Predicting a disease based on the diagnosis of another disease using machine learning

Yasmine Emad, Doaa Maher, fadia Alazab, Ahmed Younis, Ahmed Mohamed, Ahmed Labib and Mahmoud Saeed
Dr. Ahmed Hagag

Faculty of Computers & Artificial Intelligence – Benha University

Faculty of Computers & Artificial Intelligence

Benha University

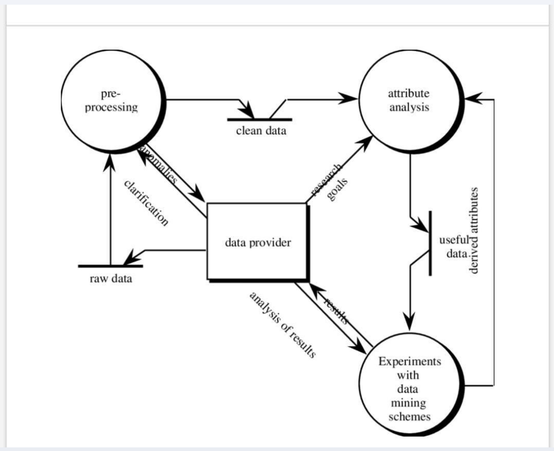
Abstract

Disease classification/detection is a crucial and challenging problem, because it helps in early diagnosis of disease by supporting the pathologists and doctors in their decision. Machine Learning technique is one of the emerging field can be used in the health sectors for the diagnosis of different diseases. This paper presents an effective approach for the diagnosis of chronic kidney disease (CKD) using artificial neural network (ANN) with back propagation algorithm, where first we fill the missing values of the dataset using mean, mode and median of attributes. Further, we have trained the NN classifier and evaluate the detection performances on separate test dataset. From the comparative analysis with other variants of classifiers like logistic regression, Naive Bayes, Decision tree, Random forest, K-NN and support vector machine (SVM), it is found that the recognition accuracy of Random forest is significantly encouraging.

Objective

Heart related diseases and chronic kidney disease (CKD) are the main reason for a huge number of deaths in the world over the last few decades and has emerged as the most life-threatening disease. So, there is a need of reliable, accurate and feasible system to diagnose such diseases in time for proper treatment. Machine Learning algorithms and techniques have been applied to various medical datasets to automate the analysis of large and complex data. Many researchers, in recent times, have been using several machine learning techniques to help the health care industry and the professionals in the diagnosis of heart related diseases.

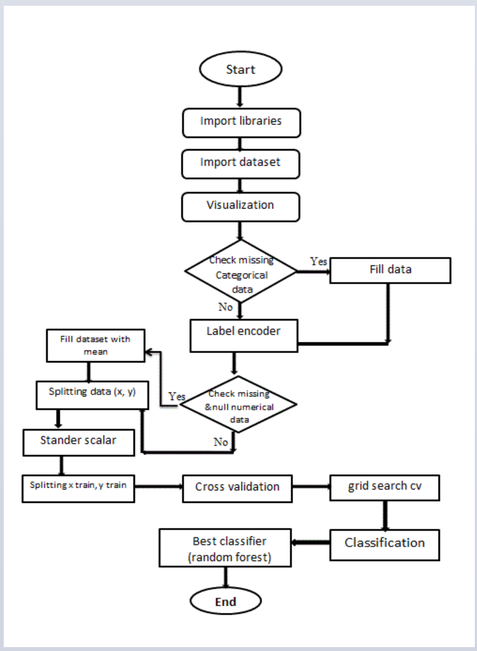
Methods



Data flow diagram

On one hand we implement an interactive Software with machine learning classifiers and web Application to help doctors and patients to record their Indication and Attachment them by using flask method

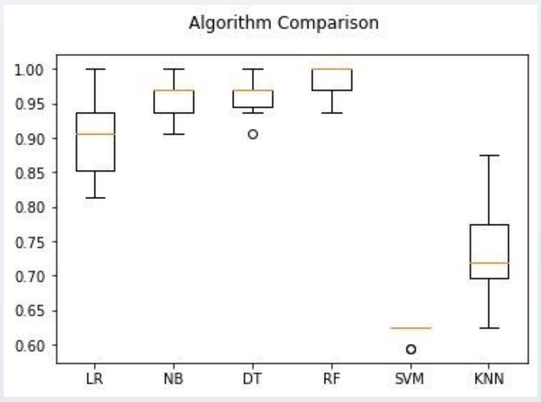
Classifiers as Logistic Regression: LR , Naive Bayes: NB, Random Forest Classifier: RFC, Extreme Gradient Boost: EGB, K-Nearest Neighbor: KNN, Decision Tree: DT, Support Vector Machine: SVM



Flow chart of machine learning model

Results

We find that random Forest Classifier is the beat classifier to our mode as like table of accuracy above



We conclude to two models and Computes accuracy to each model

classifier	accuracy	Precision score	Recall score	F-Measure
LR	97.5	0.942857142857142 8	1.0	0.970588235294179
NB	95.0	0.891891891891891	1.0	0.942857142857142 8
DT	93.75	0.917647058 8	0.939393939394	0.925373134328358
RF	98.75	1.0	0.969696969697	0.984615384 61
SVM	97.5	0.942857142857142 8	1.0	0.970588235 29
KNN	96.25	0.916666666	1.0	0.956521739 1

model1

classifier	Accuracy	F-Measure
LR	95.0	0.48717948717948
NB	52.5	0.404388714733542 6
DT	88.75	0.6232390894819
RF	96.25	0.49044585872615
SVM	95.0	0.48717948717948
KNN	93.75	0.62651727357609

Model 2

Connection between machine models and web Application by (flask).

Its function to connect between web and machine, we use flask to get data from the website and send it to machine model classifiers to get the result (if patient have CDK then predict with having HF or not) then send the result back to the website to show the result to doctor.

Used tools : Visual Studio code, Pycharm

Conclusions

As a result, we have studied Advantages and Disadvantages of Machine Learning. Also, this blog helps an individual to understand why one needs to choose machine learning. While Machine Learning can be incredibly powerful when used in the right ways and in the right places (where massive training data sets are available) and The importance of machine learning Machine learning as a technology helps analyze large chunks of data, which facilitates the task of data scientists in an automated process and gains a lot of importance and recognition. Machine learning also benefits in high-value predictions that can guide better decisions and intelligent actions in real time without human intervention. This is the use of Amazon machine learning to predict what customers want and provide it to them, which helps it generate huge profits from behind this matter.

In this project, we meet our main purpose and provide machine learning and our data to apply our project idea to help patients and doctors by firstly we make machine model then we service and Attachment them by using flask.

For the work in the future, we will work to improve and give the project more additional features in the future.

References

1.https://www.google.com/url?q=https://www.ripublication.com/ijaerspl2019/ijaerv14n10spl_30.pdf&sa=D&source=editors&ust=1625866633760000&usg=AOvVaw04Iqy1CRmttUbOMuWI_l-w

2.. https://data-flair.training/blogs/advantages-and-disadvantages-of-machine-learning/

Acknowledgements

ALLAH is the first and the last to be thanked. We want to thank all those who helped us by their knowledge and experience. We will always appreciate their efforts. First and foremost, we would like to express a sincere gratitude to our mentor and advisor Dr. Ahmed Hagag for his advice and support while working on this project. We deeply appreciate his valuable advice. Next, we want also to acknowledge all the support we have received form our faculty “Faculty of Computers & Artificial Intelligence” while working on this project. Finally, it has been a pleasure knowing everyone we worked with during that period: all the people from our department and from the other departments with whom we shared so many good experiences.