Swarm Intelligence Based Public Health Informatics and Decision Support Systems

Dr. Abhijeet Madhukar Haval[1], Md Afzal[2]

[1]Assistant Professor, Department of CS & IT, Kalinga University, Raipur, India.
[2]Research Scholar, Department of CS & IT, Kalinga University, Raipur, India.

The domain of medical informatics is concerned with the knowledge of information and computer science, engineering and technology towards health studies, medicine and its practices. The major consequence that has correlation and dependency among the attributes has not been determined over real-world dataset to explore the undermined patterns and relationship in medical data. Feature selection also, in turn, uses up by some of the authors but the relevance, and investigation of selected attributes along with its contribution to the disease has not been measured and signified. The aim of this thesis, as mentioned, is to develop a decision support model for clinical data analysis. The model has been developed using swarm intelligence and data classification techniques with statistical evaluation. A unified Intelligent Optimization Unit has been developed which suits all possible combinations of feature selection and data classification techniques.

Keywords: Health care, Decision, Health Informatics Machine learning.

1. Introduction

Over the past 30 years, medical informatics has advanced as the medical field has attempted to analyse complex data using patterns and data produced by computer systems [3]. Medical informatics research began to evolve in the 1950s with the investigation of novel results, risk behaviours, and the possibility of disease-specific symptoms. Determining aspects of gene structure, predicted disease qualities, and nucleotide polymorphism structure are all part of the bioinformatics platform [1]. Data-driven decision making can be achieved through the convergence of medical research and information technology. Integrated access to clinical data is necessary for the future community to develop computer-based data management and decision-making procedures. The majority of the current system accesses repository data and evaluates it using different classification approaches [6]. Furthermore, the algorithmic model has only been assessed in terms of correctness, meaning that not all datasets can be customised. [2].

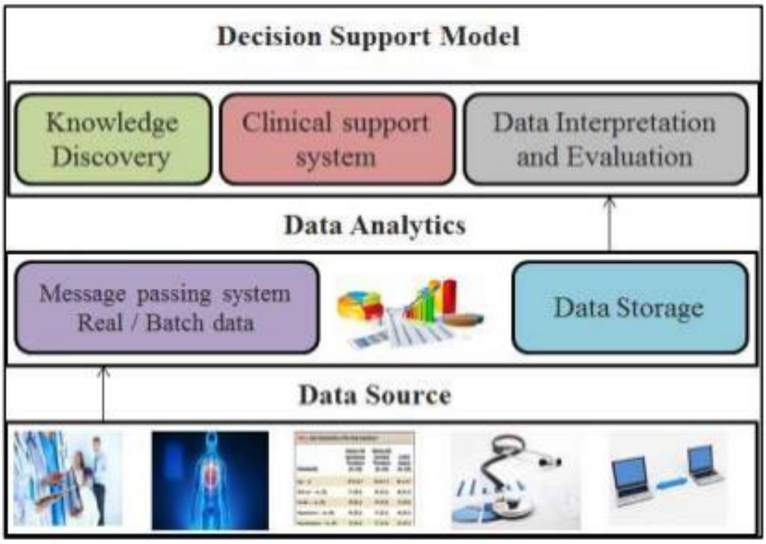

Figure 1 Healthcare analytics an overview

In order to create a decision plan, health care data analytics comprises the following activities: gathering health care data (patient behaviour, clinical outcomes, claims, and cost of incurring), interpreting the data, and looking for and analysing significant patterns in the data. Clinical data analytics is important for risk assessment, treatment cost reduction, patient behaviour prediction, and creating an analytical strategy for the relevant treatment. The significance of the hazards that typically arise as a reflection of treatment analysis is provided by the examination of trends across various forms of healthcare data [14]. Treatment costs can be significantly reduced if risk analysis is done beforehand. Healthcare data must be evaluated for decision-making at appropriate intervals due to the swift advancement of healthcare instruments and processes [9]. An overview of healthcare data collection and data-driven knowledge discovery is given in Figure 1. Core medical data is defined as reports and records that are consistent with its standards . Medical professionals are able to make decisions and understand results at the appropriate times by minimising data variability and optimising risk variables. Delivering appropriate treatment with a stabilised patient care monitoring system is the intervention in quality medicine [13]. Predictive data models could therefore provide medical professionals with access to the domain of disease-specific syndromes through processes and actions [10]. The field of healthcare data analytics strives to enhance clinical procedures while simultaneously improving risk factors [11].

The rest of the paper is organized as follows: Section 2 provides the classification scheme for the survey; Section 3 provides an overview of proposed architecture. Section 4 provides a summary and comparison of the results of the various papers discussed in this taxonomy. Finally, Section 5 concludes the paper.

2. Related Works

One new area of study for developing practical decision support systems in biomedical engineering is the application of swarm intelligence algorithms and data classification. Determining the difficulties with health care and fixing its irregularities are major issues [4]. In the field of healthcare analytics, the dataset must be analysed in order to find the method for identifying valid, usable, and eventually understandable data patterns. The high dimensional healthcare records with disease identification and investigation can be found and analysed with the aid of the data classification algorithm [5]. Prediction and data classification lie on the border between statistics reporting and the goal of finding novel, operationalizable insights. Healthcare data can be viewed as a thought of:

- Real time healthcare analysis
- Batch-based healthcare analysis

Disease-specific risk variables have varied implications that differ depending on the locale. One of the main challenges in computer-aided diagnosis in medical informatics is identifying such a disease-specific risk and its implications within a particular location [12]. The main goal of this study endeavour is to identify disease-specific risk factors with the intention of resolving their inconsistencies through the development of a decision-support model rather than region-based analysis. For each group of persons, a greater understanding of the risk factors and the best interventions are required in relation to the location, likelihood, and dietary practices.

3. Methodologies

Grey wolf optimisation is a swarm intelligence process that is well-known for its group pursuing and imitates the wolf leadership hierarchy. The Canidae family includes grey wolves, and they often prefer to live in packs. They have a strict social hierarchy, with Alpha (α)—a male or female—serving as the leader. Basic leadership is typically the responsibility of the alpha. The pack should follow the sets of the dominant wolf. The Alpha (β) is assisted in fundamental leadership duties by the Betas (β), subordinate wolves. The beta serves as the pack's disciplinarian and guide to the alpha. Omega (ω), the lower-positioned grey wolf, must present all other dominant wolves. In the unlikely event that wolves are neither alpha [7].

a. Algorithmic Steps

- Step 1: Strat the calculation the GWO parameters such as search agents (Gs), create changeable dimension (Gd), vectors a, A, C and greatest measure of iteration ($iter_{max}$).

$$\vec{A} = 2\vec{a}\mathrm{rand}_1 - \vec{a} \qquad (16)$$

$$\vec{C} = 2\mathrm{rand}_2 \qquad (17)$$

The principles of $\vec{a}$ are linearly reduced from 2 to 0 through the span of cycles.

- Step 2: Wolves are generated arbitrarily based on the size of the pack. logically, these can be communicated as,

$$\text{Wolves} = \begin{bmatrix} G_1^1 & G_2^1 & G_3^1 & \cdots\cdots & G_{Gd-1}^1 & G_{Gd}^1 \\ G_1^2 & G_2^2 & G_3^2 & \cdots\cdots & G_{Gd-1}^2 & G_{Gd}^2 \\ . & . & . & \cdots\cdots & . & . \\ . & . & . & \cdots\cdots & . & . \\ G_1^{Gs} & G_2^{Gs} & G_3^{Gs} & \cdots\cdots & G_{Gd-1}^{Gs} & G_{Gd}^{Gs} \end{bmatrix} \quad (18)$$

Where,

G_j^i is the underlying estimation of the j^th pack of the it^h wolves.

- Step 3: For each hunt agent the fitness value is estimated.

$$\vec{D} = \left|\vec{C}.\overrightarrow{G_P}\right|(t) - \vec{G}(t) \quad (19)$$

$$\vec{G}(t+1) = \overrightarrow{G_p}(t) - \vec{A}.\overrightarrow{D} \quad (20)$$

- Step 4: discover the suitable hunt agent (G_α), the subsequent suitable hunt agent (G_β) and the last suitable hunt agent (G_δ)

$$\overrightarrow{D_\alpha} = \left|\overrightarrow{C_1} \cdot \overrightarrow{G_\alpha} - \vec{G}\right| \quad (21)$$

$$\overrightarrow{D_\beta} = \left|\overrightarrow{C_2} \cdot \overrightarrow{G_\beta} - \vec{G}\right| \quad (22)$$

$$\overrightarrow{D_\delta} = \left|\overrightarrow{C_3} \cdot \overrightarrow{G_\delta} - \vec{G}\right| \quad (23)$$

$$\overrightarrow{G_1} = \overrightarrow{G_\alpha} - \overrightarrow{A_1} \cdot \left(\overrightarrow{D_a}\right) \quad (24)$$

$$\overrightarrow{G_2} = \overrightarrow{G_\beta} - \overrightarrow{A_2} \cdot \left(\overrightarrow{D_\beta}\right) \quad (25)$$

$$\overrightarrow{G_3} = \overrightarrow{G_\delta} - \overrightarrow{A_3} \cdot \left(\overrightarrow{D_\delta}\right) \quad (26)$$

- Step 5: the current hunt agent location is restored.

$$\vec{G}(t+1) = \frac{\overrightarrow{G_1}+\overrightarrow{G_2}+\overrightarrow{G_3}}{3} \quad (27)$$

- Step 6: for all the hunts the fitness values are estimated.
- Step 7: renew the value of G_a, G_β, G_δ
- Step 8: the stopping condition is checked. If the I_{ter} approximately equal to $Iter_{max}$, denote yes and print the best value of solution otherwise do again step 5.

The C4.5 decision tree algorithm is executed in Java to perform the fitness evaluation phase. Following the preparation of the training and test data using 10-fold cross-validation, the data is assessed for the creation of fitness values and advanced with the GWO algorithm's parameters [8–10].

4. Results and Discussion

Real-world and repository data have both been used in the experiments. After being gathered, preprocessed, and prepared for data processing following model development, the repository data have been processed. The experimental results from the heart disease, cancer, kidney disease and retinopathy dataset are shown in the next section.

Table 1: Performance metrics of various disease detection

Diseases	Models	Accuracy (%)	Specificity (%)	Sensitivity (%)	F1-Score	Precision (%)	Recall (%)
Chronic Kidney Disease	GWO	90	92	91	0.85	90	89
	Decision tree	95	95.16	95.15	0.97	97	89.85
	GWO + Decision tree	96.50	98	96.74	0.985	97.50	90.25
Diabetes	GWO	80	95.16	95.15	0.85	90	89
	Decision tree	85	98	96.74	0.97	97	89.85
	GWO + Decision tree	90	97.05	98	0.985	97.50	90.25
Cardio Vascular	GWO	87	92	91	0.85	90	89
	Decision tree	85	92.4	91.8	0.97	97	89.85

	GWO + Decision tree	86.50	95.16	95.15	0.985	97.50	90.25
Alzheimer's	GWO	90	95.16	95.15	0.85	90	89
	Decision tree	95	98	96.74	0.97	97	89.85
	GWO + Decision tree	91.50	90.05	91	0.985	97.50	90.25
Cancer	GWO	91	92	91	0.85	90	89
	Decision tree	92	95.16	95.15	0.97	97	89.85
	GWO + Decision tree	92.50	98	96.74	0.985	97.50	90.25

Given that the dataset was gathered in a rural area, it presents the results in terms of the risk factors that have the greatest influence and their linkage. This will assist the medical professionals in determining how the risk variables relate to one another in light of their different levels.

5. Conclusions

Medical data analysis heavily relies on data classification and prediction optimisation. It appears that gathering and analysing medical data with its range of values is a difficult task. The study endeavour as a whole denotes the creation of a decision support model featuring an enhanced PSO and Decision Tree algorithm. Fixing the operational parameter—a modified self-adaptive inertial weight with convergence logic—allows the particles to traverse the specified search space and discover a solution that is almost optimal. This represents an improvement. For analysis, a dataset including 732 records and 23 attributes, including the class label, was employed. The test result shows that, with an enhanced accuracy of roughly 98.60%, the suggested strategy generates four significant risk variables. The suggested model's efficacy has been verified by data analysis, and multiple writers have provided the methodology.

References

1. Walsh, Seán, Evelyn EC de Jong, Janna E. van Timmeren, Abdalla Ibrahim, Inge Compter, Jurgen Peerlings, Sebastian Sanduleanu et al. "Decision support systems in oncology." JCO

clinical cancer informatics 3 (2019): 1-9.

2. Cresswell, Kathrin, Margaret Callaghan, Sheraz Khan, Zakariya Sheikh, Hajar Mozaffar, and Aziz Sheikh. "Investigating the use of data-driven artificial intelligence in computerised decision support systems for health and social care: a systematic review." Health Informatics Journal 26, no. 3 (2020): 2138-2147.
3. S. Neelima, Manoj Govindaraj, Dr.K. Subramani, Ahmed ALkhayyat, & Dr. Chippy Mohan. (2024). Factors Influencing Data Utilization and Performance of Health Management Information Systems: A Case Study. Indian Journal of Information Sources and Services, 14(2), 146–152. https://doi.org/10.51983/ijiss-2024.14.2.21
4. Scott, Philip J., Angela W. Brown, Taiwo Adedeji, Jeremy C. Wyatt, Andrew Georgiou, Eric L. Eisenstein, and Charles P. Friedman. "A review of measurement practice in studies of clinical decision support systems 1998–2017." Journal of the American Medical Informatics Association 26, no. 10 (2019): 1120-1128.
5. Musen, Mark A., Blackford Middleton, and Robert A. Greenes. "Clinical decision-support systems." In Biomedical informatics: computer applications in health care and biomedicine, pp. 795-840. Cham: Springer International Publishing, 2021.
6. Madhavi, M., Sasirooba, T., & Kumar, G. K. (2023). Hiding Sensitive Medical Data Using Simple and Pre-Large Rain Optimization Algorithm through Data Removal for E-Health System. Journal of Internet Services and Information Security, 13, 177-192.
7. Martin-Sanchez, F., and V. Aguiar-Pulido. "Analytics and decision support systems in global health informatics." In Global Health Informatics, pp. 195-217. Academic Press, 2017.
8. Melton, Brittany L. "Systematic review of medical informatics–supported medication decision making." Biomedical informatics insights 9 (2017): 1178222617697975.
9. Mohamed, K.N.R., Nijaguna, G.S., Pushpa, Dayanand, L.N., Naga, R.M., & Zameer, AA. (2024). A Comprehensive Approach to a Hybrid Blockchain Framework for Multimedia Data Processing and Analysis in IoT-Healthcare. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), 15(2), 94-108. https://doi.org/10.58346/JOWUA.2024.I2.007
10. N, Ravi, and Swanand Kulkarni. 2023. "Smart Ways to Catch the Abutment DRCs at IP Level". Journal of VLSI Circuits and Systems 6 (1):51-54. https://doi.org/10.31838/jvcs/06.01.08
11. Beard, Rachel, Elizabeth Wentz, and Matthew Scotch. "A systematic review of spatial decision support systems in public health informatics supporting the identification of high risk areas for zoonotic disease outbreaks." International journal of health geographics 17 (2018): 1-19.
12. Khairat, Saif, David Marc, William Crosby, and Ali Al Sanousi. "Reasons for physicians not adopting clinical decision support systems: critical analysis." JMIR medical informatics 6, no. 2 (2018): e8912.
13. Moreira, Mário WL, Joel JPC Rodrigues, Valery Korotaev, Jalal Al-Muhtadi, and Neeraj Kumar. "A comprehensive review on smart decision support systems for health care." IEEE Systems Journal 13, no. 3 (2019): 3536-3545.
14. Bobir, A.O., Askariy, M., Otabek, Y.Y., Nodir, R.K., Rakhima, A., Zukhra, Z.Y., Sherzod, A.A. (2024). Utilizing Deep Learning and the Internet of Things to Monitor the Health of Aquatic Ecosystems to Conserve Biodiversity. Natural and Engineering Sciences, 9(1), 72-83.
15. Dixon, Brian E., Hadi Kharrazi, and Sridhar R. Papagari Sangareddy. "Public health decision support systems." Public Health Informatics and Information Systems (2020): 349-365.