International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 13 Issue: 1
Article Received: *25 December 2024* ***Revised:*** *12 January 2025* ***Accepted:*** *20 February 2025*

Performance Evaluation of Ovarian Cancer Detection Using on Machine Learning Approaches based on Feature Selection

[1]D. Savitha And [2]Dr. D. Rajakumari
[*1]Research Scholar, Department of Computer Science, Nandha Arts and Science College, Erode Tamilnadu, India
[*]E-Mail: saveethamca@gmail.com
[2]Associate Professor, Department of Computer Science, Nandha Arts and Science College, Erode, Tamilnadu, India
E-Mail: rsrajakumarid@gmail.com

Abstract

Ovarian cancer is one of the most dangerous genecology cancers because it does not show any symptoms in the early stages and there aren't any good tests to find it. Early detection is important for increasing patient survival, but traditional diagnostic methods often don't have the right level of sensitivity and specificity. This study looks into how machine learning (ML) and deep learning (DL) can be used to find ovarian cancer early and accurately. We looked at five models: Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Network (ANN), and Elman Recurrent Neural Network (ERNN). We did this before and after using the RF algorithm to select features. The results show that feature selection made all of the models work much better. The ERNN model performed the best overall, with accuracy going from 89.8% to 92.5% and AUC-ROC going from 0.94 to 0.96 after feature selection. In the same way, ANN and RF got 92.1% and 91.0% accuracy, respectively, with big improvements in precision, recall, and F1-score. These results show how important it is to optimize features to make models work better. They also confirm that intelligent ML-based systems could be used to reliably find ovarian cancer early.

1. Introduction

As the fifth most common cause of cancer-related death in women, ovarian cancer is a serious global health issue. Early symptoms are vague and nonspecific, making it a silent killer that is diagnosed at an advanced stage when treatment results are subpar. Early-stage ovarian cancer can have a five-year survival rate of over 90%, but late-stage detection has a five-year survival rate of less than 30%. In order to increase patient survival rates and quality of life, this start contrast highlighted the vital need for early and precise diagnostic techniques. Physical examinations, CA-125 biomarker testing, and transvaginal ultrasound are examples of conventional ovarian cancer diagnostic techniques. [1]. Although these methods are helpful, they usually don't have the sensitivity and specificity required for early detection. Additionally, non-malignant conditions may raise CA-125, leading to false positive results. Computational intelligence (CI) and data-driven approaches are becoming more and more popular in the medical research community. When used on large clinical datasets, machine learning models have demonstrated significant promise in improving diagnostic accuracy.

ML techniques are offered that can learn from data and make decisions without needing to be specifically programmed for particular tasks. In complex datasets, machine learning algorithms can uncover hidden patterns that traditional statistical techniques might miss [2]. Many ML methods have been applied for ovarian cancer detection, ranging from classical models such as DT, RF, SVM, ANNs, and ERNN. The DT approach is a simple and powerful method that mimics human decision-making through a series of if-then rules. Its suggestions high interpretability and is frequently used as a baseline method. Though, DTs are likely to to overfitting and may have poor generalization accuracy to unseen data. RF is an ensemble learning method that builds multiple decision trees and aggregates their prediction, leading to enhance accuracy, robustness, and struggle to overfitting. SVM is another approach which performance well in both linear and nonlinear data by utilizing kernel functions, making it suitable for biomedical applications where data is often complex and not linearly separable. But, SVMs can be

computationally intensive when taking large datasets. ANNs simulate the biological neural networks of the human brain and are talented of modeling complex nonlinear relationships. However, their performance is highly dependent on the quality of data, the architecture, and hyperparameters.

The ERNN is a kind of recurrent neural network (RNN) which is suitable for time-series analysis or sequence data [3]. ERNN is useful for applications involving dynamic and sequential patterns, like gene expression profiles or changes in biomarker levels over time, because it adds a context layer that allows it to retain a memory of prior inputs. Notwithstanding their promise, ERNNs can be difficult to train and may have problems like vanishing gradients. The use of these ML models in the detection of ovarian cancer has been investigated in a number of studies in recent years. These studies have used a variety of data types, such as blood biomarkers, gene expression data, proteomics, and medical imaging. This study's main goals are to create and assess different machine learning and deep learning models for the early detection of ovarian cancer and look into how feature selection affects model performance. Finding the best model in terms of accuracy, sensitivity, specificity, and overall diagnostic capability is the main goal of the study. The contribution of the paper is as follows,

- Five models— DT, RF, SVM, ANN, and ERNN— were implemented and evaluated on an ovarian cancer dataset.
- The study explores the effect of feature selection on each model's performance, demonstrating its role in improving predictive accuracy and model robustness.
- Detailed evaluation was performed using key metrics such as accuracy, precision, recall, F1-score, and AUC-ROC, both before and after feature selection.
- The ERNN model was identified as the best-performing classifier, particularly after feature selection, due to its superior ability to learn sequential and temporal patterns.

Remaining sections of the paper is organized as follows: Section 2 discussed recent research papers. Section 3 shows the research methods. Section 4 shows the experimental results with details results analysis. Section 5 discusses about the conclusion, limitations, and future enhancements of research work

2. Related works

Researchers have been working for detecting ovarian cancer and provided understanding information through thorough learning and research. The goal of continuing research in this area is to offer more efficient methods for getting the best results. Some recent methods that deal with the prediction of ovarian cancer are listed below. Kumar et al. (2022) [4] developed a machine learning-based ovarian cancer screening model using risk factors. Highest accuracy of 85.7% achieved using RF classifier which is an effectively integrates genetic, lifestyle, and clinical risk factors for robust prediction. Provides interpretable results aiding early identification of high-risk patients. Supports cost-effective screening by reducing unnecessary invasive tests. Enhances early detection and patient stratification in ovarian cancer prevention programs. Singh et al. (2024) [5] proposed an innovative early prediction method for ovarian cancer by combining machine learning with Explainable Artificial Intelligence (XAI) techniques. The developed method achieved a highest accuracy of 90.8% using an optimized Gradient Boosting model combined with XAI for interpretability. This method enabled clear and understandable predictions, serving clinicians comprehend model conclusions.

Chen et al. (2023) developed a novel fuzzy DL method for detection of ovarian cancer and it achieving a highest accuracy of 91.3%. The developed model efficiently handled the inherent uncertainty and inaccuracy present in data by integrating fuzzy logic with DL particularly noisy and imperfect clinical data. The developed method improves accuracy, robustness and interpretability when compared to traditional ML methods. Zhang et al. (2024) [6] proposed an advanced multi-modality DL approach that integrates imaging, clinical, and genetic data to improve diagnosis accuracy. The developed approach achieved a highest accuracy of 93.1%, significantly outperforming conventional models. The model achieved enhanced precision and sensitivity by leveraging DL methods capable of capturing complex and non-linear relationships across diverse data types. This comprehensive integration facilitated early and reliable detection of ovarian cancer, which is crucial for timely intervention and effective treatment planning.

Reddy et al. (2023) [7] presented a hybrid approach to ovarian cancer detection that blends traditional machine learning techniques with Genetic Algorithms (GA) for feature extraction and selection. By optimizing feature

subsets before classification to improve model performance, the developed method achieved the highest accuracy of 89.6%. Finding the most pertinent features in high-dimensional datasets is made possible by GA, which also effectively lowers computational complexity and increases classification accuracy. When combined with GA-optimized features, SVM and RF, which were used as classifiers, demonstrated a strong ability to differentiate between cancerous and non-cancerous cases. Rani et al. (2023) used adaptive differential evolution (ADE) for feature selection and hyperparameter tuning to create an optimal machine learning approach for ovarian cancer detection. By using ADE to dynamically adapt mutation and crossover approaches, the study aims to improve the performance of traditional classifiers and successfully identify the most discriminative features subset from complex datasets. The SVM and RF models outperformed their non-optimized approach, achieving an accuracy of 90.7%. The developed model reduced training time and model complexity while achieving the highest accuracy.

H. R. Farhan et al. (2024) [8] present an efficient technique for ovarian cancer identification using an ERNN, which can identify cancer from mass spectrometry data,. The topology of the network consists of two output nodes to show the status, five neurons for the context and hidden layers, and 100 input neurons for data reception. The suggested technique makes use of reduced-size characteristics, such as ion concentration levels at particular mass/charge values, which are trained using a variety of learning algorithms to choose the optimal one.

3. Research methods

DT, RF, SVM, ANNs, and ERNN are widely used ML models for classification and prediction tasks. DT and RF are tree-based models, with RF offering better accuracy through ensemble learning. SVM excels in high-dimensional spaces with a clear margin of separation. ANN can model complex nonlinear relationships, while ERNN, a type of recurrent neural network, is particularly effective for sequential or time-dependent data. Each model has unique strengths and limitations, making them suitable for different types of problems. The following subsection discusses about each in details.

3.1 Random forest (RF)

RF is an ensemble learning technique that builds a "forest" of multiple decision trees during training and merges their outputs to produce more accurate and stable predictions [9]. By employing distinct subsets of data and random subsets of features for every tree split, it adds unpredictability, which aids in decorrelating individual trees and minimizing overfitting. RF is made extremely robust by this ensemble approach, which also enhances generalization performance. The capacity of RF to manage high-dimensional datasets with thousands of input variables without feature selection is one of its main advantages. Additionally, it offers feature importance scores, which are helpful for figuring out which variables have the biggest impact on the prediction. But compared to a single decision tree, it can be more computationally demanding and less interpretable, which could be problematic for applications that demand model transparency. In spite of this, it is still the method of choice for many real-world issues, such as image analysis, bioinformatics, and fraud detection.

3.2 Decision tree (DT)

A supervised machine learning technique that looks like a flowchart is called a DT[10]. It begins at a root node and branches out according to feature-value-based decision rules until it reaches leaf nodes, which represent classes or results. The traits that best separate the data are selected to build the tree using metrics such as Information Gain, Gini Index, or Entropy. DT are particularly useful in industries like healthcare and finance where explainability and transparency are essential because they are simple to comprehend and visually interpret. Despite its interpretability and ease of use, DT has several disadvantages. They are prone to overfitting, especially when the data is noisy or the tree grows very deep.

3.3 Support vector machine (SVM)

SVM is an effective supervised learning model for tasks involving regression and classification. It operates by locating the best hyperplane in a high-dimensional space that maximally divides data points of various kinds [11]. SVM employs kernel functions to convert the input space into a higher dimension where a linear separator may exist when the data is not linearly separable. The crucial components of the training set that affect the hyperplane's position are the support vectors. Text categorization and gene expression data analysis are two applications that can benefit from

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 13 Issue: 1
Article Received: *25 December 2024* ***Revised:*** *12 January 2025* ***Accepted:*** *20 February 2025*

SVM's ability to perform effectively in high-dimensional spaces and even when there are more dimensions than samples. SVM, however, can be memory-intensive and slow when used on big datasets. Its performance is also heavily influenced by the kernel selection and hyperparameter adjustment, such as the kernel coefficient (gamma) and regularization parameter (C). However, its excellent classification accuracy and capacity to represent intricate boundaries make it an effective machine learning tool**.**

3.4 Artificial neural network (ANNs)

The composition and operations of the human brain serve as an inspiration for ANNs[12]. They are made up of layers of nodes, or "neurons," where each node applies a non-linear activation function after applying a weighted sum of inputs. By using a technique known as backpropagation, in which errors are propagated backward to modify weights, ANNs can learn intricate, non-linear correlations in data. They are quite adaptable and may be applied to a variety of domains for both classification and regression problems. demands significant computer resources and a significant volume of labeled data. Furthermore, ANNs may overfit, necessitating the use of regularization strategies such batch normalization and dropout. Notwithstanding these difficulties, they are essential to contemporary AI systems due to their adaptability and scalability.

3.5 Elman recurrent neural network (ERNN)

Elman proposed in 1990 a typical method called the ERNN [13]. ERNN framework shows in Figure 1. The recurrent layer based on the hidden layer of the BPNN model is added to the ERNN, one type of feedback neural network; it can be thought of as a delay operator and introduces a memory function. It maintains stability around the world and allows the network to adjust to dynamic, time-varying characteristics. The ERNN model's topological structure typically consists of four layers: The signal is passed from an input layer, whose neurons are typically linear, to the hidden layer, where it is translated or increased using an activation function. The connecting layer's responsibility is to accept the hidden layer's output and provide feedback with the information matching the previous instance to the hidden layer to create a local ring structure. The output value of the neural network is more by the real development trend of the data because the connecting layer unit has a delayed memory impact on the features contained in past data. The findings are finally output via the output layer. The ERNN is based on the BPNN's structure, but it joins the output of the hidden layer to its input automatically according to the context layer's delay and storage functions. This internal feedback process can enhance the neural network's capacity to handle dynamic information because how this joining happens is sensitive to the historical data of the neural network itself. The system can adjust to time-varying characteristics based on this mapping of the dynamics to a stored internal state. A recurrent, an output, a hidden layer, and an input layer make up an ERNN. One or more neurons are present in each layer, and they transmit data or samples by computing a nonlinear function of the weighted sum of the input samples. The definition of the input layer's mathematical model is

$$X_{it}(k) = \sum_{i=1}^{n} X_{it}(k-1) \quad (1)$$

Here, X_{it} - denote an input at time t with n number of neurons. The input model for each neuron in the hidden layer is as follows:

$$net_{jt}(k) = \sum_{i=1}^{n} W_{ij} X_{it}(k-1) + \sum_{j=1}^{p} C_j r_{jt}(k) \quad (2)$$

W_{ij} - weights between input and hidden layer. C_j - weights between hidden and recurrent layers. The hidden layer's output is as follows:

$$Z_{jt}(k) = f(net_{jk}(k) = \sum_{i=1}^{n} W_{ij} X_{it}(k-1) + \sum_{j=1}^{p} C_j R_{jt}(k) \quad (3)$$

The recurrent layer is defined as follows:

$$R_{jt}(k) = Z_{jt}(k-1) \quad (4)$$

The output layer is defined as follows:

$$Y_t(k) = f(\sum_{j=1}^{p} V_j Z_{jt}(k) \quad (5)$$

The network errors of ERNN are as follows,

$$E = \sum_{k=1}^{m} (t_t - y_t)^2 \quad (6)$$

Here, t_t – target value and y_t – predicted value. The updating of each weight matrix may be calculated using the following formula to reduce E.

$$W(t+1) = W(t) - \mu \frac{\partial E}{\partial w(t)} \quad (7)$$

Here, μ is the learning rate.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 13 Issue: 1
Article Received: *25 December 2024* ***Revised:*** *12 January 2025* ***Accepted:*** *20 February 2025*

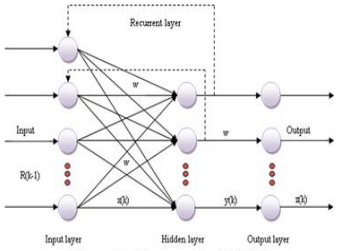

Figure 1 : Framework of ERNN

4. Experimental results analysis

The analysis of results plays a critical role in evaluating the effectiveness, reliability, and clinical relevance of machine learning models applied to ovarian cancer detection. By systematically comparing the performance of various classifiers such as DT [14], RF [15], SVM [16], ANNs [17], and ERNN across key metrics like accuracy, precision, recall, F1-score, and AUC-ROC, the study identifies not only the most promising model but also the conditions under which each model performs best. The parameters setting of ERNN shows in Table 1 and Figure 2 shows the overview of proposed method.

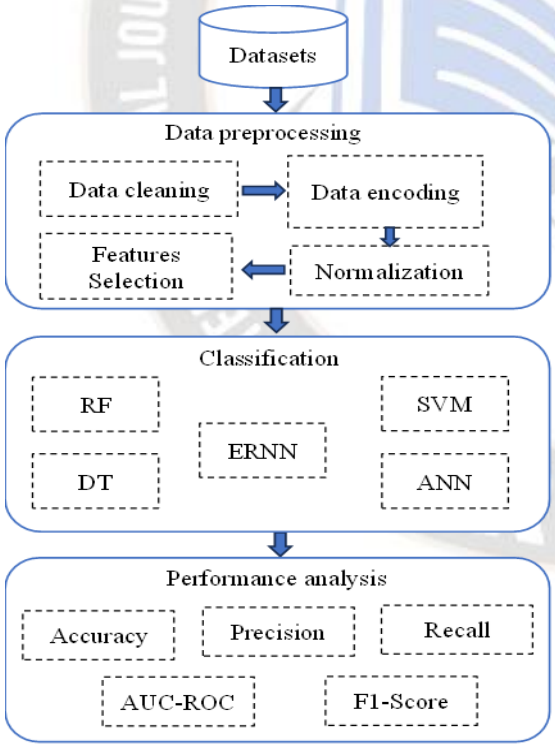

Figure 1 : Overview of proposed method

4.1 Data collection

Here, we have selected 349 individuals with OC from the Kaggle website ("https://www.kaggle.com/saurabhshahane/predict-ovarian-cancer"), comprising 178 cases of benign ovarian cancer and 171 cases of OC from the Third Affiliated Hospital of Suzhou University between July 2011 and July 2018 [2]. The data set is split into two sections: 30% of OC patients make up the testing set, while 70% of OC patients make up the training set. The selected data set consists of 49 predictor variables, including age, menopause, six tumor markers, 22 general chemical tests, and 19 normal blood tests. None of the patients had prior radiation or chemotherapy, and all patients obtained a case 146 diagnosis following their operation. The World Health Organization's criteria were used to classify the histological type of diagnosis [18]. Raw data must be normalized before the prediction phase with numerical data spanning significantly different ranges. Then, using the min-max normalization approach, all data will be turned into values between 0 and 1, as seen below [19, 20].

$$x = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (8)$$

The present value is denoted by x, whereas the maximum and minimum values are represented by $\max(x)$ and $\min(x)$.

Table 1: Performance measures

ERNN	
Parameter	**Value**
Training method	Adam
Objective functions	MSE
Learning rate	0.05
Activation	TanH, Sigmoid
Number of training epochs	1000
Error condition	0.0005
Range of weights	[-0.5- 0.5]

4.2 Performance measures

Performance analysers are employed to assess the efficacy of ML techniques. This study compares the performance of prediction algorithms using four performance indicators: accuracy, precision, recall, and F-measures. A list of them is given below [21],

- Accuracy: a model's accuracy, which measures its correctness, may be stated as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (9)$$

- Precision: The proportion of correctly recognized samples (positive) inside all identified samples

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 13 Issue: 1
Article Received: *25 December 2024* ***Revised:*** *12 January 2025* ***Accepted:*** *20 February 2025*

(positive) is known as precision [22], and it may be expressed as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (10)$$

- Recall: Recall conveys the classifier's ability to accurately categorize samples inside a specific class, which is as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (11)$$

- F-Measures: By balancing Precision and Recall, the F1-score is employed in situations where there is a class imbalance in the data, which is as follows:

$$\text{F} - \text{Measures} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision+Recall}} \quad (12)$$

False negative (FN) indicates diseases that were not correctly predicted, while false positive (FP) indicates mistakenly assumed normal. True negative (TN) indicates accurately anticipated normal, while true positive (TP) indicates effectively predicted disease.

4.3 Discussions

The comparative evaluation of machine learning and deep learning models for ovarian cancer detection, as presented in Tables 2 and 3 and Figures 4 and 5, reveals the significant impact of feature selection on classification performance. Before feature selection, the DL models—particularly the ANN and ERNN—outperformed traditional machine learning models in all performance metrics. The ERNN model achieved the highest AUC-ROC value of 0.94 and maintained high precision and recall, indicating its strong capability to learn temporal and sequential patterns from the data. ANN also demonstrated strong results with a 89 % accuracy and an F1-score of 87.1%.

After applying feature selection, all models exhibited performance improvements, confirming that irrelevant or redundant features can negatively affect model generalization and learning efficiency. The optimal features are selected based on feature importance score using RF algorithm which is shows in Figure 3. Feature selection enhanced the discrimination ability of the models by reducing noise and dimensionality. The ERNN model again stood out, with its accuracy increasing to 92.5%, and AUC-ROC improving to 0.96—the highest among all models—demonstrating its robustness and adaptability when trained on optimized feature sets. Similarly, ANN and RF showed considerable gains, with ANN reaching 92.1% accuracy and an AUC of 0.95, while RF achieved 91.0% accuracy and an AUC of 0.94. SVM and DT also showed measurable improvements, although they remained slightly behind the deep learning-based models.

These findings highlight two key insights: (1) DL models, especially ERNN, are highly effective in handling complex, high-dimensional datasets like those involved in cancer diagnosis, and (2) feature selection is a critical preprocessing step that can substantially boost model performance across all types of learning algorithms. The synergy between optimized input features and advanced learning architectures can significantly improve diagnostic reliability, making this combination highly promising for early-stage ovarian cancer detection.

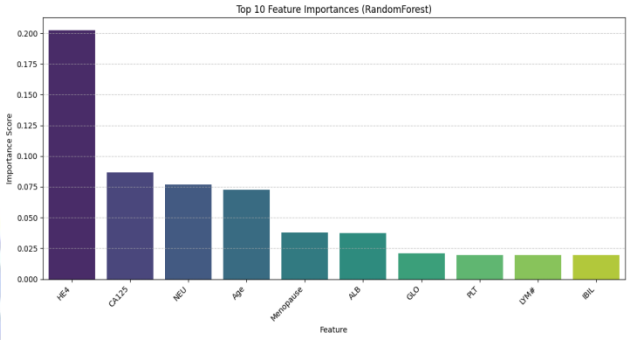

Figure 2 : Feature important score obtained by RF algorithm

Table 2 : Performance results based on different ML methods before selection

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
DT	85.2	83.5	84.5	84.0	0.86
RF	89.8	87.9	88.8	88.3	0.92
SVM	88.3	86.3	87.5	86.9	0.91
ANN	89	87.7	88.6	87.1	0.93
ERNN	89.8	88.2	89.3	88.7	0.94

Table 3 : Performance results based on different ML methods after feature selection

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
DT	87.3	85.4	86.2	85.8	0.88

RF	91.0	89.6	90.2	89.9	0.94
SVM	89.5	87.4	88.5	87.9	0.93
ANN	92.1	90.8	91.5	91.1	0.95
ERNN	92.5	91.2	91.9	91.5	0.96

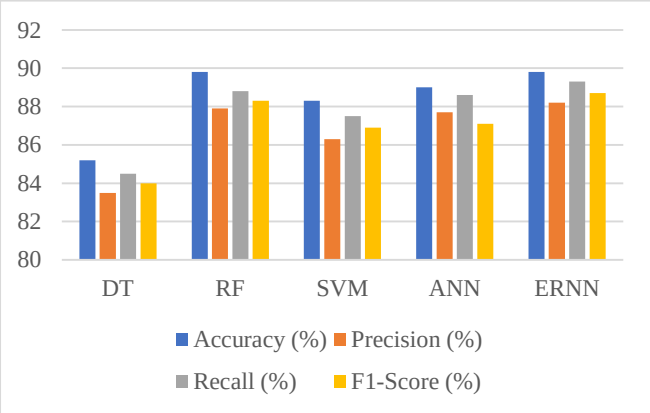

Figure 3 : Performance results based on different ML methods before selection

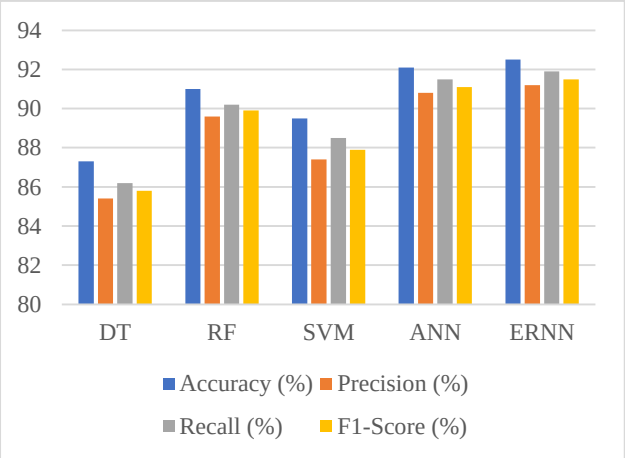

Figure 4 : Performance results based on different ML methods after feature selection

5. Conclusions

This study demonstrates the effectiveness of ML and DL models for the early detection of ovarian cancer. Among the five models evaluated—DT, RF, SVM, ANN and ERNN—the deep learning-based ERNN model achieved superior results in terms of accuracy, F1-score, and AUC-ROC. The application of feature selection techniques significantly improved the performance of all models by eliminating irrelevant attributes and enhancing learning efficiency. Notably, the ERNN model reached an accuracy of 92.5% and an AUC-ROC of 0.96 after feature selection, indicating its strong potential in medical diagnostic applications. These results highlight that integrating optimized feature selection with advanced learning models can lead to highly reliable and accurate diagnostic tools for ovarian cancer.

References

[1] M. El-Khatib, D. Popescu, O. Teodor, and L. Ichim, "New Trends in Ovarian Cancer Diagnosis Using Deep Learning. A Systematic Review," *IEEE Access,* 2024.

[2] M. Lu *et al.*, "Using machine learning to predict ovarian cancer," *International journal of medical informatics,* vol. 141, p. 104195, 2020.

[3] K. Senthilkumar, S. Sathiyabama, and D. Ayyamuthukumar, "Hate Speech Detection Using Enhanced Recurrent Neural Network Based on Elephant Herding Optimization on Social Media," in *2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, 2024: IEEE, pp. 1-6.

[4] R. Nopour, "Screening ovarian cancer by using risk factors: machine learning assists," *BioMedical Engineering OnLine,* vol. 23, no. 1, p. 18, 2024.

[5] S. L. JM and P. Subbulakshmi, "Innovative approach towards early prediction of ovarian cancer: Machine learning-enabled XAI techniques," *Heliyon,* vol. 10, no. 9, 2024.

[6] Z. Wang *et al.*, "Multi-modality deep learning model reaches high prediction accuracy in the diagnosis of ovarian cancer," *Iscience,* vol. 27, no. 4, 2024.

[7] S. Acosta-Jiménez *et al.*, "Detection of ovarian cancer using a methodology with feature extraction and selection with genetic algorithms and machine learning," *Network Modeling Analysis in Health Informatics and Bioinformatics,* vol. 14, no. 1, p. 3, 2024.

[8] H. R. Farhan, M. S. Kod, A. Taqi, and S. A. Ghazi, "Ovarian cancer detection based on elman recurrent neural network," *Periodica Polytechnica Electrical Engineering and Computer Science,* vol. 68, no. 2, pp. 202-212, 2024.

[9] A. Arfiani and Z. Rustam, "Ovarian cancer data classification using bagging and random

forest," in *AIP Conference Proceedings*, 2019, vol. 2168, no. 1: AIP Publishing.

[10] S. Piedimonte *et al.*, "Using a machine learning algorithm to predict outcome of primary cytoreductive surgery in advanced ovarian cancer," *Journal of Surgical Oncology,* vol. 127, no. 3, pp. 465-472, 2023.

[11] S. M. Ayyoubzadeh, M. Ahmadi, A. B. Yazdipour, F. Ghorbani-Bidkorpeh, and M. Ahmadi, "Prediction of ovarian cancer using artificial intelligence tools," *Health Science Reports,* vol. 7, no. 7, p. e2203, 2024.

[12] M. A. Rahman, R. C. Muniyandi, K. T. Islam, and M. M. Rahman, "Ovarian cancer classification accuracy analysis using 15-neuron artificial neural networks model," in *2019 IEEE Student Conference on Research and Development (SCOReD)*, 2019: IEEE, pp. 33-38.

[13] J. L. Elman, "Finding structure in time," *Cognitive science,* vol. 14, no. 2, pp. 179-211, 1990.

[14] A. Osmanović, L. Abdel-Ilah, A. Hodžić, J. Kevric, and A. Fojnica, "Ovary cancer detection using decision tree classifiers based on historical data of ovary cancer patients," in *CMBEBIH 2017: Proceedings of the International Conference on Medical and Biological Engineering 2017*, 2017: Springer, pp. 503-510.

[15] C. Nayak, A. Tripathy, M. Parhi, and S. K. Barisal, "Early Stage Ovarian Cancer Prediction using Machine Learning," in *2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT)*, 2023: IEEE, pp. 603-608.

[16] J. Zhou, L. Li, L. Wang, X. Li, H. Xing, and L. Cheng, "Establishment of a SVM classifier to predict recurrence of ovarian cancer," *Molecular Medicine Reports,* vol. 18, no. 4, pp. 3589-3598, 2018.

[17] M. M. Ahamad *et al.*, "Early-stage detection of ovarian cancer based on clinical data using machine learning approaches," *Journal of personalized medicine,* vol. 12, no. 8, p. 1211, 2022.

[18] X. Hu, Y. Xie, Y. Yang, and H. Jiang, "Group penalized logistic regression differentiates between benign and malignant ovarian tumors," *Soft Computing,* pp. 1-20, 2023.

[19] K. Velusamy and R. Amalraj, "Cascade correlation neural network with deterministic weight modification for predicting stock market price," in *IOP Conference Series: Materials Science and Engineering*, 2021, vol. 1110, no. 1: IOP Publishing, p. 012005.

[20] K. Velusamy and R. Amalraj, "Performance of the cascade correlation neural network for predicting the stock price," in *2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 2017: IEEE, pp. 1-6.

[21] R. Trevethan, "Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice," *Frontiers in public health,* vol. 5, p. 307, 2017.