

DIAGONISING HEPATITIS DISEASE EFFICIENTLY USING CASE BASED SWARM INTELLIGENCE TECHNIQUE

¹A. Jenita Mary, ²Dr.K. Jayavani

¹Department of Computer Applications, FSH, SRM IST

²Department of Computer Applications, Sri Vijaya College of Arts and Science

Abstract : *In this paper, an analysis framework dependent on Case Based Reasoning and a swarm knowledge strategy called Particle Swarm Optimization for hepatitis infections is presented. Right determination of a sickness is one of the most significant issues in prescription. Hepatitis malady is one of the most risky illnesses that influence a large number of individuals consistently and end man's life. Because of the structure unpredictability and poor proficiency, we need a plan that ought to have the option to assist them with making a decent choice. This technique is utilized to diminish the danger of mistake in drug field, particularly Hepatitis Diagnosis. Initial, a Case Based Reasoning technique is serviceable to preprocess the informational index in this manner a weight vector for everybody highlight is separated. A molecule swarm streamlining model is then pragmatic to amass a basic leadership framework dependent on the chose highlights and maladies perceived. The information utilized has been taken from the site UCI called hepatitis infection. This database has 155 records and 19 fields. This technique was contrasted and other characterization strategies and given the aftereffects of the proposed strategy, better outcomes were accomplished. The right conclusion execution of the proposed framework for hepatitis illness is evaluated by utilizing arrangement exactness.*

Keywords - Case Based Reasoning, Particle swarm optimization, UCI Hepatitis database.

I. INTRODUCTION

Hepatitis alludes to aggravation of the liver parenchyma and can be made for different reasons. Some of them are infectious and some of them are most certainly not. Among the elements making hepatitis, it tends to be alluded to the overabundance in liquor utilization, the impacts of certain prescriptions, and contamination with microbes and furthermore infections. Viral hepatitis brings about liver disease. The reason for the viral hepatitis malady is an infection, and at first, it can seem like a virus. Be that as it may, in contrast to a typical chilly, because of liver disappointment and trouble in treatment, constant "C" hepatitis ailment can compromise the patient's life. The vast majority of those enduring hepatitis sorts C and B have no side effects. A portion of these patients show side effects of viral contamination in nature, for example, weariness, stomach hurt, muscle agony, and sickness, and loss of hunger. In any case, indications of liver disappointment happen in cutting edge cases including swelling of the guts and appendages, jaundice, and stomach related bleedings. Over 3% percent of the people are contaminated with the infection in Iran.

A great deal of analysts have as of late utilized computational knowledge in diagnosing various sicknesses. All these wise systems can just assistance the doctor's determination as an associate and all have a modest quantity of mistake. Among these techniques, neural systems are most broadly utilized. Various types of neural systems with different particulars have been utilized in diagnosing maladies [1]. A ton of explores have been done through neural systems and fluffy framework for finding of B hepatitis infection [2, 3].

Techniques with better grouping precision will give progressively adequate data to distinguish the potential patients and to improve the conclusion exactness. Metaheuristic calculations (like hereditary calculations, molecule swarm improvements, fish swarm streamlining, and Tabu Search) and information mining instruments (neural system and choice tree) have been connected around there. Beside other conventional characterization issues, restorative information groupings are additionally connected in ailment finding. Along these lines, patients or specialists not just need to know the appropriate response (arrangement result), they additionally need to know the indications that infer this answer. Concerning other clinical finding issues, grouping frameworks have been utilized for hepatitis malady determination issue. At the point when the examinations in the writing related with this arrangement application are analyzed, it very well may be seen that an extraordinary assortment of strategies were utilized which achieved high Classification exactnesses utilizing the dataset taken from UCI AI vault.

In [4], An audit which presents ongoing discoveries on noninvasive choices for the finding of fibrosis and cirrhosis in patients co-contaminated with HIV and HCV. A programmed determination framework that incorporates PCA and ANN for grouping of HCV is proposed in [5]. Liao [6] explored of a crossover CBR strategy for disappointment instruments recognizable proof.

Yang et al. [7] incorporated CBR with an ART-Kohonen NN to improve flaw determination of electric engines. Hua Tan et al. [8] coordinated CBR and the fluffy ARTMAP NN to help chiefs in making auspicious and ideal assembling innovation venture choices. [9] presented an element based similitude measure to manage budgetary pain forecast (e.g., liquidation expectation) in China. Chang and Lai [10] incorporated the SOM and CBR for deals gauges of recently discharged books. Chang et al. [11] developed a CBR framework with hereditary calculation for distributor returning book determining. Chun and Park [12] contrived a relapse CBR for money related anticipating, which applies various loads to free factors before discovering comparative cases. Kumar and Ravi [13] exhibited an extensive survey of the works using NN and CBR to tackle the liquidation forecast issues looked by banks.

The information utilized are clarified in the accompanying and the strategies utilized for mix of case based thinking and Particle swarm insight are expressed. The test examine lastly discourse and end likewise will be managed the accompanying areas.

II. RELATED WORK

This hepatitis ailment dataset requires assurance of whether patients with hepatitis will either live beyond words. It was ISRN Artificial Intelligence 3 given by Jozef Stefan Institute, Yugoslavia. The utilized information source in this investigation was taken from UCI AI store. The motivation behind the dataset is to anticipate the nearness or nonattendance of hepatitis illness given the aftereffects of different therapeutic tests completed on a patient. This database contains 19 characteristics, which have been removed from a bigger arrangement of 155. Hepatitis dataset contains 155 examples having a place with two unique classes (32 "pass on" cases, 123 "live" cases). There are 19 traits, 13 double, and 6 characteristics with 6–8 discrete qualities. Characteristics of indications that is gotten from patient are as per the following (UCI Machine Learning Repository):

1. Age: 10, 20, 30, 40, 50, 60, 70, 80
2. Sex: male, female
3. Steroid: no, yes
4. Antivirals: no, yes
5. Fatigue: no, yes
6. Malaise: no, yes
7. Anorexia: no, yes
8. Liver Big: no, yes
9. Liver Firm: no, yes
10. Spleen Palpable: no, yes
11. Spiders: no, yes
12. Ascites: no, yes
13. Varices: no, yes
14. Bilirubin: 0.39, 0.80, 1.20, 2.00, 3.00, 4.00
15. Alk phosphate: 33, 80, 120, 160, 200, 250
16. Sgot: 13, 100, 200, 300, 400, 500
17. Albumin: 2.1, 3.0, 3.8, 4.5, 5.0, 6.0
18. Prottime: 10, 20, 30, 40, 50, 60, 70, 80, 90
19. (Histology: no, yes

III. METHOD

Case-based thinking (CBR) is a noteworthy worldview in robotized thinking and AI. On the off chance that based thinking, a reasoner takes care of another issue by seeing its likeness with one or a few recently tackled issues and by adjusting their known arrangements as opposed to working out an answer without any preparation. Case-based thinking can mean various things relying upon the proposed utilization of the thinking: adjust and join old answers for take care of another issue, disclose new circumstances as indicated by recently experienced comparable circumstances, investigate new arrangements dependent on old cases, thinking from points of reference to comprehend another circumstance, or manufacture an accord arrangement dependent on past cases. Be that as it may, these various angles can be grouped into two noteworthy sorts: interpretive (or characterization) CBR, and critical thinking CBR.

When all is said and done, given a case to tackle, case-based thinking includes the accompanying advances : 1. recovering applicable cases from the case memory (this requires to record the cases by suitable highlights);
2. choosing a lot of best cases;
3. inferring an answer;
4. assessing the arrangement (so as to ensure that poor arrangements are not rehashed);
5. putting away the recently illuminated case for the situation memory.

This exploration work takes two stages. In first stage, the case based thinking technique extricated the element dependent on weight factor and in the second stage the changed molecule swarm advancement has been utilized for expanding the precision level of the forecast.

One of highlight weighting strategy that relegates weight internationally is Gradient Descent (GD) technique, where this weighting strategy is altogether better than different models. Angle Descent is an unaided strategy that has a bit of leeway that plays out the errand of highlight weighting without bunching the component space expressly and does not have to know the quantity of groups present in the element space. In this exploration, slope drop was utilized to acquire highlight weight all inclusive and after that changed molecule swarm improvement has been utilized to streamline a model.

Input:

N: data points
K: number of classes (the same with number of cluster' centroids)
M: temporary centroids (M>K, for initial)
W: weights calculated by CBR
Procedure Stepwise Centroids PSO Clustering with CBR
M: =Weighted PSO Clustering (N, |M|, W);
Reassign M as data points (N: =M);
Reduce number of M to M
Recursive execute
Stepwise Centroids PSO Clustering with CBR until M' equals to K;
//means Re-cluster the M data points into M clusters, if M equals to K,
then final result is found
Return K centroids;
End;

Var:

j: attribute of dataset
d: dimension of each data (number of attributes)
Input: data: hepatitis disease dataset
K: number of classes
Output: Classification Result (the location of K centroids)
Procedure Weighted PSO Clustering (data, K, weights)
Generate P solutions (particles); //each solution has its own
Kcentroids selected randomly from dataset.
For each particle
Calculate Objective function
$$v_{id} = w \cdot v_{id} + c_1 \cdot \text{rand } 1 \cdot (p_{id} - x_{id}) + c_2 \cdot \text{rand } 2() \cdot (p_{gd} - x_{id})$$
$$x_{id} = x_{id} + v_{id}$$
Update p_{id}
End
Update p_{gd}
End.

EXPERIMENTAL RESULTS

The dataset used in the proposed system has been taken from UCI repository. This database has 19 fields and 155 samples. In the pre-processing phase, the Local Mean method is used to clear the noise from patient's data records and handle missing data. The Case Base Reasoning Algorithm is used to extract the most effective features on the basis of weighting method. The efficient weighting method used in this research is Gradient Descent method. These weighted feature get optimized by swarm intelligence technique called particle swarm optimization algorithm.

Hepatitis disease	PSO	Modified PSO
Best	88.32%	93.95%
Average	80.31%	90.71%
Lowest	73.51%	76.22%

Table 1: Accuracy comparisons of Modified PSO and Normal PSO.

Changed PSO technique could analyze hepatitis malady in the best state with the exactness of 93.95%, yet PSO strategy could analyze this sickness in the best state with the precision of 88.32%. The general capacity of Modified PSO strategy is better in connection to Normal PSO technique in the normal state and has higher effectiveness. So as to research the capacity of Modified PSO method (M-PSO) better, it was contrasted likewise and techniques for SVM, Naive Bayes and KNN.

Hepatitis disease	SVM	Naive Bayes	KNN	M-PSO
Best	90.11%	82.15%	88.62%	92.87%
Average	84.75%	80.12%	81.40%	90.14%
Lowest	63.11%	64.21%	70.12%	74.88%

Table 2: Accuracy comparisons of various classification algorithm

The arrangement calculations that are connected individually on the dataset are: SVM, Naive Bayes, KNN AND M-PSO. 70% of the information wear haphazardly picked for preparing while 30% of these information is picked for testing for these models with a complete number of 300 execution times. Precision esteem for various level of preparing dataset taken. Subsequent to looking at all the three classifiers it is presumed that the case based prevailing upon molecule swarm advancement (M-PSO) calculation accomplished the best result. Eventhough SVM technique could analyze hepatitis in the best state with great accuracy, there is a constraints in speed and size during both preparing and testing period of the calculation and the choice of the parameters. KNN does not function admirably with high measurements and furthermore Naive Bayes has solid element freedom assumptions. So every strategy has points of interest and disadvantages. Our proposed calculation conquer this issue and acquire the best outcome.

IV. CONCLUSION

A calculation for Diagnosing Hepatitis Disease proficiently was proposed. This exploration work takes two phases. In first phase, the case based thinking technique removed the element dependent on weight factor and in the second stage the altered molecule swarm streamlining has been utilized for expanding the exactness level of the expectation. Slope drop technique was utilized to acquire highlight weight internationally and after that adjusted molecule swarm optimization (M-PSO) has been utilized to advance a model. The proposed framework was contrasted and different characterization framework. The exploratory outcomes demonstrated the prevalence of the proposed framework which could get high grouping precision. As a future work, other streamlining procedure with fluffy rationale could be utilized in medicinal information.

REFERENCES

- [1] P. J. G. Lisboa, E. C. Ifeachor, and P. S. Szczepaniak, "Artificial Neural Networks in Biomedicine", Springer, London, UK, 2000.
- [2] M. Neshat and M. Yaghobi, "FESHDD: fuzzy expert system for hepatitis B diseases diagnosis," in Proceedings of the 5th International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control (ICSCCW '09), Cyprus, September 2009.
- [3] M. Neshat and M. Yaghobi, "Designing a fuzzy expert system of diagnosing the hepatitis B intensity rate and comparing it with adaptive neural network fuzzy system," in Proceedings of the World Congress on Engineering and Computer Science, San Francisco, Calif, USA, October 2009.
- [4] Mohammed H Af, Abdel-Rahman Hedar, Taysir H Abdel Hamid, and Yousef B Mahdy. Ss-svm (3svm): A new classification method for hepatitis disease diagnosis. International Journal, 2013
- [5] JC Bansal, PK Singh, Mukesh Saraswat, Abhishek Verma, Shimpi Singh Jadon, and Ajith Abra-ham. "Inertia weight strategies in particle swarm optimization". In Nature and Biologically Inspired Computing (NaBIC), 2011 Third World Congress on, 2011, pages 633-640. IEEE.

- [6] T. W. Liao, "An investigation of a hybrid CBR method for failure mechanisms identification," *Engineering Applications of Artificial Intelligence*, vol. 17, no. 1, 2004, pp. 123–134.
- [7] B. S. Yang, T. Han, and Y. S. Kim, "Integration of ART-Kohonen neural network and case-based reasoning for intelligent fault diagnosis," *Expert Systems with Applications*, vol. 26, no. 3, 2004, pp. 387–395.
- [8] K. Hua Tan, C. Peng Lim, K. Platts, and H. Shen Koay, "An intelligent decision support system for manufacturing technology investments," *International Journal of Production Economics*, vol. 104, no. 1, 2006, pp. 179–190.
- [9] H. Li, J. Sun, and B. L. Sun, "Financial distress prediction based on OR-CBR in the principle of k-nearest neighbors," *Expert Systems with Applications*, vol. 36, no. 1, 2009, pp. 643–659.
- [10] P. C. Chang and C. Y. Lai, "A hybrid system combining self-organizing maps with case-based reasoning in wholesaler's new-release book forecasting," *Expert Systems with Applications*, vol. 29, no. 1, 2005, pp. 183–192.
- [11] P. C. Chang, C. Y. Lai, and K. R. Lai, "A hybrid system by evolving case-based reasoning with genetic algorithm in wholesaler's returning book forecasting," *Decision Support Systems*, vol. 42, no. 3, 2006, pp. 1715–1729.
- [12] S. H. Chun and Y. J. Park, "A new hybrid data mining technique using a regression case based reasoning: application to financial forecasting," *Expert Systems with Applications*, vol. 31, no. 2, 2006, pp. 329–336.
- [13] K. Panchal, H. Kundra, and N. Kaur, "A novel approach of waves of Swarm with case based reasoning to detect ground water potential," *Journal of Technology and Engineering Sciences*, vol. 1, 2009, pp. 3–8.
- [14] K. S. Kim and I. Han, "The cluster-indexing method for case-based reasoning using self-organizing maps and learning vector quantization for bond rating cases," *Expert Systems with Applications*, vol. 21, no. 3, 2001, pp. 147–156.
- [15] S. H. Chun and Y. J. Park, "A new hybrid data mining technique using a regression case based reasoning: application to financial forecasting," *Expert Systems with Applications*, vol. 31, no. 2, 2006, pp. 329–336.
- [16] P.-C. Chang, J.-J. Lin, and C.-H. Liu, "An attribute weight assignment and particle swarm optimization algorithm for medical database classifications," *Computer Methods and Programs in Biomedicine*. In press. [26] D. W. van der Merwe and A. P. Engelbrecht, "Data clustering using particle swarm optimization," in *Proceedings of the Congress on Evolutionary Computation*, 2003, pp. 215–220.