IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

An Improved Optimization Liver Disease Prediction Using Particle Swarm Optimization (PSO)

Pratik Sutar, Dr. Gitanjali Mate, Vinayak Mhaske, Omkar Jaybhaye, Yash Wagh

Student, Assistant Professor, Student, Student, Student

Department of Information Technology

JSPM Rajarshi Shahu College of Engineering, Pune, India

Abstract - Liver disease has significantly grown recently, and in many countries, it is now one of the most dangerous diseases. One of the most active and crucial organs, the liver is in charge of producing nutrients, metabolizing them, and safeguarding a non-toxic bodily environment. Its ability to remove toxins from your blood is one of its most crucial characteristics. People with liver disease are becoming more and more common as a consequence of heavy alcohol use, drugs use, polluted gas inhalation, opioids, contaminated food, and unhealthy habits. Globally, liver disease has a significant death rate. Weight loss may be some sign of liver disease. Eyes and skin that seem yellowish and leg and ankle swelling are some other signs of liver disease. To prevent the loss of lives, it is essential to identify these disorders as soon as possible.

Using techniques for classification with machine learning in the healthcare sector is one of the solutions for liver disease prediction. We face a lot of data in the healthcare industry, which is one of the challenges in analyzing and studying the target condition. Choosing features that are more important than other characteristics is a difficulty in the field of disease prediction. To enhance the performance of the most accurate models, feature selection of subsets is done.

This study compares models to predict liver disease, including Bayesian networks, MLP, SVM, Random Forest and Particle Swarm Optimization. The goal is to find specific features. The PSO model excels the other features in terms of the criteria for specificity, sensitivity, and accuracy. Early identification of liver diseases can save the loss of human life.

Keywords – Machine Learning, Particle Swarm Optimization (PSO), Disease Prediction, Liver Disease, Healthcare

I. Introduction-

Nearly every organ system in the body depends on the liver. It helps with digestion and metabolism by interacting with the endocrine and digestive systems. The liver is a dark reddish-brown, cone-shaped organ that weighs roughly 3 pounds. The liver is a vital organ of the body of a person and is responsible to the variety of functions that support immunity, digestion, detoxification, metabolism, and vitamin storage, among many other functions. About 2% of an adult's body weight is made up of it. The liver is a unique organ because of its dual blood supply from the portal vein (about 75%) and hepatic artery (about 25%).

The majority of the chemical levels in blood are regulated by the liver, which also excretes a material called bile. This helps the liver rid itself of waste. All blood that leaves the stomach and intestines ends up in the liver. The liver breaks down, balances, and produces nutrients while processing this blood. Moreover, it breaks down medications so that the body can absorb them more readily and without harming the body. The liver is an essential organ that performs more than 500 basic body functions. These include the production of essential nutrients, the regulation of blood sugar levels, and the removal of waste products and foreign substances from the bloodstream.

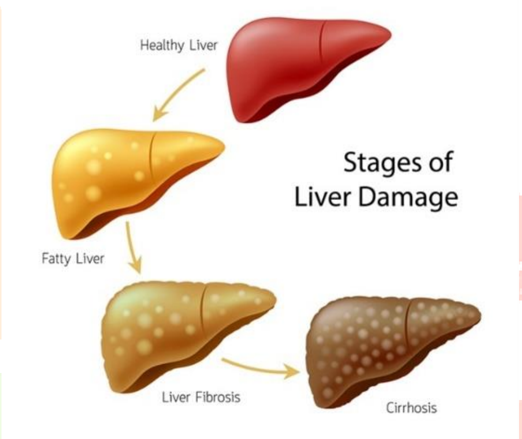


Fig 1. Levels of Liver Damage

Fig1. shows the stages of liver disease that can cause severe health conditions if not checked regularly and treated well.

Level 1: Inflammation - Early on, your liver will be swollen and occasionally painful. Perhaps it doesn't even bother you.

Level 2: Scarring and fibrosis - If treatment is not received for the inflammation, scarring will occur. Scar tissue accumulation restricts the blood flow to your liver, making the healthy parts work harder and preventing them from doing their jobs.

Level 3: Cirrhosis- The liver will function poorly or not at all as a result of the scar tissue taking over and leaving a diminishing amount of healthy tissue for it to operate on.

Level 4: Advanced liver disease or failure- This is a general term that covers a number of conditions, such as lung problems, internal bleeding, loss of renal function, enlarged liver, and fluid retention. The only treatment given is a liver transplant.

Fig2. Shows various diseases that affect people. So here we can see that after Cardiovascular diseases mostly people are dying due to liver disease and this causes the major impact on the world. Liver diseases are a big problem for public health all over the world, and they lead to a lot of people passing away. These diseases, like cirrhosis, hepatitis, and liver cancer, cause many deaths. The number of people who die from liver diseases varies in different places because of things like how much people drink, if they get infections, and how easy it is to get medical help. To overcome this problem, we need to figure out how common these diseases are and why people die from them.

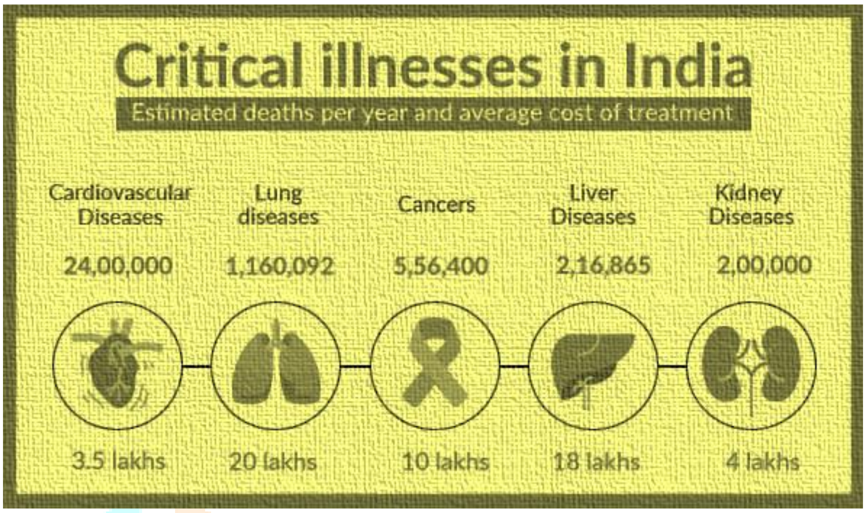


Fig2. Illness in India

II. Literature Survey –

[1] The main focus of this work is the application of machine learning algorithms for liver disease detection. J48, Artificial neural networks, and decision trees in particular have demonstrated accuracy rates with their selection depending on factors such as the dataset and features being considered. The research also acknowledges the increasing prevalence of liver diseases caused by factors like alcohol consumption and pollution. It highlights the potential of machine learning in enabling diagnosis and remote healthcare. Various machine learning techniques, including supervised, semi supervised and reinforcement learning methods like SVM, KNN, K Mean clustering, neural networks and decision trees are assessed based on metrics including specificity, preciseness, sensitiveness, and precision. The study encourages exploration of approaches and additional parameters to enhance liver disease prediction techniques, in future studies.

[2] Using machine learning approaches, this study focuses on early liver disease prediction. Accuracy, Precision, Recall, F-Measure, and AUC were among the metrics used to assess the efficacy of several machine learning models in predicting the occurrence of liver illness. These models included SVM, Logistic Regression, ANN, KNN, RT, Stacking, Bagging, and Voting. The Voting classifier performs better than other models, according to the experimental data, with an accuracy of 88.4% following SMOTE with 10-fold cross-validation, an accuracy, recall, and F-measure of 80%, precision, and 80.2%, respectively. The report makes recommendations for possible future research that could include comparing outcomes using different measures, applying deep learning techniques, and reevaluating the methodology for liver disease manifestation.

[3] This paper explores the intriguing field of machine learning algorithms and how they are used to diagnose liver disease. The researchers used the capabilities of machine learning to find significant predictors for liver illness using a dataset of 615 individuals. They used a number of methods, including principal component analysis, data visualization, and multiple imputations to handle missing data and lower the dataset's dimensionality. They used binary classification machine learning methods, such as the powerful artificial neural network, the adaptable random forest, and the dependable support vector machine, to categorize people with liver disorders. The random forest method was the one that triumphed after extensive testing. The outcome demonstrates

how well risk factors can be included when predicting liver disease.

Additionally, the study determined that vital markers for liver disease diagnosis included AST, ALT, GGT, BIL, and ALP. These discoveries open new avenues for enhanced diagnostic procedures, overcoming obstacles in the medical field, and promoting early diagnosis. In the end, there is great potential for improving patient outcomes and maximizing healthcare costs through the incorporation of machine learning approaches in detection of liver disease. In the future, these potent machine learning techniques may find use in other areas of medicine, according to the researchers' vision.

[4] This paper uses a software engineering approach to predict liver disease, addressing the pressing need for easily accessible medical services. In order to determine the risk of liver diseases, the study makes use of the Indian Liver Patient Dataset (ILPD) from the University of California, Irvine. It takes into account a number of characteristics, including age, gender, and bilirubin levels. The dataset is subjected to a number of classification algorithms, including k-nearest neighbour , Random Forest, SMO, Naive Bayes, J48, and logistic regression. Additionally, the study contrasts how well various classifiers perform when feature selection strategies are used and when they are not.

The best accuracy is achieved with shorter execution times when using techniques of feature selection in conjunction with Logistic Regression, according to the results of 10-fold cross-validation testing.

Additionally, the study assesses metrics such as mean absolute error, correctly classified instances, and the kappa statistic. Techniques used in feature selection reduce the output time of different classifiers while simultaneously improving accuracy.

A software engineering-based Intelligent Liver Disease Prediction Software (ILDPS) has also been developed and presented in this study. This highlights how software-based tools can improve liver disease prediction and diagnosis, which will ultimately increase the effectiveness and accessibility of healthcare.

This research paper concludes by highlighting the urgent need for easily accessible medical services and proving that a software engineering approach can accurately predict liver disease. The study's conclusions offer insightful information about how to increase healthcare efficiency and accessibility by creating software-based tools.

[5] The paper investigates the use of machine learning and deep learning techniques to forecast liver diseases by analyzing patient health records. They discuss the methodology and framework for predicting and diagnosing liver disorders using artificial intelligence, leveraging various patient data sources. The research evaluates algorithm performance and highlights the strengths and weaknesses of different approaches. It concludes that artificial intelligence, combined with diverse patient data, can achieve accurate predictions and diagnoses. However, the study also acknowledges the challenges in model development and algorithm evaluation. The paper suggests future directions for this field, emphasizing the potential impact on the medical industry.

The study analyses algorithm performance and identifies the benefits and drawbacks of various strategies. It comes to the conclusion that precise diagnosis and forecasts may be made when artificial intelligence and a variety of patient data are merged. The paper does recognize the difficulties in developing models and evaluating algorithms, though. The report highlights the possible influence on the medical business and recommends future options for this subject.

[6] The accurate liver disease prediction of the liver diseases through machine learning techniques is the main purpose of this study. They employ a variety of classification algorithms with patient data, including Random Forest, Perceptron, Decision Tree, K-Nearest Neighbours, and Support Vector Machine. The computational model they build reveals that KNN has the best feature selection performance. In addition to being adaptable for other diseases like diabetes and heart disease, the approach has potential applications in

community health services. Prospective avenues for investigation encompass investigating novel algorithms, eliminating noise from datasets, producing classification guidelines, and employing multiple databases for examination. The accuracy of the model has limitations, which the study acknowledges.

In future, scientists may research the integration of deep learning techniques to augment the precision of disease prognostic models. This study emphasises how machine learning and technology are still advancing medical science, with an emphasis on increasing the precision of disease prediction.

[7] The purpose of this research is to create a prototype for early detection of liver disorders utilising various machine learning methodologies. They use two datasets from the UCI ML repository and GitHub repository to test ten classifiers, including MLP, SVM, KNN, CDT, Forest-PA. The Random Forest (RF) classifier performs well with the UCI dataset, reaching an accuracy of 72.1739%. The Support Vector Machine (SVM) outperforms the others with an accuracy of 71.3551% on the GitHub dataset. These findings imply that RF and SVM have the potential to predict liver disease accurately, although their performance may differ depending on the dataset employed.

This study emphasises the need of improving accuracy and lowering error rates in liver disease prediction models. It provides valuable insights that can be used to assess the efficacy of new methodologies, models, or frameworks in predicting liver disorders. The study adds to continuing attempts to improve the early diagnosis of liver illnesses, which is critical for healthcare. Continue your excellent work!

[8] This study emphasises the vital need of detecting liver illnesses early, especially in their early stages. Healthcare workers sometimes face difficulties in predicting these disorders during their early stages, making prompt intervention a difficult endeavour. The study combines machine learning techniques, notably Support Vector Machine (SVM) and Logistic Regression, to develop classification algorithms for identifying people with and without liver diseases. These approaches are evaluated for predictive accuracy, and they achieve a significant 96% accuracy rate in predicting liver disorders. The researchers propose that future study should build on this work by comparing it to other categorization methods such as Navie Bayes or Random Forest. They also recommend investigating the implementation of parametric classifications using bio-inspired optimisation techniques.

The findings of this study contribute significantly to the early detection of liver illnesses, which is critical for advancing healthcare and improving the quality of life for those who do not have liver tumours.

[9] The Modified Convolutional Neural Network-based Liver Disease Prediction System (MCNN-LDPS) is introduced in the study to improve the accuracy of liver disease prediction. They employ dimensionality reduction approaches such as Modified Principal Component Analysis and the Score-based Artificial Fish Swarm Algorithm (SAFSA) for feature selection. SAFSA, which employs information gain and entropy values, is effective in producing precise results. The primary focus is on the Indian Liver patient dataset.

The results reveal that MCNN-LDPS beats previous approaches with a 4.05% gain in accuracy, a stunning 21.23% increase in F-measure, a noticeable 4.22% improvement in precision, and a substantial 34.26% increase in recall when compared to MLPNN. Through the integration of the evolutionary algorithm, the study effectively tackles the constraints of CNN by encoding orientational and relative spatial relationships while maintaining spatial invariance.

The CNN-LDPS technique focuses on predicting liver disease accurately. It improves classifier performance by picking relevant features first with the SAFSA method and then adjusting weight and bias values with the genetic algorithm. The combination of these strategies improves the accuracy of liver disease classification by 4.05%. This study delivers an original and significant contribution to the field of disease prediction.

Fig2. Shows the comparison of PSO with other Machine Learning algorithms. It can be seen that the PSO provides more accuracy that any other algorithm.

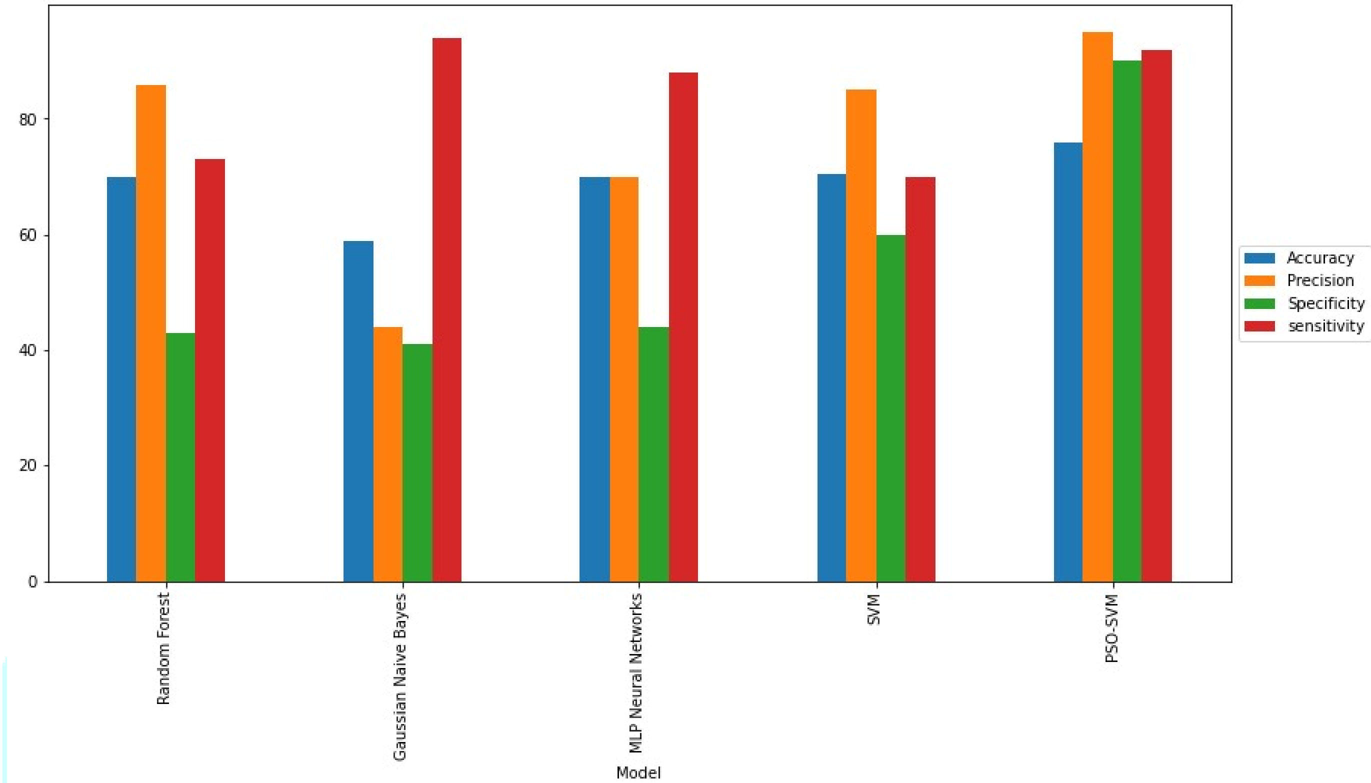


Fig2. Comparison of Several Algorithms

III. Proposed solution –

3.1 Introduction to Particle Swarm Optimization -

Within the class of swarm intelligence algorithms is the meta-heuristic known as particle swarm optimisation. Initially put forth by James Kennedy and Russell Eberhart in 1995, it finds application in a range of search and optimisation issues.

Optimisation can be defined as finding the best solution among a group of workable solutions. This is a widely accepted notion. It is essential in establishing particular

variables with specific resource constraints.

High-level processes known as meta-heuristics are independent of the problem instance or example. They may solve a variety of mathematical problems and can be considered heuristics with more generalised criteria than normal heuristics.

Swarm intelligence, on the other hand, is artificial intelligence derived from the collective actions of self-organizing, decentralised systems. These systems usually consist of a population of people, or agents, interacting with their surroundings and with one other on a local level. When used in the field of computer networks, the agents' cooperative intelligence produces impressive outcomes. Numerous well-known swarm intelligence algorithms exist, including those found in ant colonies, bird flocking, fish schools, grey wolves, ant lions, whales, and more.

3.2 How It Operates: -

PSO is a method based on populations. It makes use of several particles to create the swarm. Every particle denotes a potential resolution. The set of potential solutions cooperates and coexists at the same time. Every particle within the swarm searches the search region for the most advantageous landing spot. Thus, the set of potential solutions is the search area, and the collection of flying particles, or swarm, symbolises the evolving solutions.

Every particle maintains track of both the collective best solution (optimum) and its own best solution (optimum) during the course of the generations (iterations). Next, it adjusts two parameters: position and flight speed (velocity). More specifically, based on its own and its neighbours' flying experiences, every particle dynamically modifies its flying speed. Similar to this, it attempts to alter its location by utilising its current position, velocity, and the separation between its current position and personal and swarm optimal.

The flock of particles, resembling birds, keeps heading in the direction of a favourable region until it reaches the global optimum, which resolves the optimisation issue.

3.3 Parameters used in PSO: -

S(n)=\{s_1, s_2, \ldots, s_n\}: a swarm of n particles

S_i: an single source in the swarm with a position p_i and velocity v_i, i \in [|1,n|]

p_i: the position of a particle s_i

v_i: the velocity of a particle p*i*

$pbest_i$: the personal best solution of a particle.

gbest: the global best solution of the swarm (Global).

f: fitness function

c_1, c_2: acceleration constants (cognitive and social parameters)

r_1, r_2: random numbers between 0 and 1

t: the iteration number

3.4 Calculus-Based Models: -

The PSO algorithm has two major equations. The velocity equation is the first (equation 1), where the computed values of the individual and global best solutions are used to update the velocity of each particle in the swarm. The coefficients of c1 and c2 represent the acceleration factors related to the individual and social aspects.

These are referred to as trust parameters, where c_1 represents the particle's confidence in itself and c_2 the particle's confidence in its neighbours. Along with the random variables r_1 and r_2, they characterise the stochastic impact of social and cognitive behaviours as follows:

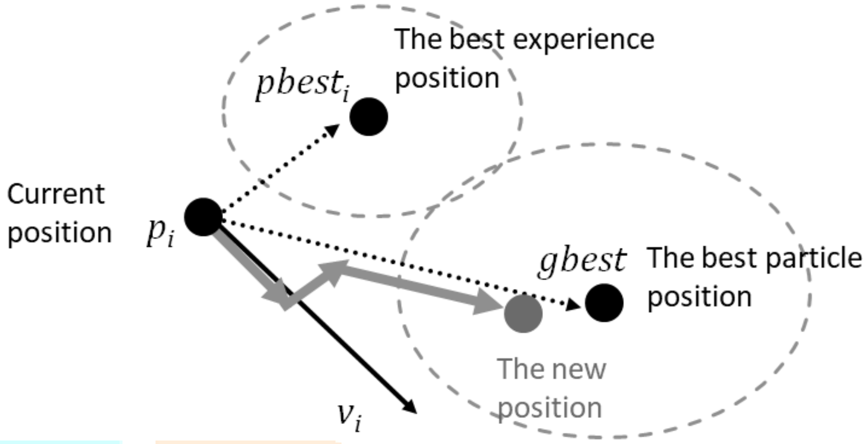
$$\boxed{v_i^{t+1} = \underbrace{v_i^t}_{\text{Inertia}} + \underbrace{c_1 r_1 (pbest_i^t - p_i^t)}_{\text{Personal influence}} + \underbrace{c_2 r_2 (gbest^t - p_i^t)}_{\text{Social influence}}} \quad (1)$$

The second (equation 2) is the position equation, where each particle updates its position using the newly calculated velocity:

:

$$\boxed{p_i^{t+1} = p_i^t + v_i^{t+1}} \quad (2)$$

The parameters of position and velocity are co-dependent, i.e., the velocity depends on the position and vice-versa. We can illustrate the moving particle in the following figure:



IV. Our Proposed Model –

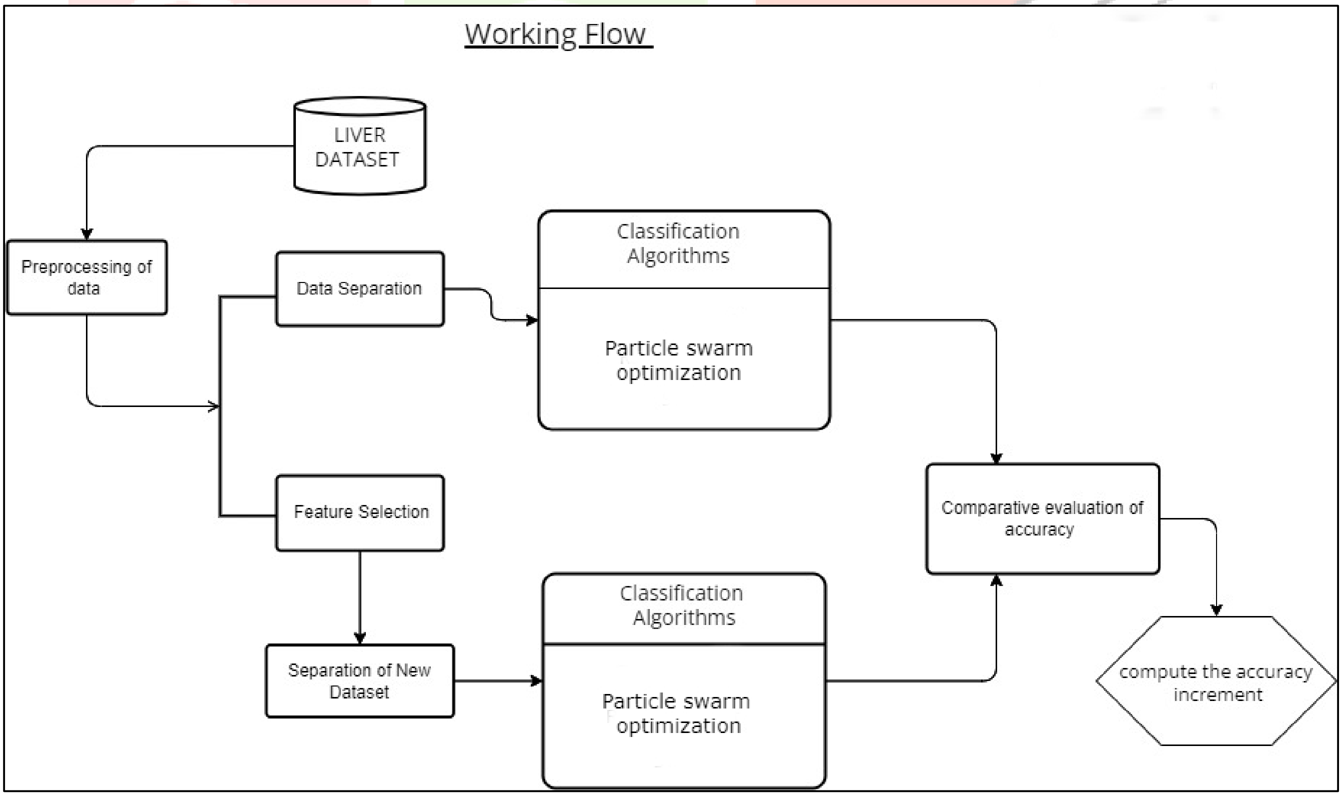


Fig3. Working Model of PSO

V. Comparative Analysis : -

Sr. No	Year	Authors	Dataset Used	Machine Learning Algorithms	Remarks	Conclusion
1	2017	Sumedh Sontake et al.	Machine Learning Repository of UCI	SVM and Back propagation	SVM with accuracy 71% and backpropagation with accuracy of 73.2%	Accuracy increased by using back propagation
2	2018	[12]	Indian Liver Patient Dataset with 10 different scenarios of 583 patients.	K-NN, Logistic regression, Support Vector Machine and ANN	Achieved Accuracy of Logistic regression with 72.23, KNN with 71.05, SVM with 75.01 and ANN with 91.7%	ANN given better accuracy than others.
3	2018	[10]	First Affiliated Hospital, Zhejiang University China's Dataset	classification algorithms	Bayesian network with accuracy of 83%	Bayesian network gives better result than others for Non- alcoholic fatty liver disease prediction.
4	2019	Durai, Vasan, Suyan Ramesh, and Dinesh Kalthireddy	UCI repository databases.	Supervised, Unsupervised and reinforcement algorithms	Practical implementation was not done, but explained how KNN increases performance of the model.	Performance of prediction is increases with help of KNN and AVM.
5	2020	G Shaheamlung, H Kaur, M Kaur	AP liver dataset and UCLA liver dataset	Decision tree, J48 and ANN	Better accuracy in liver illness diagnosis and prediction was provided by decision trees, J48, and ANN.	Accuracy increased by using a combination of Decision tree, J48 and ANN algorithms.
6	2021	Dr. AR.Arunachalam and C.Geetha	Dataset of Indian Liver Patient consisted of ten distinct parameters	SVM, Logistic Regression.	Aims to give accuracy and to compare classification algorithms. Moreover, defeat other methods like Random Forest and Naïve Bayes classification.	Accuracy increased by majorly focusing on SVM and Logistic regression.

7	2022	Zhidong Xu1*, Mostafa Babaeian Jelodar	UCLA liver dataset	Logistic Regression, Random Forest, Support Vector Machine, Deep Learning Models	Consists better result for ability to extract features. Considers factors such as potential productivity gains, reduced errors, improved collaboration, and long-term cost saving.	Deep Learning Models and each of algorithms used, contributes to increase accuracy and extracting features from medical images and handling of missing values.

VI. Conclusion : -

The paper provides us with an overview of the basic principles of earlier research on the diagnosis and detection of liver disease through different machine learning algorithms. This review and analysis have made it absolutely evident that Particle Swarm Optimisation (PSO), a machine learning algorithm, outperforms other machine learning algorithms like Decision trees, SVM, and Navie Baye's algorithm on liver disease detection. Using the proper dataset is important to predict accurate results. With more features in the dataset, we can achieve more accurate dataset as the PSO works best when used with more features. Also, our paper presents the tablular view of the survey papers that we have read and the proper comparison of the paper, with their authors, year of publication, machine learning algorithm used, datasets used and the accuracy. This table helps us to prove that the PSO algorithm has more accuracy than the existing used algorithms.

With this survey paper, we have found more accurate predictions of the liver disease, by making use of the latest algorithm i.e. Particle Swarm Optimization (PSO) which helps in accurately predicting the output.

References –

[1] A Survey on machine learning techniques for the diagnosis of liver disease.

[2] Supervised Machine Learning Models for Liver Disease Risk Prediction.

[3] Statistical Machine Learning Approaches to Liver Disease Prediction.

[4] Software-based Prediction of Liver Disease with Feature Selection and Classification Techniques

[5] Recent Developments In Machine Learning Approach For Liver Disease Prediction

[6] Prediction of Liver Diseases by Using Few Machine Learning Based Approaches

[7] Performance Assessment of Classification Algorithms on Early Detection of Liver Syndrome

[8] Evaluation based Approaches for Liver Disease Prediction using Machine Learning Algorithms

[9] Accurate liver disease prediction system using convolutional neural network.

[10] Ma, Han, Cheng-fu Xu, Zhe Shen, Chao- hui Yu, and You-ming Li. "Application of machine learning techniques for clinical predictive modelling: a cross

sectional study on non-alcoholic fatty liver disease in China." BioMed research international 2018 (2018).

[11] Sivakumar D , Manjunath Varchagall , and Ambika L Gusha S “Chronic Liver Disease Prediction Analysis Based on the Impact of Life Quality Attributes.” (2019). International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume7, Issue-6S5, April 2019

[12] Jacob, Joel, Joseph Chakkalakal Mathew, J. Mathew, and E. Issac. "Diagnosis of liver disease using machine learning techniques." Int Res J Eng Technol 5, no. 04 (2018).

[13]A.S.Aneeshkumar and C.Jothi Venkateswaran, “Estimating the Surveillance of Liver Disorder using Classification Algorithms”, International Journal of Computer Applications (095-8887), Volume 57-No.6, November 2012.

[14] Gitanjali S. Mate, Abdul K. Kureshi, Bhupesh Kumar Singh, "An Efficient CNN for Hand X-Ray Classification of Rheumatoid Arthritis", Journal of Healthcare Engineering, vol. 2021, Article ID 6712785, 10 pages,2021. https://doi.org/10.1155/2021/6712785