A Modified Grey Wolf Optimizer algorithm for feature selection to predict heart diseases

Shaik Khaja Mohiddin¹, Professor, Department of CSE, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.

Susan Peteti², Tummala Swathi³, Tambura Veera Venkata Harshith⁴, Vamshi Krishnamaneni⁵, Vatluri Hanusha⁶
²,³,⁴,⁵,⁶ UG Students, Department of CSE, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.
susanpeteti002@gmail.com², swathi.tummala12@gmail.com³, harshithram234@gmail.com⁴, vamshikrishnamaneni@gmail.com⁵, hanushavatluri57340@gmail.com⁶

Abstract
Globally, heart disease is a leading cause of illness and mortality. This impacts people from all around the world. Accurate prediction of the risk of heart disease is crucial for early detection and prevention. For this, large amounts of features/attributes need to be stored and analyzed to diagnose a patient. Storing many features can lead to substandard management of data. We need to store only the chief features. In this study, we proposed a modified grey wolf optimizer for feature selection.

The resultant subset of features is then used to predict the risk of having a heart disease using machine learning model, Support Vector Machine (SVM). We compared the proposed algorithm with the existing GWO-SVM algorithm. We evaluated the effectiveness of the proposed algorithm using accuracy, sensitivity, and specificity metrics. Our results show that, using the modified grey wolf algorithm for feature selection and using SVM we obtained an accuracy of 95.82%, specificity of 94.64%, and sensitivity of 96.86%. The results show the proposed algorithm's capability for predicting the risk of heart disease and could contribute to the development of more accurate and efficient predictive models for heart disease risk assessment.

Keywords: Feature Selection, Grey Wolf Optimizer, Heart Disease Prediction, Machine Learning, Meta-heuristic algorithm, Optimization.

Introduction
With an expected 17.9 million deaths from heart disease in 2019, it is a significant global source of mortality and morbidity (World Health Organization, 2021). Early detection and accurate prediction to the risk of heart disease can help healthcare providers design appropriate interventions to prevent or delay the onset of the disease, as well as reduce the burden of healthcare costs associated with the disease.
To diagnose a patient with heart disease, data containing various features of the patient need to be stored and analyzed. Storing such large amounts of data can lead to substandard management of data and might even lead to a wrong diagnosis. Therefore, it is important to store only the chief features so that a correct diagnosis can be made while being able to maintain good condition of data. To store only the important data, we reduce the dimensionality of the data. One method of dimensionality reduction is through the use of feature selection algorithms. These techniques can choose the optimal features in a dataset.

There are numerous feature selection methods available, but in recent years, optimization issues have been solved using Nature Inspired Algorithms (NIA). These algorithms are also known a meta-heuristic algorithm because they are used minimize or maximize a solution of a problem.

A novel meta-heuristic algorithm dubbed Grey Wolf Optimizer (GWO) algorithm was released in 2014 by Seyedali et al. [1]. The social structure of grey wolves served as an inspiration for this algorithm. The mathematical model for this algorithm is built using their hunting patterns. The two components of the GWO algorithm are exploration and exploitation. The exploration part of the algorithm is the process of exploring the search space. The exploitation part of the algorithm is the process of local searching in promising areas in exploration part.

Hou et al.[11] presented a strategy to enhance this search space exploration to determine the path of mobile robots. However, it uses a non-linear convergence factor based on normal distribution for the coefficient ‘a’, which is not needed for our application. This is taken as inspiration to the algorithm proposed in this paper.

**Literature Survey**

Many researchers in the past have used Grey Wolf Optimizer algorithm for feature selection as well as for other applications. The social structure and hunting behaviours of grey wolves served as the main inspiration for this algorithm, which was first presented by Seyedali et al. [1] in 2014.

Some of the previous works based on this algorithm where it is specifically used for feature selection in predicting heart disease is using SVM. The GWO approach is used to select features, and an SVM model is then used to predict heart disease using the resulting subset of features [2]. Singh and Kumar proposed various machine learning models that can be

A new variation of GWO known as RW-GWO that makes use of a random walk to expand the search space was put forth by Preet and Kusum Deep [6]. Another NIA for predicting both heart disease and breast cancer was suggested by Sireesha et al. [7]. It claims that one of the hybrid models that are suggested outperforms traditional models.

Turabieh [8] proposed a deep learning model ANN is used to predict heart disease. It uses GWO for finding initial weights and biases for inputs to ANN. BakrawyLamiaa shows GWO is integrated with naïve bayes to predict heart diseases. Additionally, the dataset is subjected to a 5-fold cross validation. It states that the results achieved were better than traditional naïve bayes classifier [9].

Finally, Luis Rodríguez et al. shows fuzzy logic i.e., a fuzzy hierarchical operator used in GWO for adding weights to the search agents [10]. These weights are used as an inspiration to the modified algorithm proposed in this paper.

Hou et. al. proposed an improved grey wolf optimizer for finding path in mobile robots. This algorithm overcomes the problem of convergence speed and instability in GWO and has a better optimal search capacity [11]. The methodology proposed here is modified to be used for our application. [13-21]

Problem Identification
A. Grey Wolf Optimizer
The population-based optimization algorithm known as the Grey Wolf Optimizer was modelled by the hunting strategies of grey wolves. The social hierarchy of grey wolves serves as the basis for the search agents in this algorithm. Figure 1 shows the social hierarchy of grey wolves.

- Alpha (α): This search agent acts as the best solution. It is responsible for making decision and leading the hunt. After the hunt, the alpha wolf has its share first.
- Beta (β): It search agent acts as the second-best solution. It helps the alpha wolves in hunting the prey.
- Delta (δ): It search agent acts as the third best solution. Delta wolves of the pack are the food providers.
• Omega (ω): It represents the solution that follow alpha, beta and delta solutions. It consists of elders or the weakest wolves. They often act as the scapegoat.

![Diagram of social hierarchy of grey wolves]

**Figure 1:** Social hierarchy of grey wolves

**B. Mathematical Model of GWO**

The hunting process is divided into exploration, encircling and exploitation. The mathematical model is also based on these steps. The encircling of the prey is represented as:

\[
D = |C \cdot X_p - X(t)| \quad (1)
\]

\[
X(t + 1) = |X(t) - A \cdot D| \quad (2)
\]

\[
A = 2 \cdot a \cdot r_1 - a \quad (3)
\]

\[
C = 2 \cdot r_2 \quad (4)
\]

Where ‘a’ linearly decreases from 2 to 0 and r1, r2 both are random numbers in the range [0,1]. The parameter ‘a’ regulates the transition between exploitation and exploration. It is reduced by using the equation from 2 to 0. Eq (5) is shown below:

\[
a = 2 - t \cdot \frac{2}{\text{max}\_\text{iter}} \quad (5)
\]

Where, t represents the current iteration and max_iter represents the maximum number of iterations the GWO algorithm permits.

The exploitation of the prey is represented as:

\[
D_\alpha = |C \cdot X_\alpha - X(t)|
\]

\[
D_\beta = |C \cdot X_\beta - X(t)|
\]
In terms of alpha, beta, and delta wolves, the respective positions are,

\[ X_1 = |X_\alpha - A.D_\alpha| \]

\[ X_2 = |X_\beta - A.D_\beta| \]

\[ X_3 = |X_\delta - A.D_\delta| \]  \hspace{1cm} (7)

The equation to update the position of the search agent is shown below in eq. (8);

\[ X(t+1) = \frac{X_1 + X_2 + X_3}{3} \]  \hspace{1cm} (8)

### Proposed Methodology

The system architecture of the proposed methodology consists of acquiring data, performing data pre-processing, feature selection and building classification model. Figure 2 shows the system architecture.

Nitin Mittal et al. [12] proposed two ways in which the estimate of parameter a can decrease from 2 to 0. One of those ways is
In the exploitation step, the alpha wolf which is the strongest must be closest to the prey to make the hunt successful. The weights are, say, $w_1$, $w_2$, and $w_3$ of alpha, beta, and delta search agents respectively. Hou et al. [11] proposed similar equations for updating the locations of search agents using weights.

Although, this algorithm performs well for the application of finding the path for mobile robots.

However, our application does not use the convergence factor proposed by Hou et al. [11]. Using only the equations above mentioned did not give satisfactory results.

This might be because eq. (11) uses the absolute values of $X_1$, $X_2$, and $X_3$ to give importance to the positions.

This is giving values of search agents that are outside the boundaries of the search space because the equation for coefficient ‘$a$’ proposed by Hou et al. [11] is not used. Therefore, the equations are modified according to the coefficient ‘$a$’ that decreases linearly from 2 to 0 (eq. (9)) with the condition that if the positions are still out of bounds of search space they will not be changed.

So, the equations are modified as,

$$W_a = \frac{f_1 + f_2 + f_3}{f_1}$$

$$W_b = \frac{f_1 + f_2 + f_3}{f_2}$$

$$W_\delta = \frac{f_1 + f_2 + f_3}{f_3}$$

(10)

$$P_1 = \frac{X_1 + X_2 + X_3}{X_1}$$

$$P_2 = \frac{X_1 + X_2 + X_3}{X_2}$$
The MGWO algorithm’s pseudo-code is shown below:

1. Initialize grey wolf population
2. Initialize $a$, $A$, and $C$
3. Calculate fitness values of search agents
4. $X_\alpha$ = best solution; $X_\beta$ = second-best solution; $X_\delta$ = third best solution
5. while $t < \text{max\_iter}$
6. for each search agent
7. update $a$, $A$, and $C$
8. calculate the weights and positions using (10) and (11) respectively
9. update the position of current search agent using (12)
10. calculate fitness values of search agents
11. update $X_\alpha$, $X_\beta$, and $X_\delta$.
12. $t = t + 1$
13. return $X_\alpha$

**Implementation**

**A. Experimental Setup**

The MGWO algorithm is developed in MATLAB R2022b and the machine learning models are implemented in Jupyter notebook. They are operated by a computer with an Intel Core i5 processor with 8GB RAM capacity.
B. Dataset

The dataset is publicly available and one can download it from UCI machine learning repository. The dataset is taken from Cleveland Clinical Foundation.

C. Parameter Setting

The criteria used to set up the suggested grey wolf optimization algorithm are shown in Table 1. The values of lower bound and upper bound changes according to the benchmark function that is being used.

Here, the “F7” benchmark function was used. [1] shows the various benchmark functions that can be used for grey wolf optimization.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Initialized values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum no. of Iterations</td>
<td>100</td>
</tr>
<tr>
<td>No. of Wolves</td>
<td>5</td>
</tr>
<tr>
<td>No. of Dimensions</td>
<td>13</td>
</tr>
<tr>
<td>Selection Threshold</td>
<td>0</td>
</tr>
<tr>
<td>Upper bound</td>
<td>1.28</td>
</tr>
<tr>
<td>Lower bound</td>
<td>-1.28</td>
</tr>
</tbody>
</table>

Table 1: Parameter settings for the proposed method

Results and Discussions

The preprocessing and feature selection is performed before building the machine learning models. There were no null values or missing values in the dataset. As mentioned above, Modified Grey Wolf Optimizer is used for performing feature selection and resultant set of features are used for training and testing the models. Then the data is standardized using Standard Scaler.

The models are then created and trained on the resulting set of features. The different machine learning models’ performance is determined using measures like accuracy, sensitivity, and specificity.

Table 2 conveys the performance evaluation measures of the proposed algorithm i.e., MGWO-SVM with the existing GWO-SVM. The accuracy of MGWO-SVM is 95.82% and the
sensitivity is 96.86% and the specificity is 94.64%. The same set of features are selected with each execution of the algorithm; therefore, the results are the same for each execution.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GWO-SVM</td>
<td>89.83%</td>
<td>93%</td>
<td>91%</td>
</tr>
<tr>
<td>MGWO-SVM</td>
<td>95.82%</td>
<td>96.86%</td>
<td>94.64%</td>
</tr>
</tbody>
</table>

**Table 2**: Comparison of proposed model to existing model

Other machine learning models are also built to make predictions. Their results are shown in Table 3.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MGWO-LR</td>
<td>86.91%</td>
<td>91.62%</td>
<td>81.55%</td>
</tr>
<tr>
<td>MGWO-KNN</td>
<td>87.46%</td>
<td>88.48%</td>
<td>86.31%</td>
</tr>
</tbody>
</table>

**Table 3**: Results of other models

Here, LR stands for Logistic Regression.

The better performance of MGWO compared to existing GWO algorithms might be due to:

- The improved search space exploration using proportional weighting strategy that is inspired from [11]. This helps in finding optimum solution more accurately.
- The dependable way of finding alpha, beta, and delta search agents when implemented in python.

**Conclusion**

This paper presented a modified Grey Wolf Optimizer algorithm for feature selection to find optimal features in a dataset. The features obtained from MGWO algorithm is then used to build SVM model.
It is observed that the model gives an accuracy of 95.82% and sensitivity of 96.86% and a specificity of 94.64%. When losing even one important feature reduces the accuracy of the model. The proposed algorithm MGWO gives the same set of features with each execution and all the important features are always selected. The performance of MGWO-SVM further supports the conclusion that MGWO is better for the application of feature selection.

Limitations
This study only used one dataset to perform feature selection using MGWO. This study used only a one machine learning model. Accurate predictions also depend on the type of classifier. Working with more advance models might give more accurate predictions.

Future Scope
Larger datasets can be used for feature selection using the proposed algorithm. More types of machine learning models or even deep learning models can be used for predicting heart diseases.

References


