Integrating Machine Learning Algorithms and Advanced Computing Technology Using an Ensemble Hybrid Classifier

[1] Roopashri Shetty, [2] Geetha M, [3] Shyamala G and [4] Dinesh Acharya U
[1,2,4] Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India.
[3] Department of Obstetrics and Gynecology, Kasturba Medical College, Manipal Academy of Higher Education, Manipal, Karnataka, India.
[1]roopashri.shetty@manipal.edu, [2]geetha.maiya@manipal.edu, [3]shyamala.g@manipal.edu, [4]dinesh.acharya@manipal.edu

Correspondence should be addressed to Geetha M: geetha.maiya@manipal.edu

Article Info
Journal of Machine and Computing (http://anapub.co.ke/journals/jmc/jmc.html)
Doi: https://doi.org/10.53759/7669/jmc202404068
Received 08 May 2023; Revised from 18 December 2023; Accepted 14 June 2024
Available online 05 July 2024.

Abstract – Ovarian Cancer (OC) is one of the major types of cancers in women worldwide. Despite the standardization of characteristics that can help distinguish benign from malignant ovarian masses, accurate predictive modelling following ultrasound (US) examination and biomarkers for 'progression-free survival' is lacking in the field of ovarian cancer. Important leading factors in ovarian cancer lethality are the lack of diagnostic procedures and proper screening to detect early-stage ovarian cancer, and the rapid spread of the disease over the surface of the peritoneum. Therefore, developing tools for accurate screening and prognosis, as well as the diagnosis of early stage ovarian cancer, is a current clinical need. In this study, an ensemble classifier was developed as a novel means of ovarian cancer prediction, and its effectiveness was assessed. The ensemble classifier integrates various machine learning algorithms, including support vector machines (SVM), k-nearest neighbors (KNN), decision trees (DT), naïve Bayes (NB), and logistic regression (LR). Because ensembles may integrate the benefits of numerous models, they can mitigate the limitations of each model individually and improve the overall predictive performance, making them popular in the domain of machine learning. To increase predictive performance, an ensemble hybrid approach was created by utilizing a meta-classifier to merge many base classifiers. The performance with respect to various measures of the ensemble classifier was evaluated considering a comprehensive novel dataset of ovarian cancer patients, including tumor markers as well as clinical and ultrasound features. Through extensive cross-validation studies, the hybrid model showed better prediction accuracy of 95% which is approximately 6-17% improved than the baseline classifiers and state-of-the-art ensemble approaches in predicting ovarian cancer. After comparing the performance of the ensemble classifier with other existing classifiers, the ensemble classifier outperformed the individual models and conventional diagnostic techniques in terms of sensitivity (94%) and specificity (95%) through performance evaluation.

Keywords - Ovarian Cancer, Classification, Accuracy, Data Preprocessing, Machine Learning.

I. INTRODUCTION

The human body is composed of cells, which are the smallest units. These cells generally divide and grow to form a new set of cells as our body demands. Generally, older cells die, and new cells occupy their place. Cancer begins when cells grow uncontrollably, forming a mass called a tumor. A tumor is classified as benign, which is a non-cancerous tumor, or malignant, which is a cancerous tumor. Tumors may grow in ovarian cancer and are considered to be one of the major types of cancer that can be seen in the ovaries of women, with almost 6900 new cases diagnosed every year. In women with ovarian cancer, 72% are generally diagnosed at an advanced level or stage [1]. The prognosis for ovarian cancer still remains uncertain despite improvements in treatment techniques, such as surgery, chemotherapy, and targeted medicines, especially in advanced stages when treatment options are few. Ovarian cancer (OC) is a complex disease caused by its heterogeneity, which includes a range of genetic, environmental, and lifestyle factors that influence the disease's susceptibility and course. Accurate characterization of ovarian pathology before surgery is important to guarantee that patients with ovarian cancer receive proper treatment in time [2]. Different types of tumor cells grow in the ovaries. Benign tumors do not spread and

are non-cancerous. Malignant tumors are dangerous cancer cells that require proper in-time treatment. Therefore, for any treatment of a patient with a tumor, it is important to classify ovarian tumors as benign tumors (not cancer) or malignant tumors (cancerous tumors). Different ovarian cancer types, such as epithelial ovarian cancer (EOC), metastatic, stromal cell, and germ cell ovarian cancer, are shown in **Fig 1**. Tumors are classified as either benign, borderline, or malignant based on features such as the degree to which cell division or proliferation occurs, the existence of stromal invasion, and nuclear atypia. More than 95% of OC cases are epithelial ovarian cancer.

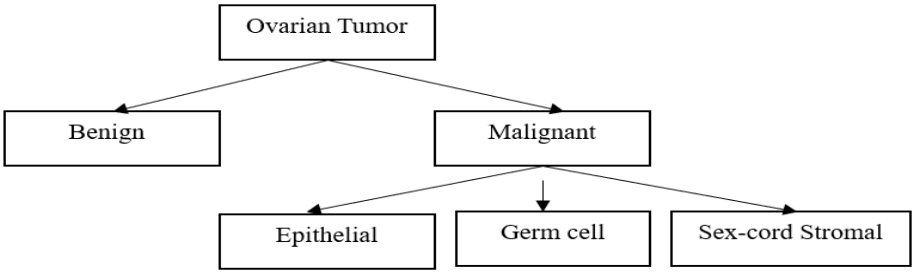


Fig 1. Ovarian Cancer Types.

Many algorithms have been developed to classify tumors as benign or malignant, but they do not show good accuracy because of the lack of symptoms associated with ovarian cancer at the early stage and the scoring systems used for classification. The International Ovarian Tumor Analysis (IOTA) is an organization that has the standard approach for ultrasound description of adnexal pathology. IOTA models provide useful tools for doctors to evaluate the potential for malignancy in ovarian tumors, marking a substantial advancement in the field of ovarian cancer prediction. These models help in decision-making for additional diagnostic and therapeutic procedures by classifying patients into low- and high-risk groups based on a combination of clinical and ultrasound-based criteria. IOTA models are strong evidence-based algorithms that incorporate multiple characteristics, including tumor morphology, vascularity, and patient demographics, to produce dependable risk estimations. These models were developed through extensive studies and validation. IOTA models provide a standardized method for ovarian tumor assessment by combining standardized protocols and criteria to improve accuracy and consistency in various healthcare settings. IOTA models provide valuable insights into treatment options and enhance diagnostic accuracy. Tumor markers such as CA125, HE4, CA 19-9 etc., and standard models such as ROMA, RMI, ADNEX, and OVA1 have been developed for the effective diagnosis of ovarian cancer. Variation in CA125 tumor markers has been found in approximately 10% to 20% of women with early stage ovarian cancer [3], and the human epididymis protein 4 (HE4) is a useful preoperative marker. High sensitivity and specificity results from CA19-9 were encouraging, indicating the test's capacity to rule out or rule out the illness. According to the ROC analysis, both CA125 and CA19-9 showed excellent diagnostic performance when used separately and in combination. [4]. The IOTA group created several popular prediction models using LR analysis, which improved the performance prediction of ultrasound diagnosis of ovarian tumor cells. These models include LR1, LR2, and IOTA ADNEX [5-7]. The important prediction models of IOTA, namely LR1 and LR2, have proven to have good diagnostic performance but misclassified fewer malignancies [8]. If the prediction of adnexal mass type is accurate, then it helps to choose an appropriate surgical method, such as immediately referring to a gynecological oncologist in case of malignant masses or laparoscopy in case of benign masses [9].

In recent years, machine learning techniques have become increasingly potent tools in the field of medical diagnostics, with the ability to identify complicated patterns within large and complex datasets. Classification is a supervised learning process that categorizes data into a set of finite classes. Classification is a method of predicting the resultant class for a given input value. To perform the prediction, the algorithm learns from a training dataset with a collection of attributes and an outcome, which is a prediction attribute [10]. Some of the simple classification techniques that are popularly used for prediction are Decision Trees DT, NB, ANN, SVM. In terms of prediction accuracy, machine learning models frequently perform better than conventional data-mining techniques. Machine learning models can identify intricate patterns in data and produce precise predictions of patient outcomes, such as the diagnosis, prognosis, and response to therapy of ovarian cancer, by utilizing advanced algorithms and optimization approaches. Ensemble classifiers are well known for their capacity to improve predictive performance through the amalgamation of the advantages of several separate models. Compared to individual classifiers, ensemble classifiers attain improved prediction accuracy by combining the predictions of many base classifiers and taking advantage of the wisdom of crowds. Ensemble classifiers can reduce the biases and errors present in individual classifiers, resulting in predictions that are more accurate and dependable. In ovarian cancer prediction, where the disease appears through a complex interaction of genetic, clinical, and demographic markers, this synergy is especially relevant.

This study proposes and evaluates an ensemble classifier to further the continuing search for a reliable and accurate method for ovarian cancer prediction. However, the ensemble classifier combines a few efficient baseline classifiers, thereby improving the prediction accuracy. The complexity and heterogeneity of the disease necessitate a multidimensional approach that transcends the constraints of single models. This study seeks to offer a thorough comprehension and prediction

of ovarian cancer through the integration of various base classifiers, each tailored to capture distinct aspects of the disease. The main contributions of this paper are, the dataset was preprocessed because it was raw and had many redundant features. The ovarian cancer dataset was classified as benign or malignant using a novel ensemble hybrid classifier specially tailored for predicting ovarian cancer by combining multiple baseline classifiers. The proposed model was compared with other baseline classifiers to verify the efficiency of the model. It is observed from the obtained results that the proposed ensemble hybrid model performed better than the other existing baseline models in terms of accuracy for classifying the ovarian dataset.

II. WORK IN THIS AREA

There are different methods to predict ovarian cancer. IOTA has developed many models and scoring systems for the diagnosis of ovarian cancer. These scoring systems use several data mining methods for prediction.

IOTA Models

Adnexal masses can be identified and classified as benign mass or malignant tumor using a variety of scoring methods and methodologies, including the risk of malignancy index (RMI) model and risk of ovarian malignancy algorithm (ROMA). The IOTA group created numerous novel algorithm-based risk-forecasting models in 2005. Simple Rules and LR sonographic features were used to create modified versions of Logistic Regression models 1 and 2 (LR1 and LR2). Compared to all earlier models, the predictive models outperformed the other models. The ADNEX model, created by IOTA in 2014, is a new model with improved performance. It is the first risk model to identify the type and stage of ovarian cancer, including borderline cancer, Stage I cancer, or stage II–IV cancer, and most importantly, whether the tumor is benign tumor or malignant tumor[11]. The ADNEX model is associated with an elevated negative predictive value and can accurately detect the features of ovarian tumors while ruling out malignancy [12]. Anton et al. [13] conducted a comparative study of different techniques that helped classify ovarian masses. They compared prediction techniques, such as CA125, ROMA, HE4, and RMI. CA125 levels are elevated only in 50% of early-stage ovarian cancers, and the threshold cut-off value is 35 U/ml. The cutoff value for HE4 was 70pM. The RMI was calculated using the ultrasound score of the patient, menopausal status, and level of the tumor marker CA125. The specificity and sensitivity of all parameters were assessed. The sensitivities of HE4, CA125, RMI, and ROMA in distinguishing malignant tumors from benign ovarian masses were 79.6%, 70.4%, 63%, and 74.1%, respectively. The sensitivity of HE4 is considered to consistently perform better in the overall evaluation of malignant ovarian tumors. ROMA [14] is an algorithm that uses HE4 and the most commonly used serum marker CA125 to classify patients with ovarian cancer into high-risk and low-risk epithelial ovarian cancer. The Receiver operator characteristic - the area under the curve (ROC–AUC) of the different tumor markers was compared, and the performance of HE4 and CA125 was the same, but for post-menopausal patients, CA125 performed better. The ROC–AUC value of HE4 was slightly higher in the premenopausal group than in the postmenopausal group. From stage I to stage IV disease, a clear trend was evident, and CA125 and HE4 were dramatically underperformed in these stages. As a result, ROMA did not perform better. An additional easy-to-use metric to eliminate the possibility of ovarian tumors is RMI. The product of the absolute value of the tumor marker CA-125, ultrasonographic score (U), and menopausal status score (M) were used to determine RMI [15]. RMI scores are four different varieties: RMI1, RMI2, RMI3, and RMI4. A small number of studies have produced findings that are essentially consistent and show that there are few statistical variations among all four categories of RMI scores [16]. The RMI approach is straightforward but offers no diagnostic benefits. The drawbacks of RMI include its inability to estimate cancer risk and its excessive reliance on serum CA-125[17].

To characterize ovarian pathology, Kaijser et al. [18] proposed the most used prediction model, RMI. RMI is the best available method of testing people with ovarian tumors and is triaged to refer to oncology units. RMI uses several ultrasound markers, such as the CA125 level, which heavily influences prediction. CA125 is a serum marker used to predict ovarian cancer. CA125 is highly expressed in malignant ovarian cancer tumors. However, as per the research data, subjective assessment by experienced doctors is better than that of CA125 serum in terms of diagnostic performance. IOTA has developed two important approaches for classifying ovarian masses. Among the two approaches, the first uses LR-1 and LR-2, which are risk prediction models. These models are regression models developed by IOTA. These models have good diagnostic performance, and the smaller number of variables used in LR2 may favour clinicians over the use of LR1, but both models seem to generate good results only for the population using which they are developed and trained. The second approach uses rules with simple descriptions that appeal to clinicians. D Timmerman et. al.[19] explained the most reliable and simple method to differentiate malignant and benign adnexal masses before surgery, which is a method of subjectively assessing ultrasound examination. According to the rules of ultrasound examination, five benign ultrasound features, such as solid areas with a maximum diameter of less than or equal to 7 mm, unilocular cyst, acoustic shadow, regular multilocular cyst with a maximum diameter of less than 100 mm, and no blood flow on color Doppler scan, and five malignant ultrasound features, such as a solid mass that is irregular, ascites, solid mass with multiple locules, which is irregular with a diameter less than 100 mm, and at least four papillary projections. This is an instant diagnostic technique; however, when the combination of benign and malignant ultrasound features is present and the absence of both benign and malignant features, the result becomes inconclusive, and a secondary test type is recommended. Nunes et al. [20] presented an analysis of simple IOTA rules that determine the risk of tumor malignancy. The results of the simple rules are then compared with historical findings and any of the pattern recognition methods. The expectation of conducting a subjective assessment is an alternative approach to simple rules. Simple rules perform slightly differently at different levels of expertise. Van Calster et al. [21]

discussed the possible practical guidance to discriminate between the different types and stages of adnexal masses by applying the model Assessment of Different NEoplasias in the adnexa of the uterus (ADNEX) developed by IOTA. The ADNEX model is a special multiclass model that differentiates the different stages of malignant tumors, such as secondary metastasis, stages 2 to 4, stage 1, or borderline cancer. The ADNEX model was developed based on a dataset of 5909 patients who were referred to undergo ultrasound examination for suspected adnexal masses between 1999 and 2007 in 24 canters in 10 countries. The ADNEX model has six ultrasound predictors, including the maximum diameter of the largest lesion, number of papillary projections, proportion of solid tissue, presence of ascites, presence of 10 or more cysts, and acoustic shadow, and three clinical predictors, such as CA125 serum marker, type of center the patient has been referred for an ultrasound, and age in years. Suyang et al. [22] examined how well the ADNEX model from the IOTA performs when combined with HE4 to diagnose ovarian cancer (OC) in its early stages. They examined receiver operating characteristic curves to determine whether using HE4 in addition to the IOTA ADNEX model improved the diagnostic accuracy compared to using IOTA ADNEX alone. Zhang et al. [23] discussed the role of tumor marker CA125 in the prediction, diagnosis, and oncogenesis of epithelial ovarian cancer. When screening for the presence of epithelial OC in women, CA125, which is also known as mucin16 (MUC16), has proven to be a reliable as well as a trustworthy biomarker. However, there has been an ongoing debate regarding the usefulness of this marker in clinical practice, especially in recent years. The FDA advises applying CA125 to track treatment response and to monitor any lingering illness or risk of recurrence following first-line therapy. Minghai et al. [24] investigated the significance of several tumor markers for the diagnosis of ovarian cancer, including CA-125, CA 19-9, CA15-3, HE4, hCG, AFP, inhibin, and LDH, as well as their effects on treatment response evaluation and disease surveillance. HE4 and CA-125 exhibit higher sensitivity and specificity, especially in early-stage detection. Furthermore, hCG is a promising prognostic biomarker that can aid in the outcome evaluation and treatment response. Combining several serum markers, such as HE4 and CA-125, improves the risk stratification and diagnostic accuracy.

Classifiers For Medical Data

Abhishek et al. [25] developed a hybrid classifier by merging decision trees and naive Bayes NB classifiers to classify the fitness data. When the dataset was classified with this hybrid classifier, the accuracy improved by 15.79% when compared with the decision tree and 3.61% when compared with Naive Bayes. To categorize ovarian cancer, k-nearest neighbor (KNN]) and support vector machine (SVM) models were applied [26]. Al Islam Bandung Hospital provided the data, which included 203 cases, 130 of which were classified as malignant and 73 as benign. The findings demonstrated that KNN outperformed SVM in terms of accuracy and F1-score, achieving 90.47% accuracy and 94.12% F1-score compared to 90.47% accuracy and 92.3% F1-score for SVM. Zeng et al.[27] discussed the idea of various data-mining and machine-learning approaches to research healthcare data. The random forest algorithm was used to classify the medical data as it avoids overfitting the data, thereby improving the prediction accuracy. Raphael et al. [28] reviewed all the existing machine learning algorithms and deep learning models used for predicting asthma in children. For the asthma dataset, it was observed that ANN and SVM performed better in terms of accuracy when compared with other models such as Naive Bayes, C4.5, KNN, Decision Tree and Random Forest. Lu et al. [29] discussed a very simple decision tree approach that can reliably distinguish between ovarian cancer and benign ovarian tumors using two important biomarkers, CEA and HE4. Using the patients' data, a machine learning method called Minimum Redundancy – Maximum Relevance (MRMR), which is a feature selection approach, was used to identify and select the most pertinent features, which were then used to build a straightforward decision tree model. Radhimeenakshi et al. [30] evaluated and compared the performances of popular ML models, namely ANN and SVM, on the statlog-based database and Cleveland HD database, which were taken from the UCI/ML Repository. The missing values in the dataset were imputed using means, medians, and mode methods. To reduce the error function, ANN uses the Gradient Descent (GD) technique. The SVM classifier uses the 2D kernel function. Chen Meng et al. [31] investigated the successful application of data mining in the prediction and diagnosis of various types of cancer. The dataset considered comprises various types of cancers along with demographic data, clinical characteristics, and treatment responses along with genomic data. Three common methods for data mining, namely association rule mining, clustering, and classification, were used to discover useful patterns. Performance measurements, including sensitivity, accuracy, F1-Score, specificity, and area under the curve (AUC), will be used to create and assess a number of prediction models, including random forests or RF, support vector machines, or SVM, as well as various deep learning networks. In the work proposed by Ismael et al. [32], a variety of datasets were utilized, incorporating clinical characteristics alongside different types of genomic features represented in various coding schemes as well as combinations thereof. Utilizing interpretable machine learning (IML) models, one linear model and another nonlinear model led to enhanced interpretability of these five important feature sets. Subsequently, nine classification techniques specially designed for binary classification were compared to assess their accuracy, Recall, and Area Under the Curve. The findings indicated that models incorporating a combination of clinical features and genomic data, particularly when gene positional coding in patients was considered, exhibited notable improvements in predictive performance. This study underscores the significance of incorporating diverse preprocessed patient data, particularly through insights provided by IML models, to aid clinicians in gaining a better understanding of the disease and facilitating informed treatment decisions.

The use of machine learning algorithms to predict the outcome of ovarian cancer was thoroughly examined by Sindhu et al. [33]. The authors investigated different machine learning approaches used for prognostic and outcome prediction in ovarian cancer patients through a comprehensive evaluation and analysis of published material. Regarding the predicted accuracy, interpretability, and clinical utility, they overcome the benefits and drawbacks of various algorithms, such as ensemble approaches, decision trees or DT, neural networks or NN, and support vector machines or SVM. The research also emphasizes the importance of combining several data sources, including imaging features, genomic profiles, and clinical indicators, to improve the capacity of machine learning models for prediction. Overall, this study provides insights into the state-of-the-art when it comes to using machine learning to treat ovarian cancer. A thorough investigation on the early detection of ovarian cancer symptoms using classifier models on imbalanced datasets was carried out and described in the publication by Ambekar, Kaul, and Hudnurkar [34]. The authors evaluated how well different classifier models handle unbalanced data, which is a frequent problem in medical diagnosis, while concentrating on feature selection approaches. They assessed the efficiency of classifiers, including NN, DT, and SVM, in correctly recognizing early stage signs of ovarian cancer through their analysis. By offering useful insights into the best method for choosing features and classifier models to increase the precision and dependability of early ovarian cancer diagnosis, these findings advance the fields of medical informatics and personalized healthcare.

Research Gaps

From the literature review of the existing ovarian cancer prediction models, it is observed that the techniques used in the prediction are not accurate because the models use only very few symptoms and family history for the prediction, whereas the clinical history is completely ignored even though clinical history plays a major role in ovarian cancer. Ovarian cancer is a disease that does not have any symptoms in the initial stage, and when the symptoms start to appear, the cancer reaches an advanced stage and treatment does not show much improvement. This results in the wrong prediction where the benign tumors are predicted as malignant and given appropriate treatments such as surgery, and the patient had to suffer due to incorrect predictions. Considering these drawbacks in the existing systems, a new hybrid ensemble classifier has been developed that uses all the important parameters related to ovarian cancer that were determined in consultation with a medical practitioner to accurately classify the tumor as benign or malignant.

III. PROPOSED METHODOLOGY

The ovarian dataset was preprocessed using existing techniques to prepare the data for further analysis to predict ovarian cancer. The ensemble classifier was then designed and developed using existing baseline classifiers, and the performance of the proposed model was compared with other baseline and ensemble classifiers.

Dataset Description

The dataset was obtained from (Cancer Data Access System) after obtaining approval from the NCI to test the developed system. The dataset is obtained only after signing the ethical clearance; hence, the name and personal details of the patient are hidden and replaced with the patient ID. The study population included patients admitted to the hospital for the removal of an adnexal mass between 2017 and 2022.

Table 1. Variables Considered for The Work

Table	**Feature**	**Description**
Medical complications	Canctype	Type of tumor – Benign / Malignant
Screening	Ca125_level	CA125 tumor marker value
	Ca19-9_level	CA19-9 tumor marker value
	Numcyst	Number of cysts present
Screening_abnormalities	Bilateral	whether lesion appears both left and right or in any one side
	Ascite	Presence of ascites
	Locules	uni-/multilocularity
	Metastasis	If the cancer is a metastasis one or originated in the ovary
	Solid	Presence of solid area in abnormality.
Main	Post_menopausal	If the patient was post-menopausal at trial entry
	Fh_cancer	Whether there exists a family history of cancer
	Age	Age of the patient
	ph osumm trial	Did the patient have a personal history of ovarian or any other type of cancer prior to this?

While creating the database, patients provided written informed consent. The dataset was raw, and there were four tables in the dataset: main table, medical complications table, screening table, and screening abnormalities table. The primary table, denoted as the main table, contains a singular record for each participant. However, the secondary tables include multiple entries per participant. The details of the variables related to the ovarian tumors considered in this study are given in **Table 1**.

The number of patients in each table differs based on the diagnostic procedures and treatments that the patient has undergone. The details of the tables present in the datasets are listed in **Table 1**. In the subsequent step, the main table is fused with the secondary tables using a LEFT OUTER JOIN operation. Not all participants in the main table had corresponding entries in the secondary tables, as shown in **Table 2**. Therefore, the LEFT OUTER JOIN operation leads to a significant number of discarded NULL entries. This reduced the dataset by over 49%, from 78,209 samples to 39,990 samples.

Table 2. Dataset Description

Table	Entries	Entries common in main table
Main	78,209	78,209
Medical_complications	68,100	39,990
Screening	1,50,993	40,721
Screening_abnormalities	60,186	39,992

Data Preprocessing

Prior to the model development, extensive preprocessing was performed to ensure the quality and consistency of the dataset. This is very important for preprocessing the dataset before applying any data-mining algorithms to achieve good accuracy in terms of mining results. The dataset considered was raw and had missing values for some of the attributes that were addressed through various imputation techniques such as mean, median, mode, and advanced imputation methods such as multiple imputation methods. Z-score normalization was applied to ensure a mean value of 0 and a standard deviation of 1. Feature selection was essential to ensure an improvement in the dataset's discriminatory strength for the ensemble classifier. Filter methods, such as correlational analysis and wrapper methods, such as forward and backward feature selection methods, have been applied to select the optimal features from the dataset, which has a greater impact on the prediction of ovarian cancer. Resampling was performed with regards to the "canctype" feature using stratified sampling to ensure the balanced distribution of the target class in the dataset so that benign and malignant cases are balanced preserving all the important information of the dataset. PCA was performed using two components to aid effective data visualization. A variety of statistical techniques, including feature importance from individual base models and correlation analysis, were used to determine which biomarkers and clinical features provided the most information for classifying the ovarian tumor dataset as benign or malignant. Each feature in the data is given a score based on the importance of the feature. The greater the score, the more significant or pertinent the feature is to your output variable. The feature importance plot is shown in **Fig 2**. The important parameters obtained from the feature importance method were again validated with the help of a clinical expert to understand the impact of every feature considered in the prediction of ovarian cancer. Along with the features selected through feature selection methods, few important parameters for the prediction system were carefully selected in consultation with a gynaecological expert, and from the literature which majorly contributes to the prediction of malignancy in ovarian tumors. For every parameter, the Gini impurity was computed, which estimated the likelihood of misclassifying the data. The preprocessed dataset is then divided into training and testing data using a 5-fold cross-validation technique to minimize overfitting and evaluate the classifier's generalization performance.

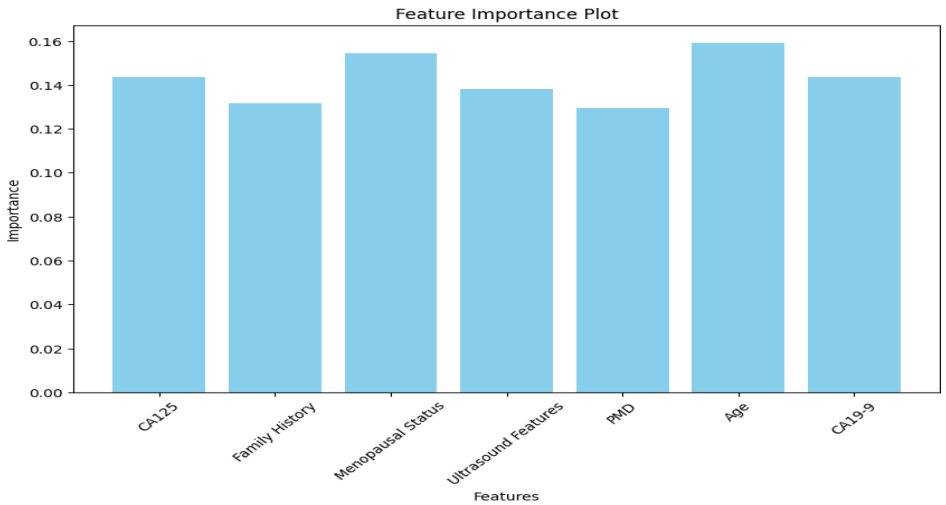


Fig 2. Feature Importance Plot.

Baseline Classifiers Used

A detailed literature review was conducted by implementing various classification models on the Ovarian Cancer dataset and other healthcare datasets [35], and the models that performed better on the dataset were considered for constructing a hybrid classifier. Various Machine learning models used in the development of the ensemble hybrid model were analyzed for their behavior with respect to the ovarian cancer dataset considered. A variety of base classifiers, such as decision trees (DT), support vector machines (SVM), K Nearest neighbor (KNN), naive Bayes (NB), and logistic regression (LR), were used in the construction of the hybrid ensemble classifier. The robustness and generalization performance are improved by the SVM by finding the ideal hyperplane that maximizes the margin between classes. Because NB require little training data and are computationally efficient, they can be used in applications with limited computational resources or training sample sizes. The KNN can be used to categorize patients according to their comparable feature profiles, such as clinical aspects or ultrasound parameters. When dealing with noisy data and extraneous characteristics, the DT is resilient and can handle both continuous and categorical features. To maximize the performance of each individual model, a pre-processed dataset was used to train each baseline classifier. For the experimental analysis, classifiers were trained and tested with 80:20, 50:50, and 70:30 data distributions; however, 80:20 performed better than other distributions and hence was considered for the classification using the ensemble hybrid classifier.

Ensemble Hybrid Model Construction

Hybrid ensemble classifiers combine the strengths and capabilities of several base classifiers to create a more reliable and accurate predictive model. To construct the hybrid ensemble classifier, the algorithm was designed with two-layered estimators. The first layer of the algorithm consists of multiple baseline classification models that are used to predict the outputs of the ovarian tumor dataset considered for testing the models. The second layer has a meta-classifier, which considers the predictions made by all the baseline classification models of the first layer as input and generates a new set of predictions. The algorithm was then modified to include various combinations of the baseline classifier with the meta-classifier, as shown in **Fig 3**. In the meta-classifier for logistic regression, the probability distribution function for the model parameters was obtained using the Bayesian technique. The ensemble classifier classifies data points as benign or malignant.

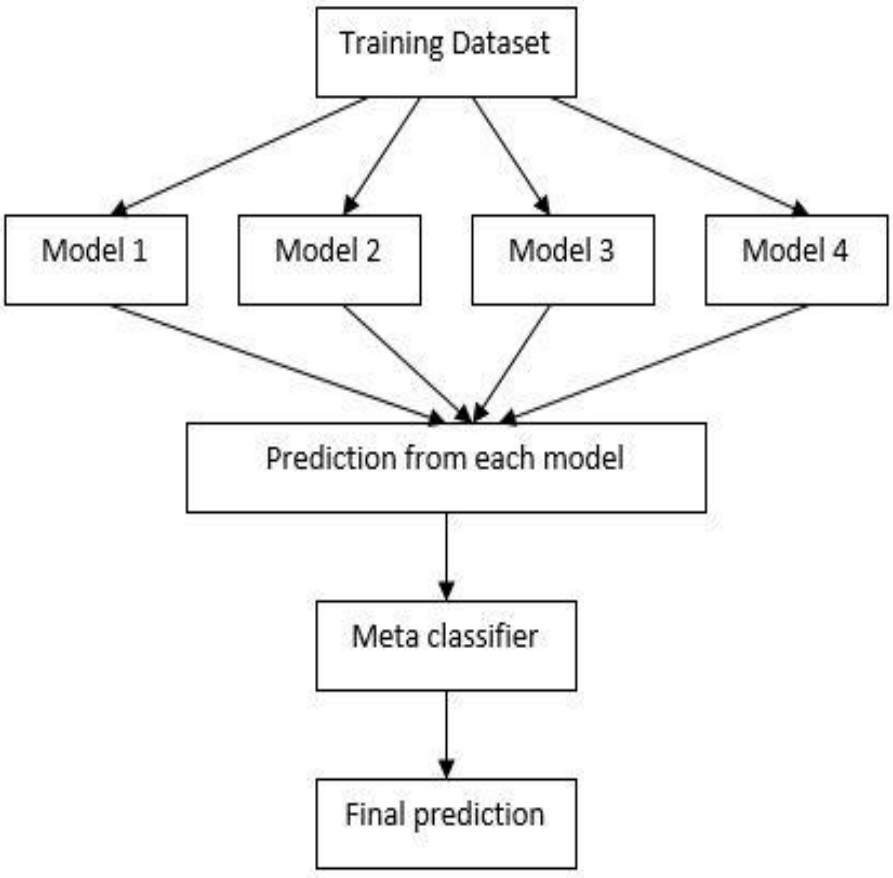


Fig 3. Proposed Methodology.

The algorithm for the ensemble classifier is given in Algorithm 1.

Algorithm 1: Hybrid ensemble Classifier

Input: Training data D = {(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)}
Base models {M_1, M_2, ..., M_k}
Meta-model M_{meta}
New data x_{new}
Output: final_prediction - classes

Step 1: Train Base Models
for each base model M_i in {M_1, M_2, ..., M_k} do
Mi.fit(D)
end for

Step 2: Generate Meta-Features
Initialize empty dataset D_{new} = [ ]
for each example (x_j, y_j) in D do
P = [ ]
for each base model M_i in {M_1, M_2, ..., M_k} do
P.append(M_i.predict(x_j))
end for
D_{new}.append((P, y_j))
end for
Step 3: Train Meta-Classifier
M_{meta}.fit(D_{new})
Step 4: Final Prediction
P_{new} = [ ]
for each base model M_i in {M_1, M_2, ..., M_k} do
P_{new}.append(M_i.predict(x_{new}))
end for
final_prediction = M_{meta}.predict(P_{new})

Different combinations of the baseline classifier and meta-classifier were tested in two layers of the ensemble classifier to obtain accurate predictions, as presented in **Table 3**.

Table 3. Accuracy Of Ensemble Classifier With Various Combinations Of Baseline Classifiers And Mata Classifier

Baseline classifiers	**Meta classifier**	**Accuracy (%)**
KNN, DT	LR	82
KNN, NB	LR	92
DT, NB	LR	82
KNN, SVM	LR	85.8
DT, SVM	LR	94.83
SVM, NB	LR	87.083

IV. EXPERIMENTAL ANALYSIS & RESULTS

IBM statistics software (SPSS for Windows version 20.0) was used for statistical analysis. Tumor ultrasonographic characteristics, ovarian cancer symptoms, a description of the clinical and demographic parameters, and tumor marker analysis were all performed in this study. Popular analysis techniques, namely the chi-square test and Fisher's exact test, were used for categorical data values, while the Mann-Whitney U test was used when the data values were continuous. All comparisons were statistically significant at $p < 0.05$.

The ensemble algorithm was compared with the existing classifiers, namely, SVM, LR, DT, Random Forest (RF), extreme gradient boosting (XgBoost), adaptive boosting (AdaBoost), and bagging algorithms to compare the accuracy. The entire dataset is split into two parts: training data and testing data in a ratio of 80:20 with 5-fold cross-validation. Cross validation is a method of assessing the overall performance and robustness of the model, and it is important in the hybrid model that helps in understanding and estimating how well the model will generalize new unseen data when multiple classifiers are combined into a meta model. 5-fold cross validation is used here, wherein the model is trained and tested five times, each time using a separate fold of data for testing when the remaining folds are considered for training. After training

the model, it was tested with a testing dataset to measure the accuracy and other performance metrics, such as sensitivity R, specificity S, precision P, and F-measure F1, which were calculated using the equations 1-6.

$$\text{Accuracy: } A = (\text{True}_{PP} + \text{True}_{NN}) / (\text{True}_{PP} + \text{True}_{NN} + \text{False}_{PP} + \text{False}_{NN}) \quad (1)$$

$$\text{Misclassification: } M = 1 - \text{accuracy} \quad (2)$$

$$\text{Sensitivity (recall): } R = \text{True}_{PP} / (\text{True PP} + \text{False}_{NN}) \quad (3)$$

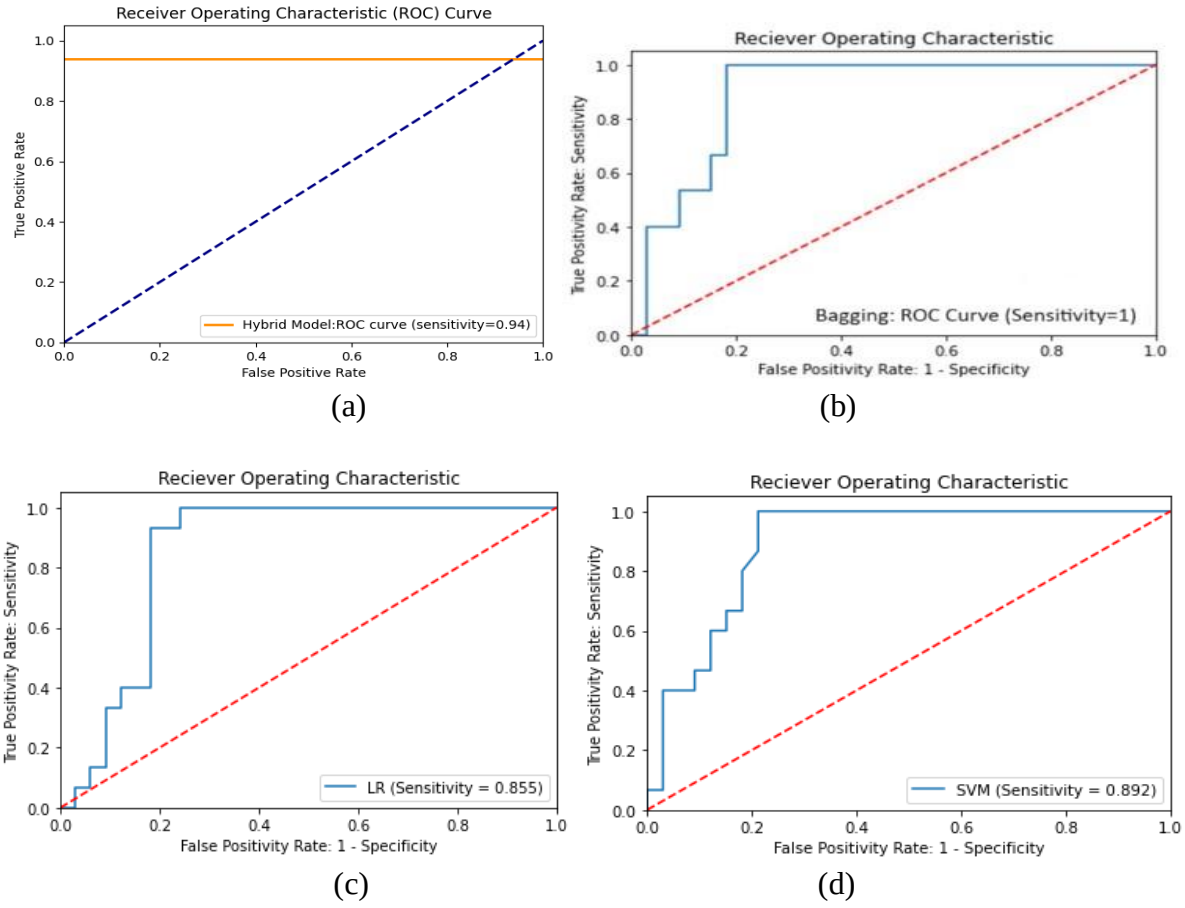
$$\text{Specificity: } S = \text{True}_{NN} / (\text{True}_{NN} + \text{False}_{PP}) \quad (4)$$

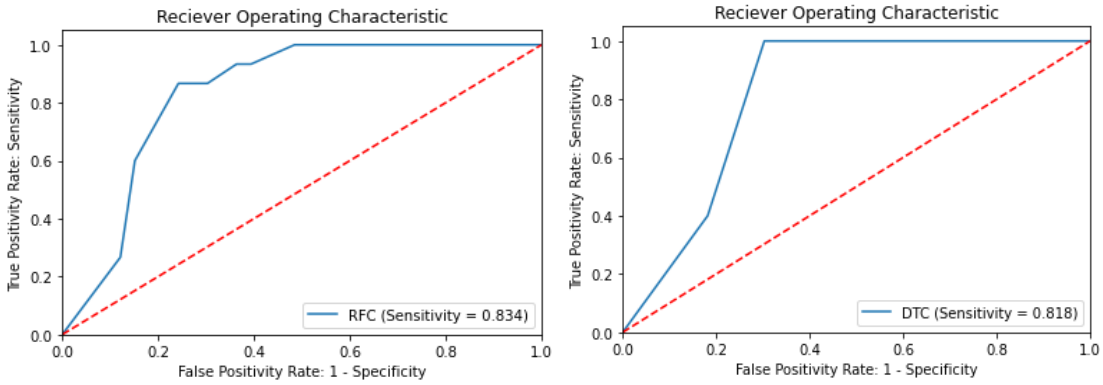
$$\text{Precision: } P = \text{True}_{Pos} / (\text{True}_{PP} + \text{False}_{PP}) \quad (5)$$

$$\text{F-measure: } F1 = 2 \times (\text{recall} \times \text{precision}) / (\text{Precision} + \text{Recall}) \quad (6)$$

where True_{PP} - True Positive (correctly classified positive instances), True_{NN} -True Negative (correctly classified negative instances), False_{PP} - False Positive (incorrectly classified positive instances), and False_{NN} - False Negative (incorrectly classified negative instances).

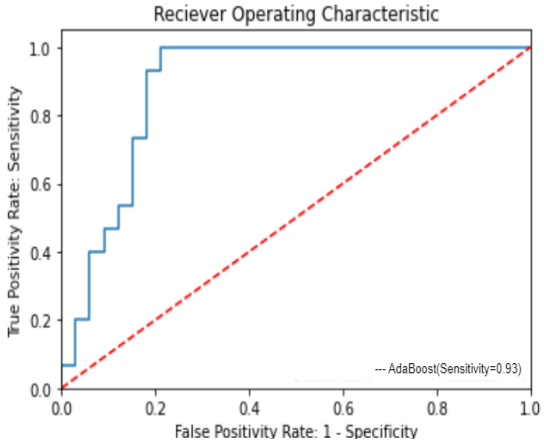
A classification model's performance across different thresholds is shown graphically by the Receiver Operating Characteristic (ROC) curve. For various threshold settings, it plots the true positive rate (TPR) against the false positive rate (FPR). ROC curves for the proposed hybrid model along with other classifiers are shown in **Fig 4**. The sensitivity achieved from the ROC graph for the hybrid model is 0.94.





(e)

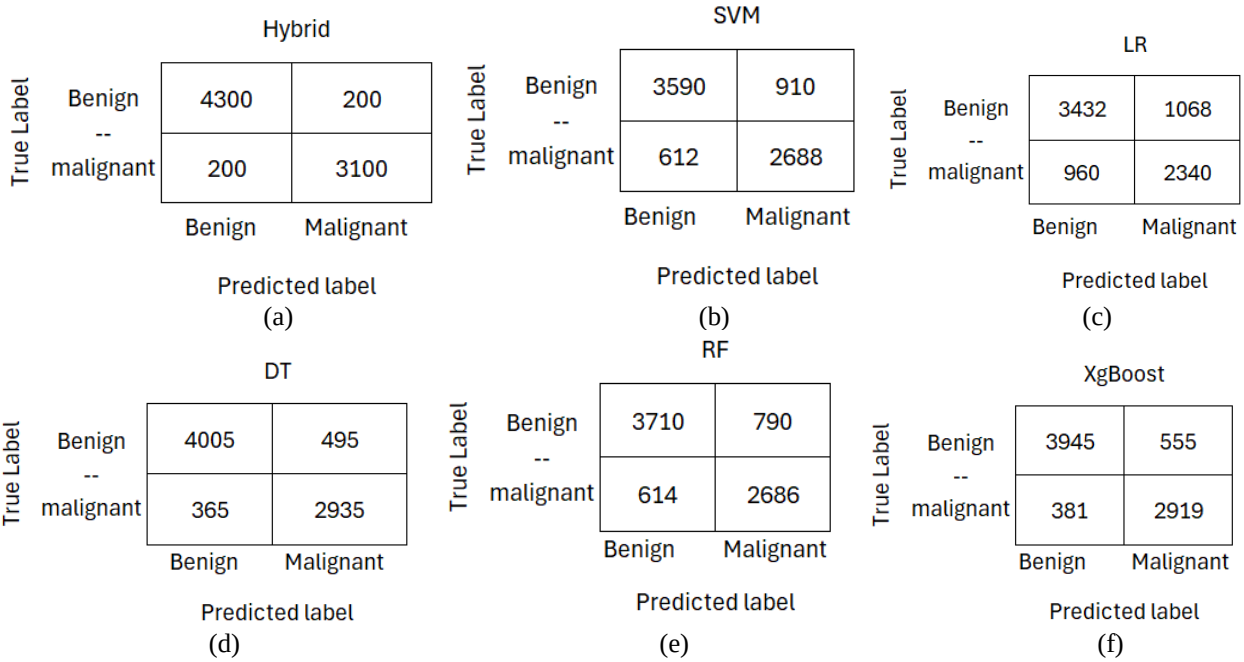
(f)

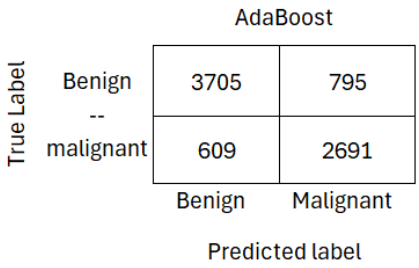


(g)

Fig 4. ROC Curve Of All The Classifiers Along With The Hybrid Classifier.

Confusion matrices offer a thorough analysis of how the model's predictions and the real class labels such as benign or malignant compare. To evaluate the performance of the classification model, examine the TP, TN, FP, and FN counts. **Fig 5** gives the confusion matrix of predicted values with actual values when the dataset is classified using all the classifiers along with the hybrid classifier.





(g)

Fig 5. Confusion Matrix Of All The Classifiers.

The confusion matrix would enhance understanding of the classifier's accuracy in ovarian cancer classification by clearly understanding how the actual values are predicted. Table 4 provides a comparative analysis of the hybrid ensemble classifier with other existing classifiers for all three variations of the developed scoring system.

Table 4. Comparison Of Hybrid Classifier with Other Existing Classifiers

	Hybrid model	SVM	LR	DT	RF	XgBoost	AdaBoost
Accuracy	0.95	0.8	0.74	0.89	0.82	0.88	0.82
Misclassification	0.05	0.2	0.26	0.11	0.18	0.12	0.18
Sensitivity	0.94	0.892	0.855	0.818	0.83	0.82	0.93
Specificity	0.95	0.72	0.89	0.93	0.66	0.46	0.66
Precision	0.94	0.72	0.89	0.93	0.66	0.46	0.66
F-Measure	0.94	0.79	0.79	0.89	0.77	0.59	0.77

From **Table 4**, it is observed that the ensemble hybrid classifier performed well when compared with other existing classifiers in terms of the various performance evaluation methods used.

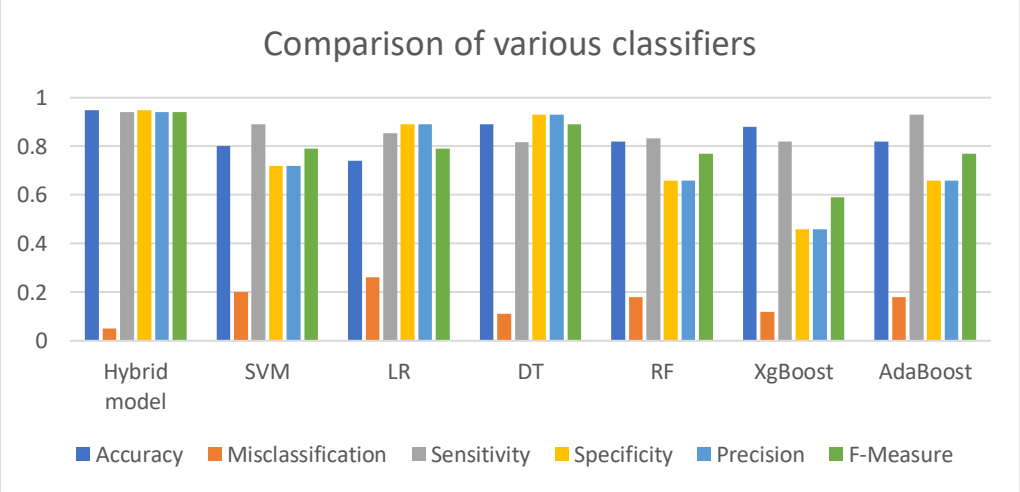


Fig 6. Comparison of Various Classifiers.

Fig 6 shows the comparison of various classifiers with respect to the features considered for the work. The important features listed in **Table 1** are considered to measure the performance of all the classifiers. However, Hybrid model performed better with an accuracy of 95% when compared with other models when tested on the test data with the same set of features.

The trade-off between recall which is the proportion of relevant instances that have been recovered over the total number of relevant instances and precision which is the fraction of relevant instances among the anticipated positives is depicted by precision-recall curves. The curve's many operational points correspond to various cutoff values for determining whether an instance is positive or negative. The precision-recall curve to highlight the hybrid model's precision and recall trade-offs is shown in **Fig 5**.

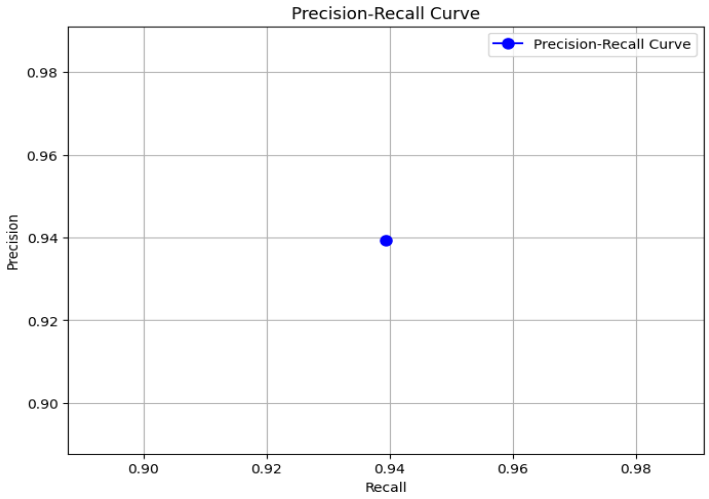


Fig 5. Precision-Recall Curve.

Several important aspects contribute to the ensemble classifier's performance in classifying ovarian cancers as benign or malignant, proving its superiority over other classifiers in predicting ovarian cancer. The ovarian dataset used in this study places greater emphasis on the characteristics that have the greatest influence on the incidence of ovarian cancer: family history, symptoms such as postmeal distention, menopausal status, ultrasonography factors, and biomarkers such as CA125 and CA19-9. By using a meta-classifier layer that dynamically balances the predictions of different classifiers, the ensemble hybrid classifier enables the model to adjust to the varied characteristics of ovarian cancers and increase the overall accuracy.

The proposed model can be used to classify the data of patients who visited the hospital for the removal of adnexal masses. The mass could be benign or malignant; hence, the data of the patient were used by the proposed model to clearly predict if the mass was benign or malignant before surgery, which helped the doctor avoid unnecessary complications of surgery and its consequences. [36-38]

V. CONCLUSION

To overcome the drawbacks of single methods and advance the field of ovarian cancer prediction, a novel ensemble classifier was developed that combines the best features of multiple models. Because ovarian cancer is very subtle and is frequently detected at a later stage because there are no early signs, predictive models that are both accurate and novel are required. Owing to its late-stage detection, high death rate, and clinical heterogeneity, ovarian cancer poses a serious global health concern. Owing to its complexity, ovarian cancer requires highly developed predictive models that can integrate various data sources and accurately represent multiple characteristics of the disease. Therefore, creating hybrid ensemble classifiers seems to be a viable way to improve the predictability, robustness, and interpretability of ovarian cancer predictions. After extensive testing and analysis, our ensemble classifier outperformed the separate base models and conventional diagnostic techniques in terms of performance. Because of the model's high sensitivity and specificity, it may be used to detect ovarian cancer early and provide focused, timely treatment measures. Because sampling is performed on the result variable of the dataset, there is no observable bias in the evaluation process. However, when the imputed values vary significantly outside of the limits during preprocessing, the results might be biased. The hybrid ensemble classifier can revolutionize the prediction of ovarian cancer with further improvement in the consideration of borderline cases and validation, leading to better patient outcomes and a more efficient method of treating this disease.

Author Contributions

Idea, R.S., G.M., U.D.A., S. G., (Roopashri Shetty, Geetha M., U. Dinesh Acharya, Shyamala G); design, implementation and paper writing R.S., (Roopashri Shetty); review and editing, G.M., U.D.A., S.G.,(Geetha M., U. Dinesh Acharya, Shyamala G.); supervision, and project administration, G.M., U.D.A., S.G.,(Geetha M., U. Dinesh Acharya, Shyamala G.). All authors have read and agreed to the published version of the manuscript.

Funding

This research receives no external funding.

Ethical Statement

The authors confirm that the research conducted for this study complies with the relevant ethical guidelines and regulations. The dataset considered in the work has been obtained after required ethical clearance.

Conflicts of Interest
The authors declare no conflicts of interest.

Funding
No funding agency is associated with this research.

References
[1]. M. Kalaiyarasi and H. Rajaguru, “Performance Analysis of Ovarian Cancer Detection and Classification for Microarray Gene Data,” BioMed Research International, vol. 2022, pp. 1–16, Jul. 2022, doi: 10.1155/2022/6750457.
[2]. J. C Citadel, J. F. Paula, F. Jacques, V. Jerome, L. Mathieu, L. Lihu, A. Maaike, B. Freddie, and S. Isabelle, “Ovarian cancer today and tomorrow: A global assessment by world region and human development index using globocan 2020”, International Journal of Cancer, vol. 151, no. 9, pp. 1535–1541, 2022.
[3]. T. K. Burki, “CA-125 blood test in early detection of ovarian cancer,” The Lancet Oncology, vol. 16, no. 6, p. e269, Jun. 2015, doi: 10.1016/s1470-2045(15)70237-8.
[4]. F. T. Ali et al., “Sensitivity and specificity of microRNA-204, CA125, and CA19.9 as biomarkers for diagnosis of ovarian cancer,” PLOS ONE, vol. 17, no. 8, p. e0272308, Aug. 2022, doi: 10.1371/journal.pone.0272308.
[5]. D. Timmerman et al., “<scp>ESGO</scp>/<scp>ISUOG</scp>/<scp>IOTA</scp>/<scp>ESGE</scp> Consensus Statement on preoperative diagnosis of ovarian tumors,” Ultrasound in Obstetrics & Gynecology, vol. 58, no. 1, pp. 148–168, Jun. 2021, doi: 10.1002/uog.23635.
[6]. D. Timmerman et al., “Logistic Regression Model to Distinguish Between the Benign and Malignant Adnexal Mass Before Surgery: A Multicenter Study by the International Ovarian Tumor Analysis Group,” Journal of Clinical Oncology, vol. 23, no. 34, pp. 8794–8801, Dec. 2005, doi: 10.1200/jco.2005.01.7632.
[7]. D. Timmerman et al., “Predicting the risk of malignancy in adnexal masses based on the Simple Rules from the International Ovarian Tumor Analysis group,” American Journal of Obstetrics and Gynecology, vol. 214, no. 4, pp. 424–437, Apr. 2016, doi: 10.1016/j.ajog.2016.01.007.
[8]. J. Koneczny, A. Czekierdowski, M. Florczak, P. Poziemski, N. Stachowicz, and D. Borowski, “The use of sonographic subjective tumor assessment, IOTA logistic regression model 1, IOTA Simple Rules and GI-RADS system in the preoperative prediction of malignancy in women with adnexal masses,” Ginekologia Polska, vol. 88, no. 12, pp. 647–653, Dec. 2017, doi: 10.5603/gp.a2017.0116.
[9]. J. A. Hall, “Adnexal Masses: When to Observe, When to Intervene, and When to Refer,” Obstetrics & Gynecology, vol. 116, no. 2, p. 440, Aug. 2010, doi: 10.1097/aog.0b013e3181ea4f15.
[10]. K. Deeba and B. Amutha, “Classification Algorithms of Data Mining,” Indian Journal of Science and Technology, vol. 9, no. 39, Oct. 2016, doi: 10.17485/ijst/2016/v9i39/102065.
[11]. B. Van Calster, K. Van Hoorde, and W. Froyman W,” Practical guidance for applying the ADNEX model from the IOTA group to discriminate between different subtypes of adnexal tumors.”, Facts Views Vis Obgyn. Vol. 7, pp. 32-41, 2015.
[12]. K. G. Araujo et al., “Performance of the <scp>IOTA ADNEX</scp> model in preoperative discrimination of adnexal masses in a gynecological oncology center,” Ultrasound in Obstetrics & Gynecology, vol. 49, no. 6, pp. 778–783, Apr. 2017, doi: 10.1002/uog.15963.
[13]. A. Cristina., “A comparison of CA125, HE4, risk ovarian malignancy algorithm (ROMA), and risk malignancy index (RMI) for the classification of ovarian masses”, ultrasound obstet gynecol, vol. 31, pp. 681-690, 2008.
[14]. Â. Melo, R. Veríssimo, M. Farinha, N. N. Martins, and F. N. Martins, “Discriminative value of CA-125, HE4, Risk of Malignancy Index II (RMI-II) and Risk of Malignancy Algorithm (ROMA) in the differential diagnosis of pelvic masses: conclusions from a referral Centre in Portugal,” Journal of Obstetrics and Gynaecology, vol. 38, no. 8, pp. 1140–1145, Jun. 2018, doi: 10.1080/01443615.2018.1457632.
[15]. E. Aktürk et al., “Comparison of four malignancy risk indices in the detection of malignant ovarian masses,” Journal of Gynecologic Oncology, vol. 22, no. 3, p. 177, 2011, doi: 10.3802/jgo.2011.22.3.177.
[16]. A. Borges et al., “2022-RA-727-ESGO Performance of the IOTA ADNEX model in differentiating between benign and malignant adnexal lesions in a Portuguese population,” Diagnostics, Oct. 2022, doi: 10.1136/ijgc-2022-esgo.157.
[17]. B. Van Calster et al., “Validation of models to diagnose ovarian cancer in patients managed surgically or conservatively: multicentre cohort study,” BMJ, p. m2614, Jul. 2020, doi: 10.1136/bmj.m2614.
[18]. J. Kaijser et al., “Improving strategies for diagnosing ovarian cancer: a summary of the International Ovarian Tumor Analysis (<scp>IOTA</scp>) studies,” Ultrasound in Obstetrics & Gynecology, vol. 41, no. 1, pp. 9–20, Dec. 2012, doi: 10.1002/uog.12323.
[19]. D. Timmerman et al., “Simple ultrasound-based rules for the diagnosis of ovarian cancer,” Ultrasound in Obstetrics & Gynecology, vol. 31, no. 6, pp. 681–690, May 2008, doi: 10.1002/uog.5365.
[20]. N. Nunes, G. Ambler, X. Foo, J. Naftalin, M. Widschwendter, and D. Jurkovic, “Use of <scp>IOTA</scp> simple rules for diagnosis of ovarian cancer: meta-analysis,” Ultrasound in Obstetrics & Gynecology, vol. 44, no. 5, pp. 503–514, Oct. 2014, doi: 10.1002/uog.13437.
[21]. B. Van Calster, K. Van Hoorde, W. Froyman, J. Kaijser, L. Wynants, C. Landolfo, C. Anthoulakis, I. Vergote, T. Bourne, D. Timmerma, “Practical guidance for applying the ADNEX model from the IOTA group to discriminate between different subtypes of adnexal tumors”, Facts Views Vis Obgyn, vol 7, no. 1, 32-41, 2015.
[22]. S. Yang et al., “Performance of the IOTA ADNEX model combined with HE4 for identifying early-stage ovarian cancer,” Frontiers in Oncology, vol. 12, Sep. 2022, doi: 10.3389/fonc.2022.949766.
[23]. M. Zhang, S. Cheng, Y. Jin, Y. Zhao, and Y. Wang, “Roles of CA125 in diagnosis, prediction, and oncogenesis of ovarian cancer,” Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, vol. 1875, no. 2, p. 188503, Apr. 2021, doi: 10.1016/j.bbcan.2021.188503.
[24]. A. Matsas et al., “Tumor Markers and Their Diagnostic Significance in Ovarian Cancer,” Life, vol. 13, no. 8, p. 1689, Aug. 2023, doi: 10.3390/life13081689.
[25]. A. Lal and C. R. S. Kumar, “Hybrid classifier for increasing accuracy of fitness data set,” 2017 2nd International Conference for Convergence in Technology (I2CT), Apr. 2017, doi: 10.1109/i2ct.2017.8226326.
[26]. V. V. P. Wibowo, Z. Rustam, S. Hartini, F. Maulidina, I. Wirasati, and W. Sadewo, “Ovarian cancer classification using K-Nearest Neighbor and Support Vector Machine,” Journal of Physics: Conference Series, vol. 1821, no. 1, p. 012007, Mar. 2021, doi: 10.1088/1742-6596/1821/1/012007.
[27]. Y. Zeng and F. Cheng, “Medical and Health Data Classification Method Based on Machine Learning,” Journal of Healthcare Engineering, vol. 2021, pp. 1–5, Nov. 2021, doi: 10.1155/2021/2722854.
[28]. R. H. Ekpo, V. C. Osamor, A. A. Azeta, E. Ikeakanam, and B. O. Amos, “Machine learning classification approach for asthma prediction models in children,” Health and Technology, vol. 13, no. 1, pp. 1–10, Jan. 2023, doi: 10.1007/s12553-023-00732-8.
[29]. M. Lu et al., “Using machine learning to predict ovarian cancer,” International Journal of Medical Informatics, vol. 141, p. 104195, Sep. 2020, doi: 10.1016/j.ijmedinf.2020.104195.

[30]. S Radhimeenakshi, "Classification and prediction of heart disease risk using data mining techniques of support vector machine and artificial neural network", In 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), p. 3107–3111. IEEE, 2019.

[31]. F. C. Meng, Dr. B. Amiya, and Dr. D. Urmisha,"Predictive Diagnosis through Data Mining for Cancer Detection and Treatment", Onomázein, vol. 61, pp. 276–283, 2023.

[32]. I. Gómez-Talal, A. Barquín, L. Bote-Curiel, M. Yagüe-Fernández, J. L. Rojo-Álvarez, and J. García-Donás, "A machine learning approach to predicting ovarian cancer: Integrating clinical and genetic data," Science Talks, vol. 10, p. 100319, Jun. 2024, doi: 10.1016/j.sctalk.2024.100319.

[33]. S. Sindhu, D. Hemavathi, K. Sornalakshmi, G. Sujatha, and S. Srividhya, "A Comprehensive Study on the Application of Machine Learning Algorithms in the Prognosis of Ovarian Cancer", The Open Biomedical Engineering Journal, vol. 17, no. 1, 2023.

[34]. S. Ambekar, N. Kaul, M. Hudnurkar, and Niharika, "Feature Selection based Performance Comparison of Classifier Models for an Imbalanced Dataset: Early Diagnosis of Symptoms for Ovarian Cancer," 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Dec. 2023, doi: 10.1109/icses60034.2023.10465317.

[35]. R. Shetty, M. Geetha, D. U. Acharya, and G. Shyamala, "A Comparative Study of Efficient Classification Models," Recent Advances in Artificial Intelligence and Data Engineering, pp. 441–447, Nov. 2021, doi: 10.1007/978-981-16-3342-3_35.

[36]. Sudhi M, Shukla VK, Shetty DK, Gupta V. Advancements in bladder cancer management: a comprehensive review of artificial intelligence and machine learning applications. Engineered Science. 2023;26.

[37]. R. Sankaranarayanan, K. S. Umadevi, N. Bhavani, B. M. Jos, A. Haldorai, and D. V. Babu, "Cluster-based attacks prevention algorithm for autonomous vehicles using machine learning algorithms," Computers and Electrical Engineering, vol. 101, p. 108088, Jul. 2022, doi: 10.1016/j.compeleceng.2022.108088.

[38]. A. Kadekar et al., "An Overview of Yeast-Based Toxicology: Unraveling Molecular Mechanisms, Advancing Drug Discovery, and Addressing Translational Complexities," Engineered Science, 2023, doi: 10.30919/es1063.