiatric clinic, either due to the unavailability of clinics close to the area in which they live or for personal reasons. Instead, they prefer attending counseling sessions online. However, it is difficult to complete all stages of diagnosis from a distance since there is another stage after the initial diagnosis, which is taking psychological tests. These tests are usually done manually, so the patient has to visit the clinic to take them, which makes the diagnosis go back to the starting point. The calculation of the tests results takes a long time to analyze and reach complete and correct diagnosis results. A number of these tests would take more than a day to analyze and understand the results. Moreover, the calculation and analysis process are prone to human mistakes since the tests have many questions and the way of calculating the results is complicated.

The main goal of the current research is to help psychologists in Saudi Arabia to overcome these problems, and hence facilitate their work. The contribution of this paper is twofold:

First, it proposed a new machine learning model to assist Arabic psychologists in diagnosing three prevalent mental disorders, namely Anxiety, Depression and OCD, based on DSM-5 [3].

Second, it automates the Arabic psychological tests based on APA [4].

This paper is structured as follows: Section II sheds light on the research background. Section III describes related work. Section IV clarifies the adopted research methodology including the results. Section V discusses the results. Section VI concludes the paper.

II. Research Background

The purpose of the research is to develop a diagnosis assistant for Arabic psychologists. Psychological tests are considered essential elements in this diagnosis process. Therefore, a brief background is provided to clarify what mental illness and psychological test are. In addition, the three most important prevalent psychological disorders that are covered in this research will be briefly explained.

A. Psychological Test

Psychological testing (psychometrics) is the systematic use of tests to quantify psychophysical behavior by allowing the person to answer some questions about a particular test. The test is determined by a psychologist, to make predictions about psychological performance. It is one of the tools that helps psychologists to measure how much of a specific psychological construct a patient has. Although psychological tests are available on some websites, they are tools like any other tools, if they are not in the hands of a trained professional, they might not achieve their intended goals [5].

B. Mental Illness

The term "mental illness" or “mental disorder” refers to a health condition that results in emotional, behavioral, or thinking changes it also can be combination of them. Mental illnesses are often accompanied by distress and/or problems functioning at home, work, or in social settings. It can affect anyone regardless of age, gender, social status, or any other aspect. Developing a mental health treatment plan requires collaboration between a mental health clinician (psychologists and psychiatrist) and the patient (and family members if desired). There are many types of treatment available, including psychotherapy (talk therapy), medication, and others. Medication and therapy are often most effective when combined. In this work, the focus is on three mental illnesses, namely Anxiety, Depression, and Obsessive-Compulsive Disorder (OCD)

Anxiety is considered one of the most common mental disorders. According to the Saudi National Survey [1], anxiety is the most prevalent disorder among individuals in Saudi Arabia. Around 12% were diagnosed with separation anxiety disorder, which is the largest percentage compared to the rest of the disorders. In the current project, Taylor Manifest Anxiety Scale[6] will be used. This scale is a test of anxiety as a personality trait to measure the severity of anxiety.

Depression is the second prevalent mental disorder in the world. The latest national survey indicates that 6% of people suffer from depression in Saudi Arabia [1]. It is a serious mental disorder that may lead to a suicide thought. Therefore, discovering the problem and knowing the severity of depression may help in creating a treatment plan that would help people to recover and get rid of this disorder. In the current work, the Beck Depression Inventory [7] will be used to measure characteristic attitudes and symptoms of depression.

Obsessive-Compulsive Disorder (OCD) is considered one of the most common mental disorders in the world [8]. It was rarely diagnosed in the past, but nowadays it is seen as a neuropsychiatric disorder mediated by specific neural circuits and closely related to neurological conditions such as Tourette's syndrome and Sydenham's chorea [8]. OCD could be observable behavior or mental rituals. The obsessions and compulsions of obsessive-compulsive disorder are qualitatively different from obsessive-compulsive personality traits such as perfectionism and excessive conscientiousness. The psychologist can professionally diagnose the patient to determine the type of disorder. In the project, Brown Obsessive Compulsive Scale will be used [5].

III. Related Work

In this section, the systems and web applications that serve the field of mental health and offer services to assist psychologists in the diagnoses process are discussed.

Labayh [9] is a Saudi mobile app approved by the Saudi Ministry of Health. It is considered as a virtual clinic that provides immediate or scheduled consultation with many different psychologists and psychiatrics allow the patient to choose any of them. Because it is not charitable or free platform, the patient must pay for each session. The app also provides two uncertified psychological tests for depression and anxiety, offered to all people who want to try.

Shezlong [10] is an Arabic website that gathers a group of therapists and presents their information in a clear manner with the cost and duration of the counseling session. This gives the users the opportunity to choose the therapist appropriate to their psychological status and budget. The website offers six psychological tests that initially diagnose the patient's condition and make recommendations for the best psychologist to follow up with. However, the patient takes a test based on what he thinks about his situation, not according to specialized advice from a therapist.

Mentalines [11] is a profit-based Arabic website that provides several psychological tests. The user can choose any of them to purchase and take online under the supervision of a psychologist. The website also provides many services related to the mental health like articles, training sessions, therapy trips, and group therapy sessions.

In [12], assistance was provided to psychologists to diagnose Anorexia using natural language processing to assess the expressed emotions by the patient through body description according to DSM criteria [3]. The diagnosis is made by processing the patients notes about their body. To achieve it, researchers used a dataset from a collection of opinions from the Stanford Amazon Dataset service and trained the model using the RNN (Recurrent Neural Network) algorithm. According to [12], the results showed the relationship between psychologists and patients had improved; writing notes made them feel safe, less resistant, and more credible. Although the good results, the model did not recognize some words, which influenced the diagnosis.

In [13], a machine learning model was built to assess five levels of three disorders, namely, anxiety, depression, and stress, without the need for a psychologist's intervention. The dataset was collected through online questionnaires filled out by different participants. The researchers applied eight algorithms that belong to four different categories: bayes, neural networks, lazy, and tree. The results of this research

showed the neural networks RBFN (Radial Basis Function Network) model was the best with depression disorders, while random forest was the best with anxiety disorders. It had a 100% accuracy rate, this occurred due to an imbalanced dataset.

A Machine Learning Approach to detect Depression and Anxiety using Supervised Learning was proposed in [14]. This paper suggested completing the diagnosis process without the need for a psychologist's intervention. It can be done by the patient choosing the disorder, answering the questionnaire, and then the system shows the result to the patient. The dataset was collected through a standard, structured questionnaire. The researchers used four algorithms to develop the model, which are: linear regression, LDA (Linear Discriminant Analysis), CNN (Convolutional Neural Networks), SVM (Support Vector Machine), and KNN (K-Nearest Neighbor). The results of this research showed that the CNN model was the best for depression with 96% accuracy and anxiety with 96.8% accuracy. However, there were limitations in in work. The questionnaire was too long, and there was a potential that the patient would not complete it. Also, according to [14] without a psychologist's help, patients will be less honest.

In [15], a predictive model was built to predict two disorders: Major depressive disorder (MDD) and generalized anxiety disorder (GAD). The researchers used an existing EHR dataset containing biometric and demographic data collected from 4184 undergraduate students. The model was trained using varied non-psychiatric input features such as blood pressure, heart rate, housing. status, and public insurance. The participants were assessed for full Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM IV). For the prediction accuracy, the sensitivity and specificity for MDD were 55% and 70%, and for GAD were 70% and 66% respectively. Additionally, the positive predictive value for MDD was 20% and for GAD it was 16% and the negative predictive value for MDD was 92% and for GAD it was 96%.

From the reviewed literature, it is clear that many different approaches have been proposed to automate the process of diagnosis in the field of mental health. However, they mainly focus on one part of the process, and sometimes without supervision from specialized personnel. In the next sections, our proposed approach is explained. It aims to help psychologists in the whole diagnosis process including the initial diagnosis phase using deep learning, and final diagnosis phase involving automated Arabic psychological testing. Table I compares the reviewed systems and proposed system.

According Table I, there are several systems and research have been created and developed to improve the diagnostic process in the field of mental health. Three of them [6] [7] [8] are actual real system platforms that connect psychologists with patients. They provide a number of psychological tests without any utilization of artificial intelligence algorithms to facilitate the initial diagnosis process. Moreover, they do not have a standard reference for the diagnosis process, but they depend only on the Psychologist's experience.

TABLE I. COMPARING THE REVIEWED SYSTEMS AND THE PROPOSED SYSTEM

System	Analyze the symptoms based on DSM-5	Using AI in the diagnosis process	Provide psychological test	Software system	Digital diagnosis assistant for psychologist
[9]			*	*	
[10]			*	*	
[11]			*	*	
[12]	*	*			*
[13]		*			
[14]		*		*	
[15]		*			
Proposed System	*	*	*	*	*

Four of the reviewed papers, [9], [10], [11], [12] were conducted with the aim of using different algorithms in the diagnosis process. However, they did not rely on DSM-5 [3] on the symptoms analysis process. All of these four researches have not been deployed as an actual system yet.

As noticed, each of the related work focused on a specific side and ignoring others. For example, helping psychologists reach beneficiaries and offering psychological tests without using any new techniques. Other systems [12], [13], [14], [15] adopted artificial intelligence algorithms on the initial diagnosis, but they did not offer any other services, such as offering and calculating psychological tests which can help psychologists to expedite and facilitate the diagnosis process.

The proposed system aims to leverage these limitations. It uses deep leaning algorithm (CNN) in the diagnosis process based on a standard reference which is DSM-5 [3]. Additionally, it offers appropriate physiological tests according to the result of the diagnosis.

IV. RESEARCH METHODOLOGY

The research methodology comprises several steps as summarized in Fig. 1 and clarified in the following subsections.

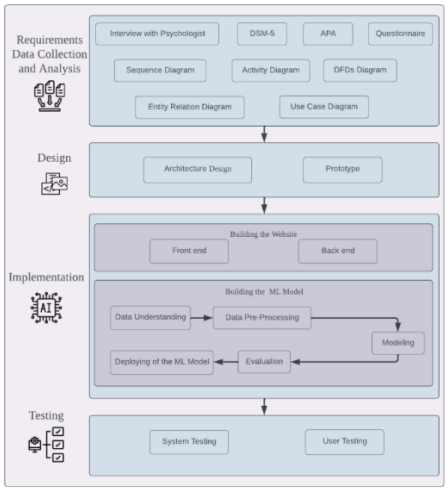

Fig. 1. Research methodology.

A. Requirements Collection and Analysis

In the process of collecting the requirements, we had several online and face-to-face interviews with psychologists from Saudi Arabia in the period between August 27, 2022 and October 18, 2022. Through these interviews, the most important needs of psychologists were identified, and the diagnostic mechanism was fully understood. Additionally, to investigate the significance of this system, an online questionnaire was designed for the experts in the field of mental health [2], and distributed to the psychologists in the Saudi Arabia in the period between 28 Sep 2022 and 8 Oct 2022. Since the target user group was precisely defined in terms of the field, country and specialization, there has been a struggle to reach them in the given timeframe. A total of 70 responses were received. They pointed out that they suffered in terms of time and accuracy to diagnose each patient. They agreed that they are in need of an Arabic system to assist them by automating the process, but at the same time not excluding them.

After full familiarity with the diagnosis process, the functional requirements, including user and system requirements, for the proposed system were identified. The requirements of three main actors namely, Admin, Psychologist, and Patient were analyzed thoroughly using data flow diagrams, use cases and scenarios, activity diagrams, sequence diagrams and Entity Relationship Diagram (ERD) [16] to clarify different aspects of the requirements. Fig. 2 illustrates scenario for the actors including psychologists and patients.

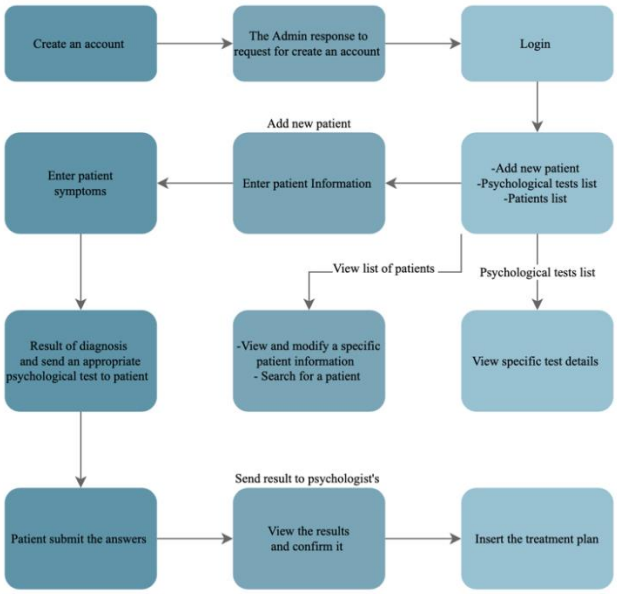

Fig. 2. System scenario to clarify the steps and the main functions of the system.

B. System Design

In this step, the system architecture was designed as illustrated in Fig. 3. Client-server architectural pattern was used to represent the architectural design of the system. Each client was considered as an end-user of the system. The functionality of the system was organized into services, with each service delivered from a separate server. Each client can access the services through the Internet; therefore, the most suitable architectural pattern is the Client-server [16].

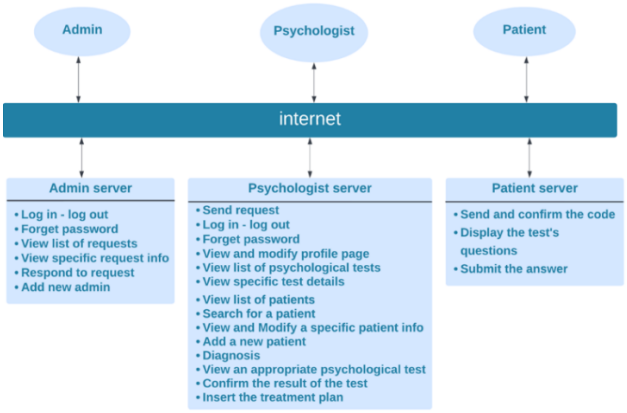

Fig. 3. Client-Server architecture for the proposed system displayed all functions for each actor.

The interfaces of the website were designed as well. Fig. 4 shows the homepage and how it looks on different devices. Fig. 5 and Fig. 6 show two pages that would assist the psychologists during the diagnosis process.

Fig. 4. The proposed website applies responsive design to work properly on different devices.

Fig. 5. This page is displayed to psycholgists allowing them to enter patient's symptoms.

Fig. 6. Result of diagnosis and suggested psychological test for patient to take.

C. Implementation

This section clarifies the implementation step and its details. This step constitutes three main activities, building the website, building of machine learning model, and deployment.

1) Building the Website: To build the website, the frontend and backend were implemented. Frontend implementation involved,

- Writing the content of the website by using HTML.
- Styling the interface by using CSS.
- Making the website interactive by using JavaScript.
- Backend implementation involved,
- Creating website's database in Amazon cloud (AWS).
- Connecting the database to Django.
- Creating tables of the website's database.
- Writing functions of CRUD (Create, Read, Update, and Delete) operations. These operations are implemented in Python. Example of such operations include, calculating the test score and saving the result to be displayed to the psychologist.

2) Building the machine learning model: To build the ML model, the following steps were followed:

- Collecting dataset.

For this research, the dataset was provided by psychologists from Saudi Arabia. There were 305 observations with 13 symptoms. The symptoms were food problems, sleep problems, conscience, communication, face features, speech, mood, behavior, fears, thoughts, focus, attention, and duration. The psychologists provided balanced observations, where around 62 observations were collected for each disorder: anxiety, obsessive-compulsive disorder, and depression. Fig. 7 shows a sample of the collected dataset.

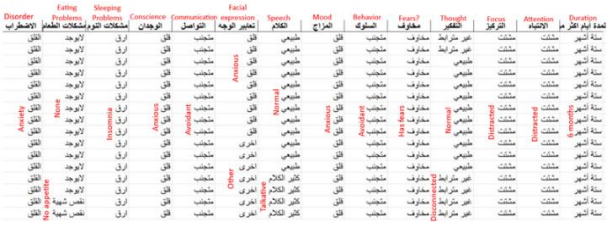

Fig. 7. Sample of the collected dataset showing the symptoms being analayzed.

Pre-processing the data. The first step was to ensure that the dataset did not have missing data. This can be done by using the Python function data.isnull.sum().sum(). A total of 38 missing cells were found as illustrated in Fig. 8.

```
path = 'withmissing.xlsx'
DF = pd.read_excel(path)
print("Number of Missing Data:", DF.isnull().sum().sum())
print("Dataset Shape: ", DF.shape)
```

```
Number of Missing Data: 38
Dataset Shape:  (305, 14)
```

Fig. 8. The dataset before solving the missing data problem.

To fix this problem using Python method fillna(method="bfill") from pandas' library and the way of filling missing data was (backward) method that uses next data point to fill the gap [17], as depicted in Fig. 9.

```
new_df = DF.fillna(method="bfill")
new_df.to_excel('dataset.xlsx')
print("Number of Missing Data:",new_df.isnull().sum().sum())
```

```
Number of Missing Data: 0
```

Fig. 9. The dataset after solving the missing data problem.

Since the dataset was categorical, the pre-processing was done by using one hot encoding technique. One hot encoding is a common approach for transforming categorical features into suitable binary vectors to be used as input in machine learning models [18]. This can be achieved by writing get_dummies Python method from Pandas' library. The dataset was converted into binary vectors and the features increased from 13 to 35 as shown in Fig.10.

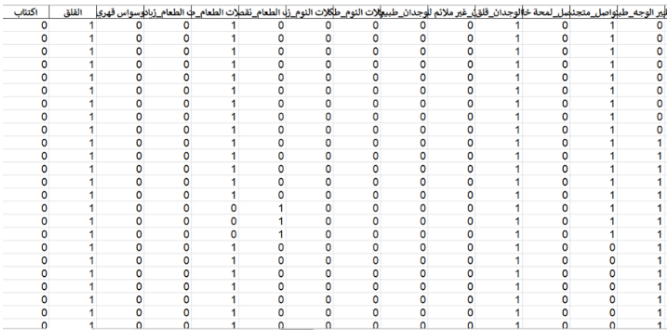

Fig. 10. The result of One Hot Encoding shows that all dataset converted into binary vectors.

The dataset was checked again to make sure that there was not any missing data, using Python Pandas data.isnull.sum().sum() and the result was printed as shown in Fig. 11.

```
data= pd.read_excel('dataset_encoded.xlsx')
print("Number of Missing Data:",data.isnull().sum().sum())
X = data.drop(["اكتئاب","القلق","وسواس قهري"],axis=1)
y = data[["اكتئاب","القلق","وسواس قهري"]]
print("X shape:",X.shape)
print("Y shape:",y.shape)
```

```
Number of Missing Data: 0
X shape: (305, 35)
Y shape: (305, 3)
```

Fig. 11. The result of data preprocessing.

- Training the model.

Two machine learning algorithms were chosen namely, SVM and KNN and one deep learning algorithm was chosen which is CNN.

Support Vector Machine (SVM) is a supervised machine learning algorithm. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that putting the new data point in the correct category in the prediction will be easy [19]. This best decision boundary is called a hyperplane [19]. SVM has multiple kernels: linear, polynomial, and RBF. Kernels are different in making hyperplanes. Linear works with linearly separable data, and polynomial and RBF work with non-linearly separable data [20]. The SVM algorithm was implemented in the current work using the Python Scikit-Learn package.

K-Nearest Neighbor (KNN) is a supervised machine learning algorithm. It is one of the simplest machine learning algorithms, it assumes the similarity between the new case/data and available cases and put the new case into the category that is most like the available categories [13]. This means when new data appears then it can be easily classified into a well suited category by using KNN. The KNN algorithm was implemented in the current work using the Python Scikit-Learn package.

Convolutional Neural Networks (CNN) is a type of artificial neural network. CNN contains multilayer convolutional, and each layer's output feeding into the next layer's input until out layer [21]. Keras model was used, which is a high-level, deep learning API developed by Google for implementing neural networks. It is written in Python and is used to make the implementation of neural networks easy. Keras is an open-source software library that provides a Python interface for artificial neural networks [22]. Keras also supports multiple backend neural network computation. Keras was chosen to implement CNN adopted 20 layers, the rectified linear unit (ReLU) as the activation function and optimized with a learning rate of 0.5.

The training process started by splitting the dataset into 70% for training data and 30% for testing data to ensure the same data is trained and evaluated in each model.

The SVM and KNN models were trained by fitting the training data to the model using fit() Python method as shown in Fig. 12 and Fig. 13.

```
model = OneVsRestClassifier(SVC(kernel="linear"))
model.fit(X_train,y_train)
```

Fig. 12. Training the SVM model.

```
knn = KNeighborsClassifier(n_neighbors=5)
knn.fit(X_train, y_train)
```

Fig. 13. Training the KNN model.

In training KNN, the neighbor that was chosen after more than one attempt which was five neighbors showed the best accuracy.

The CNN model was built with 20 layers, activation function (softmax,relu), and a learning rate of 0.5, and the dataset was trained 20 times using 20 epochs as shown in Fig. 14.

```
# Build The structure of Model
model = keras.Sequential()
# here we define 20 nodes with input shape 35 and activation function called relu
model.add(keras.layers.Dense(20, input_shape=(input_shape,), activation='relu'))
#then the output 3 node because we have multiclassifiaction (anxity/ desperation /OCD)
#and the activation function called softmax
model.add(keras.layers.Dense(3, activation='softmax'))
```

```
#compile the model
optimizer = keras.optimizers.Adam(lr=0.5)
model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])
```

Train model

```
history = model.fit(X_train, y_train, epochs=20,
          validation_data=(X_test, y_test),
          verbose=1)
history_df = pd.DataFrame(history.history)
history_df.loc[:, ['loss', 'val_loss']].plot();
```

Fig. 14. Building and training the CNN model.

- Evaluating the models

To measure the quality of predictions for each model, accuracy, precision, recall, and F1 score were calculated using the confusion matrix, as clarified in Eq. (1) to (4). The confusion matrix represents the true positives (TF), false positives (FP), true negatives (TN), and false negatives (FP) from predicted and actual values [23].

$$Accuracy = \frac{TP+TN}{(TP+TN+FP+FN)} \quad (1)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (2)$$

$$Recall = \frac{TP}{(TP+FN)} \quad (3)$$

$$F1\ Score = 2 \times \frac{Precision*Recall}{Precision+Recall} \quad (4)$$

Table II presents accuracy, precision, recall, and F1 score results of the SVM, KNN and CNN models.

TABLE II. COMPARISON OF THE MODELS' PERFORMANCE

Models	Accuracy	Precision	Recall	F1-score
SVM (Linear)	92.48%	96.8%	97.2%	97%
SVM (poly)	92.48%	96.8%	97.2%	97%
SVM (RBF)	94.48%	96.8%	97.2%	97%
KNN	89.48%	96.9%	96.9%	96.9%
CNN	96.74%	96.8%	97.2%	97%

The results are discussed in Section V below.

- Model Deployment.

Since the results were similar in terms of precision, recall, and F1 score, the CNN model was chosen as the best model to deploy it the website taking into consideration its accuracy result which was 96.74%.

The CNN model was deployed to the website by importing all the required Python libraries, the three target disorders were specified. The user-defined Python function inidiag() is used to connect the input data to the model and makes the prediction of specific diagnosis process then saves the result. After that, the input data was processed by converting it to the binary vector.

D. Testing

The proposed system was evaluated using system testing and usability testing.

1) System testing: System testing is the level of testing that validates a complete and integrated software product. To check how the components interact with each other and with the system as a whole and check the comprehensive test for each input in the system to verify the required output [15]. Four different scenarios were used to conduct the system testing. The results indicate that all errors were minor, and they were related to the front-end part. Upon completing the system testing, the errors were resolved.

2) Usability testing: A total of nine users participated in the testing, five of them represented the role of patients and the other four were psychologists. Each participant was given a number of tasks to perform on the system, and they were observed during the testing. The total time to complete each task for each participant was recorded along with the number of errors per task. After the testing session, the participants were interviewed and asked about the system and their experience in using it. The feedback was positive, and they pointed out the usefulness and ease of use of the system.

V. Discussion

Table II indicates that the results of the three models were almost similar in precision, recall, and F1 score, however, accuracy differences were noticed. SVM performed better than the KNN model, with accuracy of 92.48%. SVM with kernel RBF has performed better than poly and linear models, with accuracy of 94.48%. Among all models, the CNN model performance was the best, with accuracy of 96.74%. CNN uses epoch, and the higher the epoch, the greater the accuracy. When an epoch is executed, it compares the earlier validation result to the original result. As a result, if an issue is discovered, it attempts to minimize the problem by upgrading the layer function.

Fig. 15, 16, and 17 illustrate the curve of training loss and validation loss in the three models. They show that CNN has a better fit compared to SVM and KNN.

In the confusion matrices illustrated in Fig. 18 to 22, 0 refers to Depression, 1 to Anxiety, and 2 to OCD.

The accuracy results were averaged after 20 epochs for CNN and averaged cross-validation [24] with 10 Kfold to avoid overfitting for SVM and KNN. Overfitting is the term for a model that does not generalize properly from observed data to unobserved data and defeating its purpose [25].

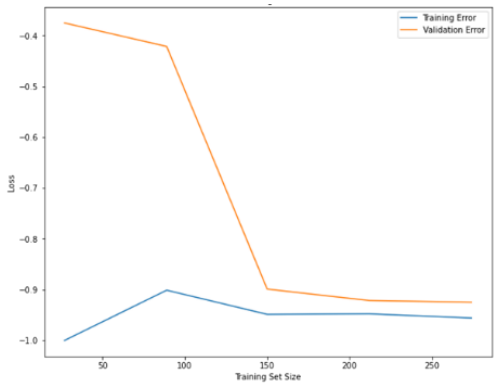

Fig. 15. Plot loss curve for the SVM model.

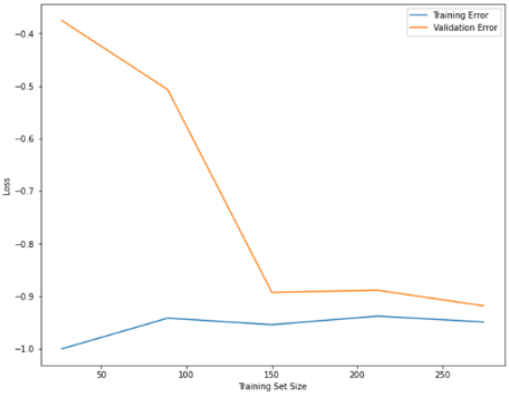

Fig. 16. Plot loss curve for the KNN model.

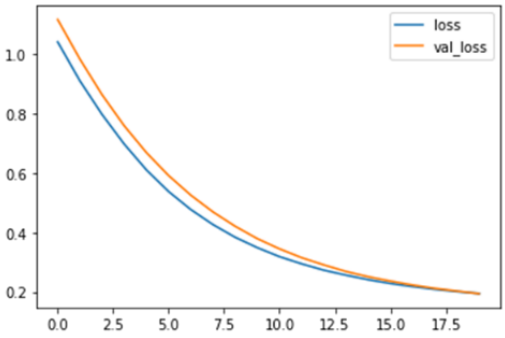

Fig. 17. Plot loss curve for the CNN model.

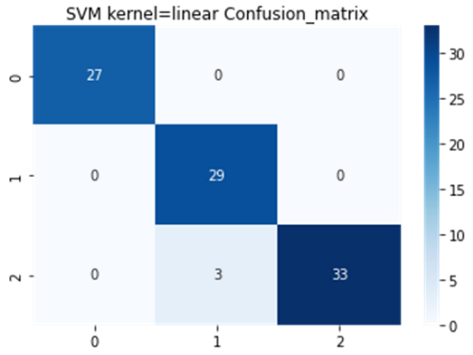

Fig. 18. The confusion matrix for SVM model with linear kernel.

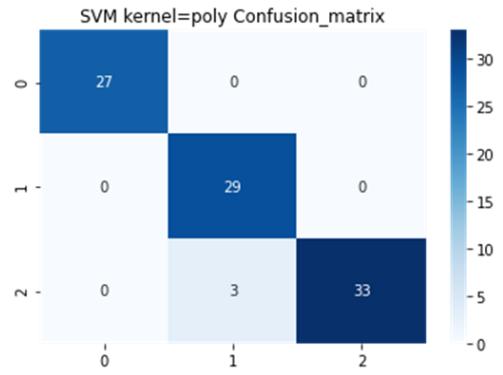

Fig. 19. The confusion matrix for the SVM Model with poly kernel.

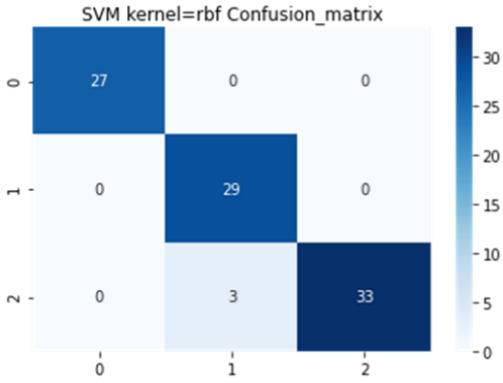

Fig. 20. The confusion matrix for the SVM model with RBF kernel.

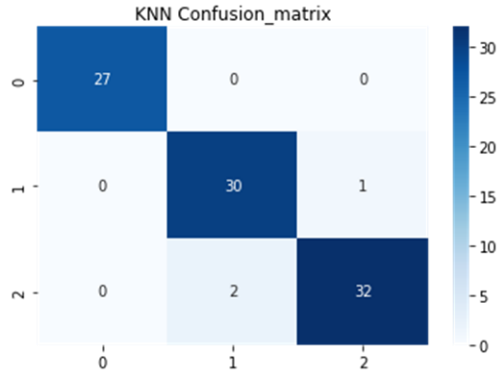

Fig. 21. The confusion matrix for the KNN model.

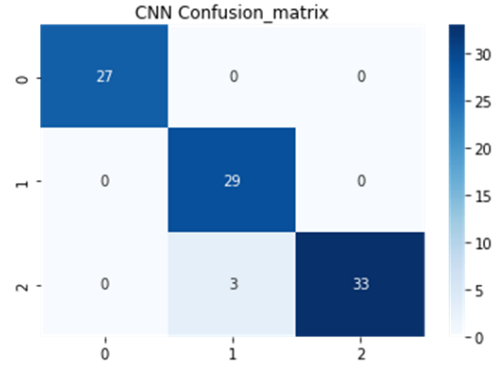

Fig. 22. The confusion matrix for the CNN model.

VI. Conclusion and Future Work

In this paper, an Arabic system was proposed to assist Saudi psychologists in making the diagnosis process of mental disorders accurate, easier, and faster. The methodology followed involved requirements collection and analysis, design, implementation, and testing. The machine learning approach was adopted to diagnose the patient based on the symptoms, prior to deciding which psychological test to take. The focus was on three popular mental disorders namely, Depression Anxiety, and OCD. The dataset was built from scratch for them by collecting anonymous data from psychologists. The ML model was trained by using three different algorithms SVM, KNN, and CNN. The CNN algorithm was chosen to deploy on the system, because it was the most accurate algorithm with a minimum number of errors. Two main limitations were encountered during the work on this research. The proposed system did not cover all mental disorders; it covered three prevalent mental disorders. In addition, the collected dataset was limited to 305 observations. The dataset was balanced with around 100 observations for each mental disorder. The main symptoms needed to diagnose the disorders were considered as features in the dataset. The dataset did not contain the less important features such as the background and history of the patient.

As a future work, we are planning to add more mental disorders that can be diagnosed, add more psychological tests that can be taken by patients, and suggest a treatment plan to the psychologist using a machine learning approach.

References

[1] Y. AlTwaijri, A. Al-Subaie and A. Al-Habeeb, “Saudi National Mental Health Survey Technical Report,” [Online], King Salman Center for Disability Research., Riyadh., 2019. Available: www.healthandstress.org.sa.

[2] A. Alayed, M. Alrabie, S. Aldumaiji, G. Allhyani, S. Siyam, and R. Qaid, “Saber: Digital Diagnosis Assistant for Psychologist,” Google Docs. [Online]. Available: https://docs.google.com/forms/d/e/1FAIpQLSdV055qRlKMLpWn07zwAc2jKY5zwf4N--W76JSdbW8fOVCIqQ/viewform?usp=pp_url.

[3] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders (DSM-5 (R))*, 5th ed. Arlington, TX: American Psychiatric Association Publishing, 2013.

[4] “American Psychological Association (APA),” APA. [Online]. Available: https://www.apa.org/.

[5] G. Domino and M. L. Domino, Psychological Testing: An Introduction, 2nd ed., vol. 2nd ed. Cambridge, England: Cambridge University Press, 2006.

[6] J. A. Taylor, “Taylor Manifest Anxiety Scale,” *PsycTESTS Dataset*. American Psychological Association (APA), 12-Sep-2011.

[7] A. T. Beck, R. A. Steer, and G. Brown, “Beck Depression Inventory–II,” *PsycTESTS Dataset*. American Psychological Association (APA), 12-Sep-2011.

[8] D. J. Stein, “Obsessive-compulsive disorder,” The Lancet, vol. 360, no. 9330, pp. 397–405, 2002.

[9] B. Albeladi, “Labayh psychiatric consultations with ease and privacy,” *Labayh App*, 15-Jun-2022. [Online]. Available: https://labayh.net/en/.

[10] “Talk to your therapist online privately anytime anywhere!,” *Shezlong*. [Online]. Available: https://www.shezlong.com/en/.

[11] Mentalines, “مينتالاينز,” *Mentalines*, 08-Aug-2021. [Online]. Available: https://mentalines.com/.

[12] K. Rojewska et al., “Natural Language Processing and Machine Learning Supporting the Work of a Psychologist and Its Evaluation on

the Example of Support for Psychological Diagnosis of Anorexia," Applied Sciences, vol. 12, no. 9, p. 4702, May 2022, doi: 10.3390/app12094702.

[13] P. Kumar, S. Garg, and A. Garg, "Assessment of anxiety, depression and stress using machine learning models," *Procedia Comput. Sci.*, vol. 171, pp. 1989–1998, 2020.

[14] A. Ahmed, R. Sultana, M. T. R. Ullas, M. Begom, M. M. I. Rahi and M. A. Alam, "A Machine Learning Approach to detect Depression and Anxiety using Supervised Learning," 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), Gold Coast, Australia, 2020, pp. 1-6, doi: 10.1109/CSDE50874.2020.9411642.

[15] M. D. Nemesure, M. V. Heinz, R. Huang, and N. C. Jacobson, "Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence," Sci. Rep., vol. 11, no. 1, p. 1980, 2021.

[16] I. Sommerville, Software engineering, 10th ed. Harlow: Pearson Education, 2019.

[17] C. M. Salgado, C. Azevedo, H. Proença, and S. M. Vieira, "Missing Data," in *Secondary Analysis of Electronic Health Records*, Cham: Springer International Publishing, 2016, pp. 143–162.

[18] C. Seger, "An investigation of categorical variable encoding techniques in machine learning: binary versus one-hot and feature hashing", Dissertation, 2018.

[19] A. Saidi, S. B. Othman and S. B. Saoud, "Hybrid CNN-SVM classifier for efficient depression detection system," 2020 4th International Conference on Advanced Systems and Emergent Technologies (IC_ASET), Hammamet, Tunisia, 2020, pp. 229-234, doi: 10.1109/IC_ASET49463.2020.9318302.

[20] P. K. Intan, "Comparison of kernel function on Support Vector Machine in classification of childbirth," J. Mat. MANTIK, vol. 5, no. 2, pp. 90–99, 2019.

[21] Z. Wang and Z. Qu, "Research on Web text classification algorithm based on improved CNN and SVM," 2017 IEEE 17th International Conference on Communication Technology (ICCT), Chengdu, China, 2017, pp. 1958-1961, doi: 10.1109/ICCT.2017.8359971.

[22] "Keras documentation: About Keras", 2020, [online] Available: Keras.io.

[23] F. Demir, "Deep autoencoder-based automated brain tumor detection from MRI data," in *Artificial Intelligence-Based Brain-Computer Interface*, Elsevier, 2022, pp. 317–351.

[24] D. Berrar, "Cross-Validation," in *Encyclopedia of Bioinformatics and Computational Biology*, Elsevier, 2019, pp. 542–545.

[25] X. Ying, "An Overview of Overfitting and its Solutions," *J. Phys. Conf. Ser.*, vol. 1168, p. 022022, 2019.