PREDICTION OF MONKEYPOX INFECTION FROM CLINICAL SYMPTOMS WITH ADAPTIVE ARTIFICIAL BEE COLONY-BASED DEEP NEURAL NETWORK

2023
MASTER THESIS
COMPUTER ENGINEERING

Ahmed MUHAMMED KALO HAMDAN

Thesis Advisor
Assist. Prof. Dr. Dursun EKMEKCI

PREDICTION OF MONKEYPOX INFECTION FROM CLINICAL SYMPTOMS WITH ADAPTIVE ARTIFICIAL BEE COLONY-BASED DEEP NEURAL NETWORK

Ahmed MUHAMMED KALO HAMDAN

Thesis Advisor
Assist. Prof. Dr. Dursun EKMEKCI

T.C.
Karabuk University
Institute of Graduate Programs
Department of Computer Engineering
Prepared as
Master Thesis

KARABUK
November 2023

I certify that in my opinion the thesis submitted by Ahmed MUHAMMED KALO HAMDAN titled "PREDICTION OF MONKEYPOX INFECTION FROM CLINICAL SYMPTOMS WITH ADAPTIVE ARTIFICIAL BEE COLONY-BASED DEEP NEURAL NETWORK" is fully adequate in scope and in quality as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Dursun EKMEKCI ....

Thesis Advisor, Department of Computer Engineering

This thesis is accepted by the examining committee with a unanimous vote in the Department of Computer Engineering as a Master of Science thesis. November 17, 2023

<u>Examining Committee Members (Institutions)</u>		<u>Signature</u>
Chairman	: Assist. Prof. Dr. Kürşat KARAOĞLAN (KBU)	....
Member	: Assist. Prof. Dr. Dursun EKMEKCİ (KBU)	....
Member	: Assist. Prof. Dr. Veli BAYSAL (BARU)	....

The degree of Master of Science by the thesis submitted is approved by the Administrative Board of the Institute of Graduate Programs, Karabuk University.

Assoc. Prof. Dr. Zeynep ÖZCAN ....

Director of the Institute of Graduate Programs

"I declare that all the information within this thesis has been gathered and presented in accordance with academic regulations and ethical principles and I have according to the requirements of these regulations and principles cited all those which do not originate in this work as well."

Ahmed MUHAMMED KALO HAMDAN

ABSTRACT

Master Thesis

PREDICTION OF MONKEYPOX INFECTION FROM CLINICAL SYMPTOMS WITH ADAPTIVE ARTIFICIAL BEE COLONY-BASED DEEP NEURAL NETWORK

Ahmed MUHAMMED KALO HAMDAN

Karabuk University
Institute of Graduate Programs
Department of Computer Engineering

Thesis Advisor:
Assist. Prof. Dr. Dursun EKMEKCI

November 2023, 57 pages

In 2022, the World Health Organization (WHO) declared an outbreak of monkeypox, a viral zoonotic disease. With time, the number of infections with this disease began to increase in most countries. A human can contract monkeypox by touching with an infected human, or even by touch with animals. In this thesis, diagnostic model for early detection of monkeypox infection based on artificial intelligence methods is proposed. The proposed method is based on training the Artificial Neural Network (ANN) with the Adaptive Artificial Bee Colony (aABC) Algorithm for the classification problem. In the study, the ABC algorithm was preferred instead of classical training algorithms for ANN because of its effectiveness in numerical optimization problem solutions. The ABC algorithm consists of food and limit parameters and three procedures: employed, onlooker and scout bee. In the algorithm standard, artificial onlooker bees are produced as much as the number of artificially employed bees and an equal number of limit values are assigned for all food sources. In the advanced adaptive design, different numbers of artificial onlooker bees are used in each cycle, and the limit numbers are updated. For effective exploitation, onlooker bees tend towards more successful solutions than the average fitness value of the solutions, and limit numbers are updated according to the fitness values of the solutions for efficient exploration. The system was trained and tested on a dataset representing the clinical symptoms of monkeypox infection. The dataset consists of 240 suspected cases, 120 of which are infected and 120 typical cases. The proposed model's results were compared with those of ten other machine-learning models trained on the same dataset. The Deep Learning model achieved the best result with an accuracy of 75%. It was followed by the Random Forest model with an accuracy of 71.1%, while the proposed model came third with an accuracy of 71%.

Key Words : Monkeypox, Monkeypox Clinical Symptoms, Machine Learning, Artificial Neural Network, Artificial Bee Colony Algorithm, Adaptive Artificial Bee Colony Algorithm.

Science Code : 92432

ÖZET

Yüksek Lisans Tezi

ADAPTİF YAPAY ARI KOLONİ TABANLI DERİN SİNİR AĞI İLE KLİNİK BELİRTİLERDEN MAYMUN ÇİÇEĞİ ENFEKSİYONUNUN TAHMİNİ

Ahmed MUHAMMED KALO HAMDAN

Karabük Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı:
Dr. Öğr. Üyesi Dursun EKMEKCİ
Kasım 2023, 57 sayfa

2022'de Dünya Sağlık Örgütü (WHO), viral zoonotik bir hastalık olan maymun çiçeği salgını ilan etti. Zamanla çoğu ülkede bu hastalığa yakalananların sayısı artmaya başladı. Bir insan, enfekte bir insanla doğrudan temas yoluyla veya hatta hayvanlarla temas yoluyla maymun çiçeği hastalığına yakalanabilir. Bu çalışmada, maymun çiçeği enfeksiyonunun erken teşhisi için yapay zekâ yöntemlerine dayalı bir tanı modeli önerilmiştir. Önerilen yöntem, Yapay Sinir Ağının (YSA) sınıflandırma problemi için Adaptif Yapay Arı Kolonisi (aYAK) Algoritması ile eğitilmesine dayanmaktadır. Çalışmada YSA için klasik eğitim algoritmaları yerine sayısal optimizasyon problemlerinin çözümündeki etkinliğinden dolayı YAK algoritması tercih edilmiştir. YAK algoritması yiyecek ve limit parametrelerinden ve üç prosedürden oluşur: işçi, gözcü ve kâşif arı. Algoritma standardında yapay işçi arı sayısı kadar yapay gözcü arı üretilmekte ve tüm besin kaynakları için eşit sayıda sınır değer atanmaktadır. Gelişmiş uyarlanabilir tasarımda, her döngüde farklı sayıda yapay gözcü arı kullanılır ve limit sayıları güncellenir. Etkili kullanım için gözcü arılar, çözümlerin ortalama uygunluk değerinden daha başarılı çözümlere yönelmekte ve verimli keşif için çözümlerin uygunluk değerlerine göre sınır sayıları güncellenmektedir. Sistem, maymun çiçeği enfeksiyonunun klinik semptomlarını temsil eden bir veri seti üzerinde eğitildi ve test edildi. Veri seti, 120'si enfekte ve 120 tipik vaka olmak üzere 240 şüpheli vakadan oluşuyor. Önerilen modelin sonuçları, aynı veri kümesi üzerinde eğitilmiş diğer on makine öğrenimi modelinin sonuçlarıyla karşılaştırıldı. Derin Öğrenme modeli, %75 doğrulukla en iyi sonucu elde etti. Onu %71,1 doğrulukla Random Forest modeli takip ederken, önerilen model %71 doğrulukla üçüncü oldu.

Anahtar Kelimeler : Maymun Çiçeği Hastalığı, Maymun Çiçeği Hastalığının Klinik Belirtileri, Makine Öğrenimi, Yapay Sinir Ağı, Yapay Arı Kolonisi Algoritması, Adaptif Yapay Arı Kolonisi Algoritması.

Bilim Kodu : 92432

ACKNOWLEDGMENT

First of all, I thank Allah and praise Him for this great achievement that He has made easy for me and made difficult times easy for me. I also especially mention my parents. My dear father and mother who supported me at every stage of my life. And to my brothers and sisters who were with me in life, supporting and supporting me. I owe a special thanks to my thesis supervisor, Assist. Prof. Dr. Dursun EKMEKCI.

I would also like to thank everyone who was helpful, including university teachers, colleagues, and academics at Karabuk University. I also extend my special thanks to my English language supervisor, Safaa Saree, for always supporting and encouraging me.

I dedicate this thesis to my beloved country, Syria. And to beautiful Turkey, which embraced this scientific experience and contributed to providing all the possibilities for graduation in this distinguished way.

CONTENTS

Page

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

PART 1

INTRODUCTION

In 2022, the monkeypox virus (MPXV) spread, causing panic among people, and causing concern among scientists due to its rapid spread [1]. Roughly 1 to 11% of cases result in fatality. [2]. The World Health Organization (WHO) Stated that the count of individuals who have been infected with this virus had increased significantly, and it was confirmed that more than 318,000 patients were infected in August 2022 [3]. This virus belongs to the genus of corticoviruses and is similar to zoonotic smallpox [4] It is caused by the orthopoxvirus and is a genera of the poxviridae family that is dangerous to humans [5]. Figure. 1. according to the World Health Organization, shows the number of monkeypox infections in each country (May 26, 2022).

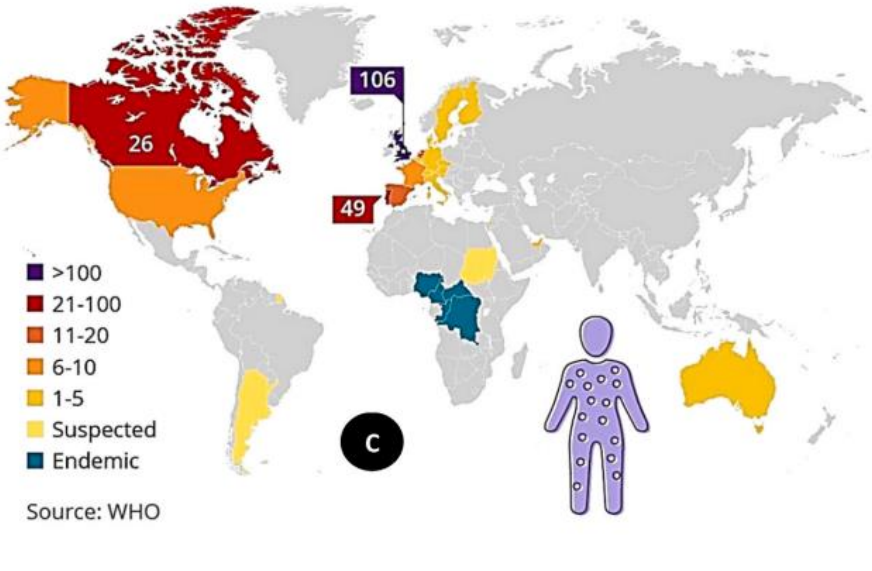


Figure. 1. Number of monkeypox infections in each country (May 26, 2022) [16]

The disease was detected for the first time in Africa, specifically in the Republic of the Congo [4]. And then it spread among the countries of the world. As of June 2022, more than 1,256 cases of monkeypox have been reported in several regions of Spain.

Most of the cases were male. At the same time, the average age was about 36 years [6]. The method of infection with monkeypox is direct contact with an infected person, animal or other material. It is also transmitted through the mucus of the nose, mouth, or eyes [7]. Monkeypox transmission is concentrated, but not exclusively, during sex [5]. Figure. 2. shows infection with monkeypox. which begin to appear after 3 days at most of the infection with the fever. At first, the symptoms are on the face, then they spread to the rest of the body. A person is considered contagious before the rash appears by five days. And it remains until a new layer of skin forms underneath. It takes between two to four weeks [8].

Figure. 2. Symptoms of infection in monkeypox disease that appear on the skin [9]

The clinical picture of monkeypox and smallpox is very similar, and the symptoms that appear upon infection differ from one case to another. However, skin rash is the most likely sign of infection, along with anogenital lesions, lethargy, and muscle pain [7]. Monkeypox symptoms last up to four weeks. Children are also the most vulnerable [10]. Patients with the disease may suffer several side effects, including bronchiolitis, hypothermia, bacterial infections and respiratory failure [11]. Diagnosing the condition based on a range of clinical features is difficult. An accurate diagnosis of monkeypox requires a molecular test in a specialized laboratory to distinguish it from other diseases, and sometimes it takes a long time to know the laboratory test results [12] [13].

1.1. THE MAIN PROBLEMS IN THE SUBJECT AREA

After the spread of the Covid-19 pandemic, and its great impact on the countries of the world, a new threat called Monkeypox appeared. that is actually spreading in the world. Although Monkeypox is a viral disease of animal origin, it did not only spread in forests in Africa, but it is also spreading in various countries of the world. Based on these data, the only way to avoid any massive spread of the monkeypox virus is early detection and determine of infected people so that they can be dealt with in an appropriate manner. Therefore, it is important to find a way to diagnose infection for patients accurately and quickly enough to give them the right treatment at the right time.

However, challenges are faced to the diagnosis of infection with monkeypox, including: A - The great similarity between monkeypox disease and smallpox disease, being from the same group of viruses. In addition to the similarity of clinical symptoms that appear after infection on patients. B - The risk of contracting it is not limited to infected individuals, but also affects health workers. C - Limited knowledge of this disease also among patients and health personnel increases the risk of its spread.

1.2. RESEARCH IDEA

Given the seriousness of human monkeypox disease, as well as the ability of machine learning and artificial intelligence to diagnose diseases, this study proposes to find a mechanism for early and rapid diagnosis of patients without contacting the medical staff, depending on the symptoms of the disease, using Artificial Neural Network (ANN) trained by Adaptive Artificial Bee Colony (aABC) Algorithm.

1.3. THE AIM OF THE STUDY

Dangerous infectious diseases such as Covid-19 and monkeypox can spread rapidly among people. Therefore, conducting a study that will help reduce the number of people suffering from these diseases through early detection has become necessary. In the study, the aABC Algorithm, an improved version of the classical ABC, was used for more efficient ANN training. The study leads to improve the accuracy of prediction for early diagnosis of monkeypox based on symptoms in a person. The aims explaining the importance of this research can be summarized as follows:

- Create a strategy for the early detection of monkeypox disease based on deep learning.
- Protecting the medical staff from the risk of monkeypox by preventing direct contact between the patient and the medical staff.
- Reducing the spread of infection between infected patients and non-infected persons.

1.4. SCIENTIFIC CONTRIBUTION OF RESEARCH

The main contribution of this study is the combination of neural network and artificial bee algorithm in establishing a new diagnostic model for early detection of monkeypox disease. The importance of this idea is to take advantage of the ability of the bee algorithm to find optimal solutions to numerical problems and apply it to train the neural network in order to obtain the best possible values for the weights of the neural network.

- Data analysis of patients infected with monkeypox virus according to the data taken.
- Design a machine learning model based on ANN
- Train this model depending on the modified bee algorithm
- Results obtained from empirical research depending on the database used

1.5. ORGANIZATION OF THESIS

As mentioned above, the thesis reviews and analyzes studies and algorithms that are used in the early detection of monkeypox. In addition, the effectiveness of the techniques is analyzed and compared to the data set. In Part 2, new techniques in the field of taxonomy (monkeypox infection) are studied. Some literature and trials were reviewed. Part 3 provides a theoretical overview of the algorithms used and comparison. The research methodologies for this dissertation are explained with a discussion of the experiments and results in Part 4. Conclusions and some future directions in Part 5.

PART 2

LITERATURE REVIEW

In this section, previous efforts related to methodologies for the diagnosis of monkeypox disease in medical systems will be presented. Considering that early detection of the disease is crucial to controlling its transmission. With the spread of AI applications, researchers have resorted to making use of it in diagnosing disease conditions in medical and biomedical applications [14]. They used it in multiple ways, depending on the dataset collected from the lesions' images or the infected clinical symptoms.

In the field of automatic virus identification in transmission electron microscopy (TEM) images, [14] relied on image datasets to characterize the monkeypox virus. It consists of 1245 micrographs of 22 viruses taken by TEM. However, this study was limited to 14 types of viruses, such as Astrovirus & Adenovirus & CCHF & Ebola...etc. The study relied on Convolutional Neural Network (CNN) Deep Learning (DL) models to build its model. In fact, the results of the proposed method were 93.1%.

[2] use different approaches in data acquisition. The data is a set of images of skin lesions. Collected by manual searches and contact with infected persons. The study focused on separating monkeypox from similar cases of different types of smallpox. The approach taken was VGG16 Deep Transfer Learning. It consists of three layers of convolutional filters to extract the features from the images and then the neural network. It was a perfect idea that he used Transfer Learning. The accuracy of the results obtained was 86%.

[15] research was divided into three separate studies. All of them were conducted on the proposed approach. Transfer Learning approaches (GRA-TLA) work on multiclass classification using Generalization and Regularization. The training dataset is the images of skin lesions. It was intended to support decision-making assistance to the hospital. Computational results showed that the proposed approach could distinguish between infected and non-infected monkeypox individuals with an accuracy of 77% to 88% in the first and second studies. At the same time, the residual network (ResNet) had the best performance for multiclass classification in the third study, with Accuracy ranging from 84% to 99%.

[10] also relied on training data consisting of images, it aimed to establish an early detection mechanism for monkeypox that would help identify infected people. In fact, the approach taken in this paper was to compare several models of ResNet50, EfficientNetB3 and EfficientNetB7 algorithms. In the end, it was concluded that the results of the EfficientNetB3 algorithm were the best.

In Ref [16], taking a different approach from previous research, it relied in the study on numerical data for clinical symptoms of the disease from 500 suspected cases from Spain and Nigeria. The dataset was not limited to monkeypox but included several other diseases. Such as: monkeypox, acne, alopecia, normal, psoriasis, and smallpox. The Human Monkeypox Diagnosis (HMD) strategy was applied in this study. In fact, it consisted of two phases. First: extracting the appropriate features, using the Improved Binary Chimp Optimization (IBCO) algorithm, which is a hybrid selection algorithm. The second stage: composed of three machine learning algorithms Weighted Naïve Bayes (WNB), Weighted K Nearest Neighbors (KNN), and deep learning. In fact, a final election is made for the output of these three algorithms.

Table 1. Summary of published studies on monkeypox disease

No	Study Year	Proposed Work	Methods	Dataset	Performance
1	2021 [14]	Distinguishing monkeypox virus from other viruses using microscopic images	CNN	Image datasets, 14 type of virus	Accuracy = 93.1%.
2	2023 [15]	An assistant in the diagnosis to distinguish between the types of infections that appear on the skin, and determine the disease according to them	ResNet	Dataset is Images of skin lesions	Accuracy ranging from 84% to 99%.
3	2022 [2]	The study focused on separating monkeypox from similar cases of different types of smallpox	VGG16 Deep Transfer Learning	Images of skin lesions	Accuracy = 86%.
4	2022 [10]	Early detection mechanism for monkeypox	ResNet50, EfficientNetB3 and EfficientNetB7	Images	Accuracy = 87 %
5	2023 [16]	Human monkeypox diagnose (HMD) strategy for detection monkeypox by clinical symptoms	Weighted Naïve Bayes (WNB), Weighted K Nearest Neighbours (KNN), and deep learning	Numerical data for clinical symptoms	Accuracy = 98.48%
6	2021 [11]	The study was conducted in the early period of this onset of the outbreak. The results obtained noted a rash in the anogenital area, the presence of lymphadenopathy, and fatigue, which greatly increases the likelihood of a positive test.	Interview patients orally and collect data	Interview with 140 patients	-
7	2019 [17]	The Study describes the largest documented human outbreak of the west African clade of the monkeypox virus	Was Collected data for study and determine clinical characteristics	Data were collected with standardized case investigation form, with a case definition of human monkeypox that was based on previously established guidelines	All 122 patients had vesiculopustular rash, and fever, pruritus, headache, and lymphadenopathy were also common

Table 2. Summary of published studies on monkeypox disease (continued)

No	Study Year	Proposed Work	Methods	Dataset	Performance
8	2022 [18]	It was conducted a retrospective observational analysis of people with polymerase chain reaction (PCR) confirmed monkeypox virus	Interview and Collect data	197 patients, Clinical data were collected through one of three electronic healthcare systems	-

PART 3

THEORETICAL BACKGROUND

The thesis presented in the last part summarized that result of research are easy for understanding the severity of disease and the mechanisms used in predicting prognosis of disease states. In this part we review the theoretical basis of the algorithms applied in this study. And what are the parameters that depend on their application.

3.1. ARTIFICIAL INTELLIGENCE

With the development of Artificial Intelligence (AI) technologies, machines have the ability to perform tasks that normally require human capabilities to complete. The field of AI includes a wide range of technologies and concepts, such as machine learning through which systems can improve their performance over time and adapt to the environment, natural language processing that allows machines to effectively understand and interpret human language, and computer vision that enables systems to analyze and comprehend images and videos. In addition to designing and developing robots that find use in a variety of applications and industries. [19,20].

Machine learning is an important and essential component of AI. This means that it involves developing and training computer algorithms to understand and analyze patterns in data. This type of learning may be necessary for machines that rely on improving their performance over time through past experiences and adaptive knowledge. It can also contribute to enabling systems to develop their capabilities in forecasting and making decisions based on data, making them able to adapt to continuous changes in the environment and move forward in their sustainable development.

AI shows a wide range of uses in various fields, including education, health, finance, sports and entertainment [21,22]. Among the uses of AI are voice assistants, image recognition used in cameras, and self-driving cars. And uses in various industries and fields. Here are some examples:
The capabilities of AI methods are based on simulating human mental processes and learning. These techniques show diverse applications in a wide range of fields. In the field of health, AI is an effective tool in analyzing radiographs, discovering medicines, and improving the delivery of personalized health care. Intelligent systems provide accurate recommendations to doctors based on medical data, which helps doctors diagnose diseases and guide treatment processes.

In the field of finance, AI is an important tool in monitoring financial security. Intelligent data analysis allows for the detection of unusual patterns that indicate potential fraud, in addition to the classification of financial transactions.

In transportation, AI controls traffic and directs autonomous vehicles. By analyzing traffic-related data, smart systems can direct drivers to the best routes and anticipate potential delays.

There are many sections and fields under the name AI, each of which can constitute a specific field that contributes to life. There are also wide practical applications for it. These applications contribute greatly to developing and accelerating our daily lives. In addition to making our lives safer and more secure [23]. Here are some of the main parts of AI:

Technology in the field of AI is based on a variety of techniques and applications. AI involves training algorithms to recognize patterns in data and make predictions based on these patterns, as well as understanding and generating human language in machine learning and natural language processing. Computer vision technology enables devices to recognize visual information, such as photos and videos.

Furthermore, AI is used to design and program robots to perform tasks that normally rely on human intervention. It is worth noting that expert systems simulate the decision-making capabilities of a human expert in a specific field, and neural networks rely on models that resemble the human brain and can be trained to recognize patterns.

Applications of AI also extend to machine learning in the field of speech recognition and interpretation, and solving research and planning problems, including the development of self-driving cars and robotics. In general, these technologies combine to enable machines to perform tasks that typically rely on human capabilities, contributing to the development of advanced technology in various fields.

3.2. META-HEURISTIC SEARCH PROBLEM

Meta-heuristic search problems are a class of optimization problems that involve looking for the best solution among the large and complex group from possible solutions. These problems typically involve searching through a vast number of potential solutions using heuristic methods to determine the best possible solution [24,25].

Meta-heuristic algorithms are inspired by natural phenomena, like genetic algorithm, swarm behavior, and simulated annealing. They can be used to fix wide range of issues in many parts, including engineering, finance, and logistics. Examples of meta-heuristic algorithms include:

1. Genetic algorithms: These algorithms mimic the process of natural selection, where solutions are evolved and improved through mutation, crossover, and selection.
2. Swarm optimization algorithm: The algorithm depends on the movement of swarms, where particles move around a search space and update their positions based on their own experience and that of their neighbors.
3. Simulated annealing: the algorithm depends on process of annealing in metallurgy, where a material is heated and cooled for improving its properties. In simulated annealing, the solution space is explored by gradually reducing the "temperature" of the system.

Meta-heuristic algorithms are useful when the solution space is large, complex, and difficult to explore using traditional search methods. It is used in several fields within research issues, finding the best solution, and improving pre-existing solutions to reach the ideal solution. [19,26–28].

3.2.1. ATIFICIAL BEE COLONY (ABC) ALGORITHM

It falls under swarm intelligence algorithms. Inspired by nature, mimics the work of bee swarms [29]. When bees go to a field of flowers, they look for the places where the flowers are most abundant. The bees use a swarm optimization algorithm to go to this area. [30]. The bee's proximity search theory follows an efficient design that takes advantage of the bees' biological strategy. In this model, bees initially scatter the area in search of flowers, and each bee records which areas contain the greatest number of flowers. The bees then move randomly in the area and update their information if they find an area containing more flowers. The effectiveness of this strategy lies in enabling the bees to explore the area comprehensively.

Upon completion of the random search, each bee announces what it has found, enabling the swarm to know the best places to collect nectar and make the most of it. This strategy represents a successful application of AI in the field of search and exploration applications where random search and intermittent updating techniques can be leveraged to efficiently achieve certain goals. [31].

It is one of the methods of AI in Meta-Heuristic research problem [32]. ABC is an effective tool in finding and improving solutions. In general, its work can be summed up in finding and exploiting sources of food and then looking for a new alternative. In ABC algorithm, food reflect the initial solutions through which bees will look for the perfect solution. The quality of the food fitness represents the assessment of the solution, while the limit factor indicates the amount of food available in the source. While Cycle represents the number of searches [33].

The bees are divided into three main groups, (a)- Scout Bees: Their mission is to find food sources, which is random. (b)- Employed bees: bees’ numbers are same sources numbers of food, each bee is entrusted to go to a specific food source from these sources, in addition, it does a local search for a new source next to her own. As described in Eq. 1. (c)- Onlooker bees: also, Onlooker bees' numbers are same sources numbers of food [34,35].

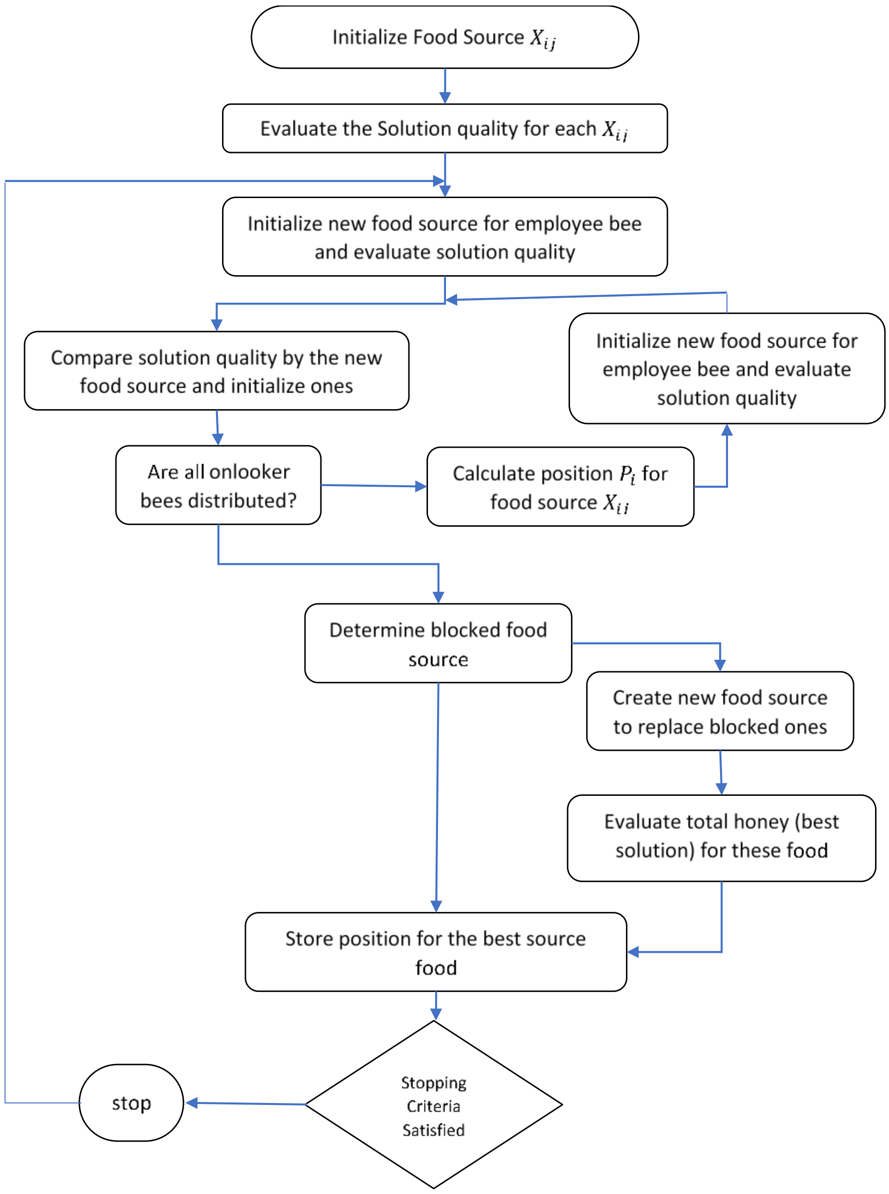


Figure. 3. The principle and stages of the ABC algorithm [33]

It monitors employed bees. when they come and by vibrating bees determine the source of appropriate food to go to.

$$v_{ij} = x_{ij} + \emptyset_{ij}(x_{ij} - x_{kj}) \quad \textbf{(1)}$$

The fundamental steps involved in creating the ABC algorithm can be outlined as follows: - Generate an initial solution group. - Sending employed bees to food sources. - Sending Onlooker bees to the most appropriate source of food. - Save the optimal source. - Repeat previous steps**.** Figure. 3. shows the working diagram of ABC algorithm.

3.3. MACHINE LEARNING (ML)

Machine Learning ML is an important branch within the field of AI, as it involves teaching computer systems to recognize recurring patterns in life that humans cannot model. These patterns can be learned based on previously collected incoming data. This helps in automatically learning the data and improving its performance on a particular task. By relying on data and finding patterns, the most prominent role of the machine is to detect recurring patterns in the data without the need to program it explicitly.

Therefore, the main goal of ML is to give automated systems the methodology to recognize patterns in data, analyze them and find hidden relationships between them, in addition to making accurate predictions to uncover new cases that were not previously trained or make decisions based on that data. [36,37].

Achieving ML requires feeding huge amounts of data into algorithms designed for this purpose. This data is fed to systems that rely on ML techniques. Statistical methods are then used to analyze this data with the goal of identifying patterns, trends, and relationships within it. There are three main types of ML:

1. Supervisor learning: Supervisory learning is one of the ML mechanisms in which algorithms rely on how to map the trained data set to pre-defined output values through defined examples. This type of learning provides a data set and an output corresponding to it in the data set, then training the algorithm to understand and find a relationship between these two series through which it can predict new cases. Therefore, it can be said that the goal of supervised learning is to enable the algorithm to accurately predict the output of new and previously unseen input data. This process

relies on the use of defined examples to analyze patterns and relationships in the data, allowing the algorithm to develop a generalization function that can predict the output of new input data based on the acquired experience. Supervised learning is a powerful tool for solving a variety of problems. [38].

2. Unsupervised learning: Unsupervised learning is a type of ML with wide importance and applications because it allows algorithms to discover patterns in input data without the need for clear output labels. In this type of learning, the attached training data is not pre-parameterized, as the dataset is provided with the data received as input, but no output corresponding to the trained data. The goal of unsupervised learning is to apply the clustering principle to sort the received training data into groups. These groups represent patterns on which the data can be segmented. In addition to finding relationships between these groups, which helps in understanding this trained data structure. The algorithm searches for structures, similarities, differences, and relationships between data points without clear guidance from the data. Unsupervised learning is a highly effective tool for exploring patterns and deeply understanding data. [39].

3. Reinforcement learning: is a distinct type of ML that involves training an agent or AI system to make decisions based on the feedback it receives from its environment. In the process of reinforcement learning, the agent interacts with the environment, learns how to interact in it, and receives that feedback in the form of rewards or punishments based on its performance. The main goal of reinforcement learning is to enable the agent to develop strategies that help it achieve higher cumulative rewards over the long time run. The agent learns by experiencing and exploring various behaviors, and monitoring the impact of those behaviors on those rewards. Over time, the agent develops plans and strategies that enable it to make the most feasible decisions to achieve its goals and sustainably increase its rewards. [40,41].

It should be emphasized that ML has a wide range of practical applications, including image and text recognition, natural language processing, recommender systems, and predictive analysis. It is a rapidly growing field that has the potential to transform many industries and fields.

3.3.1. ARTIFICIAL NEURAL NETWORK (ANN)

An ANN is a model inspired by how the human brain handles data, stores and processes data. It consists of a large number of nerve cells called neurons. These cells work together to process information, access data, and make predictions based on that information.

At the heart of an ANN are neurons, which receive signals from other cells or external sources, and process these signals to form an output signal. Each signal is assigned a weight, and this weight determines the strength of the connection between neurons. The output signals from a neuron are calculated by applying an activation function to the weighted sum of the input signals. [42].

The process of training an ANN is called back propagation. Through it, all weights between neurons are adjusted, depending on the error resulting from the prediction process. This error is calculated by the difference between the actual output and the expected output. The ANN training process is an iterative process of correcting weights. It can be considered a cumulative process of correcting weights with each training period. [36,43,44].

ANNs find many and varied applications in several fields. They are mainly used in image and speech recognition, natural language processing, and prediction modeling. It shows high effectiveness in analyzing complex patterns and relationships in data, and finds wide applications in various industries, including finance, healthcare, and the manufacturing industry. [40,45].

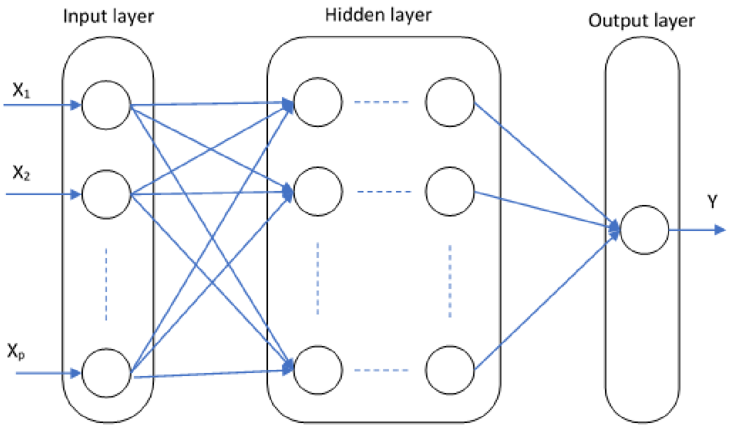


Figure. 4. Structure Artificial Neural Network (ANN) [42]

Structure of an ANN consists from layers of connected processing cells called neurons. As shown in Figure. 4. There are three types of sections in ANN:

Input Layer: is responsible for receiving incoming data and transmitting it to the next layer for processing. Each neuron in this layer represents an attribute or feature of the input data.

Hidden layer: is the main place where most of the computations in an ANN take place. They are called "intermediate" because they are not directly connected to the input or output layers. Each neuron in the intermediate layer receives signals from the previous layer, calculates the weighted sum of these signals, and then applies an activation function to produce the output.

Output layer: output layer carries the responsibility of generating the ultimate outcome of ANN. The output layer contains one neuron for each output class or prediction that the ANN is trained to produce.

ANNs can have different structures, depending on count of layers and neurons in layers. A shallow ANN has only one hidden layer, while a deep ANN has multiple hidden layers. Each layer can contain a different number of neurons in each layer. This difference depends on the complexity of the problem to be solved. It is worth noting that the structure of an ANN can significantly affect its performance and ability to generalize to new data. Choosing the right structure for a particular problem is an important part of designing an effective ANN.

The activation function is an essential component of the process by which ANNs make decisions. An activation function is applied to the outputs of each neuron in the ANN after calculating the weighted sum of its incoming signals. This function determines whether a neuron will fire and send a signal to the next layer or not.

The activation function takes into account the weighted sum of incoming signals and compares it to a certain threshold. If the sum is greater than this threshold, the neuron is activated and sends a signal to the next layer. If the sum is less than the threshold, the cell is disabled and does not send a signal.

The activation function can be different for different models and tasks. For example, the activation function could be a Sigmoid function that evaluates signals between zero and one, or a ReLU (Rectified Linear Unit) function that activates only if the sum is positive and disables it if it is negative. [36,37].

Some different activation functions are used by the neural network during the reversal learning process, each with its own strengths and weaknesses. The most commonly used activation functions in neural networks are:

1. Sigmoid function: It converts signals into a range between zero and one. If the sign is negative, the output is close to zero, and if positive, it is close to one. Used in models that require classification results.

2. ReLU function: Simply put, it activates neurons if the weighted sum is positive and deactivates them if it is negative. This function is commonly used in many models due to its high efficiency.

3. Linear Activation: It does not perform any signal conversion and simply acts as an addendum on the input. They can be used in networks that require a simple model.

4. SoftMax: is one of the common functions used in the output layer in ANNs. It is of great importance in multi-class classification problems. The SoftMax function is used to convert the output from the neural network into a probability distribution over the different classes. When the model learns from the data and arrives at the final results in the output layer. These results are usually expressed as a set of numerical values. The SoftMax function takes these values and converts them into a probability distribution that represents the probability of each class.

The activation function of the neural network is determined according to the nature of the problem to be solved and the structure of the neural network. In general, the activation function should be able to produce non-linear outputs and should be differentiable to allow for gradient-based optimization during training [39,40].

3.3.1.1. TRAINING OF ANNs

The idea of training ANNs using back propagation is a fundamental concept in the field of machine and deep learning. This process represents the process of a neural network learning from errors and differences between expected output and actual output. Training begins by feeding the network with training data. This data includes inputs and expected outputs. A neural network responds to input by processing the data and generating an output that approximates the expected output. The difference between expected output and actual output is calculated. This difference is known as error and is used as an indicator of network performance.

The back propagation algorithm is the main step in the training process. Back propagation is used to modify the weight and slip parameters in the network in order to reduce the error. This process is done so that the error is fed backwards through the network starting from the final layer down to the first layers.

The weight and slip coefficients are generated based on this error such that the error is gradually reduced. This work is repeated many times with multiple training stages until the network performance is improved and the error is significantly reduced. The number of cycles, data size, and network architecture depend on the complexity of the problem and the desired goal.

This process is called “back propagation” because it works in reverse from the process that data passes through the network during prediction. This process is repeated several times until the network learns how to reduce the error and improve its performance in predicting the data. This process enables neural networks to perform tasks such as classification, prediction, and pattern recognition with greater accuracy.[46,47].

Steps to implement the back propagation algorithm:
1. Forward propagation: Training data is fed into first layer, and then result is appeared by applying process on weights then give them to activation functions of the neurons.
2. Error calculation: The process of calculating the difference between expected output and actual output is a critical step in the training process. It is used as an indicator of how well a network's performance matches expected results. When the actual output exactly matches the expected output, the error is zero. However, in most cases, there is a small difference between the two outputs, resulting in a non-zero error value.
3. Weight update: weights between the neurons are updated in a way that minimizes the error.
4. Repeat: Steps 1-3 are repeated for multiple epochs, or passes through the training data, until the error is minimized.

The backpropagation algorithm is a technique used in ML and is one of the most important methods in training ANNs. This algorithm is used in multiple fields where we have labeled data and we want to train a model based on this data to do a particular task. It enables ANNs to perform complex tasks such as classification, prediction, and pattern recognition with high accuracy. This algorithm is key to why neural networks outperform humans on tasks in some cases. [48].

3.3.1.2. LEVENBERG-MARQUARDT (L-M) ALGOITHM

The Levenberg-Marquardt (L-M) Algorithm is an optimization algorithm commonly used to train ANNs. It is a modification of the standard Gauss-Newton algorithm and is used to solve the least-squares problem in nonlinear regression [49,50].

The L-M algorithm is an iterative algorithm that works based on the calculated error, the weight of the connections between neurons is updated. This update is done using the rule of mathematical behavior and differentiation. During each iteration, the algorithm estimates the Hessian matrix is a matrix containing second order partial derivatives. and uses it to update the weights and biases of the neural network. In Marquardt's update relationship as Eq. 2., the damping parameter λ is scaled by the diagonal of the Hessian $J^T W J$ for each parameter.

$$[J^T WJ + \lambda\, diag(J^T WJ)]\, h_{lm} = J^T W\, (y - \hat{y}) \qquad \textbf{(2)}$$

The L-M algorithm combines advantages of two other optimization algorithms, the Gradient Descent (GD) method and Gauss-Newton method. It has the ability to converge quickly and accurately to the minimum of the error surface, and it can also handle non-convex error surfaces, making it well-suited for training ANNs.

The L-M algorithm is an algorithm used in a variety of applications. This algorithm is effective in solving optimization and error correction problems in mathematical models, neural networks, and other applications.

The advantage of the L-M algorithm lies in its ability to adapt to such complex models and big data, which makes it useful in dealing with problems of high complexity and accurate prediction. However, it should be noted that they are sensitive to initial conditions and may require some tuning of their parameters to achieve optimal performance. This means that users must adjust the parameters of the algorithm and adapt it according to the specific problem and the available data. [51–53].

3.3.2. NAÏVE BAYES (NB) ALGORITHM

Naive Bayes (NB) is a probabilistic algorithm commonly used in ML for classification tasks. It is based on Bayes' theorem, a mathematical formula that describes the relationship between the probability of a hypothesis and observed evidence

Naive Bayes (NB) algorithm Its name is based on Bayes' theorem which is used to analyze the relationship between probabilities and available evidence. The algorithm is considered “naive” because it assumes independence of the variables or attributes used for classification, a strong and simple assumption that allows easy calculations.

The idea of the Naive Bayes algorithm is based on using Bayes' theorem to calculate probabilities and statistical laws to classify data. The algorithm is trained using a data set containing pre-labeled examples, and the evidence (features) associated with each example is analyzed. This evidence is then used to classify new data. As Eq. 3**.** given below states:

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \tag{3}$$

In Naive Bayes, the algorithm makes predictions by calculating probability of a given data has a specific classification. The algorithm assumes that the features of the input independence from each other gave it the name “naive”. This assumption allows the algorithm to simplify the calculations and make predictions efficiently [16,54].

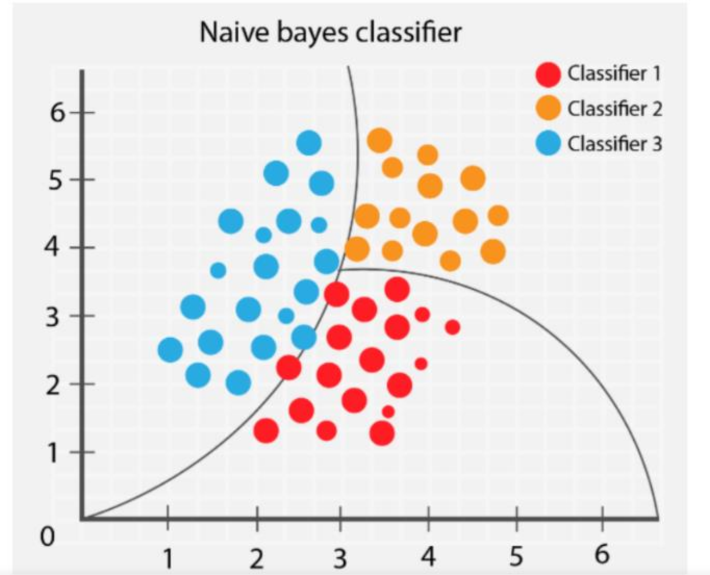


Figure. 5. Naïve Bayes Classifier [55]

NB algorithm works on the basis of analyzing data and building a classification model based on probabilities. The process first begins with a training phase where we are presented with a set of data containing attributes (features) with previously known labels. The algorithm calculates the probabilities of each feature's presence based on its frequency in the training data. These probabilities reflect the extent to which the presence of a feature is associated with the presence of a particular classification.

Once the training phase is completed, the model can be used for classification. When new data is given to the model, the algorithm calculates the probability that the new data will be in each possible classification. This is done by using the probabilities that were calculated during the training phase. Once probabilities are calculated for each classification, the classification with the highest probability is chosen as the predictor for the new data., Figure. 5.

The NB algorithm can be used in text classification tasks due to its flexibility and efficiency in dealing with text data, and in classifying electronic messages into “spam” and “non-spam” categories based on the content of the messages. In the field of natural language processing, to classify texts into expressive categories such as positive, negative, or neutral. This is widely used in sentiment analysis on social media and evaluation of public reactions [16,55].

One of the advantages of NB is a good choice in many of these fields due to its simplicity and effectiveness, and has achieved success in multiple classification tasks. It also is very efficient, even large amounts of data. However, it does make the strong assumption of feature independence, which may not be accurate in some cases [56].

3.3.3. K-NEAREST NEIGBOURS (KNN) ALGORITHM

The K-Nearest Neighbors (KNN) algorithm is one of the famous algorithms in the field of ML and can be used in classification and regression tasks. This algorithm is based on the principle of instance-based learning, which makes predictions by taking advantage of the information available about “nearest neighbors” in the training data. [57].

This letter indicates the number of relatives used to predict the class or value of a new data point. Choosing the appropriate value of K is an important step in applying KNN. The value of K can be any positive integer. If K is small (e.g., K=1), the model will rely heavily on noise in the data and will overestimate detail. If K is large (e.g., K=10), the model will depend more on the overall structure of the data.

To determine the K-nearest, the algorithm must calculate the distance between the new point and all the training points. Here different measures of distance can be used, such as Euclidean distance (or Yoruba distance) and urban distance (Manhattan distance).
After calculating the distance and selecting the K relatives, the algorithm calculates the classification ratio for each class based on the class labels of the K relatives. Different techniques can be used to calculate these ratios, such as simple majority where the class with the largest number of relatives is chosen as the classification for the new point.

The value of K affects the performance of the model. If K is too small, the model may be sensitive to noise in the data and show an underestimation of detail. If K is too large, the model will be less sensitive and show a more generalized estimate. as shown in Figure. 6. [16,58].

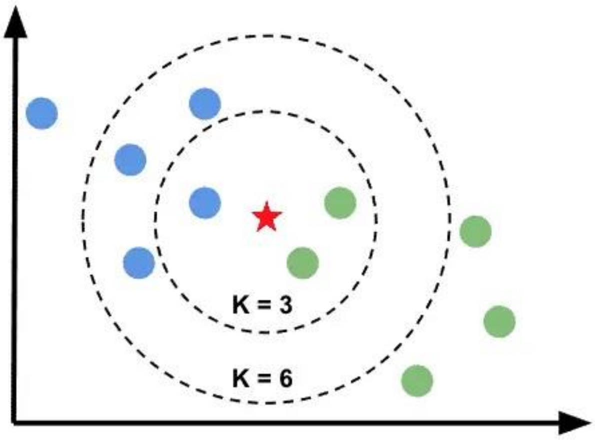


Figure. 6. KNN Algorithm [16]

In the Classification task, the goal is to classify new points into one of the designated classes. KNN works by searching for the K nearest neighbors of the new point in the training data. Once the neighbors are found, the most common class among these neighbors is chosen as the classification of the new point. For example, if you use KNN to classify flowers based on their features, you will find K nearest neighbors of the new flower and the most common class label among those neighbors will be the one assigned to the new flower.

In a regression task, the goal is to predict a numerical value based on training data. Instead of classifying points into specific classes, KNN is used to find K-nearest neighbors and calculate a numerical value such as the mean or median of the target values for these neighbors. For example, if you use KNN to predict the price of real estate based on the features of homes, it will find K-nearest neighbors and predict the median or median price of the new point home. [16]. The distance between each testing sample and each training samples is calculated using the Euclidean distance, mentioned in Eq. 4.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{4}$$

The KNN algorithm can be useful when data distribution couldn't be well-known or when the relations between features and target variables are complex. However, the algorithm can also be computationally expensive and sensitive to choice of distance metric and value of K [59,60].

3.3.4. SUPPORT VECTOR CLASSIFICATION (SVC) ALGORITHM

Support vector classification (SVC) algorithm is one of the popular ML techniques which is used in classification tasks. The main idea behind SVC is to find a hyperplane (margin) that can separate different classes in the training data with the largest possible margin to be able to identify the appropriate class. As shown in Figure. 7.

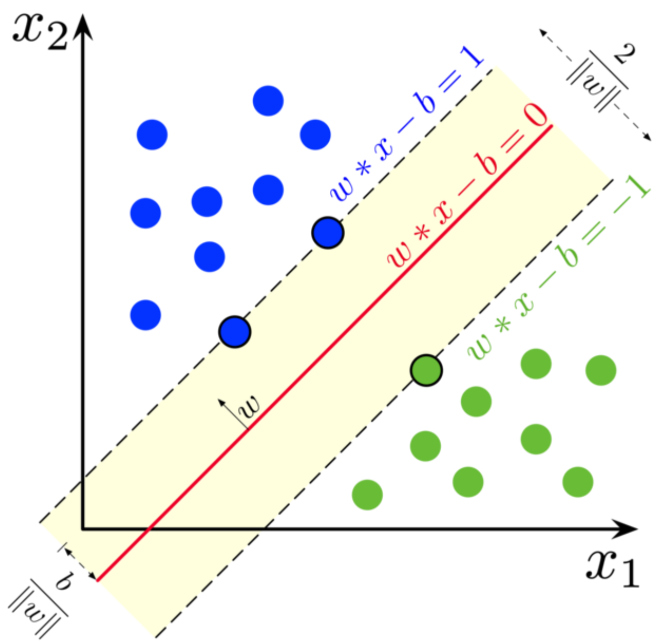


Figure. 7. SVC Algorithm [61]

In the context of binary classification, the SVC algorithm finds the hyperplane that maximizes marginal distance between points in each type of data, also known as vectors. The algorithm then assigns each new data is classified depending on which side of the plane it belongs to it falls on [61].

Kernel functions are commonly used in SVC algorithm to handle cases where classes are not linearly separable in the original feature space. The kernel function allows data to be transformed into a higher dimensional space where it becomes possible to separate classes by a hyper level. The linear kernel function is used when linear separation between classes is possible in the original space. They are simple and used when there is no need to increase dimensions. Polynomial Kernel Used when nonlinear separation is possible. [62].

The data is transformed into a higher dimensional space using periodic drivers (such as a quadratic basis) to separate classes. Radial Basis Function (RBF) is commonly used and is used when nonlinear separation is more complex. This function is particularly known for its strong impact on data transformation. It is used to deal with nonlinear classification problems. Sine Kernel is based on the sine function. It is used to deal with nonlinear classification problems and shows good results in some applications. [63,64].

The use of kernel functions allows improving the ability of the SVC algorithm to handle a variety of problems and provide accurate performance in nonlinear classification situations. These functions are an essential part of the power and flexibility of the SVC algorithm in ML.

3.3.5. RANDOM FOREST (RF) ALGORITHM

Random Forest (RF) algorithm is one of the famous algorithms in the field of ML and is considered very effective in performing classification and regression tasks. This algorithm is based on the idea of forming a set of sub-decision trees to make predictions and make decisions.

Several subtrees (decision) are created, which are completely independent of each other. The data is randomly selected from the original dataset and randomly selected features for each tree. Each tree is trained on its assigned subset. As shown in Figure. 8.

Trees use recursive partitioning, where they partition data based on a feature that you select by searching for the best partition. The partitioning process continues until all points are successfully classified into subcategories. Once all the trees are trained, they are used together to vote when the model has to make a decision. In the case of a classification task, the subtrees compute independent classifications and the classification that receives the largest number of votes is chosen. In a regression task, the mean or median value of the predictions generated by the subtrees is calculated. RF tends to reduce the problem of overtraining by collecting multiple views from a variety of trees. This allows for the construction of a robust and accurate model [65].

RF have several advantages over single decision trees, such as reducing the risk of overfitting, increasing the accuracy of predictions, and providing a measure of feature importance. The algorithm treats missing values and noisy data and uses for both binary and multiclass classification.

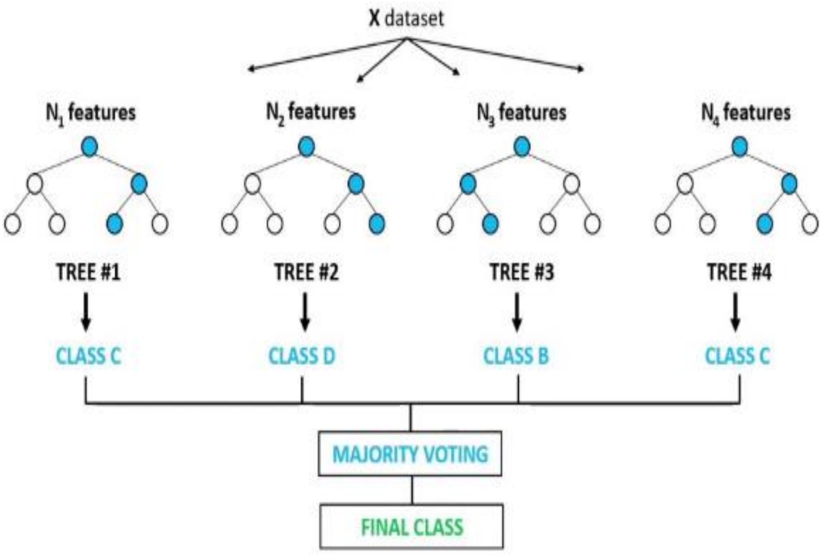


Figure. 8. Random Forest Algorithm [65]

RF is commonly used in applications such as bioinformatics, image classification, and remote sensing. However, the algorithm can be computationally expensive, and the model interpretability can be limited due to the complexity of the ensemble model [17,66,67].

3.3.6. GRADIENT BOOSTING ALGORITHM

Gradient Boosting is a powerful algorithm widely used in ML for classification and regression tasks. It is a method of learning a set of weak models and putting them together to build a strong model. It starts by creating a preliminary model, usually using a weak model that can initially make predictions about the data. The error between current predictions and actual values in the data is calculated. Next, the derivative of the error is calculated. The goal is to modify the current model to reduce this error.

Another model is created and trained on error predictions. This new model is introduced so that it helps reduce the overall error. The new model is added to the set of existing weak models. This combination is then used to predict new data. This process is repeated over and over again, with new models being created with each cycle and iteration. After several iterations, the weak models are combined to build the final Gradient Boosting model. The result is a more powerful model that can handle the data better and provide accurate predictions. A variety of functions can be used to calculate error and adjust weak models, and these include mean square error, entropy, vectorization, and many others. [68].

Learning rate is the step rate that determines the amount of modification made to weak models in each iteration. If the learning rate is low, the models will learn slowly and the process will be more stable. If it is high, the models will learn quickly but there may be a problem with bursting. The learning rate must be carefully adjusted to obtain the optimal balance. While the number of iterations is the number of times the process is repeated, new weak models are created in each iteration. This number can be critical, as it must be carefully chosen to balance performance and time. A variety of weak models can be used in a gradient boosting algorithm, including decision trees, linear models, neural networks, and many other types. The most suitable type can be selected according to the specific problem and model performance.

Different techniques can be used to improve the performance of the gradient boosting algorithm. For example, the weak model can be abstracted and dimensionality reduced to reduce complexity. It is also possible to reduce the depth of the model or limit the number of trees in the ensemble. Loss Functions Loss functions are used to estimate the error between predictions and actual values. An appropriate loss function can be selected and customized according to the particular problem.[17,69].

3.3.7. DECISION TREE ALGORITHM

Decision tree is a powerful ML algorithm used for classification and regression tasks. It can be easily understood through a similar analysis of the decision-making process undertaken by individuals in daily life. A decision tree begins by collecting all training data at the root of the tree.

It then divides the data into subsets based on a specific feature. The feature is selected based on the purity criterion, where the algorithm must search for that feature that contributes to dividing the data in a way that increases the purity in the subgroups [72].

The goal of partitioning is to increase purity. Purity is usually measured using metrics such as entropy and Gini impurity. Purity reflects how homogeneous the data are in a subset. The lower the entropy or impurity, the purer the subset. The process is repeated iteratively based on different features of each subgroup. The power lies in segmenting data interactively to create a tree structure of decisions.

The process continues until a certain condition is met, such as a certain number of divisions or a certain purity. The tree is built based on these divisions and the final decision comes at the end of the process. Once the tree is built, it can be used to classify or predict new data. This is done by passing new data through the tree based on rules developed while building the tree. [70,71].

Visualizing and understanding the workings of a tree is a relatively easy task. People who are not specialists in ML can easily understand the decisions made by the tree and why. This makes it ideal for settings that require transparency in explaining why a decision was made. Decision trees can handle a variety of data, including categorical data (categorical data) and numeric data (continuous data). This makes it suitable for classification and regression tasks. Unlike some algorithms that require complex preprocessing of data, raw data can be used without complex preparation when using decision trees. This saves time and effort in data processing. Decision trees can perform well on many tasks, especially when they are configured appropriately and the appropriate feature is chosen for partitioning [73–75].

3.3.8. BAGGING CLASSIFIER ALGORITHM

Ensemble Classifier aims to improve model performance by combining several basic classifiers to produce a more accurate and robust final model. The concept of mobilization is like assembling a group of experts to make a joint decision, where the diversity and opinions of experts are leveraged to gain a better estimate. Several basic classifiers are trained on subsets of the training data.

These subsets are created using a bootstrapping technique, where data are randomly selected with replacement, resulting in some data being duplicated and others missing. After the base classifiers are trained, they are used to predict the new data class. The predictions of the base classifiers are summed and the final class is calculated based on the majority predictions or based on the weights assigned to each base classifier if weighted summation is used [76,77].

In a bagging classifier, a set of base classifiers is trained on subsets of the training data. These subsets are created using a technique called bootstrapping, which involves selecting data at random with replacement. After training the base classifiers, the prediction process is done by each base classifier on the test data.

The predictions of the base classifiers are then summed and the final class is determined based on the majority predictions or using a weight for the base classifiers' vote if weighted summation is used. This helps reduce contrast and increase accuracy and stability. Bagging classifier is widely used in automated classification for a variety of tasks such as image classification and text classification.

Examples of popular bagging classifiers include Random Forest, Bagged Decision Trees, and Bagged Support Vector Machines, which can be tuned with different parameters to achieve optimal performance in specific applications. However, you may face some challenges in dealing with imbalanced or high-dimensional datasets, and you may need to consider a balance between accuracy and performance when choosing the number of underlying classifiers and configuration parameters. [70,78–80].

PART 4

METHODOLOGY

This section describes a proposed approach for Identifying infection early in monkeypox patients. Help diagnose monkeypox quickly. as shown in Figure. 9.

4.1. ADAPTIVE ARTIFICIAL BEE COLONY (aABC) ALGORITHM

aABC algorithm overcomes some of the problems faced by ABC algorithm. It relies on problem to be solved. In fact, aABC agrees with ABC in the main bee divisions, employed bees, scout bees, and onlooker bees. But the modification that occurs in the mechanism of action of both onlooker bees and scout bees (Inspired by [81,82]).

First Amendment: Onlooker bees select the best food source from a group of foods accessed by employed bees. However, onlooker bees can evaluate the solution based on the specific fitness of each solution according to the following Eq. 5.:

$$P_i = \frac{fitness[i]}{\sum fitness[i]} \tag{5}$$

Then, the mean fitness of all solutions is calculated. Onlooker bees only search for solutions with greater than mean fitness. The goal of this process is to search further in the set of solutions with low fitness. Hence giving it more importance. Second Amendment: Scout bees remove spent solutions from the entire population. The depletion of the solution is calculated by limit factor. In fact, each solution is assigned limit value depending on the fitness of the solution. According to the following Eq. 6.:

$$Limit[i] = \frac{fitness[i] * food * D}{\sum fitness[i]} \quad \textbf{(6)}$$

Because the bee algorithm gives high results in search issues, these changes have affected the accuracy of the algorithm [83].

4.2. PROPOSED MODEL

It aims to take advantage of the ability of aABC to search on the ideal solutions, in training the weights of the ANN [42]. Figure. 9. shows the method for training ANN weights using aABC. The proposed model is a neural network composed of one input layer, two hidden layers, and one output layer. This network is trained by aABC algorithm. which are as follows:

• Generate an initial population to search for weights. At this stage, all the neural network weights are arranged in the form of Vectors, each cell in this vector represents a specific weight of the neural network weights. This vector takes arbitrary values (the initialization of the neural network weights). However, population size is related to food factor, which determines the number of vectors that will be generated. The length of a single vector is defined by D, which is weights of the neural network.

• Each vector is evaluated using the RMSE equation as in **Eq. 7.** Vector evaluation is calculated after all the training data has been passed and the resulting error is calculated. In fact, each vector has its own fitness.

$$RMSE = \sqrt{\sum_{i=1}^{N} (y_o - y_p)^2 \Big/ N} \quad \textbf{(7)}$$

• The weights training process follows the previously mentioned bee algorithm methodology. The number of training times is subject to the Epoch factor.

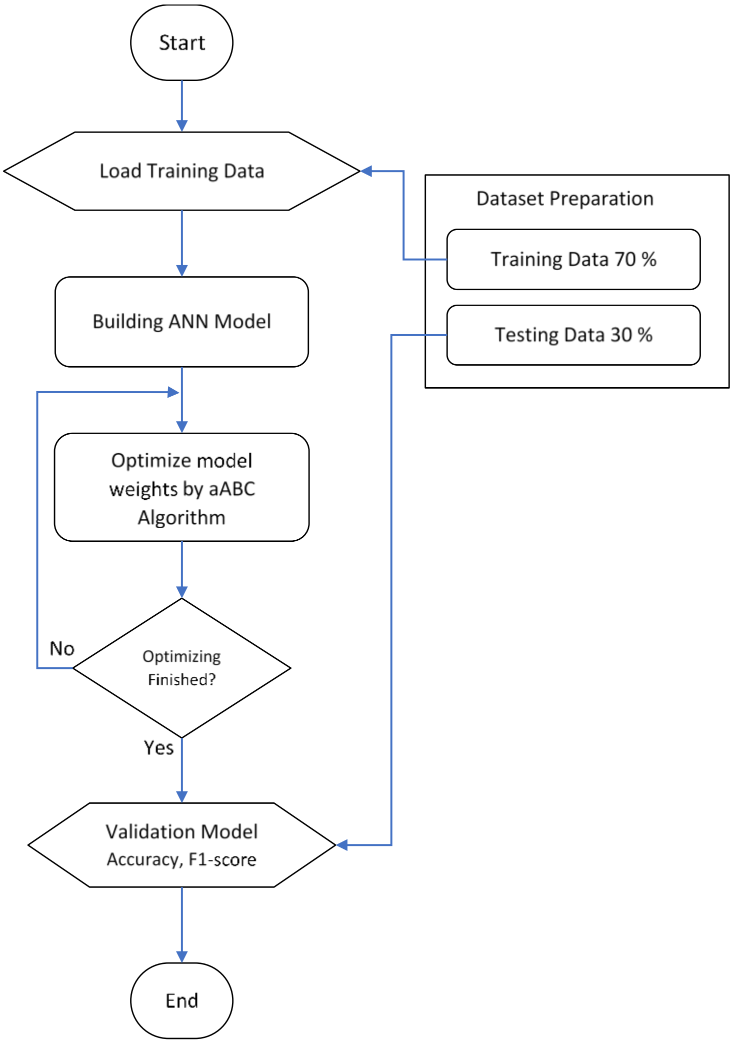


Figure. 9. Flow Diagram of Methodology

• Vector with the least error value is selected and saved. In each Epoch cycle It is considered as the ideal solution within this cycle. In fact, this cycle's best vector is compared with the previous best vector, and the vector with the lowest fitness is retained.

• At the end of the training, best vector representing the weights of the neural network is obtained among all cases.

• The final neural network is evaluated, verified and measured for accuracy. Figure. 6. shows the mechanism of the proposed model.

4.3. EXPERIMENTAL STUDY

This section describes the training data set, along with its processing and mentions some of its characteristics. And also, the performance measures for the proposal model, and the hyperparameter by which the model was set.

4.3.1. USED DATASET

The database provided for the study was obtained from Kaggle [84]. This dataset was based on patients with monkeypox, and other suspected cases. This data is published according to the bmj center. This data contains 240 diagnosed cases with 11 features which are described in Table 2. This dataset contains newly infected patients who show symptoms of monkeypox.

Table. 3. Detail of Dataset

SN	Attribute	Type	Value
1	Patient_ID	Numerical	[1 - 240]
2	Systemic Illness	Nominal	Fever, None, Swollen Lymph Nodes, Muscle Aches and Pain
3	Rectal Pain	Nominal	True, False
4	Sore Throat	Nominal	True, False
5	Penile Oedema	Nominal	True, False
6	Oral Lesions	Nominal	True, False
7	Solitary Lesion	Nominal	True, False
8	Swollen Tonsils	Nominal	True, False
9	HIV Infection	Nominal	True, False
10	Sexually Transmitted Infection	Nominal	True, False
11	MonkeyPox	Nominal	Positive, Negative

From the 240 data set we have 120 cases of monkeypox, and 120 healthy cases. Patients with monkeypox are considered positive cases, whilst healthy individuals are considered negative instances. A negative case does not always imply that the person is healthy and free from monkeypox. But, based on this information, we can tell if he merely had monkeypox. Figure. 11. Snapshot of the monkeypox dataset. There are 11 characteristics in this dataset, including the patient's clinical symptoms like fever and inflammation. These characteristics are used to describe the symptoms that a patient experiences in order to convey the patient's condition.

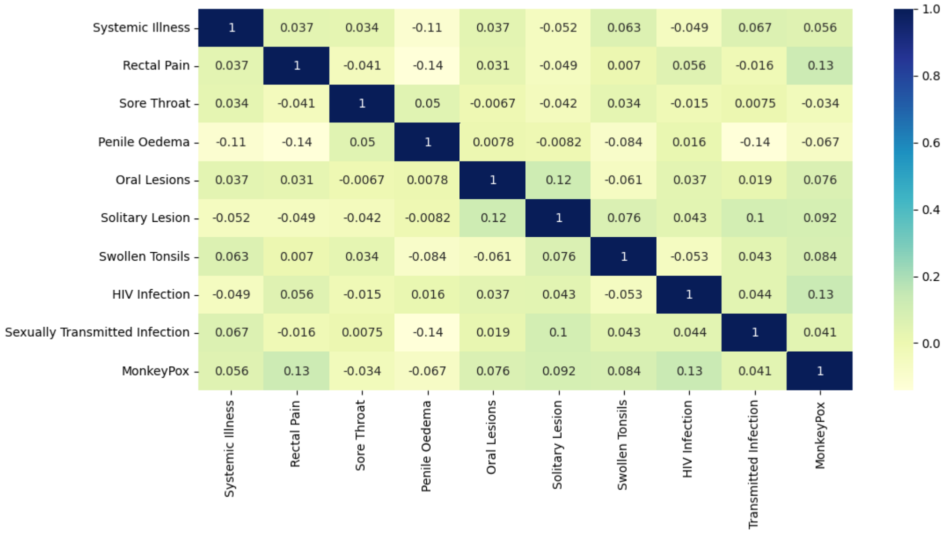


Figure. 10. heatmap displays the correlation between dataset features.

In fact, the monkeypox dataset was separated into 168 sample as a training set of data and 72 sample as a test set of data. The number of positive cases is 120 and the negative cases are 120. Figure. 10. shows how the features relate to each other.

Patient_ID	Systemic Illness	Rectal Pain	Sore Throat	Penile Oedema	Oral Lesions	Solitary Lesion	Swollen Tonsils	HIV Infection	Sexually Transmitted Infection	MonkeyPox
P0	None	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	Negative
P1	Fever	TRUE	FALSE	TRUE	TRUE	FALSE	FALSE	TRUE	FALSE	Positive
P2	Fever	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	Positive
P3	None	TRUE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	Positive
P4	Swollen Lymph Nodes	TRUE	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE	Positive
P5	Swollen Lymph Nodes	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Negative
P6	Fever	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	Positive
P7	Fever	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	Positive
P8	Muscle Aches and Pain	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	Positive
P9	Fever	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	FALSE	Negative
P10	Muscle Aches and Pain	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	Negative
P11	Swollen Lymph Nodes	TRUE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	Negative
P12	Fever	TRUE	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	Positive
P13	Swollen Lymph Nodes	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	Positive
P14	Swollen Lymph Nodes	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	Negative
P15	Swollen Lymph Nodes	FALSE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	Positive
P16	None	TRUE	TRUE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	Positive
P17	None	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	Positive
P18	Muscle Aches and Pain	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	Negative
P19	Swollen Lymph Nodes	TRUE	FALSE	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE	Positive
P20	Fever	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE	FALSE	FALSE	Negative
P21	None	FALSE	TRUE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	Negative

Figure. 11. Snapshot form used dataset.

Initial instances often contain noisy and missing values. Therefore, it need to pre-process the raw data to achieve good results. All data set used has been verified. In

(Systemic Illness) we converted the categorical attribute to numeric. The dataset does not have any missing values. Furthermore, we performed correlation analysis on these datasets, when two attributes are closely related, one of them needs to be omitted to achieve better results.

4.3.2. PERFORMANCE EVALUATION

Statistical methods are used to measure the accuracy of classification algorithms. These methods contribute to determining the standardization of the applied algorithm such as: accuracy, precision, F1-score, and sensitivity. In our dataset, Monkeypox can be classified as True Positive or True Negative if the individuals have been accurately classified. It can be classified as False Positive or False Negative if misdiagnosed. Specific statistical measures are detailed in Table. 3. [85]

Table. 4. Statistical methods for measuring the accuracy of a machine learning model

Method Name	Equation
Accuracy	$\frac{T_p + F_p}{T_p + T_n + F_p + F_n}$
Precision	$\frac{T_p}{T_p + F_p}$
Sensitivity	$\frac{T_p}{T_p + F_n}$
F1-score	$2 * \frac{Recall * Precision}{Recall + Precision}$

If the individuals have been correctly categorized, monkeypox in our dataset can be classed as True Positive or True Negative. If misdiagnosed, it may be labelled as a False Positive or False Negative. Figure. 12. illustrates these properties [16,85]. As a result, the following estimated values are provided:

- *True Positive (TP):* It predicts positive values when its true values are positive.
- *True Negative (TN):* It predicts negative values when its true values are negative.
- *False Positive (FP):* It predicts positive values when its true values are negative.
- *False Negative (FN):* It predicts negative values when its true values are positive.

Tools by which algorithm accuracy is measured depending on Confusion Matrix are:

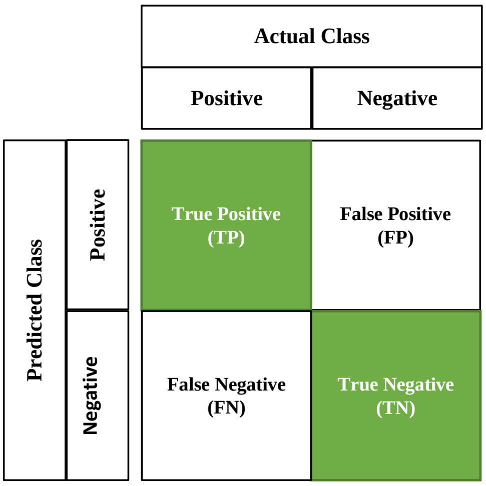


Figure. 12. Confusion Matrix [85]

4.3.2.1. ACCURACY

This way is used to describe the performance of a classifier based on the correctly predicted states versus the overall states. As in Eq. 8.

$$Accuracy = \frac{T_p + F_p}{T_p + T_n + F_p + F_n} \quad \textbf{(8)}$$

This measure is not considered sufficient to be considered the best model, if the data set is not balanced.

4.3.2.2. PRECISION

It determines the ratio between actual positive values and all projections that are positive. When the model assumes more false positives, the accuracy value drops. As in Eq. 9.

$$Precision = \frac{T_p}{T_p + F_p} \quad \textbf{(9)}$$

4.3.2.3. SENSITIVITY

The percentage of positive diagnostics that were diagnosed as positive. As in Eq. 10.

$$Sensitivity = \frac{T_p}{T_p + F_n} \tag{10}$$

4.3.2.4. F1-SCORE

The F1-Score runs from 0 to 1, and it is a harmonic recall and mean of precision. Low false negative and false positive readings produce this metric's higher value. As in Eq. 11.

$$F1 - score = \ 2 * \frac{Recall * Precision}{Recall + Precision} \tag{11}$$

4.3.3. ROOT MEAN SQUARE ERROR (RMSE)

It is a standard method for measuring model error. It is also called Loss Function. Know its equation as follows:

$$RMSE = \sqrt{\sum_{i=1}^{N} (y_o - y_p)^2 \Big/ N} \tag{12}$$

The value of this equation tells us the distance difference between the vector of expected values and the vector of observed values. In data science, this formula is used to evaluate trained models. It gives us the error rate between the training results and original results.

4.3.4. HYPERPARAMETER

The proposed model contains four layers: first is input layer, and then two hidden layers and the last one is output layer. It contains 11, 10, 10,1 neurons per layer, respectively. For the hidden layers, the activation function of the RELU Function has been set as in Eq. 13. and for the output layer, Sigmoid Function has been set as in Eq. 14. As for the parameters of aABC algorithm, they are as follows:

$$Epoch = 400$$

$$food = 50$$

$$Limit[i] = \frac{fitness[i] * food * D}{\sum fitness[i]}$$

In order to avoid falling into the problem of (Vanishing gradient), the RELU function was used for the hidden layers. It is the basic way to solve this problem. Whereas, to solve (Overfitting) K-Folds technique was relied in training the model.

$$y_j = f_j(x) = \max(0, x) \quad \textbf{(13)}$$

$$y_j = f_j(x) = \frac{1}{1 + e^{-x}} \quad \textbf{(14)}$$

4.4. RESULT AND DISCUSSON

The proposal model is tested with a sample dataset consisting of 72 values taken from the dataset, it is a mixture of monkeypox cases collected at thebmj center. The sample data consisted only of cases of monkeypox with different symptoms. It also displays the number of positive cases and the number of negative cases with different symptoms, thus proposing ANN model diagnostic method based on aABC algorithm. aABC is one of the evolutionary algorithms that contribute to the training of neural network weights.

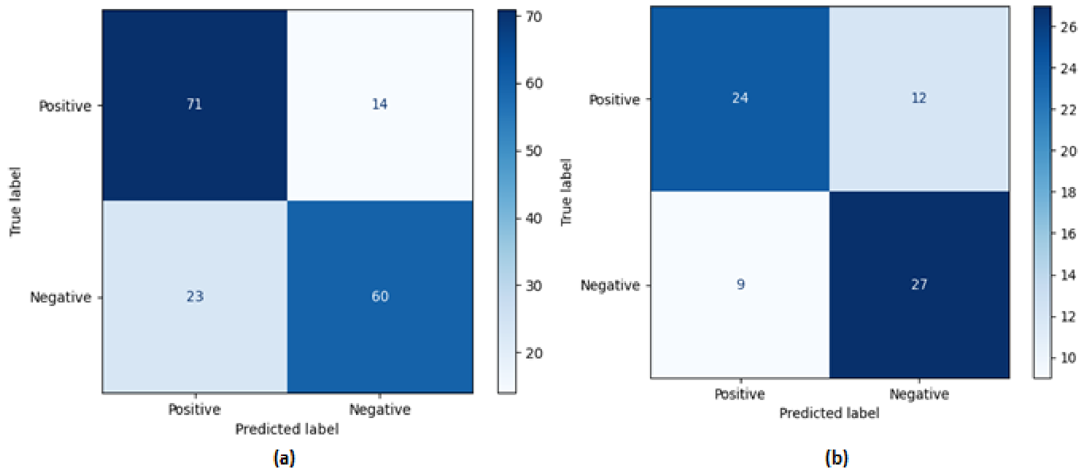


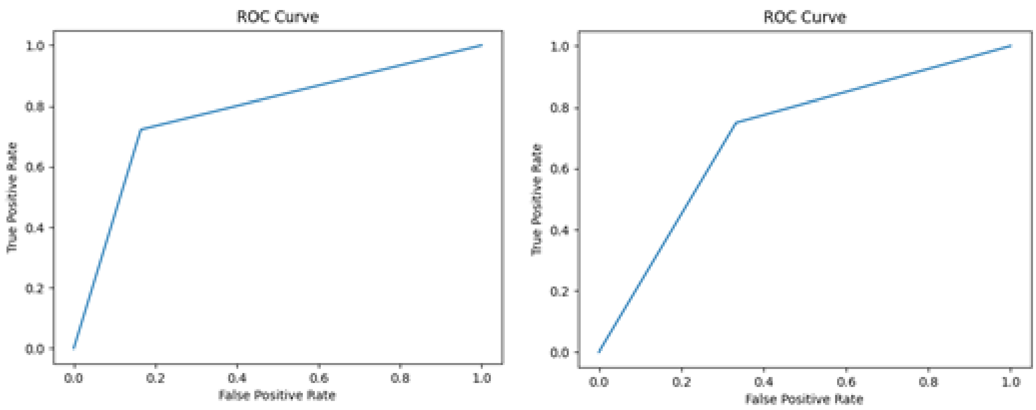
Figure. 13. (a) Confusion matrix of training phase for the proposed model. (b) Confusion matrix for testing.

Table. 4. Shows us the results of the proposed model during training and testing. The performance of the model was measured by several criteria. They appear as follows: in the training period accuracy, F1-score, precision and sensitivity take the values 78%, 76%, 81%, 84%. Respectively. When testing, accuracy, F1-score, precision, and sensitivity standards were taken as 71%, 72%, 69%, 67%. Respectively. Figure. 13. presents the confusion matrix of the proposed model during training and testing.

Table. 5. Performance of proposed model.

phase	Performance			
	Accuracy	F1-Score	Precision	Sensitivity
Training	78 %	76 %	81 %	84 %
Testing	71 %	72 %	69 %	67 %

The performance of ten models of ML and Deep Learning algorithms and the proposed model are summarized in Table 5. All ten models were trained on the same dataset. The training and validation process for all algorithms was repeated 30 times, and the accuracy of each stage was recorded separately.



(a) (b)

Figure. 14. (a) represents ROC curve during training phase. (b) ROC during testing phase.

We also used another method to measure model performance called the AUC-ROC curve, (as shown in Figure. 14.). It is one of the evaluation tools approved in the classification. This graph measures the performance of the classification model when there are only two classes, a positive class and a negative class. This curve indicates a ROC curve that plots the False Positive Rate (FPR) on the horizontal axis (X) against the True Positive Rate (TPR) on the vertical axis (Y).

AUC-ROC is the area under this graph. If the area under the curve is large, this indicates good performance of the model in distinguishing between positive and negative classes. Therefore, we notice that the threshold value is close to 1, and therefore the area covered by this curve is larger. Therefore, the model has good performance.

This curve is found based on the confusion matrix values mentioned previously. Emphasis is also placed on positive and negative values in the model's performance. Figure (a) represents a curve for the results of the training period, while Figure (b) represents the results for the validation period.

When the model's performance improves, the vertical axis values become close to 1, while the horizontal axis values become close to 0. For each of the two figures listed above, A and B.

Table 6. summarizes the performance of ten different models over 30 runs.

Trials	Methods									
	ANN Training by aABC	ANN Training by ABC	Deep Learning	SVC	KNN	Random Forest	Bagging Classifier	Decision Tree	Gradient Boosting	Naïve Bayes
1	69.0%	64.0%	72.9%	45.8%	65.1%	71.1%	59.5%	62.8%	63.8%	65.2%
2	67.0%	62.0%	68.8%	45.8%	65.1%	71.1%	66.6%	62.8%	63.8%	65.2%
3	64.0%	62.0%	64.6%	45.8%	65.1%	71.1%	60.9%	62.8%	63.8%	65.2%
4	58.0%	61.0%	68.8%	45.8%	65.1%	71.1%	60.5%	62.8%	63.8%	65.2%
5	67.0%	57.0%	62.5%	45.8%	65.1%	71.1%	52.1%	62.8%	63.8%	65.2%
6	58.0%	58.0%	60.4%	45.8%	65.1%	71.1%	45.5%	62.8%	63.8%	65.2%
7	54.0%	67.0%	72.9%	45.8%	65.1%	71.1%	62.5%	62.8%	63.8%	65.2%
8	57.0%	56.0%	68.8%	45.8%	65.1%	71.1%	66.6%	62.8%	63.8%	65.2%
9	65.0%	57.0%	66.6%	45.8%	65.1%	71.1%	62.5%	62.8%	63.8%	65.2%
10	67.0%	65.0%	70.8%	45.8%	65.1%	71.1%	60.8%	62.8%	63.8%	65.2%
11	58.0%	61.0%	70.8%	45.8%	65.1%	71.1%	61.5%	62.8%	63.8%	65.2%
12	69.0%	65.0%	68.8%	45.8%	65.1%	71.1%	50.0%	62.8%	63.8%	65.2%
13	64.0%	65.0%	66.6%	45.8%	65.1%	71.1%	52.4%	62.8%	63.8%	65.2%
14	61.0%	60.0%	66.6%	45.8%	65.1%	71.1%	65.2%	62.8%	63.8%	65.2%
15	60.0%	53.0%	64.6%	45.8%	65.1%	71.1%	68.2%	62.8%	63.8%	65.2%
16	62.0%	67.0%	64.6%	45.8%	65.1%	71.1%	65.1%	62.8%	63.8%	65.2%
17	71.0%	57.0%	68.8%	45.8%	65.1%	71.1%	62.5%	62.8%	63.8%	65.2%
18	61.0%	60.0%	70.8%	45.8%	65.1%	71.1%	69.8%	62.8%	63.8%	65.2%
19	67.0%	61.0%	66.6%	45.8%	65.1%	71.1%	50.0%	62.8%	63.8%	65.2%
20	61.0%	54.0%	64.6%	45.8%	65.1%	71.1%	63.6%	62.8%	63.8%	65.2%
21	61.0%	65.0%	70.8%	45.8%	65.1%	71.1%	53.3%	62.8%	63.8%	65.2%
22	64.0%	64.0%	66.6%	45.8%	65.1%	71.1%	57.8%	62.8%	63.8%	65.2%
23	58.0%	62.0%	72.9%	45.8%	65.1%	71.1%	51.2%	62.8%	63.8%	65.2%
24	67.0%	57.0%	68.8%	45.8%	65.1%	71.1%	60.9%	62.8%	63.8%	65.2%
25	62.0%	65.0%	68.8%	45.8%	65.1%	71.1%	57.8%	62.8%	63.8%	65.2%
26	64.0%	56.0%	75.0%	45.8%	65.1%	71.1%	55.3%	62.8%	63.8%	65.2%
27	61.0%	64.0%	68.8%	45.8%	65.1%	71.1%	56.6%	62.8%	63.8%	65.2%
28	64.0%	58.0%	58.3%	45.8%	65.1%	71.1%	58.5%	62.8%	63.8%	65.2%
29	60.0%	68.0%	66.6%	45.8%	65.1%	71.1%	57.8%	62.8%	63.8%	65.2%
30	64.0%	58.0%	70.8%	45.8%	65.1%	71.1%	58.8%	62.8%	63.8%	65.2%
Best	**71.0%**	**68.0%**	**75.0%**	**45.8%**	**65.1%**	**71.1%**	**69.8%**	**62.8%**	**63.8%**	**65.2%**
Worst	**54.0%**	**53.0%**	**58.3%**	**45.8%**	**65.1%**	**71.1%**	**45.5%**	**62.8%**	**63.8%**	**65.2%**
Mean	**62.8%**	**61.0%**	**67.9%**	**45.8%**	**65.1%**	**71.1%**	**59.1%**	**62.8%**	**63.8%**	**65.2%**
SD	**0.040**	**0.040**	**0.037**	**0.000**	**0.000**	**0.000**	**0.058**	**0.000**	**0.000**	**0.000**

The ANN deep learning model showed the best performance among all the other algorithms. The model got an accuracy of 75%. The results of the Random Forest algorithm were the second-best performers, achieving about 71.1%. While the model proposed in this paper showed the third best performance. It achieved an accuracy of 71%. However, the performance of the remaining seven models was less and unsatisfactory than SVC, as shown in Table. 5. For other statistical measures such as: F1 score, sensitivity, and Precision, their range of values did not differ much from the accuracy measure.

The performance of our custom architectural model trained from scratch using aABC algorithm can be compared to other ML models. In fact, the proposed model could not outperform Deep Learning or even Random Forest. While it was able to show better results than the rest. It is also noted from Table 5 that the proposed model with Deep Learning, Bagging Classifier changes the accuracy of each model every time it is tested. While the models SVC, KNN, Random Forest, Naïve Bayes, Decision Tree maintained one result throughout the testing period. That is, algorithms that rely on a statistical principle are more stable than learning algorithms, although they did not outperform learning algorithms in their best performance. Which is confirmed by the statistical measures in Table. 5.

In this Study, we relied on collected dataset from people with the disease and suspects. They were collected at thebmj center. Our study differs from previous studies in this respect. This appears clearly if we look at the difference between the results obtained in this study and the results of previous research. Our goal was to obtain an early detection model of the disease using the clinical symptoms that appear on the person when suspected.

PART 5

CONCLUSION

The Study provides a brief summary of the emergence of monkeypox virus, a zoonotic disease transmitted from animals to humans. This virus belongs to the highly virulent Orthopoxvirus family. The spread of this disease in societies alarms many people. Therefore, society needs an automated system for early detection that helps detect infection with this disease, if it occurs. Early prediction can prevent complications for people with the disease and save human lives.

This study aims to provide a model for distinguishing monkeypox infection by the clinical symptoms associated with the disease that appear on the infected person. The proposed model is a combination of the aABC algorithm and ANN. A comparison was made between the results obtained with several ML methods trained on the same dataset. The ANN deep learning model achieved the best performance with an accuracy of 75%, while the proposed model obtained an accuracy of 71%.

Since the proposed model is supported by several published literatures that use an AI-based diagnostic model, we hope that this article will contribute to future researchers and practitioners benefiting from the presented approach to develop a diagnostic mechanism for monkeypox disease.

REFERENCES

1. Singh, S., Rao, A., Kumar, K., Mishra, A., and Prajapati, V. K., "Translational vaccinomics and structural filtration algorithm to device multiepitope vaccine for catastrophic monkeypox virus", ***Computers In Biology And Medicine***, 153: (2023).

2. Shahyeez Ahamed, B. S. H., Usha, R., and Sreenivasulu, G., "A Deep Learning-based Methodology for Predicting Monkey Pox from Skin Sores", (2022).

3. Rimmer, S., Barnacle, J., Gibani, M. M., Wu, M. S., Dissanayake, O., Mehta, R., Herdman, T., Gilchrist, M., Muir, D., Ebrahimsa, U., Mora-Peris, B., Dosekun, O., Garvey, L., Peters, J., Davies, F., Cooke, G., and Abbara, A., "The clinical presentation of monkeypox: a retrospective case-control study of patients with possible or probable monkeypox in a West London cohort", ***International Journal Of Infectious Diseases***, 126: 48–53 (2023).

4. Yinka-Ogunleye, A., Aruna, O., Dalhat, M., Ogoina, D., McCollum, A., Disu, Y., Mamadu, I., Akinpelu, A., Ahmad, A., Burga, J., Ndoreraho, A., Nkunzimana, E., Manneh, L., Mohammed, A., Adeoye, O., Tom-Aba, D., Silenou, B., Ipadeola, O., Saleh, M., Adeyemo, A., Nwadiutor, I., Aworabhi, N., Uke, P., John, D., Wakama, P., Reynolds, M., Mauldin, M. R., Doty, J., Wilkins, K., Musa, J., Khalakdina, A., Adedeji, A., Mba, N., Ojo, O., Krause, G., Ihekweazu, C., Mandra, A., Davidson, W., Olson, V., Li, Y., Radford, K., Zhao, H., Townsend, M., Burgado, J., and Satheshkumar, P. S., "Outbreak of human monkeypox in Nigeria in 2017–18: a clinical and epidemiological report", ***The Lancet Infectious Diseases***, 19 (8): 872–879 (2019).

5. Rodríguez, B. S., Guzmán Herrador, B. R., Franco, A. D., Sánchez-Seco Fariñas, M. P., del Amo Valero, J., Aginagalde Llorente, A. H., Pérez de Agreda, J. P. A., Malonda, R. C., Castrillejo, D., Chirlaque López, M. D., Chong Chong, E. J., Balbuena, S. F., García, V. G., García-Cenoz, M., Hernández, L. G., Montalbán, E. G., Carril, F. G., Cortijo, T. G., Bueno, S. J., Sánchez, A. L., Linares Dópido, J. A., Lorusso, N., Martins, M. M., Martínez Ochoa, E. M., Mateo, A. M., Peña, J. M., Negredo Antón, A. I., Otero Barrós, M. T., del Carmen Pacheco Martinez, M., Jiménez, P. P., Pérez Martín, O. G., Rivas Pérez, A. I., García, M. S., Soria, F. S., Sierra Moros, M. J., Brandini Romersi, A. M., Lozano, C. G., Vallejo-Plaza, A., Campelli, G. S., Balader, P. S., San Miguel, L. G., Cano, E. A., Ruiz-Algueró, M., Simón, L., Arias, P., Vázquez, A., Sánchez, P., Herrero, L., Molero, F., Torres, M., Sánchez, L., Cejudo, C., Polo, R., Castellá, J. G., Koerting, A., Vazquez Rincon, I. M., Ugarriza, A. V., Remacha, C. M., Boone, A. L., Huerta, M. H., Riutort, A. N., Torres Lana, Á. L., Herrera, A. A., García, I. F., Guijarro, M. V., Duran, G. R., Marín, V. R., Saavedra, M. R. R., Arribas, S. F., Rodríguez, H. M., Calvo, N. R., Rio, V. A.,

Garzón, N. G., Martínez-Pino, I., Jesús Rodríguez Recio, M., Roig Sena, F. J., Pública, S., del Mar López-Tercero Torvisco, M., del Carmen García Bañobre, M., del Pilar Sánchez Castro, M., Martínez Soto, M. R., García, M. A., Martínez, F. M., Jose Domínguez Rodríguez, M., Morales, L. M., Navarro, A. H., Deorador, E. C., Forte, A. N., Julia, A. N., Largo, N. C., Ortíz, C. S., Marín, N. G., Díaz, J. S., Belen, M., Zarzuelo, R., Pariente, N. M., del Buey, J. F. B., Velasco Rodríguez, M. J., Peña, A. A., Baena, E. R., Benito, A. M., Meixeira, A. P., Martinez, J. I., Ordobas, M., Arce, A., Naranjo, A. S. M., Castilla, J., Casado, I., Burgui, C., Egües, N., Ezpeleta, G., Carollo, O. M., Larraitz, A., Arricibita, E. H., and Ibáñez Pérez, A. C., "Epidemiologic Features and Control Measures during Monkeypox Outbreak, Spain, June 2022", ***Emerging Infectious Diseases***, 28 (9): 1847–1851 (2022).

6. Kannan, S. R., Sachdev, S., Reddy, A. S., Kandasamy, S. L., Byrareddy, S. N., Lorson, C. L., and Singh, K., "Mutations in the monkeypox virus replication complex: Potential contributing factors to the 2022 outbreak", ***Journal Of Autoimmunity***, 133: (2022).

7. Thornhill, J. P., Barkati, S., Walmsley, S., Rockstroh, J., Antinori, A., Harrison, L. B., Palich, R., Nori, A., Reeves, I., Habibi, M. S., Apea, V., Boesecke, C., Vandekerckhove, L., Yakubovsky, M., Sendagorta, E., Blanco, J. L., Florence, E., Moschese, D., Maltez, F. M., Goorhuis, A., Pourcher, V., Migaud, P., Noe, S., Pintado, C., Maggi, F., Hansen, A.-B. E., Hoffmann, C., Lezama, J. I., Mussini, C., Cattelan, A., Makofane, K., Tan, D., Nozza, S., Nemeth, J., Klein, M. B., and Orkin, C. M., "Monkeypox Virus Infection in Humans across 16 Countries — April–June 2022", ***New England Journal Of Medicine***, 387 (8): 679–691 (2022).

8. Silva, M. S. T., Santos, D. G. dos, Coutinho, C., Ribeiro, M. P. D., Cardoso, S. W., Veloso, V. G., and Grinsztejn, B., "The first case of acute HIV and monkeypox coinfection in Latin America", ***The Brazilian Journal Of Infectious Diseases***, 102736 (2022).

9. "Monkeypox Skin Lesion Dataset | Kaggle", **https://www.kaggle.com/datasets/nafin59/monkeypox-skin-lesion-dataset** (2023).

10. Dwivedi, M., Tiwari, R. G., and Ujjwal, N., "Deep Learning Methods for Early Detection of Monkeypox Skin Lesion", (2023).

11. Patel, A., Bilinska, J., Tam, J. C. H., Da Silva Fontoura, D., Mason, C. Y., Daunt, A., Snell, L. B., Murphy, J., Potter, J., Tuudah, C., Sundramoorthi, R., Abeywickrema, M., Pley, C., Naidu, V., Nebbia, G., Aarons, E., Botgros, A., Douthwaite, S. T., Van Nispen Tot Pannerden, C., Winslow, H., Brown, A., Chilton, D., and Nori, A., "Clinical features and novel presentations of human monkeypox in a central London centre during the 2022 outbreak: Descriptive case series", ***The BMJ***, (2022).

12. Irmak, M. C., Aydın, T., and Yağanoğlu, M., "Monkeypox Skin Lesion Detection with MobileNetV2 and VGGNet Models", (2022).

13. de la Calle-Prieto, F., Estébanez Muñoz, M., Ramírez, G., Díaz-Menéndez, M., Velasco, M., Azkune Galparsoro, H., Salavert Lletí, M., Mata Forte, T., Blanco, J. L., Mora-Rillo, M., Arsuaga, M., de Miguel Buckley, R., Arribas, J. R., and Membrillo, F. J., .

14. Matuszewski, D. J. and Sintorn, I. M., "TEM virus images: Benchmark dataset and deep learning classification", ***Computer Methods And Programs In Biomedicine***, 209: (2021).

15. Ahsan, M. M., Uddin, M. R., Ali, M. S., Islam, M. K., Farjana, M., Sakib, A. N., Momin, K. Al, and Luna, S. A., "Deep transfer learning approaches for Monkeypox disease diagnosis", ***Expert Systems With Applications***, 216: (2023).

16. Saleh, A. I. and Rabie, A. H., "Human monkeypox diagnose (HMD) strategy based on data mining and artificial intelligence techniques", ***Computers In Biology And Medicine***, 152: (2023).

17. Nhat-Duc, H. and Van-Duc, T., "Comparison of histogram-based gradient boosting classification machine, random Forest, and deep convolutional neural network for pavement raveling severity classification", ***Automation In Construction***, 148: 104767 (2023).

18. Rimmer, S., Barnacle, J., Gibani, M. M., Wu, M. S., Dissanayake, O., Mehta, R., Herdman, T., Gilchrist, M., Muir, D., Ebrahimsa, U., Mora-Peris, B., Dosekun, O., Garvey, L., Peters, J., Davies, F., Cooke, G., and Abbara, A., "The clinical presentation of monkeypox: a retrospective case-control study of patients with possible or probable monkeypox in a West London cohort", ***International Journal Of Infectious Diseases***, 126: 48–53 (2023).

19. Liu, L., Lei, J., Liu, T., Shang, B., Zhao, X., He, J., Pan, F., and Li, L., "Artificial skin with fast and robust self-healing ability for durable intelligent protection of magnesium alloys", ***Chemical Engineering Journal***, 458: (2023).

20. Zhao, H., "Intelligent management of industrial building energy saving based on artificial intelligence", ***Sustainable Energy Technologies And Assessments***, 56: (2023).

21. Zhu, H., "Research on intelligent analysis strategies to improve athletes' psychological experience in the era of artificial intelligence", ***Progress In Neuro-Psychopharmacology And Biological Psychiatry***, 119: (2022).

22. Lin, B., Tan, Z., Mo, Y., Yang, X., Liu, Y., and Xu, B., "Intelligent oncology: The convergence of artificial intelligence and oncology", ***Journal Of The National Cancer Center***, (2022).

23. Li, Y., "Application Analysis of Artificial Intelligent Neural Network Based on Intelligent Diagnosis", (2022).

24. Gao, X., Liu, T., Tan, L., and Song, S., "Multioperator search strategy for evolutionary multiobjective optimization", ***Swarm And Evolutionary Computation***, 71: (2022).

25. Hu, Z., Zhou, T., Su, Q., and Liu, M., "A niching backtracking search algorithm with adaptive local search for multimodal multiobjective optimization", ***Swarm And Evolutionary Computation***, 69: (2022).

26. Najari, S., Gróf, G., Saeidi, S., and Gallucci, F., "Modeling and optimization of hydrogenation of CO 2 : Estimation of kinetic parameters via Artificial Bee Colony (ABC) and Differential Evolution (DE) algorithms", ***International Journal Of Hydrogen Energy***, 44 (10): 4630–4649 (2019).

27. Ponsich, A. and Coello Coello, C. A., "A hybrid Differential Evolution - Tabu Search algorithm for the solution of Job-Shop Scheduling Problems", ***Applied Soft Computing Journal***, 13 (1): 462–474 (2013).

28. Pierre, R. da S. de Q., "Heuristics in Design: A Literature Review", ***Procedia Manufacturing***, 3: 6571–6578 (2015).

29. Kabalci, Y., Kockanat, S., and Kabalci, E., "A modified ABC algorithm approach for power system harmonic estimation problems", ***Electric Power Systems Research***, 154: 160–173 (2018).

30. Al-Ammar, E. A., Farzana, K., Waqar, A., Aamir, M., Saifullah, Ul Haq, A., Zahid, M., and Batool, M., "ABC algorithm based optimal sizing and placement of DGs in distribution networks considering multiple objectives", ***Ain Shams Engineering Journal***, 12 (1): 697–708 (2021).

31. Naidu, K., Mokhlis, H., Bakar, A. H. A., and Terzija, V., "Performance investigation of ABC algorithm in multi-area power system with multiple interconnected generators", ***Applied Soft Computing Journal***, 57: 436–451 (2017).

32. Hancer, E., Xue, B., Karaboga, D., and Zhang, M., "A binary ABC algorithm based on advanced similarity scheme for feature selection", ***Applied Soft Computing Journal***, 36: 334–348 (2015).

33. Singh, S., Chauhan, P., and Singh, N. J., "Capacity optimization of grid connected solar/fuel cell energy system using hybrid ABC-PSO algorithm", ***International Journal Of Hydrogen Energy***, 45 (16): 10070–10088 (2020).

34. Najari, S., Gróf, G., Saeidi, S., and Gallucci, F., "Modeling and optimization of hydrogenation of CO 2 : Estimation of kinetic parameters via Artificial Bee Colony (ABC) and Differential Evolution (DE) algorithms", ***International Journal Of Hydrogen Energy***, 44 (10): 4630–4649 (2019).

35. Sagayam, K. M. and Hemanth, D. J., "ABC algorithm based optimization of 1-D hidden Markov model for hand gesture recognition applications", ***Computers In Industry***, 99: 313–323 (2018).

36. Soori, M., Arezoo, B., and Dastres, R., "Machine learning and artificial intelligence in CNC machine tools, A review", ***Sustainable Manufacturing And Service Economics***, 100009 (2023).

37. Li, J. Bin, Chen, Z. Y., Li, X., Jing, L. J., Zhangf, Y. P., Xiao, H. H., Wang, S. J., Yang, W. K., Wu, L. J., Li, P. Y., Li, H. B., Yao, M., and Fan, L. T., .

38. Sayed, Y. A. K., Ibrahim, A. A., Tamrazyan, A. G., and Fahmy, M. F. M., "Machine-learning-based models versus design-oriented models for predicting the axial compressive load of FRP-confined rectangular RC columns", ***Engineering Structures***, 285: (2023).

39. Chen, Z. S., Lam, J. S. L., and Xiao, Z., "Prediction of harbour vessel fuel consumption based on machine learning approach", ***Ocean Engineering***, 278: (2023).

40. Bu, Y., Shi, L., and Ma, B., "Method for establishing predictive models for total organic halogen based on piecewise interpolation and machine learning", ***Journal Of Environmental Chemical Engineering***, 11 (3): 109928 (2023).

41. Birfir, S., Elalouf, A., and Rosenbloom, T., "Building machine-learning models for reducing the severity of bicyclist road traffic injuries", ***Transportation Engineering***, 12: 100179 (2023).

42. Mohammedqasem, R., Mohammedqasim, H., Asad Ali Biabani, S., Ata, O., Alomary, M. N., Almehmadi, M., Amer Alsairi, A., and Azam Ansari, M., "Multi-objective deep learning framework for COVID-19 dataset problems", ***Journal Of King Saud University - Science***, 35 (3): (2023).

43. Lee, V. H., Hew, J. J., Leong, L. Y., Tan, G. W. H., and Ooi, K. B., "Wearable payment: A deep learning-based dual-stage SEM-ANN analysis", ***Expert Systems With Applications***, 157: (2020).

44. Kolbadinejad, S., Mashhadimoslem, H., Ghaemi, A., and Bastos-Neto, M., "Deep learning analysis of Ar, Xe, Kr, and O2 adsorption on Activated Carbon and Zeolites using ANN approach", ***Chemical Engineering And Processing - Process Intensification***, 170: (2022).

45. Pham, B. T., Nguyen, M. D., Dao, D. Van, Prakash, I., Ly, H. B., Le, T. T., Ho, L. S., Nguyen, K. T., Ngo, T. Q., Hoang, V., Son, L. H., Ngo, H. T. T., Tran, H. T., Do, N. M., Van Le, H., Ho, H. L., and Tien Bui, D., "Development of artificial intelligence models for the prediction of Compression Coefficient of soil: An application of Monte Carlo sensitivity analysis", ***Science Of The Total Environment***, 679: 172–184 (2019).

46. Yadav, R. K. and Anubhav, "PSO-GA based hybrid with Adam Optimization for ANN training with application in Medical Diagnosis", ***Cognitive Systems Research***, 64: 191–199 (2020).

47. Chen, S., Ren, Y., Friedrich, D., Yu, Z., and Yu, J., "Sensitivity analysis to reduce duplicated features in ANN training for district heat demand prediction", ***Energy And AI***, 2: (2020).

48. Veza, I., Irianto, Panchal, H., Paristiawan, P. A., Idris, M., Fattah, I. M. R., Putra, N. R., and Silambarasan, R., "Improved prediction accuracy of biomass heating value using proximate analysis with various ANN training algorithms", ***Results In Engineering***, 16: (2022).

49. Özdoğan, H., Üncü, Y. A., Şekerci, M., and Kaplan, A., "Estimations for (n,α) reaction cross sections at around 14.5MeV using Levenberg-Marquardt algorithm-based artificial neural network", ***Applied Radiation And Isotopes***, 192: (2023).

50. Özdoğan, H., Üncü, Y. A., Şekerci, M., and Kaplan, A., "Estimations for (n,α) reaction cross sections at around 14.5MeV using Levenberg-Marquardt algorithm-based artificial neural network", ***Applied Radiation And Isotopes***, 192: (2023).

51. de Oliveira, F. R. and de Oliveira, F. R., "A locally convergent inexact projected Levenberg–Marquardt-type algorithm for large-scale constrained nonsmooth equations", ***Journal Of Computational And Applied Mathematics***, 427: (2023).

52. Žic, M. and Pereverzyev, S., "Application of self-adapting regularization, machine learning tools and limits in Levenberg–Marquardt algorithm to solve CNLS problem", ***Journal Of Electroanalytical Chemistry***, 939: 117420 (2023).

53. de Oliveira, F. R. and de Oliveira, F. R., "A locally convergent inexact projected Levenberg–Marquardt-type algorithm for large-scale constrained nonsmooth equations", ***Journal Of Computational And Applied Mathematics***, 427: (2023).

54. Wang, S., Ren, J., and Bai, R., "A semi-supervised adaptive discriminative discretization method improving discrimination power of regularized naive Bayes", ***Expert Systems With Applications***, 225: (2023).

55. Zhao, X. and Xia, Z., "Secure outsourced NB: Accurate and efficient privacy-preserving Naive Bayes classification", ***Computers And Security***, 124: (2023).

56. Vishwakarma, M. and Kesswani, N., "A new two-phase intrusion detection system with Naïve Bayes machine learning for data classification and elliptic envelop method for anomaly detection", ***Decision Analytics Journal***, 7: (2023).

57. García-Laencina, P. J., Sancho-Gómez, J. L., Figueiras-Vidal, A. R., and Verleysen, M., "K nearest neighbours with mutual information for simultaneous classification and missing data imputation", ***Neurocomputing***, 72 (7–9): 1483–1493 (2009).

58. Mateos-García, D., García-Gutiérrez, J., and Riquelme-Santos, J. C., "An evolutionary voting for k-nearest neighbours", ***Expert Systems With Applications***, 43: 9–14 (2016).

59. Faziludeen, S. and Sankaran, P., "ECG Beat Classification Using Evidential K -Nearest Neighbours", (2016).

60. Mic6, L., Oncina, J., and Carrasco, R. C., "Pattern, Recognition Le~ers A fast branch & bound nearest neighbour classifier in metric spaces", (1996).

61. Yang, B., Tang, F., and Yuan, X., "Influence evaluation of signal transmission delay on supplementary damping controller for SVC in power system", ***Energy Reports***, 8: 19–30 (2022).

62. Liu, T., Jin, L., Zhong, C., and Xue, F., "Study of thermal sensation prediction model based on support vector classification (SVC) algorithm with data preprocessing", ***Journal Of Building Engineering***, 48: (2022).

63. Loger, B., Dolgui, A., Lehuédé, F., and Massonnet, G., "Improving the Tractability of SVC-based Robust Optimization", (2022).

64. Ruoyuan, Z. and Ma, R., "Non-invasive load identification method based on ABC-SVM algorithm and transient feature", ***Energy Reports***, 8: 63–72 (2022).

65. Li, X., Liu, M., Wang, K., Liu, Z., and Li, G., "Data cleaning method for the process of acid production with flue gas based on improved random forest", ***Chinese Journal Of Chemical Engineering***, (2023).

66. Gao, W., Xu, F., and Zhou, Z. H., "Towards convergence rate analysis of random forests for classification", ***Artificial Intelligence***, 313: (2022).

67. Cao, M., Yin, D., Zhong, Y., Lv, Y., and Lu, L., "Detection of geochemical anomalies related to mineralization using the Random Forest model optimized by the Competitive Mechanism and Beetle Antennae Search", ***Journal Of Geochemical Exploration***, 249: 107195 (2023).

68. Theerthagiri, P., "Predictive analysis of cardiovascular disease using gradient boosting based learning and recursive feature elimination technique", ***Intelligent Systems With Applications***, 16: (2022).

69. Li, Y., Feng, Y., and Qian, Q., "FDPBoost: Federated differential privacy gradient boosting decision trees", ***Journal Of Information Security And Applications***, 74: 103468 (2023).

70. Hosney, H., Tawfik, M. H., Duker, A., and van der Steen, P., "Prospects for treated wastewater reuse in agriculture in low- and middle-income countries: Systematic analysis and decision-making trees for diverse management approaches", ***Environmental Development***, 46: 100849 (2023).

71. Huyut, M. T. and Huyut, Z., "Effect of ferritin, INR, and D-dimer immunological parameters levels as predictors of COVID-19 mortality: A strong prediction with the decision trees", ***Heliyon***, 9 (3): e14015 (2023).

72. Han, X., Zhu, X., Pedrycz, W., and Li, Z., "A three-way classification with fuzzy decision trees", ***Applied Soft Computing***, 132: (2023).

73. Botha, D. and Steyn, M., "The use of decision tree analysis for improving age estimation standards from the acetabulum", ***Forensic Science International***, 341: (2022).

74. Farid, M., Palmblad, M., Hallman, H., and Vänngård, J., "A binary decision tree approach for pharmaceutical project portfolio management", ***Decision Analytics Journal***, 7: (2023).

75. Saygin, F., Şişman, Y., Dengiz, O., and Şişman, A., "Spatial assessment of landslide susceptibility mapping generated by fuzzy-AHP and decision tree approaches", ***Advances In Space Research***, (2023).

76. Wang, B., Wang, W., Meng, G., Meng, T., Song, B., Wang, Y., Guo, Y., Qiao, Z., and Mao, Z., "Selective Feature Bagging of one-class classifiers for novelty detection in high-dimensional data", ***Engineering Applications Of Artificial Intelligence***, 120: (2023).

77. Abdoli, M., Akbari, M., and Shahrabi, J., "Bagging Supervised Autoencoder Classifier for credit scoring", ***Expert Systems With Applications***, 213: (2023).

78. Subasi, A., Kadasa, B., and Kremic, E., "Classification of the cardiotocogram data for anticipation of fetal risks using bagging ensemble classifier", (2020).

79. Lin, J., Chen, H., Li, S., Liu, Y., Li, X., and Yu, B., "Accurate prediction of potential druggable proteins based on genetic algorithm and Bagging-SVM ensemble classifier", ***Artificial Intelligence In Medicine***, 98: 35–47 (2019).

80. Klikowski, J. and Woźniak, M., "Deterministic Sampling Classifier with weighted Bagging for drifted imbalanced data stream classification", ***Applied Soft Computing***, 122: (2022).

81. Yildizel, S. A. and Toktas, A., "ABC algorithm-based optimization and evaluation of nano carbon black added multi-layer microwave absorbing ultra weight foam concrete", ***Materials Today Communications***, 32: (2022).

82. Karaboga, D. and Akay, B., "A modified Artificial Bee Colony (ABC) algorithm for constrained optimization problems", ***Applied Soft Computing Journal***, 11 (3): 3021–3031 (2011).

83. Kapila, D. and Bhagat, N., "Efficient feature selection technique for brain tumor classification utilizing hybrid fruit fly based abc and ann algorithm", (2021).

84. "Monkey-Pox PATIENTS Dataset. | Kaggle", **https://www.kaggle.com/datasets/muhammad4hmed/monkeypox-patients-dataset** (2023).

85. Ali, S. N., Ahmed, Md. T., Paul, J., Jahan, T., Sani, S. M. S., Noor, N., and Hasan, T., "Monkeypox Skin Lesion Detection Using Deep Learning Models: A Feasibility Study", (2022).

RESUME

Ahmed MUHAMMED KALO HAMDAN graduated from elementary and high school education in the Syrian city of Aleppo. After that, in 2011, he started a bachelor's program in the Department of Informatics Engineering at the University of Aleppo. One year later, he was unable to complete his studies at the university due to the situation in the country. In 2016, he re-enrolled in the Bachelor's program in the Department of Informatics Engineering at Al-Sham Private University in northern Syria. He graduated first in the university. He moved to Turkey and began studying for a master's degree in the Department of Computer Engineering at Karabük University in 2020. His goal is to complete his doctoral studies, Insha'Allah.