



A Model for Prediction of Kidney Cancer Using Data Analytics Technique

Aranuwa Felix Ola

Department of Computer Science, Adekunle Ajasin University, Akungba-Akoko, Nigeria

Email address:

felix.aranuwa@aaau.edu.ng

To cite this article:

Aranuwa Felix Ola. A Model for Prediction of Kidney Cancer Using Data Analytics Technique. *American Journal of Data Mining and Knowledge Discovery*. Vol. 5, No. 2, 2020, pp. 27-36. doi: 10.11648/j.ajdmkd.20200502.12

Received: August 31, 2020; **Accepted:** October 6, 2020; **Published:** October 17, 2020

Abstract: Across the globe, kidney cancer and other cancerous diseases has been a threat to human lives. The incidence and mortality rate represent a significant and growing threat to both developed and developing countries especially in Africa, where most cancers are diagnosed at an advanced stage. This typically contributes to its complications and high rate of mortality, and has been attributed to limited awareness of early signs and symptoms of the disease, lack of detective mechanism and inaccessible cancer care in our health care centres. To preclude the harm and mortality caused by the disease, an intelligent mechanism for early prediction and prognosis of the syndrome is vital. However, early detection and prognosis requires an accurate information and analytic procedure that will assist and equip the health-care providers/public with the skills to identify early the indicators of the disease. Efforts in this work, produced a model for early prediction of kidney cancer using data analytic approach. Dataset and reports pertaining to the disease were acquired from selected private and public hospitals in fifty-two (52) selected LGA in Nigeria. A two-layered classifier system consisting of Artificial Neural Networks (ANN) and Decision Tree (DT) designed for the work was successfully employed in the model building. Waikato Environment for Knowledge Analysis (WEKA) platform was used for the experiment. The performance of the classifiers considered was compared using standard metrics of accuracy and time taken as benchmark. Experimental results show that the J48 decision tree algorithm outperform all other algorithms in the classifier family with correctly classified instances of 74.7%, F-Measure of 0.614, TP rate of 0.747, FP rate of 0.135, precision and recall of 0.687 and 0.714 respectively. It took the algorithm, 0.03 seconds to build the model. The performance of this algorithm proved its suitability as a valuable tool for the research purpose. The model will in no small measure support the efforts of the national health scheme in preventing the disease mortality rate.

Keywords: Data Analytics, Kidney Cancer, Classifier System, Algorithm, GLOBOCAN

1. Introduction

The global cancer burden as released by the International Agency for Research on Cancer (IARC) in 2018 is estimated to have risen to 18.1 million new cases and 9.6 million deaths. It is projected that one-in-five men and one-in-six women worldwide will develop cancer over the course of their lifetime, and that one-in-eight men and one-in-eleven women will die from the disease [1, 2]. Additionally, it is predicted that more new cases of cancer will occur every year and by year 2030, 70% of the world's cancer burden will be from the poor countries [3]. A number of factors appear to be driving this increase, particularly a growing and ageing global population couples with exposure to cancer risk causes linked

to social and economic development. The American Society of Clinical Oncology (ASCO) 2019 through Cancer. Net listed over 120 types of cancer and related hereditary syndromes [4]. For the purpose of this work, the main focus is on kidney cancer, the malady is rated the twelfth most common cancer in the world, the ninth most commonly occurring cancer in men and the 14th most occurring cancer in women with over 400,000 new cases in 2018 [5].

Largely, cancer is a disease in which cells in the body grow out of control, and forming a tumor [6]. When the disease starts in the kidney, it is called kidney or renal pelvis cancer or renal cell cancer. Almost all kidney cancers first appear in the

lining of tiny tubes (tubules) in the kidney. If found early it can be treated before it spread (metastasize) to other distant organs. According to Lasebikan et al., (2014), cancers caught early are easier to treat successfully, but the consequences of delayed or inaccessible cancer care are lower likelihood of survival, greater morbidity of treatment and higher costs of care [7]. Although some tumors may grow to be quite large before they are detected, however in Africa most cancers are diagnosed at an advanced stage of the disease which usually contributes to its complications and high mortality rate. These challenges has been attributed to lack of detective mechanism, limited

awareness of early signs and symptoms of the disease and inaccessible cancer care in our healthcare centres. Meanwhile, an effective way to reduce the harm and mortality caused by any disease is to detect it early [8].

However, early detection and prognosis requires an accurate information and reliable analytic procedure that will enable physicians to detect the disease early. This is what this research work is set to do. The representation in the figures 1-3 show the global case counts of kidney cancer per 100,000 people by age group, gender and ethnicity between 2013 and 2017, as sourced from Centre for Disease and Control (CDC).

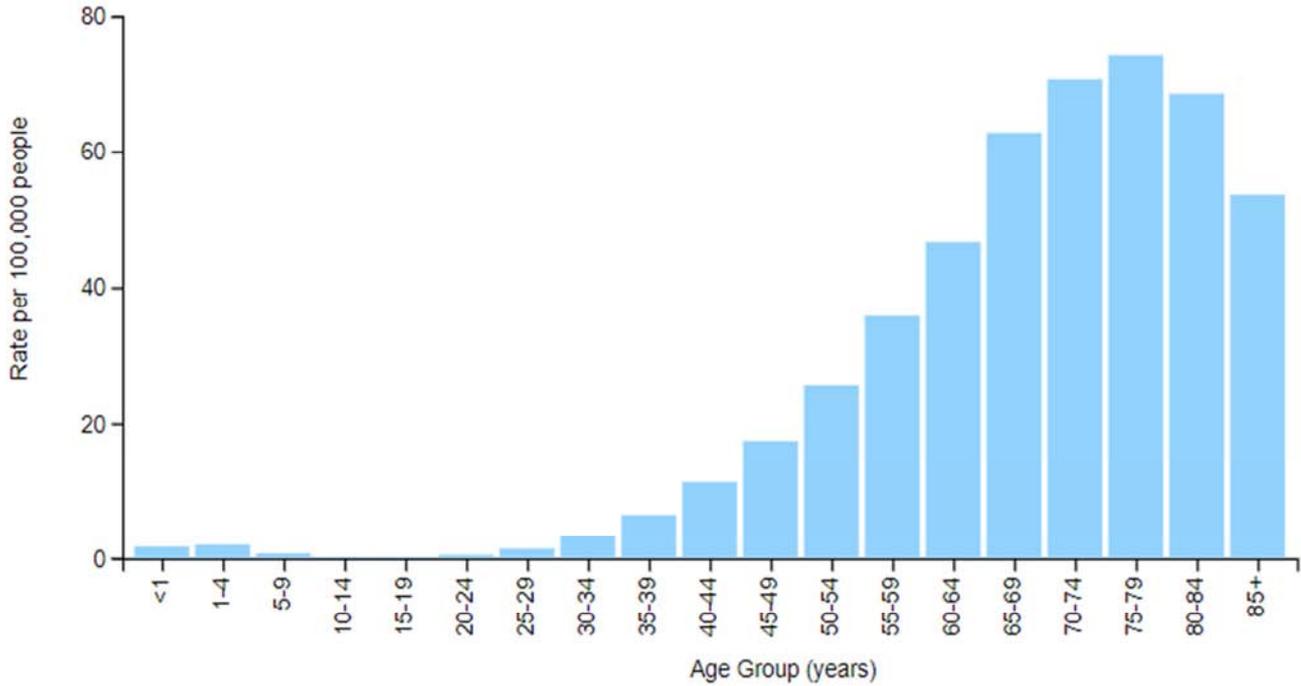


Figure 1. Case count of kidney cancer per 100,000 people by age group between 2013 and 2017.

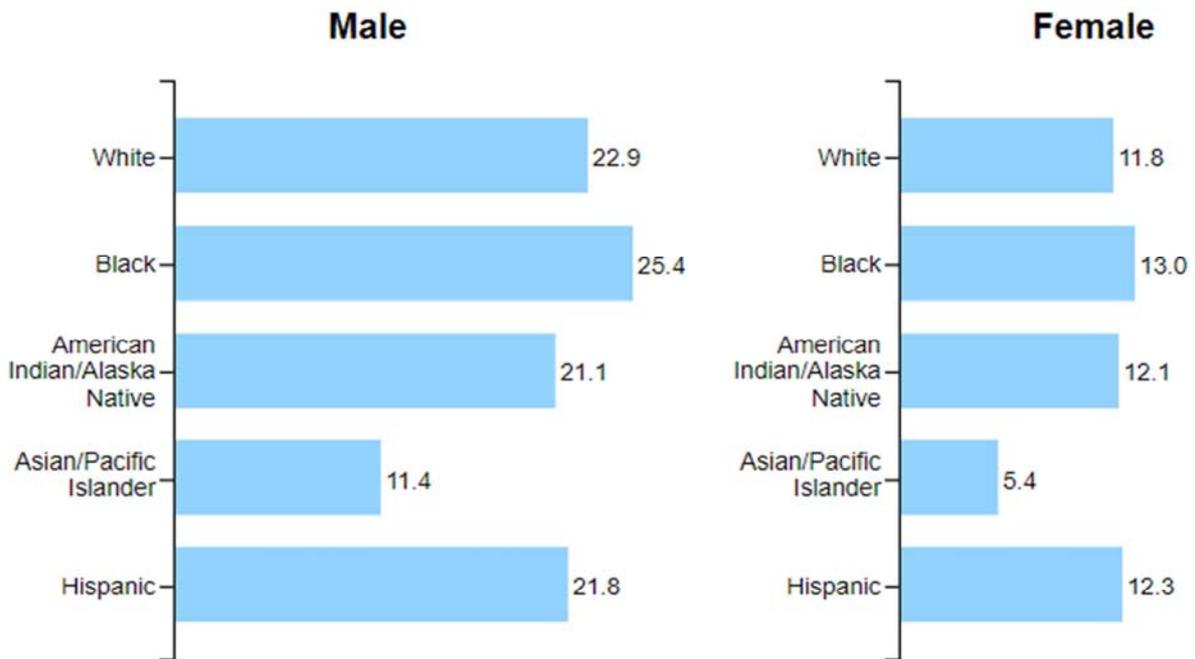


Figure 2. Case count of kidney cancer per 100,000 people by gender between 2013 and 2017.

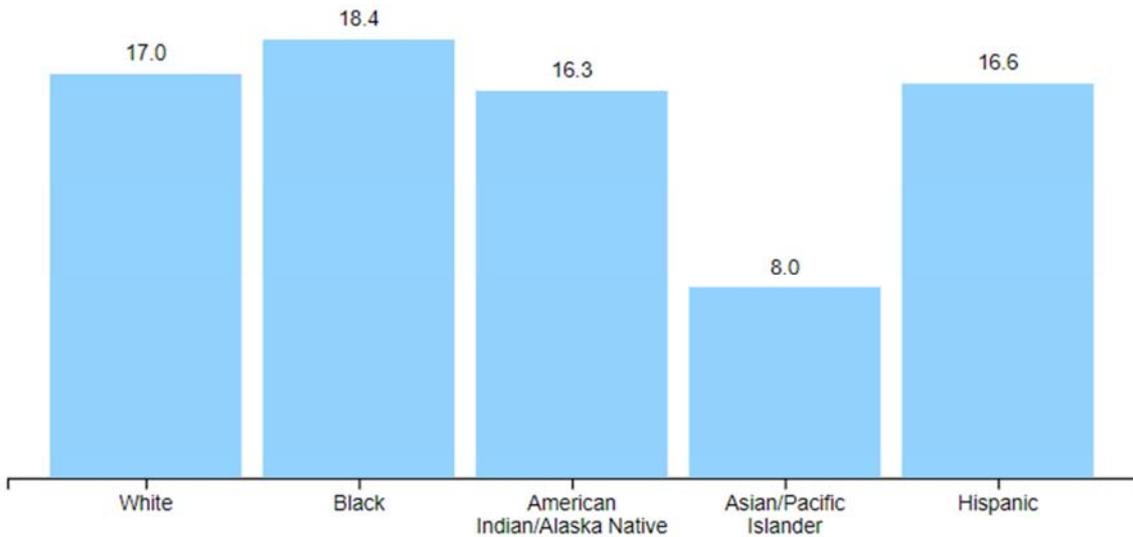


Figure 3. Case count of kidney cancer per 100,000 people by ethnicity between 2013 and 2017.

In figure 1, by analysis age 75-79 stands to suffer the highest risk of case counts, while the black suffered the highest case counts amongst gender in figure 2 and

race/ethnicity in figure 3 respectively. Death rate by age group and gender are presented in Figure 4 and Figure 5 respectively.

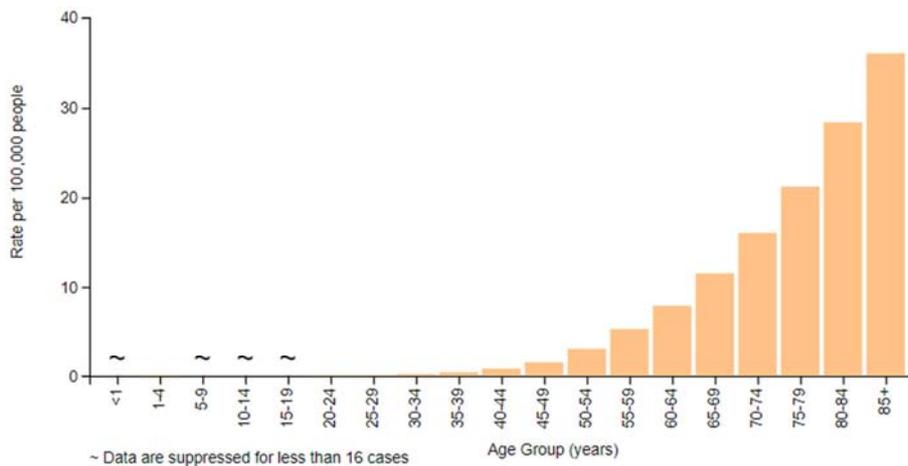


Figure 4. Death rate of kidney cancer per 100,000 people by age group between 2013 and 2017.

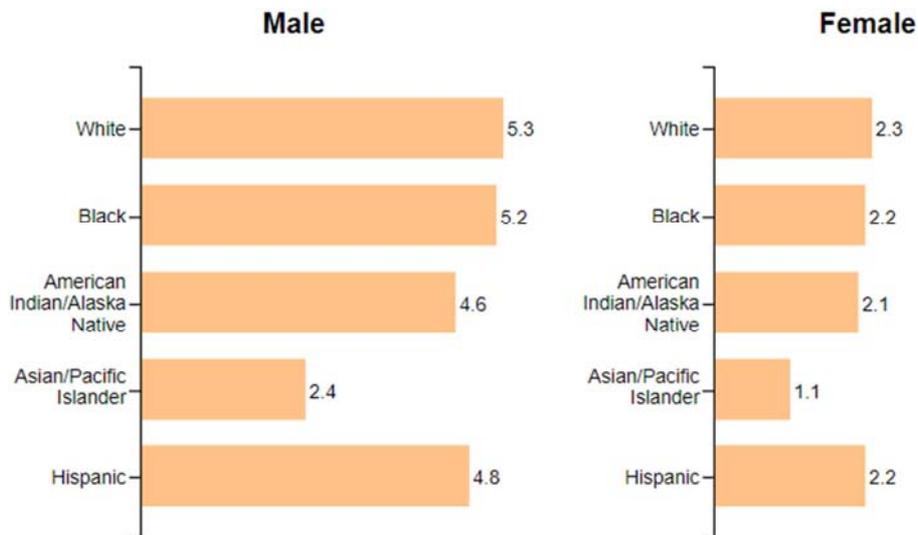
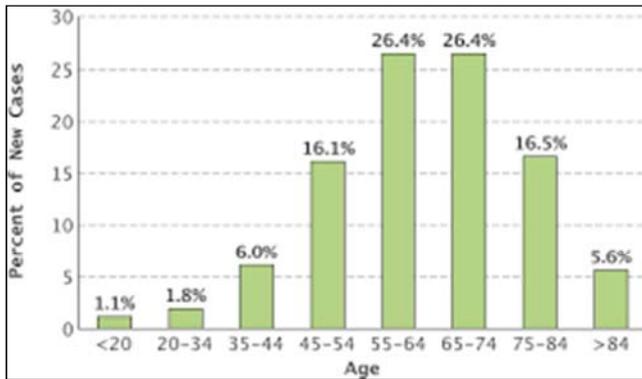


Figure 5. Death rate of kidney cancer per 100,000 people by gender between 2013 and 2017.

The percentage of new cases of the disease by age in Africa by the report of the National Cancer Registry (NCR) in 2010 and 2014 as presented in figure 6 showed that in most part of Africa, there had been a high incidence of kidney cancer between the ages of 55 and 74 years both in males and females [9, 10].



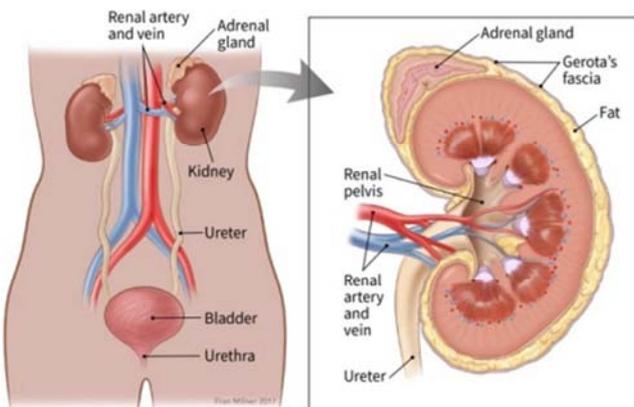
Source: [9]

Figure 6. Percentage of new cases by age in Africa.

2. Literature Review

2.1. Human Kidney Anatomy

The human body has two kidneys. They are a pair of bean-shaped organs, each about the size of a fist. They are attached to the upper back wall of the abdomen and protected by the lower rib cage. One kidney is just to the left and the other just to the right of the backbone [11]. The upper and lower portions of each kidney are sometimes called the superior pole and inferior pole. A small organ called an adrenal gland sits on top of each kidney and each kidney and adrenal gland is surrounded by fat and a thin, fibrous layer known as Gerota’s fascia [12].



Source: [12]

Figure 7. Kidney and Urinary Anatomy.

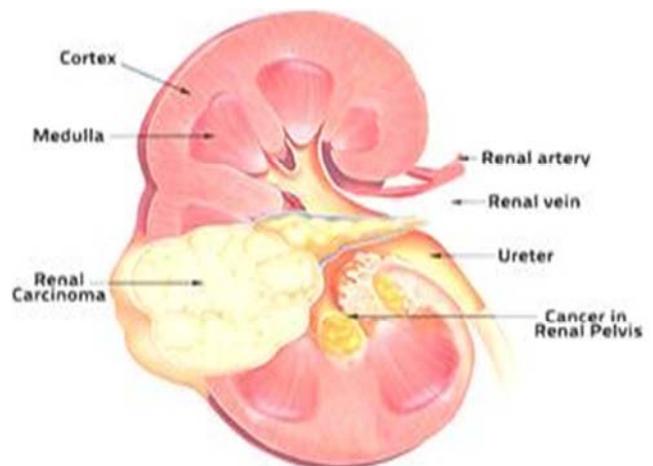
The main function of the kidney is to remove excess water, salt, and waste products from blood coming in from the renal arteries. These substances become urine, the urine collects in

the center of each kidney in an area called the renal pelvis leaves the kidneys through long slender tubes called ureters. The ureters lead to the bladder, where the urine is stored until when one urinates. Every day, the kidneys filter about 200 quarts of blood to generate 2 quarts of urine. Other function of the kidney include, the production of hormones called renin that help in controlling the body blood pressure and hormone called erythropoietin that help in the production of red blood cells by the bone marrow [4]. Figure 7 depicts the kidney and urinary anatomy.

Normally, most people have two (2) kidneys, and each kidney works independently. Some people do not have working kidneys at all, and survive with the help of a medical procedure called dialysis [13]. Dialysis is a mechanized filtering process that filters blood much like a real kidney would. It could be done through the blood, called hemodialysis, or using the patient’s abdominal cavity, called peritoneal dialysis.

2.2. Types of Kidney Cancer

The two most common types of kidney cancer are renal cell carcinoma (RCC) accounting for over 90% of cases and transitional cell carcinoma (TCC). This is also known as urothelial cell carcinomas (UCC) of the renal pelvis [11]. These two kidney cancers (RCC and UCC) develops in different way, they have different long term outcomes, and need to be staged and treated in different ways. RCC is responsible for approximately 80% of primary renal cancers, and UCC accounts for the majority of the remainder [12]. In RCC, cancerous (malignant) cells develop in the lining of the kidney tubules and grow into a mass called a tumor. Like many other cancers, the growth begins small and becomes larger over time. RCC typically grows as a single mass. But, there are cases where a kidney may contain more than one tumor. Hence, it is possible that tumors are found in both kidneys at the same time. Figure 8 shows a typical cancer of the kidney in human body.



Source: [14]

Figure 8. Typical Cancer of the Kidney in Human Body.

2.3. Symptoms and Risk Factors of Kidney Cancer

According to Centers for Disease Control and Prevention report of 2017, cancer impacts people of all ages, races, ethnicities, and sexes, but does not always affect them equally. Differences in genetics, hormones, environmental exposures, and other factors can lead to differences in risk among different groups of people. For most cancers, age increase is the most important risk factor [13]. By description, risk factor is anything that increases a person's chance of developing a disorder. In the case of cancer, different cancers have different risk factors. Although, most of them do not directly cause cancer as many people with several risk factors never develop cancer, while others with no known risk factors do. Knowing your risk factors and talking about them with your doctor may help you make more informed lifestyle and health care choices [6].

American Cancer Society's Report of 2019 listed the following factors that may raise a person's risk of developing kidney cancer:

1. Smoking. Smoking tobacco doubles the risk of developing kidney cancer. It is believed to cause about 30% of kidney cancers in men and about 25% in women.
2. Gender. Men are 2 to 3 times more likely to develop kidney cancer than women.
3. Age. Kidney cancer is typically found in adults and is usually diagnosed between the ages of 50 and 70.
4. Nutrition and weight. Research has often shown a link between kidney cancer and obesity.
5. High blood pressure. Men with high blood pressure, also called hypertension, may be more likely to develop kidney cancer.
6. Overuse of certain medications. Painkillers containing phenacetin have been banned in the United States since 1983 because of their link to transitional cell carcinoma. Diuretics and analgesic pain pills, such as aspirin, acetaminophen, and ibuprofen, have also been linked to kidney cancer.
7. Exposure to cadmium. Studies have shown a connection between exposure to the metallic element cadmium and kidney cancer. Working with batteries, paints, or welding materials may increase a person's risk as well. This risk is even higher for smokers who have been exposed to cadmium.
8. Family history of kidney cancer. People who have a first-degree relative with kidney cancer, such as a parent, brother, sister, or relatives may likely develop kidney cancer in the course of their lifetime.

Some of the earliest signs and symptoms of kidney cancer include: Haematuria (blood in the urine haematuria), low back pain on one side of the body (not caused by injury), a mass (lump) on the side or lower back, weight loss not caused by dieting, fever that is not caused by an infection and that does not go away after weeks, swelling of the ankles and legs (oedema). Although, none of these symptoms are singly or positively

indicative of kidney cancer. For example, blood in the urine may be a sign of kidney, bladder or prostate cancer, but can also be an indication of a bladder infection or a kidney stone [12].

Detecting these symptoms on symptomatic patients as early as possible generally can increase the chances for successful treatment, while the consequences of delayed and inaccessible cancer care can be costly [3]. In order to enable medical practitioners to gain valuable knowledge for early detection and proactive intervention, this work presents a model for early prediction of kidney cancer using data analytic approach to help overcome common barriers to timely diagnosis.

2.4. Overview of Data Analytics

Data analytic is the science of analyzing raw data with the goal of discovering useful information and drawing conclusions from it to support decision-making [15]. The technique has proven to be a multi-dimensional discipline that uses descriptive techniques and predictive models to uncover patterns in raw data in order to gain valuable knowledge for recommendations and decision making [16]. The technique is generally divided into descriptive statistics, exploratory data analysis (EDA), and confirmatory data analysis (CDA). EDA focuses on discovering new features in the data while CDA focuses on confirming or falsifying existing hypotheses. Data Analytics is commonly apply to business data, health, marketing mix modeling, web analysis, risk analysis, fraud analysis to communicate insights from data warehouse. The technology basically aggregates three major stages its processes: The first stage is the data acquisition, the second stage is the data transformation, while the third stage is the model building, that is responsible for obtaining knowledge through appropriate classifiers.

3. Research Method and Materials

This section describes the methodology used to achieve the objectives of the research work. It presents the data analytic model proposed and the two-layered classifier system design for the work. The method of data collection and population of the study with sampling techniques and appropriate data representation was discussed. Figure 9 and Figure 10 presents the proposed model and the two-layered classifier system designed for the work respectively.

The data analytic model aggregates three major processing phases namely: pre-processing, processing and post-processing. The pre-processing phase is an important step in data mining and machine learning processes that includes data collection, filtering, normalization and transformation. The processing phase is broadly, the manipulation of the datasets to produce meaningful information. It involves choosing algorithms, parameters for model building and evaluation of the model, while the post processing involves knowledge interpretation and representation.

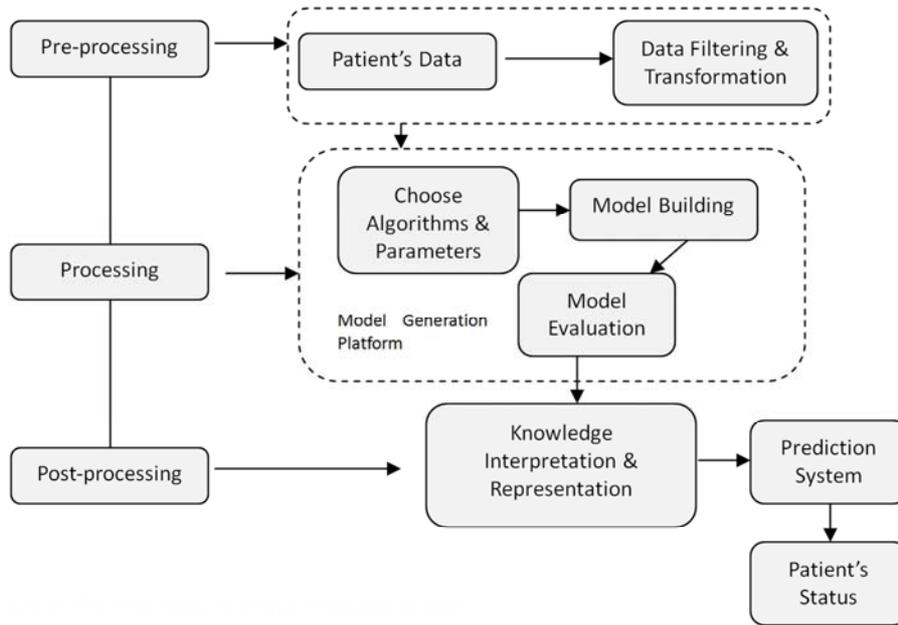


Figure 9. The Architecture of the Proposed Model.

The two-layered classifier system designed for the model building as presented in Figure 10 consists of two classification algorithms, Artificial Neural Networks (ANN) in layer 1 and Decision Tree (DT) in Layer 2. These families of classifiers have been selected because of their performances in various classification tasks. The ANN was first used to

classify the risk factors into distinctive groups. The data from the elements in these groups was later subjected to the classifier in in Layer 2 for the model building. Seven (7) different classification algorithms from this family were used to model the dataset.

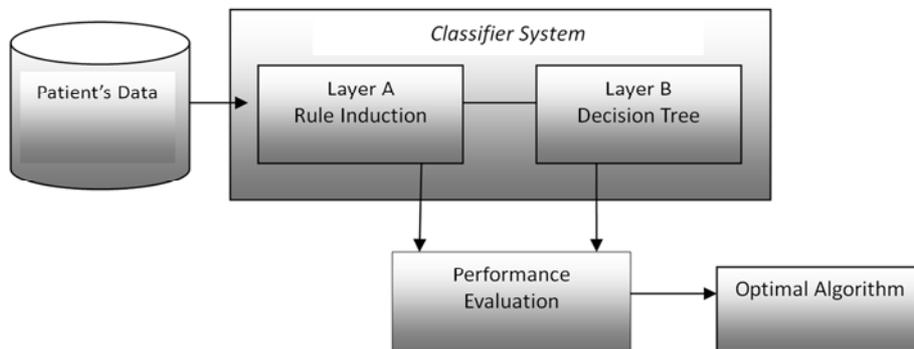


Figure 10. The Architecture of the Two-layered Classifier System.

3.1. Data Collection and Analysis

Dataset pertaining to this research work was collected from selected health centres and hospitals in fifty-two (52) selected Local Governments in Nigeria using purposive and selective sampling techniques. Sampled data totaling, 1,006 records was collected. The data collected were cleaned, normalized and organized in a form suitable for data analytic process. Table 1 and Table 2 show the data format for the data collection and statistical data for the selected attributes respectively. Out of the 1,006 patient's data captured, 55.2% are male while the remaining 44.8% are female. The data analysis on patient's lifestyle revealed that 39.5% of the patients are addicted to drugs, 29.3% to alcoholic and 13.3% are smokers. 23% are found to have been exposed to chemical and industrial contents while 10.6% of the population has

gender and hereditary disorder. Many of the patients are used to regular use of non-steroidal anti-inflammatory drug (NSAIDs) such as ibuprofen and naproxen, which usually doubles the risk of the disease. Other factors considered are obesity; faulty genes; a family history of kidney cancer, dialysis, infection with hepatitis C; and previous treatment for testicular cancer and cervical cancer.

Table 1. Shows the data format for the research data collection.

S/N	Variable Name	Variable Format	Variable Type
1	Gender	Male, Female	Categorical
2	Age	25, 30,..	Numerical
3	Lifestyle	Smoking, Obesity,	Categorical
4	G&H Disorder	Yes, No	Categorical
5	C & I Exposure	Yes No	Categorical
6	Prediction Level	One, Two, Three	Categorical

Table 2. Statistical data for selected attributes.

S/N	Attributes	Data	Percentage (%)
1	Gender		
	Male	556	55.2
	Female	451	44.8
2	Lifestyle		
	Smoking	134	13.3
	Obesity	19	1.9
	Drug Abuse	397	39.5
	HB Pressure	106	10.53
	Water Pills	40	3.98
	Dialysis	8	0.8
	Alcohol	295	29.3
	Radiation	7	0.69
3	G&H Disorder		
	Yes	106	10.6
	No	900	90.0
4	C&I Exposure		
	Yes	230	23.0
	No	776	77.6
5	Complaints		
	Blood in Urine	113	11.23
	Back pain	203	20.17
	Tumor	189	18.8
	Fibroid	131	13.02
	Stomach Ucher	144	14.31
	Kidney pain	159	15.8
6	Abdominal pain	67	6.67
	Age Group		
	20-30	38	3.8

S/N	Attributes	Data	Percentage (%)
	31-40	150	15.0
	41-50	231	23.0
	51-60	240	23.9
	61-70	211	21.0
	70-80	94	9.13
	81-90	42	4.17
	91-100	0	0

3.3. Model Building Procedure and Experimental Results

Waikato Environment for Knowledge Analysis (WEKA) version 3.8.4 (2019) platform was used for the data modeling. Seven (7) different classification algorithms from the family of Decision Tree were employed to model the dataset for the work. The datasets was first divided into two, 66% of the datasets was devoted to training while the remaining 34% was used for testing. 10-fold cross validation test and percentage split test options were also considered in the modeling. The performance comparison of the classifiers was conducted in order to determine the algorithm that models the data with best predictive accuracy. The performance analysis is presented in Table 3. Table 4 shows the standard metric accuracy details for the 10-fold cross validation mode option (the best mode) for all the algorithms in the experiment. Figure 11 and Figure 12 depict the graphs of predictive accuracy and time taken to build the model by the classifiers respectively.

Table 3. Classification Accuracy Comparison between Hold-out and 10-fold Cross Validations.

S/N	Classifiers	10-fold Cross Validation		Hold-out (Percentage Split)	
		Correctly Classified Instances	Time taken to build model	Correctly Classified Instances	Time taken to build model
1	J48 Decision Tree	74.7	0.03	74.5	0.02
2	LMT	74.6	29.25	73.7	29.03
3	M5P Algorithm	72.6	0.92	73.1	0.91
4	RepTree	71.6	2.54	72.4	2.4
5	Hoeffding Tree	70.9	0.03	70.1	0.03
6	Decision Stump	70.8	0.02	71.8	0.03
7	Random Forest	70.2	0.03	70.3	0.03

Table 4. Standard metric accuracy details for all the Classification Algorithms.

S/N	Algorithms	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Built Time (s)	Correctly classified %
1	J48 Decision Tree	0.747	0.135	0.687	0.714	0.614	0.78	0.03	74.7
2	LMT	0.746	0.239	0.73	0.746	0.733	0.863	29.25	74.6
3	M5P Algorithm	0.731	0.292	0.714	0.731	0.702	0.85	0.91	73.1
4	RepTree	0.716	0.548	0.536	0.658	0.533	0.571	0.03	71.6
5	Hoeffding Tree	0.709	0.274	0.728	0.749	0.731	0.754	0.06	70.9
6	Decision Stump	0.718	0.294	0.694	0.718	0.695	0.814	0.03	71.8
7	Random Forest	0.704	0.238	0.716	0.704	0.702	0.816	0.05	70.4

The experimental results shows that the J48 decision tree outperform all other algorithms in the layer with correctly classified instances of 74.7%, ROC Area of 0.78 and recall of 0.714 respectively. It has a lower FP rate of 0.153, F-Measure of 0.614 and took lesser time of 0.03 seconds to build the

model compared to LMT and other classifiers as shown in Table 4. J48 decision tree algorithms generally have this ability that can produce a simple tree structure with high accuracy in term of classification rate, even with huge volume of data [15]. Pruning methods have been introduced to reduce

the complexity of tree structure without any decrease in classification accuracy.

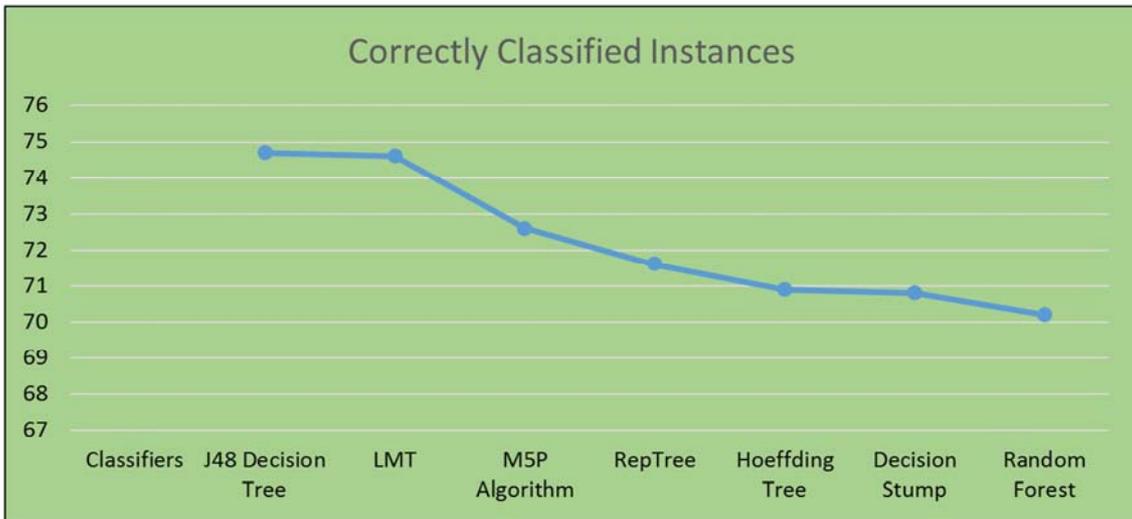


Figure 11. Predictive Accuracy of Classifiers.

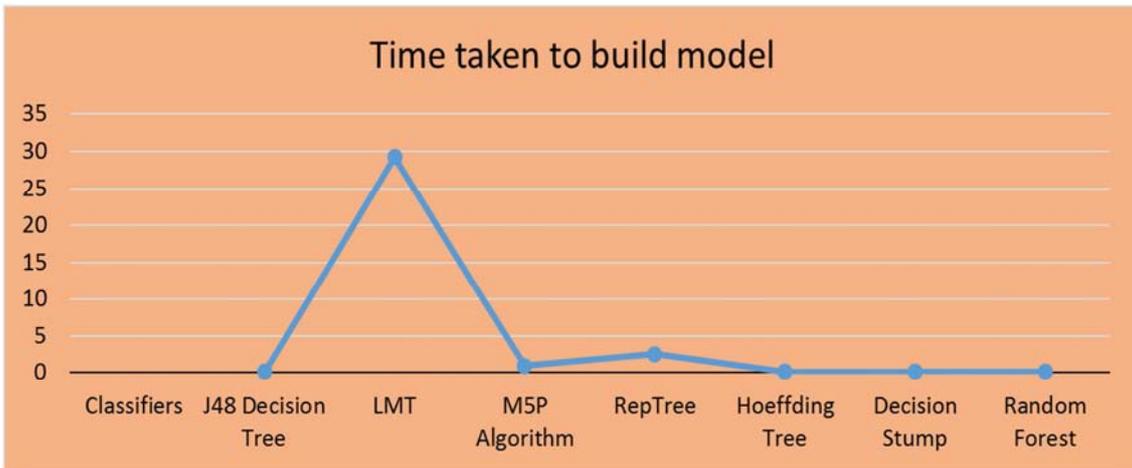


Figure 12. Time taken by the classifier to build the Model.

3.4. Rules Generation

The rules generated from the best algorithm (J48 pruned decision tree) are stated in rules 1 – 20, while the prediction levels are categorized as follows: (PL) – One, Two and Three respectively. The prediction levels showed the status of patients. Level One and Two indicates a high risk level of the disease manifestation in the patients that needs to be attended to urgently. While, level three indicates that the patient is not manifesting any symptoms of kidney cancer disease, but may suffer from other diseases. A back-end for updating the rules as the situation arises will be incorporated into the system to match other conditions.

Rule 1: IF (G&H Disorder=NO) AND (C&I Exposure=Yes) AND (Lifestyle=Smoking) AND Complaints=blood in urine: PL=One

Rule 2: IF (G&H Disorder=NO) AND (C&I Exposure=Yes) AND (Lifestyle=Smoking) AND Complaints=back pain: PL=Two

Rule 3: IF (G&H Disorder=NO) AND (C&I Exposure=Yes)

AND (Lifestyle=Smoking) AND Complaints=tumor: PL=Three

Rule 4: IF (G&H Disorder=NO) AND (C&I Exposure=Yes) AND (Lifestyle=Smoking) AND Complaints=Fibroids: PL=Three

Rule 5: IF (G&H Disorder=NO) AND (C&I Exposure=Yes) AND (Lifestyle=Smoking) AND Complaints=Stomach ucher : PL=Two

Rule 6: IF (G&H Disorder=NO) AND (C&I Exposure=Yes) AND (Lifestyle=Smoking) AND Complaints=Kidney pain: One

Rule 7 IF (G&H Disorder=NO) AND (C&I Exposure=Yes) AND (Lifestyle=Smoking) AND Complaints=Abdominal pain: Two

Rule 8 IF (G&H Disorder=YES) AND (C&I Exposure=Yes) AND (Lifestyle=Smoking) AND Complaints=blood in urine: PL=One

Rule 9 IF (G&H Disorder=YES) AND (C&I Exposure=Yes) AND (Lifestyle=Obesity) AND Complaints=blood in urine: PL=Two

Rule 10 IF (G&H Disorder=YES) AND (C&I Exposure=Yes) AND (Lifestyle=HB Pressure) AND Complaints=blood in urine: PL=Two

Rule 11 IF (G&H Disorder=YES) AND (C&I Exposure=Yes) AND (Lifestyle=Smoking) AND Complaints=Drug Abuse OR Tumor OR Fibroids: PL=Two

Rule 12 IF (G&H Disorder=YES) AND (C&I Exposure=Yes) AND (Lifestyle=Smoking) AND Complaints=Abdominal pain: PL=Two

Rule 13 IF (G&H Disorder=YES) AND (C&I Exposure=Yes) AND (Lifestyle=Smoking) AND Complaints=Kidney pain: PL=One

Rule 14 IF (G&H Disorder=YES) AND (C&I Exposure=Yes) AND (Lifestyle=Smoking) AND Complaints=stomach ucher: PL=One

Rule 15 IF (G&H Disorder=YES) AND (C&I Exposure=Yes) AND (Lifestyle=Alcohol OR Dialysis) AND Complaints=stomach ucher: PL=Two

Rule 16 IF (G&H Disorder=YES) AND (C&I Exposure=Yes) AND (Lifestyle=Radiation) AND Complaints=stomach ucher OR blood in urine: PL=One

Rule 17 IF (G&H Disorder=YES) AND (C&I Exposure=Yes) AND (Lifestyle=Water pills) AND Complaints=stomach ucher: PL=Three

Rule 18 IF (G&H Disorder=YES) AND (C&I Exposure=NO) AND (Lifestyle=Smoking) AND Complaints=stomach ucher OR kidney pain: P=One

Rule 19 IF (G&H Disorder=YES) AND (C&I Exposure=NO) AND (Lifestyle=Smoking) AND Complaints=stomach ucher: PL=Two

Rule 20 IF (G&H Disorder=YES) AND (C&I Exposure=NO) AND (Lifestyle=Smoking OR Obesity OR Drug Abuse OR Radiation OR Water Pills OR Dialysis) AND Complaints=stomach ucher: PL=Three

4. Conclusion

This work has successfully offered a model for apt prediction of kidney cancer with suitable algorithm based on data analytic technology used. The theoretical foundation for the research work was established and the research chosen design and methodology was amply discussed. The two-layered classifier system developed for the work was tested using the dataset acquired from selected hospitals in fifty two (52) different Local Governments in Nigeria. Waikato Environment for Knowledge Analysis (WEKA) platform was deployed to build the data model and seven (7) different machine learning algorithms were considered during modeling in search for the algorithm that produced the model with the best predictive accuracy. The performance of the various algorithms was compared based on standard metrics of accuracy and results presented appropriately. The experimental results show that the J48 decision tree outperform all other algorithms in the layer with correctly classified instances. The output of the research work if properly implemented will in no small measure help health practioners to detect kidney cancer early. The

predictive model will not only assist in improving the efficiency of our health care delivery system in this domain, but will also compliments the effort of the national health programmes on kidney cancer in Nigeria and other part of sub-sahara Africa.

References

- [1] Bray, F, Ferlay, J, Soerjomataram, I, Siegel, R. L, Torre, L. A and Jemal, A (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *C A Cancer Journal Clinical Oncology*. 68 (6): 394-424. doi: 10.3322/caac.21492. Epub 2018 Sep 12. PMID: 30207593.
- [2] Ferlay, J. Colombet, M. Soerjomataram, I. Mathers, C. Parkin, D. M. Piñeros, M. Znaor A. and Bray, F. (2019). Estimating the global cancer incidence and mortality. *International Journal of Cancer*. 144 (1), 1941–1953.
- [3] IARC-WHO (2018). Latest global cancer data: Press Release of 12th September, 2018. <https://www.who.int/cancer/PRGloboCANFinal.pdf?ua=1>.
- [4] The American Society of Clinical Oncology (ASCO)(2019). Kidney Cancer. Retrived on 19/08/2020 from <https://www.cancer.net/cancer-types/kidney-cancer/introduction>.
- [5] American Institute for Cancer Research (2018). Kidney cancer Statistics. Retrieved August 2020 from <https://www.wcrf.org/dietandcancer/cancer-trends/kidney-cancer-statistics>.
- [6] WebMD (2017). Understanding Kidney Cancer retrieved on 24/08/2020 from <https://www.webmd.com/cancer/understanding-kidney-cancer#1>.
- [7] Lasebikan, O. A, Nwadinigwe, C. U and Onyegbule, E. C (2014). Pattern of bone tumours seen in a regional orthopaedic hospital in Nigeria. *Niger J Med*. 2014 Jan-Mar; 23 (1): 46-50.
- [8] Kushi LH, Doyle C, McCullough M, et al. (2012). American Cancer Society Guidelines on nutrition and physical activity for cancer prevention.
- [9] Nigerian National Systems of Cancer Registries (NSCR), (2014). Newsletter of the Nigerian National Systems of Cancer Registries- 2014.
- [10] Jemal, A, Bray, F., Forman, D., O'Brien, M., Center, M., Parkin, D. M (2012). Cancer burden in Africa and opportunities for prevention. *American Cancer Society* 118 (18): 4372-84. DOI: 10.1002/cncr.27410.
- [11] Michael, C. H (2014). Fact Sheet on Kidney Cancer, Cancer Association of South Africa (CANSa) Cansa. Available online at: <http://www.cansa.org.za/files/2014/06/Fact-Sheet-Kidney-Cancer-June-2014.pdf>.
- [12] American Cancer Society, (2014). What is kidney cancer? Retrieved 24/8/2020 from <https://www.cancer.org/cancer/kidney-cancer/about/what-is-kidney-cancer.html>.

- [13] Centre for Disease Control and Prevention CDC (2020). Cancer Statistics Data Visualizations Tool www.cdc.gov/cancer/dataviz, released in June 2020.
- [14] UCLA, (2011). University of California, Los Angeles Kidney Cancer Program. <https://www.uclahealth.org/urology/kidney-cancer-program>.
- [15] Xia, B. S., & Gong, P. (2015). Review of business intelligence through data analysis. *Benchmarking*, 21 (2), 300-311. doi: 10.1108/BIJ-08-2012-0050.
- [16] Kohavi, R., Rothleder, N. J', & Simoudis, A. P (2002). Emerging Trends in Business Analytics Published by ACM Volume 45 Issue 8, Pages 45-48 August 2002.