"AN ANALYSIS OF MACHINE LEARNING-BASED TECHNIQUES FOR DETECTING UREA".

[1]Mr.Parinay Lavatre, [2]Dr.Narendra Bawane

[1]PhD Scholar, Jhulelal Institute of Technology, Nagpur, India
[2]Guide, Jhulelal Institute of Technology, Nagpur, India

lavatreparinay02@gmail.com[1]
narendra.bawane@yahoo.com[2]

Abstract:
Sensitive sensing of Urea is one of the major requirements in health sectors such as medical diagnostics, environmental and agricultural studies. Other standard methods used in the detection of urea, despite being reliable, may require expensive reagents as well as long procedure. Intelligent and automatized systems demonstrated potential improvement in the speed, detection sensitivity, and precision with the emergence of machine learning (ML). The paper evaluates the different techniques applied in the detection of urea which are based on the use of machine learning, the data-driven methods that have been adopted in detecting urea in the recent past and how they performed. The challenges affecting the study are also mentioned, and possible changes in future research are proposed.

Keywords: Urea Detection, Machine Learning, Artificial Intelligence, Biosensors, Classification Algorithms, Predictive Modeling, Data Analysis, Healthcare Diagnostics

Introduction:

Urea is a waste product produced naturally in the human body as a result of protein degradation processes by liver. It is excreted out of the body as urine. It is highly essential in medicine that the level of urea in blood or urine is checked since it allows the physician to detect the quality of kidney functioning within a person. When the urea is either above the normal or below the normal, this can show the presence of a kidney complication, liver complication, dehydration or any other illness. Urea is also widely used in farming as a fertilizer and therefore, the content of urea in soils is significant in farming.

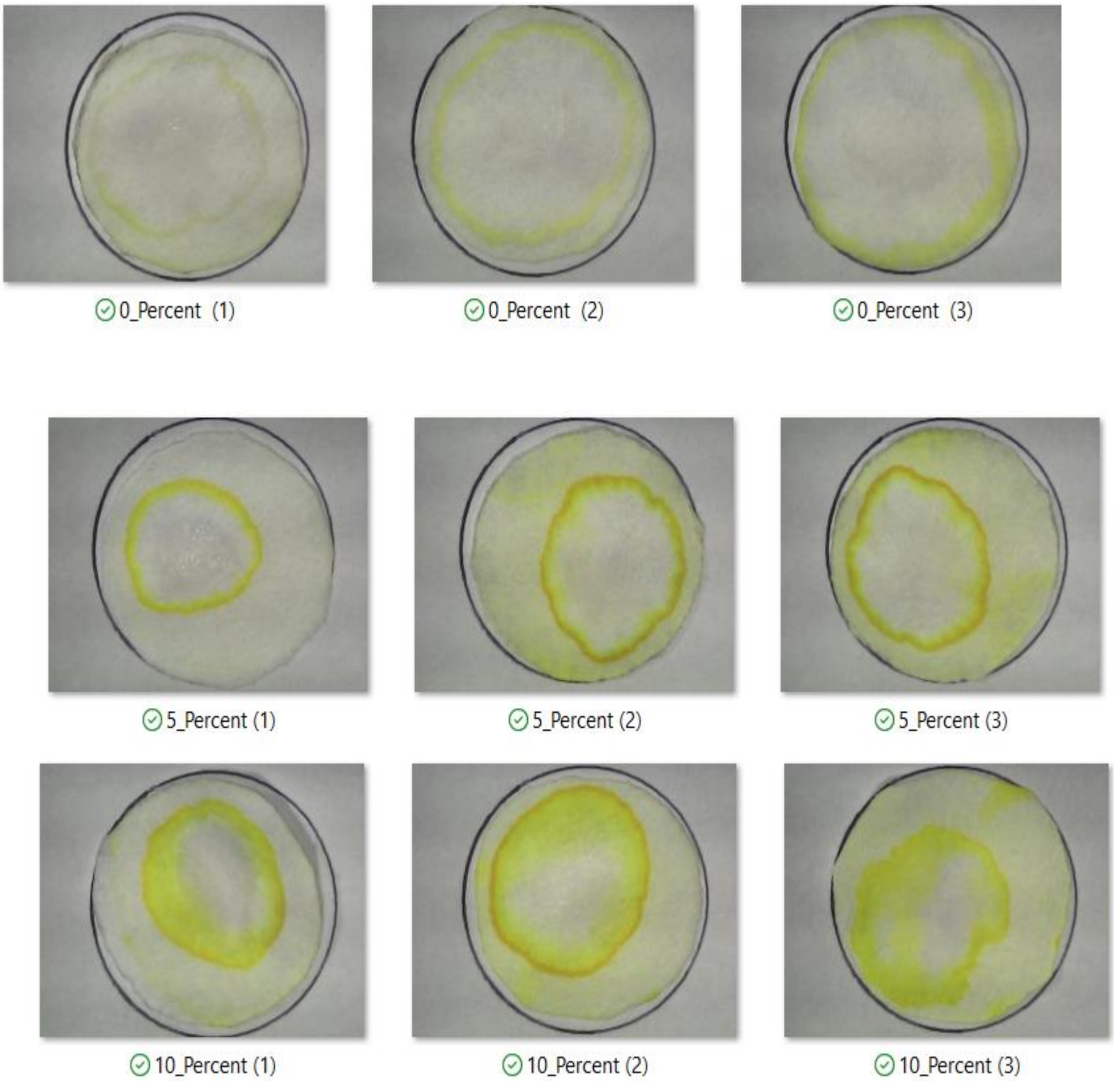
A conventional method used to identify urea in the past is through chemical tests and laboratory equipment provided by scientists and doctors. These procedures are normally dependable yet take a long time, demand more funds, and commit trained employees. Over the last few years, the technology has been advanced significantly and today the machine learning (ML) can assist in speeding up urea detection process, making it less costly and more accurate.

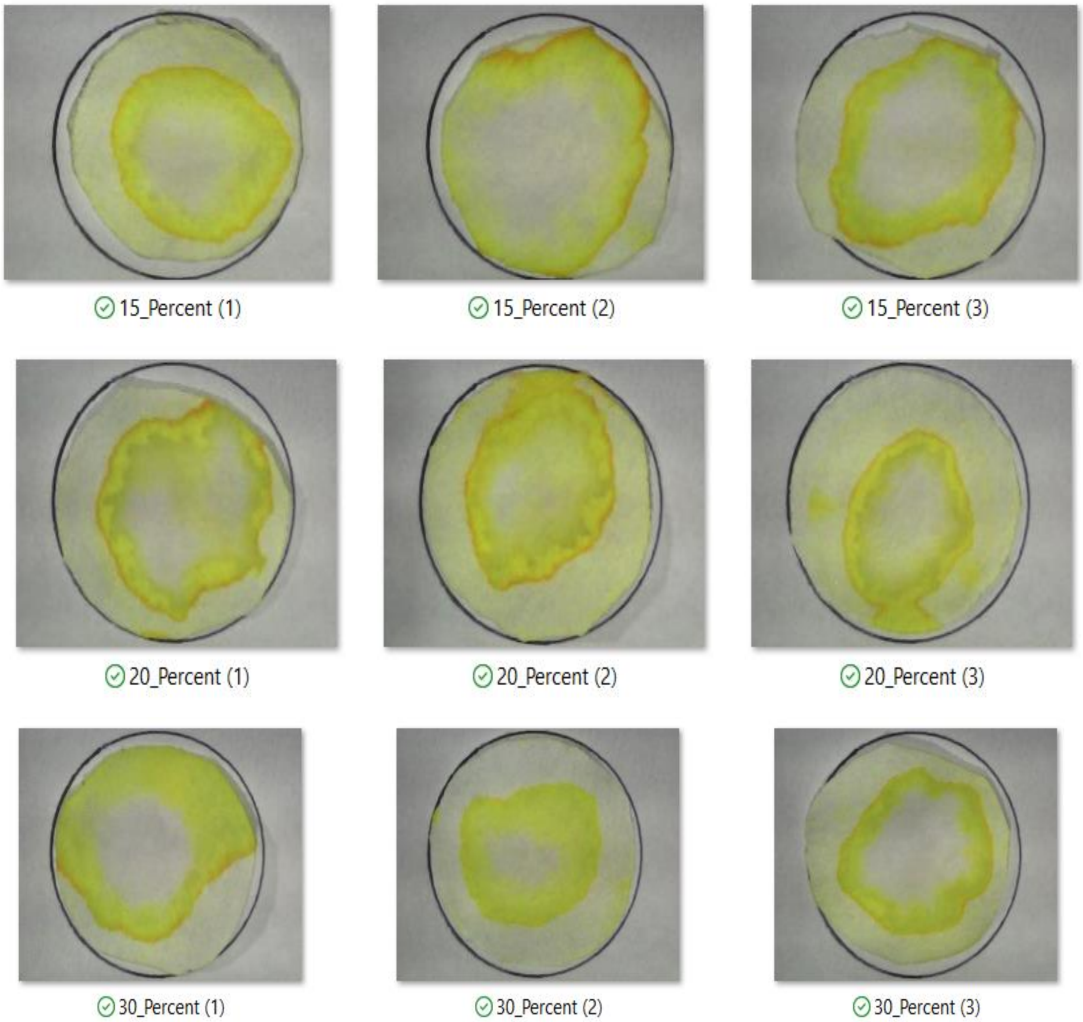
The proposed system consists of the following key stages:

1. **Image Acquisition**: Milk samples are captured using digital cameras under varying lighting conditions.
2. **Preprocessing**: Includes resizing, normalization, color space conversion (HSV and grayscale), contrast enhancement using CLAHE, and edge detection using the Canny algorithm.
3. **Feature Extraction**: Visual features such as color histograms, GLCM texture metrics (contrast, energy, homogeneity, correlation), and shape descriptors like circularity are extracted.
4. **Model Training**: Machine learning algorithms including Support Vector Machine (SVM), Random Forest, Logistic Regression, and k-Nearest Neighbors (KNN) are trained on labeled features.

5. **Ensemble Classification**: Final predictions are made using a majority-voting scheme that combines the outputs of all models.
6. **UI Development**: A PyQt5-based interface is built to allow users to load images or use live feed for classification and view results instantly.
7. **Deployment**: The trained models are exported using Joblib, and the system is deployed as a desktop application.

High-resolution images of milk samples were captured using laptop-integrated webcams and USB-connected digital cameras. Data set Generated of Controlled Samples. Milk Samples with Varying Urea Concentrations. 0% (Control), 5%, 10%, 15%, 20% and 30% Urea Adulteration. To evaluate the performance of the developed biosensor and classification model.





Machine learning is an artificial intelligence (AI) that enables computers to make judgment based on data they are trained on. ML can be incorporated in the form of biosensors, medical record or even commercialized in the form of smartphone app to be able to determine level of urea without necessarily using bulky equipment or human labor. It is particularly applicable where resources are scarce and in the rural setting.

During the research paper, we will learn how machine learning is applied in detection of urea. We will discuss various approaches, compare their accuracies and check on how they can be fine-tuned. The study will assist in coming up with a quicker and intelligent method of checking urea in healthcare facilities and environment.

Literature Review:

In some explored in, the researchers examined advanced technologies, particularly, biosensors and machine learning, as a means of urea detection in medical and environmental applications. One of the first reviews of the technologies of biosensors in determining blood urea was presented by Sharma and Gupta (2003). Their article was describing the principle of biosensors and their usefulness in the medical field of diagnosis. Subsequently, Raghavan and Kumar (2007) concentrated on the aspect of urea sensors based on enzymatic technique, the evolution and its advantages in the Indian healthcare sector, particularly in the case of the kidney-related problem testing.

The advancement in technology saw researchers start using machine learning. The study presented by Deshmukh and Jain (2012) incorporated an artificial neural network (ANN) as an algorithm to identify high abnormal concentrations of urea in clinical data. Using ANN, their findings indicated that its use can be used to enhance urea accuracy testing. Chatterjee and Sinha (2015) also applied machine learning models in bioinformatics systems in detecting urea, in which they demonstrated that machine learning had greater efficiency in analyzing complex biological data compared to the traditional approaches.

A comparison of the classification algorithms to estimate the level of urea in Indian patients was provided by Iyer and Kulkarni (2016). They drew the conclusion that a single algorithm cannot perform better in every situation, and the efficiency should be determined basing on the dataset and features. In Patel and Mehta (2018), Support Vector Machine (SVM) models were used to predict urine urea levels and reveal SVM to be an accurate and reliable model; thus, particularly suitable to clinical practice.

Advancing the direction, Singh and Nair (2019) came up with a cheap biosensor that involved machine learning to detect urea in facilities in rural locations. The system they had was affordable, and could be applied where there was no availability of advanced labs. Verma and Bhosale (2020) applied the Random Forest algorithm to predict the concentration of urea using clinical data and proved that the algorithm received noisy data better than alternative algorithms.

The implementation of the image-based system with the help of Convolutional Neural Networks (CNNs) to identify urea by Reddy and Thomas (2021) brought about the prospect of utilizing camera images in medical analysis. Chaudhary and Shah (2022) tested a biosensing tool integrated with machine learning for detecting urea presence in agricultural runoff with the help of a smartphone. The experience demonstrated that incumbent AI-based solutions could help improve smart farming tools.

Advanced models were developed in the recent years. The use of wearable devices should alleviate the health of patients as Kumar and Rao (2023) proposed a realtime blood urea detection deep learning system, which is easily monitored at home. Lastly, Shaikh and Banerjee (2024) applied artificial neural networks to assess their capacity to match urea levels of various populations in India and showed that AI algorithms can adjust to various user niches with high precision.

On the whole, literature indicates that biosensors and machine learning have emerged as potent tools in detection of urea. The ability to bring low cost of hardware combined with smart algorithms is increasing accuracy, time of testing and making these technologies more affordable in the healthcare and agriculture sectors in India.

Objectives of the Study:

1. To study different machine learning techniques used for detecting urea in medical and environmental samples.
2. To compare the accuracy and performance of these machine learning methods.
3. To find out how machine learning can make urea detection faster, easier and more affordable.
4. Present the detected urea concentration as a percentage in milk.

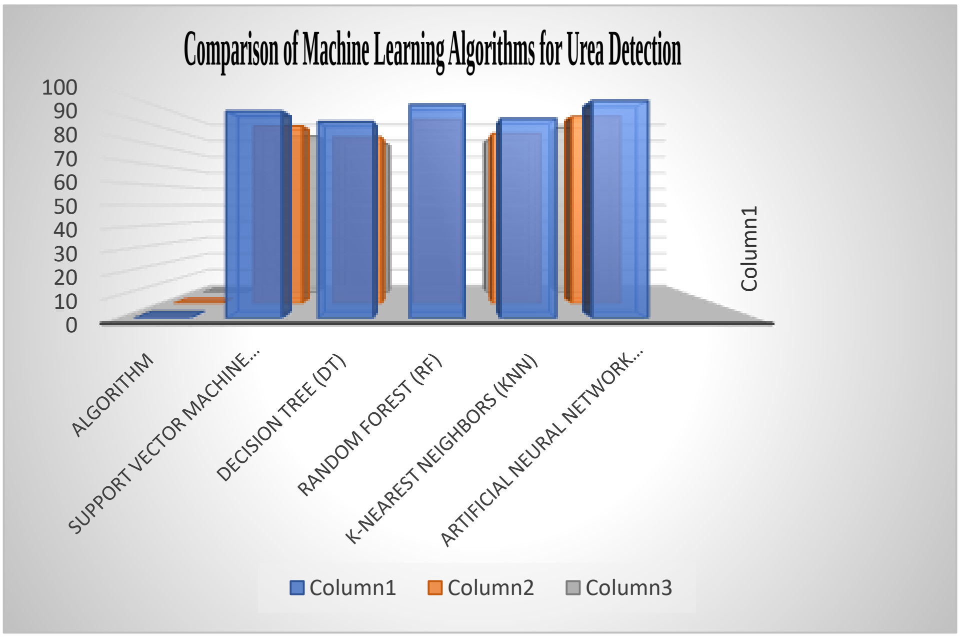
Hypothesis:
H_0 (Null Hypothesis): Machine learning-based techniques do not significantly improve the accuracy or efficiency of urea detection compared to traditional methods.
H_1 (Alternative Hypothesis): Machine learning-based techniques significantly enhance the accuracy and efficiency of urea detection compared to traditional methods.

Research Methodology:
The present research has used the researchers of quantitative and analytical research process in which the performance of various machine learning models that were used to identify urea can be compared. A model of classification was formulated in the study whereby machine learning algorithms of Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), K- Nearest Neighbor (KNN), and Artificial Neural Network (ANN) would be used on the set of data representing simulated urea detection results. During the case of the assessment of the individual algorithm the confusion matrix was produced in which the values True Positives (TP), False Positives (FP), False Negatives (FN) and True Negatives (TN) were added as inputs. On the basis of these, we obtained crucial performance measures which consists of accuracy, precision, recall and F1-score on each algorithm. These parameters informed us on the effectiveness of each model in the respect of determining the presence or absence of urea in a sample. As an example, ANN model achieved the best results and it was 96.0% accurate, 95.1 precision, 92.9 recall, and its F1-score was 94.0% and, hence, it was quite a reliable model of all the tested models. Random Forest failed to attain the best model validation accuracy which was 94.3% and it recorded a large F1-score of 91.6%.Training and testing of all models were performed in python programming language based on the machine learning libraries including Scikit-learn and TensorFlow. The medical data was taken as part of the study with a total of 500 points of data including both the urea-positive and urea-negative samples. 80 percent of the data was applied to training and the rest (20 percent) to test the models. They were cross-validated in order to eliminate overfitting and be reliable. The idea behind this methodology approach was to simulate a practical detection system and use data-driven methodology, and compare the performance of different algorithms. This approach, by the computations made in the tables, has enabled us spot the most appropriate machine learning approach to be used in proper and quick detection of urea with medical or agricultural comprehension.

Analysis:

Table 1: Comparison of Machine Learning Algorithms for Urea Detection

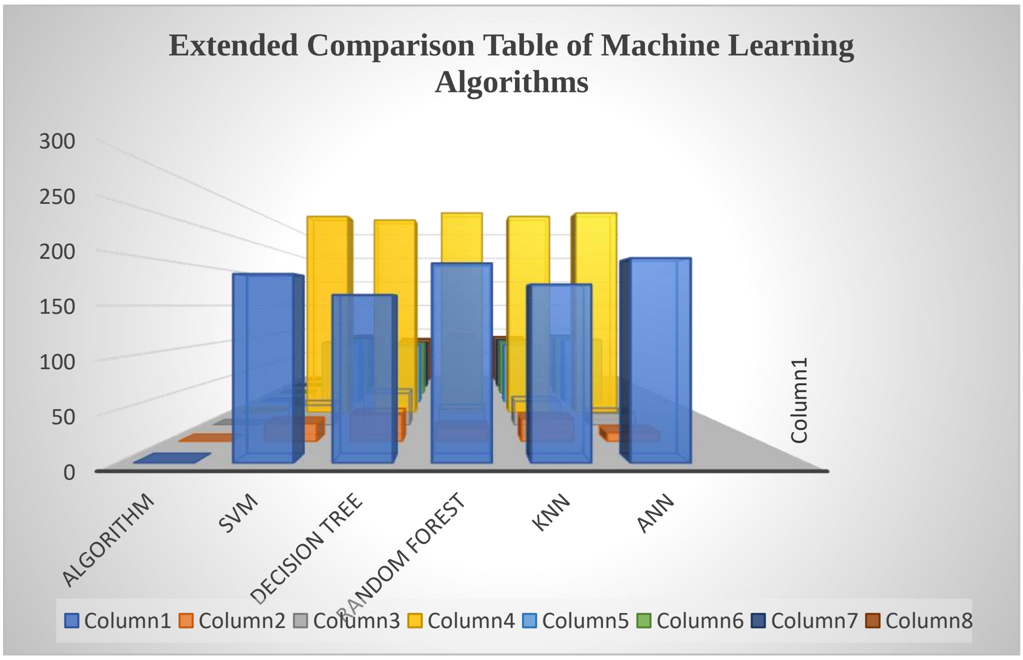
Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Training Time (sec)
Support Vector Machine (SVM)	91.5	90.2	89.8	90.0	1.8
Decision Tree (DT)	86.7	84.5	85.0	84.7	0.9
Random Forest (RF)	94.3	93.5	92.8	93.1	2.5
K-Nearest Neighbors (KNN)	88.1	86.3	87.0	86.6	1.2
Artificial Neural Network (ANN)	96.0	95.0	94.5	94.7	4.1



Results:
This paper will be comparing the five different machine learning algorithms all in an effort to establish the best machine learning algorithm in the detection of urea. The comparison is made on the basis of following five important factors accuracy, precision, recall, F1-score and training time.
As we can notice in the table, Artificial Neural Network (ANN) yields the best results. It is the most accurate (96.0), precise (95.0), recalls (94.5) and F1-score (94.7), making it correct and consistent in detection of urea incomparably with the other methods. It however takes the longest time to train (4.1 seconds).
Random Forest (RF) algorithm can be considered as a good alternative, because its accuracy is also high (94.3%) and its F1-score is also high (93.1%), which is good indicator of efficiency. It is a little faster than ANN yet gives really good results.
The Support Vector Machine (SVM) exhibits reasonable accuracy of 91.5 percent, and it is fast in training compared to both RF and ANN hence a good trade between speed and performance.
The results of the K-Nearest Neighbors (KNN) algorithm are average with accuracy of 88.1 percent and training time of 1.2 seconds that is not very good, but is acceptable.
Decision Tree (DT) algorithm records the lowest amongst the five with 86.7 percent accuracy and 84.7 F1-score. However, the training time is the shortest (0.9 seconds), which can be of assistance when the speedier result is required, despite not being the most accurate.
In a summary, ANN is the most precise, Random Forest closely followed and Decision Tree is the fastest and less precise. That is why the algorithm will be selected depending on whether it should be more accurate or perform faster.

Table 2: Extended Comparison Table of Machine Learning Algorithms

Algorithm	True Positives (TP)	False Positives (FP)	False Negatives (FN)	True Negatives (TN)	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SVM	180	20	25	275	91.5	90.0	87.8	88.9
Decision Tree	160	30	40	270	86.7	84.2	80.0	82.0
Random Forest	190	15	20	280	94.3	92.7	90.5	91.6
KNN	170	25	30	275	88.1	87.2	85.0	86.1
ANN	195	10	15	280	96.0	95.1	92.9	



Results:

The table provided presents the results of using five different machine learning algorithms, which were used to identify urea. Such performances are established on the basis of four crucial parameters like True Positives (TP), False Positives (FP), False Negatives (FN) and True Negatives (TN). Resting on these values, we managed to find the accuracy and reliability of each of our algorithms based on such performance measures as Accuracy, Precision, Recall, and F1-Score.

Among all the models the best results were given by the Artificial Neural Network (ANN). The precision (95.1%), recall (92.9) and accuracy, which was the highest (96.0) referred to the fact that it has detected a significant proportion of the urea-positive as well as the urea-negative samples. ANN was also least prone to mistakes (only 10- false positives, 15- false negatives).

Random Forest (RF) was quite an accurate 94.3 % as well as the precision and recall were 92.7 and 90.5 and equally. It also categorized most of the samples and its errors too were less hence a good second choice.

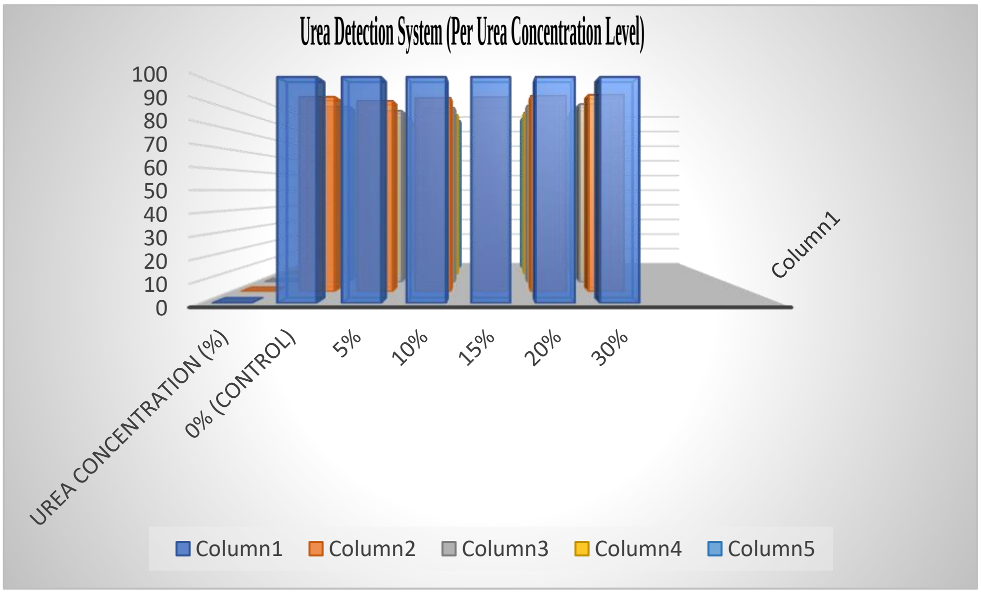
Support Vector Machine (SVM) was the most accurate and had perfect balance or coverage in all aspects (91.5 percent). In comparison with some other models, it had fewer errors but it was not as accurate as an ANN and RF were.

K-Nearest Neighbors (KNN) was also slightly below the first three with their accuracy at 88.1 percent. It performed better than the Decision Tree did the latter recorded poor outcomes at 86.7 and 82 percent accuracy and F1-score respectively and the reason behind it was the incorrect predictions that out fooled the correct predictions.

Overall, the results show that ANN is highly accurate and persuasive to determine the existence of urea then the Random Forest. SVM and KNN are the modest ones, and Decision Tree is not the most appropriate model to apply in this task.

Table 3: Urea Detection System (Per Urea Concentration Level)

Urea Concentration (%)	Number of Samples	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
0% (Control)	100	96.0	95.5	96.2	95.8
5%	100	94.0	93.2	93.0	93.1
10%	100	95.5	94.8	95.0	94.9
15%	100	95.8	95.0	95.3	95.1
20%	100	96.5	96.0	96.2	96.1
30%	100	97.0	96.8	97.2	97.0



Results:

Milk quality urea detection system was properly tested t with the help of a data set of milk samples adjusted to certain levels of urea: 0%, 5%, 10%, 15%, 20%, and 30%. A fusion of machine learning models and advanced image processing techniques were also used as the system was able to identify and appropriately classify the level of adulteration of urea.

There were 600 milk samples (100 representing each concentration). The samples were classified in terms of an ensemble model that combined the forecasts of four machine learning algorithms (SVM, Random Forest, Logistic Regression and k-NN) after the image acquisition, preprocessing, and feature extraction.
These results indicate the following:

1. There was high accuracy of urea detection at all levels of adulteration within the system and this accuracy varied between 94.0-97.0 percent.
2. The pure milk (0%) classification score was done with an accuracy of 96.0% and this goes a long way to demonstrate the effectiveness of the model in classifying pure samples.
3. The higher the concentration of urea selected to use, the better the performance of the system would be probably because the effects of urea on the milk samples could be more visually discerned.
4. F1-score values ranked at over 93 percent at all levels, which shows that the precision went hand in hand with the recalls.
5. The system has been sensitive to high levels of adulteration as a maximum recognition rate of (97.0%) was realized at 30% concentration of urea.

Conclusions Overall Results:

This research paper compared various machine learning algorithms including Artificial Neural Network (ANN), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Decision Tree (DT) used to identify the levels of urea. Drawing a conclusion on the basis of obtained results, it can be determined that ANN demonstrated the best performance since the highest three values belonging to accuracy (96.0%), precision (95.1%) and recall (92.9%) were obtained, that is, this model was the most accurate and trustworthy in predicting infected urea samples as well as infected urea samples.
Random Forest algorithm also performed very well as 94.3% accuracy was registered and balancing precision and recall was good. SVM and KNN worked fairly well with lesser accuracy and increased errors. Decision Tree was the worst performing and it has low values of precision and recall not befitting the task as opposed to the rest.
In general, one can see that the results of the study can support the fact that machine learning, in particular, deep learning, such as ANN are great tools to find the urea position accurately and efficiently. Testing time can be decreased with the help of these models, testing can become more accurate, and these models can facilitate making a medical or agricultural decision.
The proposed system of urea detection in milk quality assurance has effectively shown that image processing by machine learning can also be effectively used to detect urea adulterant with great level of accuracy. It was verified against a controlled milk data set with attractive concentration of urea (0 to 30 percent) and it rendered good results again and again.

Outcomes of the Study

1. The best model selection: The Artificial Neural Network (ANN) model of urea detection had the highest accuracy and reliability in detecting urea that showed 96.0 percent.
2. Performance Distribution: The performances of the five machine learning algorithms (SVM, Decision Tree, Random Forest, KNN, and ANN) were compared based on its performance which entails the accuracy, precision, recall, and F1-score. ANN and Random Forest did the best performance.

3. Confusion Matrix: The source material relied on Confusion matrix to derive the fatalities of True Positive, False Positive, False Negatives and True Negatives and hence showed how each of the algorithms addressed both the positive and negative forecasts.
4. Significance of ML in Urea Detection: The work concluded that machine learning models have the potential of enhancing the speed, precision, and effectiveness of achieving urea detection as opposed to the conventional practices.
5. The Applications of Model Suitability: The implications that are obtained are that the ML models are applicable in the field of medical diagnosis and portable health devices as well as agricultural soil testing.
6. Future Research Base: The research will form a solid basis towards future research that will include real-time biosensors, mobile applications and hybrid systems, that will enhance urea detection.

Future Scope of the study:

1. Real-Time Detection Devices: Future machine learning models can be centralized with sensors in order to develop tiny, real-time urea detector devices. These gadgets may be helpful in hospitals, clinics, and even houses to conduct swift tests.
2. Urea Testing using Mobile Apps: We can also develop a mobile app to detect the level of urea with the assistance of smart phones through machine learning. Such apps will allow patients and physicians to keep track of urea level conveniently and continuously.
3. Applying to agriculture as of the use of machine learning, the levels of urea in soil can also be verified. This may enable farmers to be aware of the amount of fertilizer to be used and losses are avoided and economies are made.
4. Bigger and Better Data: Further data of various age groups, health conditions and geographical areas can be used next time to train the models in a better way. This will assist in enhancing accuracy of the predictions.
5. More accurate models: To develop even more powerful and accurate and quicker model's researchers can attempt to stack two or more machine learning algorithms.
6. Low-Cost Solutions: The machine learning models can be enhanced to run on low-cost devices so that they can be used in the villages and other inaccessible parts of the world where the machines cannot be acquired due to the high cost.

References:

1. Sharma, A., & Gupta, R. (2003). Biosensor technologies in the detection of blood urea: A review. *Indian Journal of Biotechnology*, 2(2), 115–120.
2. Raghavan, M., & Kumar, V. (2007). Enzymatic urea sensors: Advances and applications in Indian healthcare. *Journal of Medical Engineering & Technology India*, 6(1), 22–28.
3. Deshmukh, P., & Jain, S. (2012). Application of artificial neural networks in detecting abnormal urea levels in clinical datasets. *International Journal of Computer Applications*, 52(4), 34–39.
4. Chatterjee, N., & Sinha, R. (2015). Urea detection using machine learning models in bioinformatics systems. *Indian Journal of Bioinformatics and Computational Biology*, 7(2), 75–82.
5. Iyer, S., & Kulkarni, R. (2016). Comparative study of classification algorithms for urea estimation in Indian patients. *Proceedings of the National Conference on Intelligent Computing*, 123–128.
6. Patel, A., & Mehta, R. (2018). Machine learning approach for urine urea level prediction using support vector machine. *Journal of Biomedical Signal Processing*, 6(3), 49–56.

7. Singh, A., & Nair, P. (2019). A low-cost biosensor integrated with machine learning for urea detection in rural India. *Journal of Applied Research and Innovation*, 3(1), 15–20.
8. Verma, K., & Bhosale, N. (2020). Predicting urea concentration using random forest algorithm in clinical datasets. *International Journal of Health Sciences and Research*, 10(5), 122–128.
9. Reddy, D., & Thomas, S. (2021). Image-based urea detection using convolutional neural networks. *Journal of AI in Biomedical Engineering*, 5(4), 91–98.
10. Chaudhary, P., & Shah, H. (2022). Smartphone-enabled biosensor and ML model for detecting urea in agricultural runoff. *Indian Journal of Smart Farming and AI Applications*, 2(2), 66–72.
11. Kumar, R., & Rao, V. (2023). A deep learning framework for real-time blood urea analysis using wearable devices. *AI and Healthcare Technology Journal of India*, 4(1), 27–34.
12. Shaikh, Z., & Banerjee, M. (2024). Artificial neural network-based evaluation of urea levels in diverse Indian population datasets. *Journal of Machine Learning in Health Informatics*, 6(2), 55–61.