Acute Cystitis and Acute Nephritis Prediction Using Machine Learning Techniques

R. Kowsalya[1] G. Sasikala[2] J. Sangeetha Priya[3]

GJCST Classification
I.2.6,K.3.2,J.3

***Abstract*-Urinary System includes kidneys, bladder, ureters and urethra. This is the major system involves electrolyte balance of the body and filters the blood and excretes the waste products in the form urine. Even the small disturbance in the renal function will step in a disasters manifestation. Among them we are considering the two diseases that affect the system are acute cystitis and acute nephritis. This paper presents the implementation of three supervised learning algorithms, ZeroR, J48 and Naive Bayes in WEKA environment. The classification models were trained using the data collected from 120 patients. The trained models were then used for predicting the acute cystitis or acute nephritis of the patients. The prediction accuracy of the classifiers was evaluated using 10-fold cross validation and the results were compared.**

Keywords-Urinary System, Ureters, Urethra, AcuteCystitis, Acute Nephritis, classification, WEKA

I. Introduction

Machine learning is a scientific discipline that is concerned with the design and development ofalgorithms that allow computers to change behavior based on data, such as from sensor data or databases. Machine learning usually refers to the changes in systems that perform tasks associated with artificial intelligence(AI). Such tasks involve recognition, diagnosis, planning, robot control, prediction, etc. [1]

A major focus of machine learning research is to automatically learn to recognize complex patterns and make intelligent decisions based on data. Hence, machine learning is closely related to fields such as statistics, probability theory, data mining, pattern recognition, artificial intelligence, adaptive control, and theoretical computer science. The attributes considered for the algorithm comprises temperature of patient , nausea , Lumbar pain, Frequency of micturation (continuous need for urination) , micturition pain, burning sensation during micturition, itch, swelling of urethra outlet. Machine Learning Techniques are effective to classify the data and to improve the predictive accuracy

About-1st R. Kowsalya, Lecturer, GR Govindarajulu School of Applied Computer Technology, PSGR Krishnammal College for Women, Coimbatore.(telephone: 9442013066 email: kowsalya@grgsact.com)
About-2nd G. Sasikala, Lecturer, GR Govindarajulu School of Applied Computer Technology, PSGR Krishnammal College for Women, Coimbatore.(telephone: 98944 13801 email: sasikala@grgsact.com)
About-3rd J. Sangeetha Priya , Lecturer , GR Govindarajulu School of Applied Computer Technology, PSGR Krishnammal College for Women, Coimbatore.(telephone: 99949 10211 email:sangeethapriya@grgsact.com)

II. Motivation

Motivation behind is to apply and analyze three different machine learning algorithm for classification of the urinary system - acute cystitis and acute nephritis. The classification models were trained using the data collected from 120 patients. The trained models were then used for predicting the acute cystitis and acute nephritis of the patients. Data set includes descriptions of hypothetical samples corresponding to 120 patients. It is identified as definitely the patient is affected withInflammation of urinary bladder or Nephritis of renal pelvis origin. [3]

III. Acute Cystitis and Acute Nephritis

Urinary System includes kidneys, bladder, ureters and urethra. This is the major system involves electrolyte balance of the body and filters the blood and excretes the waste products in the form urine. The body has two kidneys located in the lumbar region (back at about the location of the elbows). Each kidney has about 1000 nephrons that act as filter. Each nephron composed of glomeruli and tubules, which works as a filter and an absorber. The blood which carries glucose, electrolytes, and metabolic end products passes through the nephrons that filters and absorb the needed materials for the tissue and excreates the waste products along with the water in the form of urine. The cleaned blood leaves the kidney and travels throughout the body. Thus the kidney places the major role in electrolyte balance. Another elementary function of kidney is secretion of erythropoietin, which maintains the blood pressure.

The bladder is a muscular bag in which urine is stored before being discharged through the urethra. The two ureter from each side of the kidney carries urine from the kidney to the bladder. It can hold between one half to two cups of urine before it needs to be emptied. Everyday about two to five cups of urine pass through the bladder. The urine output is directly proportional to the water intake but it does not hold good in summer season because most of the body water is excreted as sweat.

About 96% of urine is water. It also contains some waste salt and a substance called urea. Urea is made during the breakdown of proteins in liver. Urea is also excreted in sweat. If urea builds up in the body, it is a sign that the kidneys are not working properly. When the kidney fails,the metabolic waste products get accumulatedin the body andlead to the consequent manifestation. For example accumulation of urea may lead on to encephalopathy [4].

Machine Learning Techniques is used to classify the presumptive diagnosis of two diseases ofurinary system.
Acute cystitis, inflammation of the urinary bladder commonly crop up due to the ascending infections by several organisms. It manifests with clinical features of burning micturation, raising temperature, hematuria (passing blood in urine), sudden occurrence of pain in the abdomen region, burning micturition and micturition pain. Symptoms decay usually within several days on proper treatment. However, there is inclination to return. A person with acute cystitis should expect that the illness will turn into protracted form, which also sequlae into hydronephrosis.Acute nephritis, inflammation of the renal parenchyma occurs considerably more often at women than at men. Sudden fever, hematuria, elevated blood pressure, lumbar pain and oliguria are the symptoms of acute cystitis. Quite not infrequently there are nausea and vomiting and spread pains of whole abdomen.

IV. Machine Learning Approach and Algorithm Basis

For analyzing the data and classification of acute cystitis and acute nephritis, the three Machine learning algorithms ZeroR, J48 Pruned tree and Naive Bayes classifier were adopted here. Zero-R algorithm is used to predict the majority class in the training data.J48 Pruned tree algorithm is an implementation of the C4.5 decision tree learner. This implementation produces decision tree models. The algorithm uses the greedy technique to bring decision tree for classification. A decision-tree model is built by analyzing training data and the model is used to classify unseen data. J48 generates decision tree, the nodes of which evaluate the existence or significance of individual features.
The Naive Bayes Classifier technique is based on Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Naïve Bayes classifiers assume that the effect of a variable value on a given class is independent of the values of other variable.
The conditional probability of attribute value given class is computed by figuring out the proportion of instances.Depending on the accurate nature of the probability model, Naive Bayes classifiers can be trained very efficiently in a supervised learning setting.

V. Experimental Setup

The data analysis and classification was carried out using WEKA software environment for machine learning. The WEKA, Open Source, Portable, GUI-based workbench is a collection of state-of-the-art machine learning algorithms and data pre-processing tools [2]. It is designed in flexible manner to try out existing methods on new dataset. In this experiment, the data set collected from UCI Repository of 120 patients with 8 features is selected as the class label. The instances in the dataset are pertaining to the two categories based on the temperature of patient, occurrence of nausea, lumbar pain, urine pushing (continuous need for urination), micturition pain, burning of urethra, itch, swelling of urethra outlet, inflammation of urinary bladder and nephritis of renal pelvis origin. The attributes are labeled as a1, a2, a3, a4, a5, a6, d1, d2. To evaluate the robustness of the classifier, the normal methodology is to perform cross validation on the classifier. In general, ten fold cross validation has been proved to be statistically good enough in evaluating the performance of the classifier. The machine learning techniques is implemented using WEKA tool. The 10-fold cross validation was performed to test the performance of the three models. The prediction accuracy of the models was compared.

VI. Result and Discussion

The results of the experiments are described in Table 1, 2 and 3. The performances of the three models were evaluated based on the three criteria, the prediction accuracy, learning time and error rate and illustrated in Figures 1, 2 and 3.

Table 1. Predictive Performance Of The Classifiers

Classifiers	Zero R	Naive Bayes	J48 Pruned tree
Time taken to build the model (in sec)	0	0	0.02
Correctly Classified Instances	70	120	120
Incorrectly Classified Instances	50	0	0
Prediction Accuracy	58.3333 %	100%	100%

TABLE 2. COMPARISON OF ESTIMATES

Validation	Zero R	Naive Bayes	J48 Pruned tree
Kappa statistic	0	1	1
Mean absolute error	0.4864	0.0602	0
Root mean squared error	0.493	0.1017	0
Relative absolute error	100 %	12.3797 %	0 %
Root relative squared error	100 %	20.6264 %	0 %

Table 3. Comparison Of Evaluation Measures ByClass

Classifier	**TP Rate**	**FP Rate**	**Precision**	**Recall**	**F-measure**	**ROC**	**Class**
Zero R	1	1	0.583	1	0.737	0.5	no
	0	0	0	0	0	0.5	yes
Naïve Bayes	1	0	1	1	1	1	no
	1	0	1	1	1	1	yes
J48 Purned Tree	1	0	1	1	1	1	no
	1	0	1	1	1	1	yes

From the confusion matrix given in Table 4, it is observed that Naive Bayes and J48 Pruned Tree produce relatively good results.

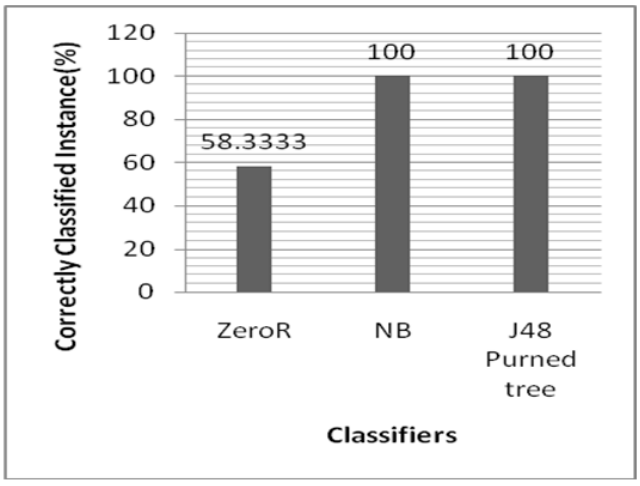
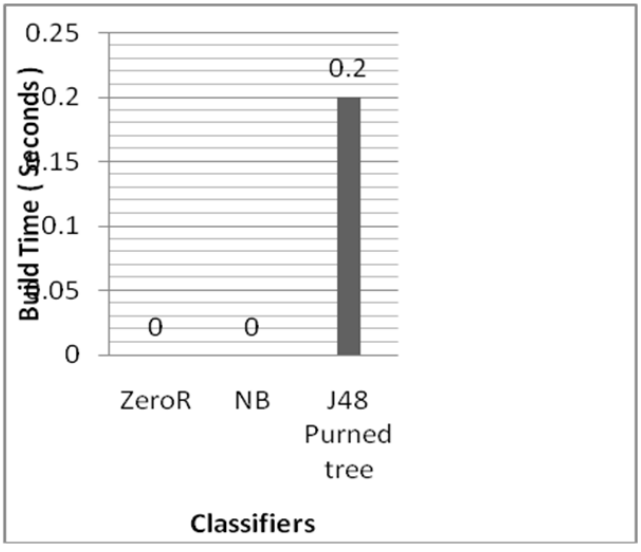


Fig.1. Prediction Accuracy



VII. Conclusion

Classifier systems play a major role in machine learning and knowledge-based systems. In this paper three supervised learning algorithm was implemented using WEKA software. By classifying each attributes to predict the accuracy of each algorithm and test the correctly classified instances of the attribute. The result percentage was compared to identify the algorithm that is well suited to classify the acute cystitis urinary bladder and acute nephritisis. The results indicate that the Naive Bayes classifier outperforms in prediction than ZeroR and J48 Pruned algorithm. Further work can be extended by repeating the experiment with other machine learning algorithms.

Figure 2 illustrates the learning time of the three schemes under consideration. J48 Pruned tree classifier takes moretime to build the model. The Naïve Bayes and ZeroR, the probabilistic classifier tends to learn more rapidly for the given dataset.

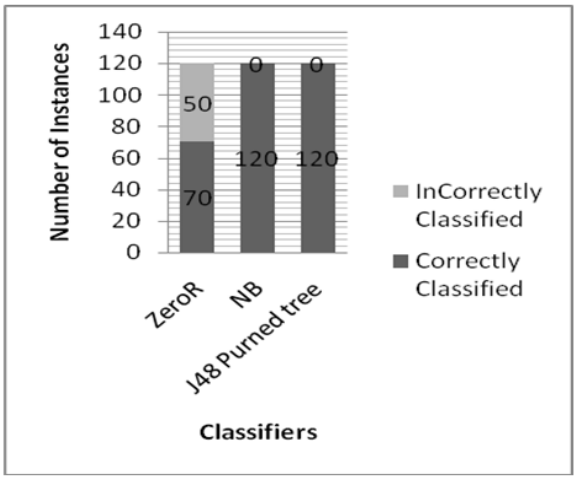


Fig.3. Error Rate

The confusion matrix was used to evaluate the classification error rate. From the confusion matrix given in Table 4, it is observed that Naive Bayes and J48 Pruned tree produce relatively good results but the time taken for the J48 Pruned tree is high when compared to the Naive Bayes.

CLASSIFER	a	b
Zero R	70 0 | a = no	50 0 | b = yes
Naive Bayes	70 0 | a = no	0 50 | b = yes
J48 Purned Tree	70 0 | a = no	0 50 | b = yes

VIII. References

1) Ethem Alpaydin.(2004). Introduction to Machine Learning, MIT Press.
2) Ian H. Witten, Eibe Frank. (2005). Data Mining – Practical MachineLearningToolsandTechniques,2nd Edition, Elsevier.
3) J.Czerniak, H.Zarzycki, (2005) Application of rough sets in the presumptive diagnosis of urinary system diseases, Artificial Intelligence and Security in Computing Systems, ACS'2002 9th International Conference Proceedings, 41-51.
4) Harrison's Principles of Internal Medicine - 17th Ed. (2008), Printed in the United States of America, The McGraw-Hill Companies, Inc.