International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4
DOI: https://doi.org/10.17762/ijritcc.v11i4.6396
***Article Received**: 04 February 2023 **Revised**: 03 March 2023 **Accepted**: 13 March 2023*

A Comparative Study on Prediction of Endometriosis Causing Infertility Using Machine Learning Techniques: in Detail

Satya Kiranmai Tadepalli[1], Prof. P. V. Lakshmi[2]
[1]Asst. professor, IT Department
CBIT(A)
Hyderabad, India
tadepallikiranmai84@gmail.com
[2]Department of CSE
GITAM (Deemed to be University)
Visakhapatnam
vpanga@gitam.edu

Abstract— The purpose of this study is to utilize Artificial Intelligence to analyse and predict endometriosis problem in women. All traditional methods are used before to develop or to predict the likelihood of endometriosis based on the symptoms presented. By identifying the symptoms of endometriosis, the machine learning algorithms can determine the type of endometriosis and the appropriate course of action for patients. This technology can be used to educate women globally on the signs and symptoms of endometriosis and help them take preventive measures to avoid this deadly disease. The results of this research demonstrate the potential of advanced technology to revolutionize healthcare by providing early detection and treatment options for endometriosis. In areas with limited access to medical care, this tool can aid in identifying ovarian cancer and reducing mortality rates. By detecting and diagnosing endometriosis at an early stage, this program can play a significant role in promoting women's health and wellbeing. The methodology proposed in this study produces classification results that are on par with cutting-edge deep learning techniques. In addition, the methodology provides visual explanations that offer valuable insights into the inner workings of each model and enhance the accuracy and reliability of the predictions.

.

Keywords- Artificial Intelligence (AI), Machine Learning Algorithms, Endometriosis, Logistic Regression, Machine Learning, Classification, Feature selection

I. INTRODUCTION

According to Ashfaque [25], endometriosis is the 3rd most common disease among women and the 10th most common disease globally, causing about 400,000 new cases and 50,000 deaths in 2018. The strength of this research lies in its potential to raise awareness of endometriosis among women worldwide, especially in rural areas with low doctor-patient ratios. By predicting the likelihood of endometriosis based on symptoms, the software can help patients identify the type of endometriosis they have and take appropriate steps to reduce mortality rates [1]. Although mortality rates for endometriosis have decreased by more than 20% since the mid-1980s due to improvements in treatment and reduction in rates, less than half of women survive beyond 5 years after diagnosis due to aggressive high-grade serous carcinomas and a lack of early detection techniques and specific early symptoms.

The field of medical science is on the cusp of a technological revolution, with computer science playing a pivotal role [3]. The use of technology will make medical treatment more convenient, effective, and accessible. Artificial intelligence (AI) has the potential to diagnose diseases with greater accuracy than human expertise, thereby reducing detection errors. This article focuses on the development of an AI model to predict the likelihood of endometriosis. Endometriosis is a major concern among women, as symptoms are not easily detectable at early stages [4,5]. By using the symptoms of endometriosis as variables in any models and techniques of machine learning algorithms to analyse the model prediction and the probability of the disease. All these methods can be compared, and the outcome is analysed based on the result. This prediction can aid in early detection and help women take necessary precautions and receive advanced medical treatment.

Endometriosis is a common disorder observed in women who are of a menstruating age. It occurs when tissues, resembling the endometrium lining, develop on the outside of the uterus and other organs in the pelvic region. Symptoms vary from mild to severe, including pelvic pain, dysmenorrhea, and infertility. While no guaranteed treatment for endometriosis exists, early diagnosis and medical or surgical interventions can reduce the risk of complications and improve patients' quality of life [6]. Predicting the likelihood of endometriosis onset by analysing the medical history of diagnosed patients could aid healthcare

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4
DOI: https://doi.org/10.17762/ijritcc.v11i4.6396
***Article Received**: 04 February 2023 **Revised**: 03 March 2023 **Accepted**: 13 March 2023*

providers in diagnosis and improve patient well-being. To achieve this, the study employed different machine learning algorithms and their comparative studies.

II. Background

To start with, this study mainly focuses on different machine learning techniques that can be used for the early detection of endometriosis and comparative evaluation has been made in this study for various datasets used with different proposed algorithms [8].

The main objectives of this article are as follows:

a. To assess different machine learning models that can accurately predict the probability of endometriosis occurrence based on medical history.
b. To determine the crucial medical events in a patient's journey that ultimately result in endometriosis diagnosis.
c. To evaluate and rank the performance of the developed models against various databases and select the best-performing ones.
d. To use the predicted scores from the models to create patient profiles based on their likelihood of developing endometriosis.

To enable this to happen, here in this article, the datasets of blastocysts containing various time lapse images, patient health claim database of US for the year 2010, PLCO dataset and a Kaggle dataset of endometriosis is considered [7,10]. These databases are tested against different algorithms of machine learning and results and observations are recorded. Various analytical techniques were utilized to analyse the dataset, ranging from rules-based patient qualification criteria to machine learning algorithms, to determine the likelihood of endometriosis [22]. Each method is explained in detail in the subsequent sections of the article.

III. Feature Reduction

The features used for analysing endometriosis using machine learning techniques can vary depending on the specific study or analysis. However, some common features that have been used in previous studies include age, menstrual history, hormonal medication use, family history of endometriosis, symptoms (such as pelvic pain and dysmenorrhea), and previous surgeries related to endometriosis. Other features that may be considered include comorbidities, medication use, and demographic information [18,14]. Based on the different datasets used and applied, some features are observed and evaluated for reduction. The result of the feature reduction process is furnished below [Table – 1].

Table 1: Feature Reduction

Name of the dataset	Number of Initial Parameters	Number of Reduced Parameters	Feature Elimination Ratio (%)
Time-Lapse Embryo	78	11	84.21
Patient Health Claim	75	15	85.52
PLCO	67	08	89.47
Kaggle	77	14	88.15

The result is analysed visually for the feature reduction is given here. [Figure – 1].

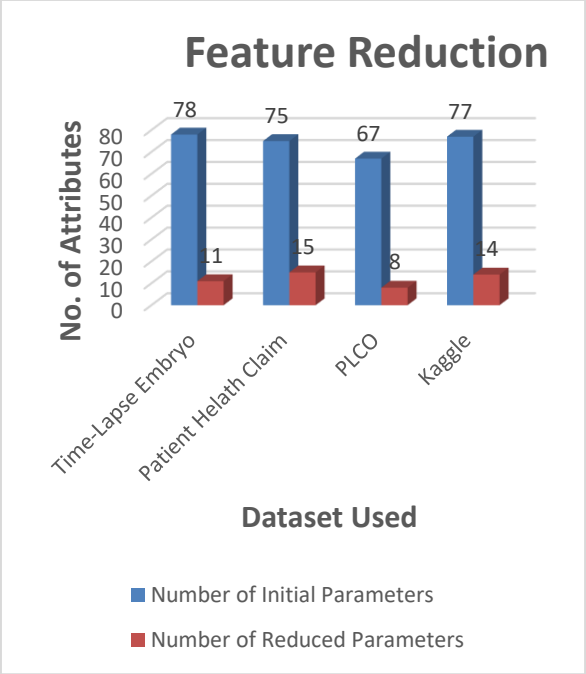

Figure 1: Feature Reduction

The number of attributes has been significantly reduced, and the accuracy of the dataset is compared to other standard methods using both the original and reduced sets of attributes.

IV. Classification Accuracy Analysis

The proposed model generates a subset of features, and when compared to the entire set of features, the classification accuracy is higher [15]. It is recommended to use the smaller set of attributes when training the model for effective prediction and improved accuracy. By using different classification methods, the results of these approaches are analysed. [Table – 2 to 5].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4
DOI: https://doi.org/10.17762/ijritcc.v11i4.6396
***Article Received**: 04 February 2023 **Revised**: 03 March 2023 **Accepted**: 13 March 2023*

TABLE 2: ACCURACY ANALYSIS OF TIME-LAPSE EMBRYO DATASET

Classification Method	**Evaluation using Training Set**			**Evaluation using Cross Validation**		
	Accuracy with the original set of attributes (%)	Accuracy with the reduced set of attributes (%) (Using MI)	Accuracy with the reduced set of attributes (%) (Using MI+RS)	Accuracy with the original set of attributes (%)	Accuracy with the reduced set of attributes (%) (Using MI)	Accuracy with the reduced set of attributes (%) (Using MI+RS)
Naïve Bayes	81.91	82.27	86.52	69.86	73.40	82.27
A-Star	100	100	98.23	58.51	68.09	76.60
Bagging	95.39	96.81	96.81	85.46	86.88	90.43
PART	99.29	99.65	99.65	87.94	92.91	95.39
Random Forest	100	100	100	73.76	89.36	95.04
Proposed Method	99.86	99.72	99.65	90.43	92.20	96.10

TABLE 3: ACCURACY ANALYSIS OF PATIENT HEALTH CLAIM DATASET

Classification Method	**Evaluation using Training Set**			**Evaluation using Cross Validation**		
	Accuracy with the original set of attributes (%)	Accuracy with the reduced set of attributes (%) (Using MI)	Accuracy with the reduced set of attributes (%) (Using MI+RS)	Accuracy with the original set of attributes (%)	Accuracy with the reduced set of attributes (%) (Using MI)	Accuracy with the reduced set of attributes (%) (Using MI+RS)
Naïve Bayes	82.65	69.39	82.99	71.77	65.31	81.29
A-Star	100	99.32	97.28	64.29	65.99	77.21
Bagging	94.90	82.65	93.88	85.37	67.35	87.76
PART	98.64	87.42	96.94	85.37	66.67	91.84
Random Forest	99.9	99.66	98.98	81.29	66.33	91.16
Proposed Method	98.98	88.10	94.56	79.25	64.63	79.93

TABLE 4: ACCURACY ANALYSIS OF PLCO DATASET

Classification Method	**Evaluation using Training Set**			**Evaluation using Cross Validation**		
	Accuracy with the original set of attributes (%)	Accuracy with the reduced set of attributes (%) (Using MI)	Accuracy with the reduced set of attributes (%) (Using MI+RS)	Accuracy with the original set of attributes (%)	Accuracy with the reduced set of attributes (%) (Using MI)	Accuracy with the reduced set of attributes (%) (Using MI+RS)
Naïve Bayes	60.16	72.36	72.36	39.84	66.67	66.67
A-Star	98.56	85.37	85.37	47.15	76.42	76.42
Bagging	85.37	83.74	83.74	46.34	71.54	71.54
PART	95.94	82.93	82.93	65.85	78.86	78.86
Random Forest	100	89.43	89.43	60.98	80.49	80.49
Proposed Method	99.19	86.18	86.18	65.04	82.93	82.93

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4
DOI: https://doi.org/10.17762/ijritcc.v11i4.6396
***Article Received**: 04 February 2023 **Revised**: 03 March 2023 **Accepted**: 13 March 2023*

TABLE 5: ACCURACY ANALYSIS OF KAGGLE DATASET

Classification Method	Evaluation using Training Set			Evaluation using Cross Validation		
	Accuracy with the original set of attributes (%)	Accuracy with the reduced set of attributes (%) (Using MI)	Accuracy with the reduced set of attributes (%) (Using MI+RS)	Accuracy with the original set of attributes (%)	Accuracy with the reduced set of attributes (%) (Using MI)	Accuracy with the reduced set of attributes (%) (Using MI+RS)
Naïve Bayes	44	76	76	27	66.5	66.5
A-Star	100	95	95	27.5	85.5	85.5
Bagging	88	92	92	81.5	90	90
PART	96.5	96	96	84.5	91.5	91.5
Random Forest	100	98	98	57.5	93	93
Proposed Method	98	91.5	91.5	57	87.5	87.5

It is evident that the accuracy has increased for all models with reduced features.

V. COMPUTATION TIME ANALYSIS

Computation time analysis in an ML model refers to the process of evaluating and measuring the time taken by the model to perform a specific task, such as training or making predictions [9]. It is an important aspect to consider when developing and deploying ML models because longer computation times can lead to slower model performance, increased costs, and reduced efficiency. Computation time analysis can help in identifying potential bottlenecks in the model's performance and optimizing it for better results. Various techniques and tools are available for performing computation time analysis, such as profiling and benchmarking, which can aid in improving the overall performance of an ML model [20]. The computational time of a model can be reduced when the feature selection and reduction of features is done based on the model evaluation. [Table – 6 to 9].

TABLE 6: COMPUTATION TIME ANALYSIS OF TIME-LAPSE EMBRYO DATASET

Classification Method	Computation time with original set of Attributes (in Sec)			Computation time with Reduced set of Attributes (in Sec)		
	Model Construction	Model Fitting	Model Usage	Model Construction	Model Fitting	Model Usage
Naïve Bayes	0.01	0.1	0.11	0	0.01	0.02
KNN	0	0	5.19	0	0	0.62
Bagging	0.27	0.39	0.01	0.05	0.12	0
Random Forest	0.52	0.44	0.06	0.26	0.21	0.09
Proposed Method	39.69	37.49	0.04	2.23	2.36	0.01

TABLE 7: COMPUTATION TIME ANALYSIS OF PATIENT HEALTH CLAIM DATASET

Classification Method	Computation time with original set of Attributes (in Sec)			Computation time with Reduced set of Attributes (in Sec)		
	Model Construction	Model Fitting	Model Usage	Model Construction	Model Fitting	Model Usage
Naïve Bayes	0.04	0.01	0.13	0.01	0	0.08
A-Star	0	0	5.18	0	0	0.53
Bagging	0.17	0.2	0.01	0.08	0.03	0.01
PART	0.1	0.08	0.01	0.05	0.05	0.01
Random Forest	0.35	0.32	0.04	0.19	0.18	0.03
Proposed Method	28.16	28.62	0.03	1.6	1.57	0.01

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4
DOI: https://doi.org/10.17762/ijritcc.v11i4.6396
***Article Received**: 04 February 2023 **Revised**: 03 March 2023 **Accepted**: 13 March 2023*

TABLE 8: COMPUTATION TIME ANALYSIS OF PLCO DATASET

Classification Method	Computation time with original set of Attributes (in Sec)			Computation time with Reduced set of Attributes (in Sec)		
	Model Construction	Model Fitting	Model Usage	Model Construction	Model Fitting	Model Usage
Naïve Bayes	0.03	0.01	0.06	0.01	0	0.04
A-Star	0	0	1.02	0	0	0.09
Bagging	0.14	0.11	0	0.04	0.01	0
PART	0.09	0.06	0	0.04	0.01	0
Random Forest	0.26	0.23	0.02	0.13	0.09	0.02
Proposed Method	11.71	11.51	0.02	0.53	0.47	0

TABLE 9 : COMPUTATION TIME ANALYSIS OF KAGGLE DATASET

Classification Method	Computation time with original set of Attributes (in Sec)			Computation time with Reduced set of Attributes (in Sec)		
	Model Construction	Model Fitting	Model Usage	Model Construction	Model Fitting	Model Usage
Naïve Bayes	0.02	0.01	0.08	0	0	0.01
A-Star	0	0	2.35	0	0	0.1
Bagging	0.18	0.19	0.01	0.05	0.02	0
PART	0.09	0.07	0	0.04	0.01	0
Random Forest	0.31	0.35	0.03	0.13	0.09	0.02
Proposed Method	18.98	20.47	0.03	0.55	0.5	0.01

Thus, it is natural to realize the efficiency of the proposed reduction method as the computation time has reduced significantly.

VI. CORRELATION MATRIX ANALYSIS

Correlation is a numerical value between zero and one that reflects the strength of the relationship between two variables. When the value is close to one, it represents a strong relationship [17]. A value close to zero represents a weak relationship. Various findings are furnished for four datasets of Infertility [Table – 10 to 13].

TABLE 10: CORRELATION MATRIX FOR TIME-LAPSE DATASET

	CP	OLDPEAK	CA	THAL	LMT	LADPROX	LADDIST	CXMAIN	OM1	RCAPROX	RCADIST
CP	*1.0*	*0.173*	*0.137*	*0.207*	*0.111*	*0.293*	*0.244*	*0.213*	*0.226*	*0.314*	*0.124*
OLDPEAK		*1.0*	*0.241*	*0.305*	*0.216*	*0.264*	*0.321*	*0.213*	*0.326*	*0.355*	*0.242*
CA			*1.0*	*0.173*	*0.169*	*0.092*	*0.252*	*0.259*	*0.261*	*0.324*	*0.194*
THAL				*1.0*	*0.157*	*0.248*	*0.302*	*0.176*	*0.198*	*0.294*	*0.157*
LMT					*1.0*	*0.065*	*0.197*	*0.248*	*0.192*	*0.182*	*0.077*
LADPROX						*1.0*	*0.12*	*0.161*	*0.233*	*0.135*	*0.149*
LADDIST							*1.0*	*0.313*	*0.274*	*0.299*	*0.2*
CXMAIN								*1.0*	*0.18*	*0.19*	*0.245*
OM1									*1.0*	*0.304*	*0.406*
RCAPROX										*1.0*	*0.105*
RCADIST											*1.0*

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4
DOI: https://doi.org/10.17762/ijritcc.v11i4.6396
***Article Received**: 04 February 2023 **Revised**: 03 March 2023 **Accepted**: 13 March 2023*

TABLE .11: CORRELATION MATRIX FOR PATIENT HEALTH CLAIM DATASET

	CP	MET	SLOPE	LMT	LADPROX	LADDIST	CXMAIN	RCAPROX	RCADIST	LVX4
CP	*1.0*	*0.159*	*0.375*	*0.233*	*0.383*	*0.185*	*0.258*	*0.262*	*0.245*	*0.257*
MET		*1.0*	*0.233*	*0.137*	*0.17*	*0.065*	*0.097*	*0.077*	*0.154*	*0.109*
SLOPE			*1.0*	*0.217*	*0.429*	*0.228*	*0.278*	*0.263*	*0.266*	*0.185*
LMT				*1.0*	*0.151*	*0.045*	*0.219*	*0.218*	*0.068*	*0.017*
LADPROX					*1.0*	*0.014*	*0.292*	*0.396*	*0.341*	*0.298*
LADDIST						*1.0*	*0.205*	*0.249*	*0.099*	*0.186*
CXMAIN							*1.0*	*0.392*	*0.185*	*0.236*
RCAPROX								*1.0*	*0.148*	*0.232*
RCADIST									*1.0*	*0.247*
LVX4										*1.0*

TABLE 12: CORRELATION MATRIX FOR PLCO DATASET

	LMT	LADPROX	LADDIST	CXMAIN	RAMUS	OM1	RCAPROX
LMT	*1.0*	*0.122*	*0.058*	*0.057*	*0.002*	*0.079*	*0.009*
LADPROX		*1.0*	*0.126*	*0.005*	*0.097*	*0.254*	*0.092*
LADDIST			*1.0*	*0.141*	*0.329*	*0.201*	*0.012*
CXMAIN				*1.0*	*0.004*	*0.189*	*0.097*
RAMUS					*1.0*	*0.337*	*0.222*
OM1						*1.0*	*0.078*
RCAPROX							*1.0*

TABLE 13: CORRELATION MATRIX FOR KAGGLE DATASET

	LMT	LADPROX	LADDIST	CXMAIN	OM1	RCAPROX	RCADIST	LVX4
LMT	*1.0*	*0.051*	*0.031*	*0.050*	*0.032*	*0.032*	*0.011*	*0.030*
LADPROX		*1.0*	*0.128*	*0.123*	*0.203*	*0.138*	*0.059*	*0.166*
LADDIST			*1.0*	*0.062*	*0.057*	*0.058*	*0.094*	*0.001*
CXMAIN				*1.0*	*0.076*	*0.195*	*0.097*	*0.277*
OM1					*1.0*	*0.041*	*0.084*	*0.021*
RCAPROX						*1.0*	*0.787*	*0.415*
RCADIST							*1.0*	*0.325*
LVX4								*1.0*

The results obtained from the proposed framework compared with the existing methods in terms of accuracy for classification and computation time.

VII. CONFUSION MATRIX ANALYSIS

A confusion matrix in ML for infertility is a table used to evaluate the performance of a binary classification algorithm that predicts infertility or non-infertility. It is a matrix that shows the number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) produced by the classification algorithm [22].

In the case of infertility prediction, a true positive would be a correct prediction of infertility, a false positive would be a prediction of

infertility when it's not present, a true negative would be a correct prediction of non-infertility, and a false negative would be a failure to predict infertility when it is present.

Confusion Matrix for Time-Lapse Embryo Dataset with reduced attributes

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4
DOI: https://doi.org/10.17762/ijritcc.v11i4.6396
Article Received: 04 February 2023 Revised: 03 March 2023 Accepted: 13 March 2023

		PREDICTED				
ACTUAL		0	1	2	3	4
	0	157	0	0	0	0
	1	0	49	1	0	0
	2	0	1	25	5	0
	3	0	0	1	31	0
	4	1	0	0	0	11

Accuracy = (157 + 49 + 25 + 31 + 11) / 282 = 96.8085 %

Confusion Matrix for Patient health Claim Dataset with reduced attributes

		PREDICTED				
ACTUAL		0	1	2	3	4
	0	175	9	2	2	0
	1	10	18	6	2	1
	2	1	5	14	5	1
	3	0	1	5	21	1
	4	0	0	0	1	14

Accuracy = (175 + 18 + 14 + 21 + 14) / 294 = 82.3129 %

Confusion Matrix for PLCO Dataset with reduced attributes

		PREDICTED				
ACTUAL		0	1	2	3	4
	0	7	1	0	0	0
	1	7	40	0	1	0
	2	0	4	26	2	0
	3	0	2	3	25	0
	4	0	0	0	1	4

Accuracy = (7 + 40 + 26 + 25 + 4) / 123 = 82.9268 %

Confusion Matrix for Kaggle Dataset with reduced attributes

		PREDICTED				
ACTUAL		0	1	2	3	4
	0	50	0	0	0	1
	1	2	51	2	1	0
	2	0	3	36	2	0
	3	0	1	1	40	0
	4	3	2	2	2	1

Accuracy = (50 + 51 + 36 + 40 + 1) / 200 = 89 %

Hence, it is proven for accuracy when compared to other methods.

VIII. NETWORK COMPARISON

This work proposes a novel network comparison using the analysis of different network characteristics. The details of the findings are listed here [Table – 14].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4
DOI: https://doi.org/10.17762/ijritcc.v11i4.6396
***Article Received**: 04 February 2023 **Revised**: 03 March 2023 **Accepted**: 13 March 2023*

TABLE 14: FRAMEWORK COMPARISON

Technique/Parameters	Accuracy (%)	Precision (%)	Sensitivity (%)	Negative predictive value (NPV %)	specificity (SPE) %
PCA	65.88	68.33	70.25	75.22	65.22
LDA	71.22	75.65	78.22	86.22	68.25
MLP	79.76	78.75	68.00	88.89	76.00
SVD+MLP	72.21	76.52	63.00	79.00	68.00
SVD+SVM	76.29	82.22	65.00	76.78	83.00
FUZZY SVD	82.24	79.98	74.00	68.34	91.50
Proposed DFKZ net	89.96	88.90	81.00	55.90	86.00

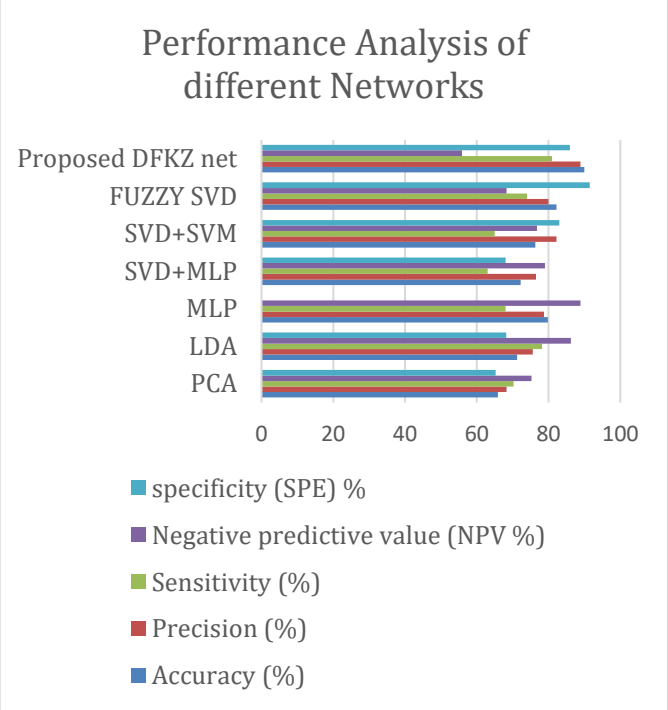

Fig.2. Network performance Comparisons

Thus, it is prominent to make a note that the proposed framework is significantly performed at its best to recent research's parallel outcomes.

IX. PROPOSED OPTIMIZER PERFORMANCE EVALUATION

Here Modified Whale Optimization has given best results when compared to other optimization techniques.

TABLE 15: OPTIMIZATION TECHNIQUES COMPARISON

Parameters	Proposed	GA	PSO	SHO	CSA	ALO
Accuracy	0.9885	0.9689	0.9446	0.8920	0.8739	0.8228
Error	0.0115	0.0311	0.0554	0.1080	0.1261	0.1772
Sensitivity	0.9886	0.9703	0.9446	0.9058	0.8860	0.8254
Specificity	0.9984	0.9956	0.9921	0.9846	0.9820	0.9747
Precision	0.9885	0.9689	0.9448	0.8921	0.8739	0.8227
FPR	0.0016	0.0044	0.0079	0.0154	0.0180	0.0253
F1_score	0.9886	0.9691	0.9444	0.8948	0.8764	0.8228
MCC	0.9869	0.9650	0.9367	0.8821	0.8607	0.7983
Kappa	0.9475	0.8578	0.7467	0.5062	0.4236	0.1899

Where here Proposed MWOA (Modified Whale Optimization) is compared with other techniques like, GA (Genetic Algorithm), PSO (Particle Swarm optimization), SHO (Spotted hyena optimization), CSA (Crow search optimization), ALO (Ant lion optimization)

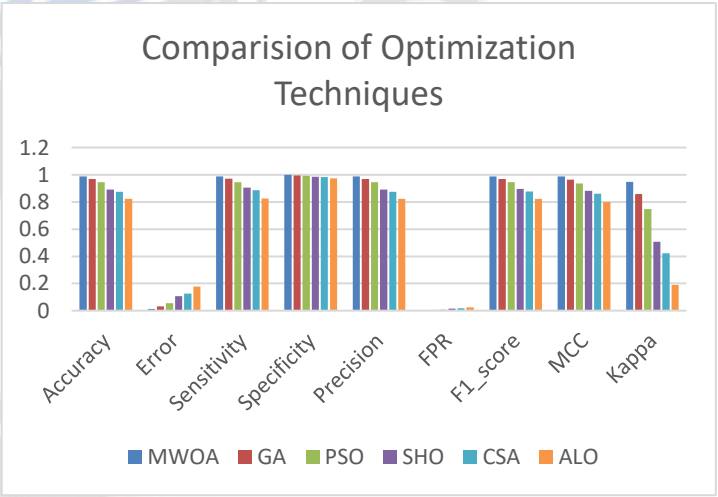

Fig.3. Optimization Techniques Comparisons

Given that, for each experiment, the feature extraction performance of the constructed model, varying the algorithm parameters as well as the number of input features presented to it, was assessed by calculating the classification accuracy (ACC) = (TP + TN)/(TP + TN + FP + FN), which denotes the probability of a correct classification;

Sensitivity (SEN) = TP/ (TP + FN), which scores the ability of the model to detect a subject with a specific disease in a population with more than one disease.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4
DOI: https://doi.org/10.17762/ijritcc.v11i4.6396
***Article Received**: 04 February 2023 **Revised**: 03 March 2023 **Accepted**: 13 March 2023*

Specificity (SPE) = TN/ (TN + FP), which scores the ability of the model to correctly rule out the disease in a disease-free population.

Precision (PREC) = TP/ (TP + FP), which defines the proportion of positive predictions.

and negative predictive value (NPV) = TN/ (TN + FN), which denotes the proportion of negative predictions.

X. CONCLUSION & FUTURE SCOPE

Overall, this article highlights the significant contribution of AI and ML in disease diagnosis, prediction, and forecasting. By analyzing the medical history of patients with endometriosis through machine learning algorithms, we were able to re-train the models on selected essential features and predict the likelihood of endometriosis occurrence in the adult female population. Early detection of the disease can enable patients to receive timely medical treatment and improve the patient's journey. We intend to develop a typing tool that can be integrated into the EHR systems, making it easily accessible to healthcare providers and aiding the diagnosis process for timely and accurate diagnosis, thereby enhancing patient care and quality of life. We plan to explore advanced deep learning algorithms in our future work to further improve the model's accuracy and performance in predicting the likelihood of the disease onset.

REFERENCES

[1] Marzyeh Ghassemi, Tristan Naumann, Peter Schulam, Andrew L. Beam, Irene Y. Chen, Rajesh Ranganath. A review of challenges and opportunities in machine learning for health. arXivLabs. 2019 v4, https://arxiv.org/abs/1806.00388

[2] Varun H Buch, Irfan Ahmed, Mahiben Maruthappu. Artificial intelligence in medicine: current trends and future possibilities. British Journal of General Practice 2018; 68 (668): 143-144. DOI: https://doi.org/10.3399/bjgp18X695213

[3] Alvin Rajkomar, Sneha Lingam, Andrew G. Taylor, Michael Blum, John Mongan. High-throughput classification of radiographs using deep convolutional neural networks. Journal of Digital Imaging 30, 95–101(2016). DOI: https://doi.org/10.1007/s10278-016-9914-9

[4] Min Chen, Yixue Hao, Kai Hwang, Lu Wang, Lin Wang. Disease prediction by machine learning over big data from healthcare communities. IEEE, 2169-3536 (2017), DOI: https://doi.org/10.1109/ACCESS.2017.2694446

[5] Adriana Gabriela Alexandru, Irina-Miruna Radu, Madalina - Lavinia Bizon. Big data in healthcare - opportunities and challenges. Informatica Economică vol.22, no. 2/2018. DOI: https://doi.org/10.12948/issn14531305/22.2.2018.05

[6] Iroju Olaronke, Ojerinde Oluwaseun. Big data in healthcare: Prospects, challenges and resolutions. IEEE, 16602629, 2016. DOI: https://doi.org/10.1109/FTC.2016.7821747

[7] Kiranmai, T. S., & Lakshmi, P. V. (2021). 3D Convolution Neural Network Based Ensemble Model to Detect Endometrium Issues at Early Stages and Enhance Fertility Chances in Women. Des. Eng, 1032-1044.

[8] Seidman, J. D., Horkayne-Szakaly, I., Haiba, M., Boice, C. R., Kurman, R. J., & Ronnett, B. M. (2004). The histologic type and stage distribution of ovarian carcinomas of surface epithelial origin. International journal of gynecological pathology, 23(1), 41-44.

[9] Kiranmai, T. S., & Lakshmi, P. V. (2021). A Comprehensive and Systematic Literature Review of Computational Intelligence Algorithms to Diagnose and Predict Female Infertility. Annals of the Romanian Society for Cell Biology, 5926-5943.

[10] S. Pathan, K. Gopalakrishna Prabhu, and P. C. Siddalingaswamy. "Automated detection of melanocytes related pigmented endometrium lesions: A clinical framework." Biomedical Signal Processing and Control 51 (2019): 59-72.

[11] Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J.A.; Van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. Med. Image Anal. 2017, 42, 60–88. [CrossRef]

[12] [18] Doupe P, Faghmous J, Basu S. Machine learning for health services researchers. Value Health. 2019;22(7):808-815. Available from: https://pubmed.ncbi.nlm.nih.gov/31277828/ [Accessed: October 1, 2020]

[13] Crown WH. Potential application of machine learning in health outcomes research and some statistical cautions. International Society for Pharmacoeconomics and Outcomes Research (ISPOR). 2015. DOI: 10.1016/j.jval.2014.12.005 [Accessed: October 1, 2020]

[14] Ghassemi M, Naumann T, Schulam P, Beam AL, Chen IY, Ranganath R. A review of challenges and opportunities in machine learning for health. arXivLabs. 2019. Available from: https://arxiv.org/abs/1806.00388 [Accessed: October 1, 2020]

[15] Buch VH, Ahmed I, Maruthappu M. Artificial intelligence in medicine: Current trends and future possibilities. British Journal of General Practice. 2018;68(668):143-144. DOI: 10.3399/bjgp18X695213 [Accessed: October 1, 2020]

[16] Rajkomar A, Lingam S, Taylor AG, Blum M, Mongan J. High-throughput classification of radiographs using deep convolutional neural networks. Journal of Digital Imaging. 2016;30:95-101. DOI: 10.1007/s10278-016-9914-9

[17] Kiranmai, T. S., & Lakshmi, P. V. A Novel Whale Optimized TGVFCMS Segmentation With Modified LSTM Classification For Endometrium Cancer Prediction.

[18] Das, S., Sanyal, M. K., & Datta, D. (2018). Advanced diagnosis of deadly diseases using regression and neural network. In Social Transformation–Digital Way: 52nd Annual Convention of the Computer Society of India, CSI

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 4
DOI: https://doi.org/10.17762/ijritcc.v11i4.6396
***Article Received**: 04 February 2023 **Revised**: 03 March 2023 **Accepted**: 13 March 2023*

2017, Kolkata, India, January 19-21, 2018, Revised Selected Papers 52 (pp. 330-351). Springer Singapore.

[19] Tadepalli, S. K., & Lakshmi, P. V. (2021). Deep Learning in IVF to Predict the Embryo Infertility from Blastocyst Images. In ICCCE 2020: Proceedings of the 3rd International Conference on Communications and Cyber Physical Engineering (pp. 1507-1515). Springer Singapore.

[20] Tadepalli, S. K., & Lakshmi, P. V. (2019). Application of Machine Learning and Artificial Intelligence Techniques for IVF Analysis and Prediction. International Journal of Big Data and Analytics in Healthcare (IJBDAH), 4(2), 21-33.

[21] Das, S., & Sanyal, M. K. (2020). Application of AI and soft computing in healthcare: a review and speculation. vol, 8, 21.

[22] Das, S., Biswas, S., Paul, A., & Dey, A. (2018). AI Doctor: An intelligent approach for medical diagnosis. In Industry Interactive Innovations in Science, Engineering and Technology: Proceedings of the International Conference, I3SET 2016 (pp. 173-183). Springer Singapore.

[23] Das, S., Sanyal, M., Datta, D., & Biswas, A. (2018). AISLDr: artificial intelligent self-learning doctor. In Intelligent Engineering Informatics: Proceedings of the 6th International Conference on FICTA (pp. 79-90). Springer Singapore.

[24] Ashfaque, J.M. (2021). Ovary Cancer Data. https://kaggle.com/ukveteran/ovary -cancer -data