

Prediction of the Risk of Being Infected by Covid-19 With Particle Swam Optimization

**Project Thesis**

**Submitted By**

|  |  |
| --- | --- |
| 19-40874-2 | Shiraji, Yeamin Rahman Shuvo |
| 17-33992-1 | Nurul Afsar |
| 17-34884-2 | Shovon, Shazzad Hossain |
| 19-40789-2 | Nafisa Binte Shahadat |

Department of Computer Science

Faculty of Science & IT

American International University-Bangladesh

January, 2023

|  |
| --- |
| Declaration |

We proclaim that this proposal is our unique work and has not been submitted in that frame of mind for one more degree or recognition at any college or other establishment of tertiary training. Data got from the distributed and unpublished work of others has been recognised in the text, and a rundown of references is given.

|  |  |
| --- | --- |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  **Nafisa Binte Shahadat**  19-40789-2  CSE | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  **Nurul Afsar**  17-33992-1  CSE |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  **Shovon, Shazzad Hossain**  17-34884-2  CSE | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  **Shiraji, Yeamin Rahman Shuvo**  19-40874-2  CSE |

|  |
| --- |
| Approval |

The thesis titled “Prediction of the Risk of Being Infected by Covid-19 With Particle Swam Optimization” has been submitted to the following respected members of the board of examiners of the department of computer science in partial fulfilment of the requirements for the degree of Bachelor of Science in Computer Science on 19th January, 2023 and has been accepted as satisfactory.

|  |  |  |
| --- | --- | --- |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  **Dr. Md. Sakir Hossain**  Associate Professor & Supervisor  Department of Computer Science  American International University-Bangladesh | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  **Md. Anwarul Kabir**  Associate Professor & External  Department of Computer Science  American International University- Bangladesh | |
|  |  | |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  **Dr. Md. Abdullah-Al-Jubair**  Assistant Professor and Head in-charge (UG)  Department of Computer Science  American International University-Bangladesh | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  **Professor Dr. Dip Nandi**  Director  Faculty of Science & Technology  American International University-Bangladesh | |
|  |  | |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  **Mashiour Rahman**  Sr. Associate Professor & Associate Dean  Faculty of Science & Technology  American International University-Bangladesh | | |
| Acknowledgement | |

This Task is our affable exertion and our manager's imagination and consistent motivation. We, first of all, might want to be appreciative to the All-powerful Allah, who conceded us the work to chip away at this task for the last two semesters. Extraordinary gratitude goes to our honourable supervisor Dr. Md. Sakir Hossain, Collaborator Teacher, American International University-Bangladesh [AIUB] for his broad help, inspiration, and creative reproach. His great oversight and persistent push made this task conceivable. We are submissively appreciative to him for offering us the chance to work with him. We also thank our honorable Vice Chancellor, Dr. Carmen Z. Lamagna, Prof. Dr. Tafazzal Hossain, our Associate Dean, Mashiour Rahman, our Director, Dr. Dip Nandi, and our Head-In-Charge of the Department, Dr. Md. Abdullah-Al-Jubair for their constant inspiration and support.

|  |
| --- |
| **Table of Contents** |

[Abstract 1](#_Toc125625831)

[Chapter 1: Introduction 2](#_Toc125625832)

[1.1 Related Work 3](#_Toc125625833)

[1.2 Research Objective 8](#_Toc125625834)

[1.3 Research Contribution 8](#_Toc125625835)

[1.4 Organization of the Thesis 8](#_Toc125625836)

[Chapter 2: Particle Swarm Optimization 9](#_Toc125625837)

[2.1 Metaheuristic Algorithm 9](#_Toc125625838)

[2.2 Particle Swarm Optimization 10](#_Toc125625839)

[2.3 PSO Algorithm 11](#_Toc125625840)

[Chapter 3: Experimental Results 13](#_Toc125625841)

[3.1 Simulation Parameters 13](#_Toc125625842)

[3.2 Objective Function Value 13](#_Toc125625843)

[3.3 Accuracy of the Covid Prediction 15](#_Toc125625844)

[3.4 Number of features selected with iterations 15](#_Toc125625845)

[3.5 List of the selected features 16](#_Toc125625846)

[Chapter 4: Conclusion 17](#_Toc125625847)

[References 18](#_Toc125625848)

|  |  |  |
| --- | --- | --- |
| **List of Tables** | | |
| **Table No.** | **Captions** | **Page No.** |
| **Table 3.1** | Simulation Parameters | 13 |
|  |  |  |
| **List of Figures** | | |
| Figure 2.1 | Metaheuristic classification | 10 |
| Figure 2.2 | Various types of POS | 11 |
| Figure 3.2 | Objective Function Value vs Iteration Graph | 14 |
| Figure 3.3 | Accuracy vs Iteration Graph | 15 |
| Figure 3.4 | No. of Features vs Iteration Graph | 15 |
| Figure 3.5 | Selected Features | 16 |

|  |
| --- |
| Abstract |

The rapid spread of COVID-19 throughout the world is the one of the major concerns of people. All the hospitals in many countries and World Health Organization (WHO) continuously tries to stop the massive spread of this pandemic virus. But till now most effective way to stop this is to wearing masks, isolation, keeping distance and avoided the crowd places. in the middle of the 2022 world has come to a recovery stage and people started living a normal life. But recently corona symptoms are again visible and new variant of corona virus is already discovered. Now this is becoming the alarming stage again. Many machine learning techniques has given a good result of detecting the covid patients. But it will be very much effective for the people to predict who can be infected by coronavirus in future. In this thesis, we propose a method for predicting the risk of being infected with Covid-19. To this end, we propose a feature selection method using particle swarm optimization (PSO) to find the best feature for improving the prediction accuracy with lower number of features. We exploit the extra tree classifier. The proposed method can predict the risk of being infected with Covid-19 with more than 99% accuracy. Furthermore, it reaches convergence by 7-15 iterations only. The required number of features ranges from 10-12. It improves the prediction accuracy by roughly 5%.

|  |
| --- |
| Chapter 1: Introduction |

In the year of 2019 suddenly, a disease spread from human to human through interaction and close contact of people. This disease was first spotted in Wuhan city of China which was deadly and unstoppable spreading massively among human. [23]. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [24] caused this disease to human which was new to human race. It was later identified Covid-19 by world health organization (WHO) which became a global pandemic rapidly. Governments across the world had to restrict human interaction by declaring lockdown so that covid-19 does not spread anymore. It attacks human immunity system first after that human respiratory system get damaged so much which leads the patients to death. Although many methods were used to predict the covid-19 diseased patients but it was time consuming and lacks clinical data, high false negative rate and delaying results as well. In the measurement of the outbreak of covid-19 antibodies has been used by many authors to predict estimation of the infection risk as well.

In this paper we proposed took an algorithmic approach to analyse the risk of covid-19 detection. Our main motivation is to analyse the risk of covid-19 and using optimization for accurate and faster results. Besides, this approach we can estimate the risk analysis of the infection as well. Here many parameters have been used for dataset to use particle swarm optimization (PSO) for the prediction. For this purpose, metaheuristic algorithm has also been used.

The infected rate was much more than expected. The countries around the world had suffered huge covid 19 infection rates. This rate was exceedingly high that the wave spread millions of people around the world. The best methods were slow against this deadly virus. Though China had controlled the covid-19 effectively but the infection rate was much more with a little bit of close contact among human.[25]

In this paper we suggested an approach of using the particle swarm optimization (PSO) to predict the risk analysis and prevent the disease by controlling it effectively.

## 1.1 Related Work

The authors [1] put forth a model that was created to analyse the effects of face mask usage in public settings. It is based on information from three datasets, including the Simulated Masked Faces in the Wild (SMFD), Labelled Faces in the Wild (LFW)., and the Real-World Masked Face Dataset (RMFD). For the purpose of identifying face masks, a hybrid model based on deep and traditional machine learning was demonstrated. The SVM classifier obtained 100% accuracy in LFW testing, 99.49% accuracy in SMFD testing, and 99.64% accuracy in RMFD testing. The main flaw is that the majority of traditional machine learning techniques do not yield the best accuracy and efficiency. Utilizing higher transfer learning models for feature extraction and the neuromorphic because it has good potential is one of the probable future projects. Barstugan et al. [2] demonstrated the use of abdominal computed tomography (CT) images for COVID-19 early phase identification. To improve the classification performance, patches were subjected to the feature extraction method. As feature extraction techniques, the Local Directional Pattern (LDP), Discrete Wavelet Transform (DWT), Grey Level Run Length Matrix (GLRLM), and Grey Level Co-occurrence Matrix (GLCM) algorithms were applied. With 10-fold cross-validation and the GLSZM feature extraction method, the best classification accuracy was attained at 99.68%. A unique convolutional neural network (CNN) model was also created by Islam et al. [4] that retrieved 100 standout characteristics from a total of 2482 CT scan pictures. Following pre-processing, the collected characteristics were used to identify targets using a number of well-known machine learning methods, as commonly used algorithms include those like Gaussian Naive Bayes, Support Vector Machines, Decision Trees, Logistic Regression, and Random Forests. The suggested model performs better than the earlier models [2], with accuracy, precision, and recall scores of 99.73%, 99.46%, and 100%, respectively. By combining the identifications from each individual ML model into an ensemble learning model, they improved the performance of the models. Compared to the conventional RT-PCR (Reverse Transcription Polymerase Chain Reaction) testing approach, the detection of COVID-19 with this model is quicker and more precise. A deep feature plus support vector machine (SVM)-based approach is suggested by the researcher [3] for the detection of coronavirus-infected individuals utilizing X-ray pictures. The 13 different CNN models' deep features are used to evaluate SVM for COVID-19 identification since it must train and verify on a big dataset. Sethy et al. [3] used pre-trained CNN models to get the best results for coronavirus COVID-19 detection. The ResNet50 plus SVM model, with a classification accuracy of 95.33%, had the highest performance. This method's drawback is that it cannot be used if the patient is in a life-threatening condition and unable to go for an X-ray scan. As pandemic was [5] highly infective the authors proposed a new way of detecting covid-19 with machine learning. As Reverse Transcription [Polymerase Chain Reaction](https://www.sciencedirect.com/topics/engineering/polymerase-chain-reaction) (RT-PCR) test has a high false negative rate and delays during the result so this study was approached using machine learning and image processing for more accurate detection. The motive of this study is to diagnose treat the covid patients and end the coronavirus pandemic as soon as possible. A 10fold cross validation was used for the implemented model. The results were with 3datasets considering 89.41% for dataset-1(CT), 98.11% for dataset-2 (CT), 99.02% for dataset-3 for (X-ray) respectively. For pneumonia, covid positive and negative patients the accuracy was around 85.96%. However, the study had its own limitation and one of them was it lacks clinical data on patients. The goal of this model was to create a model comparatively by using deep learning methods. Cabitza, Federico, et al [6] presents that rRT-PCR has been time consuming and has its shortcomings with false negative rates. Thus, development of machine learning could make faster and lesser expensive blood tests. To develop the machine learning model to complete the OSR dataset three different datasets were taken from 1624 patients. For internal-external validation CBC, Hemogas analysis values, biochemical and covid specified dataset were taken to complete the method. The developed model for OSR dataset were ranged for algorithm curve from 0.83 to 0.90 using KNN for covid specified dataset. On the other hand, CBC using RF had ranged from 0.83 to 0.87. The validation had certain good results for AUC 0.75 to 0.78 and specificity from 0.92 to 0.96. As the proposed model had faster response and cheaper but it was less accurate than the gold standard RT-PCR. It aims to facilitate the strategy of testing by faster way and has the potential to slow down the spreading of virus. This proposes method [7] represented by Brunese, Luca et al aimed to detect covid-19 by evaluating the images taken from medical dataset automatically. The proposed model is considered as a strategy for collecting datasets 85 chest x-rays using them for machine learning purposes. So, the communication system (PACS) chest X-rays were obtained for supervised classification which need to be labeled for the proposed model. The expertise for domain had to be assigned for the training dataset that is chest X-ray labels. Colour layout descriptor (CLD) features is considered to obtain numerical values to capture spatial image of colour images. The classification results show with an average precision recall of 0.965 which were acquired. A supervised machine learning method was implemented to detect Covid-19, resulting in a precision of 0.968 and a recall of 0.964. The model aims to validate the proposed method with a healthy chest X-rays using deep learning and verification technique. The author proposed [8] a passive detection for wearable sensors which is used based on decision tree for the covid1-19 infection. The adaptation of absence and symptoms are visible to the sensor data which explains the feature and post behaviour of the patients Gradient boosting prediction model based on decision tree were developed using aforementioned features. The model was trained and tested in four conditions based on symptomatic and no symptom cohort of data. A cohort of symptomatic patients were tested ranging by the AUC average of 0.83[0.81-0.85]. Analyzing the intrinsic variability of the model the approach gained 95% of confidence intervals for the results but the confidence interval accuracy was not close to 100%. This proposed model is more comprehensive of the feature importance and can further aggregate with the same category. [15] An automated machine learning-based algorithm is proposed for COVID-19 detection and grading of nine different dataset as reverse transcription-polymerase chain reaction (RT-PCR) test has been detect with a huge false negative result. This paper has proposed of image processing and machine learning to detect COVID using CXR and CT medical imaging. So, firstly images are pre-processed and remove the noises. Secondly, segmentation of images is done by fuzzy c-means clustering. At last, KNN, SRC, ANN, SVM are used whether it’s normal, pneumonia or COVID-19 patients. Performance has been measured by k (5) fold cross validation. Among KNN, SRC, ANN, SVM classifiers, SVM has showed more efficient result. The accuracy of SVM classifier is 99.14%. [16] most of the cases, machine learning method has been used for Covid detection. Such as a test which has been done with DNA sequences and identification, also CT scans and X-ray images are based on machine learning. Most recent case is COVID patients’ blood test using machine learning. In this paper, which dataset is used that is collected from Kaggle. Only CBC parameters are used for records. After collecting the chosen attributes pre-processing activities are done before classification. Then the classification is done. After that result and performances are showed. For these, classification algorithms are analysed as basis of precision, recall, a measure accuracy from confusion matrix of the dataset. [17] An analysis of the frequently employed feature extraction methods that utilize deep learning for COVID-19 classification. To obtain the accurate result some features, which are essential components of machine learning those are chosen among of deep convolution neural networks. This approach is used without data pre-processing methods so that it can be more generalized. This method is validated on publicly available COVID-19 dataset of chest X-ray and CT images. The denseNet121 feature extractor with bagging tree classifier has given best accuracy which is 99%. Second best accuracy has been given by ResNet50 feature extraction which is trained by LightGBM and accuracy is 98%. [18] In this research, a method was proposed for identifying COVID-19 cases solely based on self-reported symptoms. This approach is highly valuable as it is cost-effective and can be easily implemented at both an individual and population level. The top model achieved a sensitivity rate of 0.752, a specificity rate of 0.609, and an AUC for the ROC of -0.728. These are encouraging outcomes that support further investigations towards the development of a machine learning-based test for detecting COVID-19. [19] The ill-disposed effect of the Coronavirus pandemic has made a well-being emergency universally everywhere. The pandemic is affecting everybody truly, intellectually, and monetarily; hence, it is fundamental to examine and comprehend profound reactions during emotional well-being issues. The author fosters a brain network model and trains it utilizing physically named information to consequently identify different feelings at fine-grained marks in the Coronavirus tweets. The author made a custom Q&A RoBERTa model to separate expressions from tweets that are essentially liable for comparing feelings. The author's classification model outflanks different frameworks and accomplishes a Jaccard score of 0.6475 with a precision of 0. 8951.The custom RoBERTa back-and-forth discussion model beats different models by accomplishing a Jaccard score of 0.7865. Further, they present a verifiable feeling examination utilizing Coronavirus tweets over the USA including each state-level investigation [20]. The Coronavirus pandemic has been spreading expediently all over the planet beginning around 2019. Because of this pandemic, human existence is turning out to be progressively involutes and complex. Many individuals have passed on in light of this infection. The absence of antiviral medications is one reason for the spreading of Coronavirus infection. The principal side effect is ordinary influenza. Consequently, in the current condition, the best safeguard for this sickness is the facial covering, which covers the two areas of the mouth and nose. In this exploration, the author presents a cover identifier that utilizes an AI facial classification framework to decide if an individual is wearing a veil or not, so it could be associated with a CCTV framework to check that the main people wearing veils are permitted in. In this research, the authors have done an order by AlexNet design. It is a kind of convolutional brain organization. Essentially a grouping calculation is the most ideal for the order of RGB pictures. The authors reason that the ML-based geography gives improved results with higher exactness and is more compelling in controlling the Coronavirus pandemic. [21] Epivigila is a Chilean coordinated epidemiological observation framework with in excess of 17,000,000 Chilean patient records, making it a fundamental and extraordinary wellspring of data for the quantitative and subjective examination of the Coronavirus pandemic in Chile. By and by, given the broad volume of information constrained by Epivigila, it is challenging for wellbeing experts to group tremendous volumes of information to figure out which side effects and comorbidities relate to tainted patients. This paper intends to look at AI strategies, (for example, support-vector machine, choice tree, and irregular timberland methods) to decide if a patient has Coronavirus or not in view of the side effects and comorbidities revealed by Epivigila. From the gathering of patients with Coronavirus, we chose an example of 10% affirmed patients to execute and assess the methods. We utilized accuracy, review, exactness, F1-score, and AUC to think about the methods. The outcomes propose that the help vector machine performs better compared to choose tree and arbitrary backwoods in regards to the review, precision, F1-score, and AUC. AI methods help process and characterize enormous volumes of information even more proficiently and, accelerate medical service independent direction.

The principal commitment of this paper is that it is a scientific review contrasting AI methods on a dataset to recognize regardless of whether a patient has Coronavirus with information gotten continuously and produced from clinical cycles driven by clinicians in Chile. [26] Author of this paper has proposed a machine learning based approach of daily lifestyle to predict the percentage of infecting and prevent the disease. An online based survey dataset is used in this paper, which has 620 responses. Three machine learning based algorithm is used in this paper. These are: XGBoost Classifier, Logistic Regression, Neural Network Keras. Accuracy accordingly of these classifiers are: 92.6%, 85.5%, 94.2%.

## 1.2 Research Objective

In this paper, we experiment the PSO algorithm of covid prediction and show the accuracy percentage using this mathematical approach. Before infected by covid, the chances of a person of being infected by covid is more helpful to be careful and medical earlier. So, our objective is to predict the chances of being affected by covid in near future. If the prediction accuracy is high then this approach of PSO will be successful.

## 1.3 Research Contribution

* We present a thorough literature reviews about the progress of research of the detection of Covid-19 with machine learning
* We propose a feature selection algorithm using PSO algorithm to improve the detection accuracy and reduce computational complexity.
* Finally, we perform a detail analysis of the performance of the proposed method. We also compare the proposed method with the existing method.

## 1.4 Organization of the Thesis

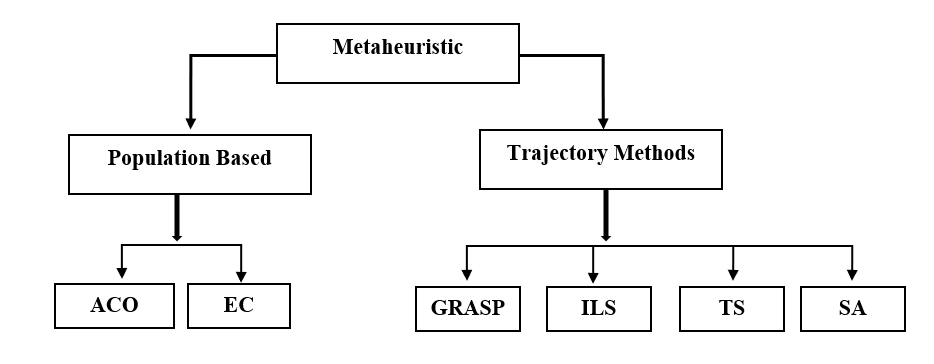
* In Abstract the whole summary of the paper is described.
* Introduction part, there is a broad description about pandemic situation and brief summary of methods that are used in this paper. And literature review is in chapter 1
* Proper algorithm of PSO and Metaheuristic are described in chapter 2
* Experimental results are shown in chapter 3
* Conclusion of the paper is in chapter 4
* At last References are mentioned.

|  |
| --- |
| Chapter 2: Particle Swarm Optimization |

## 2.1 Metaheuristic Algorithm

Meta, which stands for metaheuristic algorithm, is a synonym for "high level." The performance is typically superior to straightforward heuristics. Randomization frequently results in a substantial quantity of solution. Though metaheuristics is so much popular there is no specific definition for metaheuristics or heuristics. However, the recent trend intends naming all the stochastic algorithm with randomization. Metaheuristic algorithms are very satisfactory for non-liner modelling and optimizing. It is an efficient way to use trial and error for complex problems in a reasonable time. It is an algorithm based on computational intelligence paradigms for highly developed problem-solving optimization. Any metaheuristic algorithm has some major components those are intensification and diversification or exploitation and exploration. Intensification refers to narrowing the search space down to a particular area with the knowledge that a suitable solution has already been identified there, while diversification refers to the generation of varied solutions to explore the search space globally. To increase the pace of algorithm convergence, a good balance between intensification and diversification should be established while choosing the best solutions. Global optimality is often assured by a successful combination of these two key elements. In numerous ways we can classify metaheuristic algorithm. This classification is either population based or trajectory based. Population based metaheuristic has ant colony optimization (ACO) and evolutionary computing (EC). On the other hand, trajectory methods have greedy randomization adaptive search procedures (GRASP), iterated local search (ILS), tabu search (TS) and simulated annealing (SA) [9]. The characteristics of metaheuristic algorithm is that it has the convergence analysis with few algorithms. It has a remarkable performance of imitating the best feature of nature [10]. One of the most widely used metaheuristic optimization methods is particle swarm optimization.

A figure is shown below for the metaheuristic classification

Figure 2.1: Metaheuristic classification [9].

## 2.2 Particle Swarm Optimization

Particle swarm optimization (PSO), a population-based stochastic optimization method, Swarm intelligence is influenced by certain species' intelligent collective behaviour, such as flocks of birds or schools of fish [11]. The most common metaheuristic optimization algorithm is particle swarm optimization. PSO is a technique for finding the best selection that utilizes a collective of the population to iteratively carry out a certain function in order to identify the best choice in a search space. It was first proposed in 1995 by Kennedy and Eberhart [13]. As scientists gained knowledge of the method, they created additional iterations aimed at meeting diverse needs, created new applications in a variety of fields, published theoretical analyses of the impacts of the different parameters, and proposed several iterations of the algorithm. In order to resolve nonlinear, nonconvex, or combinatorial optimization issues that emerge across many science and engineering specialties [14]. A particle system generates a collection of moving points at random. Each particle is assigned a vector of initial velocity. It could also have other characteristics including colour, texture, and a brief lifetime. Velocity vectors are incrementally changed by a random factor. It consists of a set of spatial points and rules that control how they move and look, including rotation, size, colour, and transparency [12]. PSO demonstrated its effectiveness as an optimization technique by thoroughly exploring a high-dimensional problem space. It is a reliable stochastic optimization approach that is based on how swarms move and function. The PSO computational technique starts with a population of potential solutions, referred to as a "swarm of particles" in this context, and continually attempts to optimize a problem. Each particle is aware of the swarm's overall best position as well as its individual best position (as defined by its fitness value) thus far. Individual and collective knowledge have an impact on each iteration's velocity and position of each particle in the swarm, which are represented by d-dimensional vectors. The iteration leads repeated particle flights over the space of feasible answers to the problem in pursuit of the optimal one until a suitable stopping condition is satisfied.

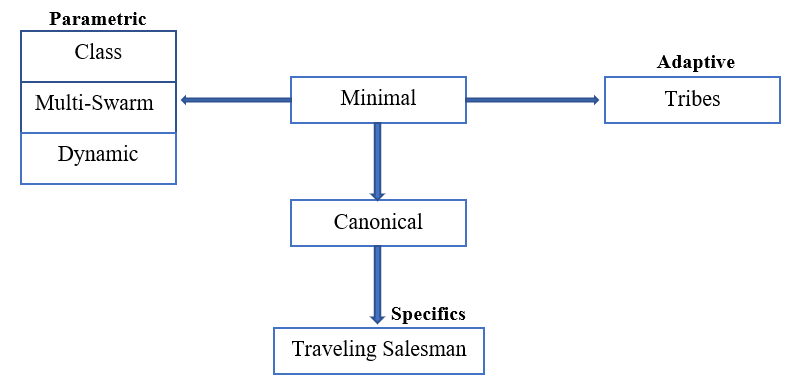


Figure 2.2: Various types of POS [22]

## 2.3 PSO Algorithm

Tracking down a worldwide least Issue (GO) presents extremely hard difficulties. All things being equal, we infer a surmised arrangement by utilizing a Particle Swarm Optimization algorithm (PSO). Since its presentation around twenty years prior by Kennedy and Eberhart50, this approach has been widely utilized for taking care of hard consistent worldwide improvement issues (See51-53 and the references in that). PSO calculation was applied in a few examination studies to manage Coronavirus expectations. In54, a powerless uncovered contaminated isolated recuperated dead (SEIQRD) model was introduced to foresee the Coronavirus spread in Italy. The model boundaries are assessed utilizing the PSO calculation. In 55, an original PSO-BLS (molecule swarm streamlining wide learning framework) was created to foresee the elements of Coronavirus. The proposed model shows higher exactness and security versus profound learning models while foreseeing the quantity of dynamic tainted cases and eliminated cases. Four phenomenological pestilence models as well as the SIR model were researched in56 to foresee the combined number of affirmed cases because of Coronavirus in Saudi Arabia and the conceivable end date of the pandemic. the thought depended on a streamlining approach utilizing the PSO calculation.

A fundamental version of the PSO algorithm operates by utilizing a group of candidate solutions, referred to as particles. These particles are navigated within the search space in accordance with a set of basic equations. The developments of the particles are directed by their own most popular situation in the pursuit space as well as the whole multitude's most popular position. At the point when advanced positions are found these will then, at that point, come to direct the developments of the multitude. The interaction is rehashed and thusly it is trusted, yet not ensured, that a palatable arrangement will ultimately be found. Officially, let f: ℝn → ℝ be the expense capability that should be limited. The capability takes a competitor arrangement as contention as a vector of genuine numbers and creates a genuine number as a result which shows the goal capability worth of the given up-and-comer arrangement. The angle of f is not known. The objective is to find an answer a for which f(a) ≤ f(b) for all b in the pursuit space, which would mean an is the worldwide least.

Allow S to be the number of particles in the multitude, each having a position xi ∈ ℝn in the hunt space and a speed vi ∈ ℝn. Allow pi to be the most popular place of molecule I and allow g to be the most popular place of the whole multitude. A fundamental PSO calculation is then, at that point.

|  |
| --- |
| Chapter 3: Experimental Results |

## 3.1 Simulation Parameters

Table 3.1: Simulation Parameters

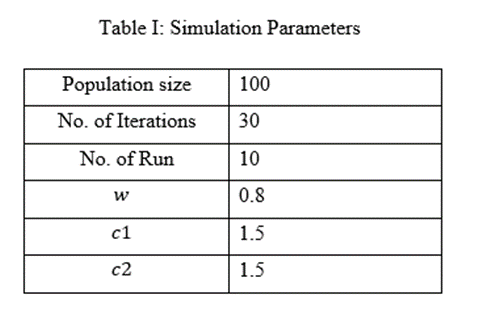


Table 3.1 indicates the hyperparameters of the PSO algorithm. In our experiment, we use a population size of 100, 30 iterations, and run it 10 times. To find a compromise between exploring new possibilities and making use of the best existing solutions, we decided on an inertia weight (w) of 0.8. The coefficients c1 and c2 are determined as an acceleration coefficient. We take 1.5 into account for both of them. The c1 hyperparameter enables determining the group's susceptibility to the best individual solutions discovered during iterations. The capacity of the group to be impacted by the best overall solution discovered during iterations may be defined using the hyperparameter c2.

## 3.2 Objective Function Value

The above figure and line measurement is based on the objective function value and iteration. The real-valued function which value is to be either decreased or increased over the set of practical options is studied as the objective function in a mathematical optimization complexity. A repetition of a series of procedures provides outcomes that are progressively closer to the desired outcome. For instance, repeating a series of computer that instructs a certain number of times until a condition is satisfied. This is also known as recursion. In the above figure there are possible results of measurement of objective function value and iteration. The iteration increases 5 times every time from the starting point 0 to ending point 30.00 and the objective function value increases 0.025 from the starting point 0.800 and ends at 1.00.

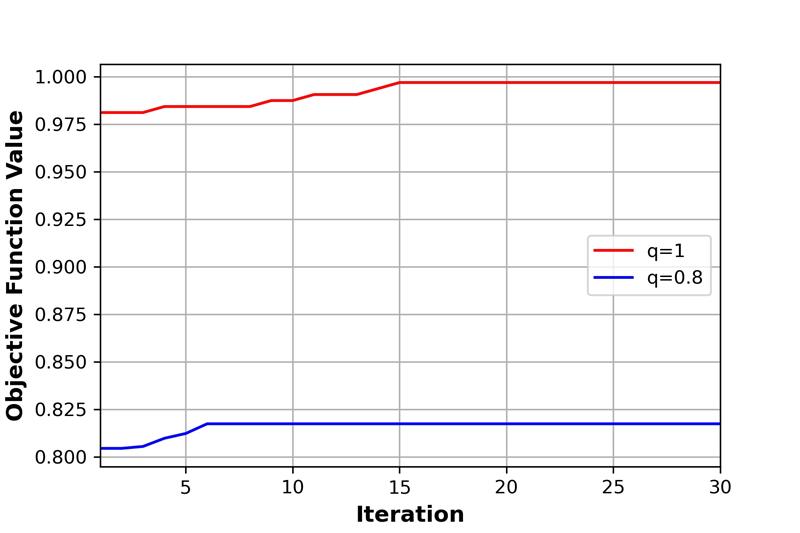


Figure 3.2: Objective Function Value vs Iteration Graph

The Objective function value is vertical while the iteration is measured in the figure is horizontal. Here we can see for the both lines the iteration is the same while the objective function value (q) which is measured from 0.800 to 1.00 is different. Keeping the iteration same, the objective function value for the blue line starts from 8.21 and suddenly increases at iteration 6 touching 0.824. The blue line remains the same for 0.824 at the end of iteration 30.00. But the Red line starts from 0.976 and slowly increases to 0.977 at iteration 4. It remains the same at value until the iteration is 8. From there it slowly increases to 0.999 at iteration 15 and after that the line remains the same objective function value till the end of the iteration 30. In this experimental result we can see that for both lines the objective function value (q) has a significant distance between them.

## 3.3 Accuracy of the Covid Prediction

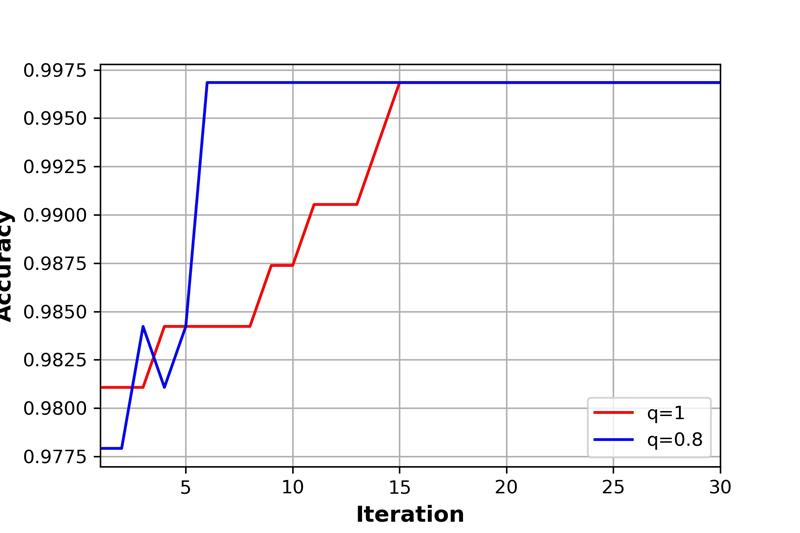
****

Figure 3.3: Accuracy vs Iteration Graph

In this experiment, total 30 times iteration has been done. When the function value is 1, accuracy is high from the first iteration. And in 15 iterations, accuracy has reached in the highest position. Highest accuracy in this experiment is 99.75%. On the other hand, when it comes to functional value 0.8, after 7 iterations, accuracy becomes 99.75% and it continues till the last iteration. So, functional value 0.8 is more consistent than functional value 1.

## 3.4 Number of features selected with iterations

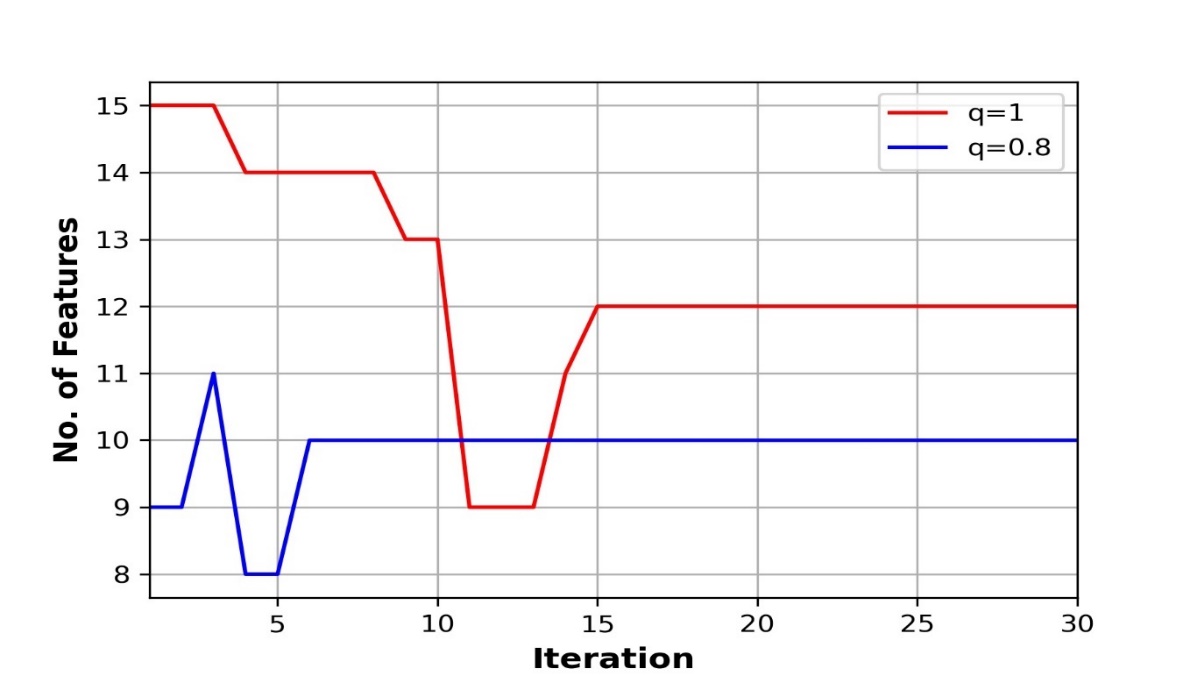
****

Figure 3.4: No. of Features vs Iteration Graph

In the context of machine learning, iteration refers to the frequency of updating the algorithm's parameters. The specific meaning can vary depending on the context. In the diagram, the number of features selected depends on the number of iterations following 15 independent numbers of features with 5, 10, versus 30 iterations, and five each iteration records the selected features. Regardless of the number of iterations, a consistent pattern emerges in regards to the number of selected features; as the number of iterations per partition decreases, the number of selected features also decreases. The number of selected features strictly goes straight line for blue line after 6 times iteration and for red line after 15 times iteration.

## 3.5 List of the selected features

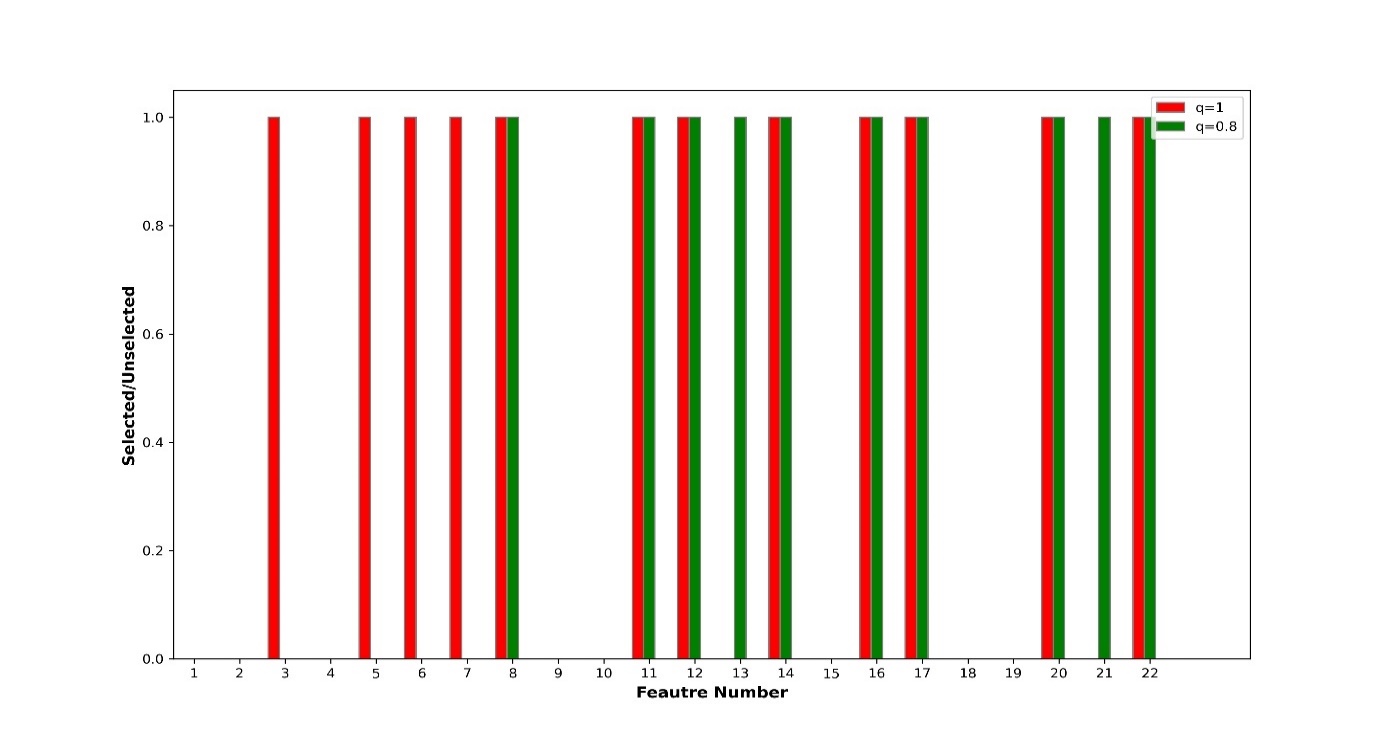
****

Figure 3.5: Selected Features

The following bar chart in Figure 3.5 illustrates the selected feature numbers for two different q values. The red bar denotes the q value of 1, and the green bar denotes the q of 0.8. The majority of the 22 features have been chosen for q=1. For the q of 1, it has been chosen 12 times. Using q of 1, the minimum 3 features are selected, and the maximum 22 features are selected. On the other side, using q of 0.8, the minimum set of 8 features and the maximum set of 22 features are selected.

|  |
| --- |
| Chapter 4: Conclusion |

In this thesis, we proposed a method for predicting the risk of being infected with Covid-19 from analyzing the data related to the life style of people. We exploited machine learning algorithm to perform the prediction. However, instead of using all 25 features of the dataset, we select the best features which improves the prediction accuracy. To select the best features, we particle swarm optimization. The objective function increases with the increase of the accuracy and decrease of the number of features. Through thorough experiments, we show that the proposed method improves prediction accuracy by roughly 5% with less than the 50% of the features.

|  |
| --- |
| References |

[1] Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M. (2021). A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement*, *167*, 108288.

[2] Barstugan, M., Ozkaya, U., & Ozturk, S. (2020). Coronavirus (covid-19) classification using ct images by machine learning methods. *arXiv preprint arXiv:2003.09424*.

[3] Sethy, P. K., & Behera, S. K. (2020). Detection of coronavirus disease (covid-19) based on deep features.

[4] Islam, M. R., & Nahiduzzaman, M. (2022). Complex features extraction with deep learning model for the detection of COVID19 from CT scan images using ensemble-based machine learning approach. *Expert Systems with Applications*, *195*, 116554.

[5] Saygılı, Ahmet. "A new approach for computer-aided detection of coronavirus (COVID-19) from CT and X-ray images using machine learning methods." *Applied Soft Computing* 105 (2021): 107323.

[6] Cabitza, Federico, et al. "Development, evaluation, and validation of machine learning models for COVID-19 detection based on routine blood tests." *Clinical Chemistry and Laboratory Medicine (CCLM)* 59.2 (2021): 421-431.

[7] Brunese, Luca, et al. "Machine learning for coronavirus covid-19 detection from chest x-rays." *Procedia Computer Science* 176 (2020): 2212-2221.

[8] Gadaleta, Matteo, et al. "Passive detection of COVID-19 with wearable sensors and explainable machine learning algorithms." *NPJ Digital Medicine* 4.1 (2021): 1-10.

[9] Abd-Alsabour, N. (2016). Hybrid Metaheuristics for Classification Problems. In (Ed.), Pattern Recognition - Analysis and Applications. IntechOpen.

[10] Gandomi, A. H., Yang, X. S., Talatahari, S., & Alavi, A. H. (2013). Metaheuristic algorithms in modeling and optimization. *Metaheuristic applications in structures and infrastructures*, *1*.

[11] Wang, D., Tan, D., & Liu, L. (2018). Particle swarm optimization algorithm: an overview. Soft computing, 22(2), 387-408.

[12] Du, K. L., & Swamy, M. N. S. (2016). Particle swarm optimization. In Search and optimization by metaheuristics (pp. 153-173). Birkhäuser, Cham.

[13] Eberhart, R., & Kennedy, J. (1995, November). Particle swarm optimization. In *Proceedings of the IEEE international conference on neural networks* (Vol. 4, pp. 1942-1948).

[14] Bansal, J. C. (2019). Particle swarm optimization. In Evolutionary and swarm intelligence algorithms (pp. 11-23). Springer, Cham.

[15] Bhargava, A., Bansal, A., & Goyal, V. (2022). Machine learning-based automatic detection of novel coronavirus (COVID-19) disease. *Multimedia Tools and Applications*, *81*(10), 13731-13750.

[16] Akhtar, A., Akhtar, S., Bakhtawar, B., Kashif, A. A., Aziz, N., & Javeid, M. S. (2021). COVID-19 detection from CBC using machine learning techniques. *International Journal of Technology, Innovation and Management (IJTIM)*, *1*(2), 65-78.

[17] Kassania, S. H., Kassanib, P. H., Wesolowskic, M. J., Schneidera, K. A., & Detersa, R. (2021). Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: a machine learning based approach. *Biocybernetics and Biomedical Engineering*, *41*(3), 867-879.

[18] Martinez-Velazquez, R., Tobón V, D. P., Sanchez, A., El Saddik, A., & Petriu, E. (2021). A machine learning approach as an aid for early covid-19 detection. *Sensors*, *21*(12), 4202.

[19] Md. Yasin Kabir, Sanjay Madria: “EMOCOV: Machine learning for emotion detection, analysis, and visualization using COVID-19 tweets” NSF, USA, [CNS-1461914](https://www.sciencedirect.com/science/article/pii/S2468696421000197#GS1).

[20] Sandeep Gupta, S.V.N. Sreenivasub, Kuldeep Chouhanc, Anurag Shrivastava, Bharti Sahue, Ravindra Manohar Potdar “Novel Face Mask Detection Technique using Machine Learning to control COVID’19 pandemic” Citation: Proceedings, ISSN: 2214-7853

[21] [Pablo Ormeño](https://sciprofiles.com/profile/2208163), [Gastón Márquez](https://sciprofiles.com/profile/1045046), [Camilo Guerrero-Nancuante](https://sciprofiles.com/profile/1438491) , [Carla Taramasco](https://sciprofiles.com/profile/790360) “Detection of COVID-19 Patients Using Machine Learning Techniques: A Nationwide Chilean Study” citation: *int. J. Environ. Res. Public Health* 2022

[22] Clerc, M., 2010. Particle Swarm Optimization. Part. Swarm Optim., pp.1-17.

[23] C.W. Stratton, Y.W. Tang, H. Lu Pathogenesis-directed therapy of 2019 novel coronavirus disease J. Med. Virol., 93 (3) (2021), pp. 1320-1342

[24] Rathore, V., Galhotra, A., Pal, R. & Sahu, K. K. COVID-19 pandemic and children: A review. *J. Pediatr. Pharmacol. Ther.* **25**, 574–585.

[25] Ding, C., Chen, Y., Liu, Z., & Liu, T. (2021). Prediction on transmission trajectory of COVID-19 based on particle swarm algorithm. *Pattern recognition letters*, *152*, 70-78.

[26] Nafiz Fuad Siam, Mahira Tabassum Khan, M. R Rownak, Md. Rejaben Jamin Jueland Ashraf Uddin. Predicting the Risk of COVID-19 Infection Using Lifestyle Data. American International University-Bangladesh, Dhaka 1229, Bangladesh.