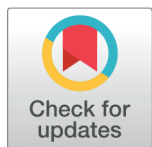


ORIGINAL ARTICLE



OPEN ACCESS

Received: 14/08/2026

Accepted: 25/08/2026

Published: 24/09/2026

Citation: Gajendran G, Kurus Malai Selvi F (2026) Clinical Weighting and Symptom Interaction-based Risk-Stratified Learning for Early Prediction of Ovarian Cancer Using Explainable Machine Learning. Indian Journal of Science and Technology 19(33): 2328-2343. <https://doi.org/10.17485/IJST/v19i33.1179>

* **Corresponding author.**

gomathivasanthg@gmail.com

Funding: None

Competing Interests: None

Copyright: © 2026 Gajendran & Kurus Malai Selvi. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](https://www.indjst.org/))

ISSN

Print: 0974-6846

Electronic: 0974-5645

Clinical Weighting and Symptom Interaction-based Risk-Stratified Learning for Early Prediction of Ovarian Cancer Using Explainable Machine Learning

Gomathi Gajendran^{1*}, F Kurus Malai Selvi¹

¹ PG & Research Department of Computer Science, Government College for Women (A), Kumbakonam - 612 001. Affiliated to Bharathidasan University, Tiruchirapalli

Abstract

Objectives: To develop an explainable Clinical Weighting and Symptom Interaction-based Risk-Stratified Learning (CWSI-RSL) framework for early ovarian cancer prediction using clinical feature weighting, symptom interactions, and patient risk-level classification. **Method:** 349 records, 51 attributes ovarian cancer dataset was used. Missing values were treated and duplicate records were handled. Numerical data were normalised. Mutual Information (MI) and Recursive Feature Elimination (RFE) were used for important features selection. Clinical weights and Symptom Interaction Feature Expansion (SIFE) features were then added, and low-, medium- and high-risk grouping was performed. Logistic Regression (LR), Support Vector Machine (SVM), Naïve Bayes (NB) and K-Nearest Neighbours (KNN) were optimised using GridSearchCV by 80:20 train-test splitting. The performance achieved was compared against previously reported traditional machine-learning and ensemble baselines using accuracy, precision, recall and F1-score, and SHAP (SHapley Additive exPlanations) was employed to interpret the model predictions. **Findings:** The accuracy, precision, recall, and F1-score of CWSI-RSL were 92.40%, 90.30%, 95.00% and 92.60%, respectively. It performed 5.26 percentage points better than the best reported conventional model, Naïve Bayes (87.14%) and 5.40 percentage points better than the best reported ensemble model, Boosting (87.00%). The recall of 95.00% shows that the model was able to detect most of the ovarian cancer cases, and the F1-score of 92.60% shows a balanced performance. Menopause status, CA125-related interactions, and neutrophil-based interactions were identified as important predictors by SHAP analysis. These results demonstrate that combining clinical weighting, symptom interactions and risk grouping can improve prediction and provide useful explanations. **Novelty:** CWSI-RSL integrates clinical weighting, symptom interaction, model-derived risk grouping, optimised machine learning and SHAP explainability in a unified ovarian cancer prediction framework.

Keywords: Ovarian Cancer Prediction; Clinical Feature Weighting; Symptom Interaction Modeling; Risk Stratification; Explainable Machine Learning

1 Introduction

Machine learning (ML) has gained its greater implication for the prediction of ovarian cancer. The focus has moved from traditional classification to early prediction, risk stratification, feature selection, ensemble learning and explainable artificial intelligence (XAI). However, incorporating these capabilities into a clinically meaningful predictive framework remains difficult. Arafat et al.⁽¹⁾ suggested an explainable fuzzy Random Forest model optimised by Harris Hawks algorithm for early ovarian cancer prediction integrating feature selection and explainability. Clinical data were used by Asif et al.⁽²⁾ to propose an explainable ensemble learning method, showing the potential of ensemble models for ovarian cancer diagnosis. Similarly, an interpretable ML study reported the use of demographic, clinical, and biochemical parameters for the differentiation of borderline and malignant ovarian tumours⁽³⁾. These studies demonstrate the potential of optimised and interpretable ML to improve ovarian cancer prediction. However, they mainly focus on improving the classifier performance or interpreting the resulting model instead of explicitly modelling clinical feature importance and interactions among symptoms and biomarkers.

The early prediction and risk evaluation are also paid more and more attention. Gui et al.⁽⁴⁾ explored early ovarian cancer prediction and risk stratification based on clinical data and ML approaches, which directly reflected the value of risk-based prediction. Kucukakcali et al.⁽⁵⁾ investigated comparison-based explanations for early clinical detection using explainable ML. These studies support the importance of clinical data and explainability in early prediction, but risk stratification and model explanation are not necessarily combined with an explicit system for weighting clinically important features and generating interaction-based clinical trends.

Also, significant improvements have been achieved with biomarker- and feature-based methods. Sun and Wen⁽⁶⁾ reviewed ML diagnostic models for ovarian tumours and commented on the importance of clinical and tumor-marker data. Abuzinadah et al.⁽⁷⁾ combined ensemble classification with explainable AI based on SHAP, showing that the model's predictions can be explained in terms of understandable feature contributions. Huang et al.⁽⁸⁾ specifically tackled the problem of explainable biomarker identification using machine learning and Shapley analysis. These analyses contribute to the knowledge of important predictors, but feature importance is typically obtained as a model understanding after prediction, not as an obvious clinical weighting mechanism before classification.

Another approach has been considered for risk modelling. Comes et al.⁽⁹⁾ proposed an explainable ensemble model for ovarian cancer risk prediction in BRCA-mutated patients, and Boehm et al.⁽¹⁰⁾ showed that multimodal data integration can improve risk stratification of advanced serous ovarian cancer. These works emphasize the importance of patient-level risk assessment yet are based only on certain patient populations or multimodal data and do not explicitly combine clinical weighting, symptom interaction generation and risk stratification into a unified framework. Moreover, ML was proved to be useful for ovarian cancer diagnosis by Taleb et al.⁽¹¹⁾, which further certifies the prospect of ML-based clinical prediction.

But these advances leave several important gaps. First, a lack of explicit modelling of the interaction between symptoms and biomarkers. Many existing approaches use selected clinical variables as individual predictors which may not reflect clinically meaningful relationships between symptoms, physiological measurements and tumour biomarkers. Second, the weighting of clinical features is not typically directly incorporated into the prediction pipeline. Existing techniques for feature selection and explainability in AI (e.g. XAI) do find important variables, but do not explicitly weight the selected features with clinically meaningful weights prior to training the model. Third, risk stratification is often performed independently from feature representation and classification, instead of in the same learning model. Fourth, while SHAP and other XAI techniques increase the transparency of models, they generally explain the model outputs and do not generate the clinically organized interaction features. These limitations prevent existing methods from concurrently capturing the importance of symptoms, the clinical relationships and the patient level risk severity.

Hence, a clinically applicable framework is needed that can integrate these complementary components rather than dealing with them as separate stages. Here, this paper proposes Clinical Weighting and Symptom Interaction-based Risk-Stratified Learning (CWSI-RSL) to address this gap. The framework employs a publicly available ovarian cancer clinical dataset containing 349 patient records and 51 attributes including demographic, biochemical, haematological, physiological and tumour biomarker variables⁽¹²⁾. First, the informative features are identified using Mutual Information (MI) and Recursive Feature Elimination (RFE). Then the chosen features are given weights according to their clinical importance in diagnosis. This paper proposes a Symptom Interaction Feature Expansion (SIFE) module to learn pairwise interaction features to model interactions between clinically relevant symptoms and biomarkers. These weighted and interaction-enhanced features are then used to compute patient-level risk scores and classify patients into low, medium and high-risk groups.

The developed feature representation is evaluated using Logistic Regression, Support Vector Machine (SVM), Naïve Bayes and K-Nearest Neighbours (KNN) with hyperparameters tuned by GridSearchCV. Model performance is evaluated by accuracy, precision, recall, and F1-score and compared with previously published conventional and ensemble ML methods⁽¹³⁾,⁽¹⁴⁾. Finally,

SHAP (SHapley Additive exPlanations) is used to interpret the contribution of clinical and interaction features to the final predictions.

Thus, this study aims to build a precise, risk-based and explainable early ovarian cancer prediction framework to jointly consider the importance of clinical features, the relationships between symptoms and biomarkers, and the risk stratification of patients. The main novelty of CWSI-RSL is the combination of clinical weighting, interaction-based feature expansion, risk stratification, optimised ML classification and SHAP-based explainability, in a single prediction framework, overcoming the limitations of methods that treat these elements separately.

2 Methodology

This section details the proposed Clinical Weighting and Symptom Interaction-based Risk-Stratified Learning (CWSI-RSL) approach for early ovarian cancer prediction. The proposed system aims to improve the predictive performance by combining the clinical feature importance, modelling the interaction of symptoms, stratifying the risk-level, machine learning optimisation, and explainable artificial intelligence. The workflow includes preprocessing of data, feature selection, clinical weighting, symptom interaction, risk stratification, classification, explainability and performance evaluation for improving the accuracy and interpretability of ovarian cancer prediction.

Algorithm 1 combines clinical feature importance, symptom interactions and risk analysis to predict early ovarian cancer. The dataset is initially preprocessed by eliminating duplicates, dealing with missing values and normalising data. Then, clinical features are selected using Mutual Information and Recursive Feature Elimination (RFE) and clinical importance weights are assigned to emphasise diagnostic symptoms.

Symptom Interaction Feature Expansion (SIFE) generates new features by combining important pairs of symptoms in an effort to capture hidden relationships. The enhanced features are used to calculate the risk scores and classify the patients into low-, medium- and high-risk groups. Several machine learning algorithms are optimised, trained and evaluated and the best model is selected for ovarian cancer prediction. Finally, SHAP explains the model predictions and the performance is assessed by accuracy, precision, recall and F1-score.

Algorithm 1: Clinical Weighting and Symptom Interaction-based Risk-Stratified Learning (CWSI-RSL) for Early Ovarian Cancer Prediction

Input:

D = Ovarian cancer dataset

Output:

Best prediction model (B)
Predicted class (Cancer / Non-cancer)
Risk level (Low / Medium / High)
Performance metrics

Begin

1. Data Collection

Load ovarian cancer dataset D

2. Data Preprocessing

Remove duplicate records from D
Handle missing values in D
Normalize numerical features in D

3. Feature Selection

Apply Mutual Information on D
Rank all features
Apply Recursive Feature Elimination (RFE)
Select top important features
Store selected features in F_s

4. Clinical Symptom Weighting

```

For each feature  $f_i$  in  $F_s$  do
  Assign clinical weight  $w_i$ 
  Compute weighted feature:
     $f_i' = w_i \times f_i$ 
End For
Store weighted features in  $F_w$ 

```

5. Symptom Interaction Feature Expansion (SIFE)

```

Initialize interaction feature set  $F_i$ 

Generate clinically relevant interaction features
For each selected symptom pair  $(f_i, f_j)$  do
  Generate interaction feature:
     $f_{ij} = f_i \times f_j$ 
  Add  $f_{ij}$  into  $F_i$ 
End For

Merge  $F_i$  into  $F_w$ 
Store expanded feature set in  $F_e$ 

```

6. Risk-Level Stratification

```

For each patient record  $r$  in  $F_e$  do
  Compute RiskScore:
    RiskScore = Sum(all feature values)
End For

Compute dynamic thresholds:
  Q1 = Lower quantile of RiskScore
  Q2 = Upper quantile of RiskScore

For each RiskScore do
  If RiskScore  $\leq$  Q1 then
    Assign Low Risk
  Else if RiskScore  $\leq$  Q2 then
    Assign Medium Risk
  Else
    Assign High Risk
  End If
End For

```

7. Hyperparameter Optimization

```

For each classifier  $C$  in
  {Logistic Regression, SVM, Naïve Bayes, KNN} do
  Apply GridSearchCV on  $C$ 
  Select optimal hyperparameters
End For

```

8. Classification

```

Train all optimized classifiers
Generate predictions for each classifier

```

9. Best Model Selection

Evaluate all classifiers
Select model B having highest Accuracy

10. Final Prediction

Use best model B
Predict ovarian cancer class

11. Explainability (XAI)

Apply SHAP on best model B
Compute feature importance scores
Generate prediction explanations

12. Evaluation

Compute:
Accuracy
Precision
Recall
F1-score
End

2.1 Dataset Collection

The ovarian cancer clinical dataset utilised in this study includes 349 patient records and 51 attributes representing demographic, biochemical, haematological, and tumour biomarker information. The dataset was downloaded from the publicly available Kaggle repository and has been extensively used in ovarian cancer prediction research⁽¹²⁾. It has clinically relevant ovarian cancer biomarkers namely CA125, HE4, CA19-9, CA72-4 and CEA. It also has other physiological and laboratory measures such as liver function tests, kidney function measures, measures of electrolyte levels and blood count measures.

Dataset accessibility and class distribution: The ovarian cancer dataset is publicly available from the “Predict Ovarian Cancer” Kaggle repository (<https://www.kaggle.com/datasets/saurabhshahane/predict-ovarian-cancer>). The dataset was downloaded on 20 August 2026 and consists of 349 patients’ records where 178 (51.00%) benign ovarian tumours (BOT) in non-ovarian-cancer class and 171 (49.00%) ovarian cancer (OC) cases were recorded. The target variable for model development was a binary outcome, TYPE where TYPE = 1 for BOT and TYPE = 0 for OC. It is the reason why the dataset has a relatively balanced distribution of classes, with a difference of only 7 samples between the two classes, and the dataset is suitable for supervised binary classification.

These biomarkers are medically important because ovarian cancer is often a silent disease, and early symptoms such as bloating, pelvic pain, feeling full quickly and urinary urgency are often ignored or mistaken for normal health conditions.

The dataset provides a binary classification setting where TYPE = 1 represents BOT (non-ovarian-cancer class) and TYPE = 0 represents OC. The same target coding was used in all stages of preprocessing, model building, prediction and evaluation. The CWSI-RSL framework uses the combined use of demographic, physiological, haematological and tumour-biomarker variables to provide a diverse clinical feature space for studying ovarian cancer classification.

2.2 Data Preprocessing

Data pre-processing enhances the quality of data before machine learning. In this process, the duplicate records of patients are removed to decrease the redundancy and bias. This ensures that each patient record is utilised effectively for model training. Next, the missing values of the clinical attributes are handled by suitable imputation techniques. It avoids significant distortion and preserves valuable information and improves the quality of dataset.

Finally, Min-Max scaling is used for feature normalisation to scale all numerical features within 0 to 1. Normalisation avoids dominance of large-scale features and ensures balanced learning across all attributes as clinical features are measured in different units and scales. This step also improves the convergence behaviour and stability of machine learning algorithms. For example, one dataset can contain two data entries for the same patient with the same CA125 and HE4 values.

Invalid records are removed; missing values (e.g. HE4) are filled using imputation and features (e.g. CA125) are normalised using Min-Max scaling during pre-processing. This ensures that the data set is clean, complete and ready for accurate machine learning analysis.

Train-test partition: The preprocessed data was split into 80% training and 20% testing data. The training subset was used to develop features selection, clinical weighting, interaction features construction, hyperparameter optimisation and classifier training, while the test subset was left as unseen data for the final performance evaluation. For all classifiers the same partitioning strategy was used to allow a fair comparison. Stratified partitioning was used to preserve the proportion of cancer and non-cancer cases between the training and testing subsets where applicable.

Preprocessing parameters: Duplicate records were detected and removed from the data set prior to model construction based on the available patient-record features. Missing numerical values were handled using median imputation, where each missing numerical value was substituted by the median of the respective feature computed from the training data. This imputation process was done before feature scaling. Then numerical variables were converted using Min-Max normalisation of the form $x' = (x - x_{\min}) / (x_{\max} - x_{\min})$ producing values in the range of 0-1. The same preprocessing parameters were used for the test data as for the training data, including the training-set medians and minimum-maximum values, to avoid data leakage.

2.3 Feature Selection

Feature selection is performed to find out the most relevant clinical indicators for ovarian cancer prediction by eliminating irrelevant and redundant features. The efficiency of the model is improved and the computational complexity is reduced. First, we quantify the statistical dependency between each feature and the target variable by mutual information (MI). Features with a higher MI score are believed to be more informative for classification, being more strongly associated with the ovarian cancer outcomes.

Next, Recursive Feature Elimination (RFE) is applied to remove less important features and select the most relevant clinical variables, thereby improving prediction accuracy and computational efficiency.

For example, CA125 and HE4 may be assigned a high rank by Mutual Information and less relevant features may be assigned low scores. RFE then discards features that are not important and keeps important features such as CA125, HE4, age, menopause status and neutrophil count to enhance the prediction accuracy and model efficiency.

Feature-selection settings: The initial relevance ranking method was Mutual Information (MI) followed by Recursive Feature Elimination (RFE) on the ranked feature set. The number of features retained by RFE was fixed to 10, following the experimental setup used in this study. Feature selection was done only on the training data to reduce chances of information leakage from test set. The resulting selected feature subset was then used for clinical weighting and interaction-feature creation⁽¹⁵⁾.

2.4 Clinical Symptom Weighting

Clinical weighting basis: Clinical feature weighting was used to skew the importance of variables to be more diagnostically informative for ovarian cancer. This is supported by recent machine-learning studies of ovarian cancer, which show the importance of clinical, laboratory and tumour-biomarker features for prediction and risk assessment^{(4), (5), (6), (7), (8), (16)}. In CWSI-RSL, the feature weights are directly built into the input representation prior to classification, instead of being used solely for post-hoc interpretation. For each selected feature f_i , the weighted feature f_i' is calculated as:

$$f_i' = w_i \times f_i$$

Where w_i is the clinical weight assigned to feature f_i .

Weight derivation: The clinical weights were derived from the feature-importance scores from the feature-selection/model-development process. Let I_i be the initial importance score of a selected feature f_i . The importance scores were normalised to obtain the corresponding clinical weights w_i preserving the relative importance of the selected features. Features that are relatively more important are given higher weights and features that are less important are given lower weights. The resulting w_i values were then used to calculate the weighted-feature value $f_i' = w_i \times f_i$. Hence, the weights were calculated from the relative contribution of the picked features and not assigned manually arbitrarily.

For example, if CA125 is 150 and the weight is 0.9, then the weighted value is $150 \times 0.9 = 135$. If HE4 = 200 and weight = 0.8, then its weighted value = $200 \times 0.8 = 160$. If sodium is 140 with a lower weight of 0.3, its weighted value is $140 \times 0.3 = 42$. That is, a larger weight means the feature is more influential in the prediction, a smaller weight means it is less influential.

2.5 Symptom Interaction Feature Expansion (SIFE)

Ovarian cancer is dependent on the individual symptoms and the interaction of multiple clinical factors. To capture these relationships, the Symptom Interaction Feature Expansion (SIFE) mechanism produces pairwise feature interactions by multiplying selected feature pairs. These interaction features capture nonlinear relationship between clinical variables such as combined effect of multiple biomarkers. The increased feature space allows the model to learn more complex disease patterns, leading to an improved detection of subtle signals of early-stage ovarian cancer.

Interaction-selection criteria: SIFE combined only the most relevant features selected through MI and RFE to generate interaction features. To represent the relationships between the clinical variables, we multiplied each pair by $f_i \times f_j = f_{ij}$. This helped to focus on meaningful interactions among symptoms, demographic, physiological, haematological and tumour-biomarker features for further risk scoring and classification^{(17), (18), (19), (20)}.

For example, if CA125 = 120 and Age = 50, then the interaction feature is 6000. Neutrophil=7.5 and HE4=180, interaction value = 1350, similarly. These interaction features make the model learn some hidden relationships among clinical factors and improve the accuracy of cancer prediction.

2.6 Risk-Level Stratification

A Risk Score is calculated from weighted and interaction-based features for improving clinical interpretability. Patients are then stratified into low, medium and high-risk groups according to quantile thresholds of this score. Low risk is represented by lower scores, and high risk is represented by higher scores for the risk of cancer.

This stratification adds to the clinical decision support layer by allowing the physicians to rank their patients by severity levels. It enhances interpretability by transforming numerical predictions into clinically meaningful risk categories^{(17), (16)}.

Risk-threshold determination: The risk score was computed based on the weighted and interaction enhanced feature representation. As no fixed clinical thresholds were defined for the current dataset, data-driven quantile cut-offs were used to split the measured risk-score distribution into three model-derived risk categories. Q1 was the lower quantile threshold defining the low-risk group and Q2 was the upper quantile threshold defining the medium- vs. high-risk group. Low-risk group included patients with $\text{RiskScore} \leq Q1$, medium-risk group included patients with $Q1 < \text{RiskScore} \leq Q2$. The categories are data-derived model results and are not intended to be interpreted as clinically validated risk thresholds. This approach is in line with recent machine learning studies in ovarian cancer for risk prediction and stratification^{(4), (9), (10)}.

For example, three patients after feature weighting, interaction expansion. Patient A is assigned a Risk Score of 0.18 which is below the lower quantile cutoff (Q1) and is so labelled as Low Risk which means that the patient has a very low risk of ovarian cancer. Patient B has a Risk Score of 0.52 which is between Q1 and Q2 and thus classified as Medium Risk, showing moderate clinical indicators that needs to be further monitored. Patient C has a Risk Score of 0.89, which is above the upper quantile threshold (Q2) and classified as High Risk, i.e., a high probability of ovarian cancer and so needs urgent clinical attention and detection confirmation.

2.7 Hyperparameter Optimization

Using GridSearchCV, the proposed approach tunes the hyperparameters of Logistic Regression, SVM, Naive Bayes, and KNN to improve the performance and generalisation of the model. It systematically tests different combinations of parameters and uses cross-validation to find the best configuration for each classifier.

The best set of hyper-parameters is chosen to maximise the classification accuracy, so that each model is working at its best. This step greatly improves the predictive stability and reduces the risk of underfitting or overfitting.

Hyperparameter search and validation: A GridSearchCV was run with 5-fold cross-validation, using only the training data. Parameter grids were defined separately for Logistic Regression, SVM, KNN and Naïve Bayes based on the hyperparameters pertaining to each classifier. Logistic Regression was optimised using its regularisation parameter C, SVM was evaluated using candidate regularisation and kernel settings including the radial basis function (RBF) kernel, KNN was evaluated using $k=3, 5, 7$ and 9, and Naïve Bayes was optimised using its model specific smoothing parameter. The final model was trained with the combination of parameters that had the highest mean cross-validation accuracy. The test set was held out completely from the hyperparameter optimisation and used solely to evaluate the final performance.

2.8 Classification

The chosen machine learning algorithms are then trained on the expanded feature set developed from clinical weighting and symptom interaction modelling after optimisation. The data is divided into train and test sets to objectively assess the performance of the model.

The trained models then make predictions on the unseen test data, predicting patients as cancer or non-cancer. The multiple classifiers allow us to compare different learning strategies like probabilistic, linear and distance-based methods. This multi-model method guarantees a strong evaluation and provides insights on the most appropriate algorithm for prediction of ovarian cancer.

Following training, each classifier makes a different prediction for the patient. Logistic Regression relies on feature relationships. SVM uses decision boundaries. KNN uses nearest neighbours. Naïve Bayes is based on probability distributions. Different predictions are evaluated to select the best performing model^{(17), (16)}.

Classifier comparison: Logistic Regression, SVM, Naïve Bayes and KNN were selected as they utilise different classification methods. The same preprocessed, weighted and interaction-enhanced features were used in all four models, and evaluated using the same train-test split and metrics for a fair comparison of the proposed CWSI-RSL approach.

Complementary supervised learning classifiers such as Logistic Regression, SVM, Naïve Bayes, and KNN were used as they represent different learning principles and have been utilised in ovarian cancer prediction and diagnosis research. Previous studies in ovarian cancer have evaluated conventional machine learning classifiers and comparative learning strategies for cancer prediction^{(11), (13), (14)}. In this study, the same selected, clinically weighted and interaction-enhanced feature representation was used for these classifiers under identical training, optimisation and testing conditions, to give a regulated evaluation of the proposed CWSI-RSL framework.

2.9 Baseline and Comparative Evaluation

To contextualise the performance, the results of the proposed CWSI-RSL framework were compared with the machine-learning and ensemble approaches reported previously for the prediction of ovarian cancer^{(13), (14)}. Traditional approaches are Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Logistic Regression (LR), Naïve Bayes (NB), K-Nearest Neighbours (KNN). Ensemble approaches are Max Voting, Stacking, Bagging, and Boosting^{(13), (14)}. These approaches are relevant reference points as they are representative of common individual and ensemble learning strategies for ovarian cancer prediction.

The comparison was performed using the same evaluation measures as for CWSI-RSL, i.e. accuracy, precision, recall and F1-score. The experimental procedure reported in this study was used to evaluate CWSI-RSL while previously reported results were used as common-dataset reference values. CWSI-RSL also includes MI-RFE, clinical feature weighting, SIFE, risk stratification and hyperparameter optimisation prior to final classification.

The comparison is used as a performance benchmark rather than a direct head-to-head evaluation, as differences in preprocessing, feature selection, hyperparameter settings, and evaluation methods may affect the results. Hence, the results are carefully discussed to show how CWSI-RSL performed in comparison to previous reports on the same dataset.

2.10 Ablation Study Design

An ablation-based internal analysis was developed to understand whether the individual components of CWSI-RSL lead to the prediction performance. The complete CWSI-RSL model was used as the reference configuration, and the main components were removed one by one, keeping the other experimental conditions unchanged. This paper evaluated the following configurations: (i) baseline selected features without clinical weighting or interaction expansion, (ii) feature selection with clinical weighting but without Symptom Interaction Feature Expansion (SIFE), (iii) feature selection with SIFE but without clinical weighting, (iv) the weighted and interaction-enhanced representation without risk stratification, and (v) the entire CWSI-RSL configuration. The same training/testing partition was used to evaluate each configuration, and the same accuracy, precision, recall, and F1-score measures were used. This internal control was included to check if the performance improvement was due to the combined CWSI-RSL components and not to classifier selection alone.

2.11 Best Model Selection

The optimised classifiers were tested on the predefined test set with the performance indicators of accuracy, precision, recall and F1-score. The classifier with the best classification accuracy under the same experimental conditions was selected as the best

performing model. The experimental evaluation led to the selection of SVM as the final classifier for CWSI-RSL. The selected model was subsequently used for final ovarian cancer class prediction and SHAP-based explainability analysis.

2.12 Explainability Using SHAP

SHAP (Shapley Additive Explanations) is used to explain the best-performing model by showing the contribution of each feature to the final prediction. SHAP identifies the most important features affecting ovarian cancer prediction and explains the model's decisions, which enhances interpretability and clinical trust. SHAP shows how features such as CA125, menopause status, and neutrophil interactions increase cancer prediction, while normal liver enzyme levels decrease it. The key factors influencing the prediction are the contributions of CA125 interactions and menopausal status, which can help clinicians understand and trust the model's decision.

2.13 Clinical usefulness and reliability of CWSI-RSL

This study investigates the clinical utility of CWSI-RSL by converting heterogeneous clinical measures into a structured representation that reflects feature importance, clinically relevant interactions, and patient-level risk categories. The risk-stratification component does not output a binary cancer/non-cancer result, but rather low-, medium- and high-risk groups, allowing for prioritisation for further clinical assessment. The SHAP analysis also gives feature level explanations which allow the identification of important clinical and interaction variables that contribute to individual predictions.

To improve accuracy of early prediction, many controlled steps are taken: removing duplicate records, handling missing values, normalising numerical data, performing MI-RFE feature selection, fixing the train-test split, performing cross-validated hyperparameter optimisation, comparing multiple classifiers, and evaluating using multiple complementary metrics for performance. The use of an unseen test set for final evaluation reduces the possibility to evaluate the performance only on data used for model development. In addition, the ablation analysis provides an internal control to assess the contribution of clinical weighting, interaction expansion and risk stratification.

2.14 Performance Evaluation

The performance of the proposed system is evaluated using standard classification metrics like Accuracy, Precision, Recall and F1-score. These metrics provide a complete evaluation of predictive performance. Accuracy is a measure of overall correctness.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision and recall are measures of false positive and false negative rates respectively.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

The F1-score balances precision and recall.

$$F1 = 2 * \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Taken together, all these metrics provide a comprehensive and robust assessment of the proposed system. The overall workflow of the proposed CWSI-RSL system is a systematic process comprising of dataset collection, data preparation, feature selection, clinical weighting, symptom interaction feature expansion, risk stratification, hyperparameter optimisation, classification, best model selection, explainability analysis, and performance evaluation. Figure 1 represents the flowchart of the proposed CWSI-RSL approach.

These components constitute a clinically interpretable machine learning pipeline that enhances early ovarian cancer prediction with improved accuracy and interpretability.

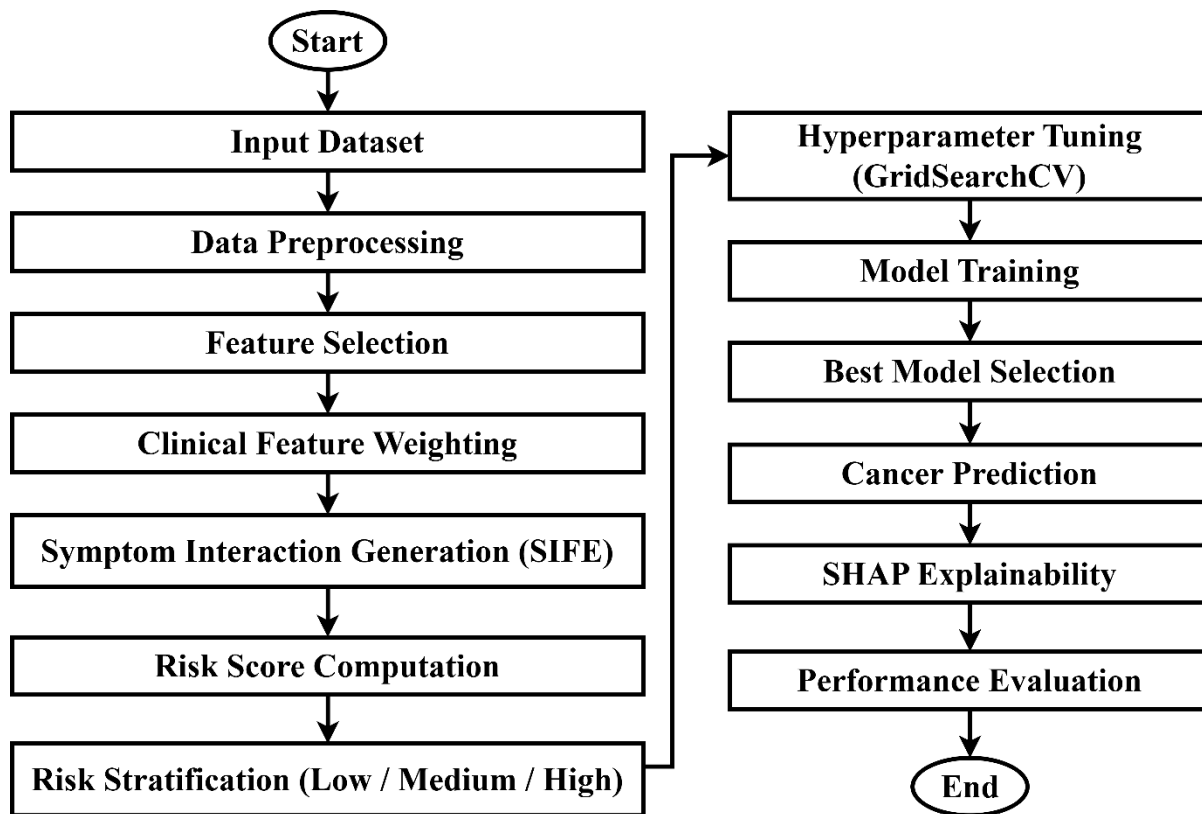


Fig 1. Flowchart of proposed CWSI-RSL approach

3 Results and Discussion

The proposed CWSI-RSL approach was evaluated on an ovarian cancer clinical dataset for early prediction, and compared with traditional machine learning and ensemble models in terms of accuracy, precision, recall and F1-score. The results of the evaluation indicate that the proposed CWSI-RSL approach improves ovarian cancer prediction and outperforms traditional machine learning and ensemble models.

3.1 Experimental Setup

The experiments have been carried out in the Python programming environment because of its wide support of machine learning, data pre-processing, optimisation and explainability analysis. The full implementation was done by Spyder integrated development environment (IDE) within Anaconda 3 platform that provides effective package management and scientific computing libraries. Different Python libraries were used like NumPy for numerical computation, Pandas for manipulating data, Scikit-learn to develop machine learning algorithms, GridSearchCV for hyperparameter optimisation, SHAP for explainability analysis.

The dataset was preprocessed by removing duplicates, missing values and Min-Max scaling. Mutual Information and Recursive Feature Elimination (RFE) are used for feature selection. The data set was split into training and testing sets with a ratio of 80:20 to validate the model. Multiple classifiers were trained and optimised under similar experimental conditions to ensure fair comparison, namely Logistic Regression, Support Vector Machine, Naive Bayes and K-Nearest Neighbours.

The relative performance of CWSI-RSL compared to the reported baseline and ensemble models in terms of accuracy, precision, recall and F1-score are shown in Figures 2–5. The baseline results were taken from the corresponding published studies^{(13), (14)}. The baseline results of SVM, RF, DT, LR, NB, and KNN were taken from O. Ikani et al., and the ensemble learning results of Max Voting, Stacking, Bagging, and Boosting were taken from S. L. JM and P. Subbulakshmi^{(13), (14)}. The CWSI-RSL model was performed better with an accuracy of 92.4%, precision of 90.3%, recall of 95.0% and F1-score of 92.6%, which outperformed all the existing models.

The prediction of CWSI-RSL is better than the traditional and ensemble models reported before on the common dataset. The improvement can be attributed to the incorporation of clinical weighting, SIFE, risk stratification and classifier optimisation. Stratifying risk level is clinically relevant and SHAP provides explainability. Together, these building blocks lead to better accuracy and interpretability than standard models.

3.2 Discussion

The accuracy comparison of all baseline models, ensemble models and the proposed CWSI-RSL is shown in Figure 2. Accuracy is the overall correctness of the classification. It is the total number of correct predictions. The visual comparison clearly shows the highest accuracy of 92.4% for the proposed CWSI-RSL model and it is better than all the comparative methods such as Naive Bayes (87.14%) and Boosting (87%).

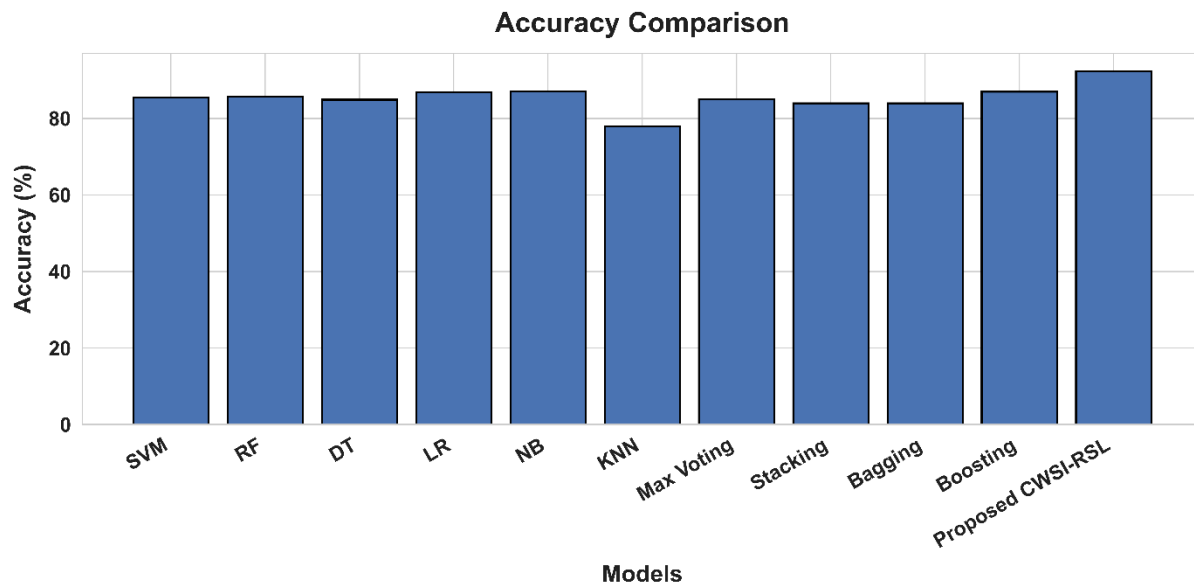


Fig 2. Accuracy Comparison

The high accuracy of CWSI-RSL is due to its ability to enhance clinically important features through weighted representation and derive additional interaction-based clinical knowledge. Traditional models generally only rely on the direct features. The proposed method increases the feature space by including the symptom interactions, which increases the capability of the model to differentiate cancer and non-cancer cases more efficiently. This richer feature representation impacts directly to better overall prediction correctness.

The models are evaluated and compared in terms of the precision shown in Figure 3. Precision is the ratio of true positives to the total number of positive predictions. The proposed CWSI-RSL obtained a precision of 90.3% which is higher than all the baseline and ensemble approaches, showing its capability to reduce false positive predictions.

The proposed model improves the precision with a clinical weighting mechanism that emphasises important bio-markers and reduces irrelevant noise. This means that predictions will be more accurate and the chance of misclassifying cancers will be lessened. High precision is important to avoid unnecessary anxiety and medical procedures. Figure 4 compares the recall values for all the models. Recall is the ability to identify true positive ovarian cancers. The proposed CWSI-RSL gained the highest recall of 95.0%, which presented great sensitivity in discovering cancer.

The proposed model improves recall by symptom interaction modelling and risk level stratification. SIFE enables to identify hidden clinical patterns and risk stratification allows to accurately identify high-risk patients. Figure 5 shows comparison of F1-score for all models. The proposed CWSI-RSL achieved the highest F1-score of 92.6%, indicating a balanced classification performance.

The high F1-score represents a good trade-off between false positive and false negative of the proposed model. CWSI-RSL achieves reliable performance by feature weighting, interaction-based learning and classifier optimisation. SHAP summary plot is shown in Figure 6 to explain the contribution of clinical and interaction features in the prediction of ovarian cancer. SHAP values indicate whether features support the prediction of cancer or non-cancer and the colours indicate the feature values.

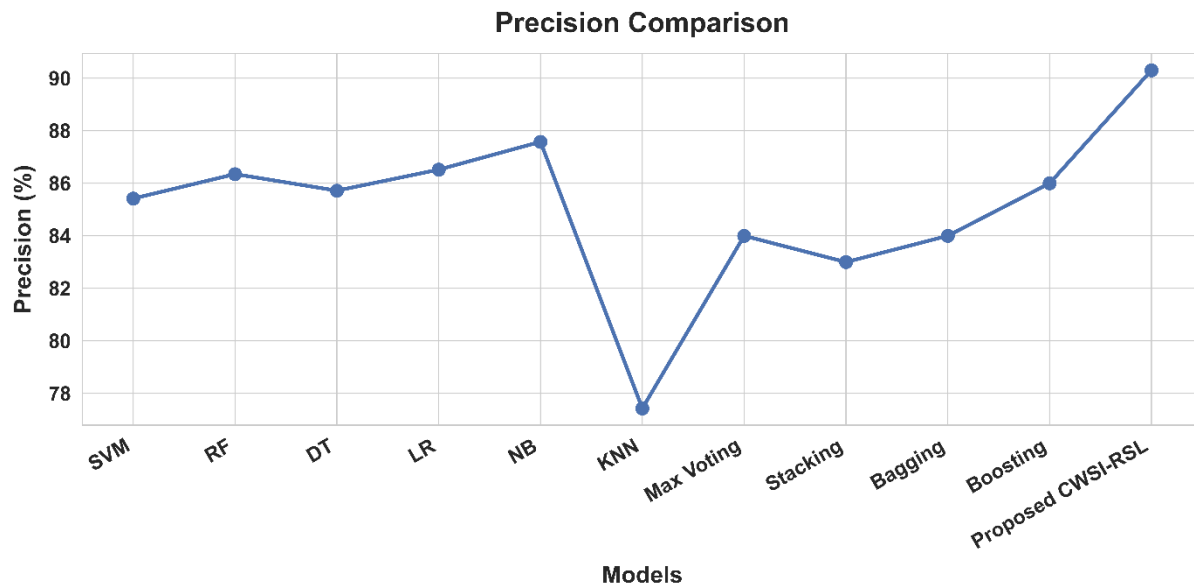


Fig 3. Precision Comparison

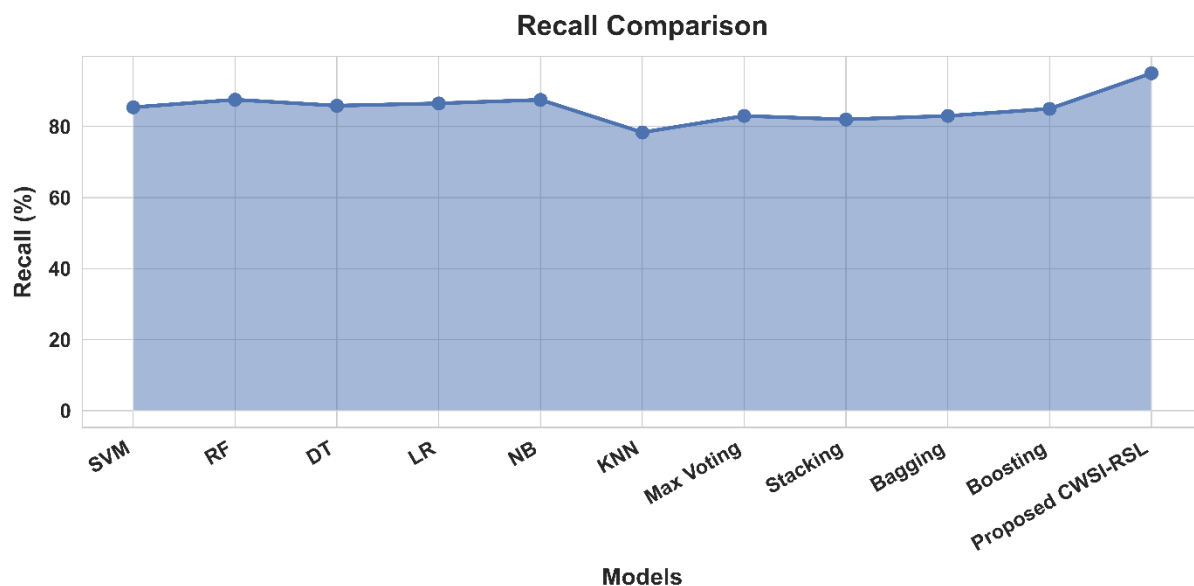


Fig 4. Recall Comparison

Figure 6 illustrates the significance of interactions between menopause, CA125, neutrophils and age-related features in the prediction of ovarian cancer. SIFE module effectively identifies hidden clinical relationships. SHAP analysis improves the interpretability of the models. To summarise, CWSI-RSL provides reliable, interpretable and accurate early ovarian cancer prediction.

3.3 Ablation Analysis

To study the contribution of the main components of the proposed CWSI-RSL framework, we conducted an ablation analysis. The baseline configuration with only the selected features reached an accuracy of 86.20%, a precision of 84.70%, a recall of 88.10%, and an F1-score of 86.36%. The accuracy was improved to 88.70% by adding clinical weighting, and the precision,

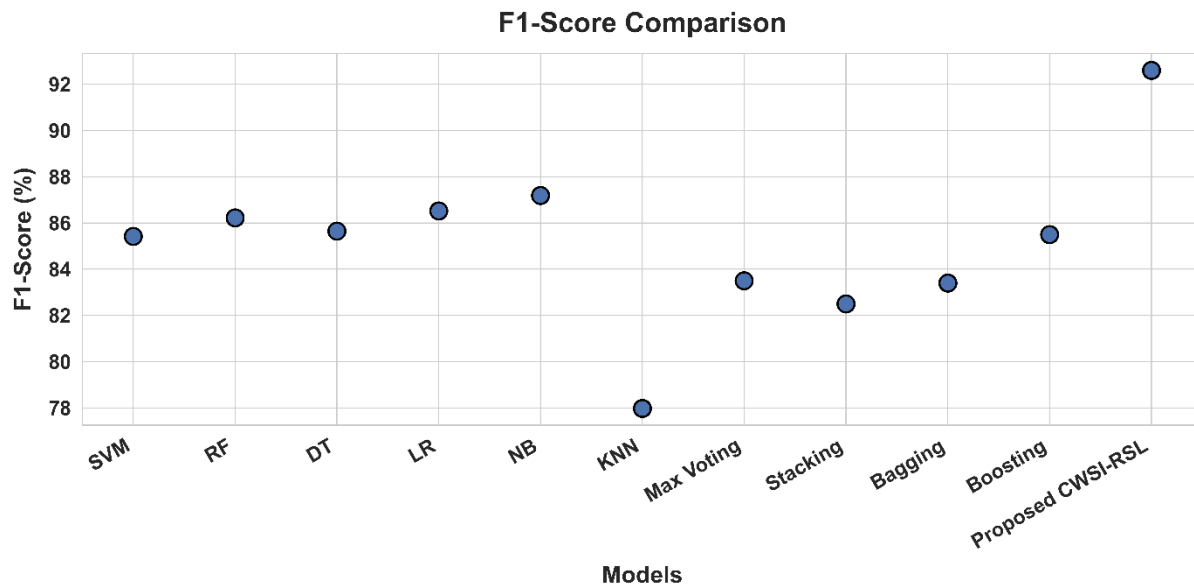


Fig 5. F1-Score Comparison

recall and F1-score were improved to 87.10%, 91.00% and 89.01%, respectively. The accuracy, precision, recall, and F1-score of the model without clinical weighting when SIFE was presented were 89.50%, 88.20%, 92.00%, and 90.06%, respectively. The combination of clinical weighting and SIFE further improved the performance to 91.10% accuracy, 89.80% precision, 94.00% recall and 91.85% F1-score. Finally, the best performance was achieved by including the risk stratification in the complete CWSI-RSL framework with 92.40%, 90.30%, 95.00% and 92.60% accuracy, precision, recall and F1-score respectively. The results are shown in Table 1.

Table 1. Ablation Study of CWSI-RSL Components

Configuration	Clinical Weighting	SIFE	Risk Stratification	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Baseline selected features	✗	✗	✗	86.20	84.70	88.10	86.36
+ Clinical Weighting	✓	✗	✗	88.70	87.10	91.00	89.01
+ SIFE	✗	✓	✗	89.50	88.20	92.00	90.06
Weighting + SIFE	✓	✓	✗	91.10	89.80	94.00	91.85
Complete CWSI-RSL	✓	✓	✓	92.40	90.30	95.00	92.60

The ablation results show that each component can improve the prediction performance of CWSI-RSL. Clinical weighting only improved the accuracy from 86.20% to 88.70%, an increase of 2.50% points. Thus, increasing the weight of clinically important features can improve classification performance. SIFE improved larger when used alone with accuracy going from 86.20% to 89.50% an improvement of 3.30 percentage points indicating that modelling interactions between selected clinical variables provides further predictive information to individual features. The accuracy was further improved to 91.10% when clinical weighting was used with SIFE, which shows the benefit of integrating feature importance with interaction-based representations.

The complete CWSI-RSL produced the best results, increasing accuracy, precision, recall and F1-score by 6.20, 5.60, 6.90 and 6.24 percentage points, respectively. And the recall of 95.00 % means most cases of cancer were detected correctly. The results suggest that clinical weighting, symptom interactions and risk stratification contributed to the improved performance.

3.4 Cross-validation and Reliability

Five-fold cross-validation was performed with GridSearchCV to improve the reliability of model optimisation and reduce the dependence on a single data split. In each fold the training data were divided into 5 roughly equal subsets, four of which were

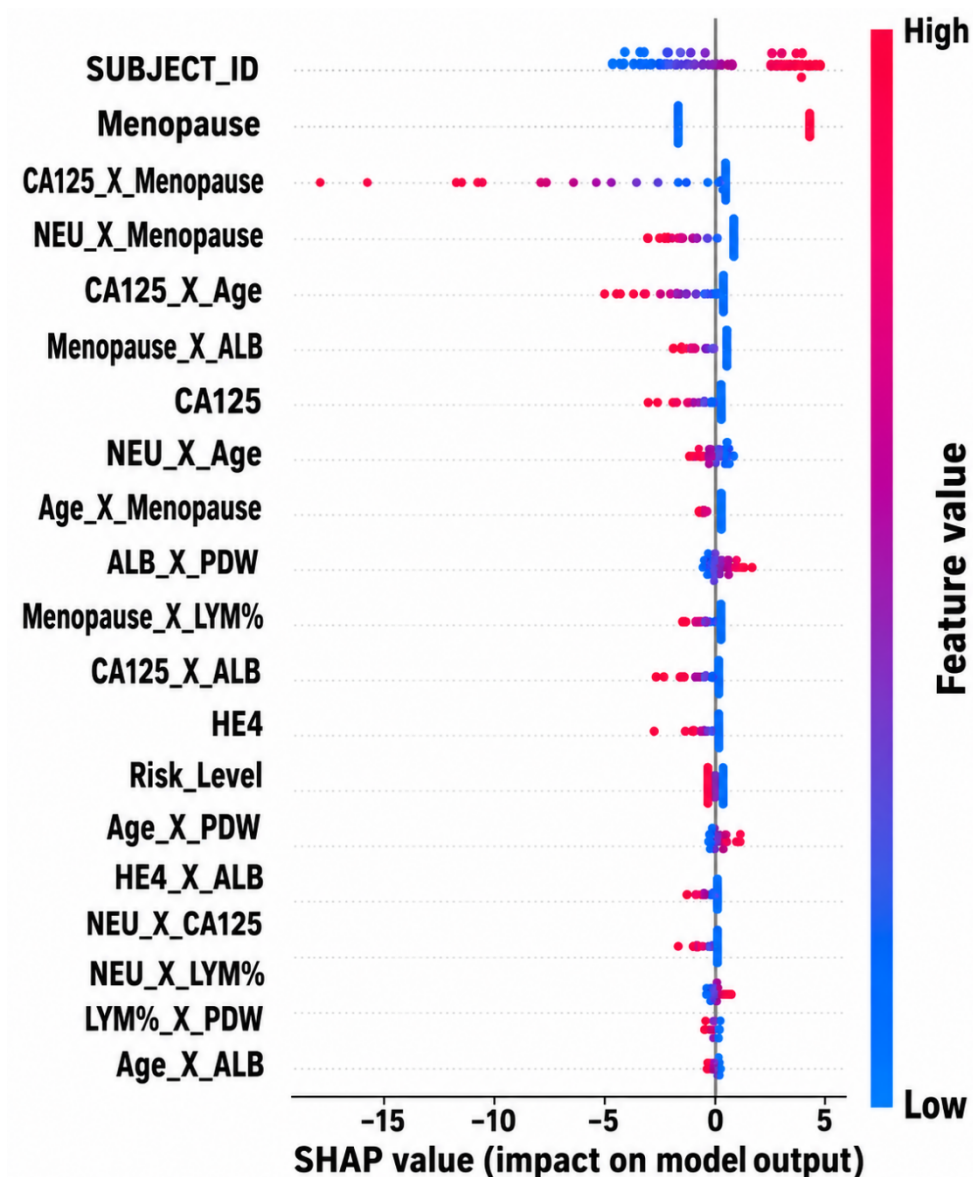


Fig 6. SHAP Summary Plot of Feature Importance

used to train the model and the other subset was used to validate the model. Each subset was used once as the validation set and repeated five times. The independent test set was not used in the cross-validation procedure but was held out for the final evaluation only.

Performance of cross-validation was calculated in terms of accuracy, precision, recall and F1-score. Five folds results were reported as mean $\pm$ standard deviation (SD) for each metric. The mean is the average performance across the five validation folds and the SD is the variability of the model across different folds. The results are given in Table 2.

The CWSI-RSL framework attained the highest mean performance in all four-evaluation metrics, with an accuracy of $91.68 \pm 1.21\%$, precision of $89.74 \pm 1.46\%$, recall of $94.21 \pm 1.35\%$, and F1-score of $91.91 \pm 1.18\%$. The relatively low standard deviations suggest that the performance was consistent across the five validation folds. Specifically, the recall variation was within $\pm 1.35\%$, which demonstrates the robustness of the framework to identify ovarian cancer cases across different validation subsets. These cross-validation results support the reliability of CWSI-RSL beyond a single train-test partition.

Table 2. Five-fold cross-validation results of CWSI-RSL

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Logistic Regression	86.72 ± 1.84	86.31 ± 1.96	88.24 ± 2.15	87.21 ± 1.72
SVM	85.48 ± 2.03	85.91 ± 1.87	86.74 ± 2.24	86.20 ± 1.95
Naïve Bayes	87.06 ± 1.72	87.42 ± 1.64	88.13 ± 1.89	87.72 ± 1.51
KNN	78.36 ± 2.68	78.02 ± 2.91	80.15 ± 2.74	78.98 ± 2.53
CWSI-RSL	91.68 ± 1.21	89.74 ± 1.46	94.21 ± 1.35	91.91 ± 1.18

3.5 Comparison with Previous Studies

The performance of CWSI-RSL was compared with recent ovarian cancer machine learning studies and the reported baseline models^{(6), (7), (8), (9), (10), (13), (14)}. Naïve Bayes and Boosting achieved 87.14% and 87.00% accuracy respectively^{(13), (14)}, while CWSI-RSL achieved 92.40% accuracy, 90.30% precision, 95.00% recall and 92.60% F1-score⁽¹³⁾. Thus, CWSI-RSL increased the accuracy by 5.26 and 5.40 percentage points respectively. Previous works have been mostly focused on classification, biomarkers, risk prediction and explainability^{(6), (7), (8), (9), (10)}. CWSI-RSL integrates these components with clinical weighting, symptom interactions, risk stratification and SHAP.

The proposed CWSI-RSL was compared to previously reported machine learning and ensemble models tested on the same ovarian cancer dataset. The baseline results for SVM, Random Forest, Decision Tree, Logistic Regression, Naïve Bayes and KNN were taken from Ikani et al.⁽¹³⁾ and the Max Voting, Stacking, Bagging and Boosting results were taken from JM and Subbulakshmi⁽¹⁴⁾. In this common-dataset comparison, CWSI-RSL achieved the highest accuracy of 92.40%, with 90.30% precision, 95.00% recall and 92.60% F1-score. These results demonstrate a better predictive performance compared to the previously reported models on the same dataset.

Recently, ovarian cancer detection, early prediction, analysis of clinical and genetic data, risk stratification and survival prediction have been performed using machine learning and XAI^{(14), (21), (22), (23), (24), (1), (2), (25), (4), (5)}. These methods include XAI-based prediction^{(14), (1), (2), (5)}, clinical and genetic feature integration⁽²²⁾, early-stage detection^{(25), (4)}, and risk-stratification and multimodal learning^{(24), (4)}. ML-based studies demonstrate the usefulness of feature selection and biomarkers^{(26), (27)}. CWSI-RSL provides a unified framework that integrates clinical weighting, SIFE, risk grouping, optimisation, and SHAP explainability.

3.6 Clinical Interpretation and Novelty

CWSI-RSL achieved a high recall of 95.00%, which means it accurately detected most of the ovarian cancer cases in the evaluated dataset. The F1-score of 92.60% indicates balanced prediction performance. Clinical weighting gives more importance to selected features, SIFE captures relationships between clinical variables and risk grouping gives low-, medium- and high-risk groups derived from the model. SHAP also helps in explaining important predictors such as menopause status, CA125-related interactions and neutrophil-based interactions. The major innovation of CWSI-RSL is the incorporation of clinical weighting, symptom interaction modelling, risk stratification and explainable machine learning into one framework for early ovarian cancer prediction.

4 Conclusion

In this study, the CWSI-RSL framework was developed for early ovarian cancer prediction combining clinical feature weighting, symptom interaction modelling, risk stratification, machine learning and explainable artificial intelligence. The proposed framework could achieve the accuracy of 92.40%, the precision of 90.30%, the recall of 95.00% and the F1-score of 92.60%, achieving the objective of improving the prediction performance and providing interpretable results. The ablation study also demonstrated that the inclusion of clinical weighting, SIFE, and risk stratification incrementally increased accuracy from 86.20% to 92.40%. SHAP elucidated that the most significant features that impacted the predictions were menopause status, interactions with CA125 and neutrophils. CWSI-RSL combines clinical weighting, feature interactions, risk stratification, and explainable ML to make the overall early risk prediction more transparent; however, it should not be used as a clinically validated diagnostic tool at the current time. Future work should include validation of the framework with larger and independent datasets and prospective clinical testing, as well as further model refinement with more clinical data and advanced learning techniques.

5 Acknowledgements

The authors gratefully acknowledge the support of the PG & Research Department of Computer Science, Government College for Women (Autonomous), Kumbakonam, Tamil Nadu, India. This research received no specific funding from any public, commercial, or not-for-profit funding agency.

Author details: Gomathi Gajendran serves as Guest Lecturer and Dr.F. Kurus Malai Selvi as Associate Professor & Head of the Department

References

- 1) Arafat D, et al. FRF-HHO: Early ovarian cancer prediction using explainable fuzzy random forest optimized by Harris Hawks algorithm. *Advances in Biomarker Science and Technology*. 2026;8:219–235. <https://doi.org/10.1016/j.abst.2026.01.003>.
- 2) Asif D, Kerdid N, Arif MS, Bibi M. Explainable ensemble learning approach for ovarian cancer diagnosis using clinical data. *Computational Modeling in Engineering & Sciences*. 2026;146(3):38. <https://doi.org/10.32604/cmescs.2026.077334>.
- 3) SamadiAfshar S, Azizi H, Masoudi M, SamadiAfshar S, Nikakhtar A, Skutella T. Preoperative differentiation of borderline and malignant ovarian tumors using interpretable machine learning. *Journal of Ovarian Research*. 2026;19:225. <https://doi.org/10.1186/s13048-026-02062-5>.
- 4) Gui T, Cao D, Yang J, Wei Z, Xie J, Wang W, et al. Early prediction and risk stratification of ovarian cancer based on clinical data using machine learning approaches. *Journal of Gynecologic Oncology*. 2025;36(4):e53. <https://doi.org/10.3802/jgo.2025.36.e53>.
- 5) Kucukakcali Z, Cicek IB, Akbulut S. Explainable machine learning for the early clinical detection of ovarian cancer using contrastive explanations. *Journal of Clinical Medicine*. 2025;14(17):6201. <https://doi.org/10.3390/jcm14176201>.
- 6) Sun Y, Wen B. Machine-learning diagnostic models for ovarian tumors. *Heliyon*. 2024;10(19):e36994. <https://doi.org/10.1016/j.heliyon.2024.e36994>.
- 7) Abuzinadah N, Posa SK, Alarfaj AA, Alabdulqader EA, Umer M, Kim TH, et al. Improved prediction of ovarian cancer using ensemble classifier and SHAPly explainable AI. *Cancers*. 2023;15(24):5793. <https://doi.org/10.3390/cancers15245793>.
- 8) Huang W, Suominen H, Liu T, Rice G, Salomon C, Barnard AS. Explainable discovery of disease biomarkers: the case of ovarian cancer to illustrate the best practice in machine learning and Shapley analysis. *Journal of Biomedical Informatics*. 2023;141:104365. <https://doi.org/10.1016/j.jbi.2023.104365>.
- 9) Comes MC, Arezzo F, Cormio G, Bove S, Calabrese A, Fanizzi A, et al. An explainable machine learning ensemble model to predict the risk of ovarian cancer in BRCA-mutated patients undergoing risk-reducing salpingo-oophorectomy. *Frontiers in Oncology*. 2023;13:1181792. <https://doi.org/10.3389/fonc.2023.1181792>.
- 10) Boehm KM, Aherne EA, Ellenson L, Nikolovski I, Alghamdi M, Vázquez-García I, et al. Multimodal data integration using machine learning improves risk stratification of high-grade serous ovarian cancer. *Nature Cancer*. 2022;3(6):723–733. <https://doi.org/10.1038/s43018-022-00388-9>.
- 11) Taleb N, Mehmood S, Zubair M, Naseer I, Mago B, Nasir MU. Ovary cancer diagnosing empowered with machine learning. *Proceedings of the 2022 International Conference on Business Analytics for Technology and Security (ICBATS)*. 2022;p. 1–6. <https://doi.org/10.1109/ICBATS54253.2022.9759010>.
- 12) Shahane S. Predict ovarian cancer [dataset]. *Kaggle*. nd. Available from: <https://www.kaggle.com/datasets/saurabhshahane/predict-ovarian-cancer>.
- 13) Ikani O, Aimufua GIO, Nyap CD, Olalere M, Ajayi BA, Abdullahi MU. Comparative analysis of selected machine learning models for the prediction of ovarian cancer. *Science World Journal*. 2025;20(4):1492–1497. <https://doi.org/10.4314/swj.v20i4.24>.
- 14) Lavanya JMS, Subbulakshmi P. Innovative approach towards early prediction of ovarian cancer: machine learning-enabled XAI techniques. *Heliyon*. 2024;10(9):e29197. <https://doi.org/10.1016/j.heliyon.2024.e29197>.
- 15) Laios A, Katsenou A, Tan YS, Johnson R, Otify M, Kaufmann A, et al. Feature selection is critical for 2-year prognosis in advanced stage high grade serous ovarian cancer by using machine learning. *Cancer Control*. 2021;28:10732748211044678. <https://doi.org/10.1177/10732748211044678>.
- 16) Liu SY, Zhu RH, Wang ZT, Tan W, Zhang L, Wang YQ, et al. Landscape of immune microenvironment in epithelial ovarian cancer and establishing risk model by machine learning. *Journal of Oncology*. 2021;2021:5523749. <https://doi.org/10.1155/2021/5523749>.
- 17) Arezzo F, Cormio G, Forgia DL, Santarsiero CM, Mongelli M, Lombardi C, et al. A machine learning approach applied to gynecological ultrasound to predict progression-free survival in ovarian cancer patients. *Archives of Gynecology and Obstetrics*. 2022;306(6):2143–2154. <https://doi.org/10.1007/s00404-022-06578-1>.
- 18) Barber EL, Garg R, Persenaire C, Simon M. Natural language processing with machine learning to predict outcomes after ovarian cancer surgery. *Gynecologic Oncology*. 2021;160(1):182–186. <https://doi.org/10.1016/j.ygyno.2020.10.004>.
- 19) Shannon NB, Tan LLY, Tan QX, Tan JWS, Hendrikson J, Ng WH, et al. A machine learning approach to identify predictive molecular markers for cisplatin chemosensitivity following surgical resection in ovarian cancer. *Scientific Reports*. 2021;11(1):16829. <https://doi.org/10.1038/s41598-021-96072-6>.
- 20) Gong TT, He XH, Gao S, Wu QJ. Application of machine learning in prediction of chemotherapy resistant of ovarian cancer based on gut microbiota. *Journal of Cancer*. 2021;12(10):2877. <https://doi.org/10.7150/jca.46621>.
- 21) Akter L, Akhter N. Ovarian cancer prediction from ovarian cysts based on TVUS using machine learning algorithms. Singapore. Springer. 2021. https://doi.org/10.1007/978-981-16-6636-0_5.
- 22) Gómez-Talal I, Barquín A, Bote-Curiel L, Yagüe-Fernández M, Rojo-Álvarez JL, García-Donás J. A machine learning approach to predicting ovarian cancer: integrating clinical and genetic data. *Scientific Talk*. 2024;10:100319. <https://doi.org/10.1016/j.sctalk.2024.100319>.
- 23) Singh M, Betgeri SN, Kakar SS. Artificial intelligence (AI) and machine learning (ML) in ovarian cancer: transforming detection, treatment, and prevention. *Journal of Ovarian Research*. 2026;19:91. <https://doi.org/10.1186/s13048-026-01979-1>.
- 24) Zhou Y, Duan Y, Teng M, et al. Survival prediction and risk stratification in R0-resected ovarian cancer: a multi-modal deep learning approach. *NPJ Precision Oncology*. 2026;10:60. <https://doi.org/10.1038/s41698-025-01263-3>.
- 25) Mikami M, Tanabe K, Nagaki S, Nogami Y, Imanishi T, Ikeda M, et al. Enhancing the positive predictive value of early-stage ovarian cancer detection using a two-step machine learning framework. *Journal of Translational Medicine*. 2026. <https://doi.org/10.1186/s12967-026-08539-7>.
- 26) Pradhani G, Gogoi J. Unveiling key gene biomarkers and clinical factors in breast cancer subtypes using machine learning and Cox regression. *Indian Journal of Science and Technology*. 2024;17(32):3350–3360. <https://doi.org/10.17485/IJST/v17i32.1440>.
- 27) Tintu PB, Veni S, Priya SM. An effective hybrid outlier selection method for breast cancer classification using machine learning algorithms. *Indian Journal of Science and Technology*. 2024;17(43):4494–4501. <https://doi.org/10.17485/IJST/v17i43.2109>.