Comparing Decision Trees to Neural Nets in Predicting Heart Disease

1st Grace Shi
Harvard University

***Abstract*—In this paper, we are empirically comparing the performance of neural nets with the performance of decision trees based on a data set for analyzing the prediction of heart disease. Heart disease is one of the leading causes of death in the world. Around one person dies of heart disease every minute in the United Sates. Researchers use several data mining techniques to help health care professionals in the diagnosis of heart disease. Data mining techniques can reduce the strain on the healthcare sector. Decision trees and neural networks are two primary techniques of data analysis and modelling, both are able to process nonlinear data and handle variable interaction. This paper compares the performance of simple neural networks and the Gini index and entropy decision trees when tested upon data of 13 factors which may affect the diagnosis of heart disease, and the prediction thereof. We find that decision trees significantly outperform the neural networks in terms of error. Subsequently, we note that this may be due to the complexity of the neural network and further research can be done in the exploring and comparing further variations of both the neural networks and the decision trees, applied to different types of data and to health care data.**

I. Introduction

Heart disease is one of the leading causes of death in the world, and the leading cause of death in the United States [1]. Therefore, the preventative prediction and diagnosis of it is important to the medical field; it could help to provide on time treatment, decrease health costs, and linked death rates. One of the primary objectives of data mining practices and algorithms in the medical field is to apply data mining algorithms to patient data by properly using the data base to discover tacit information which helps doctors make better decisions.

As a result, using data mining and finding information in cardiovascular centers could generate useful knowledge that could be used by doctors to predict the potential behaviors of heart diseases based on past records, improving the standard of service offered by managers. Some of the most significant applications of data mining and information discovery in heart patients' systems include: diagnosing heart disease from different signs and properties, determining risk factors that raise the risk of heart disease, and so on. [2]

Determining the risk factors that contribute to heart disease manifesting in heart disease is a vital contribution to society. In the past decades, the incidence of heart disease has been increasing annually and has become one of the biggest threats to human health, and they are still on the rise. The American Heart Association reports that around 17.8 million deaths were attributed to heart disease in 2017, which is a 21.1% increase from 2007. The mortality rate is expected to continue to rise in the next decade. Therefore, prediction and thereby prevention and control of heart disease are in urgent need. This problem can in fact be prevented if it is diagnosed early and effectively addressed so to have a timely intervention. [3]

Diagnosing heart disease with machine learning algorithms are receiving more and more attention from researchers. Artificial neural networks have been used for the prediction of heart disease, in which simulated data were used to train an artificial neural network in order to generate a mapping from input data to an output, and that model is then used to predict the occurrence of a heart disease. Decision tree methodology is also an efficient decision support model for this prediction system. Comparing these two data analysis and modelling techniques has been a long present debate in the data mining studies. Both methods can model data with nonlinear relationships between variables and both can handle interactions between variables, so there is much controversy between which is better for what fields. The fact of the matter is that simply each is better for different instances, and in this paper I aim to explore which is more effective for the prediction of heart disease. [4]

This paper will compare the accuracy in predicting heart disease between decision trees and neural networks. Related literature, methodologies, and results are described below. The benefits of various data analysis methods in predicting the consequences of heart disease are compared, and the findings are discussed. Results arising from this study provide important reference materials for the prediction of heart disease and selecting a more accurate approach to estimate them.

II. Literature

There have been numerous works in related literature with heart disease diagnosis using data mining techniques which have motivated our work in this paper.

Heon Gyu Lee et al. proposed a novel technique to develop the multi-parametric feature with linear and nonlinear characteristics of Heart Rate Variability (HRV). They used statistical and classification techniques to develop the multi-parametric feature of HRV, and assess the linear and non-linear properties of HR. [5]

Sellappan Palaniappan et al. developed the model Intelligent Heart Disease Prediction System with the aid of data mining techniques like Decision Trees, Naive Bayes, and Neural Networks. The results illustrated the strength of each methodology in comprehending the objectives of specified mining objectives [6].

Niti Guru et al. experimented with the prediction of heart disease, blood pressure, and sugar with the aid of neural networks on a sample database of patients' records. The neural network was tested and trained with various input variables and it was found that the supervised network should be recommended for diagnosis of heart diseases. [7]

Carlos Ordonez studied the problem of identifying constrained association rules for heart disease prediction. He assessed a data set which was encompassed by medical records of people having heart disease with attributes for risk factors such as heart perfusion measurements adn artery narrowing. It was found that these constraints reduced the number of discovered rules remarkably in addition to decreasing the running time. [8]

The work of Franck Le Duff et al. studied the data mining methods which may aid the clinicians in the predication of the survival of patients and in the consequent adaptation of the practices. It could be conducted upon each medical procedure or medical problem, and it would be feasible to build a decision tree rapidly with the patient data provided by a service or physician. Comparing traditional analysis and data mining analysis demonstrated how data mining methods which sort variables such as decision trees, and concludes the significance or effect of each variable on the study. [9]

Boleslaw Szymanski et al. proposed a novel heuristic for efficient computation of sparse kernels in SUPANOVA. When it was applied to a benchmark Boston housing market data set,they found that 83.7% predictions in the results were correct and thereby outperformed the Support Vector Machine and other previously established kernels. [10]

Lathat Parthiban et al. proposed an approach using coactive neuro-fuzzy inference system for prediction of heart disease. This model diagnosed the presence of heart disease by merging the neural network adaptive capabilities and the fuzzy logic qualitative approach and integrating other genetic algorithms. The model was found to be promising in prediction of heart disease. [11]

All of these previous studies incorporate various algorithmic or statistical models and analyze their efficacy in predicting heart disease. In this paper I aim to compare the efficacy of decision trees against that of neural networks in application to predicting heart disease, which contributes both to the heart disease prevention field as well as to the field of data science.

Heart Disease

The term heart disease encompasses all the diverse diseases and afflictions which affect the heart. It was a major cause of casualties in the United Sates, killing one person every 34 seconds [3]. Coronary heart disease, cardiomyopathy, and cardiovascular disease are all included by the term. Cardiovascular disease includes a wide range of conditions that affect the heart and the blood vessels and the manner in which blood is pumped and circulated through the body. It leads to severe illness, disability, and death. Coronary heart disease stems from narrowing of the coronary arteries which results in reduction of blood and oxygen supply to the heart, leading to myocardial infarctions, aka. heart attacks. Specifically, they are caused by a sudden blockage of a coronary artery, generally due to a blood clot. [12]

On a whole, the World Health Organization has estimated that 12 million deaths occur world wide due to cardiovascular diseases. Half the deaths in the United States and other developed countries occur due to cardiovascular diseases, and it is also a primary cause of death in numerous developing countries. It is overall regarded as the primary cause of death in adults. [12]

III. Methodology

Predictive modeling is a technique for predicting the values of one or more variables in a data set (outputs) based on the values of other variables in the data set (inputs). Neural network and decision tree models are the two most popular predictive modeling techniques. The algorithms developed in these modeling techniques arise from methodological research in various disciplines, including statistics, pattern recognition, and machine learning. These modeling techniques' algorithms are based on methodological analysis in a variety of disciplines, including statistics, pattern recognition, and machine learning.

This section will describe the techniques applied to analyze the heart disease data.

Data

This data is from an online data base for heart disease patient classification, kept up to date annually. This data is from March 2021.

The features considered in the prediction are as follows

	Factor	Description
0	age	age in years
1	sex	sex
2	cp	chest pain type
3	trestbps	resting blood pressure
4	chol	serum cholestoral
5	fbs	fasting blood sugar
6	restecg	resting electrocardiographic results
7	thalach	maximum heart rate achieved
8	exang	exercise induced angina
9	oldpeak	ST depression induced by exercise relative to rest
10	slope	slope of the peak exercise ST segment
11	ca	number of major vessels colored by flourosopy
12	thal	thalassemia inherited blood disorder

There are 303 input values each with these 13 factors. Note that we 0 index this chart for reference in later sections.

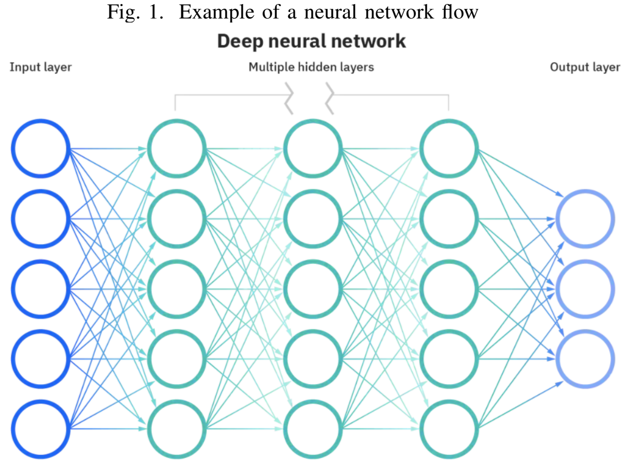
The last value of the input is target variable: the diagnosis and manifestation of heart disease, measured by the narrowing of the vessel beyond 50% which could lead to cardiac arrest, where Class 0 is defined as no heart disease and Class 1 is defined as diagnosis of heart disease.

Since we only have observational data, we divide the data such that 0.7 of the data is used for training, and 0.3 is used to test the models.

Neural Network

Theory: Researchers first created neural network models in order to simulate the neurophysiology of the human brain. The models are analytic techniques that are based on (hypothesized) learning processes in the cognitive system and neurological functions of the brain, and are capable of predicting new observations (on specific variables) from previous observations (on the same or different variables) after performing a so-called learning from existing data procedure. [13] The process is shown in Fig. 1 below.

Fig. 1. Example of a neural network flow



The feed forward network is the simplest and most commonly used type of neural network. Training a neural network describes the process of setting the best weights on the inputs of each of the units, and back propagation is the most common method for computing the error gradient for a feed forward network [14]. The neural network then adjusts weights so to minimize the error over multiple iterations.

Neural nets tend to perform well in applications with non-linear functional forms [15]. They are particularly useful for problems with uncertain mathematical formulae and no prior knowledge of the relationship between inputs and outputs.

The disadvantage of using neural networks however is that it does not provide the statistical p-value for testing the significance of the parameter estimates. Namely, the preliminary step of feature selection before any learning occurs is required [16]. The artificial neural networks which have hidden layers are better classifiers for non linear decision hyper-surfaces, but they are a veritable "black box" in that they are hard to interpret.

Implementation: As stated above, the artificial neural network is an information processing paradigm inspired by the brain. The process of neuron firing across synapses to relay information is modeled by creating networks on computers using matrices. Since the scope of this paper is to compare a decision tree to the neural net, the one we implement will be kept simple with only two layers.

The training process will consist of two steps: forward propagation, back propagation.

The forward propagation will take some input $X = x_1, x_2, ..., x_{13}$ of the data and multiply by a weights vector $W = w_1, w_2, ..., w_{13}$ (starting with a vector of random values) to get some resulting vector

$$Y = [W \times X]$$

Then we will pass this result through a sigmoid formula to calculate the neuron's output, namely,

$$\sigma(y) = \frac{1}{1 + e^{-y}}$$

The sigmoid function essentially normalizes the result between 0 and 1.

The back propagation will calculate the error, namely, the difference between the actual output and the expected output. Depending on the error calculated, it will adjust the weights by multiplying the error e with the input and again with gradient of the sigmoid curve, as follows

$$W+ = eXY(1 - Y)$$

Where $Y(1 - Y)$ is the derivative of the sigmoid curve.

This process will be repeated 10000 iterations in training.

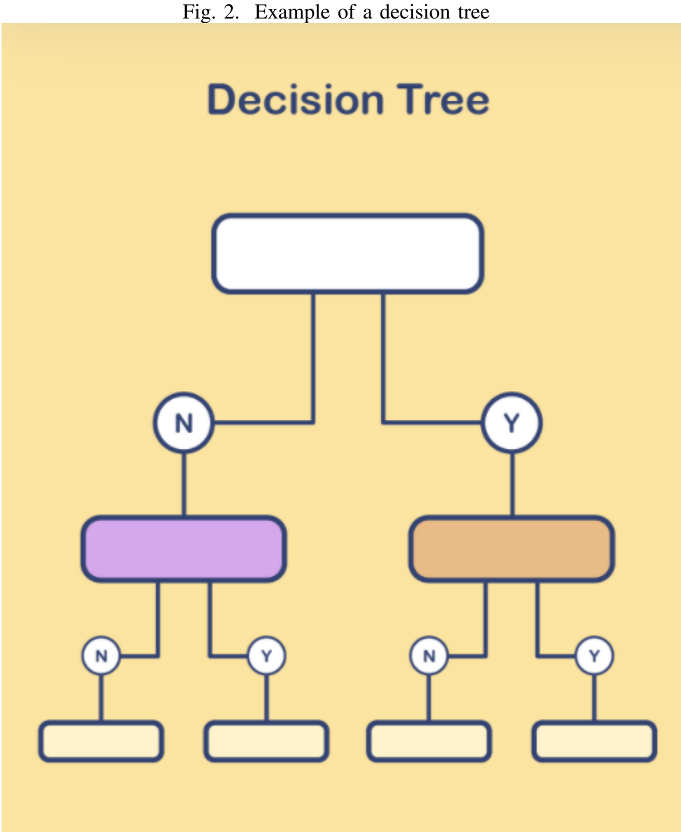
Then we test this model against the subset of data we set aside for testing. The efficacy of the model will be calculated by the final error when we compare the predicted values to the actual values of the target variable, ie. occurrence of heart attack. This will be taken as the mean absolute error. [17]

Decision Tree

Theory: In decision tree modeling, an empirical tree is a segmentation of data that is generated by applying a set of simple rules. Through the repeated process of splitting, these models produce a collection of rules that can be used for prediction as shown in Fig. 2. Chi-squared automatic interaction detection (CHAID), classification and regression trees (CART), C4.5, and C5.0 are all popular tree methods [18].

The biggest advantage of using decision trees over other modelling techniques such as neural networks is that it produces a model which represents interpretable rules and logic statements, unlike neural networks. That is, the explanation capability that exists for trees producing axis parallel decision surfaces is a crucial feature [16]. Classification in decision trees can be performed without complex computations and this can be used for continuous and categorical variables. In addition, decision tree models provide results which clearly

Fig. 2. Example of a decision tree



rank the importance of significant factors for prediction or classification.

The downside of decision trees are that they do not perform as well as neural networks for nonlinear data, and are therefore more susceptible to error if the data is noisy. Decision trees are usually considered by the academic community to be more suitable for predicting categorical outcomes, over data that is taken over time (unless visible trends and sequential patterns are available) [15].

Implementation: In implementing the decision tree, we make the following assumptions: at the beginning the whole training set is the root, attributes are assumed to be categorical for information gain and continuous for Gini index.

First we find the best attribute and place it on the root node of the tree, then we split the training set of the data into subsets. Each subset of the training dataset has around the same value. Then we find leaf nodes of all branches by repeating the above on each subset, as shown in Fig 2. More specifically, there are two phases in building the decision tree: building, and operational.

In the building phase, we pre process the data set and split it into training and test data as described in the data section above. Then we train the classifier.

We use the Gini index and information gain to select from the $n = 13$ attributes of the data set which attribute should be placed at the root node (or the internal node in a recursive call). The Gini index is as follows

$$1 = \Sigma_j P_j^2$$

where P is the probability distribution function of an object being classified to a particular class j over all classes j. The Gini index is used to measure how often a randomly chosen element will be incorrectly identified. Namely, an attribute with a lower Gini index is preferred.

We also use entropy in the building phase for information gain. Consider entropy such that if a random variable x can take N different values, the ith value x_i wil probability $p(x_i)$ will have the entropy

$$H(x) = -\sum_{i=1}^{N} p(x_i) \log_2 p(x_i)$$

Entropy is the measure of uncertainty of a random variable. Namely, it characterizes the impurity of an arbitrary collection of examples. The higher the entropy, the more information content. Then in information gain, we determine which feature gives the maximum information about a class by aiming to reduce the level of entropy from the root node to the leaf nodes.

In the operational phase, we make predictions and calculate the accuracy (absolute error). We also use a confusion matrix to understand the trained classifier behavior over the test dataset by counting the number of times the instances of some class X is misclassified as class Y. [19]

IV. Results

In order to find the factors which influence the occurrence of a heart attack, the neural networks and decision trees are built for this data. The target variable is whether or not a person has a heart attack. Factors considered for their influence over the target variable are the factors listed in methodology.

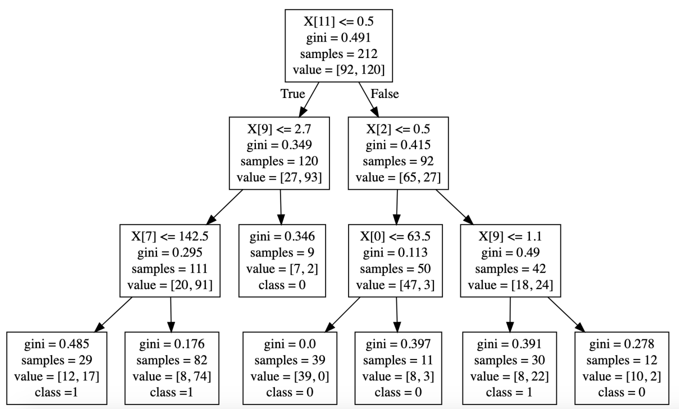
Neural network

After training the neural network model on the training data over 10000 iterations, we find that we get the final weights as follows

	Factor	Final Weight
0	age	-0.16595599
1	sex	0.44064899
2	cp	-0.99977125
3	trestbps	-0.39533485
4	chol	-0.70648822
5	fbs	-0.81532281
6	restecg	-0.62747958
7	thalach	-0.30887855
8	exang	-0.20646505
9	oldpeak	0.07763347
10	slope	-0.16161097
11	ca	0.370439
12	thal	-0.5910955

It is evident from the table above that the factors of cp, fbs, chol, restecg are among the highest influencing factors, and

Fig. 3. Gini index decision tree (reference image at the end)



oldpeak, age, slope, exang are among the lowest influencing factors. Namely, the most influential factors end up being chest pain type, fasting blood sugar, cholesterol, and resting electrocardiograph results; the least influential factors are ST depression induced by exercise relative to rest, agem the slope of the peak exercise ST segment, and exercise induced angina. The final equation for the output is therefore (approximately) given as follows

$$Y = -0.16age + 0.44sex - 0.99cp - 0.39trestbps - 0.706chol - 0.81fbs - 0.62restecg - 0.308thalach - 0.206exang + 0.07oldpeak - 0.16slope + ca0.37 - 0.59thal$$

We are given the mean absolute error of 49.45%.

Decision tree

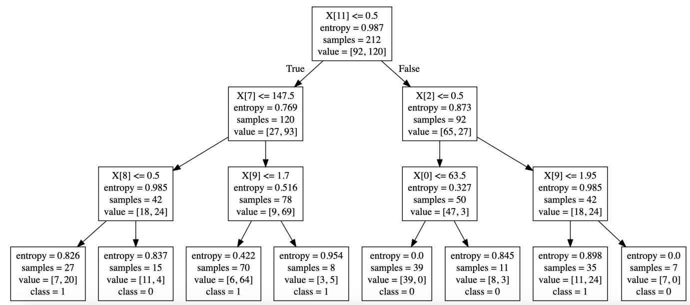
Gini Index: From the decision tree analysis, we get the classification tree as shown for Gini index in Fig. 3 with $X[i]$ corresponding to factor i in the factors table given in methodology.

It is clear that $X[11]$ is a pinnacle factor in the Gini index tree, namely the ST depression induced by exercise relative to rest. Then its clear that $X[2]$ and $X[9]$ are the next important classifying factors in the Gini index decision tree, namely age and maximum heart rate achieved.

Then consider that many factors are considered dummy variables by the tree. Namely, $X[1], X[3], X[4], X[5], X[6], X[8], X[10], X[12]$ in the Gini index decision tree. However it should be noted that there are some factors which do not change the final class assignment dependent on the value. Namely, $X[7], X[0]$. So, we have that sex, trestbps, chol, fbs, restecg, exang, slope, thal, thalach, age are not influential factors in the Gini index based decision tree.

We get the overall accuracy score of 73.62% , 82% for class 0 corresponding to no heart disease, and 68% for class 1 corresponding to diagnosis of heart disease. This corresponds to an average error of 1 - accuracy which is 26.38%.

Fig. 4. Entropy/information gain decision tree (reference at the end)



Entropy/information gain: From the decision tree analysis, we get the classification tree as shown for entropy/information gain in Fig. 4 with $X[i]$ corresponding to factor i in the factors table given in methodology.

It is clear that $X[11]$ is a pinnacle factor in the Entropy index tree, namely the ST depression induced by exercise relative to rest. $X[2]$ and $X[7]$, age and fasting blood sugar, are the next important classifiers for the entropy decision tree.

Then consider that many factors are considered dummy variables by the tree. Namely, $X[1], X[3], X[4], X[5], X[6], X[10], X[12]$ for the entropy tree. However it should be noted that there are some factors which do not change the final class assignment dependent on the value. Namely, $X[9]$ given $X[7] \leq 147.5$, $X[0]$. So, we have that sex, trestbps, chol, fbs, restecg, slope, thal, age are not significant factors in influencing the entropy based decision tree.

We get the overall accuracy score of 83.51% , 92% for class 0 corresponding to no heart disease, and 78% for class 1 corresponding to diagnosis of heart disease. This corresponds to the average error of $1-$error which becomes 16.49%.

V. Empirical Analysis

For the purposes of comparison, we have the following comparison chart

Model	Error
Neural Network	49.45%
Decision Tree - Gini	26.38%
Decision Tree - Entropy	16.49%

It is evident from this data that decision tree based on entropy/information gain is the most accurate, and neural network is the least by a significant factor of 299.87% . The decision tree overall outperforms the neural network from this comparison.

In this empirical application to heart disease study, the simple decision trees seem to be significantly more viable than the simple neural network.

For another source of comparative data, consider the following summary of the significant factors identified in the models.

For the neural network, we identify significance as having an absolute weight above 0.5. Note that this is an arbitrary division just for the purposes of this visualization.

For the decision tree, we identify significance as being a divisive node in the tree.

	Factor	Neural Net	Gini	Entropy
0	age			
1	sex			
2	cp	*	*	*
3	trestbps			
4	chol	*		
5	fbs	*		
6	restecg	*		
7	thalach			*
8	exang			*
9	oldpeak		*	*
10	slope			
11	ca		*	*
12	thal	*		

This analysis suggests that there is at least a correlation between the significant factors and later diagnosis of heart disease. We are not able to conclude conclusively causation in this study alone because that involves statistics is out of the scope of this study. However we can conclude from the errors and the models generated that there is a strong correlation between these significant factors of cp, oldpeak, ca and later diagnosis of heart disease.

VI. Analysis

We conclude that the decision tree outperforms the neural networks in this study.

We must note that this does not imply that neural networks are in general worse in predicting heart disease data than decision trees. Because of the simplicity of the neural network we implemented, its is very likely that if we had added more hidden layers (aka. "neurons") it is entirely possible that the neural networks could beat the performance of the decision trees. However a similar argument could be made for increasing the complexity of the decision trees, and it must be noted that the decision to implement the simplest version of a neural net was for the sake of comparison. Therefore, this comparison allows us to conclude that the simplest neural network performs worse compared to a simple single decision tree, and does not conclude about the most complex neural network or decision tree mechanism (ie. random forest). However, that is beyond the scope of this paper and could be an area of further study.

In general, in application to observational data, we can interpret from this study that with the most simple neural network and decision tree mechanisms, the decision tree outperforms the neural network in regards to this study of a small set of observational data.

VII. Conclusion

Overall from the comparison, the data seems to suggest that decision trees are more effective for the prediction of heart disease based on this data set in this study. There is a significant difference in error between the performance of the neural network and both variations of the decision trees such that we can conclude this just from this study.

We also find that the neural networks and decision trees after training have determined different factors to be more influential than others, and while that may very well contribute to the differences in error, we find that some factors are more influential than others in predicting the diagnosis of heart disease.

We find this result surprising in that we did not predict how significant the difference in error would have been between the decision trees and the neural networks. However, since this is observational data and limited in size, this is not surprising since in the case of having not ideal data sets, unpredictable effects may occur. As noted in the analysis section, we must note that there is further room for research in comparing different kinds of neural networks and decision trees which may perform differently.

Data mining in health care management is unlike other fields since the data present are heterogeneous, and contain certain ethical, legal, and social constraints due to the constraints that apply to private patient data and medical information. Where there is a large volume of healthcare data and from a large variety of sources, not all of them are appropriate in structure or quality. The exploitation of the knowledge and experience of specialists and clinical screening data of patients which is gathered in a database during the diagnosis procedure is widely recognized. The pre-processed heart disease data used in this study for example, is an example of a sample of this data, and this study demonstrates the difficulty as well as the fruitfulness of studying the rich data. Heart disease is one of the most significant causes of death in the world, and combining it with data analysis techniques contributes to alleviating the health burden on the entire human society.

The debate between neural networks and decision trees is a long held argument because both methods model data with non-linear relationships between variables and both are able to handle interactions between variables. There are numerous arguments for each. Neural networks are less easy to decipher, and the user must simply trust the results and that the hidden networks are performing as they were meant to. Decision trees do not perform as well on continuous data, or data that is meant to be interpreted globally, such as image recognition or natural language processing. This will continue to be a huge debate in data mining and will continue to require further studies, both in the debate of decision trees versus neural networks and the prediction of diagnosis of heart disease.

It must be noted that the study of data analysis is not purely a futile intellectual exercise; the further study and development of data mining techniques incrementally both improves on the final product, the frontier of technology which significantly contributes to the day to day life of people in society, as well as improving the techniques dynamically as each paper is released. Namely, studying data analysis is both improving the final product and dynamically improving the current methods, which ultimately help a diverse range of fields from healthcare to investment to sales to energy conservation.

Areas for further research include studying the comparison of decision trees and neural networks of different levels of complexity. Namely it is difficult to determine what intermediate levels of complexity are appropriate for comparison, and therefore we have only implemented the simplest versions in this paper. Further studies could be done on more complex versions of both, for example on neural networks with more hidden layers or random forests, or use of equal frequency discretionization gain, multivariate decision trees, or ratio decision trees . Further data research could also be done on logistic regression, linear regression, or other mechanisms of data analysis. We could also explore different rules such as association, clustering, K-means, etc for the neural network [20].

References

[1] "FastStats - Leading Causes of Death." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 1 Mar. 2021

[2] Nasrabadi, Abbas and Haddadnia, Javad. (2016). Predicting Heart Attacks in Patients Using Artificial Intelligence Methods. Modern Applied Science. 10. 66. 10.5539/mas.v10n3p66.

[3] Virani, S.S.; Alonso, A.; Benjamin, E.J.; Bittencourt, M.S.; Callaway, C.W.; Carson, A.P.; Chamberlain, A.M.; Chang, A.R.; Cheng, S.; Delling, F.N.; et al. Heart Disease and Stroke Statistics-2020 Update: A Report From the American Heart Association. Circulation 2020, 141, e139–e596.

[4] "Decision Trees Compared to Regression and Neural Networks." DTREG, www.dtreg.com/methodology/view/decision-trees-compared-to-regression-and-neural-networks.

[5] Heon Gyu Lee, Ki Yong Noh, Keun Ho Ryu, "Mining Biosignal Data: Coronary Artery Disease Diagnosis using Linear and Nonlinear Features of HRV," LNAI 4819: Emerging Technologies in Knowledge Discovery and Data Mining, pp. 56-66, May 2007.

[6] Sellappan Palaniappan, Rafiah Awang, "Intelligent Heart Disease Prediction System Using Data Mining Techniques", IJCSNS International Journal of Computer Science and Network Security, Vol.8 No.8, August 2008

[7] Niti Guru, Anil Dahiya, Navin Rajpal, "Decision Support System for Heart Disease Diagnosis Using Neural Network", Delhi Business Review , Vol. 8, No. 1 (January - June 2007.

[8] Carlos Ordonez, "Improving Heart Disease Prediction Using Constrained Association Rules," Seminar Presentation at University of Tokyo, 2004.

[9] Franck Le Duff , Cristian Munteanb, Marc Cuggiaa, Philippe Mabob, "Predicting Survival Causes After Out of Hospital Cardiac Arrest using Data Mining Method", Studies in health technology and informatics ,107(Pt 2):1256-9, 2004.

[10] Boleslaw Szymanski, Long Han, Mark Embrechts, Alexander Ross, Karsten Sternickel, Lijuan Zhu, "Using Efficient Supanova Kernel For Heart Disease Diagnosis", proc. ANNIE 06, intelligent engineering systems through artificial neural networks, vol. 16, pp:305-310, 2006.

[11] Latha Parthiban and R.Subramanian, "Intelligent Heart Disease Prediction System using CANFIS and Genetic Algorithm", International Journal of Biological, Biomedical and Medical Sciences 3; 3, 2008

[12] Shantakumar B.Patil, Dr.Y.S.Kumaraswamy, "Extraction of Significant Patterns from Heart Disease Warehouses for Heart Attack Prediction", IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.2, February 2009.

[13] Geoffrey K.F. Tso, Kelvin K.W. Yau, Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks, Energy, Volume 32, Issue 9, 2007, Pages 1761-1768, ISSN 0360-5442.

[14] M.J.A. Berry, G. Linoff, Data mining techniques for marketing, sales, and customer support, Wiley, New York (1997)

[15] S.P. Curram, J. Mingers, Neural networks, decision tree induction and discriminant analysis: an empirical comparison, J Oper Res Soc, 45 (1994), pp. 440-450

[16] P. Perner, U. Zscherpel, C. Jacobsen, A comparison between neural networks and decision trees based on data from industrial radiographic testing, Pattern Recognition Lett, 22 (2001), pp. 47-54

[17] "Implementing Artificial Neural Network Training Process in Python." GeeksforGeeks, 21 Aug. 2020

[18] J.R. Quinlan, C4.5 programs for machine learning Morgan Kaufmann, San Mateo (1993)

[19] "Decision tree implementation using Python." GeeksforGeeks, 21 Nov, 2019.

[20] Jaymin Patel, Prof.TejalUpadhyay, Dr. Samir Patel, Heart Disease Prediction Using Machine learning and Data Mining Technique, IJCSC Vol 7, Sept 2015

Fig. 5. Gini index decision tree

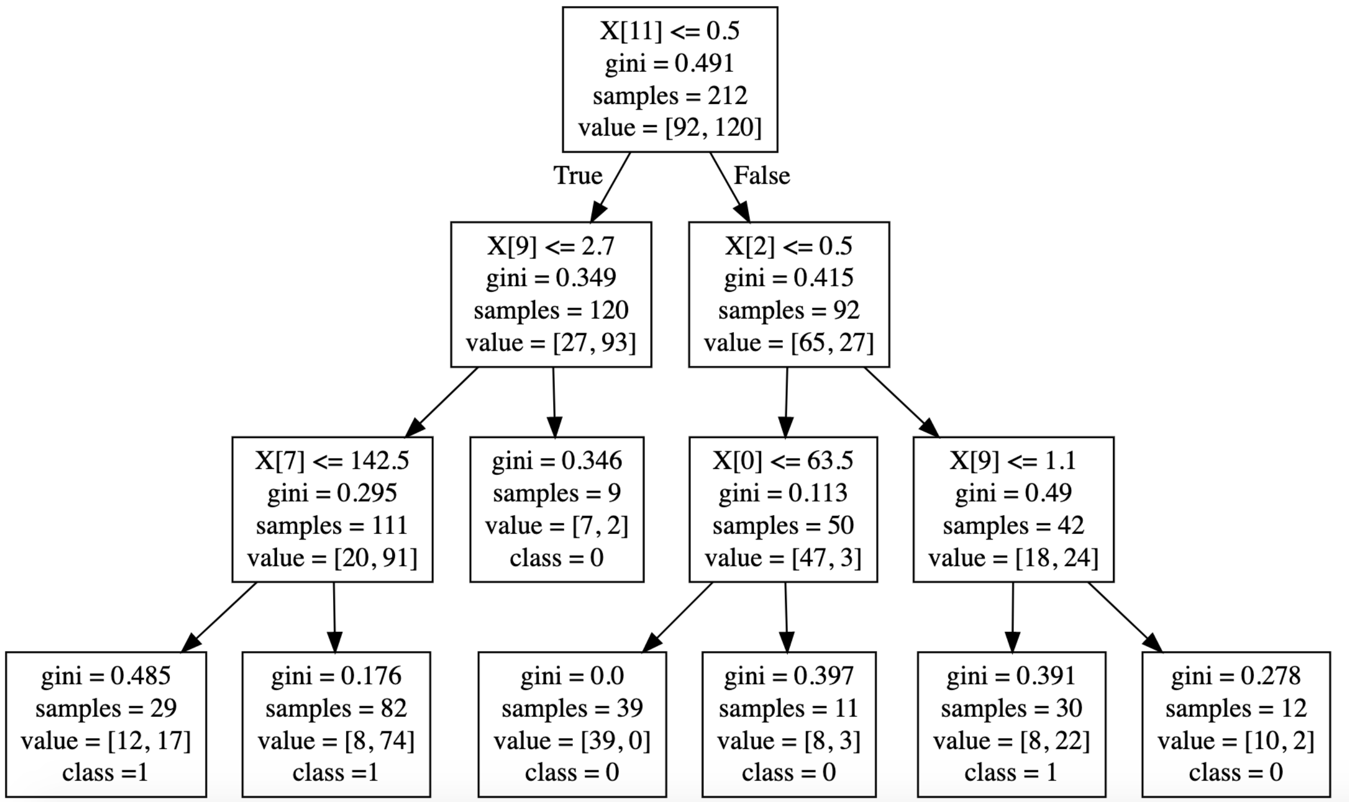


Fig. 6. Entropy/information gain decision tree