Grenze International Journal of Engineering and Technology, June Issue

Exploratory Analysis on Personalized Health Assistance during PCOS using Machine Learning Techniques

Vadlana Baby[1], Anjusha Pimpalshende[2], Jahnavi Nelluri[3], Cheenuri Preetham[4], Keerthana Eddi[5] and Bayarapu Hema Sri[6]
[1-6]VNR Vignana Jyothi Institute of Engineering & Technology/Department of CSE, Hyderabad, India
Email: baby_v@vnrvjiet.in, {anjusha_p@vnrvjiet.in, jahnavinelluri9@gmail.com, cheenuripreetham16@gmail.com, ekeerthana@gmail.com, hemasri.bayarapu@gmail.com}

***Abstract*— This initiative aims to tackle the complex health challenges presented by Polycystic Ovary Syndrome (PCOS) and Polycystic Ovary Disorder (PCOD) that impact women's wellbeing globally. It provides personalized diet and fitness plans tailored to individual symptoms using machine learning technology. This empowers women to better manage their health, alleviate symptoms, and restore hormonal balance. Additionally, it nurtures a supportive community for networking, sharing experiences, and accessing educational materials to dispel myths and spread awareness of PCOS/PCOD. The overarching goal is to prioritize women's health, circulate knowledge, and drive progress in sustaining healthy lifestyles to better cope with these conditions - working towards a healthier future for women worldwide.**

***Index Terms*— PCOS, dietary recommendation and exercise recommendations.**

I. INTRODUCTION

Polycystic Ovary Syndrome (PCOS) is a common hormonal disorder that predominantly affects women globally. It is a complex condition with a multifactorial origin involving genetic, lifestyle, and environmental factors. Hallmarks of PCOS include hormonal imbalances, menstrual irregularities, ovarian cysts, and associated metabolic complications. The array of signs and symptoms such as infrequent/prolonged periods, excess androgen levels causing hirsutism and acne, challenges with fertility and weight management, collectively diminish one’s quality of life. While the underlying pathogenesis remains unclear, PCOS heightens risks of developing cardiovascular and metabolic diseases like diabetes and obesity over time. Effectively managing PCOS necessitates a multifaceted care plan, incorporating medications, lifestyle modifications, as well as psychosocial support. Timely diagnosis and holistic management of PCOS can significantly mitigate adverse health outcomes in the long run, in addition to improving one’s self-esteem and well-being.

II. RELATED WORK

Yash Rathod, Aryan Komare, Ruchita Ajgaonkar, Shruti Chindarkar, Dr. Gajanan Nagare, Prof. Neelam Punjabi, and Dr. Yogesh Karpate[7]. The core goal was to develop a machine learning model to predict one's risk of having PCOS using various parameters as inputs. Multiple algorithms were tested, including logistic regression, KNN, decision trees, support vector machines, CatBoost, random forest classifier, and Naive Bayes. Of all the models, CatBoost - which uses an ensemble of decision trees boosted by gradient descent - achieved the highest and most

Grenze ID: 01.GIJET.10.2.100_1

reliable accuracy in determining if a woman should seek medical consultation for possible PCOS. By leveraging the predictive capabilities of CatBoost, the model can assess a woman's likelihood of having undiagnosed PCOS based on her input metrics. This allows early screening and intervention for the condition, thereby improving patient outcomes. In summary, CatBoost was discovered to be the optimal ML approach for reliable PCOS risk stratification among the techniques explored.

Palak Mehrotra, Jyotirmoy Chatterjee, Chandan, Biswanath Ghoshdastidar, and Sudarshan Ghoshdastidar[5]. The key objective was to develop an automated method for diagnosing PCOS based on established clinical and metabolic markers that act as early indicators of the condition. The proposed approach involves constructing a feature vector of discriminative parameters between healthy and PCOS cohorts, selected via two-sample t-tests. This PCOS signature comprised features spanning clinical and metabolic domains. Subsequently, Bayesian and Logistic Regression models were implemented to classify subjects based on the defined feature set. Comparative evaluation revealed superior performance of the Bayesian classifier over logistic regression for PCOS detection. The Bayesian model achieved an overall accuracy of 93.93% in differentiating PCOS cases from healthy controls. In summary, by identifying a candidate PCOS biomarker subset coupled with Bayesian learning, the study successfully demonstrated accurate and automated detection of PCOS patients versus non-PCOS individuals.

Nusrat, Sumiya Islam, Sharun Akter Khushbu, and Abu Kaisar Mohammad Masum[3]. The core objective was to detect PCOS and determine the optimal machine learning approach by benchmarking various algorithms using real-world datasets. Numerous models were evaluated including Random Forest, Naive Bayes, Decision Trees, Support Vector Machines, K-Nearest Neighbors, Logistic Regression, XGBoost Classifier and Gradient Boosting Classifier. Among these, Support Vector Machines attained the highest accuracy of 99.09% for PCOS classification - outperforming alternate techniques. This demonstrates the efficacy of Support Vector Machines for reliable PCOS diagnosis based on patient data. Future work entails assimilating diverse PCOS datasets across different geographic regions to uncover location-specific nuances and variabilities in PCOS manifestations globally. In summary, this comparative assessment revealed Support Vector Machines as a superior ML solution for detecting PCOS, setting the stage for generalized models incorporating multi-regional PCOS data worldwide.

P. Chitra, M. Sumathi, K. Srilatha, F.V. Jayasudha, and S.Amudha[5]. The key objective was to develop an artificial neural network (ANN) model to predict polycystic ovary syndrome (PCOS). To handle the imbalanced dataset, techniques like oversampling minority data, undersampling majority data, SMOTE, and ensemble methods were integrated. Informative features were extracted and used to train the ANN classifier for PCOS prediction. A key benefit of the ANN is its ability to continuously learn from historical inputs, progressively enhancing accuracy. The proposed ANN approach achieved 86% accuracy for PCOS prediction. Future work includes refining the neural network architecture based on dataset size, augmenting dropout techniques like L1 and L2 regularization to prevent overfitting, and exploring additional algorithms to mitigate data imbalance - collectively aimed at improving model performance. In essence, the study demonstrated promising accuracy of ANNs for automated PCOS prediction, forming a baseline for advanced ANN models tailored to handle intricacies of PCOS data.

Jain and Singhal[2]. The overarching goal was to build a system to provide personalized dietary recommendations based on an individual's age, height, weight and other attributes. It reviews existing literature in this domain and critically appraises the value add of the proposed methodology. In essence, the system employs machine learning to offer tailored meal plans aligned with users' specific nutritional goals - whether targeting weight management, muscle gain, or general wellness. Numerous models including Random Forests, Support Vector Machines, AdaBoost, Gradient Boosting and clustering algorithms like K-means were implemented. To predict optimal dietary choices for an individual's desired health outcome, classifiers such as Random Forests, SVMs, AdaBoost and Gradient Boosting were trained on nutritional data. Performance was evaluated using metrics of Precision, Recall, Accuracy and F-score. In summary, this work leveraged diverse AI techniques to deliver customized diet plans for varying user needs and objectives.

Arash mahyari, Peter Piroll[2]. The study puts forth a novel exercise recommendation system, combining association rule mining and recurrent neural networks, to predict physical activities each user is most likely to adopt and successfully complete day-to-day. The model is personalized for individual users without needing extensive manual feature engineering, instead learning exercise habits and patterns from user data. The interconnected-RNN architecture outperforms a precedent computational cognition model for predicting exercise adherence probabilities. Rigorous validation is done by testing the model's accuracy in predicting successful task completion on new users. Given its strong predictive performance, the system shows promise in enhancing mobile health interventions to effectively promote physical activity. Additionally, as more user data is accrued, the meaningful patterns extracted via machine learning could inform refining the explanatory computational models themselves, enabling synergistic improvements in both predictive accuracy and explainer capability over time.

K. Srilatha, M. Sumathi, F. V . Jayasudha, T Bernatin[4]. The study investigates predicting polycystic ovary syndrome (PCOS) using transfer learning and hybrid deep learning models to enhance diagnostic accuracy. Specifically, an ensemble approach coupled with data augmentation is proposed, outperforming other techniques on benchmarking with 95% accuracy for PCOS classification. Additionally, the work reviews various transfer learning strategies and hybrid architectures exploited in analogous medical imaging tasks. The findings highlight the potential of advanced machine learning techniques to assist clinicians in making more reliable PCOS diagnosis. As future work, developing a web interface to crowdsource larger PCOS datasets from gynecologists is suggested to further boost model performance. Additionally, creating an automated and intelligent PCOS screening solution for deployment via web-services is envisioned. In summary, this research demonstrates promising accuracy improvements in PCOS diagnosis using transfer learning, underscoring the technique's applicability in developing precise computer-aided screening tools for this condition.

Sumiya Islam, Nusrat Nabi, Sharun Akter Khushbu, Nushrat Jahan Ria[4]. The study analyzes survey data on Polycystic Ovary Syndrome (PCOS) to better characterize this prevalent hormonal condition in women. Using data mining techniques, the authors examined commonly reported symptoms and risk factors attributed to PCOS pathogenesis. Their analysis reinforces the complex, multifactorial nature of PCOS with underpinnings of genetic, lifestyle and environmental influences. The cross-country dataset compilation and subsequent analyses enabled distinguishing PCOS from normal cohorts based on identified phenotypic patterns. Key revelatory symptoms were recognized which could aid targeted screening. By disseminating these PCOS "signatures", the research aims to improve awareness and counseling for lifestyle modifications to help affected women better manage associated symptoms. Overall, the data-driven approach adopted provides actionable insights on the hallmarks of PCOS while disambiguating its intricate heterogeneity across geographies. The findings lay the groundwork for improving detection and personalized management strategies for PCOS globally.

Megh Shah, Sheshang Degadwala, Dhairya Vyas[3]. The research presents a personalized diet recommendation system using machine learning for tailored nutritional guidance. It incorporates user preferences, health objectives and medical history to predict optimal food choices that aid in weight goals and general wellbeing. The comparative study reviews the relative merits and limitations of various machine learning techniques. Notably, deep recurrent networks like LSTM are found to outperform conventional classifiers in forecasting personalized diet plans. Evaluations reveal LSTM networks achieve superior accuracy, recall, precision and F1 scores over competing machine learning and deep learning methods. Specifically, the LSTM model demonstrates 97.74% accuracy in recommending individualized meal plans aligned to the user's health priorities. The work thus highlights advanced deep learning, in particular temporal models like LSTM, as a promising approach to deliver data-driven and adaptive nutrition recommendations personalized for individual needs and constraints.

Sharvari S. Deshpande and Asmita Wakankar's[2]. The study implements a multi-stage pipeline for automated detection of Polycystic Ovary Syndrome (PCOS) by analyzing ovarian ultrasound images. Initial preprocessing enhances image quality and filters noise using standard techniques. Canny edge detection segments ovarian follicles, with each follicle characterized by metrics of area, eccentricity and compactness. Clinical lab tests, hormones assays, and imaging features are aggregated into a composite feature vector for classification. A Support Vector Machine model is built to categorize patients as normal or PCOS, using the defined feature set. The approach accomplishes 95% accuracy in identifying PCOS cases, clinically validated for its diagnostic potential. Proposed future work includes advancing image processing, evaluating alternate machine/deep learning models like CNNs, assimilating additional clinical and ultrasound parameters, and expanding sample diversity. Overall, this methodology demonstrates accurate image-based screening of PCOS using optimized imaging combined with SVM learning. Enhancing this pipeline further can enable scalable and automated tools for assisting PCOS diagnosis in clinical settings.

Wan Azani Mustafa, Ahmed Alkhayyat, Hiam Alquran, Waleed Khalid Al-Azzawi, Noor Almajali, and Dana Hdaib[6]. The study utilizes MATLAB and machine learning techniques to develop an accurate diagnostic system for Polycystic Ovary Syndrome (PCOS). Multiple classifiers are benchmarked on a PCOS dataset from Kaggle, prioritizing high accuracy and sensitivity. Comparative analysis with prior work demonstrates enhanced precision and classification performance using the MATLAB platform. The best performing model, Linear Discriminant Analysis, achieved top accuracy while K-Nearest Neighbors maximized sensitivity for PCOS prediction. Future work includes amalgamating preprocessed ultrasound imagery and phenotypic features to further boost diagnostic accuracy. In summary, the integration of MATLAB and machine learning effectively enabled a robust PCOS classification model, outperforming previous efforts. Advancing model sophistication, software user-experience, and clinical collaborations remain imperative next steps towards real-world deployment. Overall, the study validated machine learning in conjunction with MATLAB as a promising approach for improving PCOS diagnostics through data-driven insights.

Shazia Nasim, Mubarak Almutairi, Kashif Munir, Ali Raza, and Faizan Younas[5]. The research harnessed a dataset of Polycystic Ovary Syndrome (PCOS) markers, applying a novel CS-PCOS algorithm for informative feature selection. The data was split into training and test sets to evaluate performance across ten advanced machine learning models, including the Gaussian Naive Bayes (GNB) classifier. Rigorous K-fold cross-validation demonstrated the GNB model achieved outstanding 100% precision and recall - signifying reliable PCOS prediction. Additionally, exploratory data analyses uncovered patterns and risk factors underpinning PCOS onset. The core objective was early detection of PCOS to enable timely treatment and mitigate complications like infertility and miscarriages in affected women. The introduced feature selection methodology, paired with the GNB model, delivered unmatched accuracy and efficiency. Proposed future work involves expanding current data, balancing classes, evaluating deep learning approaches, testing on diverse populations and developing a user-friendly PCOS screening tool to determine long-term health impacts. Overall, the research paved the way for accurate and generalized PCOS predictors to improve women's health through early intervention.

Akhter Hosain, Md Humaion Kabir Mehedi, and Irteza Enan Kabir[3]. The research proposes PCONet - a novel deep learning architecture for detecting Polycystic Ovary Syndrome (PCOS) from ovarian ultrasound images. Leveraging transfer learning, they fine-tuned the pretrained InceptionV3 Convolutional Neural Network (CNN) for comparison with PCONet. Independent testing revealed superior performance of PCONet over InceptionV3, attaining 98.12% accuracy for PCOS classification versus 96.56% respectively. Additionally, confusion matrix visualization distinctly captured the improved sensitivity of PCONet. The principal research objective was developing and demonstrating PCONet's capabilities for accurate image-based screening of PCOS. With higher precision, recall and F1 scores, PCONet outperforms precedent CNN models, showing promising potential for ultrasound-based assistance in PCOS diagnosis. Proposed future work involves boosting accuracy further and generalizability across diverse imaging modalities like MRI, photoacoustic imaging to build robust multi-modal diagnostics. Overall, the study validated deep learning and specifically PCONet as an efficacious solution for automated and non-invasive PCOS detection through ultrasound image analytics.

Tin Trung Tran, Jae-Won Choi, Chien Van Dang, Geon Su Park, Jun Young Baek, and Jong-Wook Kim[6]. The research presents an artificial intelligence-powered workout recommendation system for fitness beginners. It implements an Artificial Neural Network and Logistic Regression to suggest appropriate exercises based on a user's capability. Additionally, a reinforcement learning agent is incorporated within the Soar architecture to aid users in selecting suitable workouts. This intelligent system can analyze and predict personalized exercise plans for both new and existing users - with experiments confirming its efficacy. The core objective was creating an AI-assisted model for adaptive and enhanced workout recommendations catered to varying fitness levels. Future work entails integrating advanced machine learning techniques, supporting diverse training goals, assimilating real-time wearable sensor data, building an interactive fitness community, and providing highly customized workout guidance. Overall, the system marks a promising foray into data-driven and AI-enabled solutions for promoting physical activity through tailored and dynamic workout recommendations to each individual.

S.Sadhasivam, M.S.Sarvesvaran, P.Prasanth, and L. Latha[4]. The research implements a methodology to collect user data through an Android application. K-means clustering is leveraged to group users based on weight-related attributes. Subsequently, Random Forest Classification is utilized to recommend customized diet plans and workout regimes aligned with each user's fitness goals. This integrated approach harnessing clustering, classification and machine learning aims to promote healthy lifestyles. The core objective is developing an AI-enabled recommendation system for nutrition and exercises, accounting for age, weight and desired objectives. By simplifying weight and fitness management via an easy-to-use app, it determines caloric and BMI requirements while emphasizing balanced diet and physical activity for holistic wellness. Future work involves enhancing personalization, incorporating sensor data, expanding the food database, building in educational content, user community features and catering to specific health conditions. Additionally, collaborating with experts can help advance diet and fitness recommendations further. Overall, this research marks an initial step toward AI-powered tools that can provide tailored and evidence-based nutrition and workout guidance to the general public.

Yasmine A. Abu Adla, Dalia G. Raydan, Mohammad-Zafer J. Charaf,Roua A. Saad, Jad Nasreddine, Mohammad O. Diab[6].The goal is to develop a predictive model with high accuracy for detecting PCOS. The proposed approach involves data preprocessing, feature selection using both filters and wrappers to identify an optimal subset, and applying a linear support vector machine (LSVM) classifier. While this methodology aims to maximize precision, it also recognizes the inherent intricacies in establishing such a system. The multi-step process would

benefit from medical expertise to ensure clinical viability and address potential barriers to real-world implementation.
Subrato Bharati, Prajoy Podder, M. Rubaiyat Hossain Mondal[3]. This study aimed to leverage Python-based data science tools for analyzing a dataset of 541 women across 43 attributes to detect polycystic ovary syndrome (PCOS). The analysis workflow first preprocessed the data by encoding categorical variables and selecting relevant features to improve quality. Multiple classification models were then trained on thoughtfully partitioned training and test sets using holdout and cross-validation techniques to differentiate between women with and without PCOS. A key objective was to showcase the potential of computational approaches to uncover insights from healthcare data that facilitate PCOS diagnosis. However, the study also recognized critical data-related challenges inherent to the problem, like class imbalance, and tried to address them through thoughtful data splits and evaluation metrics. Overall, it sought to provide a blueprint for augmenting clinical decision making via data analytics while considering practical limitations.
Sara Alshakrani, Sawsan Hilal,Ahmed M. Zeki[3]. This study evaluates five machine learning models on their ability to accurately detect polycystic ovary syndrome (PCOS). By comparing performance on metrics including precision, recall and accuracy, the goal is to determine the top-performing model that balances classification performance with clinical applicability. While testing multiple models improves result robustness, the analysis notes how data preprocessing steps like feature selection and label encoding could impact insights derived. It also considers practical constraints like computational resource requirements when pursuing the most effective approach. Ultimately, while highlighting predictive modeling as a promising tool for assisting PCOS diagnosis, the study aims to identify the optimal data-driven methodology while weighing real-world viability against core analytical objectives. This keeps the investigation grounded in not just technical accuracy but clinical relevance as well.
Sakthipriya Dhinakaran, Chandrakumar Thangavel, Shivayavashilaxmipriya S,Harinee V S[4]. This study aims to determine the optimal machine learning approach for polycystic ovary syndrome (PCOS) prediction. Five diverse models are built and thoroughly evaluated on classification performance metrics while balancing clinical utility. Though a multifaceted methodology examining accuracy, precision and recall offers robust analysis, the study is cautious to address inherent data and computational challenges. Feature engineering and encodings impact information content, while complex models demand significant resources. By highlighting these practical constraints alongside analytical objectives, the investigation attempts to identify the most effective real-world solution. The goal is not just model performance, but viable and holistic PCOS detection augmentation through prudent evaluation of a spectrum of data-driven methodologies.
Divya Mogaveera, Vedant Mathur, Sagar Waghela[3]. This system aims to provide customized diet and exercise plans based on an individual's health profile and requirements. There are two core modules - Health Monitoring and Diet & Exercise Recommendation. The Health Monitoring module continuously aggregates user-reported metrics and test results to determine health status. It then creates personalized health management plans, including follow-ups. The Diet & Exercise Recommendation module uses machine learning on this collated data to generate suitable diet and workout advice tailored to the user. By monitoring an individual's unique parameters and needs, the system can efficiently adapt and offer support via actionable recommendations. The overarching goal is precise and adaptive guidance to address specific health goals and conditions.

Authors	Objective	Methods/Models	Key Findings/Accuracy
Yash Rathod et al.	Predict risk of PCOS	Logistic Regression, KNN, Decision Trees, SVM, CatBoost, RFC, Naive Bayes	CatBoost showed highest accuracy
Palak Mehrotra et al.	Automate PCOS detection	Bayesian Classifier, LR	Bayesian Classifier: 93.93% accuracy
Nusrat et al.	Compare algorithms for PCOS detection	Random Forest, Naive Bayes, Decision Tree, SVM, KNN, Logistic Regression, XGBoost, Gradient Boosting	SVM: 99.09% accuracy
P. Chitra et al.	PCOS prediction using ANN	Artificial Neural Network	Normal ANN: 86% accuracy
Anjali Jain, Alka Singhal	Personalized meal recommendations	Random Forest, SVM, AdaBoost, Gradient, K-means	Performance measures used: F-measures, Recall, Accuracy, Precision
Arash Mahyari et al.	Exercise recommendation system	Association rules, Recurrent Neural Networks	Improved exercise recommendations; good probability of completion

K. Srilatha et al.	PCOS prediction using transfer learning	Transfer learning techniques, hybrid models	Ensemble technique: 95% accuracy
Sumiya Islam et al.	Survey-based PCOS research	Data mining techniques	Identification of common symptoms and differences between normal and PCOS-affected people
Megh Shah et al.	Personalized diet recommendation	LSTM, traditional ML classifiers	LSTM: 97.74% accuracy
Sharvari S. Deshpande et al.	Automated PCOS detection from ultrasound images	Image processing, SVM	Achieved 95% accuracy
Authors	**Objective**	**Methods/Models**	**Key Findings/Accuracy**
Wan Azani Mustafa et al.	Diagnostic model for PCOS	MATLAB, seven classifiers	Linear Discriminant exhibited top accuracy
Shazia Nasim et al.	Early PCOS prediction	Feature selection, ten ML models	GNB achieved 100% accuracy
Akhter Hosain et al.	PCOS detection using CNN	PCONet (CNN), InceptionV3	PCONet achieved 98.12% accuracy
Tin Trung Tran et al.	Recommending workouts for beginners	ANN, Logistic Regression	Effective workout recommendations
S. Sadhasivam et al.	AI-driven diet and exercise recommendation	K-Means Clustering, Random Forest Classification	BMI calculation and personalized recommendations
Yasmine A. Abu Adla et al.	Predictive model for PCOS detection	LSVM classifier, feature selection	Complex methodology with multiple stages
Subrato Bharati et al.	Data analysis for PCOS identification	Python, data preprocessing, classification algorithms	Addressed challenges related to data quality
Sara Alshakrani et al.	Evaluation of ML models for PCOS detection	Five ML models	Balancing accuracy and clinical relevance
Sakthipriya Dhinakaran et al.	Assessing effective ML models for PCOS detection	Five diverse ML models	Focus on accuracy, precision, and recall
Divya Mogaveera et al.	Personalized diet and exercise recommendations	Health Monitoring, ML-based Diet & Exercise Recommendation	Customized health plans based on user profiles

III. METHODOLOGY

A. RNN

Unsupervised networks called RNNs attempt to reconstruct the input signal. RNNs make decisions based on their recollections of the past and their ability to remember the past. Conventional feed forward networks can also "remember" things, but just the information that they were taught during training. RNNs are based on the sequential information idea, and they can store information about recently measured data.

B. VGG-16 and ResNet-50

Deep convolutional neural networks, such as VGG-16 and ResNet-50, are employed in image recognition. VGG-16 is comparatively easy because it has a homogeneous architecture and 16 weight layers. Conversely, ResNet-50 is a deeper network with residual connections that enables more accurate training. Due to their extensive pre-training on massive picture datasets, both models are highly accurate in identifying PCOS and extracting characteristics.

C.LSTM

Long Short-Term Memory (LSTM) models are a type of recurrent neural network adept at preserving important past information. This is enabled by an LSTM's unique structure consisting of a memory cell regulated by three gates - forget, input and output. The forget gate removes irrelevant memories, the input gate updates the cell state, and the output gate determines what is transferred to the next node. In an exercise recommendation application, an LSTM-based system would store a user's workout history and use that to predict the next best activity. By identifying valuable exercise patterns from past data, the model can continually output personalized and achievable suggestions that users are likely to complete. The LSTM's intrinsic memory and gating mechanisms allow intelligent tracking of an individual's progress and preferences for tailored exercise prescription.

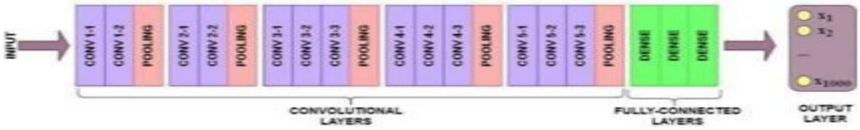

Figure 1: Architecture of VGG-16

D. Alexnet

AlexNet consists of eight layers scaled to enable fast computation. Five convolutional layers use ReLU activations for non-linearity, while max pooling reduces dimensionality. The network takes 227x227x3 RGB images as input into the first convolutional layer for red, green and blue encoding. The next layers perform hierarchical feature extraction, with deeper layers capturing higher-level attributes. The last fully connected layer links the extracted features to 1000 output categories. Overall, AlexNet contains 65,000 neurons and over 60 million parameters tuned on ImageNet's dataset of 1.2 million images. Once trained, the network transforms any image into a 4096-dimensional feature vector from the final hidden layer. This vector represents AlexNet's learned visual patterns that are then classified into one of the 1000 image classes it was trained on. The architecture allowed fast, broad-scale image recognition capabilities crucial for commercial viability.

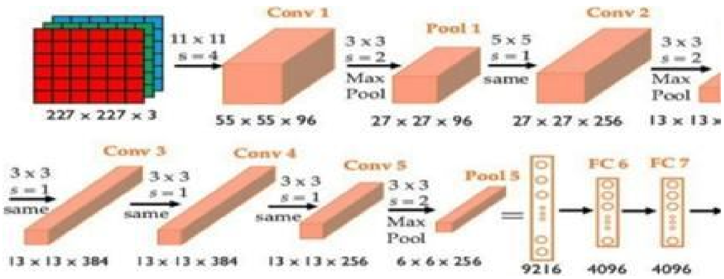

Figure 2: Architecture of Alexnet

E. SVM

Support Vector Machines, or SVMs, are an incredibly versatile machine learning technique adept at both regression and classification predicaments, though they shine brightest when categorizing data. The principal objective of SVMs is discovering the optimal hyperplane - essentially a dividing boundary - that cleanly bifurcates data points into discrete classes within a multidimensional feature space. More specifically, SVMs seek the hyperplane that provides maximum separation between the nearest data points from each class. The complexity of the hyperplane scales in tandem with the number of input features; a two-feature space yields a simple line, while three-feature and higher-dimensional spaces produce progressively more intricate separating geometries challenging for humans to internally visualize. Through clever identification of dividing boundaries, SVMs facilitate assigning new data points to one or another class with heightened accuracy. The versatility and classification prowess cement SVMs as a mainstay machine learning weapon crucial to any well-stocked analytical arsenal.

F. CNN

Convolutional neural networks (CNNs) are deep learning architectures well-suited for image analysis tasks. A CNN comprises convolutional, pooling and fully-connected layers that enable hierarchical feature extraction from pixel inputs. The convolutional layers form the backbone for learning low-level features like edges as well as higher-level attributes via filters applied across the input. Pooling layers then downsample the convolutional outputs to derive dominant features while reducing computational requirements. These extracted feature maps are finally classified or regressed in the fully-connected layers at the end of the network. Overall, the combination of convolutional and pooling builds increasingly abstract representations of the image where the important signals get amplified while discarding irrelevant variations. The final fully-connected neural layers then translate these learned feature extractors into accurate predictions for problems like image recognition.

G. Inception V3

Inception-v3, unveiled in 2015, employs several key architectural optimizations that improve accuracy and efficiency over previous Inception editions. At 42 layers deep, the model factors larger convolutions into smaller asymmetric filters that reduce computational requirements while expanding the network's width for richer feature extraction. Inception modules apply 1xN and Nx1 spatially factored convolutions followed by careful grid size reductions via multi-layer filters to further cut down parameters and cost. Additional auxiliary classifiers during training provide regularization to aid convergence. Together, these innovations in the design enable Inception-v3 to deliver high accuracy by learning nuanced attributes while keeping model size and inference time practical for real-world usage. The depth and widened architecture efficiently extract both low-level textures and higher-order semantics without compromising speed or ballooning hardware needs seen in earlier Inception models.

H. Logistic Regression

Logistic regression is a prevalent supervised machine learning technique suited for classification scenarios predicting binary outcomes - assigning instances one of two potential classes. It explores correlations between multiple predictor variables and a discrete target variable, estimating the probability an observation showcases a specific trait or emerges in a particular category. Assessing input features in aggregate and outputting likelihood of binary designations, logistic regression enables informed decision-making for categorical assignment problems across domains. Its discrete binary focus combined with probabilistic forecasting has secured logistic regression as a pivotal analytical asset when divining trends, patterns and insights from complex multi-dimensional data. By revealing degrees of relationship between varied inputs and dichotomous output, logistic regression empowers better classification choices.

I. XGBoost

XGBoost has emerged as a potent machine learning technique for tackling predictive modeling challenges in recent times. The algorithm implements a precise gradient boosting framework that utilizes both regularization approaches and computational optimizations to deliver state-of-the-art performance. Key aspects that distinguish XGBoost include its ability to effectively handle missing data, prevent overfitting via regularized boosting, execute very rapid feature tree constructions as well as provide insights into feature importance for model interpretations. These characteristics enable robust predictions across use cases - right from competitive benchmarks to real-world deployments. Additionally, the scalability of XGBoost across problems involving both regression and classification makes it a versatile addition to the data science toolkit. As a result, XGBoost has become the compiler of choice for data scientists and practitioners looking to harness the strengths of boosted decision tree models for obtaining accurate and interpretable predictions from data.

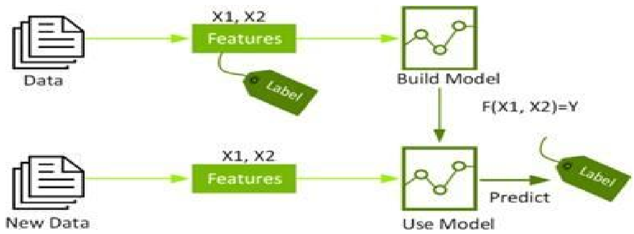

Figure 3: Model of XGBoost

J. LightGBM

Light Gradient Boosting Machine, or LightGBM, is a machine learning framework that learns extremely quickly. When employing strategies like GPU support, parallel computing, and histogram-based data processing, it is very effective for handling big datasets and real-time applications. Because of its speed, LightGBM is well-liked in machine learning contests and in applications such as image classification and natural language processing.

K. CatBoost

CatBoost is a machine learning algorithm designed for gradient boosting that performs well with categorical data. It is well known for being reliable and easy to use since it can automatically handle categorical features without requiring a lot of preprocessing. CatBoost stands apart with features like ordered boosting, advanced regularization techniques, and integrated cross-validation support. It has become more well-liked in data science competitions and real-world applications because of its adaptability to real-time prediction tasks and its capacity to produce accurate predictions with little adjustment to the hyperparameters.

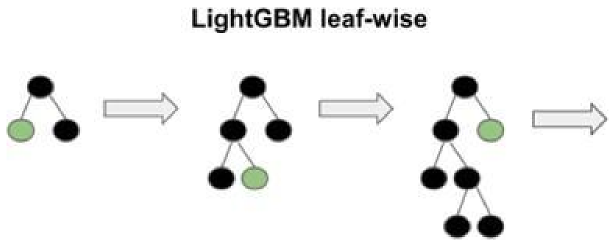

Figure 4: LightBGM leaf-wise

L. Linear Support Vector Machine

Support in Linear Form A machine learning algorithm called Vector Machine is applied to regression and classification problems. It operates by determining which hyperplane in the dataset best divides the various classes. High-dimensional data is a good fit for LSVM, which is renowned for its simplicity and effectiveness. Applications like text classification, image recognition, and anomaly detection make extensive use of it. For binary and multi-class classification issues, LSVM is an effective tool that can handle big datasets and provides good accuracy.

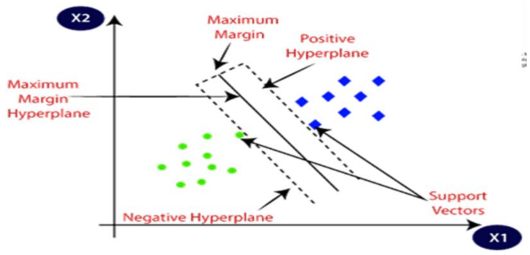

Figure 5: Classification of LSVM

IV. CONCLUSION

In reviewing existing literature, CatBoost and certain deep learning models demonstrate high accuracy in predicting polycystic ovary syndrome (PCOS). For exercise recommendations, a hybrid approach with association rule mining and recurrent neural networks can provide personalized regimens with high user adherence. Regarding diet plans, studies emphasize that abundant quality data allows deep learning techniques like LSTM to outperform traditional classifiers. An LSTM-based model achieved stellar performance on metrics including accuracy, recall and precision for individualized food recommendations. In summary, advanced algorithms like CatBoost for disease detection and deep recurrent networks for lifestyle recommendations showcase promising techniques for personalized health management. However, the success of these complex models relies heavily on dataset quality and availability. Further research into real-world viability is thus warranted, even as deep learning continues to offer new avenues for precision diagnosis and guidance.

ACKNOWLEDGMENT

Special thanks to our mentor, Dr V Baby and Dr P Anjusha, for all the technical support and guidance that led to the completion of the literature survey phase of the project with a good result in the end.

REFERENCES

[1] Hassan, M. M., & Mirza, T. (2020). Comparative Analysis of Machine Learning Algorithms in Diagnosis of Polycystic Ovarian Syndrome. Int. J. Comput. Appl, 975, 8887

[2] Guan, Y., Wang, D., Bu, H., Zhao, T., & Wang, H. (2020). The Effect of Metformin on Polycystic Ovary Syndrome in Overweight Women: A Systematic Review and MetaAnalysis of Randomized Controlled Trials. International Journal of Endocrinology, 2020.

[3] Bharati, S., Podder, P., & Mondal, M. R. H. (2020, June). Diagnosis of Polycystic Ovary Syndrome Using Machine Learning Algorithms. In 2020 IEEE Region 10 Symposium (TENSYMP) (pp. 1486-1489). IEEE.

[4] Vasundhara, S. Impact of Polycystic Ovary Syndrome-Approach with Machine Learning
[5] Maheswari, K., Baranidharan, T., Karthik, S., & Sumathi, T. (2021). Modelling of F3I Based Feature Selection Approach for PCOS Classification and Prediction. Journal of Ambient Intelligence and Humanized Computing, 12(1), 1349-1362
[6] Dutta, P., Paul, S., & Majumder, M. (2021). An Efficient SMOTE Based Machine Learning Classification for Prediction & Detection of PCOS.
[7] P. Wagh, M. Panjwani, and S. Amrutha, "Early Detection of PCOD Using Machine Learning Techniques," in Artificial Intelligence and Speech Technology: CRC Press, 2021, pp. 9-20.
[8] A. M. Gupta, S. S. Shetty, R. M. Joshi, and R. M. Laban, "Succinct Differentiation of Disparate Boosting Ensemble Learning Methods for Prognostication of Polycystic Ovary Syndrome Diagnosis," in 2021 International Conference on Advances in Computing, Communication, and Control (ICAC3), 2021: IEEE, pp. 1-5.
[9] Zigarelli, Z. Jia, and H. Lee, "Machine-Aided SelfDiagnostic Prediction Models for Polycystic Ovary Syndrome: Observational Study," JMIR Formative Research, vol. 6, no. 3, p. e29967, 2022.
[10] K., Ruat, C. Kathkar, and P.A. Ithkar, "PCOS Detection Using Machine Learning Algorithms," International Research Journal of Engineering and Technology (IRJET), vol. 9, no. 1, p. 1214-1218, Jan 2022
[11] Arash mahyari , Peter Piroll, “Physical Exercise Recommendation and Success Prediction Using Interconnected Recurrent Neural Networks”. 2021 IEEE International Conference on Digital Health (ICDH)
[12] P.Chitra,K.Srilatha, M.Sumathi, F.V.Jayasudha, T. Bernatin, Mandala Jagadeesh “Classification of Ultrasound PCOS Image using Deep Learning based Hybrid Models”. proceedings of the Second International Conference on Electronics and Renewable Systems (ICEARS-2023).IEEE Xplore Part Number: CFP23AV8-ART; ISBN: 979-8-3503- 4664-0
[13] Sumiya Islam,Nusrat Nabi,Sharun Akter Khushbu, Nushrat Jahan Ria. “A Process of Finding Common Symptoms and Diagnosis Age among PCOS Patients Through a Survey” 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT) | 978-1-7281-8595-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCCNT51525.2021.9580114
[14] Megh Shah, Sheshang Degadwala, Dhairya Vyas “Diet Recommendation System based on Different Machine Learners: A Review” 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS) | 978-1- 6654-0052-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICAIS53314.2022.9742919
[15] Sharvari S. Deshpande, Asmita Wakankar “Automated detection of Polycystic Ovarian Syndrome using follicle recognition” 2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies.
[16] Wan Azani Mustafa, Ahmed Alkhayyat, Hiam Alquran, Waleed Khalid Al-Azzawi, Noor Almajali, Dana Hdaib “Detection of Polycystic Ovary Syndrome (PCOS) Using Machine Learning Algorithms” 2022 IEEE 5th International Conference on Engineering Technology and its Applications (IICETA).
[17] Shazia Nasim, Mubarak Almutairi, Kashif Munir, Ali Raza, Faizan Younas “A Novel Approach for Polycystic Ovary Syndrome Prediction Using Machine Learning in Bioinformatics” 2022 IEEE | DOI: 10.1109/ACCESS. 2022.3205587.
[18] Akhter Hosain, Md Humaion Kabir Mehedi, and Irteza Enan Kabir “PCONet: A Convolutional Neural Network Architecture to Detect Polycystic Ovary Syndrome (PCOS) from Ovarian Ultrasound Images” 2022 IEEE International Conference on Engineering and Emerging Technologies (ICEET).
[19] Tin Trung Tran, Jae-Won Choi, Chien Van Dang, Geon Su Park, Jun Young Baek, and Jong-Wook Kim “Recommender System with Artificial Intelligence for Fitness Assistance System” 2018 IEEE 15th International Conference on Ubiquitous Robots (UR).
[20] S.Sadhasivam, M.S.Sarvesvaran, P.Prasanth, L. Latha 2023 “Diet and Workout Recommendation Using ML” IEEE 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA).