NVEO
Natural Volatiles & Essential Oils

Detection And Prediction System For Polycystic Ovary Syndrome Using Structural Normalized Square Similarity Detection Approach

[1]Dr.L.Thara , [2]T.M. Divya

[1]Associate Professor, Department of Computer science PSG College of Arts & science Coimbatore, India

[2]Research scholar, Department of Computer science PSG College of Arts & science Coimbatore, India

Abstract: Poly Cystic Ovary Syndrome (PCOS) is the most common endocrine disorders affected to female in their reproductive cycle. It seriously affects women's health. It is observed that PCOS is a condition seen among the women of reproductive age is having a major influence in the cause of infertility. Over five million women worldwide in their reproductive age got affected by PCOS. It is an endocrine disorder characterized by changes in the female hormone levels and the abnormal production of male hormones. This condition leads to ovarian dysfunction with increased risk of miscarriage and infertility. The symptoms of PCOS include obesity, irregular menstrual cycle and excessive production of male hormone, acne and hirsutism. It is extremely difficult to diagnose PCOS due to the heterogeneity of symptoms associated and the presence of a varying number of associated gynaecological disorders. The objective is detecting the follicles using object growing method. It consists of two major stages including pre-processing phase and follicle identification based on object growing method. To address this problem, this paper proposes a system for the early detection and prediction of PCOS, by applying the Structural Normalized Square Similarity Detection Approach (SNSSDA).

Keywords: Medical data, Polycystic Ovary Syndrome, Prediction, SNSSDA, ANN, feature extraction.

I. INTRODUCTION

PCOS (Polycystic Ovary Syndrome) is the disorder occurs in female's reproductive stage identified by formation off ollicular cysts in the ovary. These follicular cysts are observed in ultrasound image which is obtained by scanning the ovary. Prevalence of PCOS in India ranges from 3.7 to 22.5 percentage depending on the population studied and the criteria used for diagnosis. Abnormalities inleptin-adiponectin (adipocyte biology), oxidative stress and autoimmunity are among the mechanisms studied regarding pathogenesis of PCOS [1]. The main cause of this disorder in females is due to menstrual problems, hirsutism, endocrine abnormalities, acne, obesity etc [2].

The detection of ovarian follicle is done using ultrasoundimages of ovaries. Object recognition in an ultrasound image is a challenging task which includes the detection offollicles in ovary, growth of the foetus, monitoring of proper development of the foetus and presence of tumor [3].Nowadays, the diagnosis performed by doctors is to manually count the number of follicular cysts in the ovary, which is used to judge whether PCOS exists or not. This manual counting may lead to problems of variability, reproducibilityand low efficiency. Automating this mechanism will resolve these problems. Over five million women worldwide in their reproductive age PCOS. It is an endocrine disorder characterized by changes in the female hormone levels and the abnormal production of male hormones. This condition leads to ovarian dysfunction with increased risk of miscarriage and infertility.

PCOS is diagnosed by gynaecologists and good and better understanding required analysing the disorder with a bunch of symptoms. The calculation of PCOS prevalence depends on the parameters used to determine this syndrome. Because the PCOS symptoms emerge surreptitiously and correlate with normal puberty progression, ingenious features may not be examined in the early stages. This may be due to the inability to identify the young girl's disorder[4].

The symptoms of PCOS include obesity, irregular menstrual cycle, and excessive production of male hormone, acne, and hirsutism. It is extremely difficult to diagnose PCOS due to the heterogeneity of symptoms associated and the presence of a varying number of associated gynaecological disorders. The time and cost involved in in numerous clinical tests and ovary scanning has become a burden to the patients with PCOS. PCOS is diagnosed by exclusion of irrelevant symptoms or test results, mainly because of lack of knowledge of its complex patho-mechanism[5].

The diverse symptoms of this condition force medical practitioners to call for large number of clinical test results and unnecessary radiological imaging procedures. The early detection and diagnosis of PCOS with minimal tests and imaging procedures is of utmost importance and of great significance as the condition directly leads to ovarian dysfunction with an increased risk of miscarriage, infertility or even gynaecological cancer and mental agony for the patients due to wastage of time and money[6].

To address this problem this paper proposes system for the early detection and prediction of PCOS from an optimal and minimal but promising clinical and metabolic parameters,which act as an early marker for this disease. The objective of the present work is to find an ovary as normal or PCOS using ultrasoundimage by using SNSSDA -ANN based feature extraction.

II. LITERATURE REVIEW

Gopalakrishnan, C., and M. Iyapparaja[7] Polycystic ovary syndrome (PCOS) disorder is identified by the presence of a number of follicles present in the ovary of female reproductive system. Ultrasound imaging of the ovary contains essential information about the size, number of follicles and its position.

In real time, the detection of PCOS is a difficult task for radiologists due to the various sizes of follicles and is highly connected with blood vessels and tissues. This often results in error diagnosis. For pre-processing various standard filtering techniques are applied on ovary image. Based on the performance, appropriate filter is chosen to remove the noise from the image[16].

This paper presents an effectual active contour with modified Otsu threshold value to automated discovery of follicles from the ultrasound images. The performances of the proposed method illustrate the betterments of the proposed approach over other techniques[14].

Deng, Yinhui, Yuanyuan Wang, and YuzhongShen[8] proposed work consists of two major functional blocks: pre-processing phase and follicle identification based on object growing. In the pre-processing phase, speckle noise in the input image is removed by an adaptive morphological filter, then contours of objects are extracted using an enhanced labelled watershed algorithm, and finally the region of interest is automatically selected.

The object growing algorithm for follicle identification first computes a cost map to distinguish between the ovary and its external region and assigns each object a cost function based on the cost map. The object growing algorithm initially selects several objects that are likely to be follicles with very high probabilities and dynamically update the set of possible follicles based on their cost functions.

The proposed method was applied to 31 real PCOS ultrasound images obtained from patients and its performance was compared with those of three other methods, including level set method, boundary vector field (BVF) method and the fuzzy support vector machine (FSVM) classifier.

Denny, Amsyet al [9] proposes system for the early detection and prediction of PCOS from optimal and minimal but promising clinical and metabolic parameters, which act as an early marker for this disease. The data sets required for this system development are obtained through patient survey of 541 women during doctor consultations and clinical examinations.

Out of the 23 features from clinical and metabolic test results, 8 potential features are identified using SPSS V 22.0 based on their significance. Classification of PCOS with the feature set transformed with Principal Component Analysis (PCA) is done using various machine learning techniques such as Naïve Bayes classifier method, logistic regression, K-Nearest neighbour (KNN), Classification and Regression Trees (CART), Random Forest Classifier, Support Vector Machine (SVM) in Spyder Python IDE.

Pushkarini H and M A Anusuya[10] proposes a symptom evaluating and monitoring system using which the risk of developing PCOS can be predicted and the patients can be advised to consult a doctor and undergo tests and scanning only when the risk is high[10]. Parameters selected to evaluate the risk of developing PCOS are optimal and low cost.

Training and testing of models are performed on the PCOS dataset using the following features i.e., Testosterone, Hirsutism, Family history, BMI, Fast food, Menstrual disorder, Risk (target). Linear regression, KNN and Random forest models are implemented and evaluated on various performance metrics like R^2, MAE, and RMSE using python 3. Random forest algorithm outperforms the other two algorithms by achieving less error values and highest R^2 value.

Sumathi, M., P. Chitra[11] Analysis recent studies show that about 18% of Indian women suffer from this syndrome. Doctors were manually examining ultrasound images and conclude the affected ovary but unable to find whether it is a simple cyst, PCOS, or cancer cyst. In this paper, CNN based algorithms proposed and coding developed in Python programming for classification of cysts, and they are filled with blood or fluid using ultrasound images.

The study is performed on CNN based image processing feature extraction to classify cysts in the dataset. That is the study is carried out using an independent trained dataset of the same PCOS related diseases. Finally, the test dataset is used for performing the feature extraction process and the results are met with 85% accuracy using performance factors[12].

III. PROPOSED WORK

Polycystic ovary syndrome (PCOS), is one of the relevant, most prevalent hormonal disorder seen among the women of childbearing age. This is a heterogeneous endocrine disorder which is highly prone to infertility, anovulation, cardiovascular disease, type 2 diabetes, obesity etc. PCOS is a common condition detected in nearly 12-21% of women of reproductive age and among them 70% is remain undiagnosed. PCOS condition can be treated to some extend by controlled medication and bringing alterations in life style. This includes the treatment methods with pills for birth control, diabetes, fertility, anti-androgen medicines and scanning procedures like ultrasound scan. PCOS detection is still operated manually by a gynaecologist by counting the number and size of follicles in the ovaries, so it takes a long time and needs high accuracy. The process is done by adding image data, pre-processing and segmenting, to remove unwanted data and to detect the disease with high accuracy.

In general, PCOS can be detected by calculating stereology or feature extraction and classification. This work designed a system to classify PCOS by using the feature extraction (SNSSDA method) and ANN. Initially the image pre-processing is done such as resizing and noise removal. After that, segmentation method is developed for follicle detection in ultrasound images using Adaptive k-means clustering algorithms.

After segmentation, the SNSSDA based feature extraction and further ANN based classifier is used to identify the follicles in the ovary image. The proposed algorithm is tested on sample ultrasound images of ovaries for identification of follicles and with the ANN, the ovaries are classified into two, normal ovary and abnormal ovary.

A. Image pre-processing

In this phase pre-processing is done by using image resizing and noise removal. Image Resizing using object carving an enhancement is designed using Object carving. In results, for minimizing salience distortions and artifacts, this enhancement will coordinate with other resizing algorithms. For image resizing, interactive performance is preferred by the users. Substantially faster performance is shown by this framework, when compared with original Multi-Op techniques and better results can be produced using this. If there is a need to resize both the dimensions, image height and width should be handled separately in this algorithm and there will be a resizing of longer dimension in the initial stages.In this phase, noise removal is performed using de-noising process; input image corresponds to resized US images. A statistics based nonlinear signal processing technique is median filter.

B. Feature Extraction

Feature extraction is the process to identify main parameters or attributes of the data. The basic idea behind feature selection is to build a subset of significant attribute which is used to train the classifier and modify its coherency. Feature extraction help us to convert image to numerical parameters we extract features like GLCM, Texture, Geometric parameters like area, perimeter, compactness, smoothness, entropy of segmented image and train table is created for dataset. GLCM is featured by two dimensionless histograms for each image pixel pair using gray levels which are divided by conceptual relationship.

The texture differentiating problem is to examine a set of concurrent matrices. It consists of a statistical method which involves an indicator of x rows and y columns. The indexed data depicts the given range of the gray level image content and the value of P(x,y) in a particular position shows the frequency of gray levels x and y have particular distance and direction.

C. Classification

Classification is done using ANN algorithm. It is one of the supervised classification algorithms. The idea in ANN is to search the closest match of the test data in feature space. It works on a basis, assuming every data point falling in near to each other will be in same class 22, 23. Here classification is done using various parameters of feature extraction along with comparing trained dataset.

ANN with Structural normalized square similarity detection approach is presented by using Structural normalized similarity to compute the similarity between test samples with neural network samples. It uses a Similarity judgment algorithm and calculates the similarity value for each training sample. Then it calculates the maximum between these values.

Algorithm: SNSSDA-ANN

Step 1: Image pre-processing of US image data.
Step 2: Input the image data for training, the interrelated values of input and output execute for training using feed forward back propagation neural network algorithms.
Step 3: Initialize weights and biases plays an important role in determining the final model of follicle detection

Step 4: Calculate the neurons of output, every neuron output data calculated using $on_j = \sum_{i=1-m} w_{ij} x_i + b_j$

where onj and wij are output neurons and connection weight neurons, respectively, while xi and bj are theinput data, and bias neurons.

Step 5:Data of output layers' calculation using,

$$on_k = TV_k + \delta_K^L$$

where TVk is target value of output neurons and δ_L^K is the error of neuron.

Step 6: Compute the error E of neuron k and step 3 and step 6 were repetitive until network was congregated, and the error was computed using,

$$E_j = on_j(1 - on_j) \sum_k E_k w_{jk}$$

The proposed model is primarily comprised of two procedures: the inner parameter optimization and the outer classification performance evaluation. During the innerparameter optimization procedure, the ANN parameters are dynamically adjusted by the IFFOA technique. Then, the obtained optimal parameters are fed to the ANN prediction model to perform the classification task for folliclediagnosis.

IV. EXPERIMENTAL RESULTS

This section shows performance measures and the results are compared to find the best model for PCOS risk prediction. The quality of an image is important when processing medical ultrasound images. The visual perception is improved in the de-noised images which makes the diagnosis easier.

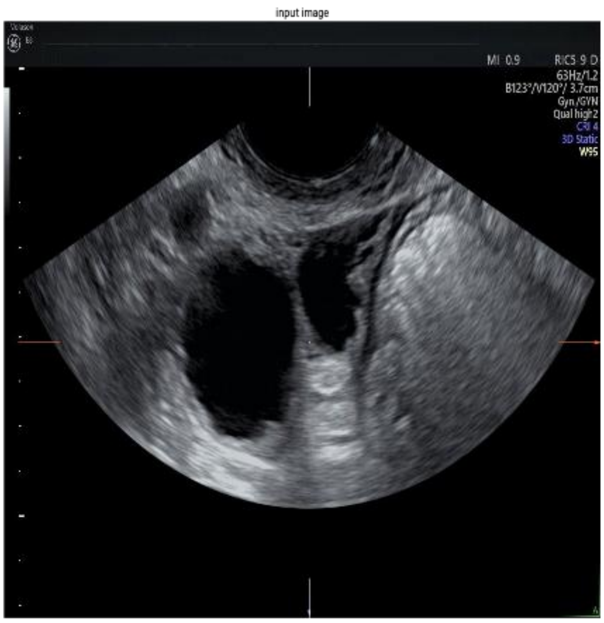
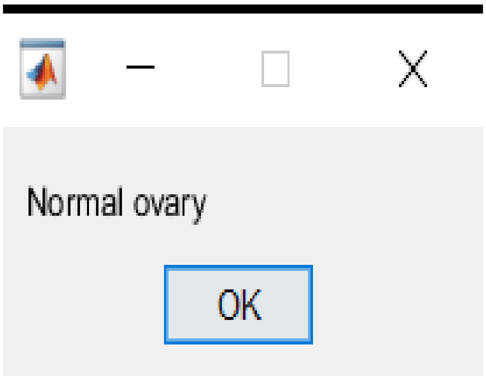
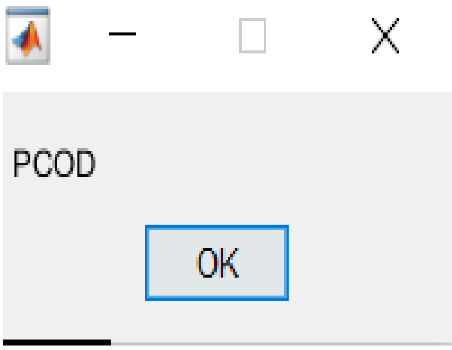


Fig 1: Input image

Fig 2: Filtered Image



Naive Bayes result



ANN result

Table 1: Training and testing data using NB

Training data Validation parameters NB classifier	Testing data Validation parameters for NB classifier
Accuracy 7.241379e+01 precision 8.947368e+01 recall 7.391304e+01 Fmeasure 8.095238e+01 Specificity 6.666667e+01	Accuracy 5.454545e+01 precision 60 recall 8.571429e+01 Fmeasure 7.058824e+01 Specificity 0

Table 2: Training and testing data using ANN

Training data Validation parameters for ANN classifier	Testing data Validation parameters for ANN classifier
Accuracy 100 precision 100 recall 100 Fmeasure 100 Specificity 100	Accuracy 6.363636e+01 precision 6.666667e+01 recall 8.571429e+01 Fmeasure 75 Specificity 25

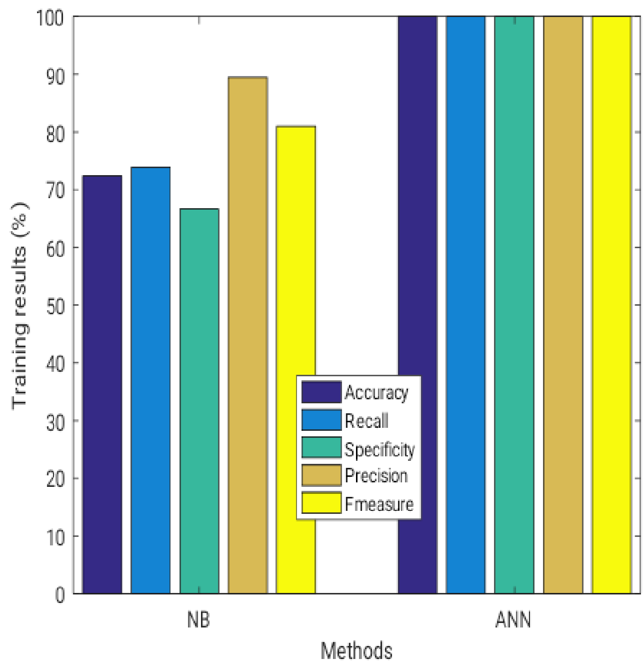


Chart 1: Training result

Chart 1 shows result obtained for classification of data of the trained data set.

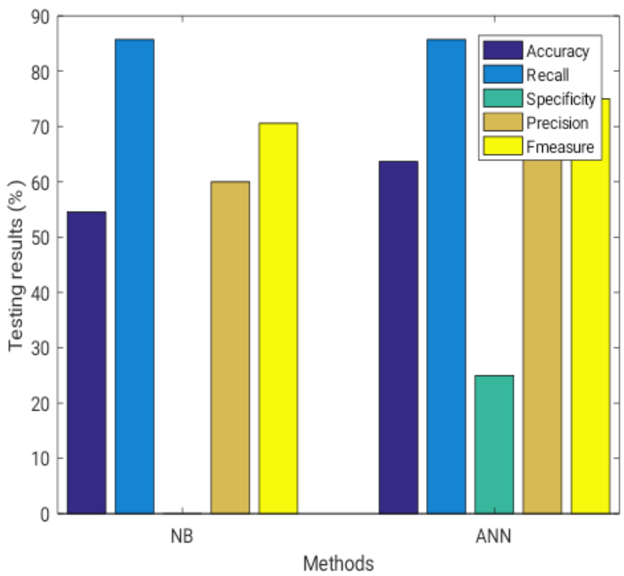


Chart 2: Testing Result

Chart 2 shows result obtained for classification of data of the testing data set

Recall: Recall quantifies the number of positive class predictions made out of all positive examples in the image dataset as given below

$$\mathrm{RECALL} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

Precision: Precision quantifies the number of positive class predictions that actually belong to the positive class and it is estimated as follows

$$\mathrm{PRECISION} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

F-measure: F-Measure provides a single score that balances both the concerns of precision and recall in one number and it is estimated as follows:

$$\mathrm{F} - \mathrm{MEASURE} = \frac{(2 * \mathrm{PRECISION} * \mathrm{RECALL})}{\mathrm{PRECISION} + \mathrm{RECALL}}$$

Accuracy: It is one of the most commonly used measures for the classification performance, and it is defined as a ratio between the correctly segmented samples to the total number of samples as follows

$$\mathrm{ACCURACY} \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

The experimentalresults are in good agreement with the manual detection of the ovarian classes by medical experts and, thus,demonstrate efficacy of the proposed method. The ovarian classification performance is 97.5% using our proposed classifier. Hence, the proposed method serves as the effective basis classification of ovariesduring entire female cycle. It helps to study the ovarian morphology of the patients and also significantly improvesthe quality of diagnosis and treatment of patients.

V. CONCLUSION

Polycystic ovary syndrome (PCOS) disorder is identified by the presence of a number offollicles present in the ovary of female reproductive system. Ultrasound imaging of the ovary contains essential information about the size, number of follicles and its position. In real time, the detection of PCOS is a difficult task for radiologists due to the various sizes of follicles andis highly connected with blood vessels and tissues. This often results in error diagnosis. Forpre-processing various standard filtering techniques are applied on ovary image, based on the performance. However, few studies have tried to develop a diagnostic model based on gene biomarkers. In this paper proposes a Structural normalized square similarity detection approach (SNSSDA) to predict the PCOS problem. The performances of the proposed method, illustrate the betterments of the proposed approach over other techniques.

VI. REFERENCES

[1]. Purnama, Bedy, Untari Novia Wisesti, Fhira Nhita, Andini Gayatri, and Titik Mutiah, "A classification of polycystic Ovary Syndrome based on follicle detection of ultrasound images." In 2015 3rd International Conference on Information and Communication Technology (ICoICT), pp. 396-401. IEEE, 2015.

[2]. A.S. Ambekar, D. S. Kelkar, S. M. Pinto et al., “Proteomics of follicular fluid from women with polycystic ovary syndrome suggests molecular defects in follicular development,” The Journal of Clinical Endocrinology & Metabolism, vol. 100, no. 2, pp. 744–753, 2015.

[3]. H. Liu, L. Zeng, K. Yang, and G. Zhang, “A network pharmacology approach to explore the pharmacological mechanism of xiaoyao powder on anovulatory infertility,” Evidence-based Complementary and Alternative Medicine: Ecam, vol. 2016, article 2960372, 13 pages, 2016.

[4]. C. Lu, X. Liu, L. Wang et al., “Integrated analyses for genetic markers of polycystic ovary syndrome with 9 case-control studies of gene expression profiles,” Oncotarget, vol. 8, no. 2, pp. 3170–3180, 2017.

[5]. T. S. Domingues, T. C. Bonetti, D. C. Pimenta et al., “Proteomic profile of follicular fluid from patients with polycystic ovary syndrome (PCOS) submitted to in vitro fertilization (IVF) compared to oocyte donors,” JBRA Assisted Reproduction, vol. 23, no. 4, pp. 367–391, 2019.

[6]. S. Joseph, R. S. Barai, R. Bhujbalrao, and S. Idicula-Thomas, “PCOSKB: A Knowledgebase on genes, diseases, ontology terms and biochemical pathways associated with Polycystic Ovary Syndrome,” Nucleic Acids Research, vol. 44, no. D1, pp. D1032–D1035, 2016.

[7]. Gopala krishnan, C., and M. Iyapparaja, "Active contour with modified Otsu method for automatic detection of polycystic ovary syndrome from ultrasound image of ovary." Multimedia Tools and Applications 79, no. 23 (2020): 17169-17192.

[8]. Deng, Yinhui, Yuanyuan Wang, and Yuzhong Shen, "An automated diagnostic system of polycystic ovary syndrome based on object growing." Artificial intelligence in medicine 51, no. 3 (2011): 199-209.

[9]. Denny, Amsy, 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." IEEE 2019.

[10]. Pushkarini H and M A Anusuya, “A Prediction Model for Evaluating the Risk of Developing PCOS”, IRJET, 2020

[11]. Sumathi, M., P. Chitra, R. SakthiPrabha, and K. Srilatha. "Study and detection of PCOS related diseases using CNN." In IOP Conference Series: Materials Science and Engineering, vol. 1070, no. 1, p. 012062. IOP Publishing, 2021.

[12]. C.-H. Ho, C.-M. Chang, H.-Y. Li, H.-Y. Shen, F.-K. Lieu, and P. S.-G. Wang, “Dysregulated immunological and metabolic functions discovered by a polygenic integrative analysis for PCOS,” Reproductive Bio Medicine Online, vol. 40, no. 1, pp. 160–167, 2020.

[13]. X. Z. Zhang, Y. L. Pang, X. Wang, and Y. H. Li, “Computational characterization and identification of human polycystic ovary syndrome genes,” Scientific Reports, vol. 8, no. 1, article 12949, 2018.

[14]. J. J. Cheng and S. Mahalingaiah, "Data mining polycystic ovary morphology in electronic medical record ultrasound reports," Fertility Research and Practice, vol. 5, no. 1, p. 13, 2019.

[15]. B. Vikas, B. Anuhya, K. S. Bhargav, S. Sarangi, and M. Chilla, "Application of the apriori algorithm for prediction of Polycystic Ovarian Syndrome (PCOS)," in Information Systems Design and Intelligent Applications, pp. 934–944, Springer, 2018.

[16]. D. K. Meena, D. M. Manimekalai, and S. Rethinavalli, "A novel framework for filtering the PCOS attributes using data mining techniques," International Journal of Engineering Research & Technology, vol. 4, no. 1, pp. 702–706, 2015.

[17]. B. Vikas, B. Anuhya, K. S. Bhargav, S. Sarangi, and M. Chilla, "Application of the apriori algorithm for prediction of Polycystic Ovarian Syndrome (PCOS)," in Information Systems Design and Intelligent Applications, pp. 934–944, Springer, 2018.