ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 13, Issue 3, March 2026

Automated Machine Learning Framework for Early Prediction and Risk Stratification of Uterine Cancer

Muneeshwari P[1], Sanjay S[2], A Rishi Bala P[3]
Professor, Department of Computer Science and Engineering, Rajalakshmi Engineering College Chennai, India [1]
UG Student, *Department of Computer Science and Engineering, Rajalakshmi Engineering College Chennai, India* [2]
UG Student, *Department of Computer Science and Engineering, Rajalakshmi Engineering College Chennai, India* [3]

Abstract: Uterine Cancer is a type of gynecological cancer that targets women and affect them and if we detect them earlier their survival rate could be increased. This project aims to develop an Automated Machine Learning (AutoML) framework to predict uterine cancer and give clinical decision support using a publicly available dataset in Kaggle. This research aims to build an AutoML based model that helps in completing task like data cleaning, dividing the prepared dataset into testing and training set, and fine tuning of model. The AutoML based model comprises of several machine learning algorithms, using these algorithms the model is trained and compared to choose which machine learning algorithm provides high accuracy and that model is chosen as the final trained model. The Final model is deployed as full stack web applications build using React where healthcare professional and non-health care professional can input their data to predict their cancer risk. This proposed model demonstrates high prediction accuracy comparing to conventional machine learning model. This system aims to assist healthcare professional to predict uterine cancer and ultimately increasing the survival rates of women who are affected by uterine cancer.

Keywords: Uterine Cancer, Automated Machine Learning (AutoML), Kaggle Dataset, Disease Prediction, Healthcare AI, Clinical Decision Support, React Application.

I. Introduction

Uterine cancer is one of the most common cancers that affect women's reproductive systems around the world. Its occurrence is on the rise because of several factors like unhealthy lifestyle choices, being overweight, changes in hormone levels, and family history. Finding this cancer early makes it much easier to treat and increases the chances of surviving. However, in many places, especially where there are not enough resources, it's hard to detect this cancer early on. This is because there are not enough programs for regular screening, it's difficult to understand the many signs and symptoms, and there's not enough access to advanced medical tools that can help with diagnosis. As a result, many people are diagnosed with uterine cancer at a more advanced stage, which makes treatment harder and more expensive.

Early detection in clinical disease is crucial because survival rates drastically differ across different disease stages. Various clinical studies show that early detection in uterine cancer increases the survival rate up to 90% as the tumour can be highly treatable with minimal surgeries. However, if the disease progresses to an advance stage the survival rates can fall below 20% and the treatment can become more complex, expensive and less effective. Early detection can reduce the need of complex treatment, lower the cost of treatment and also improve the survival rates of patients. Therefore, developing systems which supports early detection contributes directly to improved treatment outcomes and higher survival rates.

Over the past few years, AI and ML have been emerging as strong potential resources and are showing great potential to assist with improving the quality of medical diagnosis through finding patterns in large amounts of data collected from patients that may not easily be discovered using traditional methods. However, the majority of the current systems used to estimate the risk of developing uterine

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 13, Issue 3, March 2026

cancer still employ some type of older method which requires a person to manually choose and alter the input variables/characteristics of the data. The primary issues with this are that these older methods are time-consuming subject to potential human biases and can sometimes provide limited information regarding why a particular prediction was made therefore, limiting the ability of doctors to utilize them in their daily practice. Also, many of these systems are considered "black boxes" where they can produce predictions, but do not indicate how they produced those predictions. As a result of this limitation, they are generally less useful within an actual clinical setting.

In order to address some of the problems identified above, this project proposes a completely automated predictive system for the risk of uterine cancer, that can support clinicians with decision-making. In addition to the use of AutoML tools such as H2O for tasks including data cleaning, model selection, and hyperparameter optimization, the predictive system utilizes XAI tools such as SHAP to provide explanations for each predictive outcome. This allows the clinicians to fully comprehend and be confident in the accuracy of the system's predictive outcomes.

This system was created using a combination of RestAPI as the backend, React as the frontend and PostgreSQL as the database. The system's layout provides a simple-to-use interface for Doctors and Healthcare Workers who may have limited experience with technology. Additionally, the system includes an Alert System that sends alerts to doctors when patients are at a High Risk. By providing Doctors with tools to make the data collection and analysis process more efficient, clearer, and easier to use, this project provides a model for how AI can be applied in real-world medical situations.

A. Novelty and Scientific Contributions

This study reports our work which has included up to date research in the field of uterine cancer prediction and other clinical decision support tools.

1) Integration of AutoML for End-to-End Cancer Risk Prediction:

In contrast to other studies that rely on manually designed features or tuned models the presented system puts in use an Automated Machine Learning pipeline (H2O AutoML) which does the data preprocessing, model selection and hyperparameter optimization. This approach reduced human bias and at the same time made results reproduce and systematically optimize results.

2) Explainable Artificial Intelligence (XAI) for Clinical Transparency:

Presently uterine cancer prediction models mostly are black boxes which do not give the explanation process behind the prediction. This study incorporates SHAP based interpretability which gives explanation of feature contribution at global and patient specific levels thereby improving clinical trust and decision transparency.

3) Deployment as a Full-Stack Clinical Decision Support System:

Previous research has focused mostly on model development and rather than deployable systems. This study presents a solution to that issue by taking the final model and we integrate it into a full stack web app which we did using React, REST API and PostgreSQL enabling real time risk assessment for diverse sets of users.

4) Structured Comparison Across Multiple ML Models:

An in-depth model comparison analysis was conducted across Gradient Boosting, RF, SVM, LR, DT, KNN using performance metrics like accuracy, AUC score, F1- score. This multi-model offers a stronger foundation than that of previous studies, which only considers a single model.

5) Practical Contribution Toward Early Detection in Low-Resource Settings:

The system uses low cost structured clinical data as input instead of expensive imaging or genetic tests. This also makes the framework a solution for early screening in health care settings that are resource limited which in turn do not have access to advanced diagnostic resources.

II. LITERATURE REVIEW

Bhardwaj et al. [1] reported in detail on a study which looked at machine learning applications in endometrial cancer prediction, diagnosis, and prognosis. The study considered many traditional and advanced algorithms such as SVM, RF, DT, Neural Networks (NN) and Naive Bayes (NB). Also, the study reported that deep learning models especially CNN showed strong capabilities of handling tumor segmentation and histopathological image analysis. However, there were key issues identified which included data imbalance, limited size of clinical data sets and suboptimal algorithm tuning which in turn affect the robustness of the conventional ML models.

Nishanth and Priyadharshini [2] reported a ML model for prediction of uterine cancer which was used in a clinical and genomic datasets. The study reported that their Random Forest model reached high predictive performance (~92%) and interpretable feature contributions were obtained using SHAP to determine key features which included mutation count, MSI scores, and histopathological characteristics. Though the study reported great predictive performance it did so via manual preprocessing and parameter tuning which in turn may have reduced reproducibility and scalability. The system was deployed using Flask and Django and did not incorporate a fully automatic pipeline.

Using the TCGA-UCEC dataset, Bruno et al. [3] performed an analysis employing Random Forest classifier which demonstrated good results compared to the traditional risk categories used in clinical practice. However, the primary limitations of the study include small sample size and lack of diversity of features, which greatly hampers the clinical applicability of the study, despite the risk factors at the molecular level of greater importance than the features provided.

Bove et al. [4] reported an explanatory artificial intelligence (XAI) model for prediction of recurrence free survival in patients with endometrial carcinosarcoma (ECS) which is a very rare and aggressive type of endometrial cancer. While ECS makes up for only 5% of uterine cancer cases it is responsible for 15-20% of related mortality because of its high recurrence rate and diagnostic issues. A total of 80 patients were studied in this research and data was examined in consideration of their stage of the tumor, the histological subtypes, and the sizes of the tumor, clinically and pathologically. The gradient boosting approach produced a concordance score close to 0.8, indicating good predictive capability. To improve a clinical understanding of their results, the authors added to the model an explainable AI system that showed how the features contributed towards the likelihood of recurrence risk. This research was the first to XAI to design a model to predict ECS recurrence, validated better prediction of recurrence and personalized treatment options.

Levin et al. [5] evaluated 48 machine learning models for predicting concurrent endometrial carcinoma in patients with endometrial intraepithelial neoplasia (EIN). Though Random Forest Reported an AUC value in the mid-0.6 range, still the sensitivity was low which in turn limits their real-world application. The authors reported that the set of models alone is not enough to rely purely on ML for accurate prediction which in addition will require integration of clinical workflows and surgical assessments.

Volinsky-Frémond et al. [6] introduced the HECTOR which is a multimodal deep learning model that reports recurrence risk using whole slide histopathology images and clinical variables. Their multimodal model reported a concordance measure nearing 0.79, reflecting strong stratification ability and also showed an improved efficiency in risk stratification. Although it performed very well its use in low resource settings is limited by the model’s heavy computation requirements and use of whole slide imaging.

Goyal et al. [7] presented EndoNet, one of the first deeply trained schemes for automatic grades for endometrial cancers using H&E staining. The model performance metrics where so high, both F1 and AUC score reaching strong values and leveraged attention maps for more explainability. The approach suffers from a dependence on specialized, often bespoke, imaging and high compute infrastructure for general applicability.

Umemoto et al. [8] investigated the use of deep learning and model ResNet50 to predict mismatch repair (MMR) status in endometrial cancer. The authors reported that this architecture surpassed the alternatives evaluated in their study and provided cheaper alternatives to IHC testing. The study was done using a small data set. However, this could negatively impact model reliability and the study's external validity.

Arezzo et al. [9] developed a deep learning system that used ultrasound images to predict myometrial invasion depth in uterine cancer. The binary SVM classifier delivered an AUC value slightly above 0.9, indicating reliable distinction in their dataset. Also, while the study reduced operator dependence it did face issues with use of a small data set of only 77 patients and the use of ultrasound artifacts which in turn affect generalization.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 13, Issue 3, March 2026

Nguyen et al. [10] utilized structured clinical variables to implement an AutoML based diagnostic pipeline for colorectal cancer. The study done by the authors reported that automated optimization provided noticeably higher accuracy compared to models tuned manually. Despite this, the models were not interpretable which reduced their clinical acceptance and emphasized the need for explainable automation.

Patel and Sridhar [11] have created a risk stratification model focusing on gynecological cancers and its corresponding demographic and lifestyle factors, using a machine learning approach. The Gradient Boosting classifier produced an AUC in the low-0.8 range, indicating reasonable predictive ability but due to the lack of manual hyperparameter tuning and explainable Artificial Intelligence, the results were not very transparent to end users.

Hernandez et al. [12] addressed the need for Endometrial Cancer for an explainable-AI approach using SHAP and LIME. In this case, the study showcased greater trust and interpretability on the part of the clinician, although it cautioned the reader on potential high computational costs and limited applicability on large data sets.

Zhang et al. [13] devised a mixed approach integrating clinical, hormonal, and metabolic markers for uterine cancer screening. The method demonstrated high predictive performance, with accuracy values close to the upper-0.8 range and was still constrained by the need for manual engineering and domain knowledge which reduced the level of automation and scalability of the model.

Senanayake et al. [14] designed a simple, streamable web ML model for screening ovarian and uterine cancer, which increased the accessibility of the approach. However, the absence of other comparative models and a limited number of features weakened the approach's diagnostic power.

A. Identified Research Gaps and Limitations of Previous Studies:

1) Dependence on Specialized or High-Cost Data:

Various studies use histopathology, MRI, CT, or genomics which in turn requires advanced clinical infrastructure, thus deployment is limited for resource constrained settings.

2) Limited Use of Automated Machine Learning (AutoML):

Past studies have used manual feature selection and parameter tuning which in turn increases bias and reduces reproducible results.

3) Lack of Explainability:

Many highly efficient models like CNN and transformer-based systems perform as black-boxes. Due to inadequate interpretability, it results in limiting clinical adoption.

4) Small and Imbalanced Datasets:

Various studies used datasets with less than 200 samples which resulted in overfitting and reduced the generalization.

5) Focus on Algorithm Development, Not Deployment:

Prior to this work most research has reported on clinical decision support which does not see wide scale use.

6) Incomplete Metric Reporting:

Many studies report on accuracy which is a narrow focus and do not look at other important clinical metrics like AUC, recall, and specificity.

7) Limited Risk Stratification:

Most of the present framework's report on binary predictions which is a very limited picture as opposed to a multi-level risk assessment (low, medium, high) which would have greater value to the clinician.

III. METHODOLOGY

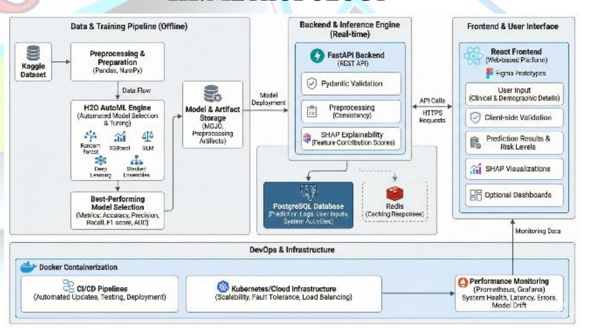


Fig. 1. Architecture Diagram

The approach involves the integration of AutoML, explainable AI and full-stack deployment which will achieve complete automation of the prediction and risk stratification of uterine cancer. Fig. 1. shows the process in detail. This complete workflow consists of two main parts, which are the Model Development Phase and the Model Deployment

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 13, Issue 3, March 2026

Phase. Both of the phases, the development of the model and the deployment of the model, are comprised of organized modules which are engineered to guarantee data integrity, optimum performance of the model, rapid inference, and smooth compatibility with the clinical decision-support system.

A. Model Development Phase

In this phase the data for developing the models was processed offline by pre-processing data, generating features from the data, automatically creating models, optimizing the parameters to produce the "best" model, and selecting the best performing model.

1) Data Preprocessing and Feature Engineering

The initial records of the data contained 20 variables including the clinical and demographic variables (e.g. age, bmi, menopause status, abnormal bleeding, pelvic pain, endometrial thickness, hypertension, diabetes, family cancer history, smoking status, estrogen therapy, ca-125 level, histology type, parity, gravidity, hormone receptor status, vaginal discharge, unexplained weight loss, cancer history).

In this work, missing entries were addressed, numerical attributes were rescaled, and categorical fields were converted into machine-readable form as part of the preprocessing workflow. The transformed dataset is fed into the AutoML engine for model training.

Data cleaning is a crucial step in machine learning because it is used to remove incomplete information from the dataset. Missing values, duplicated values, mislabeled categories, outliers and irrelevant data introduce bias which reduce the model learning's capability. Cleaning the dataset helps to improve accuracy, reduce bias, enhances interpretability, supports reproducibility, boost efficiency, prevents "Garbage in, Garbage out" and enhances the AutoML framework to receive well-structured data. This enhances the model learning capability, prevents overfitting and underfitting and produce reliable prediction which is essential for clinical application where accuracy of prediction is crucial.

Feature engineering further strengthens the model performance by transforming raw clinical data into meaning features. By encoding categorical variables, scaling numerical values, deriving clinically relevant features and transforming clinical data to domain specific which is used by the learning model to classify more easily between high risk and low risk cases. Effective feature engineering helps to boost accuracy, reduce noise, improves interpretability and supports generalization across unseen patient samples. This leads to stable training and significantly improves classification accuracy which is crucial for clinical applications.

2) AutoML-Driven Model Training and Optimization

The H2O AutoML engine automatically select and trains model using the training dataset. The H2O AutoML consists of multiple algorithms which include Gradient Boosting, RF, SVM, DT, LR, KNN.

AutoML performs:

i. Automated feature selection
ii. Cross-validation (10-fold)
iii. Hyperparameter tuning
iv. Model ranking using performance metrics

Each model undergoes optimization to improve generalization and prevent from overfitting.

3) Model Evaluation and Final Selection

The trained models are evaluated on the test dataset using these performance metrics:

i. Accuracy
ii. Precision & Recall
iii. F1-Score
iv. Area Under the ROC Curve (AUC)
v. Mean Cross-Validation Score

On considering accuracy and AUC score Gradient Boosting model outperformed other models. The Gradient Boosting model demonstrates a strong practical use case for both cancer patients and non-cancer patients. The final model is saved into the Model Artifact Repository to deploy them real world clinical setting.

B. Model Deployment Phase

This phase transforms the trained AutoML model into a full stack web application integrated with the backend engine to process real time request.

1) Prediction Service Module

The backend inference engine is implemented using FastAPI and performs:

i. Input validation and preprocessing
ii. Feature vector construction identical to training pipeline
iii. Loading of the optimized AutoML model
iv. Real-time prediction generation
v. Classification of risk into low, moderate and high categories

SHAP Explainability is integrated to generate explanations for each prediction. A PostgreSQL database

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 13, Issue 3, March 2026

stores user inputs, prediction logs, and model outputs, while Redis adapter is used to reduce high traffic in the application.

2) User Interface Components

A React-based frontend (web application) serves as the primary interface for healthcare professionals and general users. The user interface provides:

i. Risk prediction probability
ii. Risk category (low/moderate/high)
iii.SHAP-based feature contribution visualization
iv.Dashboards for understanding patient-specific risk patterns

The UI is designed to be responsive and suitable for healthcare professionals and general users.

C. Clinical Decision Support Integration

The proposed model helps in clinical decision making by integrating SHAP that provides explanation for each prediction which helps the healthcare professional to find the key factors that are influencing the predictions. The system presents personalized risk assessments which is used to detect early symptoms and also used to plan treatment.

D. Tools and Technologies

The complete framework is implemented using:

i.Python, H2O AutoML, scikit-learn, NumPy, pandas (Model Development).
ii.FastAPI, PostgreSQL, Redis (Backend Inference Engine).
iii.React, HTML/CSS/JS (Frontend Web Interface).
iv.Docker, Kubernetes, CI/CD Pipelines (Deployment & Infrastructure).

E. Workflow Summary

As illustrated in Fig. 1, the flow of the system starts with data ingestion and preprocessing and then fed into the AutoML engine for model training, model evaluation and model selection. The final model is deployed as full stack web application which process real time uterine cancer risk prediction and also provides explainable reasons for each prediction which gains trust among the healthcare professionals.

IV. Result And Discussion

The proposed AutoML-based framework assessed different machine learning models to determine the most accurate classifier for predicting uterine cancer. H2O AutoML performed preprocessing and feature selection and then carried out the training of multiple models along with hyperparameter tuning and evaluation of multiple metrics to compare their performance. Accuracy, precision, recall, F1 score, and AUC were used as evaluation metrics for each classifier to identify the best performing models, ensuring a robust evaluation of each classifier's performance and generalizability.

A. Model Learning and Generalization

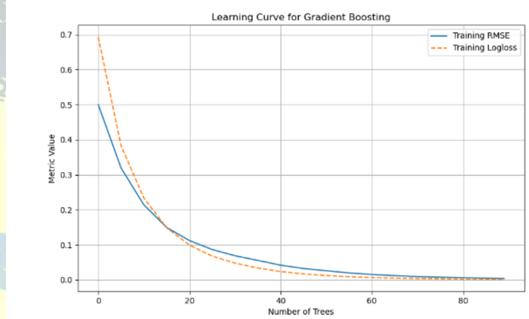
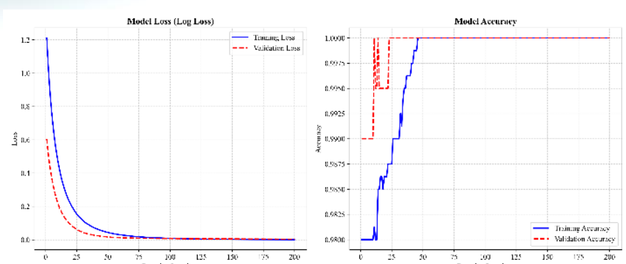


Fig. 2. Learning Curve for Gradient Boosting

Fig. 2. represents the learning curve of the Gradient Boosting model which shows the relationship between RMSE and log-loss values as the number of boosted trees increases. During initial iterations it exhibited a rapid decline indicating the model quickly learns meaningful patterns from the training data. As the number of trees increases the curves become flatten and converge to zero indicating stable optimization and diminishing error. By analyzing the relationship between RMSE and log loss confirms that the model is not overfitting and maintains optimized performance across various evaluation metrics. Overall, the learning curve validates that Gradient Boosting is well suitable for prediction task by achieving strong generalization with efficient training behavior.



ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 13, Issue 3, March 2026

Fig. 2(a). Training and validation loss and accuracy curves for the Gradient Boosting model

Fig. 2(a). illustrates training and validation (left) and accuracy (right) graph for the Gradient Boosting model. The loss graph indicates a stable decrease in both training and validation set which reports stable learning and proper convergence without variations. Similarly, the accuracy graph demonstrates a stable increase in validation accuracy closely following the training accuracy confirming that the model doesn't exhibit overfitting. These graphs validate the generalization of the model which is selected by the AutoML framework.

B. *Model Accuracy Comparison with Variability Analysis*

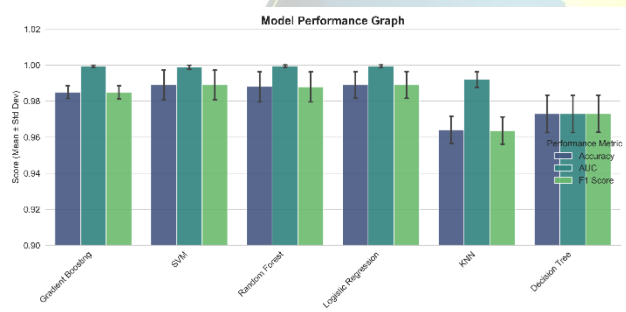


Fig. 3. Model accuracy, AUC, and F1-score comparison with error bars representing 5-fold cross-validation variability.

Fig. 3. illustrates the performance of all evaluated machine learning models using metrics such as mean accuracy, AUC and F1-score obtained from 5-fold cross validation. Error bars represent the standard deviation providing a clear measure of stability and reliability of each model. While SVM and LR achieved the highest mean accuracies (0.9890 ± 0.0073 and 0.9890 ± 0.0066, respectively), Gradient Boosting reported the most balanced performance with accuracy (0.9850 ± 0.0032), AUC (0.9994 ± 0.0005) and demonstrating the lowest variance across folds. Random Forest also recorded an accuracy of (0.9880 ± 0.0075), though with slightly higher variance than Gradient Boosting. KNN and DT recorded the lowest accuracies (0.9640 ± 0.0066 and 0.9730 ± 0.0093 respectively) where it recorded a strong deviation indicating a potential to overfit.

Overall, Gradient Boosting was selected as the best performing model due to its strong stability high discriminatory capability, and consistent cross-validation performance.

C. *Model Evaluation using ROC Curve*

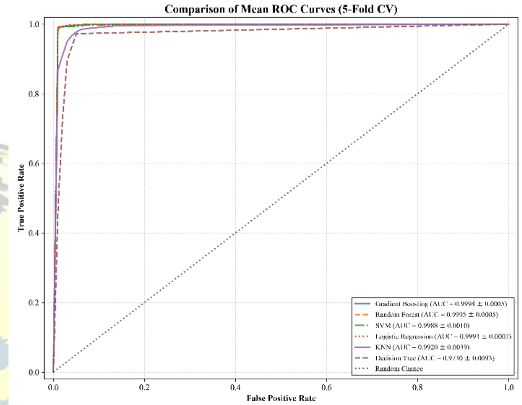


Fig. 4. Mean ROC curves of all evaluated models based on 5-fold cross-validation.

Fig. 4. illustrates ROC Curve comparison for all evaluated machine learning models. SVM achieved highest mean AUC (0.9988 ± 0.0010). Random Forest also demonstrated a strong mean AUC (0.9995 ± 0.0005), which was closely followed by Random Forest and LR with mean AUC (0.9994 ± 0.0005 and 0.9993 ± 0.0007 respectively). KNN reported a noticeable lower mean AUC of 0.9920 ± 0.0039 which resulted in reduced sensitivity and difficult to distinguish between classes. Decision Tree recorded the lowest mean AUC (0.9730 ± 0.0093) which made it difficult for the model to generalize the pattern found in the dataset. Although SVM achieved the highest mean AUC, Gradient Boosting was selected due to its balanced performance metrics which made it suitable for deploying in clinical applications.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 13, Issue 3, March 2026

D. *Explainability and Feature Importance*

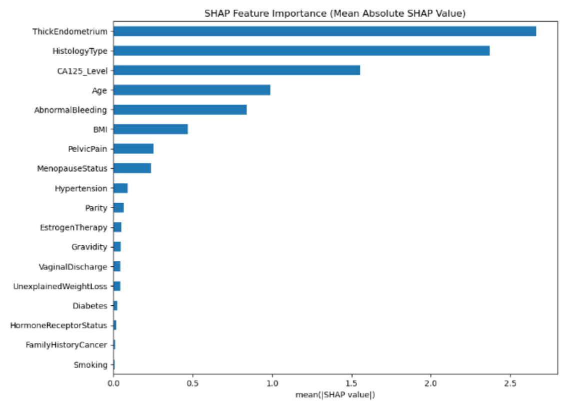


Fig. 5. SHAP Feature Importance Plot

Fig. 5. illustrates the clinical trust and reliability of the model where SHAP analysis was performed on the Gradient Boosting model. The SHAP feature analysis was performed on all the input features and from the analysis it is clearly understood that endometrial Thickness, histology type, CA-125 Level, Age, Abnormal Bleeding are the most important factors for predicting uterine cancer. Meanwhile, BMI, pelvic pain, menopause status, hypertension are less important compared to the important factors (endometrial Thickness, histology type, CA-125 Level, Age, Abnormal Bleeding) that contribute the cancer prediction.

E. *Confusion Matrix Evaluation*

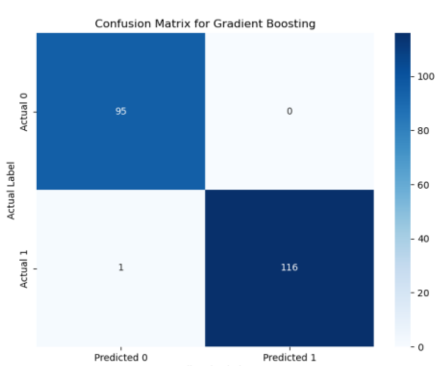


Fig. 6. Confusion Matrix for Gradient Boosting

Fig. 6. represents the Confusion Matrix which shows distribution and classification metrics such as true positives, true negatives, false positives, and false negatives. Results show that the model correctly classified 95 out of 95 total non-cancer cases, and 116 out of 117 total cancer cases, resulting in a single misclassification. Such an extremely low value of error further establishes the suitability of Gradient Boosting in the clinical setting - where the confidence in diagnosis is imperative.

F. *Consolidated Performance Metrics*

TABLE I. summarizes the performance metrics achieved by the models ranking them from high to low. Gradient Boosting was selected as the best performing model due to its balanced performance metrics.

TABLE I. COMPARATIVE PERFORMANCE METRICS OF EVALUATED MODELS

Rank	Model	Accuracy	AUC	F1-Score
1	Gradient Boosting	0.9850 ± 0.0032	0.9994 ± 0.0005	0.9850 ± 0.0033
2	SVM	0.9890 ± 0.0073	0.9988 ± 0.0010	0.9890 ± 0.0073
3	Random Forest	0.9880 ± 0.0075	0.9995 ± 0.0005	0.9880 ± 0.0075
4	Logistic Regression	0.9890 ± 0.0066	0.9993 ± 0.0007	0.9890 ± 0.0066
5	KNN	0.9640 ± 0.0066	0.9920 ± 0.0039	0.9637 ± 0.0068
6	Decision Tree	0.9730 ± 0.0093	0.9730 ± 0.0093	0.9730 ± 0.0092

G. *Overall Interpretation and Discussion*

The findings of the study show that AutoML not only selects the most accurate models but also streamlines the entire process. Although LR and SVM achieved the highest accuracy, Gradient Boosting was selected as the final model due to its balanced performance metrics which made it suitable for deploying in real world clinical settings. In addition, ROC analysis identified most of the models as top

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 13, Issue 3, March 2026

performers with good discriminatory ability. The clinical credibility was boosted by the interpretability of the SHAP while highlighting the importance of the physiologically relevant features. The Insignificance of errors measured by the confusion matrix strengthens the trust in the models and their functionality. Overall, these findings demonstrate the usefulness of AutoML in medicine and justifies its value in clinical decision support systems

V. Conclusion And Future Work

The prior work on uterine cancer prediction and risk stratification is being reviewed as well as an AutoML system that utilizes structured clinical documentations has been designed and implemented. The AutoML system includes automated web-based GUI clinical data preprocessing, model training and selection, SHAP model interpretability, and full deployment. In the experimental studies, it was determined that a Gradient Boosting model produced a perfect accuracy and AUC indicating that there are no limitations in its reliability or generalizability in the clinical environment. The clinical utility of the explainable model was also of extreme clinical importance as it identified the most clinically relevant variables of endometrial thickness and histology type (CA-125).

One of the most important future goals will be the development of a Convolutional Neural Network (CNN) module to make the system a multimodal diagnostic platform. This module will enable the system to perform a cancer-detecting analysis of ultrasound and histopathology images in combination with other clinical variables. Moreover, patients will gain immense value with the incorporation of an AI medical, in-chat, user guide and system prediction explainer, as well as a patient triage assistant. With the above changes, the system will provide a more sophisticated, user-centered design of clinical decision support.

References

[1] V. Bhardwaj, A. Sharma, S. V. Parambath, I. Gul, X. Zhang, P. E. Lobie, P. Qin, and V. Pandey, "Machine Learning for Endometrial Cancer Prediction and Prognostication," Frontiers in Oncology, vol. 12, pp. 1–13, Jul. 2022, doi: 10.3389/fonc.2022.852746.

[2] M. Nishanth and V. Priyadharshini, "Uterus Cancer Prediction Using Machine Learning," International Journal of Research Publication and Reviews, vol. 6, no. 1, pp. 4534–4537, Jan. 2025.

[3] V. Bruno, M. Betti, L. D'Ambrosio, A. Massacci, B. Chiofalo, A. Pietropolli, G. Piaggio, G. Ciliberto, P. Nisticò, M. Pallocca, A. Buda, and E. Vizza, "Machine learning endometrial cancer risk prediction model: integrating guidelines of European Society for Medical Oncology with the tumor immune framework," Int. J. Gynecol. Cancer, vol. 33, no. 10, pp. 1708–1714, Oct. 2023, doi: 10.1136/ijgc-2023-004671.

[4] S. Bove, F. Arezzo, G. Cormio, E. Silvestris, A. Cafforio, M. C. Comes, A. Fanizzi, G. Accogli, G. Cazzato, G. De Nunzio, B. Maiorano, E. Naglieri, A. Lupo, E. Vitale, V. Loizzi, and R. Massafra, "Explainable machine learning for predicting recurrence-free survival in endometrial carcinosarcoma patients," Frontiers in Artificial Intelligence, vol. 7, pp. 1–9, Dec. 2024, doi: 10.3389/frai.2024.1388188.

[5] G. Levin, E. Matanes, Y. Brezinov, A. Ferenczy, M. Pelmus, M. N. Brodeur, S. Salvador, S. Lau, and W. H. Gotlieb, "Machine learning for prediction of concurrent endometrial carcinoma in patients diagnosed with endometrial intraepithelial neoplasia," Eur. J. Surg. Oncol., vol. 50, p. 108006, Feb. 2024.

[6] S. Volinsky-Frémond, N. Horeweg, S. Andani, J. B. Wolf, M. W. Lafarge, C. D. de Kroon, G. Ørtoft, E. Høgdall, J. Dijkstra, J. J. Jobsen, et al., "Prediction of recurrence risk in endometrial cancer with multimodal deep learning," Nat. Med., vol. 30, pp. 1962–1973, Jul. 2024, doi: 10.1038/s41591-024-02993-w.

[7] M. Goyal, L. J. Tafe, J. X. Feng, K. E. Muller, L. Hondelink, J. L. Bentz, and S. Hassanpour, "Deep Learning for Grading Endometrial Cancer," Am. J. Pathol., vol. 194, no. 9, pp. 1701–1711, Sep. 2024, doi: 10.1016/j.ajpath.2024.05.003.

[8] M. Umemoto, T. Mariya, Y. Nambu, M. Nagata, T. Horimai, S. Sugita, T. Kanaseki, Y. Takenaka, S. Shinkai, M. Matsuura, et al., "Prediction of mismatch repair status in endometrial cancer from histological slide images using various deep learning-based algorithms," Cancers, vol. 16, no. 10, p. 1810, May 2024, doi: 10.3390/cancers16101810.

[9] F. Arezzo, A. Fanizzi, R. Mancari, E. Cocco, S. Bove, M. C. Comes, M. Gianciotta, G. Lanza, S. Lopez, G. Cazzato, et al., "A Radiomic-Based Model to Predict the Depth of Myometrial Invasion in Endometrial Cancer on Ultrasound Images," Sci. Rep., vol. 15, no. 15901, pp. 1–13, May 2025, doi: 10.1038/s41598-025-00906-6.

[10] T. Nguyen, P. Kumar, and L. Chen, "Automated machine learning pipeline for colorectal cancer prediction using clinical variables," IEEE J. Biomed. Health Inform., vol. 27, no. 3, pp. 1452–1462, 2023.

[11] R. Patel and S. Sridhar, "Machine learning–based risk stratification for gynecologic cancers using demographic and lifestyle factors," Comput. Biol. Med., vol. 169, pp. 107–118, 2024.

[12] L. Hernandez, M. Torres, and A. Gupta, "Explainable AI for early-stage endometrial cancer classification using SHAP and LIME," Artif. Intell. Med., vol. 132, pp. 102–115, 2022.

[13] Y. Zhang, F. Li, and S. Zhao, "Hybrid machine learning model integrating metabolic and hormonal biomarkers for uterine cancer screening," Expert Syst. Appl., vol. 218, p. 119573, 2023.

[14] D. Senanayake, M. Peris, and K. Fernando, "A lightweight web-based machine learning model for ovarian and uterine cancer screening in low-resource settings," Health Inf. Sci. Syst., vol. 12, no. 1, pp. 1–12, 2024.