HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 4 (2024)

Using Machine Learning to Optimize Kidney Transplants

Shreyas Das

JW. Mitchell High School, USA

ABSTRACT

Currently, 100,000 people are waiting on the organ donation list in the United States, 17 of which die each day. This is due to the demand for organs far outweighing their supply. There needs to be objective criteria regardless of race, economic status, or sex for the distribution of organs. The most common types of transplants are kidney transplants. For this reason, the allocation of kidneys must be given to the patients who require the kidneys the most. The operational definition for needing kidneys the most should be the patients who are most likely to progress to chronic kidney disease (CKD). Artificial intelligence can help predict which patients will progress to ESRD and has shown promise in doing so. However, many different types of AI models can be used, and many of them have stark differences in how they operate. Comparing these models can allow researchers to understand which models are most effective for diagnosing CKD. The models featured in this study were the logistic regression, ridge classifier, and decision tree models. All three models had a mean accuracy of 0.975. The logistic regression model had a mean precision of 0.960, a mean recall of 1.00, and a mean F1 score of .980. The ridge classifier model and the random forest classifier model both had a mean precision of 1.00, a mean recall of 0.958, and a mean F1 score of 0.979.

Introduction

As of February 1st, 2024, 103,323 people are waiting on the organ transplant waiting list in the United States. This is because the current demand for organs far outweighs the supply. Unfortunately, this causes 17 people to die every day waiting for an organ transplant (Health Resources and Services Administration [HRSA] , 2024). Among the organs that can be transplanted, kidneys are the most transplanted organs (European Commission, n.d.). As of March 1st ,2024, out of 106,441 patients in the United States who need an organ transplant, 89,101 need a kidney transplant (HRSA, 2024). In other words, approximately 83.7% of the patients on the organ donation waiting list require a kidney transplant. The main cause of needing a kidney transplant is End Stage Renal Disease (ESRD). ESRD is the fifth and final stage of chronic kidney disease (CKD) - the gradual decline of the kidney (MayoClinic, September 2023). In this stage, the kidneys cannot meet the filtration requirements of the body, which causes dangerous waste to accumulate. Due to the severity of ESRD, the only way to treat it is through dialysis and kidney transplants (“MayoClinic, October 2023) .

Due to the previously mentioned demand for organs far outweighing the supply, organs need to be allocated to the patients that need them the most. However, in many cases, organs are not allocated based on whomever needs them the most. A study by Wesselmann and colleagues (2021) found that African American patients were less likely to receive organ transplants compared to white patients (It is plausible that in many cases patients who desperately need organs may not receive the organs. To prevent this, factors such as race, gender, and economic status should not be considered in the allocation of organs (Buinnik, 2023). In terms of kidneys, the patients who need kidney transplants the most will be the patients who have CKD, as those patients are the most likely to progress to ESRD. The objective of this study is to use machine learning to predict whether a patient would progress to CKD; to compare different AI models and see which one is the best at predicting progression to CKD, and to determine which factors are most important in determining progression to CKD.

Background and Literature Review

In a research study published in BMC Nephrology, an open access journal for publishing peer reviewed kidney research, a study by Segal and colleagues used Artificial Intelligence to predict progression to ESRD. The researchers used a decision tree algorithm (Gradient Boosting Trees (XG Boost Implementation)) and the word2vec algorithm. The decision tree algorithm analyzed a database of 10 million insurance claims from 550,000 patient records. The AI had a sensitivity of 0.715 and a specificity of 0.958. In other words, the model had 71.5% accuracy when predicting that a patient would progress to ESRD and 95.8% accuracy when predicting a patient would not progress to ESRD. Furthermore, the study highlighted that age, highest CKD stage diagnosed during the initial eligibility, annual count of hypertensive crisis diagnosis, and presence of new hypertension as important factors when diagnosing ESRD (Segal et al, 2020). This study highlights important aspects of applying machine learning to medical situations, including vectorization and a potential model that can be used in diagnostic situations. The word2vec algorithm mentioned in the study is used to turn words into vector representations of the words. This is necessary because computers and machine learning algorithms cannot understand the context behind words just with their string form. The words need to be translated into numerical values known as vectors. Then the vector in question is understood by the algorithm, by using other similar vectors. This allows the algorithm to understand the word's meaning within the problem's context.

A decision tree algorithm uses various features to data features to make decisions. Based on the different data features, decision tree algorithms make different decisions and go down different data paths, causing different predictions ("1.10 Decision…", n.d.). However, the main issue is that claims data only has medical data that can be billed, and therefore doesn't have the patient's entire medical data. Therefore, since the model relied on billable diagnosis codes and not pure medical data, the diagnosis can't be properly verified. Furthermore, the reason for progression to ESRD may not be listed in the insurance claims which could cause false negatives.

Additionally, a study by Lee and Colleagues in 2022 utilized Artificial Intelligence to predict whether sepsis survivors were at risk for ESRD. Lee and Colleagues used five different AI models in their study, those being "namely logistic regression, random forest, extra trees, extreme gradient boosting (XGBoost), light gradient boosting machine (LGBM), and gradient boosting decision tree (GBDT)… (Lee et.al ,2022, para 8)." The data for this study came from Taipei Veterans General Hospital and comprised 142,624 sepsis survivors aged 20 or older. 11,661 sepsis survivors were in the final cohort for the study. The GBDT model had the highest area under the curve (AUC) of 0.879. The second highest was the LGBM model with 0.868. The extra trees and the random forest algorithm had AUCs of 0.865 and 0.864. The GBDT model performed the best with higher accuracy, F1 score, precision, and recall when compared to the other models (Lee et.al, 2022). This study has key differences when compared to the study done by Segal and colleagues. Namely, this study used many more AI models and was nuanced in the sense of mainly looking at Sepsis patients, while Segal and colleagues focused on ESRD. The numerous AI models used and the fact that this study noted which factors were important to the AI models when making their decisions allowed for greater interpretability for the AI.

Data Set Selection

The data set used in this research study was found on the UC Irvine website, as it was donated to the university. This data comes from Bangladeshi patient data from Enam Medical College (*Islam and Aktar, 2020*). The dataset comprises 200 patients, with 127 having CKD and 73 not having it. This distribution is shown by the graph below with class referring to whether a patient has CKD or not.

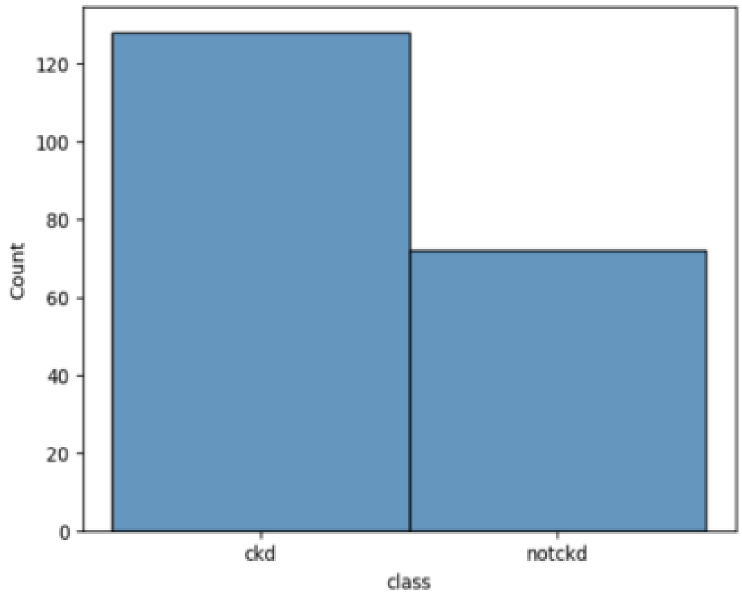

Figure 1. CKD status of patients

Furthermore, the dataset initially had both integer and string data and required preprocessing. The dataset provides 29 risk factors associated with CKD as variables. The numerous variables given by the dataset can be used to observe which variables are more associated with chronic kidney disease. This would help predict whether a patient could progress to kidney failure if they have a certain factor strongly associated with CKD. Notable variables include age and stage, both of which were shown to be important factors when predicting if a patient would progress to ESRD, along with Urine Specific Gravity, hemoglobin, and many more.

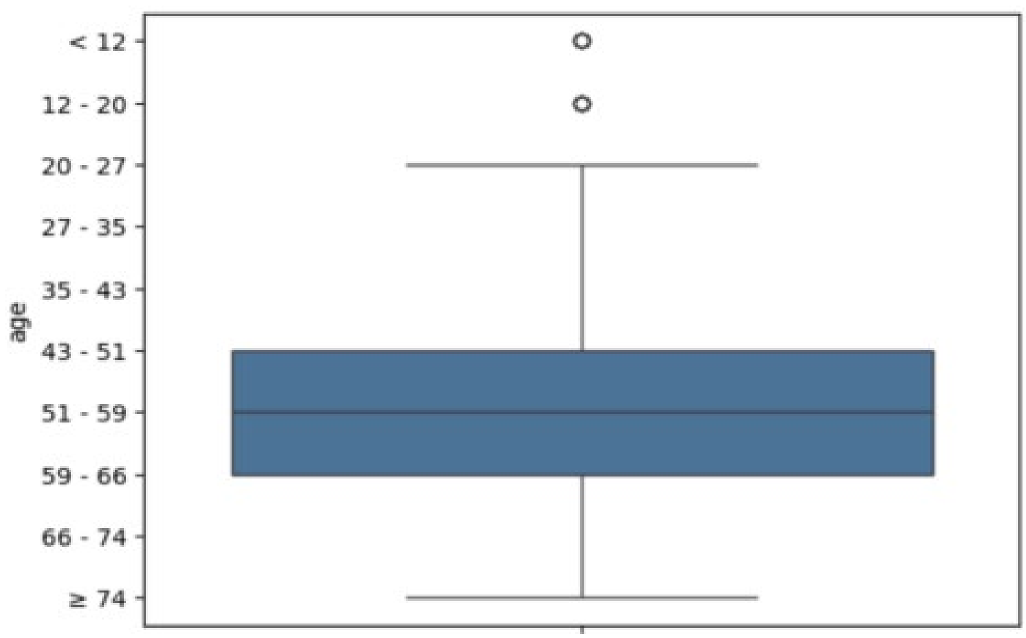

Figure 2. Boxplot of the Distribution of Ages in the Dataset

The distribution of ages in the dataset is skewed to the left, with most patients being older. The median change is between 51-59, and the maximum age is >74. There are two outliers on the lower end of the distribution.

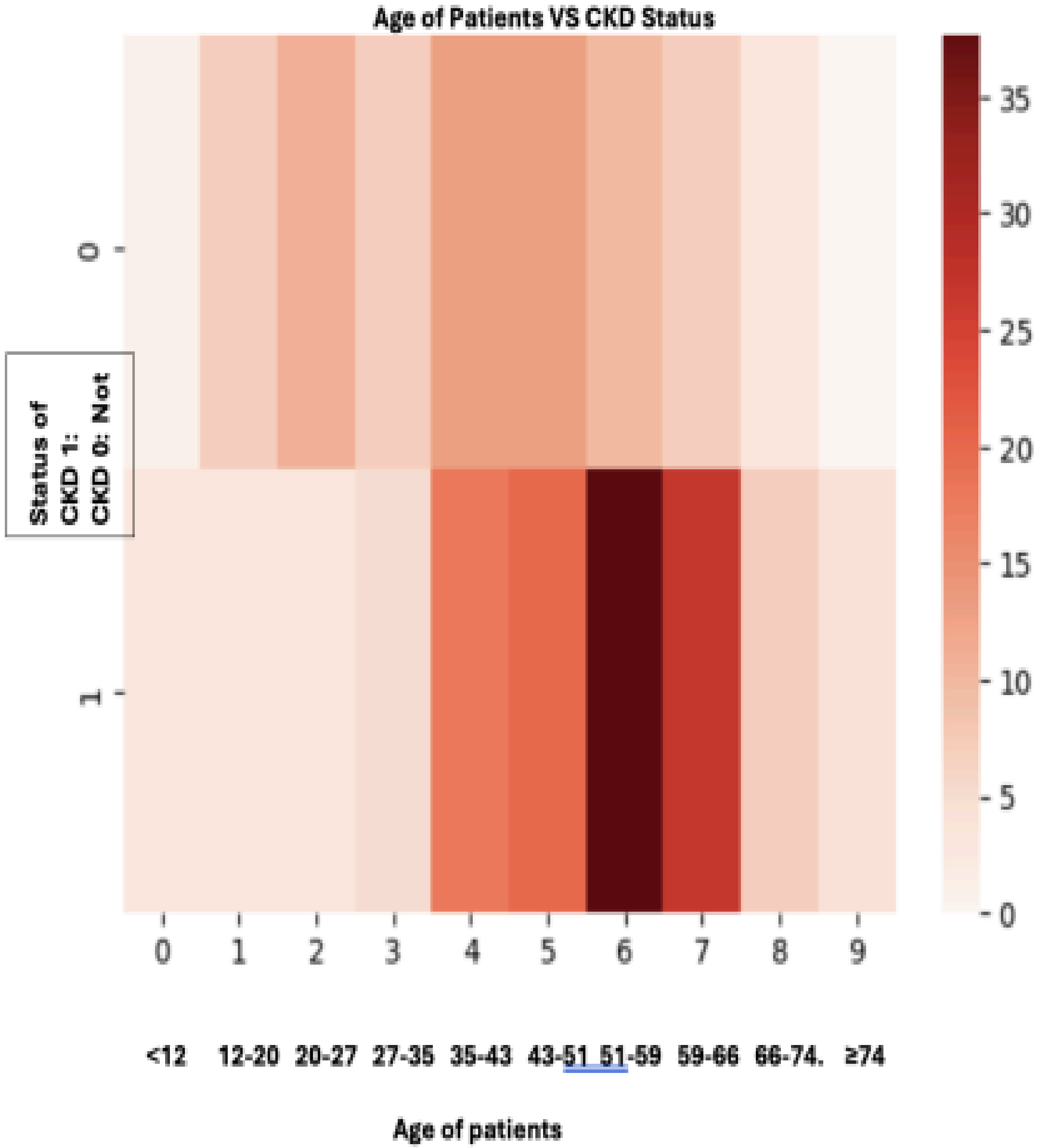

Figure 3. Heat Map of Age of Patients vs CKD status (0= Patient does not have CKD; 1= Patient has CKD)

Like the research study done by Segal and colleagues, there appears to be a correlation between age and having CKD, as the age increases, the number of patients that have CKD also increases. In the heat map below on the y axis, 1 represents having CKD and 0, represents not having it. On the y-axis, 0 through 9 represents the various age ranges, where 0 is less than 12, and 9 is greater than or equal to 72. From 0-6 (less than 12 to 51-59) there is an increase in the number of patients that have CKD per age group, indicated by the darker shading of red. While the shading of red gets lighter from 7-9 (59-66 to greater than or equal to 74), this is due to less patients in the study being in those age ranges.

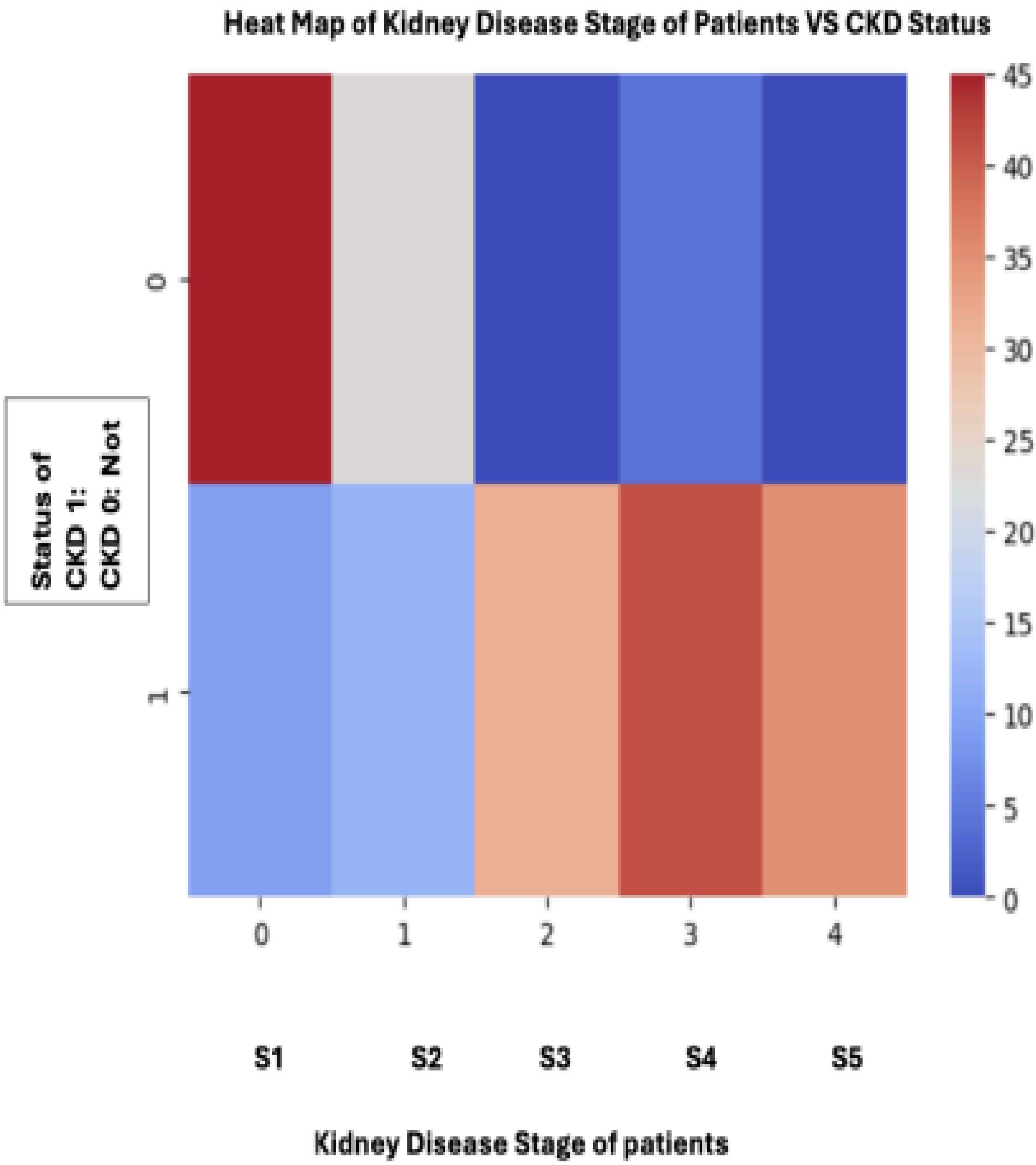

Figure 4. Heat Map of Stage of Kidney Disease for Patients vs CKD status. (0= Patient does not have CKD; 1= Patient has CKD)

There is a similar trend when looking at stage versus CKD. As the initial stage of diagnosis increases, the more patients in that group have CKD. In this case, 0-4 represents stages 1-5 of CKD. The shift from blue to red in the 1 section indicates that as the stage increases, there are more patients with CKD. The opposite is also true, the higher the stage, the fewer patients there are that don't have CKD. This is a plausible trend, because the higher the stage of kidney disease, the worse the kidneys are able to filter substances and this would indicate kidney failure.

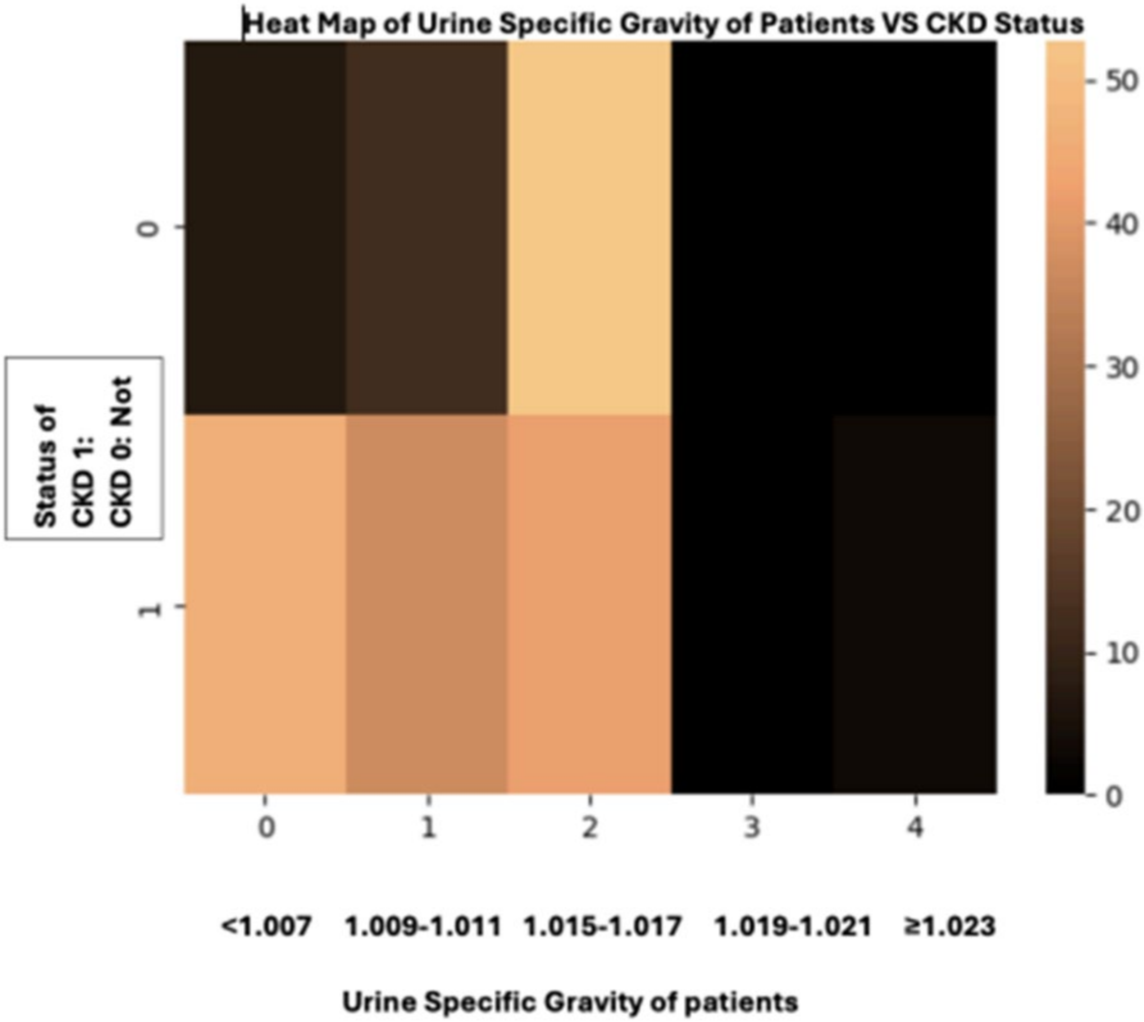

Figure 5. Heat Map of Urine Specific Gravity of Patients vs CKD status. (0= Patient does not have CKD; 1= Patient has CKD)

Finally, the opposite trend is present with Urine Specific Gravity. 0-4 represents <1.007 to ≥1.023. In this case as urine specific gravity decreases, the number of patients with CKD increases, as the lighter color is on the side with less urine specific gravity. Urine specific gravity describes the total concentration of chemicals within the urine (University of California San Francisco [UCSF], 2019). Lower urine specific gravity being associated with more CKD patients makes sense, as it could indicate that the kidney's filtering has been compromised. The purpose of the kidney is to filter waste from the blood that is released as urine. The more chemical particles in the urine, the more waste was filtered from the kidneys. Ergo, less waste in the urine means, the waste is being accumulated within the body. However, this correlation may not be as strong, because lowered urine specific gravity can also be caused by increasing fluid consumption (UCSF, 2019), which would reduce the concentration of chemicals by diluting the urine solution.

Methods and Materials

As previously stated, the data comes from Bangladeshi patient data from Enam Medical College. The dataset comprises 200 patients, with 127 having ESRD, and 73 not having it. The reason for using this data was because there were many medical factors in this data set. Therefore, it would ensure that the reason for CKD progression would most likely be because of those factors, rather than an unknown factor not seen on billable data. To actively interact with the data, the Pandas library was used. Pandas is an open-source library for making data structures known as data frames, which allow users to interact with the data ("pandas…", n.d). The first step when working with this data was to preprocess it. This data did not come preprocessed and had multivariate data in the form of strings and integers. This is a problem, because a machine learning algorithm cannot interpret pure strings as data and needs numerical data. To start, all null values in the dataset needed to be removed, as they served no purpose in the training of the model. The original dataset had 202 rows, and the first two were removed, as they only contained null values. Secondly, many columns had abbreviations that may not make sense to people without a medical background, so many of them were renamed. For example,

- sg: Urine Specific gravity
- al: Aluminium Toxicity

- rbc : Red Blood Cell
- su: Serum urate
- ba: Highly elevated Plasma levels
- rbcc : Red Blood Cell Calcium

Then several data visualizations were made using the seaborn library. Seaborn is a library used to visualize data through statistical graphics ("seaborn…", n.d.). This library was first used to make a histogram of stages to see the distribution of the various stages in the dataset. Next it was used to make box plots for various variables to see the distribution of values for these variables. These variables include Blood Urea Nitrogen, Blood Sugar Random, Red Blood Cell Calcium, Hemoglobin, Syncope and Collapse, pcv, sod, and age.

After these box plots were made, all the above factors were preprocessed. All of the above-mentioned factors used string data to represent a range of values for the patient. Those strings were replaced with various numbers to represent the data ranges. For example, with age, <12:0, 12-20:1, 20-27:2,...., greater than or equal to 74: 9. Other factors that were preprocessed like this included class, Urine Specific Gravity, and glomerular filtration rate. It was important to do this after making the boxplots, because otherwise the boxplots would have a range of numbers, instead of labels for the values, and they would be uninterpretable. For example, with age, <12 makes more sense than 0, as it is not clear what 0 means.

After that seaborn was used to make heat maps of stage and class (having CKD or not), age and class, and finally Urine specific gravity and class. Heatmaps are used to show the correlation between two variables. These variables were sorted earlier to see if they could be relevant to CKD, and heat maps were made to check whether these 3 factors correlated with CKD. The reason why heatmaps were made after data preprocessing is that the heatmaps required numerical data and couldn't be made with the string labels.

Now that the data was preprocessed and visualized, the coding of the models could begin. There were three models used in this study. A logistic regression model, a ridge regression model, and a decision tree classifier. All three of these models were imported from sci-kit learn. Sci-kit learn is a library for machine learning models used to handle data analysis problems such as classification and regression ("sci-kit…", n.d.).

These 3 models were chosen, as they are used in classification studies. The main aim of this research study is to classify whether patients will progress to ESRD or not, so these models would be helpful. Logistic regression is used to highlight the relationship between predictor variables and a categorical response variable ("12.1- Logistic…", n.d.). The logistic regression model shows the probability of a variable being classified a certain way based on the values of other variables used to determine its value. Ridge classification puts a restriction of the values of certain parameters to lower the mean sum of squares to ensure lower error and higher accuracy (" 5.1 -Ridge…", n.d.). Additionally, the

Ridge Classifer was not used in the study by Lee and Collegues. As stated above,a decision tree algorithm uses various features to data features to make decisions. Based on the different data features, decision tree algorithms make different decisions and go down different data paths, causing different predictions ("1.10 Decision…", n.d.). In this case, the algorithm would predict whether the patient will progress to ESRD or not.

For all three of these models, the train-test split function was used to have 80% of the data be the training data, and 20 % of the data is the testing data. The data was separated into X_train, X_test, y_train, and y_test. Y was the class column, and X was a dataframe of every column except 'class', 'Aluminum Toxicity', 'Serum urate','pot','glomerular filtration rate' and 'affected'.

The final part of this study was determining the metrics used to evaluate these models. The first metric used was accuracy. Accuracy refers to the number of correct predictions out of the total number of predictions. Additionally, bar graphs were made of each model, highlighting how important a certain factor was when predicting ESRD. This would allow the developer to see what factors may be important in predicting ESRD. However, accuracy doesn't show the whole picture, and other classification metrics needed to be used. The first two being precision and recall. Precision is defined as the number of relevant items among those retrieved. In this case, it would be the number of patients correctly predicted as progressing to ESRD out of all patients that were predicted to progress to ESRD. Recall is the exact opposite of precision and is defined as the number of retrieved items among those relevant. In this case, that would be the number of patients predicted progressing to ESRD out of all patients that progressed to ESRD. The formulas to calculate precision and recall are shown below.

Equations 1 and 2: Formulas of precision and Recall

$$\text{Precision} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Positive})}$$

$$\text{Recall} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Negative})}$$

In this study, a true positive would be that a patient progressed to ESRD and the model correctly predicted this. A false positive would be a patient did not progress to ESRD, but the model predicted that they would. A true negative would be that the patient did not progress to ESRD and our model correctly predicted that they would not progress. A false negative, the most dangerous situation in this study, is that the patient progressed to ESRD, but the model predicted that they would not. The next metric used in this study was an F1 score. An F1 score weighs precision and recall equally to evaluate the performance of the model. This prevents overestimating the model's performance just because of either high precision or high recall. The formula to calculate it is shown below.

The program,including the three AI models, the data preprocessing, and the functions to calculate the metrics, was run five times for a total of five trials. At the end of each trial, the accuracy, precision, recall, and F1 score were recorded for each model. Finally, the mean scores for each of these metrics were calculated. The reason for multiple trials is due to the data set only having 200 patients, and therefore there is not much data for the performance of the models. The presence of five trials not only provides more data, but it can allow us to see whether the AI performed consistently.

Equation 3: Formula for F1 score

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

Results and Discussion

To start with accuracy, the logistic regression model's mean accuracy was 97.5%; the ridge classifier's mean accuracy was 97.5%; the decision tree classifier's mean accuracy was 97.5%. From the standpoint of accuracy, there is no indication of which model is better.

Trials and Mean Score	Logistic Regression	Ridge Classifier	Decision Tree Classifer
Trial 1	0.975	0.975	0.975
Trial 2	0.975	0.975	0.975
Trial 3	0.975	0.975	0.975
Trial 4	0.975	0.975	0.975
Trial 5	0.975	0.975	0.975
Mean	0.975	0.975	0.975

Figure 6. Accuracy scores for all three models

The most common factors that the logistic regression model used in the diagnosis of CKD are listed below. The importance of each factor was the same for all five trials.

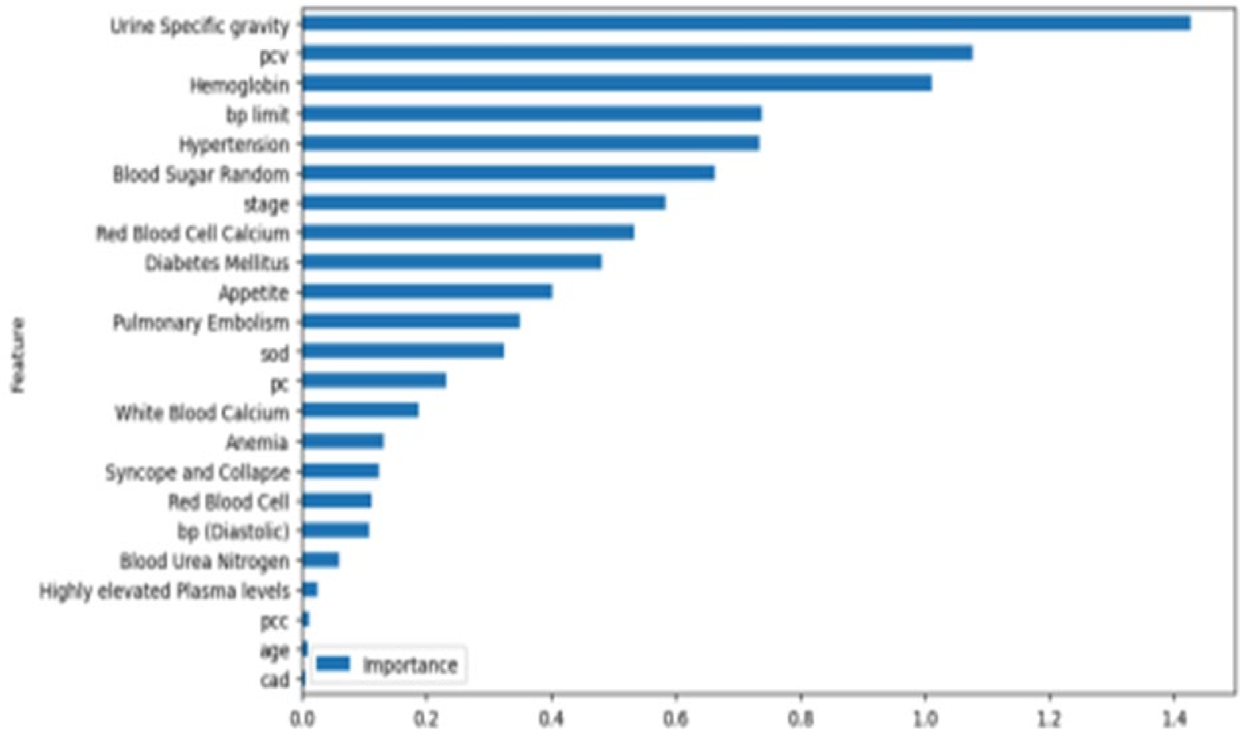

Figure 7. Importance of various features for the logistic regression model

The top 3 factors used were Urine Specific Gravity, pcv, and Hemoglobin. This elucidates that Urine Specific Gravity could be a strong indicator of CKD, as a decrease in it could mean that the kidneys aren't properly filtering waste, as it isn't leaving the body through urine.

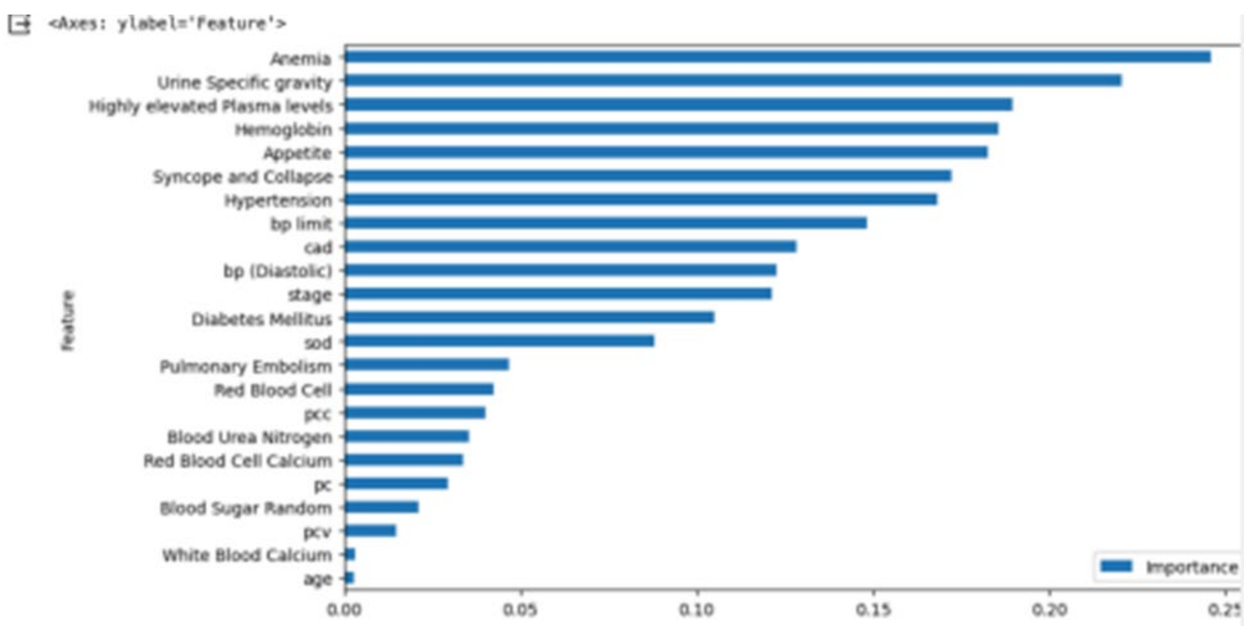

Figure 8. Importance of features for the Ridge Classifier Model

While Urine Specific Gravity was important, for the Ridge Classifier, it was no longer the most important factor, as anemia was. Anemia has been shown to occur with CKD. According to the Mayo Clinic (2023) , Anemia occurs when there are not enough healthy red blood cells to carry oxygen to the blood. This can occur with CKD, because if a person has CKD, in many cases their bodies cannot make erythropoietin, which causes less healthy blood cells to form (Shaikh et al, 2023). While the most important factors are different for both, it still highlights that CKD can be found in a myriad of ways and it can be helpful for a doctor to consider all of these.

The Logistic Regression model had a mean precision of .96, a mean recall of 1.0, and a mean F1 score of .980. This indicates that the Logistic regression model mainly struggled with false positives due to the lower recall meaning, that the model identified CKD in cases it was not supposed to.

Trials and Mean Score	Precision	Recall	F1 Score
Trial 1	0.960	1.00	0.980
Trial 2	0.960	1.00	0.980
Trial 3	0.960	1.00	0.980
Trial 4	0.960	1.00	0.980
Trial 5	0.960	1.00	0.980
Mean	0.960	1.00	0.980

Figure 9. Precision, Recall, and F1 scores for the Logistic Regression Model

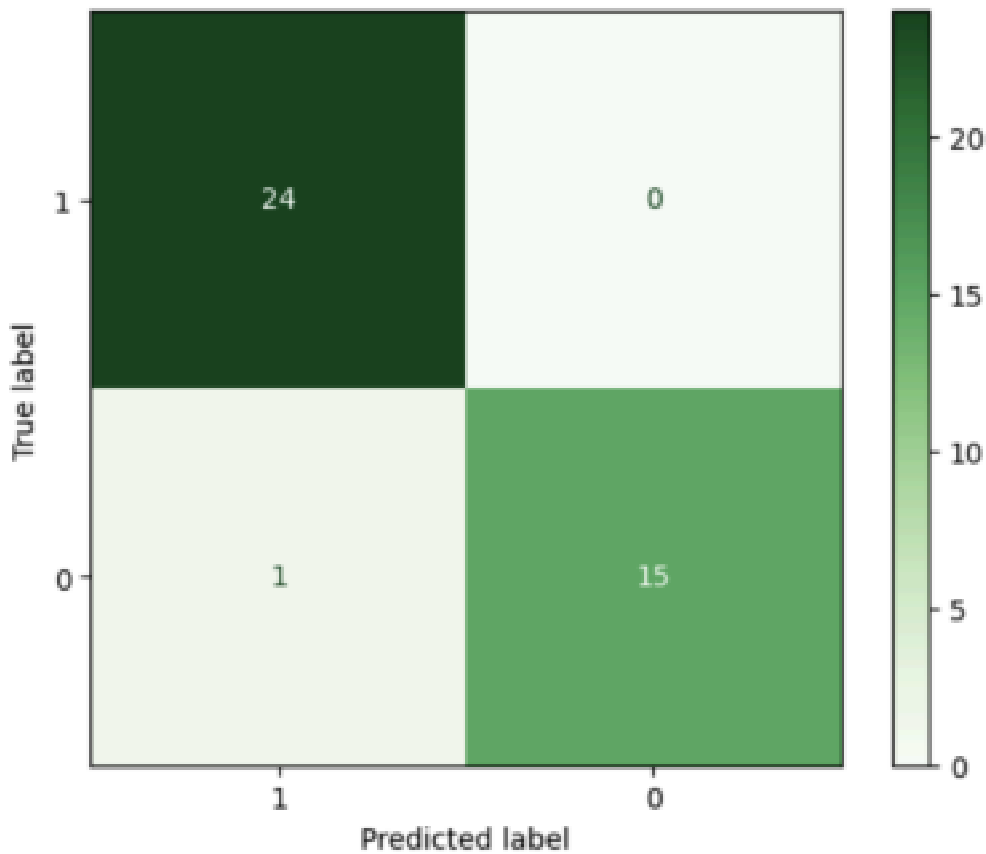

Figure 10. Confusion Matrix for Logistic Regression Model. (0= Patient does not have CKD; 1= Patient has CKD)

This is highlighted by the confusion matrix, as the model only had a single false positive. This indicates that the logistic regression model is an exceptional model when predicting progression to CKD, as it had no false negatives. False negatives are much more detrimental compared to false positives in this case, as a false negative means that the patient has CKD, but the model predicted that they do not. This is extremely dangerous, as if a doctor uses the false negative prediction, this could prevent the patient from getting treatment. In the case of CKD, the longer the patient has it, the more likely they will progress to ESRD, and the lower the probability is that they get successful treatment.

The ridge classifier model had a mean precision of 1.00, a mean recall of 0.958, and a mean F1 score of 0.979. This is indicated by the confusion matrix, as the Ridge Classifier has 1 false negative.

Trials and Mean Score	Precision	Recall	F1 Score
Trial 1	1.00	0.958	0.979
Trial 2	1.00	0.958	0.979
Trial 3	1.00	0.958	0.979
Trial 4	1.00	0.958	0.979

Trial 5	1.00	0.958	0.979
Mean	1.00	0.958	0.979

Figure 11. Precision, Recall, and F1 scores for the Ridge Classifier Model

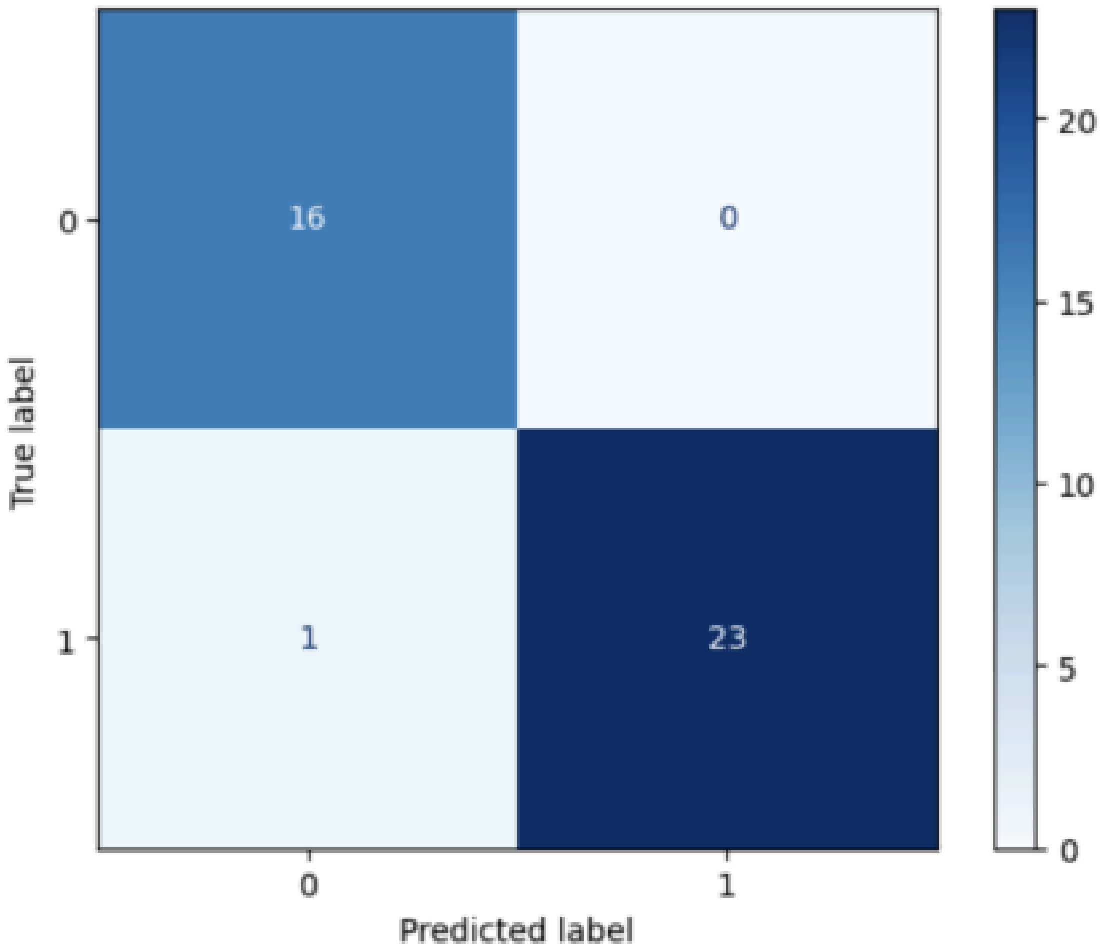

Figure 12. Confusion Matrix for the Ridge Classifier Model. (0= Patient does not have CKD; 1= Patient has CKD)

The decision tree classifier model had a mean precision of 1.00, a mean recall of 0.958, and a mean F1 score of 0.979. This is indicated by the confusion matrix, as the Random Forest Classifier has 1 false negative.

Trials and Mean Score	Precision	Recall	F1 Score
Trial 1	1.00	0.958	0.979
Trial 2	1.00	0.958	0.979

Trial 3	1.00	0.958	0.979
Trial 4	1.00	0.958	0.979
Trial 5	1.00	0.958	0.979
Mean	1.00	0.958	0.979

Figure 13. Precision, Recall, and F1 scores for the Decision Tree Classifier Model

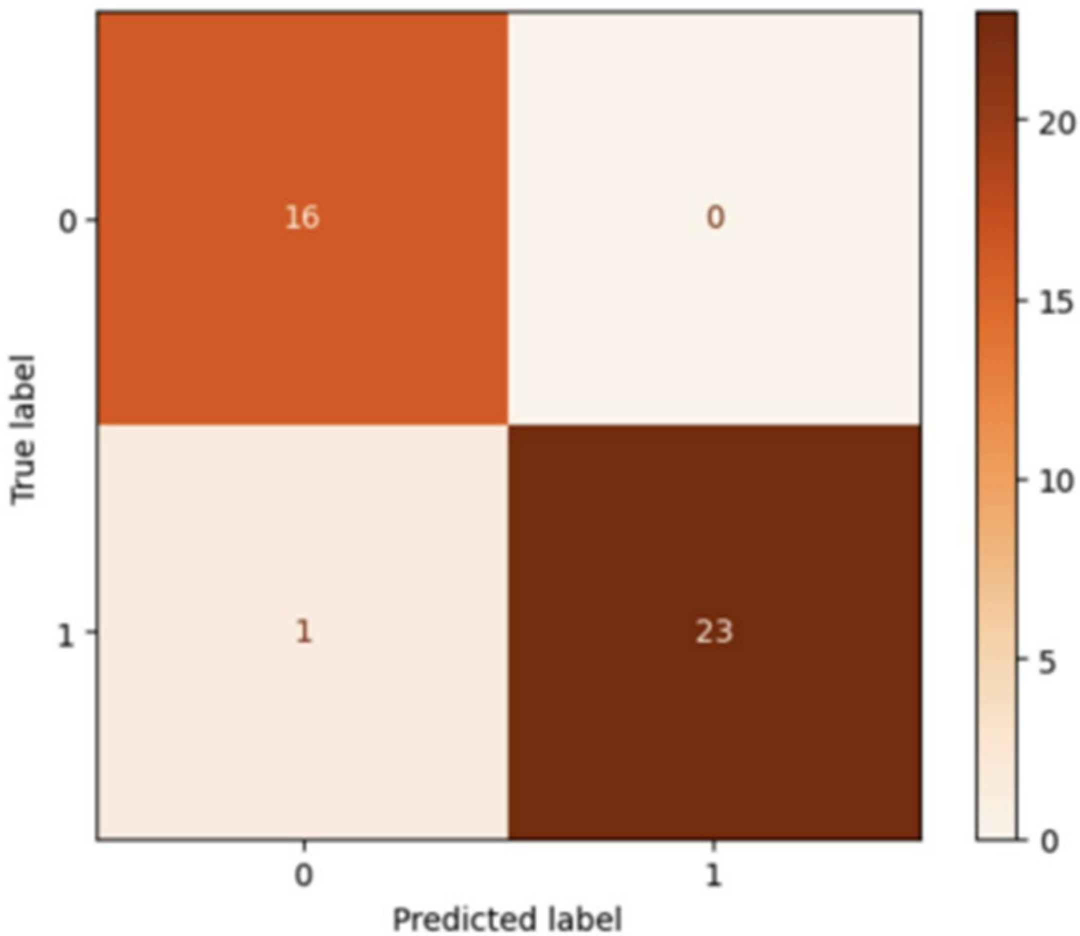

Figure 14. Confusion Matrix for the Random Forest Classifier Model

Overall, while all three models are successful at predicting CKD, the logistic regression model is slightly more successful. This is not only indicated by the slightly higher f1 score, but the overall scores in precision and recall. While the Ridge Classifier and the Random Forest Model have higher precision, the logistic regression model has higher recall. This means that it is less likely to predict a false negative. False negatives are much more dangerous, as they can lead to delayed detection, meaning from a medical standpoint, the logistic regression model is less likely to endanger the lives of patients.

Conclusions

Ultimately, while all three models are successful at predicting CKD, the logistic regression model was the most successful model due to it having the best recall. Any of these models and the criteria they used can make the organ donation system more efficient by allowing for early detection of CKD and by determining which patients need a kidney transplant the most. This will help solve the problem of the growing gap between demand and supply in the organ donation system. By eliminating factors such as race, sex, and economic status, and only determining allocation based on which patients are more likely to progress to ESRD, more lives will be saved.

Limitations

However, despite the success of the models in this study, this study has several limitations. The main limitation of this study is the size of the dataset. Unfortunately, the dataset has only 200 patients. This means that when the data is split, there are only 40 patients for the AI model to be tested on. Due to the limited number of patients, the accuracy, precision, and recall of the models are likely inflated, as the more patients there are in the study, the more likely the model will make mistakes. Also, with fewer patients, the data is less likely to be consistent in the long run. Furthermore, this model was tested only on a specific group of patients from Bangladesh, so to generalize the results of the study, it needs to be tested on more datasets to better generalize the results, and to get a more accurate indication of the models' performance. Regardless of these limitations, these models and the results of these studies can be greatly implemented into organ donation policies. These models and the criteria used by these models can help doctors predict which patients will most likely have CKD. If CKD can be detected early, then the likelihood of progression to ESRD is greatly reduced. As stated before, the only way to treat ESRD is through dialysis and a kidney transplant. Therefore, by reducing the likelihood of progression to ESRD, increases the likelihood of other treatment options to reduce the severity of the disease. These include lifestyle changes and medicine to reduce blood pressure and cholesterol ("Treatment…). This would in turn reduce the number of patients in severe condition and would also reduce the demand for an immediate kidney transplant.

Along with closing the gap between the supply and demand for kidneys, these models and criteria can also be applied by determining which patients need kidney transplants the most. By predicting which patients have CKD, along with analyzing factors such as stage, doctors can determine which patients are most likely to progress to ESRD. This determination will provide objective criteria for determining the allocation of organs.

Acknowledgments

I would like to thank Stanford graduate Alaisha Alexander for being an amazing mentor during this research process. This research wouldn't have come as far as it did without your guidance.

References

Anemia - symptoms and causes. (2023, May 11). Mayoclinic. Retrieved March 23, 2024, from https://www.mayoclinic.org/diseases-conditions/anemia/symptoms-causes/syc-20351360#:~:text=Overview,weakness%20and%20shortness%20of%20breath.

Bunnik, E. M. (2023). Ethics of allocation of donor organs. *Current Opinion in Organ Transplantation*, *28*(3), 192-196. https://doi.org/10.1097/mot.0000000000001058

European Commision. (n.d.). [Organs]. health.ec.europa.eu. Retrieved August 31, 2024, from https://health.ec.europa.eu/blood-tissues-cells-and-organs/organs_en#:~:text=Kidneys%20are%20the%20most%20frequently,transplantation%20are%20continuously%20bein

Health Resources and Services Administration. (2024, February 1). *Organ Donation Statistics* [Organ Donation Statistics]. organdonor.gov. Retrieved March 18, 2024, from https://www.organdonor.gov/learn/organ-donation-statistics

Lee, K.-H., Chu, Y.-C., Tsai, M.-T., Tseng, W.-C., Lin, Y.-P., Ou, S.-M., & Tarng, D.-C. (2022). Artificial Intelligence for Risk Prediction of End-Stage Renal Disease in Sepsis Survivors with Chronic Kidney Disease. *e Acute Kidney Injury to Chronic Kidney Disease: Pathophysiology and Therapy Targets*. https://doi.org/10.3390/biomedicines10030546

MayoClinic. (2023, September 6). *Chronic kidney disease - symptoms and causes*. Mayoclinic. Retrieved March 23, 2024, from https://www.mayoclinic.org/diseases-conditions/chronic-kidney-disease/symptoms-causes/syc-20354521

MayoClinic. (2023, October 10). *End-stage renal disease - symptoms and causes*. Mayoclinic. Retrieved March 23, 2024, from https://www.mayoclinic.org/diseases-conditions/end-stage-renal-disease/symptoms-causes/syc-20354532

Md. Ashiqul Islam, Shamima Akter. *Risk Factor Prediction of Chronic Kidney Disease*. UCI Machine Learning Repository, 2020, https://doi.org10.24432/C5WP64

1.10. Decision Trees [1.10. Decision Trees]. (n.d.). scikit-learn.org. Retrieved March 19, 2024, from https://scikit-learn.org/stable/modules/tree.html

pandas documentation [pandas documentation]. (2024, February 23). pandas.pydata.org. Retrieved March 20, 2024, from https://pandas.pydata.org/docs/

seaborn: statistical data visualization [seaborn: statistical data visualization]. (n.d.). seaborn.pydata.org. Retrieved March 20, 2024, from https://seaborn.pydata.org/

Segal, Z., Kalifa, D., Radinsky, K., Ehrenberg, B., Elad, G., Maor, G., Lewis, M., Tibi, M., Korn, L., & Koren, G. (2020). Machine learning algorithm for early detection of end-stage renal disease. *BMC Nephrology*, *21*(1). https://doi.org/10.1186/s12882-020-02093-0

Shaikh, H., Hashimi, M. F., & Aeddula, N. R. (2023). Anemia of Chronic Renal Disease [Anemia of Chronic Renal Disease]. *National Instiutes of Health*. https://www.ncbi.nlm.nih.gov/books/NBK539871/

Treatment Chronic Kidney Disease [Treatment Chronic Kidney Disease]. (n.d.). nhs.uk. Retrieved March 23, 2024, from https://www.nhs.uk/conditions/kidney-disease/treatment/#:~:text=There's%20no%20cure%20for%20chronic,stay%20as%20healthy%20as%20possible

University of Calinforina San Francisco. (2019). *Urine Specific Gravity Test* [Urine Specific Gravity Test]. ucsfhealth.org. Retrieved March 20, 2024, from https://www.ucsfhealth.org/medical-tests/urine-specific-gravity-test

Wesselman, H., Ford, C. G., Leyva, Y., Li, X., Chang, C.-C. H., Dew, M. A., Kendall, K., Croswell, E., Pleis, J. R., Ng, Y. H., Unruh, M. L., Shapiro, R., & Myaskovsky, L. (2021). Social determinants of health and race disparities in kidney transplant. *Clinical Journal of the American Society of Nephrology*, *16*(2), 262-274. https://doi.org/10.2215/cjn.04860420

The Word2Vec Algorithm [The Word2Vec Algorithm]. (2018, November 27). datasciencecentral.com. Retrieved March 19, 2024, from https://www.datascience