Journal of Theoretical and Applied Information Technology
15th January 2025. Vol.103. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PREDICTION OF POLYCYSTIC OVARIAN DISEASE IN MEDICAL DATA USING DEEP LEARNING MODELS

[1]R. PARVATHI,[2]DR.P.GEETHA

Research scholar, Department of Computer Science, Alagappa University, Karaikudi.
parvathikrishnamoorthi@gmail.com
Associate Professor & Head, PG Department of Computer Science,
Dr,Umayal Ramanathan College for Women, Karaikudi.
Email: geeth.ganesan@gmail.com

ABSTRACT

Objectives/Backgrounds: In modern times, polycystic ovarian disease, or PCOD, is a major factor in women's lives. The primary causes of PCOD are a mix of genetic predispositions and hormone imbalance. Every month, the two ovaries in a typical menstrual cycle will release mature, fertilized eggs in turn. **Methods/Statistical Analysis:** For the prepossessing, the PCOD dataset is obtained in a.csv file format from the Kaggle repository. Pre-processing involves removing irrelevant data, adding missing values, and so on. The prediction then receives the final product as its input. **Findings:** Basic characteristics like age, height, and weight are considered for the prediction, along with specific features like I beta and II beta HCG, FSH, LH, endometrial thickness, and screening to see if the patient is pregnant. To determine the precision of the algorithms, the data set that has been processed is categorized using Deep Learning Models like DNN, RNN, and CNN. Classification metrics including precision, recall, and f-measure values are used to compare the performance of the three techniques; DNN performs better than the other two. **Improvement:** Subsequently, further categorization techniques were employed to locate vast amounts of data.

Keywords: *Polycystic Ovarian Syndrome, DNN, RNN, CNN, Polycystic Ovarian Disease Prediction., Precision, f-measure, recall, Accuracy.*

I. INTRODUCTION

Innovation and human beings working together could lead to improved health care services. Within artificial intelligence, deep learning allows to automatically learned system, get better without needing to be explicitly designed. It primarily focuses on developing new algorithms for deep learning that grant access to designated datasets and make use of the information for open-access research and study. Deep learning applications help bring about major change, particularly in the health industry where they are employed for data identification as well as prediction, photo recognition, diagnostics, and other tasks.

Polycystic ovarian syndrome is an endocrine disorder that typically affects women in their adolescent years. It was first told by Leventhal along with Stein in 1935. Hormone imbalance is one of the main symptoms of polycystic ovarian syndrome in women. Serious health problems result from it, including irregular menstrual cycles and trouble getting pregnant. Women with PCOS are susceptible to a number of diseases, such as high blood pressure, coronary artery disease, type 2 diabetes, overweight or obese women, gynecological cancers, high-risk pregnancies, and Mellitus. PCOS symptoms include acne, high blood pressure, period inconsistencies, weight gain, elevated androgen hormone levels, and more. Since PCOS inhibits the follicles' developmental process, which describes the ovaries' maturation, we view it as the main cause of infertility [1].

A new study has shown that there is an elevated chance of miscarriages in the first trimester. Of women of reproductive age, 12–21% have PCOS, and 70% of instances go untreated. This condition can be healed by following the doctor's prescription and altering one's lifestyle. Medication includes things like contraception pills, diabetes tablets, anti-androgen drugs, fertility testing, and ultrasounds. The diagnosis of PCOS is made by excluding out irrelevant symptoms or test findings, which are usually the product of an unskilled composite pathomechanism. Medical practitioners are forced

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to do numerous clinical tests and pointless radiological imaging procedures as a result of these various symptoms [2].

Women's reproductive systems are entirely dependent on mismatched hormones, which must be in balance for the processes required for ovulation, conception, and the growth of a child inside their wombs. The four hormones needed are follicle stimulating hormone (FSH), luteinizing hormone (LH), progesterone, and estrogen. Progesterone and estrogen are produced by the ovaries, while FSH and LH are produced by the pituitary gland. For women to have an adequately balanced reproductive system, progesterone and estrogen are both essential. Women with PCOS worry about a number of issues, including sleep apnea, being infertile, excessive bleeding from the uterus, high cholesterol, elevated lipids, a condition known as nonalcoholic fatty hepatic disease, anxiety and depressive symptoms, elevated blood pressure, obesity, metabolic syndrome, premature deliveries, and also cardiac dangers. Amenorrhea, which affects 30 to 40% of women, gaining weight around the waist, swollen breasts prior to menstruation, and excessive or undesirable facial or body hair development are all signs of PCOS. Neuralgic pain is present. ovarian cysts, vulvar and vaginal irritation, and hysteria [3]. The architecture of this research project is depicted in Figure 1.

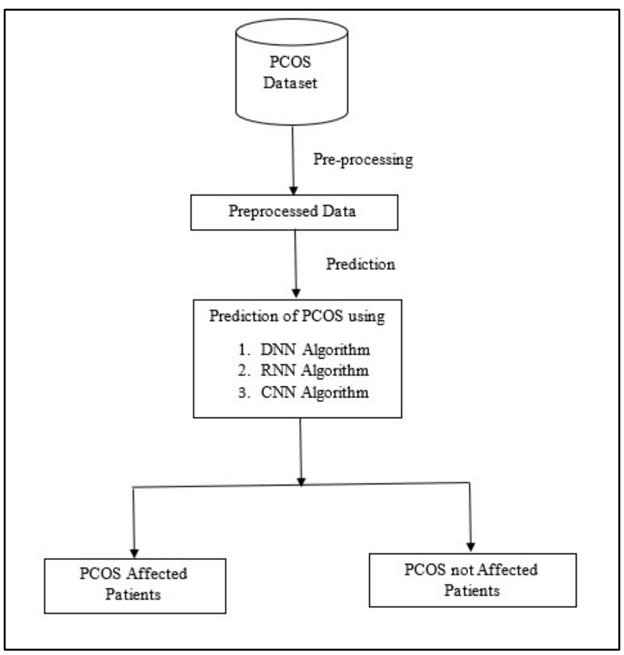

Figure 1: Architecture Diagram

As follows is the arrangement of the remaining study. While Section II covers a scan of the literature, Section III describes the tools and techniques employed in this study, including the usage of deep neural networks (the DNN), recurrent neural networks (the RNNs), and convolutional neural networks (the CNNs) to determine accuracy. The experiment's results are displayed in section IV. Section V concludes this study with the innovative information it contains.

2. LITERATURE SURVEY

To predict PCOS, Bharati S et al. used machine learning algorithms from Kaggle. For instance, the authors used an univariate feature selection (UFS) approach to apply gradient boost RF, LR, as well as a hybrid RFLR model that combined RF and LR using the PCOS dataset in [4]. To train and test the models, they divided the dataset using holdout and cross-validation techniques. According to the results, RFLR coupled UFS worked the best. Principal Component Analysis (PCA) was used by the authors of [5] to lower the overall number of characteristics. To predict PCOS, they used SVM, NB, KNN, LR, RF, as well as RF with characteristics. The outcome demonstrated that RF has the highest accuracy level. Using correlated feature selection techniques, the authors in [6] selected a selected group of characteristics from the database. SVM, LR, RF, DT, KNN, QDA, LDA, GB, AdaBoost (AB), XGBoost (XB), and CatBoost were among the many machine learning models they employed. Based on correlation levels, they were able to select the optimal model. The results showed that RF had the most effective model.

In [7], the authors compared several models, such CNN, ANN, SVM, DT, also KNN, and employed feature selection strategies to detect PCOS. RF created the model that performed the best. According to Pearson correlation, the top qualities were identified in [8]. When compared to their SVM, the method used SVM, RF, as well as XG boost perceptron with multiple layers with chosen features has the greatest accuracy rate. The authors of [9] suggested a feature selection method that combines filters and wrappers to lower the feature count. Additionally, they used a range of machine learning algorithms with certain characteristics to anticipate PCOS. Most accurate model was SVM.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The development of an automated PCOS identification method based on medical and metabolic markers is the aim of the research conducted by Palak et al. Bayesian along with logistic regression algorithms are used in the study strategy to classify features. Out of the two systems that were examined, the Bayesian classifier has the highest accuracy (93.93%) and is the best-designed model [10]. According to Denny et al. [11], The intention is to avoid the costs and time related to clinical diagnostics such as ovarian scanning. The research design transforms PCOS characteristics using PCA by utilizing machine learning techniques such as KNN, SVM, RF, etc. With an accuracy of 0.89, Random Forest generated the most efficient and precise model for PCOS detection.

Using a dataset from Subrato et al.'s Kaggle library, data-driven diagnosis of PCOS was conducted [12]. Among the classifiers used are logistic regression, RFLR, random forest, and gradient boosting. The approaches used in the study design include holdout and cross validation. With a 90% recall value, 91.01% RFLR was the best testing accuracy. Ning-Ning Xie and colleagues stated that the goal is to identify gene biomarkers and create a diagnostic model[13]. A computational strategy is to combine several machine learning algorithms, like ANN and Random Forest. A unique diagnostic model was developed having an AUC of 0.7273 within the transcriptome data and 0.6488 for the RNA-seq dataset.

Sonograms that just display the physical symptoms will be used to classify PCOS, in accordance with Priyanka et al. [13]. used a range of model-finding and classification methods, such as Random Forest, Random Tree, M5 rules, Decision Table, K-star, IB1 instance-based, and locally weighted learning. K-star did the best in comparison to other algorithms. Tanwani Namarat [14], A model based on the root causes and manifestations of PCOS is created, and the output is a prediction as to whether PCOS will manifest or not. Two popular machine learning guided classification techniques are KNN and Logistic Regression. With a 92% accuracy rate, the logistic regression approach is the most accurate model ever developed.

Early evaluation and management of this condition, Pijush et al. [15]. To identify PCOS early, SMOTE was paired with five other algorithms: Random Forest, Decision Timber support vector machine KNN, Logistic Regression, and SMOTE. Results from the most effective model were as follows: AUROC: 95.6%, F1 score: 0.010 seconds, recall: 98%, accuracy: 98%, training time: 97.11. Using a probabilistic technique, Khan Inan et al. [16] selected statistically significant parameters connected to PCOS cases. The Chi-Square test, the ENN, ANOVA, and SMOTE tests were used to identify important features. Classifiers such XG Boost, which is SVM, KNN, NB, MLP, RF, and AdaB were used. Compared to all other classifiers, G Boost performed better with 0.96 precision and 0.98 recall.

To get over these limitations, Mutinda et al. in [17] introduce the LeBERT sentiment categorization model, which integrates an emotional vocabulary, N-grams, BERT, and CNN. The model vectorizes words selected from a subset of the source text using BERT, N-grams, and sentiment lexicon. CNN, a type of deep neural network classification algorithm, assigns an output sentiment category after feature mapping. The suggested method is assessed using three publicly available datasets: Yelp restaurant reviews, Amazon product reviews, and IMBD movie reviews. Text mining along with text processing approaches were described in depth by S. Vijayarani et al. [18] in order to find knowledge in text material that was contributed by members in social media. They talked about stemming, stop word removal, the TF/IDF approach, and the first text preprocessing phase. We also looked at each group's stemming algorithms, which demonstrated the benefits and drawbacks of different stages. This work provided step-by-step instructions on text preparation, which is necessary before performing sentiment evaluation or text mining.

3. MATERIALS AND METHODS

This section discusses the problem definition for this research project. Organizing the many forms of unorganized information in a medical record is the main challenge associated with healthcare data mining activities. It is necessary to comprehend the patterns and important phrases in the medical record of a person, which can vary greatly, in order to accurately detect disease from medical data. The PCOS dataset, which was preprocessed to remove redundant data, missing data, and unnecessary attributes, was then submitted to the DNN, RNN, and CNN algorithms for prediction purposes in order to determine whether or not the affected patients would be identified. The PCOD dataset is available in a.csv file format from the Kaggle repository. The dataset is named PCOS_data.csv in the file. Over 5000 records make up this research project, which also included an analysis of 541 patient records.

A. Deep Neural Network

Deep learning has been successful in various application fields in recent years. The science of machine learning is rapidly expanding and has found application in both traditional and novel fields. Based on many learning categories, including supervised, semi-supervised, and unsupervised instruction, numerous ways have been created. A subset of machine learning known as deep learning (DL) was created utilizing deep learning architectures or hierarchical learning techniques. The process of learning involves estimating model parameters to enable the trained model to complete a task. For example, in an Artificial Neural Network (ANN), the parameters that need to be estimated are the weight matrices. DL, on the other hand, has several layers between the layers of input and output in its design [29] [30]. Pattern recognition and feature learning can benefit from the multi-stage nonlinear information processing made possible by this hierarchical structure design. In this context, representation learning refers to a data-driven learning approach. Recent research indicates that hierarchical structures are used in DL-based representation learning, whereby higher-level concepts are described in terms of lower-level concepts and vice versa.

The training cost of DNNs is attributed to the massive number of multiply-and-accumulate primitives that are needed to calculate the weighted sums of the input to the neurons. Methods such as low-precision arithmetic and sparse connectivity [19] [20] [21] are widely studied to tackle this problem. For example, compared to 32-bit fixed-point encoding, using 8-bit fixed-point decoding for AlexNet prediction on the Cifar-10 data has demonstrated a 6 reduction in energy consumption [22]. On the other hand, an extremely large neural network, such as LSTM with a mix of experts, will require about 137 billion parameters to use 32-bit precision [23]. Due to its ability to address a broad range of problems in a wide range of contexts, deep learning (DL) is referred to as a global learning technique in certain research papers. Stated differently, DL is not specific to the task. There are other DNN types described in the literature, such as LSTM (long short-term memory) and convolutional neural networks (CNNs). This section discusses the most effective and promising types of DL architectures that have been suggested for healthcare image localization in the literature.

B. Convolutional Neural Network (CNN)

A deep learning networks is usually constructed by stacking the convolutional and pooling layers together. One layer at a time, the intricate details of sample data are dynamically learned. Consequently, the acquired features possess layer characteristics and exhibit enhanced generalization and mapping skills [26]. The convolutional layer is composed of two extraction of features layers and two feature mapping layers. The local features may be extracted because every neuron in the extraction of features layer has an interaction from the input it receives to the local sense domain of the preceding layer. Because it performs better than other ML models in terms of accuracy, CNN has become one of among the most popular deep learning models (DL models) in computer vision. This is especially true for picture classification. The CNN requires greater computational power and memory even though it performs better than previous DL models [29] [30]. There, centralized high-performance systems are used to conduct CNN training. Since Krizhevsky et al. [24] became victorious in the 2012 ImageNet competition, CNNs have been incredibly well-liked across a wide range of fields as a useful method for classifying images.

The primary benefits of CNN stem from its ability to function as an autonomous network that doesn't require supervision [25]. Three different types of layers make up the CNN architecture: convolutional, pooling, and completely connected. Additionally, a CNN employs convolution operations in at least one layer, in contrast to other DL models. There is a thorough explanation of the CNN's regular layers. (a) Convolutional layers: These layers consist of a group of filters for which parameters must be learned. Both the height and weight of the filter are lower than the input. To calculate an activation map, a filter moves through the input in both height and width, and the dots that result that exist between input with the filter are taken into

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

account at each spatial position. The activation maps of each filter are then stacked to calculate the convolutional layer output. The pooling layer reduces the number of parameters that must be computed by downsampling the representation. (b) Pooling layer: Using smaller grid regions as input, pooling processes generate a single integer for each region. For computation, the mean function (average pooling) or the maximum function (max-pooling) are usually used. (c) Dropout layer: dropout randomly turns the output edges of hidden units, which are neurons that form hidden layers, to 0 during the training phase.

C. Recurrent Neural Network (RNN)

The recurrent connections in the RNN allow the network to remember prior input patterns. Because it is expected that the ROI is dispersed over numerous adjacent medical imaging slices (such as CT or MRI), consecutive slices are linked. Data is sequentially captured by the RNN from the input slices. Novel recurrent architectures have recently been suggested in the literature as an improvement over the original RNN since both past and future values for input are able to be used in many ways to affect the output value.

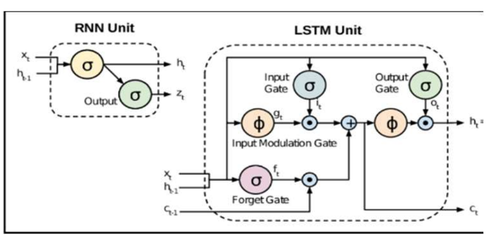

Figure 2. General Architecture Of LSTM

The most popular type of RNN is the LSTM. Compared to RNNs and conventional neural feedforward networks, LSTMs have a number of advantages. Their capacity to actively recall patterns over extended periods of time is the reason behind this. One of the issues with RNNs is that the LSTM's chainlike construction allows it to retain knowledge for extended periods of time [11]. As seen in Figure 3, the LSTM incorporates an internal memory that contains information about its inputs and is managed by a range of fully connected gates. For this reason, localized spatiotemporal features can be extracted by the LSTM. The following are the LSTM's three main components: (A) Forget gate: expunges information that is not anymore necessary to finish the objective. The improvement of network efficiency requires the completion of this stage. (a) The input gate is in charge of supplying inputs to the cells. (c) Output gate: configures the required inputs and outputs. The LSTM has been explored for several applications, including automatically localizing ROIs in medical videos, because to its capacity to examine spatiotemporal information from videos.

4. EXPERIMENTAL RESULTS

The objective of this research work is to predict the heart disease using text dataset (Categorical data). Various methods are used to find the issues in the health care ie., medical data sets and for the prediction of individual expenditures and disease risks for patients. There are two main process in this research work (1) Data preprocessing and (2) Prediction. The duplicate record, missing data, noisy in the consistent data will be removed from the database in preprocessing. For prediction, PCOS dataset given in the prediction process by means of Deep Neural Network Algorithm, Recurrent Neural Network Algorithm and Convolutional Neural Network Algorithm to predict the affected patients or not.

A. Data Set

The PCOD dataset used in this study was obtained as a.csv file from the Kaggle repository. The dataset is named PCOS_data.csv in the file. Over 5000 records make up this research project, which also included an analysis of 541 patient records.

.

Sl. No	Patient File No.	Age (yrs)	Weight (K	Height(Cn	BMI	Blood Gro	Pulse rate	RR (breatl	Hb(g/dl)	Cycle(R/I)	Cycle leng	Marraige S	Pregnant(	No. of abc	I beta-H	II beta-H	FSH(mIU/	LH(mIU/m
1	1	28	44.6	152	19.3	15	78	22	10.48	2	5	7	0	0	1.99	1.99	7.95	3.68
2	2	36	65	161.5	24.9	15	74	20	11.7	2	5	11	1	0	60.8	1.99	6.73	1.09
3	3	33	68.8	165	25.3	11	72	18	11.8	2	5	10	1	0	494.08	494.08	5.54	0.88
4	4	37	65	148	29.7	13	72	20	12	2	5	4	0	0	1.99	1.99	8.06	2.36
5	5	25	52	161	20.1	11	72	18	10	2	5	1	1	0	801.45	801.45	3.98	0.9
6	6	36	74.1	165	27.2	15	78	28	11.2	2	5	8	1	0	237.97	1.99	3.24	1.07
7	7	34	64	156	26.3	11	72	18	10.9	2	5	2	0	0	1.99	1.99	2.85	0.31
8	8	33	58.5	159	23.1	13	72	20	11	2	5	13	1	2	100.51	100.51	4.86	3.07
9	9	32	40	158	16	11	72	18	11.8	2	5	8	0	1	1.99	1.99	3.76	3.02
10	10	36	52	150	23.1	15	80	20	10	4	2	4	0	0	1.99	1.99	2.8	1.51
11	11	20	71	163	26.7	15	80	20	10	2	5	4	1	2	158.51	158.51	4.89	2.02
12	12	26	49	160	19.1	13	72	20	9.5	2	5	3	0	1	1.99	1.99	4.09	1.47
13	13	25	74	152	32	17	72	18	11.7	4	2	7	1	0	1214.23	1214.23	2	1.51
14	14	38	50	152	21.6	13	74	20	12.1	2	5	15	0	0	1.99	1.99	4.84	0.71
15	15	34	57.3	162	21.8	13	74	22	11.7	2	5	9	0	0	1.99	1.99	7.45	3.71
16	16	38	80.5	154	33.9	13	78	22	11.4	2	5	20	0	0	1.99	1.99	9.51	2.51
17	17	29	43	148	19.6	13	80	20	11.1	2	5	2	1	0	8104.21	91.55	2.02	0.65
18	18	36	69.2	160	27	13	72	18	10.8	2	5	7	0	0	1.99	1.99	4.86	2.96
19	19	31	52.4	159	20.7	17	72	18	12.7	2	5	7	0	0	1.99	1.99	6.05	1.05
20	20	30	85	165	31.2	16	72	18	12.5	4	7	7	0	0	23.58	1.99	1.89	0.81
21	21	25	64	156	26.3	11	70	18	11.2	2	6	6	0	0	1.99	1.99	2.82	1.3

Figure 3. Sample Dataset

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The dataset used in this research project is in.CSV format. There are 43 attributes in the dataset, and 541 items are examined. This pertains to patients who may or may not have PCOS. Figure 3 displays the dataset sample. Each patient has different characteristics.

B. Preprocessing

Data preparation is the process of transforming raw data into a comprehensible format. Since raw data is unusable for data mining, it also serves as a crucial stage. Prior to using deep learning or other data mining techniques, one needs evaluate the quality of the data. Preprocessing involves removing duplicate records, missing data, and noisy data from the input dataset so that the consistent data may be stored in the database. The four primary tasks in data preprocessing are cleaning the data, integrating the data, reduction of data, and data transformation.

1. Data Cleaning

The practice of eliminating erroneous, lacking, and deceptive information from datasets and substituting the absent values is known as data cleaning. Here are a few techniques for data cleansing:

Managing absent values

- Regular values such as "Not Available" or "NA" could be utilized to bridge the gaps.
- If a dataset is huge, it is not recommended to manually fill in the missing values.
- The attribute's median amount can be utilized to replace the value that is missing if the data has an irregular distribution, but the attribute's mean value may be used when the data is normally distributed.
- Regression or decision trees methods can be applied to replace a missing variable with the most likely value.

Handling noisy data

Generally speaking, noisy describes random errors or having additional data points. Here are a few methods for working with noisy data. One of the most important procedures is handling noisy data as it optimizes the model we're using.

- Binning: This method is applied to handle or reduce noise in data. After sorting the data, the sort values are divided and stored as bins. Three methods can be used to smooth the information in the bin. Using the bin mean approach to smooth: By using this technique, the mean of the bin is used to replace each value in the bin. Smoothing via bin median: This technique replaces the median value with every value in the bin. Smoothing by bins border: This method substitutes the closest border value for values instead of the lowest and highest bins value [35] [38].
- Regression: Which is a useful tool for handling and smoothing data that contains superfluous information. Identifying a variable that is acceptable for our investigation is made easier with the help of purposeful regression.
- Clustering: This is a technique for organizing data and identifying data anomalies. Usually, clustering is used in unsupervised learning.

2. Data Integration

assembling data from multiple places to produce a single dataset. One of the key elements of data management is the process of data integration. When integrating data, there are a few things to keep in mind.

- Schema integration: integrates metadata from those sources, which is a collection of data from several sources that characterises other data.
- Entity identification problem: locating entities across several databases. The system or user, for instance, needs to be aware of the student id from one database and the studentname from another database that are both part of the same entity.
- Detecting and resolving data value concepts: When databases are combined, various data may be used. One database's attribute values might not match those of another. The date format, for instance, can be different, such as "MM/DD/YYYY" or "DD/MM/YYYY".

3. Data Reduction

The results are almost identical while analysis is facilitated by reducing the amount of data. As a result of this reduction, storage space is also reduced. Data reduction can be achieved through dimensional reduction, numerosity reduction, and data compression.

- Real-world applications require dimensional reduction because of the size of the data. The

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

data set's dimensionality can be decreased by decreasing random variables or qualities in this procedure. Retaining the original properties of data attributes while combining and merging them. In addition, reducing this leads to a reduction in storage space and computation time. When data is very dimensional, the 'Curse of Dimensionality' arises [33].

- Numerosity Reduction: With this technique, the volume of the representation of the data is decreased. No data will be lost during this reduction.
- Data compression is the process of minimizing data. Compression can be achieved either without loss or with loss. The use of lossless compression is when there is no data loss during compression. The opposite of lossy compression is that it only removes information that is not necessary.

4. Data Transformation

The process of changing the structure or format of data is known as data transformation. The requirements determine how complex this phase will be [40] [42]. There are a few methods for transforming data.

- Smoothing: Using algorithms, we can remove distortion from a dataset so that its essential characteristics are more easily discerned. Smoothing makes it possible to see even minute changes that help in forecasting.
- Aggregation: This method saves the data and presents it as an overview. The data analysis description includes the information set, which is sourced from multiple sources. This is an important step because the accuracy of the data is determined by both its amount and quality. When both the quantity and quality of the data are high, the outcomes are more relevant [34].
- Discretization: In this process, the continuous data is divided into intervals. Discreteization reduces the amount of data. Rather than stating the class hour, we could, for example, specify a range, such as 3–5 pm or 6–8 pm.
- Normalization: It is a method of scaling data to fit inside a smaller range for display. An illustration using a -1.0 to 1.0 range.

C. PerformanceMetrics

Performance metrics provide information about the dataset's performance. The evaluation criteria used to evaluate the suggested scheme's presentation are Precision, Recall, and F-measure. Here, conventional count values are taken advantage of, including True Positive (Tp), True Negative (Tn), False Positive (Fp), and False Negative (Fn).

$$Precision = \frac{T_p}{T_p+F_p} \times 100 \quad (1)$$

$$Recall = \frac{T_p}{T_p+F_n} \times 100 \quad (2)$$

$$Fmeasure = 2 * \frac{Precision*recall}{Precision+recall} \quad (3)$$

These measurements are precisely taken in order to determine the algorithms' performance based on the examination of the information set selected for the study.

D. Results and Discussions

This section provides a thorough report on the outcomes of the three current PCOS disease algorithms, which are written in the Python programming language and include CNN, RNN, and DNN. The total number of patients with PCOS disease is displayed in Table 1, and the dataset's graphical form is shown in Figure 4.

TABLE 1: Number of Patients affected by PCOS Disease

PCOS	Number of Patients
Affected	364
Not Affected	177
Total	541

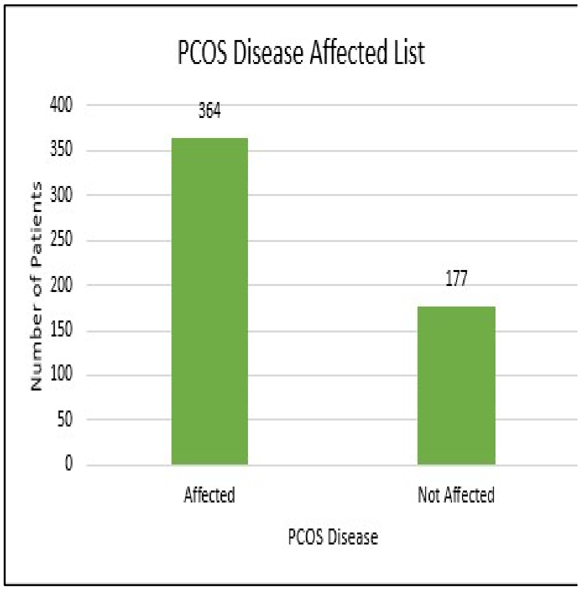

Figure 4. Number Of Patients Affected By PCOS Disease

All three methods' efficacy is evaluated using f-measure, accuracy, precision, and recall. The percentages of recall, f-measure, and precision for each of the three algorithms are displayed in figure 5.

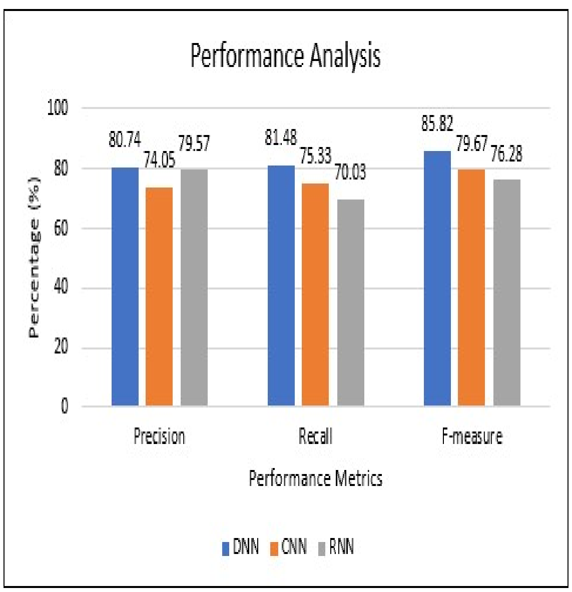

Figure 5. Performance Analysis

Figure 5 shows the results of the performance analyses of the three algorithms. The DNN algorithm achieves the highest results, with precision (80.74%), recall (81.48%), and F-measure (85.2%). The accuracy, recall, and f-measure values attained by the CNN algorithm are 74.05%, 75.33%, and 79.67%. The RNN Algorithm is achieved with precision of 79.57%, recall of 70.03%, and F-measure of 76.28%. It is evident from the data that the algorithms DNN algorithm performs better than the other two current approaches.

TABLE 2: Performance Analysis Of DNN, CNN And RNN

Algorithms	Precision	Recall	F-measure
DNN	80.74	81.48	85.82
CNN	74.05	75.33	79.67
RNN	79.57	70.03	76.28

The performance analysis is displayed in Table 2, and the time and memory usage of each algorithm is displayed in Table 3.

TABLE 3: Average Computational Time And Memory Utilization Of Algorithms

Algorithms	Execution time (ms)	Memory utilization (bits)
DNN	2925	145579
CNN	3442	199366
RNN	4211	257399

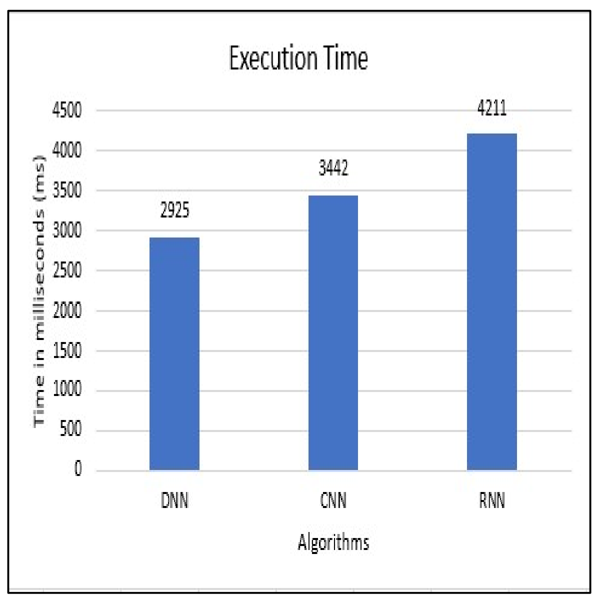

Figure 6. Results Based On Run Time

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

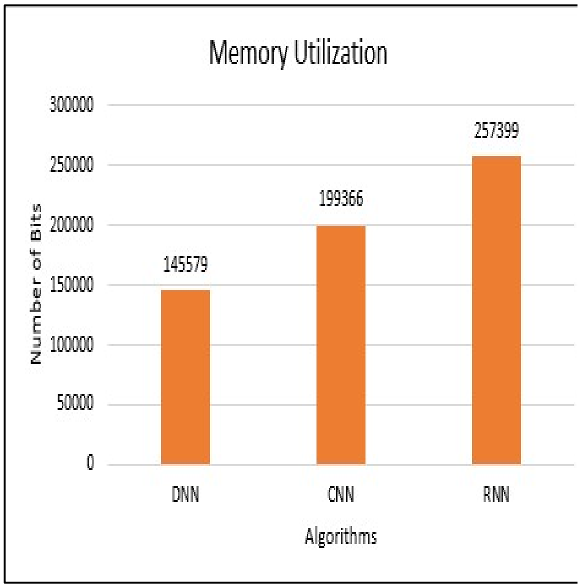

Figure 7. Results based on Memory Space

The final dataset of all three classification algorithms' execution times is graphically represented in Figure 6.A visual representation of the storage space used by the dataset produced by the three classification methods is displayed in Figure 7. Figure 6 illustrates how much faster the DNN algorithm computes is than the CNN and RNN algorithms. Figure 7 illustrates how, for the chosen dataset, the DNN algorithm uses comparatively less memory than the RNN and CNN algorithms.

V. CONCLUSION

In general, it is impossible to anticipate which method would perform best for any given type of data set. However, the outcomes of various categorization algorithms vary depending on the type of data used for analysis. In real-world applications, the classification algorithms are essential for assessing various types of data. The PCOS Disease Dataset is used in this study. Two different output formats are produced, one of which lists the patients with PCOS disease and the other of which indicates that, according to the dataset, 3/4 of the patients have the disease. The dataset's several attributes are used to obtain it, and the f-measure, precision, and recall are used to determine the effectiveness metrics of each of the three algorithms. When contrasted with each of the two exiting methods, the DNN achieves the highest prediction accuracy value of 80.74%. Additionally, the DNN approach achieves the highest f-measure value, recall, and precision. It is verified that, in terms of accuracy, the DNN technique predicts a very high prevalence of PCOS disease, whereas the RNN and CNN techniques predict a lower prevalence when compared to the DNN method. Without the requirement for explicit annotations, the DNN can identify entities of different sorts on different data genres. In the Future the researcher has ample opportunity to use other prediction algorithms in subsequent work to increase forecast accuracy.

REFERENCES

[1] Aroni Saha Prapty and Tanzim Tamanna Shitu, "An efficient decision tree establishment and performance analysis with different machine learning approaches on polycystic ovary syndrome", 23rd International Conference on Computer and Information Technology (ICCIT), pp. 1–5, 2020. doi:10.1109/ ICCIT51783.2020.9392666.

[2] Amsy Denny, Anita Raj, Ashi Ashok, C Maneesh Ram, and Remya George, "i-hope : Detection and prediction system for polycystic ovary syndrome (pcos) using machine learning techniques", TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), pp. 673–678, 2019. doi:10.1109/ TENCON.2019.8929674.

[3] Palvi Soni and Sheveta Vashisht, "Exploration on polycystic ovarian syndrome and data mining techniques", 3rd International Conference on Communication and Electronics Systems (ICCES), pp. 816–820, 2018. doi:10.1109/CESYS.2018.8724087.

[4] Bharati S., Podder P., Mondal M.R.H., "Diagnosis of polycystic ovary syndrome using machine learning algorithms", Proceedings of the 2020 IEEE Region 10 Symposium (TENSYMP), Dhaka, Bangladesh, pp. 1486–1489, 5–7 June 2020.

[5] N. P. Jouppi, C. Young, N. Patil, D. Patterson, G. Agrawal et al., "In- Datacenter Performance Analysis of a Tensor Processing Unit TM," pp. 1–17, 2017.

[6] Tiwari S., Kane L., Koundal D., Jain A., Alhudhaif A., Polat K., Zaguia A., Alenezi F., Althubiti S.A., "SPOSDS: A smart Polycystic Ovary Syndrome diagnostic system using machine learning", Expert Syst. Appl., 2022, doi: 10.1016/ j.eswa.2022.117592

[7] Anda D., Iyamah E. Comparative Analysis of Artificial Intelligence in the Diagnosis of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Polycystic Ovary Syndrome. [(accessed on 17 March 2023)]. Available online: https://www.researchgate.net /publication/366320486_Comparative_Analysis_of_Artificial_Intelligence_in_the_Diagnosis_of_Polycystic _ Ovary_Syndrome.

[8] Bhardwaj P., Tiwari P., "Manoeuvre of Machine Learning Algorithms in Healthcare Sector with Application to Polycystic Ovarian Syndrome Diagnosis", Proceedings of Academia-Industry Consortium for Data Science: AICDS 2020. Springer New York, pp. 71–84, 2022.

[9] Adla Y.A.A., Raydan D.G., Charaf M.Z.J., Saad R.A., Nasreddine J., Diab M.O., "Automated detection of polycystic ovary syndrome using machine learning techniques", Proceedings of the 2021 Sixth International Conference on Advances in Biomedical Engineering (ICABME), Werdanyeh, Lebanon. pp. 208–212, 7–9 October 2021.

[10] Palak Mehrotra, Jyotirmoy, Chatterjee, Chandan Chakraborty, "Automated Screening of Polycystic Ovary Syndrome using Machine Learning Techniques", IEEE, 2012.

[11] N. Shazeer, A. Mirhoseini, K. Maziarz, A. Davis, Q. Le et al., "Outrageously large neural networks: The sparsely-gated mixture-of-experts layer", arXiv preprint, arXiv:1701.06538, 2017.

[12] Subrato Bharati, Prajoy Podder, M. Rubaiyat Hossain Mondal, "Diagnosis of Polycystic Ovary Syndrome Using Machine Learning Algorithms", IEEE Region 10 Symposium (TENSYMP), 5-7 June 2020, Dhaka, Bangladesh.

[13] Ning-Ning Xie, Fang-Fang Wang, Jue Zhou, Chang Liu, Fan Qu, "Establishment and Analysis of a Combined Diagnostic Model of Polycystic Ovary Syndrome with Random Forest and Artificial Neural Network", Hindawi BioMed Research International Volume 2020.

[14] Priyanka R. Lele, Anuradha D. Thakare, "Comparative Analysis of Classifiers for Polycystic Ovary Syndrome Detection using Various Statistical Measures", International Journal of Engineering Research & Technology (IJERT), Volume 9(3), March-2020, ISSN: 2278-0181.

[15] Namrata Tanwani, "Detecting PCOS using Machine Learning", IJMTES | International Journal of Modern Trends in Engineering and Science, Volume 7(1), 2020, ISSN: 2348-3121.

[16] Pijush Dutta, Shobhandeb Paul, Madhurima Majum-der, "An Efficient SMOTE Based Machine Learning classification for Prediction & Detection of PCOS", Research Square, November 8th, 2021.

[17] Muhammad Sakib Khan Inan, Rubaiath E Ulfath, Fahim Irfan Alam, Fateha Khanam Bappee, Rizwan Hasan, "Improved Sampling and Feature Selection to Support Extreme Gradient Boosting for PCOS Diagnosis".

[18] Mutinda, James, Waweru Mwangi, and George Okeyo, "Sentiment analysis of text reviews using lexicon-enhanced bert embedding (LeBERT) model with convolutional neural network", Applied Sciences, Volume 13(3), pp.1445, 2023.

[19] S. Vijayarani, J Ilamathi and Nithya, "Preprocessing Techniques for Text Mining- An Overview," International Journal of Computer Science & Communication Networks, Volume 5, pp. 7-16, 2015.

[20] S. Han, H. Mao, and W. J. Dally, "Deep Compression – Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding," Iclr, pp. 1–13, 2016.

[21] S. Wu, G. Li, F. Chen, and L. Shi, "Training and inference with integers in deep neural networks," arXiv preprint arXiv:1802.04680, 2018.

[22] E. Chung, J. Fowers, K. Ovtcharov, M. Papamichael, A. Caulfield et al., "Serving dnns in real time at datacenter scale with project brainwave," IEEE Micro, volume 38(2), pp. 8–20, 2018.

[23] S. Hashemi, N. Anthony, H. Tann, R. Bahar, and S. Reda, "Understanding the impact of precision quantization on the accuracy and energy of neural networks," Proceedings of the Conference on Design, Automation & Test in Europe. European Design and Automation Association, pp. 1478–1483, 2017.

[24] N. Shazeer, A. Mirhoseini, K. Maziarz, A. Davis, Q. Le et al., "Outrageously large neural networks: The sparsely-gated mixture-of-experts layer", arXiv preprint arXiv:1701.06538, 2017.

[25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Advances in Neural Information Processing Systems, Volume 25, 2012.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[26] T. Vaiyapuri, A. K. Dutta, and I. S. Punithavathi, "Intelligent deep-learning-enabled decision-making medical system for pancreatic tumor classification on ct images", Healthcare, Volume 10(4), pp. 677, 2022.
[27] Zheng Weifa, "Research on intrusion detection Algorithm based on CNN-LSTM Hybrid Model", Network Security Technology and Application, pp. 61-64, 2020.
[28] Liu Yuefeng, et al., "Network intrusion detection method combining CNN and BiLSTM", Computer Engineering, Volume 45(12), pp.127-133, 2020.
[29] Li Wenhui, ZHANG Yingjun, and Pan Lihu, "Aclassification method for improving biLSTM network", Computer Engineering and Design, Volume 41(3): P.880-886, 2020.
[30] B. Reagen, P. Whatmough, R. Adolf, S. Rama, H. Lee et al., "Minerva: Enabling low-power, highly-accurate deep neural network accelerators", Proceedings of the 43rd International Symposium on Computer Architecture, IEEE Press, pp. 267–278, 2016.
[31] A. Mishra and D. Marr, "Wrpn& apprentice: Methods for training and inference using low-precisionnumerics", arXiv preprint arXiv:1803.00227, 2018.
[32] S. Han, H. Mao, and W. J. Dally, "Deep Compression – Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding" Iclr, pp. 1–13, 2016.
[33] S. Wu, G. Li, F. Chen, and L. Shi, "Training and inference with integers in deep neural networks", arXiv preprint, arXiv:1802.04680, 2018.
[34] Altman, E.I., 1968. "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", The journal of finance, Volume 23(4), pp. 589-609.
[35] Hamori, S., Kawai, M., Kume, T., Murakami, Y. and Watanabe, C., "Ensemble Learning or Deep Learning? Application to Default Risk Analysis", Journal of Risk and Financial Management, Volume 11(1), pp. 12, 2018.
[36] Hinton, G.E. and Salakhutdinov, R.R., "Reducing the dimensionality of data with neural networks. Science", 313(5786), pp. 504-507, 2006.
[37] Larochelle, H., Mandel, M., Pascanu, R. and Bengio, Y., "Learning algorithms for the classification restricted Boltzmann machine", Journal of Machine Learning Research, Volume 13, pp. 643-669, 2012.
[38] Mazumder, R., Hastie, T. and Tibshirani, R., "Spectral regularization algorithms for learning large incomplete matrices", Journal of machine learning research, Volume 11, pp. 2287-2322, 2010.
[39] Luo, C., Wu, D. and Wu, D., "A deep learning approach for credit scoring using credit default swaps", Engineering Applications of Artificial Intelligence, Volume 65, pp. 465-470, 2017.
[40] Tomczak, J.M. and Zięba, M., "Classification Restricted Boltzmann Machine for comprehensible credit scoring model", Expert Systems with Applications, Volume 42(4), pp. 1789-1796, 2015.
[41] Rousseeuw, P.J. and Driessen, K.V., "A fast algorithm for the minimum covariance determinant estimator", Technometrics, Volume 41(3), pp. 212-223, 1999.
[42] West, D., "Neural network credit scoring models", Computers & Operations Research, Volume 27(11-12), pp. 1131-1152, 2000.
[43] S. H. F. Langroudi, T. Pandit, and D. Kudithipudi, "Deep learning inference on embedded devices: Fixed-point vs posit", 1st Workshop on Energy Efficient Machine Learning and Cognitive Computing for Embedded Applications (EMC2), pp. 19–23, March 2018.