How to Cite:
Saritha, B., & Eswaran, K. (2022). Accurate estimation of liver disease using the new enhanced machine learning algorithm. *International Journal of Health Sciences*, *6*(S2), 5144–5150. https://doi.org/10.53730/ijhs.v6nS2.6328

Accurate estimation of liver disease using the new enhanced machine learning algorithm

B. Saritha
MVSR Engineering College

K. Eswaran
Altech power and energy systems pvt. Ltd. & Sreenidhi Institute of Science and Technology

***Abstract*---**Chronic liver syndrome is a major cause of death and disease worldwide. It ensues all over the world regardless of age, gender, area or race. Cirrhosis is an final consequence of a lot of liver sicknesses characterised with fibrosis and architectural alteration of the liver with the development of the liver with the formation of renewing swellings and can have different medical exhibitions and impediments. Liver is the largest solid organ of the Human body. It plays an important role in transferring blood throughout our body. It helps in the metabolism of alcohol, drugs and destroys toxic substances. It's been doctors concern to diagnosis the irregular functionality of liver at its initial stage which can increase the patient's survival rate.Skilled physicians are needed for varied examination tests to diagnose the liver unhealthiness, however it cannot assure the accurate diagnosis. Machine learning offers a guarantee for improving the detection and prediction of disease that has made an interest in the biomedical field. The aim of this paper is to recommend and prove best machine learning models which characterize the most productive one. For this, a variant of the new machine learning algorithm " KE Sieve " is proposed. It's been observed that the 'Enhanced KE Sieve ' machine Learning algorithm has given best accuracy of 95% with different split ratios when compared to KES in very less time and fewer leftovers.

***Keywords*---**KES, non-iterative, pattern recognition, machine learning, LFT.

Introduction

The liver plays a vital role in various bodily functions because of the production of protein and blood clotting to fats or fatty acids, glucose levels in the body and iron

Manuscript submitted: 27 Feb 2022, Manuscript revised: 18 March 2022, Accepted for publication: 09 April 2022

metabolic rate. Liver gets upset due to various reasons, for instance, intake of excessive medicine which contains acetaminophen and its combination drugs like few pain relief tablets named Norco, Vicodin etc., as well as tablets that can regulate the types of cholesterol (good and bad), cirrhosis, alcohol addicts, hepatitis A, B, C, D, and E, communicable mononucleosis, excess iron in the body and nonalcoholic fatty liver disease.

Liver disease is defined as any disruption of liver function. Usually, more than 75% or three quarters of liver tissue would be affected before a decrease in function occurs. The common causes of liver inflammations are Alcoholic liver disease, Autoimmune hepatitis, Chronic viral hepatitis, Hemochromatosis, Medication toxicity, Nonalcoholic fatty liver disease. Symptoms of liver sicknesses encompass weakness and exhaustion, loss of weight , nausea, vomiting, and yellow jaundice (skin discoloration), abdominal pain and inflammation, irritated skin, puffiness in the ankles and legs, loss of appetite, dark urine color, tar-colored and blood colored stool and easy bruising. The liver is the only body part that can certainly substitute the cells which are damaged. In absence of adequate cells, the requisites of the body cannot be met by the liver. Therefore in many cases liver disease can be either manageable or irreversible. The medication of liver sickness is subjected to its cause and the condition of the patient.

According to W.H.O India's 3.00% of total deaths in the year 2018 is due to Liver Disease. In the world, India stands at 62nd place whose age adjusted death rate is twenty three per each lakh of population [9][10]. Hence, this paper proposes a new AI-based classification Algorithm, 'Enhanced KE Sieve model' which is the modification of a new algorithm 'KE Sieve' [1] [2] for the detection of unhealthy liver which will enable the process of early findings more accurately.

Need of early diagnosis:

1. Early diagnosis of liver disease is certainly not easy, though it is one of the most important causes of deaths on earth.
2. At the advanced stage of liver disease the survival probability is very less, hence it has become essential to diagnose the disease at the early stage of suffering.
3. Artificial Intelligence techniques like Machine learning offer a guarantee for improving the detection and prediction of disease that has made an interest in the biomedical field .
4. Classification and prediction model supports medical diagnosis which helps in rapid initial diagnosis, reducing the risk factor.

So, an intuitive screening would help not only India but also countries around the world to rapidly screen and avoid the worst stages of suffering that can occasionally lead to an irreversible stage that may result in death in a very short period. The uniqueness of this paper is précised as follows:

1. The proposed model has an end-to-end scheme with no human intervention [3][1].
2. Liver function tests are the blood testing tool used to detect liver disease.
3. The proposed model classifies healthy and unhealthy liver patients.

4. In the future, on a daily basis, if the training data is available, the model can be incrementally trained avoiding the necessity to train from the start. This leads to a better model.

Data Set Description

Liver based clinical parameters are considered as the dataset. Pre-processing has been done on it as per the requirement. The real-time dataset of LFT is taken from 'Medcis Path Labs', Hyderabad.

Existing Model

The KE Sieve algorithm [1] [2] is applied to the identified Datasets [3] [4]

Architecture Of The KE Sieve Algorithm

Basically the KE Sieve algorithm is a Non-Iterative, where each point is run only once to find its quadrant.

- Initial planes are computed according to 2^n >=Number of dimensions or attributes of dataset Where n is number of initial planes to compute and draw
- Consider 'G' and 'S' space as referred to in paper[1] , as described in that, Orientation Vectors (OV's) are to be computed.
- In the process of finding suitable quadrants in 'S' Space[1] for a data point, If a similar point is discovered, then this data point is to be moved to an array which is of n dimensions. Else point is placed in the same quadrant and each quadrant would accommodate at most 2 data points only.
- After the above mentioned array is filled with the size of dimension n, a new plane is constructed and drawn in the S plane that passes through the midpoint of each pair of neighbors collected in the above mentioned array. This new hyper plane is enough to separate all neighbors in S space.
- As and when a new plane is constructed, OV's[1] of each data point is to be updated as the existing size of the number of planes.
- Step Number 3, 4 and 6 continues till the last plane .
- This training phase of KES completes with the outcome of a certain number of leftovers.
- From this step the testing phase starts ,
- Consider one test data point, substitute it in all planes of training phase to derive its OV
- Find nbest value [1][2][3], where the difference between test data point OV and all OVs in the training phase is calculated and sorted.
- Now fix the nbest value , for instance if nbest is 0.1, and outcomes are 100, then first 10 values are considered
- Fix Nearest Neighbor value and computation goes on,

  - Euclidean distance of test data point and 'nbest' mapped trained data points are calculated to find suitable quadrants for test data point.
  - Sort the outcomes
  - The voting process will go on and the maximum voted or numbered class is assigned to the test point, hence Predicted.
- Step 8 to 11 is iterated for every test sample of the test dataset.

Proposed Model: "Enhanced KE Sieve" Algorithm

The KES separates every 2 points by at least one plane, whereas the enhanced KES separates the two points by more than one plane. This depends on the number of passes. If there are 5 passes then the two points are separated by at least 5 planes. In higher dimensions of data, this makes it easier to find the nearest neighbor with more planes separating the points[9]. This new algorithm has been added with a new feature to the existing algorithm [1][2], that is 'the concept of pass'. The existing algorithm, as described in the above process, will be able to run training and testing phases for once only. In the training phase, all the data points are not trained to build the training model as expected, resulting in few leftovers. Therefore these leftovers are not used in the constructing process of the training model. Because of this the accuracy in testing phase may come down which is evident in table 1. To overcome or to handle this issue 'Pass' the new feature is added to this existing algorithm.

This new feature 'Pass' is ensuring that all the samples are used to construct the train model. When this enhanced KES algorithm is executed, the number of left out samples are utmost one or two only, eventually improving the performance. This process of 'Enhanced KES' algorithm is shown in figure 1 in detail as a flowchart.

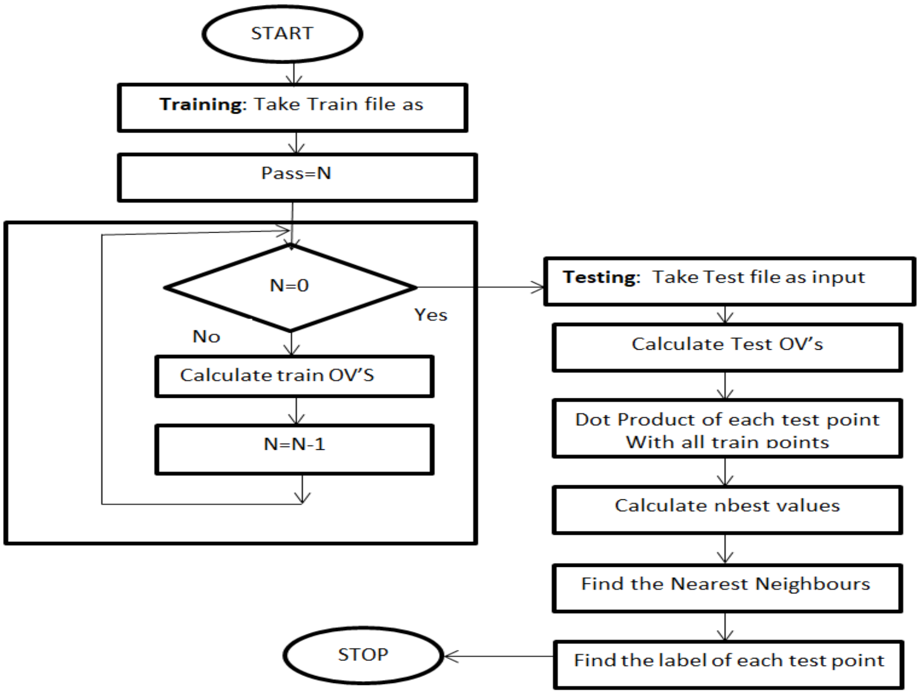

Figure 1: Flowchart of Enhanced KES Algorithm

Architecture of the "Enhanced KES" Algorithm,

- In this algorithm the user is supposed to declare the number of passes to be executed. Pass I, runs from step 2 to 6 of KES, resulting in a certain number of plane constructions and few left overs. The same steps are repeated for every Pass.
- According to the number of Passes specified, each pass run in the training phase gives planes and leftover data points as outcome. Every Pass results in construction of 'm' planes, after every pass the planes are added to the previous pass's plane. By the end of the last pass, these planes get consolidated and the number of leftovers will be very less. At this stage it's understood that the algorithm doesn't need to run another pass in the training phase.
- To execute each pass , the steps 9 to 12 of the original algorithm KES as described above are to be followed.
- In this 'Enhanced KES' the final outcome 'planes' can be assumed like a mesh as seen[1], This is an incremental property of the KES algorithm which has the probability of accommodating next samples if added for training.

Experimental Results

In order to prove the improvement of model performance in "Enhanced KES " algorithm when compared to "KE Sieve" algorithm, Training evaluation is carried out on both the algorithms using the Dataset "LFT"[6].

The Accuracy performance of the KES Algorithm is 92.56% . When the same data set is applied to "Enhanced KES Algorithm" the results acquired is 95%. The KE Sieve algorithm [1][2] has already been compared with other algorithms [][] therefore in this paper we are comparing this new enhanced algorithm only with the KE Sieve algorithm[1][2].

Table. 1: comparison of Outcomes like final planes, Leftovers and Accuracies between 'KE Sieve' and 'Enhanced KE Sieve'

KE Sieve						Enhanced KE Sieve					
Initial planes	Nearest Neighbour	nbest	Final planes	leftovers	Accuracy(%)	Passes	Nearest Neighbour	nbest	Final planes	leftovers	Accuracy(%)
4	5	0.3	30	36	90.77	2	5	0.2	60	1	94.77
5	5	0.3	29	26	91.39	2	5	0.3	89	1	93.41
6	5	0.2	29	29	91.18	2	5	0.1	59	1	94.54
7	3	0.2	29	30	91.39	2	5	0.1	57	1	95.00
8	3	0.4	31	27	90.77	2	5	0.3	59	1	94.09
9	5	0.7	31	27	92.00	2	3	0.2	62	2	94.77
10	5	0.5	30	26	92.56	2	5	0.1	62	3	95.00
11	7	0.5	28	26	92.00	2	5	0.2	64	2	94.32

Conclusion

This paper focuses on the variant of the existing algorithm KE Sieve in the direction of enhancing it with a new feature called 'Pass' in the training phase. With this enhancement, the minimum number of planes separating the train points increases with the number of passes defined. If the passes are n, then the two train points are separated by at least n planes, where n is the user defined value. This when applied to data with more number of passes, it is observed that almost all data points were involved in building a proper training model except one or two which can be seen in the latter part of table 1, Resulting in boosting the accuracy. In this scenario the highest accuracy observed is 95% at 2nd pass itself. Hence it is concluded that the 'Enhanced KES' has better results at its 2nd pass when compared to 'KES' algorithm which is ensuring the concept of Non-Iterative algorithm.

Acknowledgement

I would like to thank Medcis Path Labs, Hyderabad for providing the real time data, ALPES.ai (Altech Power and Energy Systems Pvt. Ltd.) for supporting in building the proposed model and Sai Thejeshwar (ALPES.ai) for his continuous assistance and inputs during building the project.

References

1. K Eswaran. "A non iterative method of separation of points by planes in n dimensions and its application". arXiv:1509.08742v5, 2015.
2. K Eswaran, "An Universal Classifier for Learning and Classification of data with uses in Machine Learning". Publication Number - WO 2017/191648 A1 and Published on 09 November 2017. Filed International Patent with International Patent Geneva, Switzerland.

3. K Eswaran. “On non-iterative training of a neural classifier Part-I: Separation of points by planes.” 10.1109/IntelliSys.2017.8324238, 2017.
4. Anju Gulia, Dr. Rajan Vohra , Praveen Rani “Liver Patient Classification Using Intelligent Techniques“, International Journal of Computer Science and Information Technologies, Vol. 5 (4) , 2014, 5110-5115.
5. Paritosh Khubchandani, Kundan Jha , Rohit Bijani , Suraj Lala , Pallavi Saindane ,“Medical Prediction using Artificial Intelligence”, ISSN XXXX XXXX © 2017 IJESC ,Vol. 7,issue no.3
6. B.Saritha, S.V. Ramana NarraManaswini, Rama Priyanka D.Hiranmayi, K.Eswaran. “Classification of Liver Data using a New Algorithm”, www.ijetsr.com ISSN 2394 – 3386 Volume 4, Issue 9 September 2017.
7. C. Kavya Reddy, G. Gayatri Likhita, D. Hiranmayi, B.Saritha, K.Eswaran. “Classification of Diseased Plants using Separation of Points by Planes”, International Journal of Engineering Technology Science and Research, www.ijetsr.com, ISSN 2394 – 3386 Volume 4, Issue 9 September 2017.
8. S Sai Thejeshwar, Chaitanya Chokkareddy, K Eswaran.’Precise Prediction of COVID-19 in Chest X-Ray Images Using KE Sieve Algorithm’, https://doi.org/10.1101/2020.08.13.20174144.
9. B.Saritha , K Eswaran, “A novel approach using Machine Learning Algorithm on healthy liver” , International journal of Design Engineering, ISSN: 0011-9342 | Year 2021 Issue: 8 | Pages: 7020- 7032.
10. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7376596/
11. https://www.worldlifeexpectancy.com/india-liver-disease
12. https://api.alpes.ai/liverprediction/