PROJECT REPORT

On

Liver Disease Prediction

Submitted to Centurion University of Technology& Management
in partial fulfillment of the requirement for award of the degree of

B. TECH.

in

COMPUTER SCIENCE & ENGINEERING

Submitted By

1. **Vicky kumar** **210101120004**
2. **Shubham kumar** **210101120017**
3. **Manish kumar** **210101120013**
4. **Mukesh kumar Pandey** **210101120003**

Under the Guidance of

Dr. Dhableshwar Rao

DEPT. OF COMPUTER SCIENCE & ENGINEERING

SCHOOL OF ENGINEERING &TECHNOLOGY,
CUTM, Paralakhemundi-761200

CERTIFICATE

This is to be certified that the minor project entitled "**Liver Disease Prediction**" has been submitted for the Bachelor of Technology in Computer Science Engineering of School of Engineering & Technology, CUTM, Paralakhemundi during the academic year 2021-2022 is a persuasive piece of project work carried out by "**VICKY KUMAR(210101120004), Shubham kumar(210101120017), Manish kumar(210101120013), Mukesh kumar Pandey(210101120003)**" towards the partial fulfillment for award of the degree (B.Tech.) under the guidance of "Dr. Dhableshwar Rao" and no part thereof has been submitted by them for any degree to the best of my knowledge.

Signature of HOD
Dr. Debendra Maharana

Signature of Project Guide
Dr. Dhawaleswar Rao

EVALUATION SHEET

1. Title of the Project: Liver Disease Prediction
2. Year of submission: 2022
3. Name of the degree: B-Tech
4. Date of Examination / Viva: 14th JULY 2022
5. Student Name with Reg No.:

6. Vicky kumar	210101120004
7. Shubham kumar	210101120017
8. Manish kumar	210101120013
9. Mukesh kumar Pandey	210101120003

10. Name of the Guide: **Dr. Dhawaleswar Rao**
11. Result:

[APPROVED/REJECTED]

Signature of HOD
Dr. Debendra Maharana

Signature of Project Guide
Dr. Dhawaleswar Rao

Signature of External Examiner

CANDIDATE'S DECLARATION

We "**VICKY KUMAR (210101120004), Shubham kumar (210101120017), Manish kumar (210101120013), Mukesh kumar Pandey (210101120003)**", B-Tech in CSE (Semester- II) of School of Engineering &Technology, CUTM, Paralakhemundi, hereby declare that the Project Report entitled "LIVER DISEASE PREDICTION" is an original work and data provided in the study is authentic one. This report has not been submitted to any other Institute for the award of any other degree by me.

Signature of Student

ACKNOWLEDGEMENT

I would like to thank my teacher Dr. Dhawaleswar Rao who gave me this opportunity to work on this project. I got to learn a lot from this project about Liver Disease Prediction.

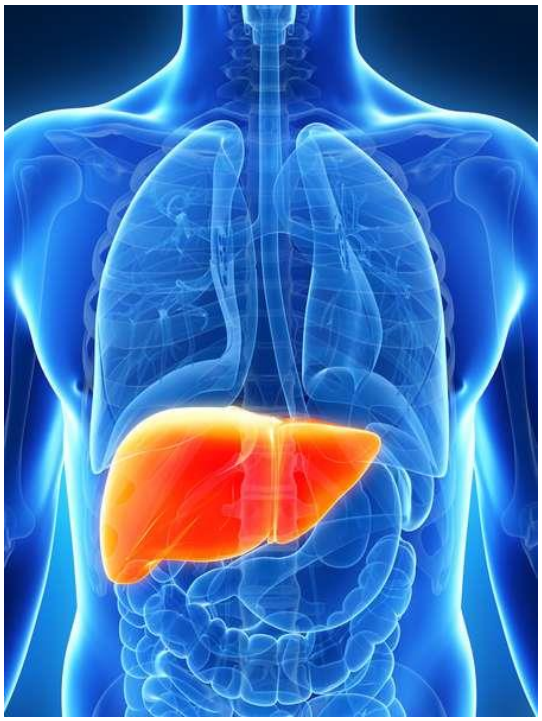
I would also like to thank our Branch H.O.D Debendra Maharana

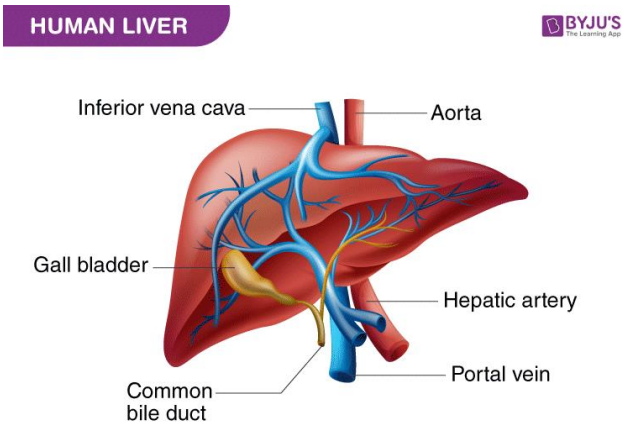
At last, I would like to extend my heartfelt thanks to my parents because without their help this project would not have been successful. Finally, I would like to thank my dear friends who have been with me all the time.

STUDENT NAMES & REG No.

Vicky kumar	210101120004
Shubham kumar	210101120017
Manish kumar	210101120013
Mukesh kumar Pandey	210101120003

Liver Disease Prediction





INDEX:

Abstract: -

In recent years in healthcare sectors, data mining became an ease of use for disease prediction. Data mining is the process of dredge up information from the massive datasets or warehouse or other repositories. It is a very challenging task to the researchers to predict the diseases from the voluminous medical databases. To overcome this issue the researchers use data mining techniques such as classification, clustering, association rules and so on. The main objective of this research work is to predict liver diseases using classification algorithms. The algorithms used in this work are Naïve Bayes and support vector machine (SVM). These classifier algorithms are compared based on the performance factors i.e., classification accuracy and execution time. From the experimental results it is observed that the SVM is a better classifier for predict the liver diseases.

The liver is the largest solid organ and the largest gland in the human body. It carries out over 500 essential tasks. It is the only organ that can regenerate self-cells.

The liver is essential for digesting food and ridding your body of toxic substances.

It sits just under your rib cage on the right side of your abdomen. The liver is essential for digesting food and ridding your body of toxic substances.

Introduction: -

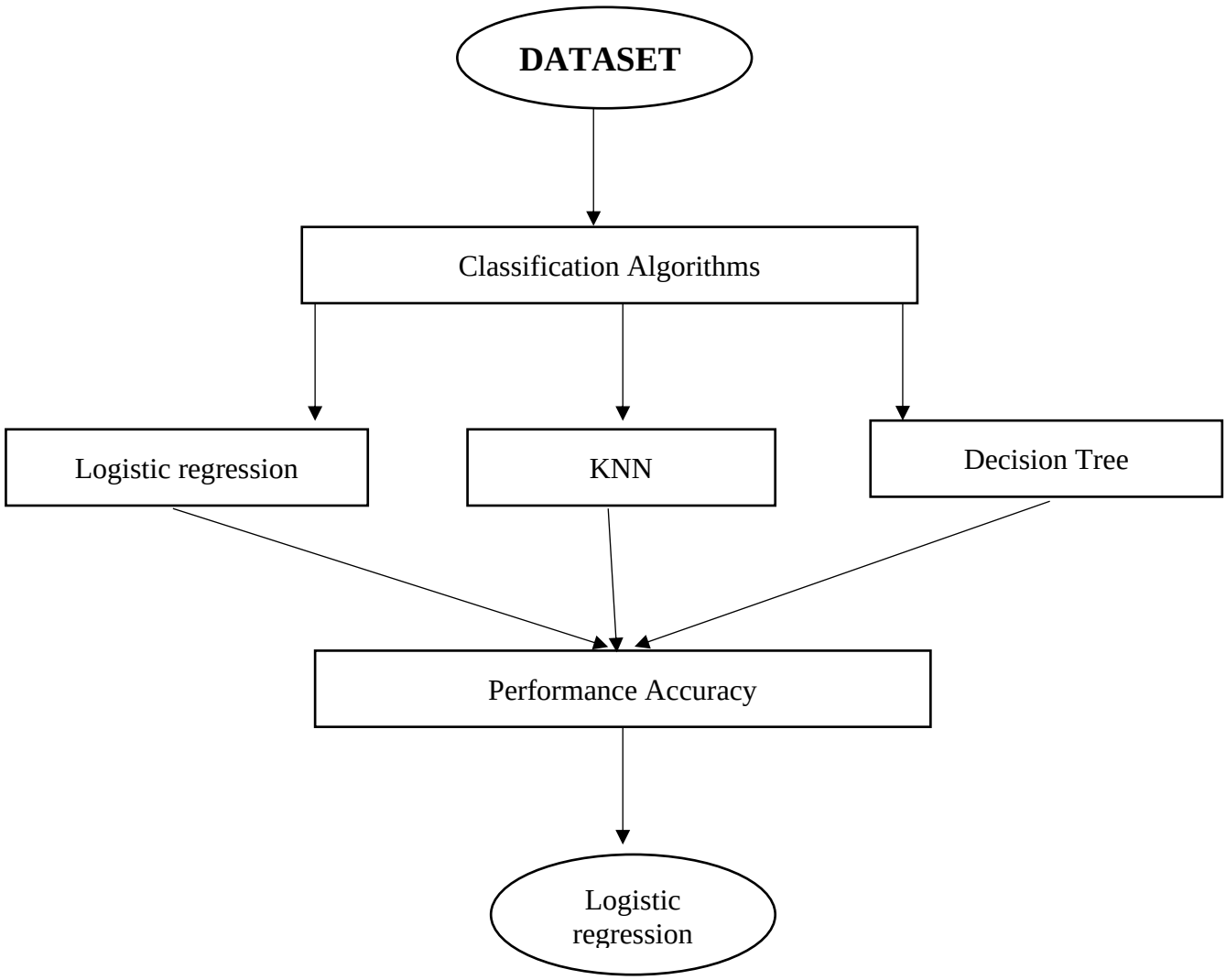
Researchers faces more challenging task in healthcare sectors to predict the diseases from the voluminous medical databases. Nowadays data mining became more essential in healthcare sectors. Data mining techniques which includes classification, clustering, association rule mining for finding frequent patterns are applied to medical data for disease prediction. In data mining, classification techniques are much popular in medical diagnosis and predicting diseases [1]. In this research work, Naïve Bayes and Support Vector Machine (SVM) classifier algorithms are used for liver disease prediction. There are several numbers of liver disorders that required clinical care of the physician [3]. The main objective of this research work is to predict liver diseases such as Cirrhosis, Bile Duct, Chronic Hepatitis, Liver Cancer and Acute Hepatitis from Liver Function Test (LFT) dataset using above classification algorithms. The liver is the second largest internal organ in the human body, playing a major role in metabolism and serving several vital functions, e.g. Decomposition of red blood cells, etc.,. [7] Its weight comes around three pounds. The liver performs many essential functions related to digestion, metabolism, immunity,

and the storage of nutrients within the body. These functions make the liver as an important organ, without this, body tissues would quickly die from lack of energy and nutrients. There are number of factors which increase the risk of liver disease. Some of them are listed below:

- Family history of liver disease
- Smoking
- Consumption of alcohol
- Intake of contaminated food
- Obesity
- Diabetes

The remaining portion of the paper is organized as follows. Related works are discussed in Section 2. The proposed methodology is given in Section 3. Section 4 analyzes the experimental results. Section 5 gives conclusion..

Methodology: -



Related Work: -

Here some related work done by others authors of different country. These works are very similar to our project.

Sl. No.	Title	Author(s)
1	Liver disease prediction using machine learning	Vasan Durai Suyan Ramesh Dinesh Kalthireddy
2	A Comparative Study On Liver Disease Prediction Using Supervised Machine Learning Algorithms	A.K.M Sazzadur Rahman, F. M. Javed Mehedi Shamrat, Zarrin Tasnim, Joy Roy, Syed Akhter Hossain
3	Liver Disease Prediction using SVM and Naïve Bayes Algorithms	Dr. S. Vijayarani , Mr.S.Dhayanand
4	LIVER DISEASE PREDICTION BY USING DIFFERENT DECISION TREE TECHNIQUES	Nazmun Nahar Ferdous Ara
5	Prediction of Liver Disease using Classification Algorithms	Thirunavukkarasu K, Ajay S. Singh, Md Irfan, Abhishek Chowdhury.
6	Liver Disease Prediction Using Machine Learning Algorithm	Sambit Mohanty, Pradosh Kumar Gantayat, Sachikanta Dash, Bhabani P. Mishra, and Shiba Ch. Barik
7	A Comparative Analysis of the Ensemble Method for Liver Disease Prediction	Vicky Barua,Nazmun Naha, Ferdous Ara,Md. Arif Istiek Neloy, Karl Andersson, Mohammad Shahadat Hossain
8	Prediction of Liver Diseases Based on Machine Learning Technique for Big Data	Engy A. El-Shafeiy, Ali I. El-Desouky and Sally M. Elghamrawy
9	Prediction of Liver Diseases by Using Few Machine Learning Based Approaches	Md. Shafiul Azam1 , Aishe Rahman1 , S. M. Hasan Sazzad Iqbal1 , and Md. Toukir Ahmed
10	Survey of Machine Learning Algorithms for Disease Diagnostic	Meherwar Fatima, Maruf Pasha

11	Pathophysiology and Mechanisms of Nonalcoholic Fatty Liver Disease	1. Staels, B.
12	Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques	Mohan, S., Thirumalai, C., & Srivastava, G
13	Prediction and Analysis of Liver Disorder Diseases by using Data Mining Technique: Survey	Kefelegn, S., & Kamat, P.
14	Prediction of nonalcoholic fatty liver disease via a novel panel of serum adipokines	Jamali, R., Arj, A., Razavizade, M., & Aarabi, M. H
15	An Extended Fatty Liver Index to predict Nonalcoholic Fatty Liver Disease	4. Chatterjee, A., Basu, A., Chowdhury, A., Das, K., Sarkar-Roy, N., Majumder, P. P., & Basu, P.
16	An Extended Fatty Liver Index to predict Nonalcoholic Fatty Liver Disease	Kantartzis, K., Rettig, I., Staiger, H., Machann, J., Schick, F., Scheja, L., ... & Stefan, N.
17	A Data Mining Approach to Prediction of Liver Diseases	Razali, N., Mustapha, A., Abd Wahab, M. H., Mostafa, S. A., & Rostam, S. K.
18	Magnetic Resonance Elastography Predicts Advanced Fibrosis in Patients With Nonalcoholic Fatty Liver Disease:	Loomba, R., Wolfson, T., Ang, B., Hooker, J., Behling, C., Peterson, M., ... & Sirlin, C.
19	Diagnosis of Nonalcoholic Fatty Liver Disease: Invasive versus Noninvasive	Wieckowska, A., & Feldstein, A. E
20	Predictive factors of early postoperative gra function in human liver transplantation	9. González, F. X., Rimola, A., Grande, L., Antolin, M., Garcia-Valdecasas, J. C., Fuster, J., ... & Rodés, J.
21	A Comparative Analysis of the Ensemble Method for Liver Disease Prediction	Nazmun Nahar

22	Liver Disease Prediction using SVM and Naïve Bayes Algorithms	Dr. S. Vijayarani1 , Mr.S.Dhayanand2
23	LIVER DISEASE PREDICTION BY USING DIFFERENT DECISION TREE TECHNIQUES	Nazmun Nahar1 and Ferdous Ara2
24	A Comparative Study On Liver Disease Prediction Using Supervised Machine Learning Algorithms	A.K.M Sazzadur Rahman, F. M. Javed Mehedi Shamrat, Zarrin Tasnim, Joy Roy, Syed Akhter Hossain
25	Performance Analysis of Liver Disease Prediction Using Machine Learning Algorithms	M. Banu Priya1, P. Laura Juliet2, P.R. Tamilselvi3
26	Strategic Analysis in Prediction of Liver Disease Using Different Classification Algorithms	Binish Khan1* , Piyush Kumar Shukla 2 , Manish Kumar Ahirwar 3
27	Model for End-Stage Liver Disease-Lactate and Prediction of Inpatient Mortality in Patients With Chronic Liver Disease	Naveed Sarmast,1 * Gerald O. Ogola
28	Evaluation and comparison of six noninvasive tests for prediction of significant or advanced fibrosis in nonalcoholic fatty liver disease	Katharina Staufer1
29	Prediction and validation of nonalcoholic fatty liver disease by fatty liver index in a Japanese population	Satoko Takahashi
30	A case-control study on insulin resistance, metabolic co-variates & prediction score in non-alcoholic fatty liver disease	S. Bajaj, P. Nigam* , A. Luthra
31	Liver disease prediction using machine learning	Rabindra Lamsal and Ayesha Choudhary
32	International Journal of Advance Research, Ideas and Innovations in Technology	Vasan Durai

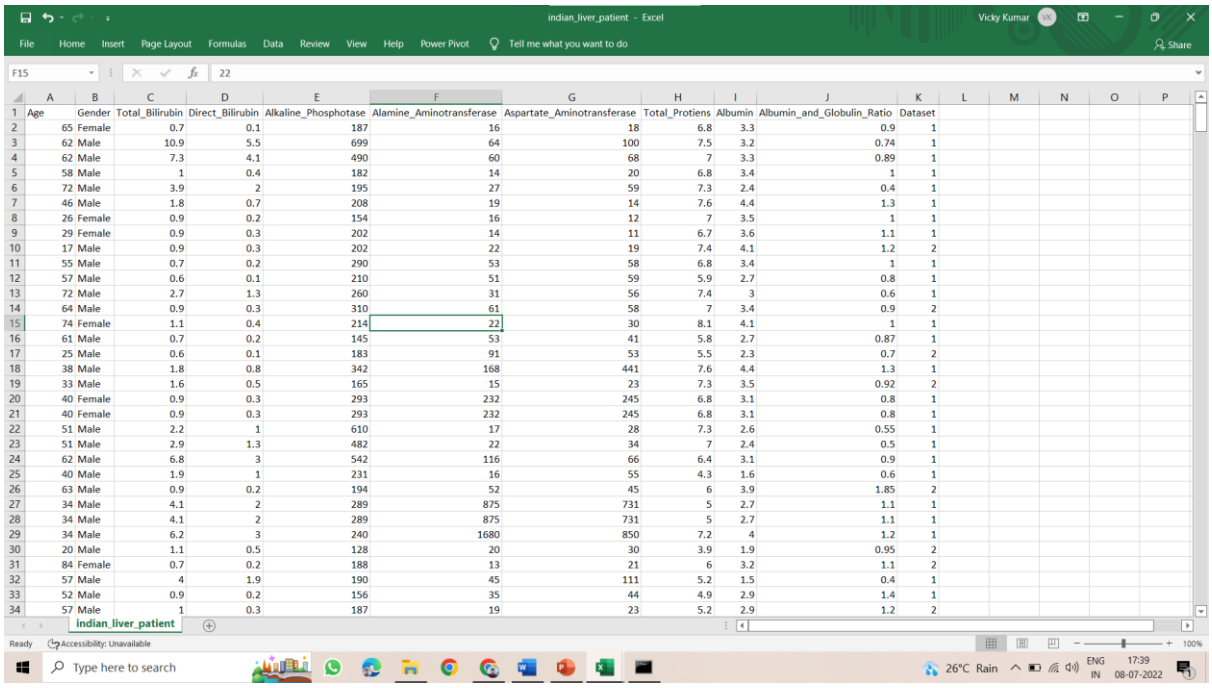
33	REVIEW OF TRENDS AND RELATED LITERATURE	Anu Sebastian
34	Prediction of liver cancer using Conditional probability Bayes theorem	N. Ramkumar
35	“Rule Optimization of Boosted C5.0 Classification	Mafazalyaqeen Hassoon
36	A Comparative Study On Liver Disease Prediction	Sambit Mohanty
37	DESCRIPTION OF THE CLASSIFICATION ALGORITHMS	A.K.M Sazzadur Rahman,
38	Performance Analysis of Liver Disease Prediction Using Machine Learning Algorithms	M. Banu Priy
39	propagation neural network and radial basis	Ebenezer Obaloluwa Olaniyi
40	Prediction of fatty liver disease using machine learning algorithms	Mafazalyaqeen Hassoon

Data set: -

This data set contains 416 liver patient records and 167 non liver patient records collected from North East of Andhra Pradesh, India. The "Dataset" column is a class label used to divide groups into liver patient (liver disease) or not (no disease). This data set contains 441 male patient records and 142 female patient records.

Columns:

- Age of the patient
- Gender of the patient
- Total Bilirubin
- Direct Bilirubin
- Alkaline Phosphatase
- Alamine Aminotransferase
- Aspartate Aminotransferase
- Total Proteins
- Albumin
- Albumin and Globulin Ratio
- Dataset: field used to split the data into two sets (patient with liver disease, or no disease)



indian_liver_patient - Excel

Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protiens	Albumin	Albumin_and_Globulin_Ratio	Dataset
65	Female	0.7	0.1	187	16	18	6.8	3.3	0.9	1
62	Male	10.9	5.5	699	64	100	7.5	3.2	0.74	1
62	Male	7.3	4.1	490	60	68	7	3.3	0.89	1
58	Male	1	0.4	182	14	20	6.8	3.4	1	1
72	Male	3.9	2	195	27	59	7.3	2.4	0.4	1
46	Male	1.8	0.7	208	19	14	7.6	4.4	1.3	1
26	Female	0.9	0.2	154	16	12	7	3.5	1	1
29	Female	0.9	0.3	202	14	11	6.7	3.6	1.1	1
17	Male	0.9	0.3	202	22	19	7.4	4.1	1.2	2
55	Male	0.7	0.2	290	53	58	6.8	3.4	1	1
57	Male	0.6	0.1	210	51	59	5.9	2.7	0.8	1
72	Male	2.7	1.3	260	31	56	7.4	3	0.6	1
64	Male	0.9	0.3	310	61	58	7	3.4	0.9	2
74	Female	1.1	0.4	214	22	30	8.1	4.1	1	1
61	Male	0.7	0.2	145	53	41	5.8	2.7	0.87	1
25	Male	0.6	0.1	183	91	53	5.5	2.3	0.7	2
38	Male	1.8	0.8	342	168	441	7.6	4.4	1.3	1
33	Male	1.6	0.5	165	15	23	7.3	3.5	0.92	2
40	Female	0.9	0.3	293	232	245	6.8	3.1	0.8	1
40	Female	0.9	0.3	293	232	245	6.8	3.1	0.8	1
51	Male	2.2	1	610	17	28	7.3	2.6	0.55	1
51	Male	2.9	1.3	482	22	34	7	2.4	0.5	1
62	Male	6.8	3	542	116	66	6.4	3.1	0.9	1
40	Male	1.9	1	231	16	55	4.3	1.6	0.6	1
63	Male	0.9	0.2	194	52	45	6	3.9	1.85	2
34	Male	4.1	2	289	875	731	5	2.7	1.1	1
34	Male	4.1	2	289	875	731	5	2.7	1.1	1
34	Male	6.2	3	240	1680	850	7.2	4	1.2	1
20	Male	1.1	0.5	128	20	30	3.9	1.9	0.95	2
84	Female	0.7	0.2	188	13	21	6	3.2	1.1	2
57	Male	4	1.9	190	45	111	5.2	1.5	0.4	1
52	Male	0.9	0.2	156	35	44	4.9	2.9	1.4	1
57	Male	1	0.3	187	19	23	5.2	2.9	1.2	2

Conclusion: -

Classification is the major data mining technique which is primarily used in healthcare sectors for medical diagnosis and predicting diseases. This research work used classification algorithms namely Naïve bayes and Support Vector Machine (SVM) for liver disease prediction. Comparisons of these algorithms are done and it is based on the performance factors classification accuracy and execution time. From the experimental results, this work concludes, the SVM classifier is considered as a best algorithm because of its highest classification accuracy. On the other hand, while comparing the execution time, the Naïve Bayes classifier needs minimum execution time.

REFERENCES: -

1) *Durai, Vasan, Suyan Ramesh, and Dinesh Kalthireddy. "Liver disease prediction using machine learning." Int. J. Adv. Res. Ideas Innov. Technol 5.2 (2019): 1584-1588.*
2) *Rahman, AKM Sazzadur, et al. "A comparative study on liver disease prediction using supervised machine learning algorithms." International Journal of Scientific & Technology Research 8.11 (2019): 419-422.*
3) *Vijayarani, S., and S. Dhayanand. "Liver disease prediction using SVM and Naïve Bayes algorithms." International Journal of Science, Engineering and Technology Research (IJSETR) 4.4 (2015): 816-820.*
4) *Nahar, Nazmun, and Ferdous Ara. "Liver disease prediction by using different decision tree techniques." International Journal of Data Mining & Knowledge Management Process 8.2 (2018): 01-09.*
5) *Singh, Ajay S., Md Irfan, and Abhishek Chowdhury. "Prediction of liver disease using classification algorithms." 2018 4th international conference on computing communication and automation (ICCCA). IEEE, 2018.*
6) *Mohanty, Sambit, et al. "Liver Disease Prediction Using Machine Learning Algorithm." Data Engineering and Intelligent Computing. Springer, Singapore, 2021. 589-596.*
7) *Nahar, Nazmun, et al. "A comparative analysis of the ensemble method for liver disease prediction." 2019 2nd International Conference on Innovation in Engineering and Technology (ICIET). IEEE, 2019.*

8) *El-Shafeiy, Engy A., Ali I. El-Desouky, and Sally M. Elghamrawy. "Prediction of liver diseases based on machine learning technique for big data." International conference on advanced machine learning technologies and applications. Springer, Cham, 2018.*
9) *Azam, Md Shafiul, et al. "Prediction of liver diseases by using few machine learning based approaches." Aust. J. Eng. Innov. Technol 2.5 (2020): 85-90.*
10) *Fatima, M. and Pasha, M. (2017) Survey of Machine Learning Algorithms for Disease Diagnostic. Journal of Intelligent Learning Systems and Applications, 9, 1-16. doi: 10.4236/jilsa.2017.91001.*
11) *Staels, B. (2016). Pathophysiology and mechanisms of nonalcoholic fatty liver disease.* Annu. Rev. Physiol*,* 78*, 181-205.*

12) *Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques.* IEEE access*,* 7*, 81542-81554..*

13) *Kefelegn, S., & Kamat, P. (2018). Prediction and analysis of liver disorder diseases by using data mining technique: survey.* International Journal of pure and applied mathematics*,* 118*(9), 765-770.*

14) *Jamali, R., Arj, A., Razavizade, M., & Aarabi, M. H. (2016). Prediction of nonalcoholic fatty liver disease via a novel panel of serum adipokines.* Medicine*,* 95*(5).*

15) *Chatterjee, A., Basu, A., Chowdhury, A., Das, K., Sarkar-Roy, N., Majumder, P. P., & Basu, P. (2015). Comparative analyses of genetic risk prediction methods reveal extreme diversity of genetic predisposition to nonalcoholic fatty liver disease (NAFLD) among ethnic populations of India.* journal of genetics*,* 94*(1), 105-113.*

16) *Kantartzis, K., Rettig, I., Staiger, H., Machann, J., Schick, F., Scheja, L., ... & Stefan, N. (2017). An extended fatty liver index to predict non-alcoholic fatty liver disease.* Diabetes & metabolism*,* 43*(3), 229-239.*

17) *Razali, N., Mustapha, A., Abd Wahab, M. H., Mostafa, S. A., & Rostam, S. K. (2020, April). A data mining approach to prediction of liver diseases. In* Journal of Physics: Conference Series *(Vol. 1529, No. 3, p. 032002). IOP Publishing.*

18) *Loomba, R., Wolfson, T., Ang, B., Hooker, J., Behling, C., Peterson, M., ... & Sirlin, C. (2014). Magnetic resonance elastography predicts advanced fibrosis in patients with nonalcoholic fatty liver disease: a prospective study.* Hepatology*,* 60*(6), 1920-1928.*

19) *Wieckowska, A., & Feldstein, A. E. (2008, November). Diagnosis of nonalcoholic fatty liver disease: invasive versus noninvasive. In* Seminars in liver disease *(Vol. 28, No. 04, pp. 386-395). © Thieme Medical Publishers.*

20) *González, F. X., Rimola, A., Grande, L., Antolin, M., Garcia-Valdecasas, J. C., Fuster, J., ... & Rodés, J. (1994). Predictive factors of early postoperative graft function in human liver transplantation.* Hepatology*,* 20*(3), 565-573.*

21) *Vijayarani, S., & Dhayanand, S. (2015). Liver disease prediction using SVM and Naïve Bayes algorithms. International Journal of Science, Engineering and Technology Research (IJSETR), 4(4), 816-820.*
22) *Nahar, N., & Ara, F. (2018). Liver disease prediction by using different decision tree techniques. International Journal of Data Mining & Knowledge Management Process, 8(2), 01-09.*
23) *Rahman, A. S., Shamrat, F. J. M., Tasnim, Z., Roy, J., & Hossain, S. A. (2019). A comparative study on liver disease prediction using supervised machine learning algorithms. International Journal of Scientific & Technology Research, 8(11), 419-422.*
24) *riya, M. B., Juliet, P. L., & Tamilselvi, P. R. (2018). Performance analysis of liver disease prediction using machine learning algorithms. International Research Journal of Engineering and Technology (IRJET), 5(1), 206-211.*
25) *Nahar, N., Ara, F., Neloy, M. A. I., Barua, V., Hossain, M. S., & Andersson, K. (2019, December). A comparative analysis of the ensemble method for liver disease prediction. In 2019 2nd International Conference on Innovation in Engineering and Technology (ICIET) (pp. 1-6). IEEE.*
26) *Khan, B., Shukla, P. K., Ahirwar, M. K., & Mishra, M. (2021). Strategic analysis in prediction of liver disease using different classification algorithms. In Handbook of Research on Disease Prediction Through Data Analytics and Machine Learning (pp. 437-449). IGI Global.*
27) *Sarmast, N., Ogola, G. O., Kouznetsova, M., Leise, M. D., Bahirwani, R., Maiwall, R., ... & Asrani, S. K. (2020). Model for end-stage liver disease-lactate and prediction of inpatient mortality in patients with chronic liver disease. Hepatology, 72(5), 1747-1757.*
28) *Staufer, K., Halilbasic, E., Spindelboeck, W., Eilenberg, M., Prager, G., Stadlbauer, V., ... & Stauber, R. E. (2019). Evaluation and comparison of six noninvasive tests for prediction of significant or advanced fibrosis in nonalcoholic fatty liver disease. United European gastroenterology journal, 7(8), 1113-1123.*
29) *Takahashi, S., Tanaka, M., Higashiura, Y., Mori, K., Hanawa, N., Ohnishi, H., & Furuhashi, M. (2022). Prediction and validation of nonalcoholic fatty liver disease by fatty liver index in a Japanese population. Endocrine Journal, 69(4), 463-471..*
30) *Bajaj, S., Nigam, P., Luthra, A., Pandey, R. M., Kondal, D., Bhatt, S. P., ... & Misra, A. (2009). A case-control study on insulin resistance, metabolic co-variates & prediction score in non-alcoholic fatty liver disease. Indian Journal of Medical Research, 129(3), 285.*