International Journal of Advance Research in Science and Engineering
Volume No. 10, Issue No. 12, December 2021
www.ijarse.com
IJARSE
ISSN 2319 - 8354

A Systematic Study on views and perception about Optimization and Classification using various algorithms for Predictive Analytics

Nithya G[1], Dr. P S S Akhilashri[2]

[1]*Department of Computer Science, Kongu College of Arts and Science, Karur, India*

[2]*Department of Computer Science, National College, Tiruchirapalli, India*

Abstract

This paper emphasis on the "Data Analytics Concepts". Albeit, various authors dealt with this, the underlining principle is that to eliminate the real-world problems and the uncertainties in every stage, how the predictive analytics is used for optimization problems using soft computing and swarm intelligence algorithms. The application of Swarm, prediction has been used in the context of forecasting problems. Predictive analytics give importance to forecasts, which is uncertain, whereas the Swarm Intelligence Approaches with high dimensionality and dynamic language of data. In this concept, data are being analyzed properly, where exploration and exploitation are fully optimized. There are few characteristics that provide single optimization and multi-objective optimization viz., for many domains. The objective of this paper is to identify the different perceptions and views that made to optimization and accuracy for COVID19 prediction especially using clinical findings which is the need of the hour due to severity increases one after another in covid which is becoming more challenges for health care community.

Keywords: Classification, Covid 19, Data Analytics, Fuzzy, Meta Heuristics, Prediction, Swarm Intelligence

1.0 Introduction

Problem needs either optimized or accurate solution. But in the process of finding solution for a problem it is better to get highly achievable performance with most cost effective and less time-consuming process. This can be achieved by minimizing the unwanted factors and maximizing the most desirable factors. Real life problems are having complexities like conventional, linear, nonlinear, discontinuous, discrete, equality and inequality constraints etc. Always conventional methods may face difficulties in finding optimal solution. There are various optimization problem types like continuous or discrete, unconstrained or constrained, single or multiple-objective, deterministic or stochastic problems. Mostly real-world problems are continuous, constrained, multi-objective and stochastic in

nature. To focus on an extensible data analytics[17] platform Data Analytics will help find the insight hidden in the data.

It gives us an analysis based on historical/statistical data[18]. For any type of problem, albeit, in predictive analysis, the applicability of the Methodologies using either Mathematical and Statistical Tool, Soft Computing, or Social Computing Model, may provide us with the optimal solution. In recent technologies, data has become more accessible and ubiquitous therefore there are many tools to convert data into useful and actionable information. To analyze the data, the basic step[1] is to perform certain cautions on the available related historical data.

Steps for analysis are

1. Collect the data
2. Clean up the data
3. Analyze the data using a suitable model
4. Predict the future
5. If more data is available then analyze the data frequently to increase the confidence, then the more accurate prediction will be identified.

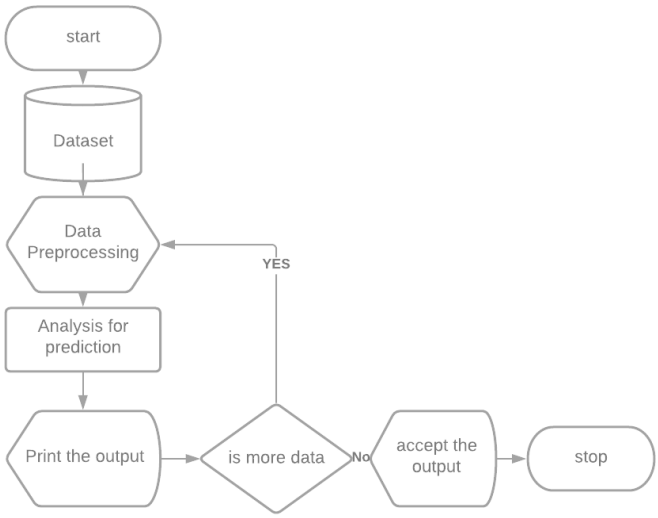

Fig 1

The above image Fig 1 explain the steps for analysis. From the transformed data, by applying proper interpretation, useful knowledge like pattern analyzing, pattern and knowledge discovery etc., are obtained for prediction and prescription about the future. In the domain of covid 19 the data are available and according to the objective, suitable predictive model is applied on it and able to predict the output for the problem.

Predictive analytics[16] used many optimization techniques from data mining statistics, machine learning, soft computing, and artificial intelligence especially swarm intelligence[22]. It analyzes the current data to make predictions of the future. It gives accurate insight into any question and allows users to create forecast. For every

individual objective value is calculated, then make the relationship between the individual and it will choose which solution is better among them.

Analytics can be done by creating a model. Predictive analytics is also executed with the help of predictive models. Every model needs standardized data where data is the feed to predict the future. It can be a historical or transactional data. Every data set can be categorized & belonging under either structured, unstructured and semi structured. Initially it is necessary to identify where the problem occurs, what kind of problems in which domain, then data is to be collected according to its domain. Then it must find the analytics techniques suitable for it to find the solution that will be applied. So that, using current data, analyze and make predictions about future.

2.Related Works:

The objective of this study is to understand the design and application of a prediction algorithm that is used for early prediction of covid 19 for any corner of the world. Such that the list of objectives that are observed in this study are discussed for further analysis.

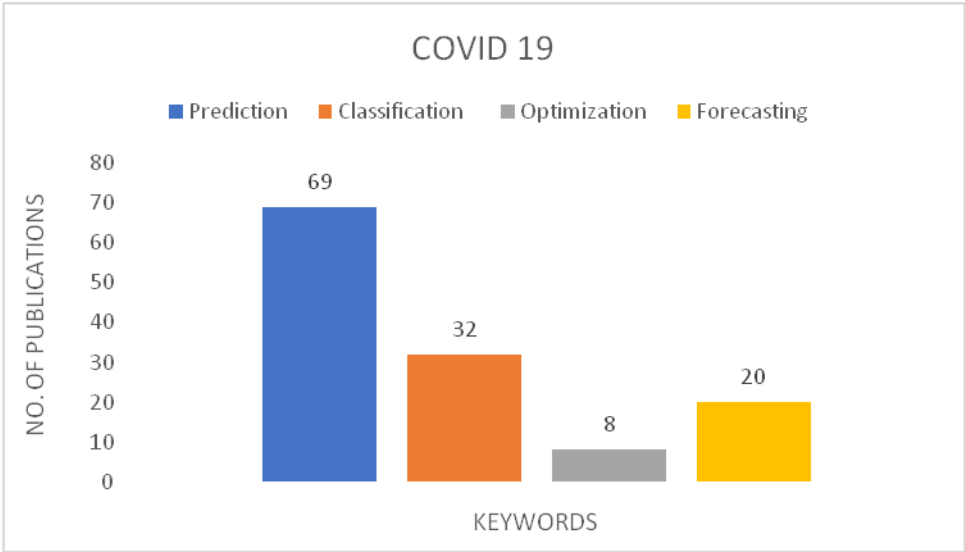

Fig 2

This paper is going to discuss the survey with the following factors.

a. Various authors and their perspectives on covid 19 analysis and their objectives.

b. Different types of datasets used for covid 19 analysis.

c. Role of feature selection on datasets in covid 19 analysis.

d. Different bio-inspired algorithms which may be suitable for covid 19 analysis.

e. Performance analysis of all the works based on accuracy as metrics in analysis.

f. Number of papers which has been using bio-inspired, fuzzy logic, standard classification model.

While searching with the keywords Classification, Prediction, Forecasting and Optimization, around 9000 papers were collected. We tried to classify the papers with filters especially with covid 19, machine learning models,

datasets etc the following graph represent the complete details about the publications. More than 120 were finally selected. Among them, 69 papers were selected under prediction, 32 papers under classification, 8 papers under optimization and 20 papers under Forecasting. These papers impacting the following concepts with certain objectives. They split under various sections as below (Fig 2)

2.1 To identify a different motive in the perspectives of different authors

Various objectives made by the authors on covid 19 analysis are:

1. With covid positive, predict cured, death using Random Forest
2. Prediction of covid high risk using clinical findings with the help of classification algorithm
3. Prediction of covid using clinical findings with the help of classification models
4. Prediction of survival and death rates using classification algorithm by identifying correlation between attributes
5. To predict covid death using deep learning, features (where combining features converted into categorical values 0 or 1)

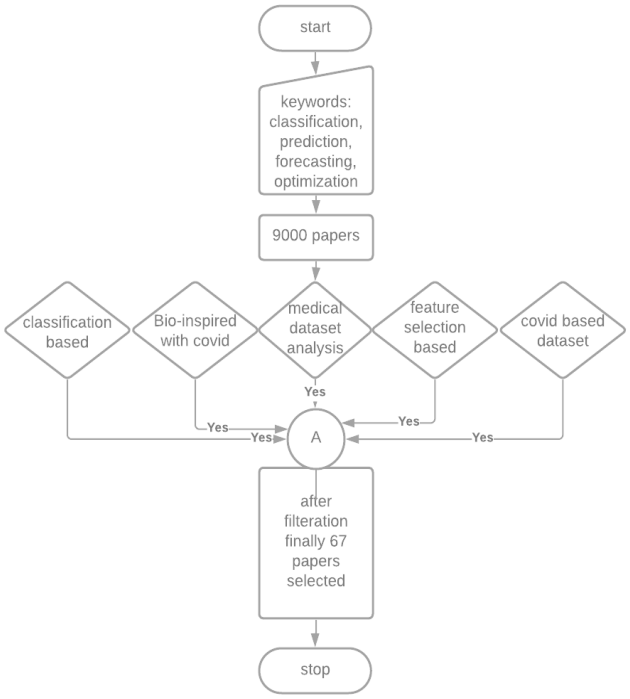

Fig 3

The above flowchart (Fig 3) clearly explains how the paper carried out the survey

6. Due to infection spread, how covid became positive using optimization algorithm framed as, hybridizing with a deep learning algorithm.

7. Early prediction of covid patients and identifying dangerous features that made the patients to death using descriptive analytics
8. Forming clusters of clinical features and identifying effects of it on patients using descriptive approaches.
9. To improve the reliability of prediction with the help of impact scenarios in past used for future prediction using Particle swarm optimization
10. To identify the right treatment for the right patient(covid) using fuzzy set theory.

2.2 Different types of datasets used for covid 19 analysis

Various researchers have studied the influence of Demographic, Cohort, Epidemiological, Clinical factors on the domain covid 19 and its prediction. Every dataset has various kinds of features Outcomes and accuracy will be changing according to the model that is applied to it. The following table(Table 1) clearly explains how various authors handled the dataset and their work for feature extraction/ feature selection.

S.No.	Source	Features of covid in the dataset	Keywords	Outcome / Accuracy
1	Kaggle[39]	Date,time,Confirmed state,confirmed national, confirmed international	coronavirus; COVID-19; respiratory tract; multi-class classification; random forest	RF- 83.54
2	Github Wolfram[50]	Patient id , age, sex, province, country, dataonset symptoms, data_admission_hospital, date_confirmation, symptoms, lives_in wuhan, travel_history, travel_history_location, chronic_disease, outcome, date_death_or_discharge, notes for discussion, location,country_new, admin_id	COVID-19, machine learning, deep learning, pandemic, rare event, fatality prediction	Deep learning, metrics- accuracy. Number of hospital resources required in the city are identified.

3	Github https://github.com/burakalakuss/COVID-19-Clinical[40]	Patient_id, age, hematrocit, hemoglobin, platelets, red blood cells. Lymphocytes, Leukocytes, Basophils, Eosinophils, Monocytes, serum, Glucose,Neutrophils, Urea , Protein, C-reactive, Creatinine, Potassium, Sodium Alanine transaminase, Aspartate transaminase,label	SARS-CoV2, COVID-19, coronavirus, deep learning, artificial intelligence	92.30% (identifying the severity of illness is done)
4	Hospital dataset (china)[41]	Myo, CD8, age, LDH, LMR, CD45, Th/ts, dyspnea, NCR, D-Dimer, CK	COVID-19, Patient outcome, Descriptive study, Random forest algorithm	Accuracy predicted as 86.6
5	PubMed, IEEE Xplore, Google Scholar, Science Direct[42]	Types of Treatment , Around 7 types of treatment were discussed through Allopathi and 5 types of approach were discussed through Ayurvedha.	SARS-CoV-2; hesitant fuzzy linguistic term sets; computational based, AHP-TOPSIS method, Course of Treatment, Drug Prioritization, Treatment selection. Multi Criteria Decision Making, Computational Modeling. Disease, COVID-19	The results obtained by HFLTS-AHP-TOPSIS and Fuzzy-Delphi-AHP-TOPSIS have a linear and continuous value calculation without any sensitive fluctuation
6	Hospital , directly from patients @ china[43]	Age, Comorbodities, RALE score and biomarkers of Systematic hyper inflammation namely Lymphopenia, elevated LDH, D-Dimer	COVID-19 ARDS Infection RALE score	Identified risk factors of the patients
7	Github China[44]	Clinical- 33 symptoms were notified	COVID-19 Epidemiology Clinical features Pneumonia SARS-CoV-2	Patients need ICU or NON ICU care is pointed out

8	Hospital @China[45]	Id, country, date, summary, location, country, gender, age, symptom _onset, Hospital_visit_date, International Traveller, Domestic traveller Exposure, start date, Exposure _end, visiting wuhan, death, recovered, symptom and source	COVID-19; linear and logistic regression; DT; KNN; SVM; SVM with grid search	Required High PIH and old age leads to death.
9	PSO USA[46]	Socio-economic, political factors	SARS-CoV-2; COVID-19; SEIR modeling; Italy; stochastic modeling; swarm intelligence	State wise prediction is local best and nation wide prediction is global best when dealt with PSO
10	Hospital, Spain [47]	P-die, superspreader, superspreader rate, P_reinfection, isolation, travel, socialdistance, pandemic duration , strains	metaheuristics; soft computing; deep learning; big data; coronavirus	Coronovirus optimization algorithm
11	Hospital, Newyork[52]	40 features (both clinical, and other health characteristics	13 characteristics only selected	Hospital course of the patients.

Table:1 Data set and its features based on authors handling on covid 19.

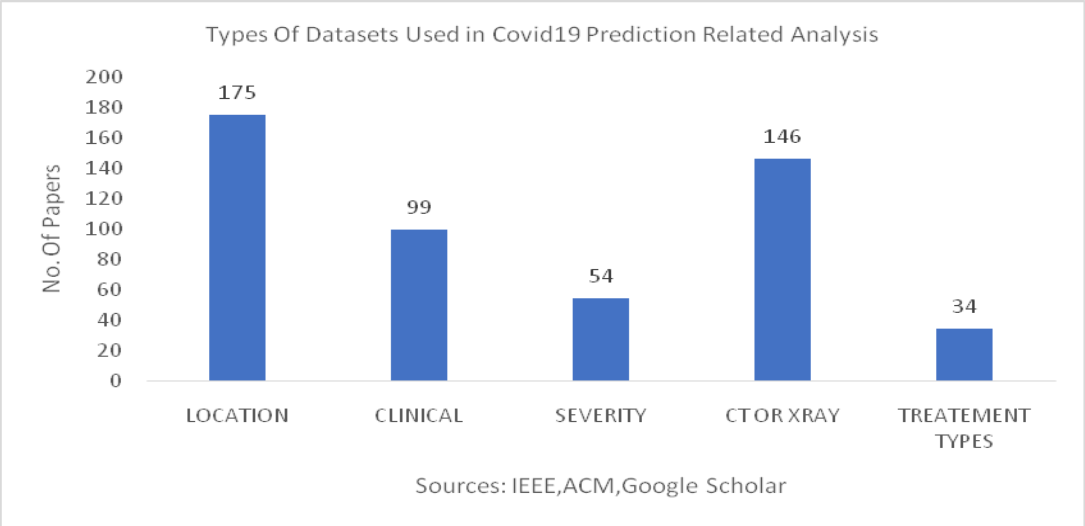

Fig 2

In standard datasets, the dataset related to Location based, clinical based, covid severity based, treatment types and also X ray or CT images related to dataset for Prediction, ICU requirement, Mortality Prediction etc were analyzed for survey analysis. Every dataset and the analysis have different motive to perform. But every motive leads to the prediction of covid and severity. Number of papers found based on above distributed information are expressed in graph (Fig 2)

2.3 Role of Feature Selection on datasets in Covid 19 Analysis

In a dataset, before applying classification / clustering it is very much necessary to do feature selection [23] or feature extraction. Since feature selection is needed for two main purposes like 1. Less number of features 2. High dimension to analyze data.

When there is a smaller number of features then less time for training the algorithm will be taken. If dimensions are high, then scatter plots cannot be seen clearly. In covid 19 analysis, every dataset contains more than 10 number of features. Some times its exceeding 30 to 40. It is possible to do feature selection which can resultant in better output in prediction.

Normally data sets include umpty number of attributes, and many attributes are irrelevant and redundant. Using filter based/embedded based / wrapper-based technology feature selection can be applied. But it requires some intelligent method like swarm intelligence [27] for evaluation functions. The swarm search can provide the best possible length of the feature set. Selection is one which requires more desired step for optimization in every problem. Every feature is to be identified which incorporates or influence more on other features or data fields, so that the output will be triggered and able to get the best solution. Analytics and statistical models and methodologies have applied on data, it can be made it as KPI which is used for interpretation and decision making. All recent years, analytics methods are applied on massive datasets instead of small datasets.

In covid 19 analysis, feature selection is playing the important role. Before applying the model on the [39] dataset, feature selection is applied using random forest algorithm. Among 8 features, 5 important features were selected for classification model to predict death cases. In anomaly detection problem of covid 19[50], the original dataset is filtered by feature extraction, and autoencoding model is applied to offer a solution for the problem. It is training the model using normal data, later it is able to identify the rare event occurring. While doing covid 19 prediction using clinical findings[40], feature extraction can be done at convolution layer in CNN a deep learning model learned automatically. So that dimension of the data is reduced. Random forest classification model is applied on dataset at [41], in which variable selection (feature selection) made as optimal selection which increases the prediction rate and reducing the error rate.

2.4 Different Bio Inspired algorithms which may be suitable for Medical Dataset Analysis

There are various evolutionary and bio-inspired algorithm are available for global level optimization. When the prediction on state-wise is predicted as local search space, it can be applied to nation-wise and made as global search space to get global optimal solution. This kind of prediction model comes under swarm intelligence. Swarm Intelligence was initialized introduced by Gerardo Beni and Jing wanPSg in 1989. Mainly swarm intelligence approaches are playing the important role in optimization.

Its basic features are

1. Nature-inspired
2. Population-based
3. Dynamic
4. Uncertainty
5. Real-time
6. Multiple Neighborhood
7. Memory Usage Methods

Nature-inspired algorithm: A Central concept in Multi-Objective Optimization is Pareto optimality representing a set of non-dominated optimal solutions. Different studies of Computer Science, Mathematics and Biology takes the advantage of Bio-Inspired Computing. It is also an emerging approach apart from traditional approach. Bio-Inspired approach is used for solving complex optimization problem. There is list of various Nature-inspired algorithm like [5] [8]GA, FSA, CSO, WOA, AAA, ESA, CSOA, MFO, GWO.

Artificial Bee colony algorithm [23] is used to apply for both data learning and feature selection. And its performance is evaluated with parameters such as accuracy and complexity. It is used to select related/ linked features. The feature selection is used to remove irrelevant feature so that it is able to reduce the dimension of the model. By reducing space complexity, computation time will be reduced.

PSO [3] is applied on particle and able to find the local best and global best. Global best value is updated from the best among the local best solution satisfied by the fitness function. This process is repeated until the cheapest value is received. Cuckoo Search Algorithm (CSA [13]) is able to solve multi objective problem. For example, in prediction to get minimum cost and lead time, it is necessary to select the resources like multi product and multi delivery. This algorithm improves and produce powerful result for pareto solution. Modified CSO also proposed by adding certain variable in random to improvise the solution at the tracing mode. So that it gives more optimized solution than just CSO/ GA solution.

In Cat Swarm Optimization [10] (CSO), there are nodes of cats like searching and tracing where memory is needed to store the updated best solution among all the cats. It is re-evaluated using fitness function if the termination node is reached it reports the best position, if not again the cats move towards the seeking mode and continued till the required criteria is satisfied. Cat generally search for prey randomly and move towards the search space. The

optimal solution is achieved, as the best position of one of the cats. It is the collaboration of multi swarm algorithm, and it behaves intelligently, to find the best solution of the optimization problem.

The Elephant search swarm optimization[7] represents the search agents to perform the tasks of exploration and exploitation which is needed in Meta Heuristic [19]Algorithm to find the optimal solution. In EHO, the elephants can be updated by updating operator followed by separating operator. A novel nature-inspired algorithm paradigm is called as Moth-Flame Optimization (MFO) algorithm.[11] For very long distance, moths fly in a straight line for a particular angle to find its prey. MFO belongs to different fields, like the power and energy systems, domains of optimization, medical, engineering, economic dispatch, and engineering design, image processing and medical applications clustering and data mining.

Grey Wolf Optimization (GWO)[6] can be applied on Complex Optimization Problem, and used as training algorithm for multilayer perception and also used for solving economic dispatch problems. This optimization procedure has four different levels as Alpha, Beta, Gamma and Omega to locate the particle in an optimized way. When compared to the conventional methods, this Grey Wolf Swarm Optimization[6] is an algorithm, which provides us the optimal solution, for the Real-World Problems. Population-based with single point search algorithm example is Fire Fly Algorithm (FFA)[12],[21]. It is based with single point search in which two features are important that is self-organization and decentralized decision making. The meta heuristic [31],[19]algorithm is to provide a balanced trade-off between local search and global search in order to get better result. The authors suggest that the levy flight Firefly algorithm is more efficient than PSO and GA to get the success rate. Advantages of Firefly algorithm is handling the multi-modal and multi objective problem, fast convergence rate to fix local and global problems.

Chicken swarm optimization[9][25], follows hierarchy orders with rooster, hens, chicks and behaves as a team. Its called multi-swarm. It can bealance the randomness and determinacy while finding the optimized solution. If the parameters are tuned significantly this algorithmic performance will be enhanced, so that it can solve many optimization problems. Whale optimization algorithm[51] executed the algorithm with 3 main steps as encircling prey, spinal bubble net attacking and search for their food. It has wide application on feature selection. Especially in medical datasets, various irrelevant and redundant data features are involved. It can reduce efficiently using this algorithm with acceptable accuracy for medical diagnosis.

Owl search algorithm [49] is suitable for black box optimization with unknown search space. Using various experimental studies that OSA provides optimal solution with higher accuracy and fast convergence rate. This algorithm imitates the hunting mechanism of owls in complete darkness.

Validation should be done using various bench mark functions. The swarm intelligence is applied on clustering [20] [28] for medical dataset and able to prove that some of the algorithms like (PSO,BAT,CUCKOO, FIREFLY) are resulting more valuable output. Especially, PSO AND BAT are executed the concept of clustering with less time complexity. The author utilized the metaheuristic algorithm to handle the complex nonlinear multi-objective [15]

constraints optimization problem. The following table table 2 is a sample which represents how swarm intelligence helped to handle the problems in medical field.

	N-I algorithm	Year	Proposed by	Full form	Feature selection	Classification	Clustering	Datasets	Supported with other method
1.	ABC	2005	Teresh ko and Loenga rov	Artificial Bee colony [11]	Yes	Yes		1. Computed Tomography (CT Sc 2. UCI - Machine Learning Repository 2013.http://archive.ics.uci.edu/ml/ .	With SVM
2.	PSO	1995	Kenne dy and Elberh art	Paritcle swarm	Yes		Yes	Human brain microarray data, Breast cancer data	With ACO
3.	GWO	2014	Mirijal ili	Grey wolf [6]	Yes	Yes		Facial Emotion Recognition[6]	With PSO
4.	ALO	2015	Mirijal ili	Ant Lion	Yes			UCI data sets-Breast cancer	With PSO
5.	EHA	2015	Wang	Elephant herding[7]	Yes			Stock Market Data [7]	With PSO, ANOVA
6	MFO	2016	Mirijal ili	Moth flame[11]	Yes			UCI – e.g. Breast cancer	Better than PSO
6.	WOA	2016	Mirijal ili	Whale optimization [51]	Yes	Yes		UCI Data set	
7.	EOA	2018	Wang	Earthworm optimization	Yes	Yes	Yes	Microarray data set	

8.	MSA	2018	Wang	Moth Swarm	Yes	Yes	Yes	Breast Cancer ,medical and other fields	
9.	HHO	2019	Mirijalili	Harris hawks	Yes	Yes		https://www.openml.org/search?type=data UCI repository dataset	With Simulated Annealing
10.	MBO	2012	Durman	Migrating Birds	Yes	Yes		Microarray data set	

Table 2

In the health care community, the bio-inspired algorithm plays the role in feature selection/feature extraction or prediction. In predictive analytics is playing the important role in decision making. In medical domains like heart disease, cancer, covid prediction, and its accuracy is based on feature selection. Both can be improved when the bio-inspired algorithm is applied than standard feature selection methods like filter, wrapper, and embedded methods. In our study [63] [64], the BAT algorithm is applied to the heart disease and breast cancer dataset. This leads to increase in accuracy of disease prediction to 95.71 with 0.32 RMSE value.

In [65], BEFS model is developed & hybridized with DM & ML algorithm. So that the total number of features is 32 which is reduced to 6 for disease prediction. In another study, for breast cancer and other most dangerous health care diseases Cat Swarm, Kill herd and B F optimization method is hybridized with SVM for accuracy improvisation. It is almost nearer 95% and above. It's a better example for the bio-inspired algorithm. After analyzing various medical disease predictions, when covid prediction is considered in [66] [67]. The authors considered the advantages of various bio-inspired algorithms for feature selection. They are ACO, BH, FF, MFO, ABC. Hybridization of bio-inspired with SVM increases the accuracy and the performance of prediction. Based on erroneous data, the performance may change. If the error is less the performance will increase. If the RMSE (error) value is 0.32, the inference of this value is the difference between observed values and predictive values. As per our study, the evaluation metrics and erratically value represents the power of prediction and are well utilized for any decision making.

2.5 Performance analysis based on confusion matrix

Optimization is the process of finding the solution for the problem which is more cost effective and performance wise reaching the highest result. It can be handled by minimizing the most unwanted impact factors and maximizing the required important factors. Feature selection is one of the concepts which requires more desired step for optimization in every problem. The problems are handled with artificial intelligence techniques, evolutionary algorithm approaches and swarm intelligence algorithms to get optimal solution. Various problems are belonging to

various domains and all problems coming under optimization, which can be with the proper selection of approaches and algorithms under any of the above-mentioned techniques. There is a concept called multi swarm techniques which is used to enhance the performance of population-based algorithm.

Using metaheuristics multi-objective [15] optimization, the basic necessity is to evaluate and rank the algorithm which can be done using fitness strategies and performance evaluation strategies. There are many fitness functions for every algorithm, by satisfying these fitness strategies, it can be concluded that whether that algorithm produces the required optimized result or not.

There are various models available for prediction. The model performance analysis can be calculated on the list of metrics are called evaluation measures. Various metrics [50] in predictive analytics are available as Accuracy, Recall or Sensitivity, Specificity, Precision, F measures.

In this domain, true positive and true negative are identified to calculate the performance for binary classification or multi classification.

The formula for the above listed metrics are

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

$$\text{Recall/sensitivity} = \frac{TP}{(TP+FN)}$$

$$\text{Precision/Specificity} = \frac{TN}{(TN+FP)}$$

$$\text{F Measures} = 2* \frac{(Precision*Recall)}{(Precision+Recall)}$$

The results are evaluated with the above metrics in machine learning, standard classification model for prediction. Every metrics has its own importance. Accuracy holds the value of accurate prediction whether it is positive or negative among total prediction. Recall and Precision are giving importance only to the positive result in prediction. It depends on how the problems are defined and established in the confusion matrix. Such that specificity gives importance to the negative prediction alone. F1 score or F measures is giving the output related to both recall and precision. When accuracy is nearing 100%, that will be helpful to the health care community to handle the problem in a better manner. Already many authors made the result nearer to 100 %. But in medical field its always needed to consider patients detailed clinical findings for prediction. That work was handled in a very few manners. If the detailed attempt made on it then that will really helpful to this community in near feature. The researchers can apply and proceed in future.

2.6 a) On Prediction of covid 19, various authors discussion and their analysis-A related works

In this survey analysis, there are several studies and analyses made for early and accurate prediction for covid 19 are listed in the literature. Every author handled this domain for forecasting or prediction in their way.

In the study [39], the Random Forest prediction model, predicted the multiple classes with feature selection technique. Objective of this author is to predict the number of covid patients, cured and death using the available data in state wise or nation wise. This author applied the classification model of Decision Tree, Random Forest, Logistic regression, Neural network and support vector machine for evaluation. By comparing the evaluation metrics and K fold cross-validation, it is able to confirm that random forest is best model using the performance with more accuracy.

In the study [40], this author predicted the fatality rates of a prediction using the autoencoder model. Here this author used the common classification model like logistic regression, random forest, support vector machine. Logistic regression model is used for categorical variable, support vector machine for maximal distance metrics model between classes. Random Forest is a combination of a number of trees used in the decision tree and some of the trees were selected randomly to predict the output. The author is applied various performance metrics as accuracy, specificity, sensitivity, and AUC. A negative correlation between certain features leads to fatality. The prediction rate of this autoencoder model is 97%. With this autoencoder model, deep learning is applied to rare events for covid prediction or identification. It is more inaccurate than any other model. The fatality rate is based on certain features like demographics, symptoms, and comorbidities.

Using clinical findings, the author [41], had a dataset with 600 patients records in Israel hospital, and by applying deep learning models prediction and evaluation made. The steps are encoder, feature extractor and decoder. The models that are followed by this author are CNN for dimensionality reduction, RNN, and LSTM. Each has its own advantages. Finally, the hybrid model CNNLSTM and its metrics were noted as accuracy 86.6, F1 score 91.89 precision 86.75, Recall 99.42 and AUC 62.5. Around 18 features were selected and considered for prediction. When accuracy increases, the model became reliable and valid and able to support health care system.

According to [42], the descriptive analytics as chi-squared and many tests for categorical field taken for research. In the dataset, there were above 18 features. Among them, only 11 features were selected. Random forest predicts and provides the proper list where certain spotting high risk provide the best result. This author selects the high risk of patient and various resources should be reported with the proper result. AUC value is 0.9905, the 5-fold error rate is all calculated and provides the proper imputation and to increase the prediction and feature selection.

While performing the Empirical analysis, by developing a tree structure to select the right parameters, using fuzzy [35][36][37] [38] set theory weight/ rank is assigned to the factors and selected treatment is able to adapt. The author [43] applied the fuzzification and defuzzification, the suitable course of treatment is identified where he concludes that convalescent plasma is assigned with the highest priority and Yashtimadhu is with least priority. Marginal mean

is used to evaluate and it is more validated by one or more analyses called sensitivity metrics. Basically, in the fuzzy model[37], ranking can be done through Expert's opinion.

The author performs [44], the study and process or clinical findings, and tried to identify which features that are in high increased the risk of mortality. He observed that in multi variate analysis, certain features are very risk factors by applying chi squared / F test on those categorical clinical set and Kruskal Wallis on continuous set. P value is used for findings the risk of patients and the metrics applied like sensitivity, Specificity, Positive and negative predictivity with 95% confidence interval. Major fields like lymphocyte, high RALE etc.

In [45], among 200 records, 172 records with 6 files were taken for observations. Data normally belong to various types in every dataset like continuous, categorical, discrete, constrained, and unconstrained. T-test with mean, the median was applied. Among various features, P-value is used to identify every seriousness of every feature in patients. Mainly clusters are formed according to the features by using prediction, it can do incurred. C- reactive, D-dimer features made the patients take care in a serious manner like ICU. It can be grouped together as clusters and the prediction can be done for survival or death.

The author [46], took the dataset with certain features like Age, PIH, and number of days. When the author studies the problem with prediction using a standard classification model with machine learning. One of the approaches is correlation between the attributes, so that prediction of survival and death rates was held. He concludes that SVM with a grid approach provides more accuracy than other algorithms. Metrics for evaluating the result with the formula describe that precision can refer better result. Prediction based on dataset may provide the best result, when the dataset is large in number. The system gets trained, some of the outputs we can observe are a) correlation between the attributes b) Comparison of algorithm that provide solution.

When data is real-world application and dynamic in nature, it is better to apply the methodology with a metaheuristic algorithm. It is the optimization approach. In Covid prediction, this author [47] handled the problem with the PSO algorithm where exploration and exploitation were held. It is stochastic in nature. Local best and global best are the two different solutions in every approach and in every execution of the adaptive behavior the convergence and stability of the final solution are to be finalized from a certain number of iterations. NRMSE is the error function that is able to reduce the error in the solution. PSO has the advantage of having the characteristics of decentralized, with different policies that provide many better solutions.

This author [48] developed the coronavirus optimization algorithm (CVOA) using the metaheuristic concept for optimal covid prediction according to the input parameters, fitness, and execution time. The terms related to the disease like reinfection, superspreading, social distancing, and traveling rate made the infection in number. By adding vaccination parameter, the infection is reduced and that have taken into the account for optimization of healthy people. The author concluded that deep learning combined with CVOA gave the best training phase for the algorithm and finally it shows the best remarkable performance than single CVOA algorithm. All the terms and features that are defined were considered as parameter/ feature in the dataset. Due to practical parameters as dataset features in CVOA, very limited number of errors were occurred in prediction. The performance is calculated with

the metrics called sensitivity based on number of strains, number of parameters, and number of social distancing measures.

In an objective of this survey, among the list of papers, predicting mortality and critical events with clinical data were handled with logistic regression and decision tree and generalized additive model in [52] and produces the 97% accuracy. From 40 features it's classified and made as 13 features.

In the paper [62], the presence of 7 symptoms to determine the test eligibility, it would maximize the prediction rate in the community. The author applied this method initially where the prediction was based on symptoms. The countries like UK, US symptoms were considered for more accurate prediction.

2.6. b) The Research status of covid 19 using artificial intelligence

With the keywords artificial intelligence, covid and predictive, the authors already approached and handled the problem for prediction in the domain covid 19. In IEEE, around 30 papers only published with these keywords and pros and cons of some of those papers are analyzed here for future work. The research status on prediction in covid .

Choujun Zhan et al [53] proposed in March 2021, a machine learning model with broad learning system. To identify key values, random forest approach applied to forecast the covid 19. The performance of random forest is compared with basic standard classification model like logistic regression, decision tree, k nearest neighbor,etc… and proved that it is better. The RMSE, Radj, R2, MAD, MAPE. In evaluation, the performance of various models with the predicted value can be done and found that RF BAGGING. The dataset referred in this paper contained only epidemic spread data, geographic and economic information data, population etc. The author stated that its required to include weather and air quality data in geographic for more accuracy.

Afshar Shamsi in Feb 2021 [54], proposed a framework for covid 19 diagnosis. The dataset here used and given as input are CT image and Xray image. The performance is calculated using standard metrics and proved that the predictive uncertainty measures are higher in CT data than x ray data. Author stated that a state of the art could be applied on certain uncertainty measures.

Sina Ardabili in Nov 2020[55], proposed a novel model for time series prediction in covid 19 domain. The algorithm utilized for this prediction are artificial neural network with grey wolf optimizer. Results were evaluated with the error method MAPE AS 6.23 in training, 13.15 in testing and 11.4% in validating the result. There are various mathematical models like SIR, SEIRU, SEIR, SLIAR, SIRD, ARIMA & SIDARTHE, but these are very much time consuming, complexity and lower reliability in prediction. And can be say that machine learning is more better than mathematical model.

Nanning Zheng et al in May 2020 [56], proposed a hybrid AI model. Initially, all individuals are same infection rate, secondly using natural language processing and LSTM were embedded with ISI model and proposed this hybrid model, that reduces the error in prediction, and it is calculated by MAPE method to verify. It does not predict the virus using any clinical data.

Hua Ye et al in Jan 2021[57], based on certain patient's details like diseases existence, symptoms, immunity and various other comorbodities, this author proposed a prediction model by applying Harris Hawk s algorithm for better optimization of fuzzy KNN. Especially feature subsets were also formed to identify severe and mild effect of virus in addition to the classification performance on prediction. Author stated that more diagnostic centres and more clinical features must include for more accurate prediction.

Tawsifur Rahman in Aug 2021[58], dealt this covid 19 domain problem to develop and validate the method for early scoring tool to identify the risk of death using blood parameters of the patients. The patients may vary in 3 types as (low, middle, high risks), when CBC is low (<=0.5%),middle (>5% to 55%) and high (>= 50%) mortality risk and also AUC played the role of validation and development cohort. This approach can mainly used for low-resource countries. The author suggested that research on covid 19 clinical data will result better prediction about individual patient's and performance will still be reliable.

Saifynaz Abdel- Fattah sayed et al in sep 2021[59], related to covid 19 prediction in the view of identifying the severity on patients were identified using X-ray. It has denoted multiple number of features; still key features are to be identified. The author applied PCA, RFE and combining these two methods to get better performance by classification method. Where XGboost classifier achieved 97%,98%,95% and 96% in accuracy, presicision, recall and f1 score respectively. At the same time SVM with handcrafted technique of RFE produced 99.99% accuracy. The limitation in this approach is listed as need to consider large datasets, external validation and lacks of interpretability and transparency.

Choujun Zhan et al in April 2021[60], established a general model with historical data in 184 countries. It applied 3 steps where (S-G) filter, PSO-SIR and finally PSO-BLS. The result of this algorithm is compared with deep learning and proved that it produces more accuracy and stability in prediction. According to various countries historical data, in prediction, the evaluation metrics for better performance is measured with RMSE, MAE, R2. Using Croatia PSO-BLS is little less accuracy metrics, when compared with spain, korea and Germany. Where RMSE as 11150.418, 1177.622, 6064.660,7228.758,599.195,3429.111 in MAE, 0.890,0.658,0.908 in R2.

EI Sayed M, EI – kenawy et al in sep 2020[61], approached this covid 19 with 3 different approaches. Using CNN, features are extracted from CT images, SFS- Guided WOA applied on selected features, Guided WOA-PSO had a method as voting classifier and achieved AUC with 0.995. However, the algorithm does not consider dataset which contains clinical findings on swarm intelligence algorithm.

3.0 Limitations

Such that various authors handled the covid 19 prediction problem using artificial (swarm) intelligence algorithm in a different manner according to the dataset availability. Still it has various limitations related to prediction. Already many authors made the result nearer to 100 %. But in medical field its always needed to consider patients detailed

clinical findings for prediction. That work was handled in a very few manners. If the detailed attempt made on it then that will really helpful to this community in near feature. The researchers can apply and proceed in future.

4.0 Conclusion

This review paper mainly focuses on predictive analytics concepts. Swarm intelligence approaches and fuzzy logic approaches are able to handle the problems based on real-life applications. There are many classification models to predict covid 19. According to the survey, Prediction of covid is the need of the hour, especially in the health care community. In fuzzy concepts on covid 19 analysis, doctors' advices are accepted as expert knowledge in fuzzy set theory for ranking purposes of features. At the same time, nature-inspired algorithms are able to predict with a high value of accuracy and are able to achieve the required objective by improving these applications. So according to our objective of this paper, with the dataset of clinical findings, fuzzy set and bio-inspired are able to apply for future prediction with more accuracy. With certain implications it is possible to say that particle swarm optimization is the best algorithm when compared with all other algorithms and also hybridizing this with fuzzy may lead to the best pathway for a decision-making situation.

References

[1] M. Shatnawi, M. B. Yassein,(2019) Big data analytics tools and applications: survey, 19 2019

[2] Satya Shah, Anastasia Theodosoulaki (2019) An Exploratory Study to Examine Big Data Application on Services and SCM , 2018 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD)

[3] Daniel Bratton James Kennedy Defining a Standard for Particle Swarm, (2007) IEEE Swarm Intelligence Symposium

[4] Zhaohao Sun Kenneth Strang Rongping Li 2018, Big Data with Ten Big Characteristics , Proceedings of the 2nd International Conference on Big Data Research

[5] Sharma, Vivek. (2016). A Review of Genetic Algorithm and Mendelian Law. International Journal of Control Theory and Applications

[6] S, Ninu & G, Brammya & R, Ramya & S, Praveena & Binu, D. & B.R., Rajakumar. (2018). Grey Wolf Optimization-based Feature Selection and Classification for Facial Emotion Recognition. IET Biometrics. 7. 10.1049/iet-bmt.2017.0160.

[7] Li, J.; Lei, H.; Alavi, A.H.; Wang, G.-G. Elephant Herding Optimization: Variants, Hybrids, and Applications. Mathematics 2020, 8, 1415. https://doi.org/10.3390/math8091415

[8] Ashraf Darwish (2018),Bio-inspired computing: Algorithms review, deep analysis, and the scope of applications, Future Computing and Informatics Journal

[9] Meng, Xian-Bing & Liu, Yu & Gao, Xiaozhi & Zhang, Hengzhen. (2014). A New Bio-inspired Algorithm: Chicken Swarm Optimization. 86-94. 10.1007/978-3-319-11857-4_10.

[10] David, Tri Widayanti , Muhammad Qadafi Khairuzzahman , Performance Comparison of Cat Swarm Optimization and Genetic Algorithm on Optimizing Functions 2019 1st International Conference on Cybernetics and Intelligent System (ICORIS)

[11] Mirjalili, Seyedali. (2015). Moth-Flame Optimization Algorithm: A Novel Nature-inspired Heuristic Paradigm. Knowledge-Based Systems. 89. 10.1016/j.knosys.2015.07.006.

[12] Łukasik, Szymon & Żak, Sławomir. (2009). Firefly Algorithm for Continuous Constrained Optimization Tasks. Lecture Notes in Computer Science. 5796. 97-106. 10.1007/978-3-642-04441-0_8.

[13] Elkazzaz, F., Mahmoud, A., & Maher, A. (2018). Cuckoo search algorithm for multi-objective supply chain problem. In MATEC Web of Conferences (Vol. 189, p. 06001). EDP Sciences

[14] Iztok Fistera, Iztok Fister Jr.a , Xin-She Yangb , Janez Bresta (2013) A comprehensive review of firefly algorithms, Swarm and Evolutionary Computation

[15] Al-Muhaideb S., El Bachir Menai M. (2013) Hybrid Metaheuristics for Medical Data Classification. In: Talbi EG. (eds) Hybrid Metaheuristics. Studies in Computational intelligence, vol 434. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-30671-6_7

[16] D Kalaivani, T Arunkumar - 2018 Multi process prediction model for customer behaviour analysis International Journal of Web Based Communities.

[17] Shah, S., & Theodosoulaki, A. (2018, November). An exploratory study to examine big data application on services and SCM. IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD) (pp. 155-159).

[18] C. P. Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data," Information Sciences, vol. 275, p. 314--347, 2014.

[19] Almufti, S. M. (2019). Historical survey on metaheuristics algorithms. International Journal Of Scientific World, 7(1), 1

[20] Rafael D Tordecilla 1 2, Angel A Juan 1 3, Jairo R Montoya-Torres 2, Carlos L Quintero-Araujo 4, Javier Panadero 1 3 (2020) Simulation-optimization methods for designing and assessing resilient supply chain networks under uncertainty scenarios: A review

[21]An Enhanced Firefly Algorithm Using Pattern Search for Solving Optimization Problems Fazli Wahid;M. Sultan Zia;Rao Naveed Bin Rais;Muhammad Aamir;Umair Muneer Butt;Mubashir Ali;Adeel Ahmed;Imran Ali Khan;Osman Khalid IEEE Access , 2020

[22] X. Gong, L. Liu, S. Fong, Q. Xu, T. Wen and Z. Liu, “Comparative Research of Swarm Intelligence Clustering Algorithms for Analyzing Medical Data” in IEEE Access, vol. 8, pp. 110251-110251, 2020, doi: 10.1109/ACCESS.2020.3000875.

[23] Feature selection for classification with Artificial Bee Colony Programming By Sibel Arslan and Celal Ozturk , Book Swarm Intelligence – Recent advances New perspectives and applications, 2019.

[24] Luque-Chang, Alberto, et al. "Social spider optimization algorithm: Modifications applications, and perspectives." Mathematical Problems in Engineering 2018 (2018).

[25] An innovative approach for feature selection based on chicken swarm optimization AI Hafez, HM Zawbaa, E Emary, HA Mahmoud, AE Hassanien 2015 7th International Conference of Soft Computing and Pattern Recognition …

[26] Nyoman Gunantara | Qingsong Ai (Reviewing editor) (2018) A review of multi-objective optimization: Methods and its applications, Cogent Engineering, 5:1, DOI: 10.1080/23311916.2018.1502242

[27] From Swarm Intelligence to Swarm Robotics- Beni, Gerardo-2004/07/17-Conference in Swarm Robotics

[28] S. Bejinariu, H. Costin, F. Rotaru and R. Luca, "Data Clustering by Nature-inspired Algorithms and Chaotic Maps," 2019 E-Health and Bioengineering Conference (EHB), Iasi, Romania, 2019, pp. 1-4, doi: 10.1109/EHB47216.2019.8970068.

[29] Revisiting evolutionary algorithms in feature selection and nonfuzzy/fuzzy rule based classification Satchidananda Dehuri Ashish Ghosh ,20 February 2013.

[30] Feature selection, optimization and clustering strategies of text documents.AK Nikhath, K Subrahmanyam International Journal of Electrical & Computer Engineering (2088-8708) 9 (2)

[31]A history of metaheuristics K Sorensen, M Sevaux, F Glover arXiv preprint arXiv:1704.0085

[32] https://www.analyticsvidhya.com/blog/2016/12/introduction-to-feature-selection-methods-with-an-example-or-how-to-select-the-right-variables

[33] Brezočnik L, Fister I Jr., Podgorelec V. Swarm Intelligence Algorithms for Feature Selection: A Review. Applied Sciences. 2018; 8(9):1521. https://doi.org/10.3390/app8091521

[34] Challenges of feature selection for Big data ,special Issue on Big Data, IEEE Intelligent Systems, 2016. arXiv admin note: text overlap with arXiv:1601.07996

[35] A.J. Jetter, K. Kok, Fuzzy Cognitive Maps For Futures Studies - A methodological assessment of concepts and methods, Futures (2014)

[36] E. I. Papageorgiou, N. I. Papandrianos, G. Karagianni, G. C. Kyriazopoulos and D. Sfyras, "A fuzzy cognitive map based tool for prediction of infectious diseases," 2009 IEEE International Conference on Fuzzy Systems, Jeju, Korea (South), 2009, pp. 2094-2099, doi: 10.1109/FUZZY.2009.5277254.

[37] Jetter, Antonie J., and Kasper Kok. "Fuzzy Cognitive Maps for futures studies—A methodological assessment of concepts and methods." Futures 61 (2014): 45-57.

[38] I. Papageorgiou, C. D. Stylios and P. P. Groumpos, "An integrated two-level hierarchical system for decision making in radiation therapy based on fuzzy cognitive maps," in IEEE Transactions on Biomedical Engineering, vol. 50, no. 12, pp. 1326-1339, Dec. 2003, doi: 10.1109/TBME.2003.819845.

[39] V. K. Gupta, A. Gupta, D. Kumar and A. Sardana, "Prediction of COVID-19 confirmed, death, and cured cases in India using random forest model," in Big Data Mining and Analytics, vol. 4, no. 2, pp. 116-123, June 2021, doi: 10.26599/BDMA.2020.9020016.

[40] Alakus TB, Turkoglu I. Comparison of deep learning approaches to predict COVID-19 infection. Chaos Solitons Fractals. 2020 Nov;140:110120. doi: 10.1016/j.chaos.2020.110120. Epub 2020 Jul 11. PMID: 33519109; PMCID: PMC7833512.

[41] Wang J, Yu H, Hua Q, Jing S, Liu Z, Peng X, Cao C, Luo Y. A descriptive study of random forest algorithm for predicting COVID-19 patients outcome. PeerJ. 2020 Sep 9;8:e9945. doi: 10.7717/peerj.9945. PMID: 32974109; PMCID: PMC7486830.

[42] W. Alosaimi, R. Kumar, A. Alharbi, H. Alyami, A. Agrawal et al., "Computational technique for effectiveness of treatments used in curing sars-cov-2," Intelligent Automation & Soft Computing, vol. 28, no.3, pp. 617–628, 2021.

[43] Ciceri, F., Castagna, A., Rovere-Querini, P., De Cobelli, F., Ruggeri, A., Galli, L., ... & Zangrillo, A. (2020). Early predictors of clinical outcomes of COVID-19 outbreak in Milan, Italy. Clinical immunology, 217, 108509.

[44] Yang, Luhuan et al. “Epidemiological and clinical features of 200 hospitalized patients with corona virus disease 2019 outside Wuhan, China: A descriptive study.” Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology vol. 129 (2020): 104475. doi:10.1016/j.jcv.2020.104475.

[45] A. K. Dubey, S. Narang, A. Kumar, S. M. Sasubilli and V. García-Díaz, "Performance estimation of machine learning algorithms in the factor analysis of covid-19 dataset," Computers, Materials & Continua, vol. 66, no.2, pp. 1921–1936, 2021.

[46] Godio A, Pace F, Vergnano A. SEIR Modeling of the Italian Epidemic of SARS-CoV-2 Using Computational Swarm Intelligence. International Journal of Environmental Research and Public Health. 2020; 17(10):3535. https://doi.org/10.3390/ijerph17103535.

[47] Martínez-Álvarez F, Asencio-Cortés G, Torres JF, Gutiérrez-Avilés D, Melgar-García L, Pérez-Chacón R, Rubio-Escudero C, Riquelme JC, Troncoso A. Coronavirus Optimization Algorithm: A Bioinspired Metaheuristic Based on the COVID-19 Propagation Model. Big Data. 2020 Aug;8(4):308-322. doi: 10.1089/big.2020.0051. Epub 2020 Jul 22. PMID: 32716641.

[48] Learning from past respiratory failure patients to triage COVID-19 patient ventilator needs: A multi-institutional study. Carmichael H, Coquet J, Sun R, Sang S, Groat D, Asch SM, Bledsoe J, Peltan ID, Jacobs JR, Hernandez-Boussard T.J Biomed Inform. 2021 Jul;119:103802. doi: 10.1016/j.jbi.2021.103802. Epub 2021 May 27.PMID: 33965640

[49] Owl search algorithm: A novel nature-inspired heuristic paradigm for global optimization Jain, Mohit , Maurya, Shubham , Rani, Asha , Singh, Vijander , Thampi, Sabu M. , El-Alfy, El-Sayed M. , Mitra, Sushmita , Trajkovic, Ljiljana Volume:34 ,Journal of Intelligent & Fuzzy Systems,DOI:10.3233 / JIFS-169452,March, 2018

[50] Individual-Level Fatality Prediction of COVID-19 Patients Using AI Methods, Li, Yun and Horowitz, Melanie Alfonzo and Liu, Jiakang and Chew, Aaron and Lan, Hai and Liu, Qian and Sha, Dexuan and Yang, Chaowei, VOLUME 8, Frontiers in Public Health, , 30 September 2020 | https://doi.org/10.3389/fpubh.2020.587937, ISSN 2296-2565

[51] Zamani, Hoda & Nadimi-Shahraki, Mohammad H.. (2016). Feature Selection Based on Whale Optimization Algorithm for Diseases Diagnosis. International Journal of Computer Science and Information Security,. 14. 1243-1247. 10.13140/RG.2.2.29065.88161.

[52] Vaid A, Somani S, Russak A, De Freitas J, Chaudhry F, Paranjpe I, Johnson K, Lee S, Miotto R, Richter F, Zhao S, Beckmann N, Naik N, Kia A, Timsina P, Lala A, Paranjpe M, Golden E, Danieletto M, Singh M, Meyer D, O'Reilly P, Huckins L, Kovatch P, Finkelstein J, Freeman R, Argulian E, Kasarskis A, Percha B, Aberg J, Bagiella E, Horowitz C, Murphy B, Nestler E, Schadt E, Cho J, Cordon-Cardo C, Fuster V, Charney D, Reich D, Bottinger E, Levin M, Narula J, Fayad Z, Just A, Charney A, Nadkarni G, Glicksberg B,Machine Learning to Predict Mortality and Critical Events in a Cohort of Patients With COVID-19 in New York City: Model Development and Validation,J Med Internet Res 2020;22(11):e24018, https://www.jmir.org/2020/11/e24018,DOI: 10.2196/24018

[53] C. Zhan, Y. Zheng, H. Zhang and Q. Wen, "Random-Forest-Bagging Broad Learning System With Applications for COVID-19 Pandemic," in IEEE Internet of Things Journal, vol. 8, no. 21, pp. 15906-15918, 1 Nov.1, 2021, doi: 10.1109/JIOT.2021.3066575.

[54] A. Shamsi et al., "An Uncertainty-Aware Transfer Learning-Based Framework for COVID-19 Diagnosis," in IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 4, pp. 1408-1417, April 2021, doi: 10.1109/TNNLS.2021.3054306.

[55] S. Ardabili, A. Mosavi, S. S. Band and A. R. Varkonyi-Koczy, "Coronavirus Disease (COVID-19) Global Prediction Using Hybrid Artificial Intelligence Method of ANN Trained with Grey Wolf Optimizer," 2020 IEEE 3rd International Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE), 2020, pp. 000251-000254, doi: 10.1109/CANDO-EPE51100.2020.9337757.

[56] N. Zheng et al., "Predicting COVID-19 in China Using Hybrid AI Model," in IEEE Transactions on Cybernetics, vol. 50, no. 7, pp. 2891-2904, July 2020, doi: 10.1109/TCYB.2020.2990162.

[57] H. Ye et al., "Diagnosing Coronavirus Disease 2019 (COVID-19): Efficient Harris Hawks-Inspired Fuzzy K-Nearest Neighbor Prediction Methods," in IEEE Access, vol. 9, pp. 17787-17802, 2021, doi: 10.1109/ACCESS.2021.3052835.

[58] T. Rahman et al., "Development and Validation of an Early Scoring System for Prediction of Disease Severity in COVID-19 Using Complete Blood Count Parameters," in IEEE Access, vol. 9, pp. 120422-120441, 2021, doi: 10.1109/ACCESS.2021.3105321.

[59] S. A. -F. Sayed, A. M. Elkorany and S. Sayed Mohammad, "Applying Different Machine Learning Techniques for Prediction of COVID-19 Severity," in IEEE Access, vol. 9, pp. 135697-135707, 2021, doi: 10.1109/ACCESS.2021.3116067.

[60] C. Zhan, Z. Wu, Q. Wen, Y. Gao and H. Zhang, "Optimizing Broad Learning System Hyper-parameters through Particle Swarm Optimization for Predicting COVID-19 in 184 Countries," 2020 IEEE International Conference on E-health Networking, Application & Services (HEALTHCOM), 2021, pp. 1-6, doi: 10.1109/HEALTHCOM49281.2021.9399020.

[61] C. Zhan, Z. Wu, Q. Wen, Y. Gao and H. Zhang, "Optimizing Broad Learning System Hyper-parameters through Particle Swarm Optimization for Predicting COVID-19 in 184 Countries," 2020 IEEE International Conference on E-health Networking, Application & Services (HEALTHCOM), 2021, pp. 1-6, doi: 10.1109/HEALTHCOM49281.2021.9399020.

[62] Elliott J, Whitaker M, Bodinier B, Eales O, Riley S, Ward H, Cooke G, Darzi A, Chadeau-Hyam M, Elliott P. Predictive symptoms for COVID-19 in the community: REACT-1 study of over 1 million people. PLoS Med. 2021 Sep 28;18(9):e1003777. doi: 10.1371/journal.pmed.1003777. PMID: 34582457; PMCID: PMC8478234.

[63] Doreswamy and M. U. Salma, "BAT-ELM: A bio inspired model for prediction of breast cancer data," 2015 International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), 2015, pp. 501-506, doi: 10.1109/ICATCCT.2015.7456936.

[64] Dr. G N Beena Bethel, N. Sree Sandhya. (2020). Heart Disease Prediction Using Bio Inspired Algorithms – Review. International Journal of Advanced Science and Technology, 29(06), 6331 - 6336. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/19919

[65] S. J. Pasha and E. S. Mohamed, "Bio inspired Ensemble Feature Selection (BEFS) Model with Machine Learning and Data Mining Algorithms for Disease Risk Prediction," 2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA), 2019, pp. 1-6, doi: 10.1109/ICCUBEA47591.2019.9129304.

[66] International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 10 (2018) pp. 7895-7902 © Research India Publications. http://www.ripublication.com 7895 Hybrid Bio-inspired Approach for Feature Subset Selection Badra Khellat-kihel, Mohamed Benyettou Laboratory of Modeling and Optimization of Industrial Systems (LAMOSI), Department of Informatics, Faculty of Mathematic and Informatics, Université des Sciences et de la Technologie Mohamed-Boudiaf, B.P. 1505 EL M'NAOUER 31000 ORAN – Algeria.

[67] Menad, Hanane & Amine, Abdelmalek. (2018). Bio-Inspired Algorithms for Medical Data