**RAJKIYA ENGINEERING COLLEGE AMBEDKARNAGAR**

Department of Information Technology

**B.TECH. (IT) FINAL YEAR PROJECT SYNOPSIS**

Project Group :No. 9 allotted project no.29 Allotted Project No.29

Name of Group Members:

1.Yashvendra Kumar Roll no: (1873713059)

2.Diwakar Shubham kr Ashok Roll no: (1873713019)

* Title of the Project: Clustering in medical Diagnosis Using quantum inspired grey wolf optimizer
* Name of Supervisor(s): **Mr Amit kumar**
* Brief Introduction of the project:

**Medical Diagnosis :** In order to make the best medical decisions, medical diagnosis plays a very important role for medical institutions. As everyone knows, false medical diagnoses will lead to incorrect medical decisions, which are likely to cause delays in medical treatment or even loss of patients’ life. Recently, a number of computer aided models have been proposed for diagnosing different kinds of diseases, such as diagnostic models for Parkinson’s disease [, breast cancer heart disease and Alzheimer’s disease . As a matter of fact, medical diagnosis could be treated as a problem of classification.

**Clustering of Medical Data** Clustering, which is an unsupervised learning technique, has been widely applied in diverse field of studies such as machine learning, data mining, pattern recognition, image analysis, and bioinformatics, medical diagnosis.

Clustering is a very useful and significant exploratory technique for the analysis and Diagnosis of Medical data. Many classical and heuristic clustering algorithms are being used for clustering Such as GA

,grey wolf optimization (GWO) is a new EC technique proposed recently . GWO mimics the social hierarchy and hunting behavior of grey wolves in nature. Due to its excellent search capacity, it has been successfully applied to many real-world problems since its introduction, like optimal reactive power dispatch problem , parameter estimation in surface waves , static VAR compensator controller design , blackout risk prevention in a smart grid , capacitated vehicle routing problem , nonconvex economic load dispatch problem , and so on. However, it should be noted that the initial population of original GWO is generated in a random way. It may result in the lack of diversity for the wolf swarms during the search space. Many studies have shown that the quality of the initial population may have a big impact on the global convergence speed and the quality of final solution for the swarm intelligence optimization algorithms, and initial population with good diversity is very helpful to improve the performance of optimization algorithms.

**4 .Brief review of the related work**: The project topic is very interesting we are going to form Clustering of Medical Diagnosis Data with the help of machine learning by using GWO algorithm for better perfomance and result of Clustering .

5.**Tech Stack/Tools/Resources required/to be used:** Grey wolf Optimizer algorithm, Machine Learning, Python IdE with good internet connectivity.

1. **For Medical Diagnosis Clustering:**

a)Data Analysis

b)Machine Learning

c)Medical diagnosis Data Set.

d)Clustering Algorithm:

* GWO/IGWO

Hardware Requirements :

**8GB** RAM, **CORE i5** PROCESSOR, **10GB** HARDISK SPACE

6**.Brief description of methodology of the project work:**

**Algorithm models/Grey Wolf Optimizer**

**Social hierarchy**

In order to mathematically model the social hierarchy of wolves when designing GWO, we consider the fittest solution as the alpha (α). Consequently, the second and third best solutions are named beta (β) and delta (δ) respectively. The rest of the candidate solutions are assumed to be omega (ω). In the GWO algorithm the hunting (optimization) is guided by α, β, and δ. The ω wolves follow these three wolves

### [https://upload.wikimedia.org/wikipedia/commons/2/2c/Grey_wolf_social_hierarchy.jpg](https://en.wikiversity.org/wiki/File:Grey_wolf_social_hierarchy.jpg)

**Search for prey (exploration)**

grey wolves mostly search according to the position of the alpha, beta,

and delta. They diverge from each other to search for prey and converge

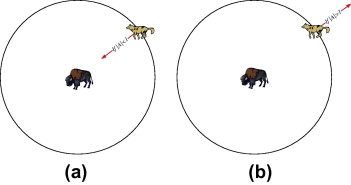
to attack prey

In GWO it is modeled by utilizing A with random values greater the 1

or less than -1 so the search agent could diverge from the pray

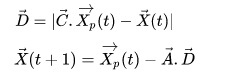
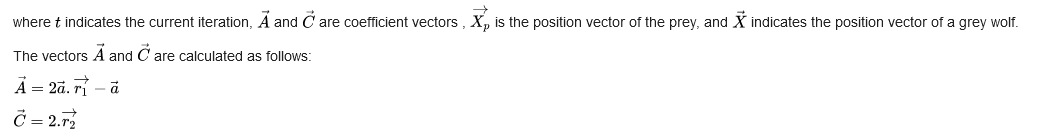
When |A| > 1 wolves are forced to diverge from the pray and find better

one



**Encircling prey**

As mentioned above, grey wolves encircle prey during the hunt. In order to mathematically model encircling behavior the following equations are proposed

****

**Attacking prey**

Grey wolves finish the hunt by attacking the

prey when it stops moving.

In order to mathematically model sere

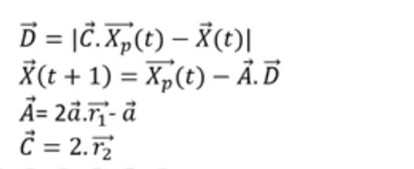
the prey we decrease the value of a.

The fluctuation range of A is also decreased by

G. In other words A is a random value in the

interval [-2a,2a] where a is decreased from 2

to 0 over the course of iterations.



ALGO

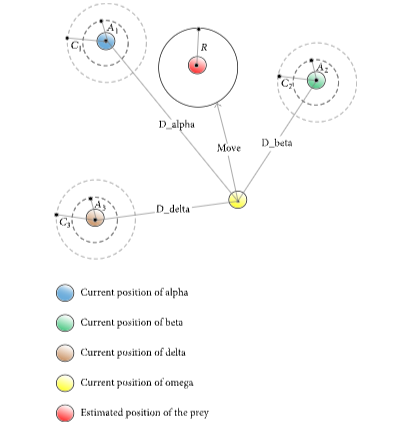
Initialize the grey wolf population Xi (i= 1, 2, .... n)

* Initialize a, A, and C .
* Calculate the fitness of each search agent
* Xalpha =the best search agent
* Xbeta=the second best search agent
* Xgama =the third best search agent
* Effects of A on the exploration and
* while (t< Max number of iterations) exploitation of GWO
* for each search agent
* Update the position of the current search agent by above equations
* end for
* Update a, A, and C
* Calculate the fitness of all search agents
* Update X alpha, X beta, and X gama
* t=t+1
* end while

* return X alpha

To see how GWO is theoretically able to solve optimization problems, some points may be noted:

* The proposed social hierarchy assists GWO to save the best solutions obtained so far over the course of the iteration
* The proposed encircling mechanism defines a circle-shaped neighbourhood around the solutions which can be extended to higher dimensions as a hyper-sphere
* The random parameters A and C assist candidate solutions to have hyper-spheres with different random radii
* The proposed hunting method allows candidate solutions to locate the probable position of the prey
* Exploration and exploitation are guaranteed by the adaptive values of a and A
* The adaptive values of parameters a and A allow GWO to smoothly transition between exploration and exploitation
* With decreasing A, half of the iterations are devoted to exploration (|A|≥1) and the other half are dedicated to exploitation (|A|<1)
* The GWO has only two main parameters to be adjusted (a and C)



* Major inputs (infrastructure) required:

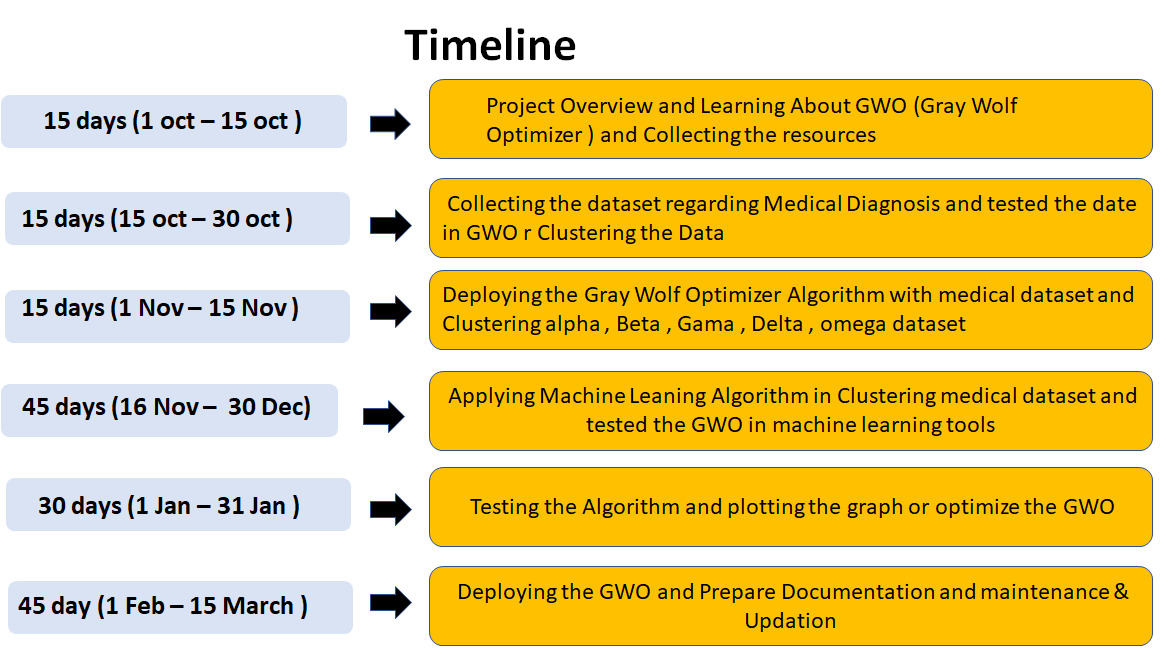
1. Dataset for training algorithms, Sequence of Unknown Protein, **protein structures in the CSV format.**

* List of references:
* <https://www.researchgate.net/publication/260010809_Grey_Wolf_Optimizer>
* <https://en.wikiversity.org/wiki/Algorithm_models/Grey_Wolf_Optimizer>
* <https://ieeexplore.ieee.org/document/7867415>
* <https://www.nature.com/articles/s41598-019-43546-3>

<https://www.kaggle.com/search?q=medical+disease+data+analysis>

(<https://www.kaggle.com/search?q=medical+disease+data+analysis> for data set collection)

* Schedule of Activities (PERT Chart):



Name & Signature of the project member(s):

1.Yashvendra Kumar

2.Diwakar Shubham kr Ashok

Name & Signature of the Supervisor(s):

1.Mr Amit Kumar