DAYANANDA SAGAR UNIVERSITY

Devarakaggalahalli, Harohalli

Kanakapura Road, Ramanagara - 562112, Karnataka, India

SCHOOL OF ENGINEERING

Bachelor of Technology

In

Computer Science and Engineering

Minor Project Report

AI-Driven Urine Analysis for Non-Invasive Kidney Stone Detection: A Machine Learning Approach

Batch: 59

By

Srujan S Shetty - ENG22CS0469.

Tejas M I - ENG22CS0481.

Venugopal - ENG22CS0498.

Tejasvi D - ENG22CSO482.

Under the supervision of

Dr. Praveen Kulkarni

Associate Professor

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING,

SCHOOL OF ENGINEERING

DAYANANDA SAGAR UNIVERSITY

(2025-2026)

DAYANANDA SAGAR UNIVERSITY

Devarakaggalahalli, Harohalli

Kanakapura Road, Ramanagara - 562112, Karnataka, India

SCHOOL OF ENGINEERING

Department of Computer Science & Engineering

CERTIFICATE

This is to certify that the Minor Project titled **"AI-Driven Urine Analysis for Non-Invasive Kidney Stone Detection: A Machine Learning Approach"** is carried out by **Srujan S Shetty (ENG22CS0469), Tejas M I (ENG22CS0481), Venugopal (ENG22CS0498), Tejasvi D (ENG22CS0482)** bonafide students of Bachelor of Technology in Computer Science and Engineering at the School of Engineering, Dayananda Sagar University, Bangalore in partial fulfillment for the award of degree in Bachelor of Technology in Computer Science and Engineering, during the year **2025-2026**.

Dr. Praveen Kulkarni	**Dr. Girisha G S**	**Dr. Uday Kumar Reddy K R**
Associate Professor Dept. of CSE, School of Engineering Dayananda Sagar University	Chairman, CSE School of Engineering Dayananda Sagar University	Dean School of Engineering Dayananda Sagar University
Date:	Date:	Date:

Name of the Examiner **Signature of Examiner**

1.

2.

DECLARATION

We, **Srujan S Shetty (ENG22CS0469), Tejas M I (ENG22CS0481), Venugopal (ENG22CS0498), Tejasvi D (ENG22CS0482)** are students of the sixth semester B.Tech in **Computer Science and Engineering**, at School of Engineering, **Dayananda Sagar University**, hereby declare that the Minor Project titled **"AI-Driven Urine Analysis for Non-Invasive Kidney Stone Detection: A Machine Learning Approach"** has been carried out by us and submitted in partial fulfilment for the award of degree in **Bachelor of Technology in Computer Science and Engineering** during the academic year **2025-2026**.

Student **Signature**

Srujan S Shetty
ENG22CS0469

Tejas M I
ENG22CS0481

Venugopal
ENG22CS0498

Tejasvi D
ENG22CS0482

Place:

Date:

ACKNOWLEDGEMENT

It is a great pleasure for us to acknowledge the assistance and support of many individuals who have been responsible for the successful completion of this Minor Project.

First, we take this opportunity to express our sincere gratitude to School of Engineering & Technology, Dayananda Sagar University for providing us with a great opportunity to pursue our Bachelor's degree in this institution.

We would like to thank **Dr. Uday Kumar Reddy K R, Dean**, **School of Engineering & Technology**, **Dayananda Sagar University** for his constant encouragement and expert advice. It is a matter of immense pleasure to express our sincere thanks to **Dr. Girisha G S, Chairman**, **Department of Computer Science, and Engineering**, **Dayananda Sagar University,** for providing the right academic guidance that made our task possible.

We would like to thank our guide **Dr. Praveen Kulkarni Associate Professor**, **Dept. of Computer Science and Engineering**, **Dayananda Sagar University**, for sparing his valuable time to extend help in every step of our Minor Project, which paved the way for smooth progress and the fruitful culmination of the project.

We would like to thank our Minor Project Coordinators, **Dr. Gousia, Dr. Damodharan D** and all the staff members of Computer Science and Engineering department for their valuable support.

We are also grateful to our family and friends who provided us with every requirement throughout the course. We would like to thank one and all who directly or indirectly helped us in the Special Topic.

TABLE OF CONTENTS

NOMENCLATURE USED

ML	Machine Learning
SVM	Support Vector Mission
XG Boodt	Extreme Gradient Boosting

LIST OF FIGURES

ABSTRACT

The urological problem of kidney stones becomes a serious health concern when left untreated that may lead to unbearable pain and important medical complications. Patients often face two issues when using traditional diagnosis methods such as CT scans and ultrasounds: these tests cost a large amount of money and expose people to radiation while also being unavailable in areas that lack resources. The research evaluates a machine learning solution which examines urine test results to identify kidney stones because it aims to develop a cost-effective diagnostic method that operates through non-invasive methods with open accessibility. The research examines four machine learning prediction models which use urine analysis data to predict kidney stones by evaluating Logistic Regression, Support Vector Machine and Random Forest and XGBoost. Six important urine characteristics including specific gravity, pH, osmolality, conductivity, urea and calcium are tested in this analysis. A preprocessing method normalized the data while performing exploratory analysis to maintain high data quality for the 1000 samples within the dataset which contained controlled labels for kidney stone identification. Models used 80% of the available data to establish training networks that were evaluated on the unutilized 20% portion. The Random Forest model achieved the superior performance among all models with an accuracy of 0.980 and precision of 0.989 then recall of 0.968 and F1 score of 0.978 and AUC-ROC of 0.986. The XG Boosting model had slightly better results but still had an Accuracy of 0.9750, Precision of 0.9684, Recall of 0.9787, F1 Score of 0.9735 and AUC-ROC of 0.844. The research presents promising findings regarding affordable non-invasive laboratory diagnostic tests for kidney stones which lead to more efficient medical care.

Keywords - non-invasive diagnosis, Random Forest, Logistic Regression, Support Vector Machine, XGBoost, urine analysis, kidney stone detection, machine learning techniques.

CHAPTER 1
INTRODUCTION

The human kidneys create solid deposits of minerals and salts identified as kidney stones because of dehydration together with metabolic imbalances and dietary choices. Prolonged uncertainty about stones in the kidneys results in intense pain along with urinary tract infections which potentially cause significant harm to the kidneys unless quick medical treatment occurs. Traditional methods for kidney stone examination include CT scans together with ultrasounds and X-rays. These diagnostic techniques are expensive while additionally they subject patients to radiation exposure and also face limited availability especially in places with limited resources. The requirement exists for affordable and non-invasive approaches for diagnosis.

The current development of artificial intelligence enables image-based analysis to become an advanced tool that detects kidney stones. Medical imaging data analysis performs well for kidney stone detection and classification using Convolutional Neural Networks which combines machine learning and deep learning techniques on CT scans and ultrasounds data. Automated technology systems employ their programming to detect stone features as well as stone types which results in better diagnosis accuracy and less human interpretation dependence. AI-powered solutions provide a practical imaging methods replacement since they enhance the detection accuracy of kidney stones by using image processing methods that involve segmentation and feature extraction and image enhancement techniques.

In this research project a machine learning solution detects kidney stones through processing urine analysis data instead of conventional imaging systems. The analysis applies Random Forest as well as Logistic Regression alongside Support Vector Machine (SVM) and XGBoost to identify patients based on urine parameter tests involving pH, specific gravity, osmolality, conductivity, urea and calcium measurement results. The dataset went through extensive normalization routines for data preparation before researchers executed exploratory data analysis to reach maximum standards of data quality. Performance evaluations of patient outcome predictions emerged as an outcome of optimization processes conducted on all developed prediction models. The research reveals machine learning tools demonstrate prospects to function as an

economical alternative to knife diagnostics that enhances preliminary diagnosis outcomes alongside providing trustworthy alternatives to conventional testing approaches. Clinical staff will add further development to execute clinical validation tests before implementing the system at standard healthcare clinics.

CHAPTER 2
PROBLEM DEFINITION

The project aims to design and develop a non-invasive, cost-effective, and AI-powered diagnostic system for the early detection of kidney stones based on urine analysis data. Traditional methods such as CT scans and ultrasounds, though effective, are expensive and expose patients to radiation, making them less feasible for routine or rural diagnostics. This project leverages machine learning algorithms to analyze biochemical urine parameters—such as pH, specific gravity, osmolality, conductivity, urea, and calcium levels—to classify the presence or absence of kidney stones.

The goal is to build a reliable and accurate machine learning model that can assist healthcare professionals in making informed decisions without requiring imaging techniques. By using various classification models including Logistic Regression, Support Vector Machine (SVM), Random Forest, and XGBoost, and evaluating their performance using metrics like accuracy, precision, recall, and AUC-ROC, the project seeks to determine the most effective model for clinical use. The final system aims to be easily integrated into a lab environment or web-based interface, facilitating rapid and non-invasive kidney stone detection.

CHAPTER 3
LITERATURE REVIEW

Medical researchers created a kidney stone detection model based on deep learning through the combination of CNNs and SVMs to analyze CT scan images that did not require contrast dye for better accuracy. [1] This detection system used detailed CT imaging with hybrid modeling technologies but currently exists solely within imaging environments. The collective application of image processing techniques with classification methods brings value to this study because it enables the development of adaptable non-imaging data processing methods.

The combination of Guided Bilateral Feature Detector technology with SVM enabled automated kidney stone detection from CT images producing 98.56% accuracy along with 99% F1 score performance. [2] Diagnostic reliability receives significant power enhancement through image enhancing approaches according to this study. The clinical achievement of this method is restricted by high-resolution CT imaging needs which makes it unfeasible for cost constrained settings while the research focuses on urine examination techniques.

The utilization of CNNs in kidney stone detection automation allowed medical specialists to classify images from large datasets while improving diagnosis through reduced dependence on interpretation-based methods. [3] The method demonstrates the advantages of deep learning algorithms in workflow automation but such approaches introduce limitations due to their dependency on medical imaging systems. The automatic functioning of CNNs guides developers in building models for data-dependent approaches which include this proposed method.

The analysis of renal stone detection using deep learning models involving ResNet, DenseNet and EfficientNet found DenseNet to achieve an accuracy level of 0.86. [4] The research findings show different performance outputs between various DL architectures so it is essential to select models which deliver optimal results. The

methodology used at present relies on these findings as they guide the process of selecting ML models suitable for urine analysis tasks.

The researchers develop a kidney stone diagnostic system using machine learning which analyzes urine test parameters without invasive procedures. Research showed that Random Forest provided the best results with 94% accuracy from osmolality and urea features. [5] The effectiveness of biochemical information for diagnostic modeling verifies itself in the current research objectives and presents an economical replacement for CT-based detection.

The developed ML model showed a moderate ability to predict kidney stone recurrence using 24-hour urine information with an AUC value of 0.65. [6] Although the model's accuracy needs improvement the analysis of urine data provides essential clinical value for developing preventive machine learning applications in nephrology. The research findings demonstrate how the work can progress to conduct extended long-term observation beyond first detection abilities.

This research presented a non-invasive machine learning detection system which uses ensemble classifiers to analyze routine laboratory results together with urine analysis tests for kidney stone examinations in Saudi Arabia. [7] Standard clinical data proves diagnostically valuable for diagnostic purposes because it reduces healthcare costs and limits radiation exposure of CT imaging technology which aligns perfectly with the research goal of accessible AI-driven medical examination.

Kavoussi et al. demonstrated the application of ML for predicting urinary abnormalities that contributes to indirect stone diagnosis through factor-based modeling. [8] The predictive uses show how ML can develop beyond detection capabilities into preventive and monitoring functions which future versions of the proposed system could integrate.

Degadwala and Rathva provided a thorough evaluation of ML and DL approaches for kidney stone detection while examining image-based and data-powered system strengths in their work. [9] The authors validate mixed detection systems yet indicate that urine-based analytical models require further investigation because of their research importance.

Alqahtani et al. enhanced urinary-based kidney stone detection through optimization achieved by merging Binary Particle Swarm Optimization (BPSO) with XGBoost algorithms while performing feature selection. [10] The research directly affects the present work because it shows that combining traditional machine learning algorithms with metaheuristics enables significant diagnostic improvement in non-imaging settings.

CHAPTER 4
PROJECT DESCRIPTION

Kidney stones, or renal calculi, are solid masses formed from crystals in the urinary tract, often causing excruciating pain, urinary infections, and long-term renal complications if left untreated. Traditional diagnostic methods such as CT scans, X-rays, and ultrasounds are effective but suffer from significant limitations including high costs, radiation exposure, and limited availability in rural or under-resourced regions. This calls for a shift toward more affordable, accessible, and non-invasive diagnostic approaches that can be deployed widely. In this project, we aim to develop a machine learning-based diagnostic model that uses urine analysis data to predict the presence of kidney stones. The objective is to construct an intelligent system that leverages the biochemical properties of urine samples to classify patients effectively, minimizing dependency on expensive and invasive diagnostic methods.

To achieve this, we utilize data obtained from clinical urine analyses, specifically focusing on six key biochemical parameters: specific gravity, pH, osmolality, conductivity, urea, and calcium. These features serve as significant indicators of metabolic changes and are instrumental in differentiating between healthy individuals and those affected by kidney stones. The dataset used in this project comprises 1000 samples, each labeled with the presence (1) or absence (0) of kidney stones. A thorough preprocessing pipeline is applied to the raw data, which includes handling missing values, performing normalization, and conducting exploratory data analysis (EDA) to visualize distributions, spot outliers, and understand relationships between variables. This preprocessing ensures that the machine learning algorithms receive clean, consistent input data to learn effectively.

Following data preparation, we apply multiple supervised machine learning models to the processed dataset. The chosen models include Logistic Regression, Support Vector Machine (SVM), Random Forest, and XGBoost. Each model is trained on 80% of the dataset, while the remaining 20% is used for testing and performance evaluation. Hyperparameter tuning and cross-validation techniques are employed to enhance the generalizability and robustness of the models. Performance metrics such as accuracy, precision, recall, F1-score, and AUC-ROC are calculated to compare the effectiveness

of the models. The Random Forest model achieved the best performance with an accuracy of 98%, precision of 98.9%, recall of 96.8%, and an AUC-ROC of 0.986, indicating a strong balance between sensitivity and specificity. XGBoost also performed remarkably well with slight variations in metrics, showcasing the effectiveness of ensemble learning methods for this classification task.

An essential component of our system design is model interpretability and explainability. Through the use of feature importance rankings and SHAP (SHapley Additive exPlanations) values, we identify the most influential urine parameters contributing to the prediction. This step not only adds credibility to our model but also helps clinicians understand the rationale behind each decision, promoting trust in AI-based diagnostic tools. In the future, we aim to integrate this model into a web or mobile-based application that enables healthcare professionals and diagnostic labs to input urine test data and receive instant diagnostic feedback regarding the presence of kidney stones. Furthermore, this platform could store historical patient data and provide risk analysis over time, turning it into a comprehensive kidney health monitoring tool.

To enhance the clinical applicability of our approach, we will collaborate with medical institutions for real-world testing and validation of the system. With an increasing focus on point-of-care diagnostics, this AI-driven model could be adapted to work with portable urine analyzers or automated lab systems, thereby extending its use beyond hospital settings into rural clinics, telemedicine platforms, and home-based care systems. Moreover, the system can be trained further with larger and more diverse datasets to improve its adaptability to different populations and subtypes of kidney stones, making it a more inclusive and scalable solution.

Future work includes extending the project from binary classification (stone/no-stone) to multi-class classification, identifying specific types of kidney stones such as calcium oxalate, uric acid, or struvite stones. This could open avenues for personalized treatment plans based on stone type. In addition, combining urine analysis data with other low-cost diagnostic modalities such as blood tests or patient symptom profiles could further improve diagnostic precision. The current focus on non-invasive diagnostics through biochemical markers sets a precedent for similar work in other areas of nephrology and metabolic health.

In conclusion, this project offers a novel and practical solution to a pressing medical challenge. By combining the precision of machine learning with the accessibility of urine analysis, we present an affordable, non-invasive, and scalable system for early

detection of kidney stones. Our multi-model approach, supported by robust preprocessing and evaluation strategies, has demonstrated strong predictive capabilities, with the Random Forest and XGBoost models standing out in terms of accuracy and reliability. The integration of this system into clinical workflows has the potential to revolutionize diagnostic practices, particularly in under-resourced regions where access to advanced imaging technologies is limited. This work paves the way for intelligent, data-driven healthcare systems that can significantly enhance patient outcomes while reducing the burden on healthcare infrastructure.

CHAPTER 5
REQUIREMENTS

- **Programming Language:** Python 3.8+ (core language for ML implementation)
- **Libraries/Frameworks:**
    - **Scikit-learn:** For implementing ML algorithms (Logistic Regression, Random Forest).
    - **Pandas/NumPy:** For data preprocessing and analysis.
    - **Matplotlib/Seaborn:** For data visualization (heatmaps, distribution plots).
    - **Jupyter Notebook:** For interactive development and testing.

CHAPTER 6

METHODOLOGY

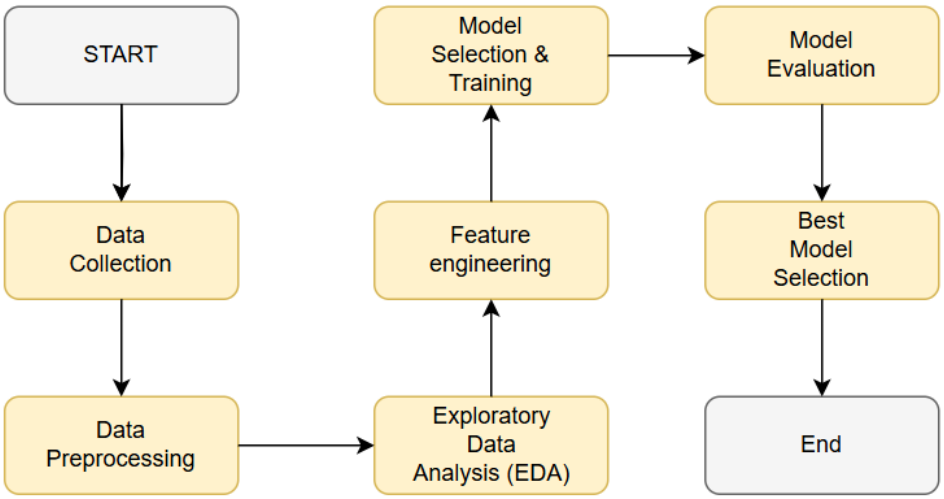


Figure 6.1: Work flow of the methodology

Our approach to solve the mentioned problems in our report are explained briefly with help of following steps:

1. **Data Collection**: Collected urine test records which contained biochemical measurements derived from patient samples. The six essential physiological aspects that are measured in the analysis are specific gravity along with pH, osmolality, conductivity as well as urea and calcium. The dataset contains a two-value target field that shows whether a particular urinary condition exists or not. (Figure 6.2)

	id	gravity	ph	osmo	cond	urea	calc	target
0	1	1.028	6.76	631	11.2	422	2.15	1
1	2	1.019	5.47	760	33.8	199	0.81	0
2	3	1.025	5.68	854	29.0	385	3.98	1
3	4	1.015	5.35	559	8.1	301	3.98	0
4	5	1.019	6.13	594	27.6	418	1.49	0
5	6	1.028	6.28	970	35.9	382	4.49	0
6	7	1.020	5.94	774	29.0	325	3.98	1
7	8	1.011	7.01	395	26.0	95	1.53	0
8	9	1.028	6.76	631	11.2	422	2.15	1
9	10	1.021	5.53	775	29.0	302	3.34	0

Figure 6.2. Features and information about the dataset

2. **Data Preprocessing**: Preliminary processing techniques were extensively applied to the dataset to achieve data cleaning along with transformation, normalization and imbalanced data resolution as well as several other techniques because real-world data typically does not match ideal machine learning requirements. The input data contains noise together with missing data that needs processing before developing accurate predictions.

3. **Exploratory Data Analysis**: The dataset underwent a thorough exploratory data analysis for extracting vital information about its nature.

The figure shown in Figure 6.3 displays the target variable distribution where samples with (0) or (1) values represent kidney stone presence or absence

respectively. The diagram displays information demonstrating 52.8% of the data points belong to class 0 and 47.2% belong to class 1. The proportional distribution between classes indicates that no significant changes need to be applied to the dataset for training effective machine learning models.

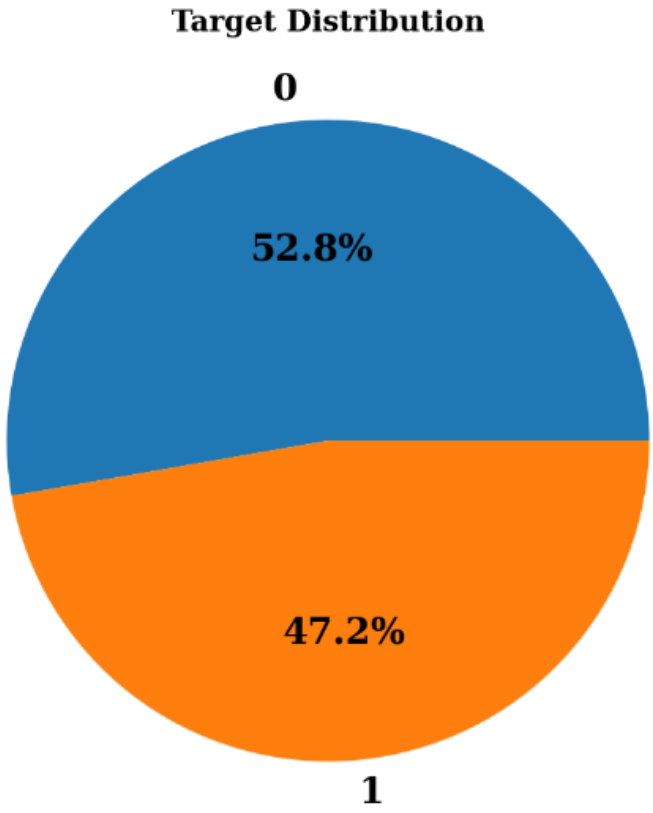


Figure 6.3. Stone distribution in dataset

The boxplot analysis in Figure 4 indicates that kidney stone patients demonstrate higher median values for all variables except pH including gravity, osmolality (osmo), urea and calcium concentrations (calc) in their urine samples. The difference between calcium as well as the urea_calc_interaction variable stands out as the most substantial indicating a potential link to stone development. The overall pattern in research data demonstrates that kidney stone occurrence appears linked to rising urinary concentration levels and distinct biochemical indicators although pH, conductivity (cond) and osmotic ratio measurements also vary mildly.

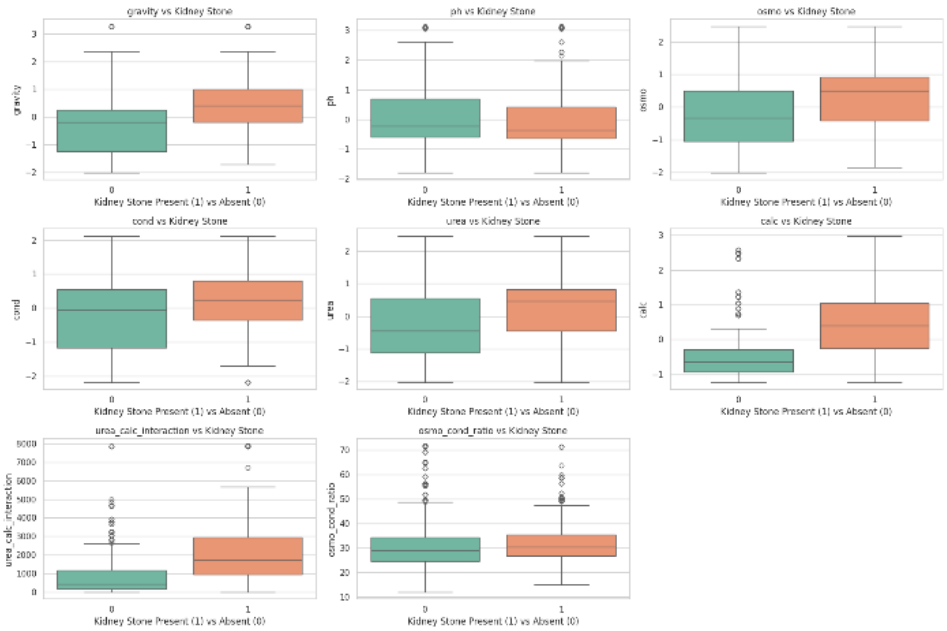


Figure 6.4. Boxplots of dataset feature (0-no stone and 1-stone)

The correlation heatmap (Figure 5) enabled us to determine how different urinary parameters relate to kidney stones. Data shows that urine calcium (calc) presents the highest positive association (r = 0.48) with kidney stone formation thus acting as a critical predictor. The risk factors for kidney stones appear to rise with higher concentrations of urine according to correlations with urea (r = 0.28) and specific gravity (r = 0.34). Urine pH demonstrates a minimal inverse connection (r = -0.12) so it contributes to stone formation to a lesser degree. Further evidence supporting the stone development process through concentrated urine came from the heatmap's display of strong connections between osmolality and conductivity as well as urea. The research results enable us to discover essential features which will be useful for developing exact machine learning model systems.

Figure 6.5. Heatmap correlation of dataset features

The pairplot analysis (Figure 6) indicates that urea, osmo and cond reveal the best capability for distinguishing target classes because class 1 displays separate clustering areas. The distributions of calc and gravity display average discrimination but ph data lacks effective separative capacity because both target classes overlap in the same areas. The correlation analysis shows osmo, cond and urea products tend to influence each other.

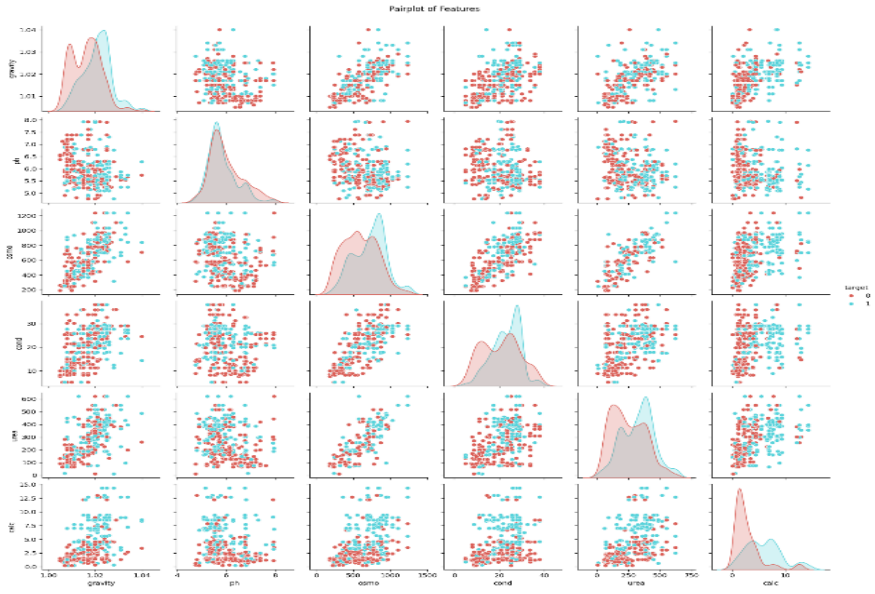


Figure 6.6. Pairplot of feature distribution

4. **Feature Engineering**: The model needed to better identify data patterns through implementation of multiple feature engineering approaches. The research incorporated interaction terms because it needed to analyze non-linear patterns by utilizing urea_calc_squared (urea × calcium²), osmo_urea_interaction (osmolality × urea), and cond_calc_ratio (conductivity ÷ calcium). The model combinations formed from clear correlations and their medical usefulness during the previous exploratory stage analysis.

   Urine pH obtained three clinically relevant categories from ≤5.5 (acidic), 5.6–6.5 (neutral) to >6.5 (alkaline) which were converted to one-hot encoding for model processing. We developed two additional markers to measure urine concentration called total_solid_score and calc_osmo_ratio which calculated the mean osmolality value with specific gravity and conductivity measurements and the calcium amount compared to osmolality measurements respectively. Reasoning for non-linear relations required the addition of features containing calc_squared and urea_sqrt terms.

   The irrelevant identifiers were eliminated prior to splitting the data into training and testing parts for subsequent scaling work to avoid data leakage effects. StandardScaler transformed all continuous numerical features to achieve consistent contributions during model training operations.

5. **Model Training**:

5.1 Logistic Regression.

Logistic Regression functions as a linear binary classification model which applies the logistic (sigmoid) function to evaluate binary outcome probabilities. The boundary decision exists as:

$$P(y = 1|X) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}} \qquad (1)$$

The model utilized linear solver alongside different values of C regularization strength and penalty type l1 or l2 for small dataset analysis. The model balanced class distribution through the use of class_weight = 'balanced'. Recall served as the primary metric for evaluation because the diagnostic system needed to detect all possible medical risks.

5.2 Support Vector Machine

With SVM we gain a hyperplane which creates the largest possible distance between two classes within a dimensional space. The optimization objective is:

$$\min \frac{1}{2}||w||^2 \quad \text{subject to} \quad y_i(w^T x_i + b) \geq 1 \quad (2)$$

SVM with rbf and poly kernels was used for detecting non-linear patterns while performing classification. Standardization ran as a pipeline to make all features equally impacting the result. A grid search method selected the regularization parameter C alongside the kernel coefficient γ (gamma). The introduction of balanced class weights served as a solution to deal with uneven distribution of samples.

5.3 Random Forest

Random Forest implements ensemble learning through the generation of numerous decision trees which combine results by using majority vote. The method serves to improve generalization abilities along with minimizing overfitting. The prediction is:

$$\hat{y} = \text{mode}\,(T_1(x), T_2(x), \ldots, T_n(x)) \quad (3)$$

Where the prediction from decision tree number n can be represented by $T_n(x)$. A process of optimizing model parameters including n_estimators, max_depth, and min_samples_split was performed. The system applied Balanced class weighting for distributing data points among positive and negative classes to achieve fairness.

5.4 XGBoost

XGBoost operates as a high-speed gradient boosting system which constructs sequential tree structures to find the minimum value in a loss function with incorporated regularization components.

$$L = \sum_i l(y_i, \hat{y}_i^{(t)}) + \sum_i \Omega(f_k), \;\; \Omega(f) = \gamma T + \frac{1}{2}\lambda ||w||^2 \quad (4)$$

Where:

The chosen criteria of l stands as the loss function (i.e. log loss for binary classification).

The value of Ω serves as a penalty which controls the complexity of the model.

T is the number of leaves

λ and γ control regularization strength

The model parameters max_depth and learning_rate together with subsample and colsample_bytree required optimization during the modelling process. The training process employed early stopping and evaluation metrics for the purpose of avoiding overfitting.

6. **Model Evaluation**: Testing and learning sections composed 20% and 80% respectively of the complete dataset for assessing all classifier models. Following metrics were used to evaluate the behavior of classification systems in the study.

- **Accuracy:**

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

- **Precision:**

$$\text{Precision} = \frac{TP}{TP+F\,P} \quad (6)$$

- **Recall:**

$$\text{Recall} = \frac{TP}{TP+F\,N} \qquad (7)$$

- **F1 Score:**

$$\text{F1 Score} = 2 \times \frac{(Precision \times Recall)}{(Precision \times Recall)} \qquad (8)$$

- **ROC and AUC Score:**

The receiver operator characteristic (ROC) curve shows the true positive rate (TPR) and false positive rate (FPR) at different threshold levels through its probability curve. A binary classifier performs best when depicted through ROC visualization. The area under the curve (AUC) determines a summary of the ROC curve. A model shows enhanced performance when it maintains a greater AUC value. A perfect classifier has an AUC value of 1. Within the same framework the random prediction yields 0.5 as its AUC measurement.

7. **Best Model Selection**: Model performance evaluation depends on key performance metrics which help establish effective medical diagnosis. Medical diagnosis requires Recall (Sensitivity) for detecting true positive cases together with precision for ensuring accurate positive predictions. Both F1-score and AUC-ROC serve to measure model reliability by balancing recall with precision and by determining class separation accuracy respectively. A confusion matrix serves as a detailed system that reports precise numbers of both correct and incorrect predictions. The best model selection relies on detection of missed diagnoses through higher recall values and better classification ability reflected by higher AUC-ROC scores and balanced F1-score and fewer false negatives combined with medical safety and practical utility for clinical use based on speed and interpretation ease for doctors.

CHAPTER 7

EXPERIMENTAL RESULT AND DISCUSSION

Figure 7.1 presents confusion matrices for model classification that displays the important FP (False Positive) and FN (False Negative) metrics which are vital to medical diagnosis. Kennel stone detection outcomes with false positives forces patients through unnecessary treatment and anxiety but wrong negatives might delay correct diagnoses thus worsening patient results. The Random Forest (RF) classifier achieved the most outstanding performance metrics by maintaining only three false positive (3) and one false negative (1) results. The Logistic Regression (LR) model produced 28 false positives and 30 false negatives which stood as the highest numbers compared to other models and this indicates its lower reliability. The performance of Support Vector Machine (SVM) and XGBoost (XGB) fell between other models because SVM produced 15 FPs and 8 FNs while XGBoost generated 2 FPs and 3 FNs.

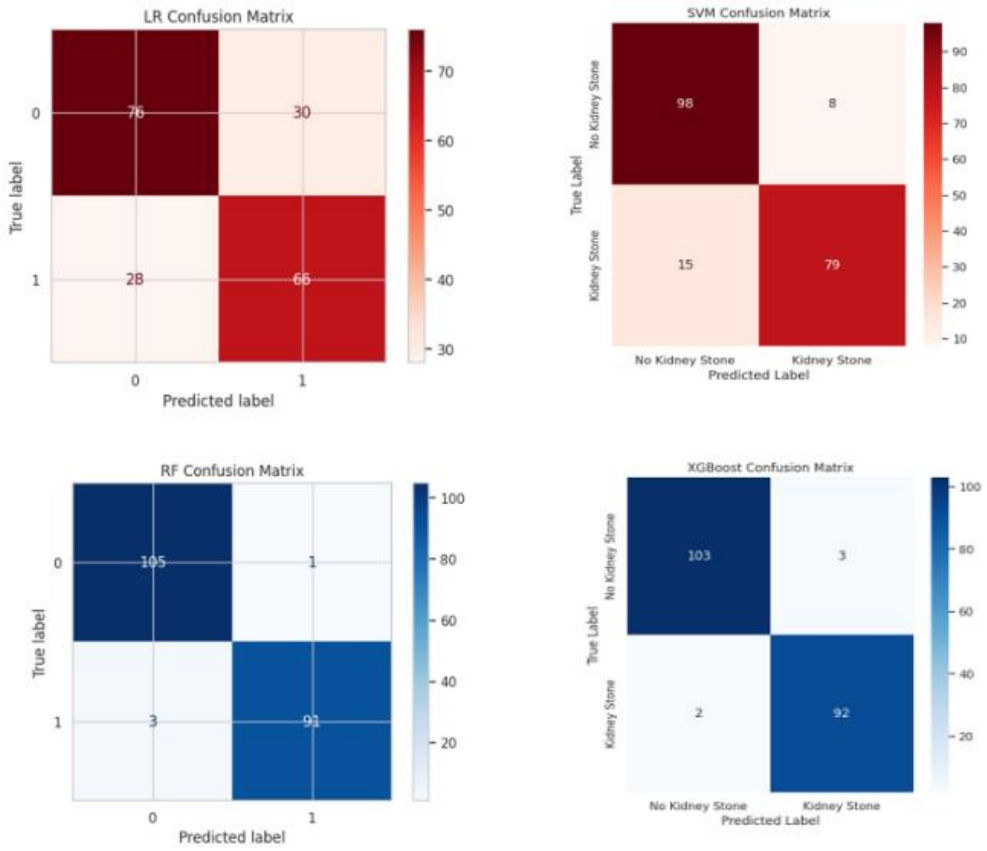


Figure 7.1 Confusion matrices of Models

Table I provides five important classification metric results including Accuracy along with Precision and Recall and F1-Score and AUC (Area Under the Curve). the formulas outlined in Equations (5) to (8) yield these numbers. The Random Forest model reached a performance score of 98% with Accuracy surpassing all other models including XGBoost at 97.50% and SVM at 88.50% as well as Logistic Regression at 71.00%. Random Forest achieved 0.96 Precision from its ability to identify and classify true positives from all positive predictions. The sensitivity of RF model exceeded 0.98 in its ability to detect positive cases during the Recall phase.

Table 1. Comparison of Models performance using valuation metrics

Model	Accuracy	Precision	Recall	F1 Score	AUC
Logistic Regression	0.7100	0.6875	0.7021	0.6947	0.783
SVM	0.8850	0.9080	0.8404	0.8729	0.826
Random Forest	0.9800	0.9891	0.9681	0.9785	0.986
XGBoost	0.9750	0.9684	0.9787	0.9735	0.844

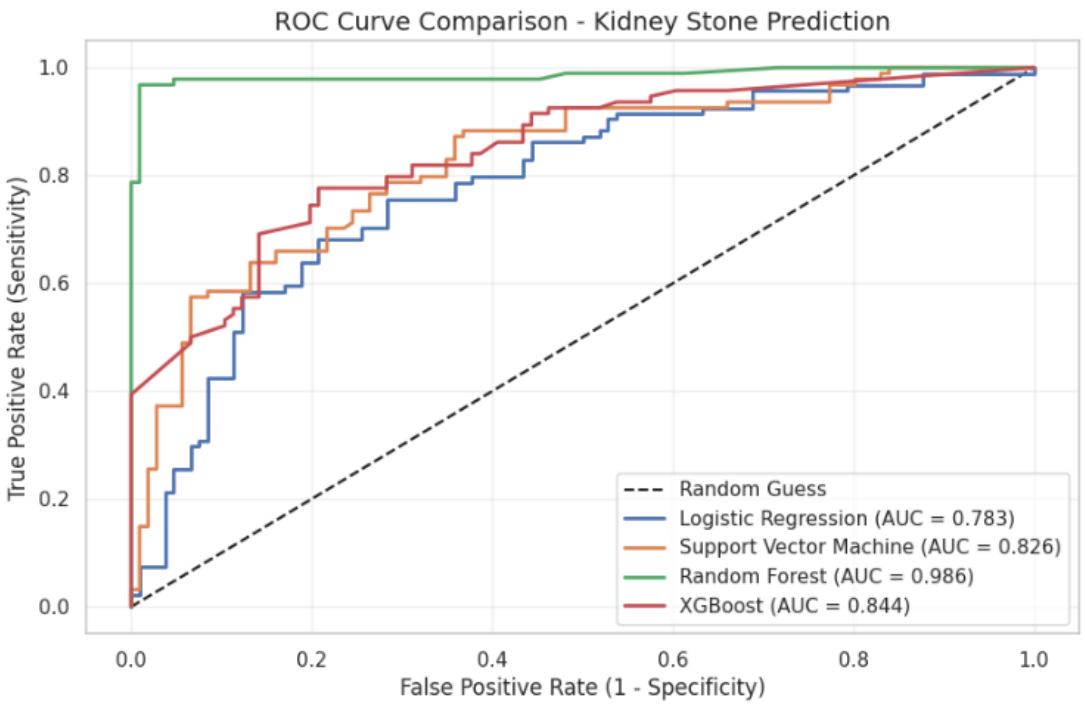
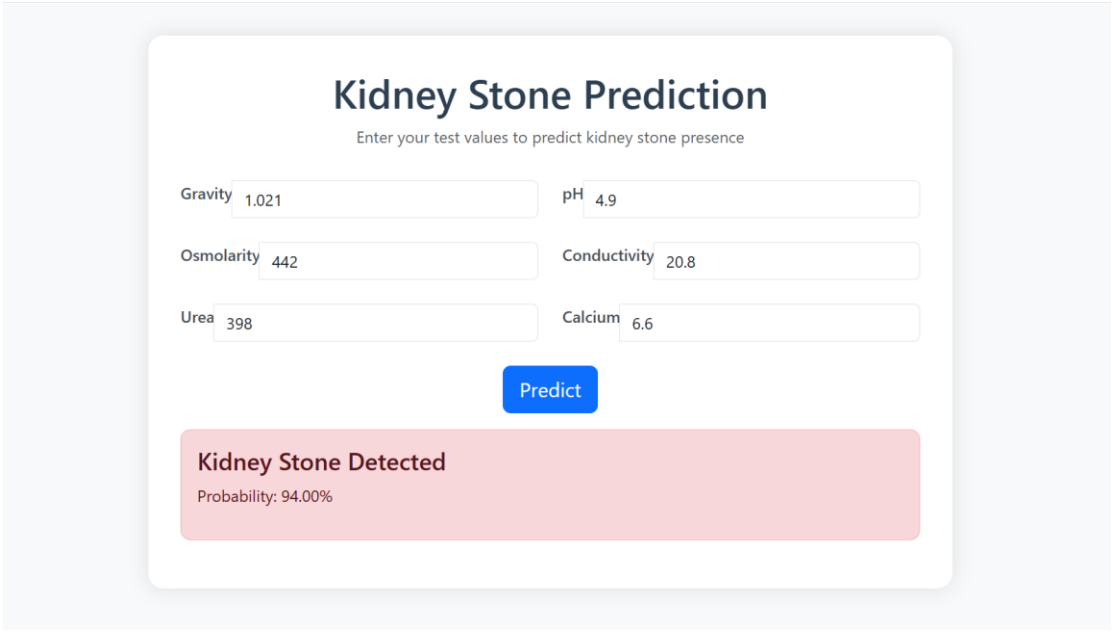


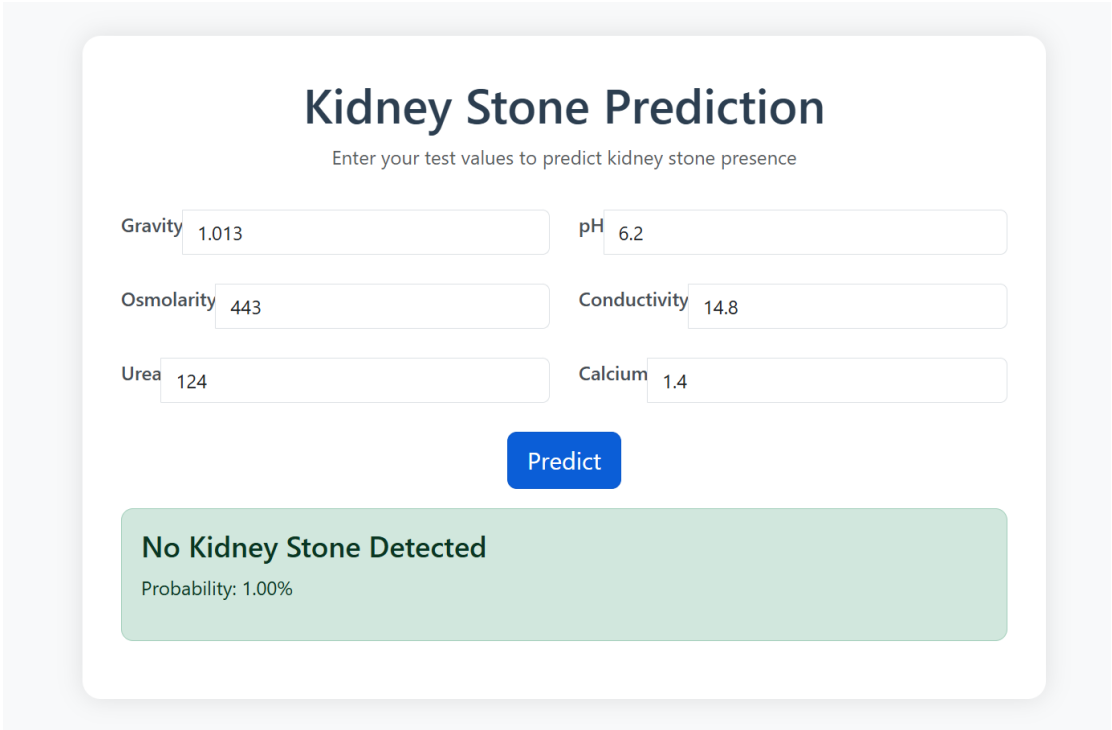
Figure 7.2 ROC Curve comparison of Models

The F1-Score achieved maximum value at 0.97 by RF which demonstrates that this model maintains optimal precision and recall values. The AUC value of XGBoost (0.844) compared to RF (0.986) does not translate into practical effectiveness measurement. The medical classification performance metrics demonstrate RF's strong suitability as the best solution for this task because it achieved superior results across all measures. SVM achieved 0.90 precision however its recall was at 0.84 which resulted in an F1-score of 0.87 while LR trailed behind with 0.68 precision and 0.70 recall and 0.69 F1-score.

When considering all evaluated metrics coupled with false negative reduction priorities the Random Forest approach excels as the most effective urine-based kidney stone detection solution.

Implementation





Conclusion

A research study performed assessments between Logistic Regression and Support Vector Machine while comparing them to Random Forest and XGBoost machine learning models which predicted kidney stone conditions through urine test results. The assessment of all models used confusion matrices together with standard classification metrics comprising Accuracy and Precision and Recall and F1-Score and AUC measure values.

The Random Forest model exhibited the highest achievement among these models with 98% accuracy together with a precision value of 0.989 and recall at 0.96 and F1-score at 0.978. A low level of misclassification errors according to the confusion matrix confirmed outstanding performance at clinical testing protocols. Random Forest provided superior diagnostic capability compared to XGBoost since it achieved higher performance rates and enhanced false negative detection points.

The method of choice for urine-based early detection of kidney stones turns out to be Random Forest based on its exceptional reliability and accuracy scores. The future work will be focused on merge deep learning methods with expanded data samples while establishing clear explanations of model functions and conducting laboratory-based medical tests to verify practical use of these models in healthcare institutions.

References

[1] S, S., Ds, A., S, P., & Pai, V. (2023). Deep Learning Based Kidney Stone Detection Using CT Scan Images. 1–7.

[2] Manoranjitham, R., Punitha, S., Ravi, V., Stephan, T., Mazroa, A. A., Singh, P., Diwakar, M., & Gupta, I. (2024). Automatic Kidney Stone Detection System using Guided Bilateral Feature Detector for CT Images. *The Open Public Health Journal*, *17*(1).

[3] Nisha, N., Shrieya, S., Shreyas, R., Kulkarni, K., & Munavalli, J. R. (2024). Kidney Stone Detection using CNN Algorithm. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, *10*(6), 1814–1823.

[4] Priyadharsini, R., Balaji, V., & S., S. (2024). Deep Learning-Based Renal Stone Detection: A Comprehensive Study and Performance Analysis. *Applied Computer Systems*, *29*(1), 112–116. https://doi.org/10.2478/acss-2024-0014.

[5] Osei, I., Baafi-Adomako, A., & Boadu, D. (2024). Detecting Kidney Stones Using Urine Test Analysis: A Machine Learning Perspective. *International Journal of Research and Scientific Innovation*, *XI*(X), 754–771.

[6] Shee, K., Liu, A. W., Chan, C., Yang, H., Sui, W., Desai, M., Chi, T., & Stoller, M. L. (2024). A Novel Machine-Learning Algorithm to Predict Stone Recurrence with 24-Hour Urine Data. *Journal of Endourology*, *38*(8), 809–816.

[7] Alghamdi, H. S., & Amoudi, G. (2024). Using Machine Learning for Non-Invasive Detection of Kidney Stones Based on Laboratory Test Results: A Case Study from a Saudi Arabian Hospital. *Diagnostics*, *14*(13), 1343.

[8] Kavoussi NL, Floyd C, Abraham A, Sui W, Bejan C, Capra JA, Hsi R. Machine learning models to predict 24 hour urinary abnormalities for kidney stone disease. Urology. 2022 Nov 1;169:52-7.

[9] Degadwala, S., & Rathva, V. (2024). A Review on Kidney Stone Detection using ML and DL Techniques. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, *10*(5), 101–112.

[10] Alqahtani, A., Alsubai, S., Binbusayyis, A., Sha, M., Gumaei, A., & Zhang, Y. (2023). Optimizing Kidney Stone Prediction through Urinary Analysis with Improved Binary Particle Swarm Optimization and eXtreme Gradient Boosting. *Mathematics*, *11*(7), 1717.