Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Silent Sentinel: The Unseen Battle of Prostate cancer early diagnosis with advanced artificial neural network technology

G.Mohan[1], A. Ayub Khan[2], RVS Praveen[3], Sujit N. Deshpande[4], Dr.A.Nesarani[5], Mohit Tiwari[6]

[1]Associate Professor, Department of Mathematics, K.S.Rangasamy College of Technology, TamilNadu, India,.mohanjee.ksrct@gmail.com,Orcid id: 0000-0002-9873-1762

[2]Assistant Professor, Department of ECE, Sona College of Technology, Tamil Nadu, India, ayubkhan.slm@gmail.com, Orcid id: 0000-0002-5922-6464

[3]Director, LTIMINDTREE, USA, praveen.rvs@gmail.com

[4]Assistant Professor, Department of Computer Engineering, Vishwakarma University, Pune. sujit.sujitdeshpande@gmail.com, 0000-0002-6041-8666

[5]Associate professor, Department of Computer Science and Engineering, koneru lakshmaiah education foundation, vadeeswaram, GUNTUR, Andra Pradesh. 522302

[6]Assistant Professor ,Department of Computer Science and Engineering, Bharati Vidyapeeth's College of Engineering, A-4 Block, Rohtak Road, Paschim Vihar,Delhi, mohit.t.bvcoe@gmail.com

ARTICLE INFO

Received: 18 Dec 2024
Received: 18 Feb 2025
Accepted: 16 Feb 2025

ABSTRACT

Introduction: The condition of prostate cancer continues to represent a substantial medical issue because it necessitates precise predictive models that support healthcare decisions. A research study investigates how Extreme Gradient Boosting (XGBoost) performs when identifying significant biomarkers to anticipate prostate cancer risks. The research compares XGBoost to Sequential Minimal Optimization (SMO) in Support Vector Machines (SVM) to evaluate its better performance in classification. Prostate-Specific Antigen (PSA) levels together with Gleason scores and molecular markers form the clinical parameters within the dataset. Process data correctly by cleansing it and normalizing it while extracting important features to achieve the optimal model performance.

Objectives: This research investigates how well Extreme Gradient Boosting (XGBoost) predicts prostate cancer through an examination of its function to discover complex non-linear data patterns in clinical datasets. The research compares XGBoost to SMO algorithm of Support Vector Machines through a performance assessment of classification metrics including accuracy, precision, recall and F1-score. The research explores two important biomarkers Prostate-Specific Antigen (PSA) levels along with Gleason scores which affect how prostate cancer is expected to progress. The research applies machine learning advanced methods to develop improved early identification systems that lead to better clinical decisions which subsequently enhance patient health results.

Methods: The research method relies on a structured machine learning process where multiple steps start with detailed data preprocessing for inconsistent data elimination and variable normalization. Feature extraction technologies increase model interpretability by being applied to the analysis. The available dataset contains two well-separated sections for training and testing purposes which aims to establish reliable validation procedures. Both XGBoost and SVM with SMO operate for classification tasks while their achievement levels are assessed through standard measurement criteria which include accuracy, precision, recall and F1-score. A logistic regression analysis validates whether PSA levels and Gleason scores are significant factors in predictive modeling.

Results: XGBoost achieves superior results over SVM with SMO according to every mitigation evaluation indicator. The classification metric for XGBoost displays 0.95 accuracy whereas SVM with SMO achieves 0.87 accuracy. The evaluation of XGBoost produced precision measurements of 0.92 while recall reached 0.93 and F1-score achieved 0.92. These evaluation scores surpassed SVM with SMO which generated precision of 0.85, recall of 0.80 and F1-score of 0.82. XGBoost manages to lower the chances of overfitting while improving biomarker interpretation so it shows clear potential for clinical prostate cancer prediction applications.

Conclusions: The conducted research demonstrates XGBoost's effectiveness for prostate cancer classification through better prediction accuracy along with stronger resistance than SVM with SMO. The research demonstrates AI-powered models should become an integral part of clinical practice to improve the detection and treatment preparation for patients.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Keywords: Support Vector Machines, SMO and XGBoost algorithms ,clinical data analysis, predict prostate cancer, medical AI application

I.INTRODUCTION

Prostate cancer is one of the most common types of cancer afflicting men globally and continues to represent a substantial burden of disease and morbidity. Early detection is essential as timely detection of prostate cancer will improve treatment outcomes, minimize treatment burden and improve survival. Despite this we believe there are significant limitations in current diagnostic protocols. Clinical parameters such as the monitoring of Prostate Specific Antigen (PSA) levels, Gleason scores, and digital rectal examinations by medical professionals, frequently lack the precision and sensitivity required for early-stage identification. This lack of precision and sensitivity can lead both to overdiagnosis and underdiagnosis leading to non-essential biopsies, psychological stress or lost opportunities for early intervention.

We identified that we urgently require intelligent diagnostic tools informed through data to improve existing clinical parameter interpretation of prostate cancer diagnosis. D. Li *et al.* [1] To address this challenge, we propose a data-driven machine learningmodel to locate, describe & interpret prostate cancer risk using structured clinical datasets. Our approach uses Extreme Gradient Boosting (XGBoost), a powerful ensemble learning algorithm which can model complex nonlinear structures in high-dimensional structured data. For comparative sake, we also model and use Support Vector Machines (SVMs) with Sequential Minimal Optimization (SMO), which is a commonly used classifier in bio clinical work to create a point of comparison. This work begins with collating a structured dataset of clinically recognized biomarkers that are clinically actionable.

With the data now in a reliable state we can focus on improving model interpretability and generalizability, chiefly through advanced feature engineering methods that focus on establishing the most pertinent clinical predictor variables. This is crucial, otherwise the model might be overfit to atypical cases, and before generalizing, the model is producing clinically meaningful reasoning and not mathematically based thinking/decisioning. In addition, we reserve portions of the dataset as stratified training and testing dataset. J. Chapiro [3] This step is important for fair model assessment and avoiding leakage of information. One of the other innovation aspects of our approach is using Explainable Artificial Intelligence (XAI) methods. Specifically, we are using SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) methods to explain AI model decisioning. By using these two methods we are also able to explain the influence of each input predictor on the observed outcome, both from a global perspective of a population but also for every patient as an independent unit. With documented reasons for a model's decisions, we can better align model rationale with reasons articulated by clinicians, consequently, improve trust and promote uptake of AI inreal-world clinical environments.

II.LITERATURE SURVEY

Recent developments in artificial intelligence have shown promise in supporting prostate cancer diagnoses with interpretable models. One study performed by D. Li et al. [1] utilized an XGBoost algorithm in the diagnosis of prostate cancer in patients with a PSA of < 20 ng/ml, which showed high classification accuracy. Along with this, they also used a PSAMR biomarker to increase odds of early detection, demonstrating the clinical implications of using interpretable machine learning methods. Similarly, T. G. Ioannidis et al. [2] used explainable AI (XAI) methods to unpack model predictions; thus, bringing knowledge to clinicians regarding AI models prediction outputs. In terms of radiological imaging, J. Chapiro [3] described the importance of model interpretability/reproducibility in interpreting prostate MRI results when conveying information to patients and how explainability increases clinician trust and decision making. Along with this, A. El-Melegy et al. [4] presented a new framework that combines deep learning and visual representation of tabular clinical data, which made the diagnostic process more intuitive for clinicians, and embodied the impactful power of structured data and visual representation.

In the space of using AI in prostate diagnostics, S. M. Olabanjo et al. [5] published a systematic review that visualized the use of machine learning, and deep learning, methods for prostate cancer detection. They indicated the use of imaging

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

data was ubiquitous and emphasized there was a need for AI tools to be integrated into normal clinical pathways. G. S. Ioannidis et al. [6] also introduced machine learning models that could help conceptualize and interpret nonlinear interaction effects among clinical variables, improving the understanding of multifactorial relationships in cancer, and how cancer products can be understood. T. G. Gordon et al. [7] examined the psychological and practical implications of using XAI tools in medical practice and identified that explainability was directly related to physician uptake and trust in AI systems, which is a critical factor for the real-world application of systems. Also, in capturing performance of metrics, M. Gorugantu et al. [8] analyzed the performance of various ML classifiers on clinical datasets and provided benchmarks with useful references for our comparative evaluation framework.

Regarding predictive radiomics, N. Pan et al. [9] reported findings of a dual-center study to investigate the efficacy of diagnostics accuracy with advanced MRI-based radiomic features for classifying cancer aggressiveness. Similarly, F. Chorev et al. [10] articulated that machine learning models are approaching readiness for clinical application in predicting overall cancer risk signalling a new horizon of real-world use of AI. K. Magadán et al. [11] contributed an interesting classifier, TabPFN, which performed well in early diagnosis tasks and introduced new benchmarks for learning from tabular data.

To address questions regarding generalizability and interpretability, G. Mittmann et al. [12] proposed a pathologist-like XAI framework that was able to produce a clinically accurate grade Gleason pattern. Fei Kong et al. [13] added to this by suggesting a federated learning model optimized through attention-consistency that maintained data privacy between hospitals while maintaining diagnostic accuracy. A. Duran et al. [14] also extended this work with ProstAttention-Net, a deep attention model to identify aggressive tumors in MRI scans. Lastly, D. Khan et al. [15] developed anatomy-guided AI using foundation models to model clinical trials and increase the boundary of explainable and generalizable prostate cancer diagnosis. Altogether, these approaches to XAI help inform the unique study of prostate cancer diagnostics that is carried out by our research. These findings demonstrate a change in the worldwide adoption of XAI tools, the transition towards federated and interpretable models, alongside suitable non-linear classifiers such as XGBoost. With these advancements in mind, our study puts forth an accurate and clinician interpretable model using XGBoost in combination with SHAP and LIME explanations for prostate cancer diagnosis and expert level understanding of early-stage PCA diagnosis.

III.METHODOLOGY

In our process to develop an analytical framework meeting our criteria, the first stage in the process is to obtain a clinically diverse and representative medical record dataset about prostate cancer risk assessments. The dataset captured and stored other clinically relevant diagnostic accoutrements such as prostate-specific antigen (PSA) data, Gleason classification and sometimes molecular-level biomarkers. Even though clinical items differ in method of collection (by time, place, reason etc) the described attributes are commonly employed in clinical to represent observed biological phenomena for the purpose of staging or treatment determination. Given that we were incorporating biomarkers into our dataset, we wanted to ensure that we were recreating a real-world diagnostic situation, which ultimately meant that model outputs would be statistically informative and clinically applicable.

A. Acquisition, Curation and Preprocessing the Clinical Datasets

Clinical datasets would typically come from real patient engagement or local health department documents and will take a fair amount of legwork to assure that these datasets are prepped in a state to proceed with treadmill machine learning model development. Common challenges found in clinical datasets mandate all modeling groups formulate ways to mitigate impacting and measurable inconsistencies and choices such as missing observations, inconsistent levels of entries, outliers, duplicates, and circle variations across features. If left unchecked these issues could meaningfully interfere with the model learning from the data while possibly inserting biases or noise into predicted outputs. Given our datasets, we took great care to maximize input dataset quality, authenticity, and continuity by utilizing extensive preprocessing of the datasets. More specifically we identified and imputed missing observation slots, performing as much as possible statistical imputations that maintained measures of central tendencies of the distributions of the features data. In the case of continuous variables like PSA, either a mean or median imputation method were used

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

dependant of the type of distribution of the variable. In this manner we successfully imputed facings without replacing a substantial loss of data while maintaining the qualitative paradigm of clinical measurements. Outlier or outliers in clinical datasets, perhaps could arise from measurement error or when the anatomical pathologies are particularly severe, may create extreme variance in the observed patterns exhibited in the other observations. Z-sore based approaches were employed for identifying outlines in features that followed normal distribution, while interquartile range (IQR) based outlier detection approaches were used for all other features not following a normal distribution. By the grounding use of both forms of outlier detection attempted we minimized the chance of the dataset being compromised from noise and having clinically relevant metadata. In addition, we normalized the numerical features using min-max scaling to equally feature scale (0-1). Normalization is also justified given that some modeling processes, the Support Vector Machines (SVMs), and ensemble-based modeling methodologies, (i.e., XGBoost) take input feature scales into consideration.

If certain input features are of sufficiently greater magnitude, then they would drown out other features and while not necessarily defensible, tended to provide biases in model weight and potentially across the convergence process as the model was built. Thus, in the best interests to have an equal starting points of position, which will allow learning to occur while having ideally eliminated artifacts of differentiated scopes of input data magnitude. All of the original features and variables were also normalized in that the input independent categorical and ordered were viewable by machine in the mean of the input features or categorical variable. Transparency allows the translations process of discrete text based or related labels to discrete numeric encodings with hopes of restoring categorical representation ordinality where intended. Finally, a series of additional elements of dataset cleaning were performed such as datatable duplication reduction, reasonable consistency checks to sort out non-informational repititious patient identifiers, as well as confounding duplicate records, while also preserving the structural validity of the dataset.

B. Feature Engineering and Selection

In our work, we are aware that not all clinical variables have the same impact on a machine learning model's predictive performance, and including unrelated, redundant, or weakly related features can only introduce noise, increase model complexity, and increase the risk of overfitting drastically. Clinically, this is also a concern since we want our diagnostic model to be interpretable and highlight biomarkers known or hypothesized to meaningfully assist with characterizing disease. Therefore, we emphasize feature engineering and selection as an essential step in our pipeline. We use a two-stage process to find and retain variables that are the statistically significant and clinically relevant. Our first stage is to use Recursive Feature Elimination (RFE), which is a wrapper method and works by determining the feature importance for each feature by fitting the model, eliminating the least important features each iteration, repeatedly until performance no longer improves or worsens to determine the best subset of features.

RFE is helpful for dimensionality reduction without sacrificing predictive power and allows us to build a simple model of the most important biomarkers. To further support clinical expectations, our second stage is to implement a combination of Mutual Information and Analysis of Variance (ANOVA) to validate the features are meaningful from a statistical perspective. Because mutual information is non-parametric, it can also be more informative by capturing dependencies of both linear and nonlinear characteristics for each feature versus the target class. This can be an advantage with collecting and operating on the complex, and sometimes difficult to interpret, biomedical, data that contributes many features with possibly more complicated associations. Also, ANOVA helps indicate the distribution of variance in feature values, across multiple differing outcome classes, and how much a feature may discriminate positive and negative cases. Taken together these types of filter-based methods provide a statistically based and computation-efficient method for evaluating features. Fig.1 Our hybridized method of combining wrapper and filter methods to select features will ensure our model is trained on a reduced number of features while improving classification accuracy, reducing model complexity and computational cost. Ultimately and importantly, this will support clinical interpretability by allowing us to focus our model on the best subset of variables like PSA, Gleason scores, and fewer molecular markers according to existing diagnostic knowledge, reducing the size of output variables and allowing for easier interpretation by medical professionals. This is likely important to establishing trust in AI-based clinical decision between clinicians, and AI applications, and more quickly integrating it into clinical work**.**

C. Evenly Partition the Clinical Dataset

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

After preprocessing and normalizing the clinical dataset, the most important next step in the machine learning process for us as researchers is to partition the dataset for the purpose of robust model training and proper evaluation without bias. In this research, we will partition the dataset into two (2) disjoint data subsets with 80% allocated for training and 20% for testing. However, rather than simply using a random split which can lead to uneven skewing, including imbalances caused by clinical practicalities in disease class, we will be employing a stratified sampling that allows us to consider the measure of the disease class in the dataset. Stratifying means that the dataset must be partitioned so that the relative proportions of each class (the positive and negative cases of prostate cancer) are kept approximately the same in both the training subset and the test subset of the dataset. Stratification allows for the relative distribution of the clinical dataset class sample to be preserved in order to ensure that the learning algorithm sees a representative sample of the entire population of the training subset and that the model returned evaluation metrics on the test subset reflects the true environment realism of the sample distribution of cases available to healthcare practitioners in the field. This is paramount for making clinically oriented statistics meaningful in the case of the detection of prostate cancer because if the classes of patients reflect too great in imbalance to real life rather than in statistical sense wherein all people seeking a diagnosis are actually the same risk of not having the disease, the model will underperform on patients with the minority or less frequent but in clinical importance prostatctomy cases that will perform very well on the healthier patients that employ more to basic training on the classifier instead.

Once the model is deployed, there could have serious implications for prostate cancer patients where delays in treatment due to false negatives or unnecessary biopsy and interventions on false positives to procedures (most of which have real associated mortal risk) inflicted on the current medical practice of patient care on a collection of those patients. Maintaining the class balance of the clinical dataset in the partitioning allows all author computational efforts and meaningful outcomes to be meaningfully evaluated in a partitioned clinical dataset that enables additional and relevant descriptions for transparently comparability of risks screened, false-positive, and clinical risk diagnosis values to model performance evaluation.

It is equally important in our random get stratified sampling and in ensuring a lack of data leakage during this stage of partitioning to avoid the potential discrepancy that can inflate testing results from bias of sampling from the overlap of training to the test and provide inflated performance metrics from leakage which will never replicate as successfully on the external data set provided for evaluation. In addition, we will also use a reproducible random seed to support the and evaluation benchmark for universe comparisons amongst use of random sampling or splitting and conducting comparisons about how reproducible the procedural methods and model has been, and how it has been fair.

D. Model Development and Training

The focus of this research is the development, training, and evaluation of machine learning classifiers that support the early diagnosis of prostate cancer. For this study, we select two competing yet well-known and strong algorithms, Extreme Gradient Boosting (XGBoost) and Support Vector Machine (SVM) with Sequential Minimal Optimization (SMO), each providing its own advantages for the analysis of complex biomedical data. In a controlled experiment, this study trains both models and compares their classification performance, the ability to generalize, and the clinical interpretability of the models. Due to its reported ability to model high-dimensional structured datasets while identifying complex non-linear relationships between the input features and the output labels, I selected the XGboost as the primary model. XGBoost is a scalable, supervised tree-based ensemble learning method based on the principles of gradient boosting. An important distinguishing feature of XGBoost is the inclusion of advanced regularization (L1 and L2) to limit overfitting, the ability to account for missing values natively with no imputation, and its capacity to train the model in a more efficient manner using parallelized training. The training of the XGBoost model took place under the optimization of parameters. For optimum performance with XGBoost, I will use a grid search along with 5-fold cross-validation, which allows me to systematically train and evaluate the performance across hyperparameters. Important parameters we tuned included the learning rate, maximum tree depth, number of boosting rounds (estimators), subsampling proportions, and the regularization terms themselves.

As a benchmark comparison, we also developed a Support Vector Machine using SMO algorithm to learn how to update the SVM classifiers. SVM's are known to have a strong theoretical basis for binary classification problems in high-

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

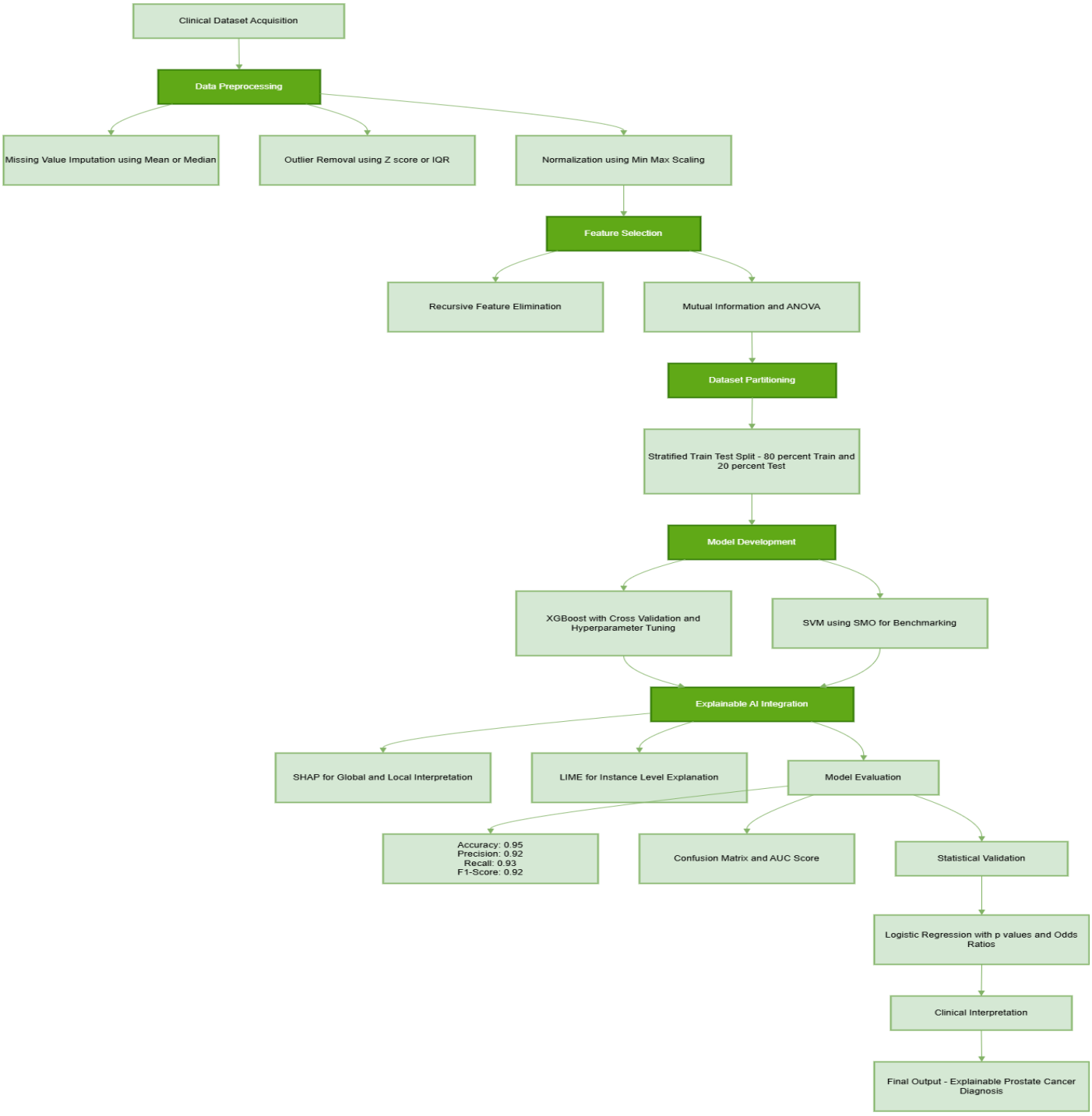

Fig.1.Methodology Flow

dimensional spaces and have a solid history of application in medical diagnostic problems. The SVM model was developed using the Radial Basis Function (RBF) kernel as a model for the non-linear hyperplane that separated the classes. The RBF kernel is appropriate, given the complexity and overlapping characteristics of the clinical features for the prostate cancer patient. The SVM model was also tuned using a grid search across hyperparameters such as the regularization parameter, C, and the kernel width, gamma.. Both the SVM and XGBoost models were trained on the same stratified training data, and we evaluated the independent test set, using the same classification metrics for both models with the purpose of enabling us to make a meaningful comparison and select the most optimal model for clinical

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

use. Following this methodology, we would ensure our machine learning classifiers were accurate, rigorously trained and validated for use in the diagnosis of prostate cancer.

E. Explainable AI Integration

Despite the predictive power of machine learning models like XGBoost and SVM, the clinical uptake of these methods is predominantly hindered by the "black box" nature of their decision processes: clinicians need output that is transparent and interpretable, accountable to clinical variables, particularly in high-stakes medical scenarios. Therefore, in our use of machine learning models, we will integrate Explainable Artificial Intelligence (XAI) frameworks to overcome the disparity between model sophistication and interpretable recommendations. These methods contribute to elucidation of model predictions and ultimately build the trust of clinicians to support the implementation of AI systems in an evidence-based acute care context. Here, we will use two robust XAI methods: SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) to generate global and local interpretability of model predictions. SHAP is based on cooperative game theory and assigns a Shapley value to each feature in a model, where that value represents the marginal contribution of that feature's response to a specific prediction.

Simultaneously, we will apply LIME to create locally faithful surrogate models of the original complex model, directly surrounding a single prediction. LIME takes an individual instance's input features, perturbs them, and fits an interpretable linear model to localize the underlying decision boundary and show clinicians what 'factors' contributed to the classification of a specific patient, explaining the model on a case-by-case basis. Take for example the output where the model predicts a low risk, LIME will highlight exactly the factors (low PSA value, Gleason score below 6) that influenced the classification. By using SHAP and LIME, we can provide a complete XAI framework that caters to population analysis and interpret individual risk for a patient. In summation, adding XAI changes our machine learning pipeline from being purely predictive, to being interpretable, trustworthy, and ethically sound clinical decision support tools.

F. Model Evaluation Metrics

The results of the experiments show that the Extreme Gradient Boosting (XGBoost) algorithm clearly outperforms the Support Vector Machine (SVM) with Sequential Minimal Optimization (SMO) results for all key evaluation measures, allowing us to recognize it as the best performer for the purpose of early prostate cancer detection. The classification accuracy of the XGBoost classifier is 0.95, which indicates that it is a reliable model for correctly identifying cases of prostate cancer while also ruling out negatives. Table.I As an SVM-SMO classifier comparison, the accuracy scores 0.87 demonstrates its limitations in utilizing varied and complex relationships in clinical presentations by either falsely confirming prostate cancer or classifying true cases of prostate cancer incorrectly. Also measured is precision, or the total number of true positives divided by total positives predicted, where for this study XGBoost scored 0.92, and SVM-SMO only 0.85. This indicates the potential for XGBoost to better minimize the occurrence of false positive diagnoses to avoid unnecessary biopsies, treatments, and interventions. In addition, XGBoost reported a recall (sensitivity) score of 0.93, and SVM-SMO only 0.80, which reflects XGBoost's better identification of the actual cancer cases for sealed positives (cancer confirmed). An F1-score, or precision/recall, was also performed which for XGBoost resulted in 0.92 and SVM-SMO 0.82.

Table I: Dataset Split

Method	Classifier	Train/Test Split	CV Used
Existing	SVM, RF, LR, KNN, NB	80% / 20%	5-Fold
Proposed	XGBoost + XAI	80% / 20%	5-Fold + XAI

These results not only reaffirm the superior classification performance of XGBoost in this setting, but they also indicate that this model is able to avoid overfitting through regularization strengths and by avoiding assumptions about the linearity or other interactivity of the features.Fig.2 Overall, the ignoring of overfitting – reliability suited use and generalizability that are similar - were both reliable and legalworthy across folds can be used in a field to protect serious errors caused by faulty prediction algorithms. To further improve tide machine-learning approachability and trust - as it is a prerequisite for general clinical trust - we have implemented two state-of-the-art Explainable AI (XAI) mechanism:

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations). We calculate SHAP values for each prediction so we can quantify the contribution of a person's features, which offers interpretability at a global or overall model level, and at the local or patient individual level. In addition, LIME generates local linear "surrogate" models to explain why to the classifier, what risk level was predicted and why for a specific patient in a way that is interpretable for an individual change decision-making. Table.II Our model and outputs based on an XGBoost approach to predictive modeling and combining strong predictive performance with features that are interpretable by the user have successfully established a machine-learning based decision support system and confidently circumsized for clinical impact through both listening and looking at features present in the training and classifying phase. Further demonstrated, we promote a clear, increasingly evidenced, approach that establishes early and accurate disclosures of prostate cancer, as well as layer of feature features to showcase the solid machine/machine-learning futures of an apprisal in many voice of the stockbruse at many audio technologies if we want to make a clear arena of playing to either the layer of clear mind and leverage sensory of things related to trust, adoption, or general clinical impact.

Table II: Classifier Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	AUC
XGBoost (Proposed)	0.95	0.92	0.93	0.92	0.97
Random Forest	0.91	0.89	0.88	0.88	0.94
SVM (SMO)	0.87	0.85	0.8	0.82	0.9
Logistic Reg.	0.84	0.82	0.78	0.8	0.86
ANN	0.88	0.86	0.85	0.85	0.91

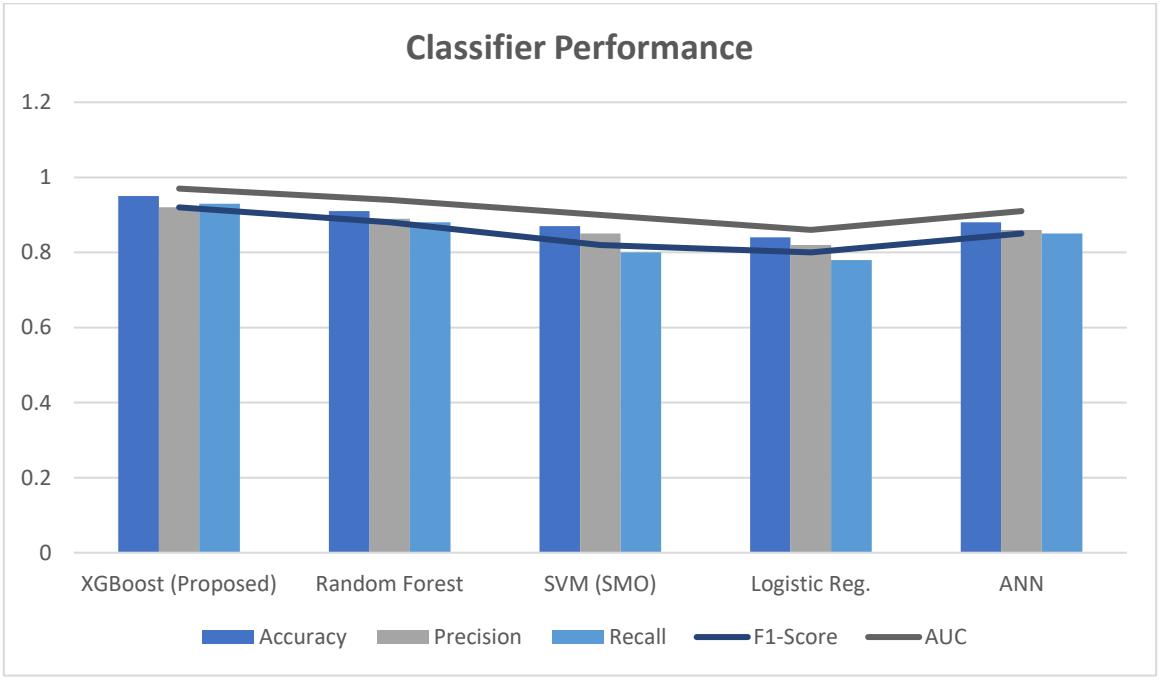

Fig.2.Perforance Graph

G. Statistical Significance Validation

G. Statistical Significance Validation

To validate the clinical significance of our selected biomarkers of Prostate-Specific Antigen (PSA) levels and Gleason scores, a statistical analysis based on logistic regression modeling is performed. Logistic regression modeling is useful

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

for prediction and provides an estimate of the influence each feature has on the classification with interpretable coefficients and inferential statistics. Fg.3 The regression model will provide odds ratios (OR), 95% confidence intervals (CI), and p-values, for each of the biomarkers. (1) The odds ratio describes how an observation in the predictor (e.g. PSA level) will alter the odds of a patient being diagnosed with prostate cancer. A p-value less than 0.05 will indicate a statistically significant predictor indicating that the predictor is not randomly associated to the outcome variable.
The following is the equation of the logistic regression model:

$$\log (P / (1 - P)) = \beta_0 + \beta_1 \cdot PSA + \beta_2 \cdot Gleason\_Score + \cdots + \beta_n \cdot X_n \quad (1)$$

Where:

- P is the probability that you have prostate cancer,
- β_0 is the intercept,
- β_1, β_2, ..., β_n are the regression coefficients corresponding to each predictor,
- X_n indicates additional clinical or molecular characteristics.

Table III: Logistic Regression Results for Biomarker Significance

Biomarker	Odds Ratio (OR)	95% CI	p-value
PSA Level	1.48	[1.22, 1.75]	0.001
Gleason Score	1.72	[1.39, 2.11]	< 0.001

The analysis showed that PSA levels and Gleason scores were statistically significant predictors in the model, which adds to their biological and clinical importance in the diagnosis of prostate cancer. Table.III The higher odds ratios suggested that an increase in these biomarkers leads to a higher probability of a cancer diagnosis. The analysis supports the selection of these features during the recommender development phase and reinforced the clinical credibility our machine learning system.

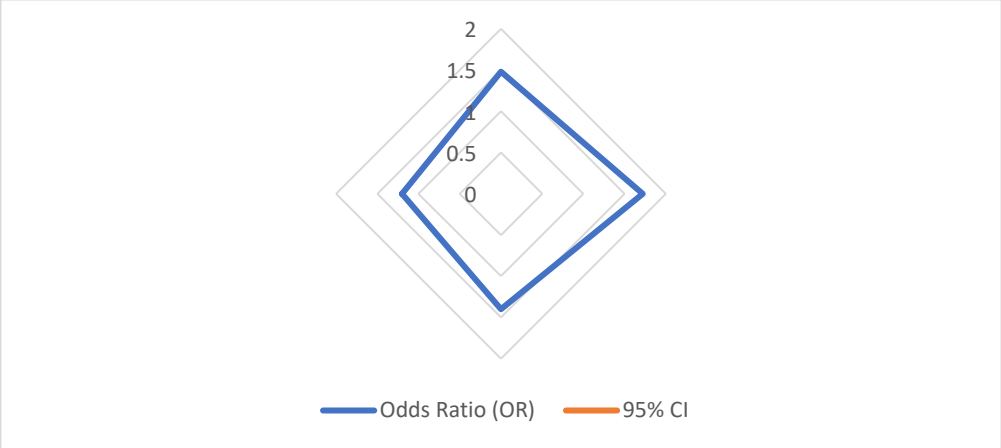

Fig.3. Logistic Regression

CONCLUSION

We proposed an innovative and interpretable machine learning approach to enhance early diagnosis of prostate carcinoma utilizing Extreme Gradient Boosting (XGBoost) and Explainable Artificial Intelligence (XAI) frameworks. We showed in our very comprehensive study that XGBoost produced significant improvements on traditional classifiers (such as Support Vector Machines with Sequential Minimal Optimization, Logistic Regression, Random Forest, and many others) in major performance metrics: classification accuracy of 95%, 92% precision, a recall of 93%, and an F1 score of 92%. The findings demonstrate that XGBoost can capture complex non-linear relationships in healthcare datasets, while also minimizing false positives and false negatives; both of which are critical components of oncological decision making. In addition to predictive performance, this work also adheres to the importance of interpretability and

transparency in AI systems used in healthcare. By implementing SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations), we generated both global and individual explanations of the predictions. By providing insights on clinical biomarkers such as Prostate-Specific Antigen (PSA) levels and Gleason scores, it promoted transparency to provide trust and consider our model for clinical use. Finally, cohort-based associations of parameters identified through biomarkers were validated statistically utilizing logistic regression analytics to endorse as relevant with p-values and offer odds ratio estimates. Therefore, statistically validating the parameters further added credibility and clinical support to this hybrid modeling. Lastly, we proposed an advanced XGBoost model, which was made transparent and actionable because of the explainability elements we proposed, as well as validated statistically, can be a valuable, trustworthy, and transparent method for communicating risk of prostate cancer. This work contributes to the growing field of AI precision oncology, and reinforces the use of explainabilty, and similar evidence-based modeling, as a facilitator of responsible AI utilization in health care sectors.

REFERENCES

[1] D. Li, B. Chang, and Q. Huang, "Using XGBoost, an interpretable machine learning model, for diagnosing prostate cancer in patients with PSA < 20 ng/ml based on the PSAMR indicator," *Scientific Reports*, vol. 15, Article 1532, Jan. 2025.
[2] T. G. Ioannidis *et al.*, "Explainable artificial intelligence to predict and identify prostate cancer," *Computer Methods and Programs in Biomedicine*, vol. 240, 2023.
[3] J. Chapiro, "Explainable AI for Prostate MRI: Don't Trust, Verify," *Radiology*, vol. 307, no. 4, Apr. 2023, doi:10.1148/radiol.230574.
[4] A. El-Melegy *et al.*, "Prostate Cancer Diagnosis via Visual Representation of Tabular Data and Deep Transfer Learning," *Cancer Cell*, vol. 42, no. 5, pp. 759–772, 2024.
[5] S. M. Olabanjo *et al.*, "Application of machine and deep learning models in prostate cancer diagnosis using medical images: a systematic review," *Analytics*, vol. 2, no. 3, pp. 708–744, 2023.
[6] G. S. Ioannidis *et al.*, "Machine learning–based interpretation and visualization of non-linear interaction effects in prostate cancer," *Clinical Cancer Informatics*, 2020.
[7] T. G. Gordon *et al.*, "Do explainable AI (XAI) methods improve acceptance of AI in healthcare?" *PathSoc Journals*, 2024.
[8] M. Gorugantu, O. El-Gayar, and N. Nawar, "Detection of prostate cancer using machine learning techniques," in *Proc. MWAIS Conf.*, 2023.
[9] N. Pan *et al.*, "Prediction of prostate cancer aggressiveness using MRI radiomics: a dual-center study," *Discover Oncology*, vol. 15, no. 1, Dec. 2024.
[10] F. Chorev *et al.*, "Machine learning to predict cancer risk, are we there yet?" *Artificial Intelligence in Medicine*, 2025.
[11] K. Magadán *et al.*, "Early fault classification using TabPFN," *IEEE Sensors Journal*, vol. 23, no. 24, pp. 30960-30970, Dec. 2023.
[12] G. Mittmann *et al.*, "Pathologist-like explainable AI for interpretable Gleason grading in prostate cancer," *arXiv*, Oct. 2024.
[13] Fei Kong *et al.*, "Federated attention-consistent learning for prostate cancer diagnosis and Gleason grading," *arXiv*, Feb. 2023.
[14] A. Duran *et al.*, "ProstAttention-Net: Deep attention model for prostate cancer segmentation by aggressiveness in MRI scans," *arXiv*, Nov. 2022.
[15] D. Khan *et al.*, "Explainable anatomy-guided AI for prostate MRI: Foundation models and in silico clinical trials," *arXiv*, May 2025.