International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-7 Issue-4, September 2018

A Diagnosis System for Multi class Primary Headaches using Ant Miner plus Algorithm

Nissankara Lakshmi Prasanna, Pujari Jeevana Jyothi, Nallamekala Rajeswari, R Tejaswini

Abstract— This article proposes a diagnosis system for dealing with primary headache classification using variant of ant colony optimization algorithm for classification. The diagnosis system proposed is an expert system completely designed with information gathered from patients suffering from headaches. Randomly chosen patients who are visiting various hospitals in India are addressed to collect data for expert system. They are provided with detailed questionnaire about the symptoms to analyze the headache types. With help of neurologist the headaches are compared with the results of ant colony optimization algorithms for classification. Ant miner plus algorithm is enforced for better results. The algorithm is observed for its accuracy levels of classification and is analyzed. This article addresses 4 types of headaches as multi-class classification problem with an in detailed report of symptoms gathered from patients. This kind of expert system is useful to neurologists to track symptoms of their patients and to provide mediation.

Keywords: primary headaches, ant colony optimization, diagnosis system, ant miner plus algorithm, multi-class classification problem.

I. INTRODUCTION

Classification is important because it concentrates to identify an objects group and name depends on the similarities found with existing categories. It plays crucial role in decision making. A classification problem is defined as the assignment of a unknown sample to a predefined class according to its characteristics [1], [2]. Classification problems occur in various fields. Medical, engineering, scientific, robotics, banking, in every domain multiple classification problems will be found. In medical field, diseases can be broadly categorized based on the symptoms, majorly classification models are used to identify the disease of a new patient using existing data. Such automation system of disease identification and rectification procedures makes the neurologists work simple and efficient.

Diagnosing the headaches is one of the prominent challenging classification problems in medical diagnosis. Mostly headaches are classified as primary and secondary type of headaches. The most common disease affecting the working people are primary headaches and also shows high impact on their working capability as well as productivity also[3, 4]. In 2004 The International Classification of Headache Disorders second Edition (IHCD-2) was published [5], and then headache classification became an active research area. The primary headaches are not related to any medical condition of a person. The current revision of IHS criteria is ICHD-3 beta was published in 2013 [6-8]. Secondary headaches are due to some sort of medical problems in human body, and these are considered as symptom of life threatening disease without quick and accurate diagnosis [9, 10].

Primary headaches were classified as four different types of headaches such as tension-type, migraine, sinus and cluster-type headaches. Different headaches has some similar symptoms or features for classification, which creates complications to diagnose. Sometimes, physicians may also experience some difficulties to accurately differentiate them due to the similarities of features exist. But accurate diagnosis of cause is necessary for more than 200 diseases for optimal treatment. Due to lack of proper treatment, most of the people are continuously suffering from headaches and their symptoms.

For diagnosing primary headache disorders more accurately, several classification algorithms are studied [11-19] to generate rules.

One of the more familiar meta heuristic algorithm proposed by M.Dorigo[20] in the early 90's inspired by the natural behaviour of ant colonies is ant colony optimization(ACO). ACO is one of the developed optimization algorithms in swarm intelligence. Because of the strategy of ACO learning style suits for solving complex combinatorial optimization problems and used in most of the real time applications such as Travelling salesman problem(TSP)[21],quadratic assignment problems [22,23], vehicle routing problems[24], scheduling problems[25,26], sequential ordering problem[27], open-shop scheduling problem[28] and many other applications are involved. Ants exchange information in an indirect way via their search environment by retaining the chemical substance named as pheromone.

Revised Manuscript Received on 30 September 2018.
***** Correspondence Author
Nissankara Lakshmi Prasanna, Assoc. prof, dept. of Computer Science & Engineering, Vasireddy Venkatadri Institute of Technology, Guntur (A.P), India
Pujari Jeevana Jyothi, Assoc. prof, dept. of Computer Science & Engineering, Vasireddy Venkatadri Institute of Technology, Guntur(A.P), India
Nallamekala Rajeswari, Asst. prof, dept. of Computer Science & Engineering Vasireddy Venkatadri Institute of Technology, Guntur (A.P), India
R Tejaswini, Asst. prof, dept. of Electronics & Communications Engineering, Vasireddy Venkatadri Institute of Technology, Guntur (A.P), India

However, we used the variant of ant colony optimization algorithm is Ant Miner+ to diagnostic and predictive models for generating more accurate classification rules on Real world primary headache patient dataset consists of multi class problem(4-class problem). The remaining sections of this paper are organized as follows. Section II is review of literatures which covered various classification algorithms applied for headache classification. Method applied in each algorithm and results accuracy levels of classification including drawback of each algorithm is studied. Section III covers dataset and basic introduction of ant colony algorithms. Attributes of data set, their identification importance of those attributes and used expert system to gather the data set is explained. Behavior of ant colony algorithms, ant miner algorithm used for classification and its variant ant miner plus algorithm is explained. Our approach applied on the data set is explained. Section IV discusses about the results and accuracy of classification. Section V concludes this paper and listed scope for future research.

II. REVIEW OF LITERATURE

This section mainly focuses on the algorithms to handle primary headache classification problem. Here we are presenting many classification algorithms are proposed in past to solve the problem of primary headaches. The computerized headache diagnosis system [10] shows different clinical records that shows the various patients symptoms. The web-based headache diagnosis expert system [11] uses Microsoft access database to store the patient's records by manually. Pryse-Philips et al. [12] provides decision making system based on the Rule-based algorithm and decision tree algorithm only for the migraine type of headaches. Maizels M, Wolfe and et al. [13] accurately identified the various headaches by using the CHAT (Computerised Headache Assessment Tool). The results shown in different categories of headache like 1. Episodic Headache 2. Chronic Headache. CHAT correctly identified the Tension, Cluster, Migraine Headaches based on the category also. Less accurate results were shown by the system in case of transformed migraine and patients shown less interest to use this tool are the disadvantages of this system.

D. Kernick, S. Stapley, and W. Hamilton et al.[14] can accurately diagnosed the headache categories as 1. Primary headaches 2. Secondary headaches. Results in this method showing that, headaches are more likely to present in adults because of their hormonal changes in that age. The proportion of the patients suffering with indifferent headaches is increasing by age factor. The main problem with this study is migraine is highly wrong diagnosed due to the reason that women are more sufferers with the migraine. Olesen, Jes, and T. J. Steiner te al. [5] identified various types of headaches, need different diagnosis systems and treatments. Major headaches need medical practitioner along with neurologist consultation of the patient. In this methodology they evaluated a rule based fuzzy system for diagnosing the patient's symptoms. This methodology is useful in accurately classifying the migraine and tension-type of headache but not for cluster type of headache.

Yin, Ziming et al. [15] proposed a system which can accurately identify the probable migraine (PM) and probable tension-type headache (PTTH) which confuses even the doctors. Along with that, authors proposed another solution namd rule based case-based decision making to identify headache disorders. Yin, Ziming, et al. [16] shows in another research uses a guideline based headache diagnosis algorithms for classification. Krawczyk et al. [17] uses advanced machine learning techniques to recognize the primary headache types. The computer decision support system applied on the large datasets results high accuracy in recognizing the primary headaches.

In this system authors used another approach to [18] [19] classify the headache is interval-valued intutionist fuzzy sets and aggregate operator. Dorigo, Marco et al. [20] introduced a decision support system for diagnosing the headaches. The results shown can accurately identify the primary tension type and migraine type of headaches. Świątnicki, et al. [21] shows various surveys on machine learning classifiers on headache datasets. Maniezzo, Vittorio et al. [22] developed a methaheuristic algorithm based on the biological metaphor of ANTS. The two algorithms used in this methodology are 1. Heuristic Technique, used to identify an Ant system metaphor and that is used as lower bound in every constructive step. 2. Branch and bound exact approach, is has some elements defined which are used in the Ant algorithm.

III. PROPOSED METHODOLOGY

We collected data records of 500 patients suffering with headaches that they are not aware of various types of headaches. Website questioner system is given to each patient which gathers information about various symptoms like pain area, Nausea, Throbbing pain or not, Severity of pain, When they are feeling the headache, Dizziness or not and etc. Patients were chosen randomly from various cities and both sexes who are visiting neurologists for treatment of their headaches. These patients belong to various age groups from 15 to 70. These patients addressed a set of questions in web site under the observation of a neurologist.

We considered information from the International Headache Society and the International classification of headache disorders. Based on the doctor's survey, Headaches are classified as follows. 1. Normal headaches 2.Chronic headaches. Headaches have high chances to wrong diagnose by the doctor if patient cannot identify their proper symptoms. So, for that reason we concentrated mainly on the identification of symptoms of patients during headaches. Classification of the patients symptoms are depending upon the doctor's suggestions. The previous system is revised depends on the classification of the patient's symptoms.

The major parameters that we consider to decide headache type are Pain area, Throbbing, Eye redness, Time of pain, Nasal congestion, Severity of pain, Sensitivity to light and sound, Nausea or vomiting sensation, Visual disturbance, Trouble speaking, Paralysis, Menstruation details, Hormonal change, Caffeine consumption, regarding physical activities, Blood pressure details, Any head injury in past.

Apart from these medical symptoms, patients are oral examined about some personal parameters also like name of the patient, Age, Sex, injury in past, paralysis, any mental imbalance.

Based on all the parameters we consider some of them as main attributes to decide the type of headache. We consider these parameters as main attributes to identify the class headache: Pain area, Nausea, Time of pain, Severity of pain, Sensitivity of light and sound, Throbbing, Pain duration, Dizziness.

The symptoms which we collected are the parameters of our database. Attributes can be considered as the primary symptoms. Various types of headaches can be considered as classes. Our Project results shows that identifying the classes depends on the attributes selected from the patients.

These questions are designed to identify various types of headaches which are not concerned in previous studies. This research aims to identify exact type of headache with in deep information provided by patients. The questionnaire having various questions asked to the patients about their feelings when they get the headaches. And also general questions like their name, age, sex, sleep hours, habits and so on. 30 plus attributes are gathered from patients for classification algorithm. PHP with MySQL database is used for designing the expert system and to prepare database.

After preparing data set various forms of ant colony algorithm is applied and compared with their accuracy levels of classification. These algorithms generated a rule based classification system which resembles the usual ICHD-2 procedure used by the neurologists. Ant colony algorithm, Ant Miner Algorithm and ant miner

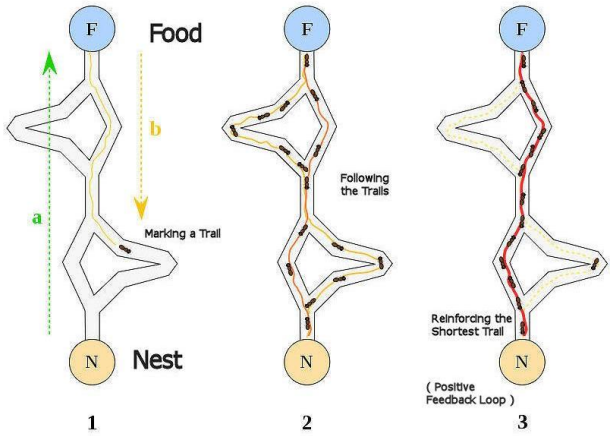

Fig 1: Working of ACO algorithm

Plus algorithms are applied on these collected samples of data set. Each algorithm is compared with their levels of accuracy for classification.

A. ACO for Classification

Inspired by social insects' ants, ant colony algorithm was designed. Ants forms ant colonies and they are very intellectual in search of their food. They always use indirect communication known as stigmergy. Multiple ants start their journey in search of food. Each ant deposits a substance called pheromone in its path. Paths with higher pheromone levels turns as shortest path to their food source. Paths not chosen will decrease their pheromone levels by evaporation as shown in Fig1.

With an inspiration of this natural mechanism Ant colony algorithm uses artificial ants to find optimized solution for discrete problems. Multiple ants were initialized to start the search process. Each ant travels in its own path towards the solution by adding pheromone in its path. Pheromone and heuristic values are used for selecting best path. Pheromone can be used to identify the number of ants required for obtaining

solution and heuristic a problem dependent measure for identifying promising path. Once the ant arrives destination pheromone level of the path is increased accordingly. The remaining ants followed in that path automatically identifies shortest path to the destination. Evaporation of pheromone level on remaining paths also helps in choosing fruitful paths towards solution.

Requirements to Ant system are identifying problem domain, heuristic function, pheromone updating rule, a probabilistic function that strengths pheromone level by using heuristic function and clear specification of solution. Representing problem, which allows the ant to traverse entire domain. An heuristic function that was designed based on problem which is used for identification of partial solutions. For updating pheromone a rule is need to be specified. A probabilistic function is required which is constructed over pheromone value and heuristic function. Traditional ACO can be combined with various ideas for better results.

A variation of ant colony optimization is ant miner which can be used for learning classification rules. This mines classification rules from data set. All the rules mined are applied on trained data set. This is used to filter some rules mined by ant miner algorithm. The rules mined can be expressed in the form if rule then action format. All the attributed assumed are categorical. The algorithm produces a list of rules for classification in a sequential order.

$$P_{ij}(t) = \frac{\mathcal{T}_{ij}(t).\eta_{ij}}{\sum_{k=1}^{n} x_k \sum_{l=1}^{p_k} (\mathcal{T}_{kl}(t).\eta_{kl})}$$

The problem domain is a classification of multi class variable to identify the type of headaches in data set. Ant miner initializes multiple ants. Problem can be represented as directed graph G(V, E) where V is the set of vertices which are the attributes of considered data set. Each vertex can be a defined in terms of multiple values which belongs to list of categorical set.

Class	True Positive Rate	False Positive Rate	Precision	Recall	F-Measure	ROC Area	PRC Area
Migraine	0.972	0.026	0.921	0.972	0.946	0.985	0.962
Sinus	0.916	0.009	0.974	0.916	0.944	0.98	0.97
Cluster	0.966	0.034	0.945	0.966	0.903	0.984	0.974
Tension	0.926	0.059	0.852	0.926	0.888	0.974	0.915

Table 1: Results of ant Miner Plus Algorithm Over Proposed Expert System

Are initialized by using number of values of each vertex. The graph constructed is a symmetric and bidirectional. Let n be the total number of variables, pi is the number of values of each variable, xi a binary variable denotes

Algorithm 1 Classification

```
Construct Directed Acyclic Graph
While (not early stopping)
    Initialize heuristics, pheromones and probabilities of edges.
    While (not converged)
        Create ants;
        Let ant run from source to sink;
        Evaporate pheromone on edges;
        Prune rule of the best ant;
        Update path of best ant;
        Adjust pheromone levels if outside boundaries kill ants;
        Update probabilities of edges;
    End
    Extract rule
    Flag data points covered by extracted rule
End
Evaluate performance on data set
```

Visited status of each variable by current ant, and start is the start vertex. Pheromone is tij and heuristic value is nij the probability function pij is calculated with equation 1.

Each ant begins at start vertex and traverses each vertex of graph by adding terms to partial rule. According to user requirements each rule can be selected. The final rule will be pruned with irrelevant terms. This process is repeated by all ants. Each ant constructs a rule in its path. Best rules will be chosen from the collected set. By using training set these rules will be observed checked till it meets the user specified threshold value.

Though ant miner can be effectively used for generating classification rules an improvement is proposed to ant miner which is ant miner+. With few modifications in ant miner algorithm ant miner+ is proposed which is rather effective performance than ant miner. The modifications proposed are

- the problem is represented with directed acyclic graph
- nominal and ordinary variables are distinguished
- weight parameters are included
- Max-Min ant system is included
- heuristics are defined foe accurate class specifications rather major class
- Early stopping criteria are included.

B. Sample Rule List Generated

If Nausea = No AND Pain area = Behind eye brow then class= Sinus

If Nausea = No AND Pain area = Nose, eyes then class= Sinus

If Nausea = No AND Pain area = All over head then class=Tension

If Pain area = Both sides AND Nausea = Yes then class=Migraine

If Nausea = No AND Pain area = Both sides AND Pain duration = 2-3 hrs class=Sinus.

IV. RESULTS

Detailed accuracy of classification by class for ant miner plus algorithm is shown in Table 1. In this study, we collected 500 samples of headache patients to identify four class labels (Migraine, sinus, tension, clustered).

We presented the results of AM+ algorithm in the following table. The rows of the table represent the specified class labels and columns represent the classification accuracy details.

In this AM+ algorithm, we used 10-fold cross validation method to test the reliability of the proposed algorithm. We iterated AM+ of 200 iterations and 100 ants initially to obtain the best accuracy in the classification of primary headaches and also generate the classification rules to predict new instances accurately. We attain the accuracy of 97.2% of accuracy in the migraine, 91.6% for sinus, 96.6% in cluster typed, and 92.6% in tension type headaches.

V.CONCLUSION AND FUTURE SCOPE

Patients suffering from headaches need proper approach for their medical treatment. The expert system used provides detailed information of the symptoms and rather provides precise information to the doctors. This kind of approach in hospitals makes the environment more clear and simple for information collection.

In this paper we discussed ant miner+ classification algorithm. By using ant miner+ algorithm we classified head ache types occurred in patients. We collected samples of 500 patients belongs to various cities and with ant miner+ algorithm we obtained 97.132% of accuracy. The results of algorithm are simple rule generation for classification with accuracy. The problem considered is multi class problem which identifies the type of head ache as migraine, tension type, sinus or cluster type with ant miner+ algorithm. Though our system produced accurate results this can be improved further.

REFERENCES

1. Freitas, Alex, and André Carvalho. "A tutorial on hierarchical classification with applications in bioinformatics." In Research and Trends in Data Mining Technologies and Applications, pp. 175-208. IGI Global, 2007.
2. Duda, R. O., P. E. Hart, and D. G. Stork. "Pattern Classification, New York: Wiley Interscience." (2000).
3. Pop, P. H. M., C. M. Gierveld, H. A. M. Karis, and H. G. M. Tiedink. "Epidemiological aspects of headache in a workplace setting and the impact on the economic loss." European journal of neurology 9, no. 2 (2002): 171-174.
4. Lipton, Richard B., Walter F. Stewart, Seymour Diamond, Merle L. Diamond, and Michael Reed. "Prevalence and burden of migraine in the United States: data from the American Migraine Study II." Headache: The Journal of Head and Face Pain 41, no. 7 (2001): 646-657.
5. Olesen, Jes, and T. J. Steiner. "The International classification of headache disorders, 2nd edn (ICDH-II)." (2004): 808-811.
6. Headache Classification Committee of the International Headache Society (IHS). "The international classification of headache disorders, (beta version)." Cephalalgia 33, no. 9 (2013): 629-808.

7. Scottish Intercollegiate Guidelines Network. Diagnosis and Management of Headache in Adults: A National Clinical Guide. Scottish Intercollegiate Guidelines Network, 2008.
8. Ravishankar, K. "The art of history-taking in a headache patient." Annals of Indian Academy of Neurology 15, no. Suppl 1 (2012): S7.
9. Kernick, David, Sally Stapley, and William Hamilton. "GPs' classification of headache: is primary headache underdiagnosed?." Br J Gen Pract 58, no. 547 (2008): 102-104.
10. Morgan, Myfanwy, Linda Jenkins, and Leone Ridsdale. "Patient pressure for referral for headache: a qualitative study of GPs' referral behaviour." Br J Gen Pract 57, no. 534 (2007): 29-35.
11. Al-Hajji, AhmadA. "Rule-Based expert system for diagnosis and symptom of neurological disorders "Neurologist Expert System (NES)"." In Proceedings of the 1st Taibah University International Conference on Computing and Information Technology, Al-Madinah Al-Munawwarah, Saudi Arabia, vol. 1214, p. 6772. 2012.
12. Kinabalu, Kota. "2012 International Symposium on Computer Applications and Industrial Electronics (ISCAIE 2012)." (2012).
13. Yin, Ziming, Zhao Dong, Shengyuan Yu, Xudong Lu, Guanjun Feng, and HuilongDuan. "A guideline-based decision support system for headache diagnosis." Studies in health technology and informatics 192 (2013): 1022-1022.
14. Dong, Zhao, Ziming Yin, Mianwang He, Xiaoyan Chen, XudongLv, and Shengyuan Yu. "Validation of a guideline-based decision support system for the diagnosis of primary headache disorders based on ICHD-3 beta." The journal of headache and pain 15, no. 1 (2014): 40.
15. Yin, Ziming, Zhao Dong, Xudong Lu, Shengyuan Yu, Xiaoyan Chen, and HuilongDuan. "A clinical decision support system for the diagnosis of probable migraine and probable tension-type headache based on case-based reasoning." The journal of headache and pain 16, no. 1 (2015): 29.
16. Yin, Ziming, Lingtong Min, Xudong Lu, and HuilongDuan. "A clinical decision support system for primary headache disorder based on hybrid intelligent reasoning." In Biomedical Engineering and Informatics (BMEI), 2014 7th International Conference on, pp. 683-687. IEEE, 2014.
17. Krawczyk, Bartosz, Dragan Simić, Svetlana Simić, and MichałWoźniak. "Automatic diagnosis of primary headaches by machine learning methods." Open Medicine 8, no. 2 (2013): 157-165..
18. Mendes, Karina Borges, Ronald Moura Fiuza, M. Teresinha, and A. Steiner. "Diagnosis of headache using artificial neural networks." J. Comput. Sci 10, no. 7 (2010): 172-178.
19. Celik, Ufuk, NiluferYurtay, EmineRabiaKoc, NerminTepe, HalilGulluoglu, and Mustafa Ertas. "Migraine, tension-type and cluster-type of headaches classification by using immunos algorithms." Journal of Medical Imaging and Health Informatics 6, no. 5 (2016): 1173-1177.
20. Dorigo, Marco. "Optimization, learning and natural algorithms." PhD Thesis, Politecnico di Milano (1992).
21. Świątnicki, Zbigniew. "Application of ant colony optimization algorithms for transportation problems using the example of the travelling salesman problem." In Advanced Logistics and Transport (ICALT), 2015 4th International Conference on, pp. 82-87. IEEE, 2015.
22. Maniezzo, Vittorio. "Exact and approximate nondeterministic tree-search procedures for the quadratic assignment problem." INFORMS journal on computing 11, no. 4 (1999): 358-369.
23. Maniezzo, Vittorio, and Alberto Colorni. "The ant system applied to the quadratic assignment problem." IEEE Transactions on Knowledge & Data Engineering 5 (1999): 769-778.
24. Gambardella, Luca Maria, ÉricTaillard, and Giovanni Agazzi. "Macs-vrptw: A multiple colony system for vehicle routing problems with time windows." In New ideas in optimization. 1999.
25. Blum, Christian, and Michael Sampels. "An ant colony optimization algorithm for shop scheduling problems." Journal of Mathematical Modelling and Algorithms 3, no. 3 (2004): 285-308.
26. Den Besten, Matthijs, Thomas Stützle, and Marco Dorigo. "Ant colony optimization for the total weighted tardiness problem." In International Conference on Parallel Problem Solving from Nature, pp. 611-620. Springer, Berlin, Heidelberg, 2000.
27. Gambardella, Luca Maria, and Marco Dorigo. "An ant colony system hybridized with a new local search for the sequential ordering problem." INFORMS Journal on Computing 12, no. 3 (2000): 237-255.
28. Blum, Christian. "Beam-ACO—Hybridizing ant colony optimization with beam search: An application to open shop scheduling." Computers & Operations Research 32, no. 6 (2005): 1565-1591.
29. Khan, Salabat, and Abdul Rauf Baig. "Ant colony optimization based hierarchical multi-label classification algorithm." Applied Soft Computing 55 (2017): 462-479.
30. Stützle, Thomas, and Holger H. Hoos. "Improving the Ant System: A detailed report on the MAX–MIN Ant System." FG Intellektik, FB Informatik, TU Darmstadt, Germany, Tech. Rep. AIDA–96–12 (1996).