ISSN No:-2456-2165

Performance Analysis of Selected Machine Learning Algorithms for Prediction of Mortality and Survival Chances of Viral Hepatitis and Hepatocellular Carcinoma Patients

Mba Obasi Odim , Uchechukwu Frederick Ekpendu, Bosede Oyenike Oguntunde, Adeniyi Samson Onanaye
Departemnt of Computer Sceince Redeemer's University Ede, Nigeria

Abstract:- Investigating the mortality/survival chances of Viral Hepatitis and Hepatocellular Carcinoma (HCC) patients could provide informed knowledge for planning and implementation of efficient and effective strategies for curtailing the mortality rate of the disease and at the same time providing more information about the relationship between HBV/HCV and HCC. This study, modelled and assessed the performance of some selected machine learning algorithms (Artificial Neural Networks (ANN), Decision Tree, K-nearest neighbours (K-NN) Logistic Regression, Naïve Bayes, Random Forest and Support Vector Machine (SVM) for the prediction of the mortality/survival chances of HCC and Hepatitis patients. The data were collective from UCI machine learning repository, consisted of clinical test result of 155 hepatitis patients of 20 attributes with 123 survived patient and 32 mortalities. There were 13 instances with missing values, which was removed while cleaning the dataset leaving 142 instances with 116 survivors' class and 26 death class. The HCC dataset contained 165 instances with 50 attributes, 102 survivals and 63 death instances. The algorithms were deployed within the WEKA environment and the findings revealed that the Support Vector Machine recorded the highest classification performance on the both datasets. This was followed respectively by the Naïve Bayes on the Hepatitis and the Random Forest on the Hepatocellular carcinoma. The Decision Tree recorded the least accuracies on both datasets. The result therefore suggests that the Support Vector machine, could be a most appropriate algorithm for developing a classification system for survival of Hepatitis and Hepatocellular carcinoma. Hoverer, the performance of these algorithms could as well be improved with more dataset.

Keywords:- *Machine Learning, Viral Hepatitis, Hepatocellular Carcinoma, Patients, Survival chances.*

I. INTRODUCTION

Hepatocellular carcinoma (HCC) is a common cancer that is globally threatening the lives of many people around the world, with yearly growingoccurrence[1].Ithas been reported to be one of the major cancers and the third most deadly in the world [2]. The HCC prevalence is continuouslygrowing withmore than 900 thousand freshincidencesand as numerous deaths documentedglobally in 2020 [3]. HCC rates differlargely across the globe with the highest rates conveyed in Southeast Asia, East Asia, and sub-Saharan Africa.It was estimated that almost 80% of the illness and death due to HCC wasassociated with developing countries[4].Hepatitis B virus (HBV) contagion and hepatitis C virus (HCV) contagionhave been reported as essential causes of HCC[1]. They have, indeed, been credited as the major instigating or causal agents to the development of HCC [5], with HBV contributing more than 50% and HBV adding to more than 25%. In addition, Hepatitis D Virus (HDV) and occult hepatitis B in the development of HCC have also been reported[6][7]. It has been also reported that Hepatitis B virus (HBV)/ Hepatitis C virus (HCV) coinfections and HBV, HCV, and Hepatitis virus D (HDV) infections are connected with an increased risk of developing HCC[8].

The risk of HCC can be actively addressed by taking necessary measures. Effective antiviral treatment of HBV/HCV has the potentials of mitigating their degenerating to HCC [1]. Investigating the mortality/survival chances of HCC, HBV and HCV could provide informed knowledge for planning and implementation of efficient and effective strategies for curtailing the mortality rate of the disease and at the same time providing more information about the relationship between HBV/HCV and HCC. Thus, this study, modelled and assessed the performance of some selected machine learning algorithms for the prediction of the mortality/survival chances of HCC, HBV and HCV.

ISSN No:-2456-2165

II. RELATED WORK

Effective dictation and prevention strategies in reducing the risk of HBV/HCV and HCC is an active research area. A review of related work is hereby presented in this section.

In [9], ten Machine learning algorithms, namely, the K-Nearest Neighbor (KNN), Logistic Regression (Reglog), Naive Bayes Classifier (NB), Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Random Forest (RF), Multilayer Perceptron (MLP), SVM (nu-SVC) and Linear SVM), were studied for the detection of HCC with 165 sampled patients.Parameter optimization and feature selection were based on the genetic algorithm combined with stratified 5-fold cross-validation method. The findings showed that SVM (type C-SVC) with new 2level genetic optimizer (genetic training) and feature selection recorded 0.8849 accuracy and 0.8762F1-Score.

An optimal method for predicting HBsAg seroclearance was investigated in [10]. The dataset was composed of 2,235 patients with CHBcollected from the South China HepatitisMonitoring and Administration (SCHEMA) cohort. There were 106 patients with HBsAg seroclearance. Four algorithms, consistingof the extreme gradient boosting (XGBoost), random forest (RF), decision tree (DCT), andlogistic regression (LR) were used to develop the models. The area under the receiver operating characteristic curve (AUC) was employed to determine the optimal model. The findings showed AUCs of 0.891, 0.829, 0.619, and 0.680, respectively for XGBoost, RF, DCT, and LR models, with XGBoostshowing superiority in performance.

An evaluation of recurrent neural networks and regression models was performed in[11]. It examined whether deep learning recurrent neural network (RNN) models that use raw longitudinal data obtained directly from electronic health records outperform conventional regression models in predicting the risk of developing hepatocellular carcinoma (HCC). The study considered 48,151 patients with hepatitis C virus (HCV)–related cirrhosis in the national Veterans Health Administration who have been managing the disease for at least 3 years after the diagnosis. Patients with at least one positive HCV RNA tested between January 1, 2000, to January 1, 2016 were used and were monitored from the diagnosis of cirrhosis to January 1, 2019, for the development of incident HCC. Three predictive models were formulated and compared during the 3-year period; Logistic Regression with cross-sectional inputs (cross-sectional LR); LR with longitudinal inputs (longitudinal LR); and RNN with longitudinal inputs.

Thedeep learning RNN models showed superior performances over the conventional LR models, implying that RNN models can be used to diagnose the disease.

III. METHODOLOGY

A. Data Collection and Description

The Hepatitis Dataset was collected from UCI machine learning repository. It consists of clinical test result of 155 hepatitis patients of 20 attributes with 123 survived patient and 32 mortalities. There were 13 instances with missing values, which was removed while cleaning the dataset leaving us with 142instances with 116 survivors' class and 26 death class. The HCC Dataset was also collected from UCI machine learning repository containing 165 instances with 50 attributes, 102 survivals and 63 death instances. Table 1 shows the summarized the datasets.

Table 1: Distribution of the hepatitis and HCC datasets

Cases	Attributes	Instances	Survived	Mortality	Survival rate	Mortality rate
Hepatitis	20	142	116	26	0.82	0.18
HCC	50	165	102	63	0.62	0.38

In this study, 114 instances of the Hepatitis Dataset were taken after manually eliminating instances with missing values. The class is a nominal type; the values of the class; patients who died as a result of Hepatitis and patients who survived. The dataset consists of 20 attributes. The HCC dataset contains 165 instances, having no missing values. The class is a nominal type consisting of patients who died as a result of HCC and patients who survived. The dataset consists of 50 attributes.

B. Algorithms

The supervised learning algorithms used to generate predictive models for performance analysis using the hepatitis datasets provided are all under the classification algorithms tab in WEKA; classification in machine learning is involved with recognising which class an object belongs to by training a set of objects whose class is already known—in this research's case, assigning a given patient to either "live" or "die." Classification is essentially an instance of supervised machine learning, the other being regression.

- *Random Forest*

The random forest containsmultiple decision trees that work as an ensemble. individual decision tree is a candidate for a class expectation, and the class with the most prediction turns into the entire model's prediction.

The main idea driving the random forest is based on the intelligence of groups. According to [12]in a decision standard tree, each node is divided using the best split among all variables, but in a random forest, each node is split using the best among a subset of predictors randomly chosen at that node.

WEKA uses a classifier, weka. classifiers. trees. Random Forest (or simply Random Forest), for random forest algorithm on a datasets.

➢ *Logistic Regression*

Logistic regression is a modelling technique that describes the relationship of variables, Xsof dependentand independent variables, [13]. Logistic regression is special case of linear regression, which uses logistic function, a more complex function as opposed to a linear function. The logistic function's hypothesis limits its cost function between 0 and 1, unlike linear functions which have values greater than 1 or less than 0. Logistic regression models the probability of an object occurrence based on individual characteristics. The logarithm of the probability is modelled as in equation is given by:

$$log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_{1\chi 1} + \beta_{2\chi 2} + \cdots \beta_{m\chi m} \quad (1)$$

Equation (1) is so because the probability is a ratio; π denotes the likelihood of an object occurrence while β_i are regression coefficients concerned with the reference group and the x_i denotes the independent variables.

In WEKA, logistic regression functions with datasets with both numeric and nominal classes uses a classifier, weka. classifiers. functions. Logistic (or simply Logistic).

➢ *Support Vector Machine (SVM)*

A support vector machine (SVM) is a supervised machine learning method that takes data and sorts it into one of two categories (Noble, 2006). It. It is a model that is suited for two-classed classification problems.In support vector machines, there are mainly two methods involved;

- Linear classification/classifier
- Non-linear classification/classifier by using the kernel trick

Linear classification/classifier calculates on a boundary that is a straight line. It is simpler and more computationally less intensive than the non-linear classifier.

Consider building a classifier that distinguishes patients who survived from Hepatitis and patients who could not.We find a line (a hyperplane or decision boundary) that would separate the two data points; patients who survived and those who could not. The equation involving the linear classifiers is obtained from the linear equation; $y = mx + b$. With the linear classifier, m replaced with the vector w, we have the linear classifier equation given as:

$$y = w^T x + b \quad (2)$$

where:

$y$denotes positive or negative classes (in this case, the class of survivors which is the positive class and the class of non-survivors which is the negative class). $w^T x$is the parameters for the planes between two classes, while b is the movement of the parameters out of the origin.The equation of the hyperplane (the decision boundary) is then given as:

$$w^T x + b = 0 \quad (3)$$

$y = 0$ because any data point which lies on the hyperplane is neither on the left or right of the origin. The data point which lies after the hyperplane would have $y =$ 0 and the data points which lies before the hyperplane would have $y = -1$

There are several options for obtaining the hyperplane, but the hyperplane that provides the maximum margin is chosen; the width from the nearest data point to the hyperplane on both sides.In a situation where we have outliers, we used a soft margin, which allows for errors; deviations in the margin.

The non-linear classifier is used when the data points are not linearly separable, or there are too many outliers to ignore. The non-linear classifier is created by applying the kernel trick. One of the vitalcomponents of SVMs is the kernel trick for the computation of dot products (i.e., data points) in high-dimensional feature spaces using simple functions defined on pairs of input patterns.

WEKA supports SVM with both numeric and nominal classes. WEKA uses a classifier, weka. classifiers. functions. SMO (or simply SMO), for SVM algorithm on datasets. Sequential Minimal Optimization (SMO) is an SVM algorithm for quadratic programming (QP) issue that emerges during the preparation of SVMs. SMO is broadly utilised for preparing SVMs.

➢ *K-nearest neighbours (K-NN)*

K-Nearest neighbour algorithm is used to perform classification (or regression). The decision rule provides a simple non-parametric procedure for the assignment of a class label to the input pattern based on the class labels represented by the K-closest neighbours of the vector [14]. The classification of an object is based on the majority vote of its neighbours, and then the object is appointed to the class that is most common in its K-nearest neighbour. In the event that k = 1, at that point the item is essentially assigned to the class of that solitary closest neighbour.When k-NN is used to perform the regression, the output is the property value for the object.A similarity (distance) function is needed to search a training set similar to it:

$$\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \quad (4)$$

$$\sum_{i=1}^{k}|x_i - y_i| \quad (5)$$

where x and y are vectors starting from the origin of the initial point to the endpoint (when considering two points).

Equation (4) is the Euclidean distance function, while equation (5) is the Manhattan distance function. Both functions are used exclusively for continuous variables.

In WEKA supports K-Nearest neighbour functions with datasets with both numeric and nominal classes, using the classifier, weka. classifiers. lazy. IBk (or simply IBk). IBk is capable of performing distance weighting.

➢ *Naïve Bayes*

The naïve Bayes classifier greatly simplifies learning by assuming that features are independent given classes. Although independence is generally a flawed assumption, in practice, naïve Bayes often competes with more

sophisticated classifiers. Naves Bayes classifier is primarily based on the Bayes theorem[15].Naïve Bayes algorithm originates from the Bayestheorem given that;

$$P[c|x] = \frac{P[X|C]P[c]}{P[x]} \quad (6)$$

From equation 3.8, c is the class and x is the instance, P[c|x] is the posterior probability, P[x|c] is the probability of predictor in a given class, P[c] is the prior probability, and finally, p[x] is the predictor's probability.

In WEKA, Naïve Bayes functions with datasets with both numeric and nominal classes. WEKA has a classifier, weka. classifiers. bayes. Naive Bayes (or simply Naive Bayes), which runs the naïve Bayes algorithm on datasets. Class for this Naive Bayes classifier uses estimator classes.

➢ *Decision Tree*

A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm that only contains conditional control statements [16]. Essential terminologies associated with decision trees include: the root node, decision node, splitting, terminal node (leaf), pruning, branch (sub-tree), parent and child node [17].

In WEKA, Decision Trees function with datasets with both numeric and nominal classes. WEKA has a classifier, weka.classifiers.trees.J48 (or simply J48), which runs the Decision Trees algorithm on datasets. J48 is a decision tree algorithm developed by Ross Quinlan and is used for classification. It functions on both continuous and discrete attributes, missing values, and pruning trees; it is an improvement on the ID3 algorithm.

➢ *Artificial Neural Networks (ANN)*

Artificial Neural Networks are physically cellular systems, which can obtain, store, and utilise experimental knowledge. They are a set of parallel and distributed computational elements classified according to topologies, learning paradigms, and at the way.The node is the elementary unit of the ANN,each node is capable ofadding many inputs x1, x2, …, xn form the environment or from other nodes, with each input modified by an adjusted node weight. The sum of these weighted inputs is added to an adjustable threshold for the node and then passed through a modifying (activation) function that determines the final output [18].

WEKA classifier uses weka. classifiers. functions. Multilayer Perceptron (or simply Multilayer Perceptron), for running the Artificial Neural Networks algorithm on datasets. According to Pal and Mitra (1992), the multilayer perceptron (MLP) containsseveral layers of simple, two-state, sigmoid processing elements (nodes) or neurons that interact through weighted connections. After a lowermost input layer, there is usually any number of intermediate or hidden layers followed by an outer layer at the top.

In [19] and [20], it was reported that the total input, x_j^{h+1}, in MLP received by neuron j in layer h+1 is given as

$$x_j^{h+1} = \sum_i y_i^h w_{ji}^h - \theta_j^{h+1} \quad (7)$$

where y_i^h is a state of ith neuron in the hth layer

w_{ji}^h is the heaviness of association from an ith neuron in h layer to jth neuron in h+1 layer

θ_j^{h+1} is the threshold of the j[th] neuron in h+1 layer

C. Performance Metrics

Those following metrics, computed from a confusion matrix, were employed to assess the performance of the algorithms:

There are four essential parameters in the confusion matrix:

- True Positives (TP): Instances of the dataset which the model predicted positive and the actual output was positive
- True Negatives (TN): Instances of the dataset which the model predicted negative and the actual output was negative
- False Positives(FP): Instances of the dataset which the model predicted positive and the actual output was negative
- False Negatives (FN): Instances of the dataset which the model predicted negative and the actual output was positive

➢ *Accuracy*

Accuracy is the average of the true parameters and the total number of instances of the data set.

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}} \quad (8)$$

➢ *Recall*

Recall (also called sensitivity, hit rate, or true positive rate) is the number of correct positive predictions divided by the all positive instance of the dataset.

$$\boldsymbol{recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \quad (9)$$

➢ *Precision*

Precision (also called positive predictive value) is the number of correct positive predictions divided by the number of predicted positives.

$$\boldsymbol{precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}} \quad (10)$$

➢ *Fall-Out*

Fall-out (also called false positive rate) is the number of incorrect positive predictions divided by the number of predicted negatives.

$$\boldsymbol{Fall-out} = \frac{\boldsymbol{FP}}{\boldsymbol{FP+TN}} \quad (11)$$

ISSN No:-2456-2165

➢ *F1 score/measure*

F1 Score/Measure is the Harmonic Mean between precision and recall. It denotes how many instances it classifies correctly, as well as how well it does not miss a significant number of instances.

$$\boldsymbol{F1} = \frac{1}{\frac{1}{precision}+\frac{1}{recall}} \quad (12)$$

➢ *Matthews correlation coefficient (MCC)*

The Matthews correlation coefficient (MCC) is the measure of the quality of a two-classed predictive model.

$$\boldsymbol{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (13)$$

IV. RESULTS AND DISCUSSION

This section presents and discusses the results of the assessment.

A. The Hepatitis dataset

Table 2 shows the performance of the algorithms on the Hepatitis dataset. The Support Vector Machine recoded the highest accuracy of 0.87, Fall out of 0.39 and MCC of 0.52 followed by the Naïve Bayes of accuacy of 0.86, Fall ou of 0.27 and MCC of 0.56, etc.

Table 2: Algorithms' classification performance on the Hepatitis dataset

Algo	Accuracy	Recall (TP Rate)	Fall-out (FP Rate)	Precision	F1 score	MCC
Random Forest	0.84	0.84	0.45	0.83	0.83	0.42
LR	0.82	0.82	0.43	0.82	0.82	0.40
SVM	0.87	0.87	0.39	0.86	0.86	0.52
K-NN	0.82	0.82	0.37	0.83	0.82	0.42
Naïve Bayes	0.86	0.86	0.27	0.87	0.86	0.56
Decision Tree	0.82	0.82	0.52	0.80	0.81	0.33
ANN	0.83	0.83	0.40	0.83	0.83	0.44

Although the accuracy of the SVM, which was marginally higher than the accuracy of the Naïve Bayes, The Naïve Bayes has the highest MCC score and the low Fall out. This implies that the Navies Bayes recorded the most qualitative two-classed predictive model. This is also collaborated with the least Fall out.

Figure 2 depicts the visual representation of the algorithms performance on the Hepatitis dataset.

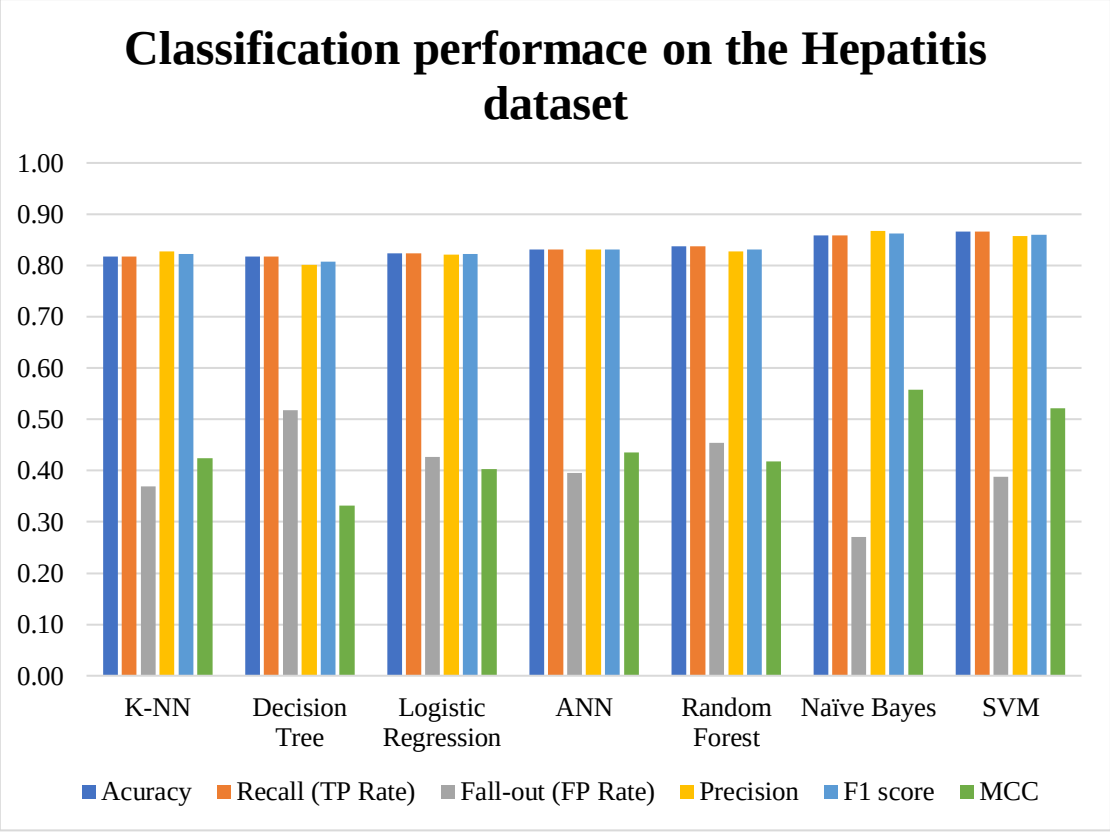

Fig. 1: Algorithms' performance on the Hepatitis dataset

ISSN No:-2456-2165

B. HCC dataset

Table 3 presents the algorithms' performance on the Hepatocellular carcinoma (HCC), while Figure 3 provide a visual representation of the dataset. Again, the SVM recorded the highest accuracy of 0.76, Fall out of 0.29 and MCC of 0.48 on the Hepatocellular carcinoma datasets; followed, this time, by the Random Forest with accuracy of 0.75, Fall out of 0.36 and MCC of 0.44. The SVM, this time, does not only has the highest accuracy but also proved to be the most qualitative two-classed predictive model

Table 3: Algorithms' Performance on the Hepatocellular carcinoma

Algorithms	Accuracy	**Recall (TP**	**Fall-out**	**Precision**	**F1 scor**	**MCC**
Decision Tree	0.59	0.59	0.44	0.60	0.60	0.15
K-NN	0.64	0.64	0.47	0.62	0.62	0.19
Naïve	0.67	0.67	0.42	0.66	0.65	0.26
Logistic	0.68	0.68	0.38	0.67	0.68	0.31
ANN	0.74	0.74	0.30	0.74	0.74	0.44
Random Forest	0.75	0.75	0.36	0.75	0.73	0.44
SVM	0.76	**0.76**	**0.29**	**0.75**	**0.76**	**0.48**

The Decision Tree recorded the least accuracy of 0.59 (0r 59%).

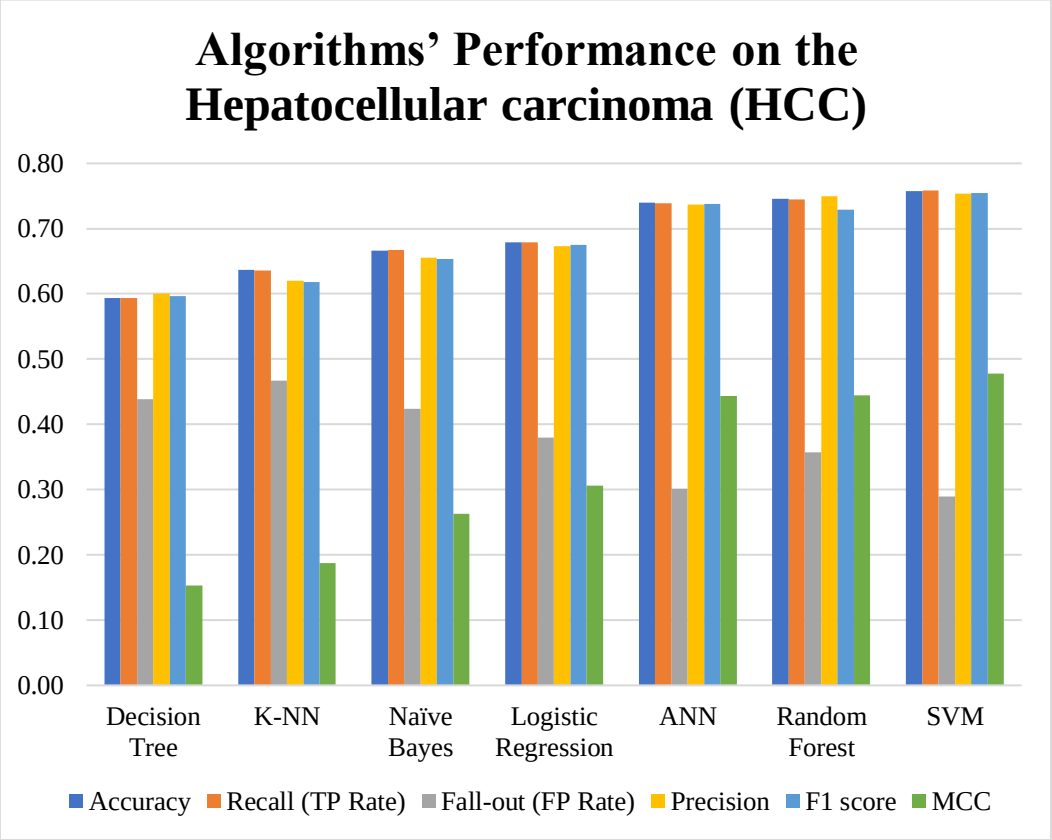

Fig. 2: Algorithms' performance on the Hepatocellular carcinoma dataset

ISSN No:-2456-2165

Table 4, depicts the algorithms' average classification accuracy performanceon both the Hepatitis and Hepatocellular carcinoma datasets. It could be sees that SVM recorded the highest accuracy on both datasets.

Table 4: Average percentage classification accuracy of the two datasets

Algorithm	Hepatitis Dataset	HCC Dataset	Average Accuracy
ANN	83.10	73.94	78.52
Decision Tree	81.69	59.39	70.54
K-NN	81.69	63.64	72.66
Logistic Regression	82.39	67.88	75.14
Naïve Bayes	85.92	66.67	76.29
Random Forest	83.80	74.55	79.17
SVM	**86.62**	**75.76**	**81.19**

The visual representation of the algorithms' accuracy on the two datasets is shown in Figure 4.

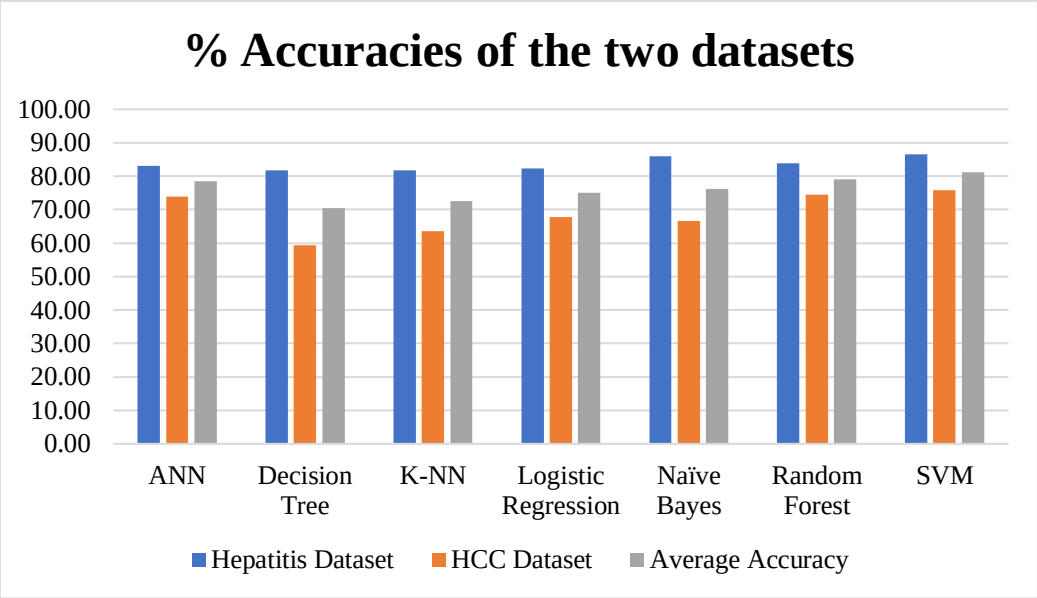

Fig. 3: Algorithms' Average classification accuracy on both datasets

V. CONCLUSION

This study assessed the performance of seven selected machine learning algorithm for classification of survival of Hepatitis and Hepatocellular carcinoma. The finding showed that the Support Vector Machine highest classification performance on the both datasets. This was followed respectively by the Naïve Bayes on the Hepatitis and the RandomForest on the Hepatocellular carcinoma. The Decision Tree recorded the least accuracies of both datasets. The result therefore suggests that the Support Vector machine, could be a most appropriate algorithm for developing a classification system for survival of Hepatitis and Hepatocellular carcinoma. Hoverer, the performance of these algorithms could as well be improved with more dataset.

ACKNOWLEDGMENT

The authors wish to appreciate Redeemer's University for the support and provision of state of the art facilities provided to us during the course of this work.

REFERENCES

[1.] X. ,. L. Zhang, H. Tian, Z. Zeng, J. Chen, D. ,. S. Huang, Ji, J. Guo, Cui and L. Y. Huipeng, "Risk Factors and Prevention of Viral Hepatitis-Related Hepatocellular Carcinoma," *Frontier in Oncology*, vol. 11, 2021.

[2.] J. Ferlay, M. Colombet and I. M. C. P. D. M. Soerjomataram, "Estimating the global cancer incidence," *GLOBOCAN sources and methods*, vol. 144, p. 1941–1953, 2019.

[3.] H. Sung, J. Ferlay and R. Siegel, "Global cancer statistics 2020: GLOBOCAN estimates of incidence," 2021.

[4.] P. Boyle and B. Levin, "World Cancer Report 2008.," 2008.

[5.] C. de Martel, D. Georges, F. Bray, J. Ferlay and G. Clifford, "Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis," *Lancet Glob Health*, vol. 8, p. e180–e190, 2020.

[6.] D. Alfaiate, S. Clement, D. Gomes, N. Goossens and F. Negro, "Chronic hepatitis D and hepatocellular carcinoma: A systematic review and meta-analysis of

observational studies," *Journal of Hepatology,* vol. 73, p. 533–539, 2020.

[7.] Shi, Y, Y. Wu, W. Wu, W. Zhang, J. Yang and e. al., "Association between occult Hepatitis B Infection and the Risk of Hepatocellular Carcinoma: a Meta-analysis.," *Liver International,* vol. 32, p. 231–240, 2012.

[8.] D. S. Mbaga, S. Kenmoe, C. Kengne-Nde, T. Ebogo-Belobo, G. Mahamat, J. R. Foe-Essomba, M. Amougou-Atsama, S. Tchatchouang, I. Nyebe, A. F. Feudjio, G. I. Kame-Ngasse and Magoudjou-Pe, "Hepatitis B, C and D virus infections and risk of hepatocellular carcinoma in Africa: A metaanalysis meta-analysis studies comparable for confounders," *PLoS ONE,* vol. 17, no. 1, p. e0262903, 2022.

[9.] W. Książeka, M. Abdarc, U. R. Acharyad and P. Pławiaka, "A Novel Machine Learning Approach for Early Detection of Hepatocellular Carcinoma Patients," *Cognitive Systems Research,* p. 10.1016/j.cogsys.2018.12.001, 2019.

[10.] X. Tian, Y. Chong, Y. Huang, P. Guo, M. Li, W. Zhang, Z. Du, X. Li and Y. Hao, "Using Machine Learning Algorithms to Predict Hepatitis B Surface Antigen Seroclearance," *Computational and Mathematical Methods in Medicine,* vol. 2019, 2019.

[11.] G. N. Ioannou, W. Tang, L. A. Beste, M. A. Tincopa, G. L. Su, T. Van, E. B. Tapper, A. G. Singal, J. Zhu and A. K. Waljee, "Assessment of a Deep Learning Model to Predict Hepatocellular Carcinoma in Patients With Hepatitis C Cirrhosis," *JAMA Network Open: Gastroenterology and Hepatology,* vol. 3, no. 9, p. e2015626, 2020.

[12.] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R News,* vol. 2, no. 3, pp. 18-22., 2002.

[13.] D. G. Kleinbaum, K. Dietz, M. Gail, M. Klein and K. M. , New York:: Springer-Verlag, 2002.

[14.] J. M. Keller, M. R. Gray and J. A. Givens, "A fuzzy k-nearest neighbor algorithm," *IEEE transactions on systems, man, and cybernetics,* vol. 4, p. 580., 1985.

[15.] M. O. Odim, A. O. Ogunde, B. O. Oguntunde and S. A. Phillips, "Exploring the Performance Characteristics of the Naïve Bayes Classifier in the Sentiment Analysis of an Airline's Social Media Data," *Advances in Science, Technology and Engineering Systems Journal* , vol. 5 , no. 4, pp. 266-272, 2020.

[16.] S. B. Rajesh, "Decision trees a Simple Way to Virtualise a Decision," 2018.

[17.] N. S. Chauhan, "Decision Tree Algorithm Explained," 2019.

[18.] M. O. Odim and V. C. Osamor, "Required Bandwidth Capacity Estimation Scheme For Improved Internet Service Delivery: A Machine Learning Approach," *international journal of scientific & technology research,* vol. 8, no. 8, pp. 326-332, 2019.

[19.] M. O. Odim, J. A. Gbadeyan and J. S. Sadiku, "A Neural Network Model for Improved Internet Service," *British Journal of Mathematics & Computer Science,* vol. 4, no. 17, pp. 2418-2434,, 2014.

[20.] M. O. Odim, J. A. Gbadeyan and J. S. Sadiku, "Modelling the Multi-Layer Artificial Neural Network for Internet Traffic Forecasting: The Model Selection Design Issues," in *CoRI'16,* Ibadan, 2016.