# Diagnosing Urinary System Diseases

Katie Briggs

Bellevue University

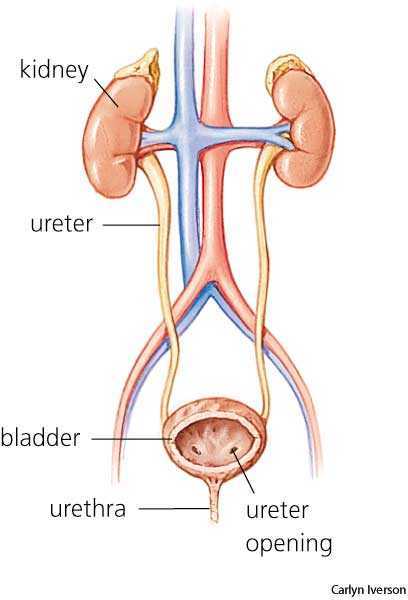
7 February 2021

**Abstract**

The urinary system works to filter blood and create urine as a waste by-product. The organs within the urinary system include the kidneys, renal pelvis, ureters, bladder, and the urethra. Our bodies take food components that it requires, and the waste is what is left behind in our bowels and in our blood. The kidney and urinary systems purge the liquid waste, which in known as urea. It also helps keep our potassium, sodium and water levels balanced. [3]

Nephritis is a condition where the nephrons, which are in the kidneys, become inflamed. This inflammation can adversely affect the kidney function [2]. Acute Nephritis of the renal pelvis occurs more often in women than in men. The typical symptoms of this disease are fever and lumbar pain.

Inflammation of the bladder, which is commonly referred to as Cystitis. It can be caused by a bacterial infection and spread into the kidneys. The symptoms are frequent urination, urethra discomfort, pelvic discomfort, and low-grade body temperature.



Urinary System – Appendix A

**Background**

With symptom of Nephritis and Inflammation of Bladder being similar it can become difficult to diagnosis the correct disease for an ailing patient. The reasoning behind this data set was for performing a diagnosis on two of the mentioned diseases through an algorithm and a decision tree. Symptoms were gathered from patients and placed into the data set. Acute Inflammation of the Bladder is described as a sudden onset of pain in the lower abdomen, frequent urination, and micturition pain. Low grade fever is present, but usually not above 38 degrees Celsius. Nephritis also is present with a fever, pain in the back, frequent urination, micturition pain and nausea.

By analyzing the relationships between the patient’s symptoms, it is possible to have a symptom checker accurately aid in the diagnosis.

**Data**

The data set for the analysis is the Acute Inflammation Data Set from UCI Machine Learning Repository. The data was gathered by a medical expert as a data set to test expert systems, that attempt to perform the presumptive diagnosis of two urinary system diseases: Nephritis and Inflammation of Bladder. [1] The data set was split into two different files. One named diagnosis names and the other diagnosis data. For data preparation, the two ASCII files where merged and the columns renamed to the symptoms that occurs with both illnesses in a CSV file. I renamed it diagnosis.csv. The data set characteristics are multivariate, and the attributes are categorical and integer. The data set consists of eight columns and 120 rows. Six of the eight columns are the attributes/symptoms. Missing values were removed.

Attribute Information:

a1 Temperature of patient { 35C-42C }

a2 Occurrence of nausea { yes, no }

a3 Lumbar pain { yes, no }

a4 Urine pushing (continuous need for urination) { yes, no }

a5 Micturition pains { yes, no }

a6 Burning of urethra, itch, swelling of urethra outlet { yes, no }

d1 decision: Inflammation of urinary bladder { yes, no }

d2 decision: Nephritis of renal pelvis origin { yes, no }

Once the columns were created, the data was easier to read and analyze.

|  | **Temp** | **nausea** | **Lumbarpain** | **urination** | **Micturitionpains** | **urethrasymptoms** | **Inflammation** | **Nephritis** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 35.5 | no | yes | no | no | no | no | no |
| **1** | 35.9 | no | no | yes | yes | yes | yes | no |
| **2** | 35.9 | no | yes | no | no | no | no | no |
| **3** | 36.0 | no | no | yes | yes | yes | yes | no |
| **4** | 36.0 | no | yes | no | no | no | no | no |
|  |  |  |  |  |  |  |  |  |

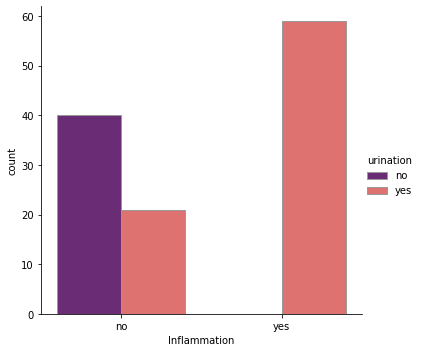
DataFrame Diagnosis.CSV – Appendix B

**Methods**

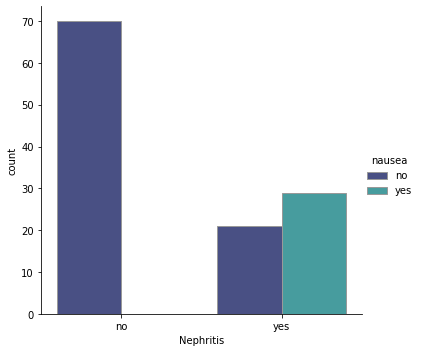
For the analysis, I followed the Cross-Industry Standard Process for Data Mining (CRISP-DM). CRISP-DM is a process encompassing six steps that guide us to develop and implement models from data. These steps are, in order, business understanding, data understanding, data preparation, modeling, evaluation, and deployment. During the business understanding portion, the problem was identified, and a solution was purposed.

For the data understanding step, many references were read and viewed to aid in the symptom collection for the data set. The background information was attributed to the medical professional that also gathered the data into the ASCII file. During the data preparation stage, missing values were removed, and the data sets were merged. Also, in this step of CRISP-DM the data was explored further with plotting the symptom relationships with seaborn.

The most insightful relationships are below. What is tells me is that the symptoms of frequent urination are seen in Inflammation of Bladder and not seen as much in patients with Nephritis.

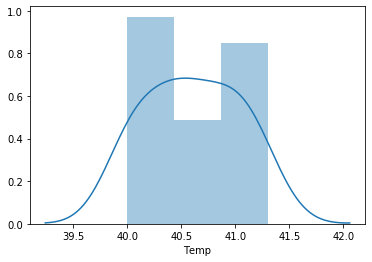
 Plot of Frequent Urination and Inflammation – Appendix C

In this plot we see that nausea is seen more frequently with Nephritis than with Inflammation of the Bladder.



Plot of Nausea Symptoms in Nephritis – Appendix D

No outliers were identified within all variables. Temperature was the only categorical variable within the data set. I looked at the relationship between patient’s with inflammation only and their reported body temperature. The range for all patients was between 35.5 and 41 degrees Celsius. That is 95.9- and 106-degrees Fahrenheit. The distplot shows the temperature range of the people who have Inflammation of the Bladder, but not Nephritis. It can be seen the body temperature does vary between 36-38 degrees Celsius, 96.8-100.4 degrees. I also plotted the temperature for Nephritis patients and the results were a higher fever. I also wanted to know if the inflammation of bladder had move to the kidneys to cause Nephritis. With both diseases, the patient’s body temperature would be 104 F and above.



Plot of Body Temp in Nephritis and Inflammation – Appendix A

It shows that the patients with inflammation and acute nephritis have a body temp range between 40-41.5. Which is higher than Inflammation alone.

**Results**

Once I completed data analysis, I created a model for diagnosing Nephritis. I used label encoder and standard scaler to fit predictors from the symptoms, which are my column headers. Then, I trained, tested, and split the data with a test size of 25 percent. After training and testing, best parameters were determined. The criterion is Friedman\_MSE with a learning rate of 10 percent. The model was boosted with the gradient boosting classifier and the accuracy was 100 percent. The confusion matrix was 16, 0 and 0, 14. The decision classifier was completed with max\_features that were the square root and a gini. It was fitted and then predict was applied. Again the accuracy was checked at 100 percent.

DecisionTreeClassifier

(ccp\_alpha=0.0, class\_weight=None, criterion='gini',

max\_depth=None, max\_features='sqrt', max\_leaf\_nodes=None,

min\_impurity\_decrease=0.0, min\_impurity\_split=None,

min\_samples\_leaf=1, min\_samples\_split=2,

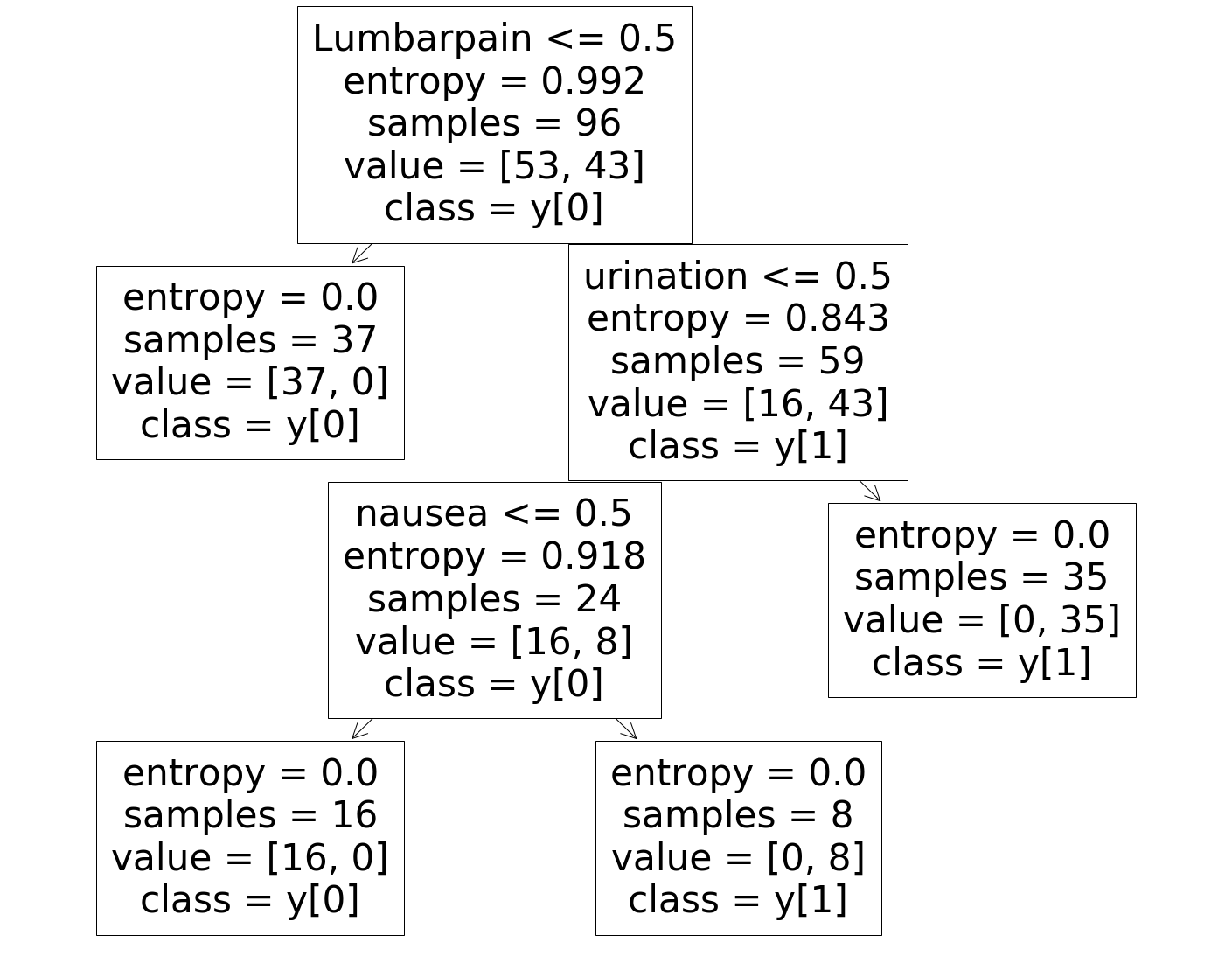
min\_weight\_fraction\_leaf=0.0, presort='deprecated',

random\_state=None, splitter='best')

Once the model was complete, I wanted to create a decision tree. I had to turn the ‘No’ to ‘0’ and the ‘Yes’ to ‘1’. The independent and dependent variable were identified. Once again, the data was trained, tested, and split, with a test size of 20 percent. The model was fitted, and the max depth set to 5. The prediction accuracy of the model came back at 100 percent.

**Conclusion**

Lumbar pain symptoms are in more than half of the reported cases with a diagnosis of Nephritis. Urination symptoms such as, blood in urine and frequent voiding are also high in patients with diagnosed Nephritis. Nausea and pain while urinating are found more in patients with Inflammation of Bladder.



**Acknowledgements**

UCI Machine Learning Repository, Bellevue University

**References:**

[1] UCI Machine Learning Repository, Acute Inflammation data set, Classification, with multivariate data characteristics, Attributes are categorical and integer. <https://archive.ics.uci.edu/ml/datasets/Acute+Inflammations>

[2] Medial News Today, lists symptoms of acute nephritis, the different types of nephritis and causes. Symptoms will be used as a guide to diagnose in model. Murrell, D., (2018), What to know about nephritis, Retrieved from: <https://www.medicalnewstoday.com/articles/312579>

[3] Grover, S., Srivastava, A., Lee, R., Tewari, A. K., & Te, A. E. (2011). Role of inflammation in bladder function and interstitial cystitis. Therapeutic advances in urology, 3(1), 19–33. Symptoms of acute inflammation in the bladder to use for diagnosis model. Retrieved from: <https://doi.org/10.1177/1756287211398255>

[4] J.Czerniak, H.Zarzycki, Application of rough sets in the presumptive diagnosis of urinary system diseases, Artifical Inteligence and Security in Computing Systems, ACS'2002 9th International Conference Proceedings, Kluwer Academic Publishers, 2003, pp. 41-51

[5] Kaggle, data sets of attributes with symptom information. https://www.kaggle.com/subh86/acute-nephritis

[6] UTDallas, Medical Diagnosis with using decision trees, Understanding how doctor’s diagnose

medical conditions based on symptoms. Retrieved from:

<https://www.utdallas.edu/~nrr150130/cs6375/2015fa/lects/Lecture_6_Decision.pdf>

[7] Decision Tree Classification, An overview of using decision tree classifiers. Chakure, A. (2019),

Retrieved from: <https://medium.com/swlh/decision-tree-classification-de64fc4d5aac>

[8] Gradient Boosting for trees. Brownlee, J.,(2020), A Gentle Introduction to the Gradinet Boosting

Algorithm for Machine Learning, Retrieved from:

[https://machinelearningmastery.com/gentle- introduction-gradient-boosting-algorithm-machinelearning/](https://machinelearningmastery.com/gentle-%20%20%20introduction-gradient-boosting-algorithm-machinelearning/)

[9] Diagnose urinary inflammations using Neural Designer, an example of machine learning diagnosis.

<https://www.neuraldesigner.com/learning/examples/urinary-inflammation-diagnosis>

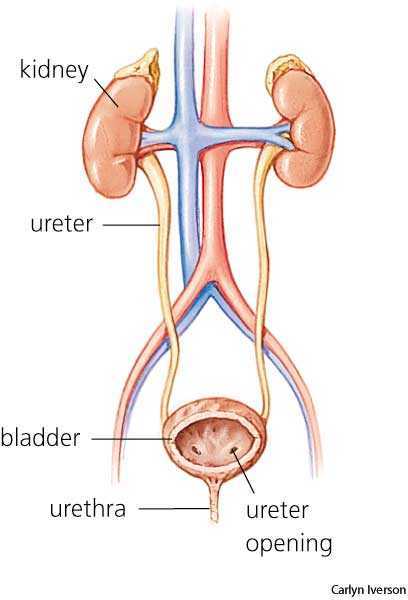
[10] Predicting urinary tract infections in the emergency department with machine learning, Taylor, R.,

Moore, C., Brandt, C., (2018), <https://doi.org/10.1371/journal.pone.0194085>

**Appendix**

Appendix A:

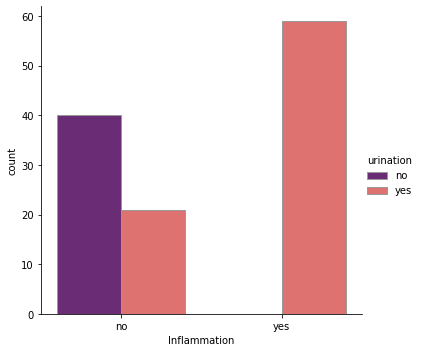
Anatomy of the Urinary System



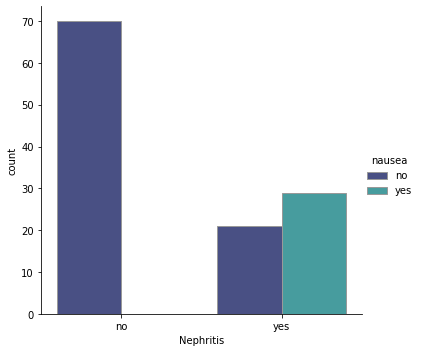
Appendix B: DataFrame of diagnosis.CSV

|  | **Temp** | **nausea** | **Lumbarpain** | **urination** | **Micturitionpains** | **urethrasymptoms** | **Inflammation** | **Nephritis** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | 35.5 | no | yes | no | no | no | no | no |
| **1** | 35.9 | no | no | yes | yes | yes | yes | no |
| **2** | 35.9 | no | yes | no | no | no | no | no |
| **3** | 36.0 | no | no | yes | yes | yes | yes | no |
| **4** | 36.0 | no | yes | no | no | no | no | no |
|  |  |  |  |  |  |  |  |  |

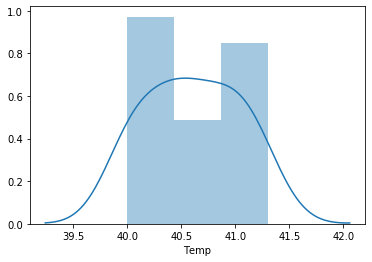
Appendix C: Plot of Urination in Inflammation



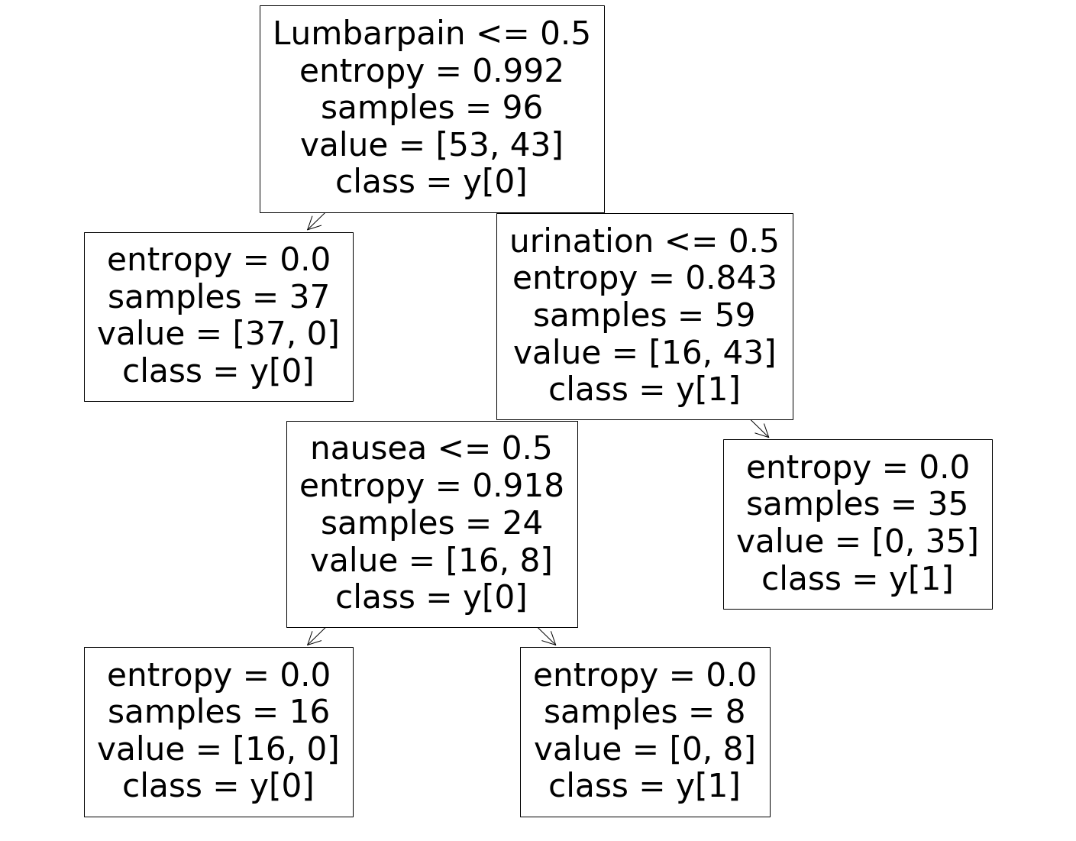
Appendix D: Plot nausea in Nephritis



Appendix E: Plot of Body Temp of Patients with Inflammation and Nephritis



Appendix F: Decision Tree of Nephritis



**Ten Questions:**

What symptoms are most associated with Acute Nephritis?

What symptoms are most associated with Inflammation of the Bladder?

Why did you choose this project?

Did anything surprise you in the data analysis?

What did not go as planned?

What steps did you take to overcome those obstacles?

Are the results what you expected?

What did you lessons did you learn in this project?

Would you use something like this to help identify other diseases.

What did go right?