Grenze International Journal of Engineering and Technology, Jan Issue

A Classification of COVID-19 Cases using Fine-Tune Different Model of Machine Learning

Kiran Dehariya[1], Dr. Rajesh Boghey[2] and Mrs Manjari Singh Rathore[3]
[1]M.Tech. Scholar, Department of CSE, TIT Excellence, Bhopal (M.P), India
[2]Professor, Department of CSE, TIT Excellence, Bhopal (M.P)
[3]Asst. Professor, Department of CSE, TIT Excellence, Bhopal (M.P)

***Abstract*—The Machine Learning Techniques are used for finding upcoming or futuristic Prediction. It's very powerful tools in finding of future prediction. From Last decade to recent days Data mining and Machine Learning play vital roles in many Industries. In last 2-3 Years overall World is suffering from a new Virus i.e., COVID-19. The first covid case was detected in Month of December. In very less time it spread out throughout the world. In 2-3 months, this illness is declared as Pandemic by WHO (World Health Organization). The behaviour of spreading of COVID-19 was very abnormal. it was growing Exponentially. Here In this Paper as a Researcher's we are trying to find out overall Analysis of COVID-19. Here our main motto to finding How many patients gets infected, how many of them get cure in sometimes & finally How many of them goes to death. Here our main concentration is to finding the features of the COVID. We will be applying Machine Learning Algorithms like LR, SVM, DT and many more. After applying the above methods, we will find the performance measures of each one and compare with their values. Finally, we will try to Implement some new Enhancement in Existing Algorithms.**

***Index Terms*— Keywords: Machine Learning, COVID-19, Training & Testing Data, Model or classifier, NB, SVM, DT, Performance Measures.**

I. INTRODUCTION

This is World Known facts that Viruses having very huge family that may causes different types of illness. COVID-19 has also different variants which causes different types of illness. In this illness patient may faces different symptom like common cold and throat infection [6]. If we are talking about the variation that is MERS & MERS-Cov. SARS-Cov is also a different variation of COVID-19. if we talked about initiation of above two diseases one is originated from China in 2002 i.e., SARS-Cov-2 and another one originated from Middle East with the name MERS in 2012 [1,8]. The very new and recent one was originated from city Wuhan which is situated in China. From here only this COVID-19 starting spreading into different Region of the World. If we talk about official claim then [5] this is claimed by China government and Reported to WHO to December 2019. In very less time this is spread to overall world and infection rate was very high. With infection death toll was also very high. As we all know this disease is spreading through touch so Here Artificial Intelligence Enable Devices i.e., Robotics play very important and Effective Role [13]. Electronics devices i.e. Robotics works upon Data gather and implement different Algorithms from Data mining and Machine Learning. In recent days many hospitals and organization start using the devices which play base upon ML Techniques [14]. In Health Care we

Grenze ID: 01.GIJET.9.1.503_4

are using these techniques in very effective manner. Here Researcher's used a dataset with record values 44,25,485 cases. out of them 3,02,059 went dead. and 203 countries get infected in very less time. Due to huge number of increments in cases all hospitals get over burden and this industry is going to collapse in very less time in throughout the world [43]. In even Developed countries they face the same problem. In developing countries like India and other faces more issues in this time. Due to collapse Machine like ventilators and different apparatus play very important roles in mechanical monitoring of patient.
Prediction means to analyse the future steps of any events depends upon the previous data and their behaviour. In this we have Different categories which is given below:
Predictive models explain for sure short relationships and some behavioural patterns that usually edge to a very certain behaviour.
Descriptive models explain by the creating partition or segment depends upon the given dataset.

A. Machine Learning Techniques used in covid-19 Analysis

Since we know that in recent days Machine Learning algorithm play important roles in different industries. In this section we worked for analysis or finding in covid-19 pataient and their behaviour. We know that we have number of algorithms to solve our problem out of them we are explain some algorithms in figure 1:

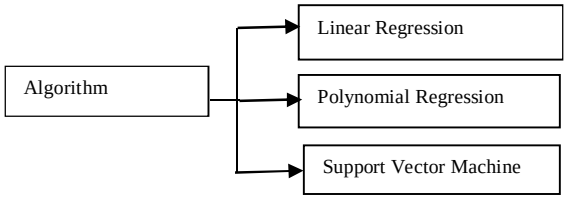


Figure 1: Used Algorithms in Sentiment Analysis

II. Related Work

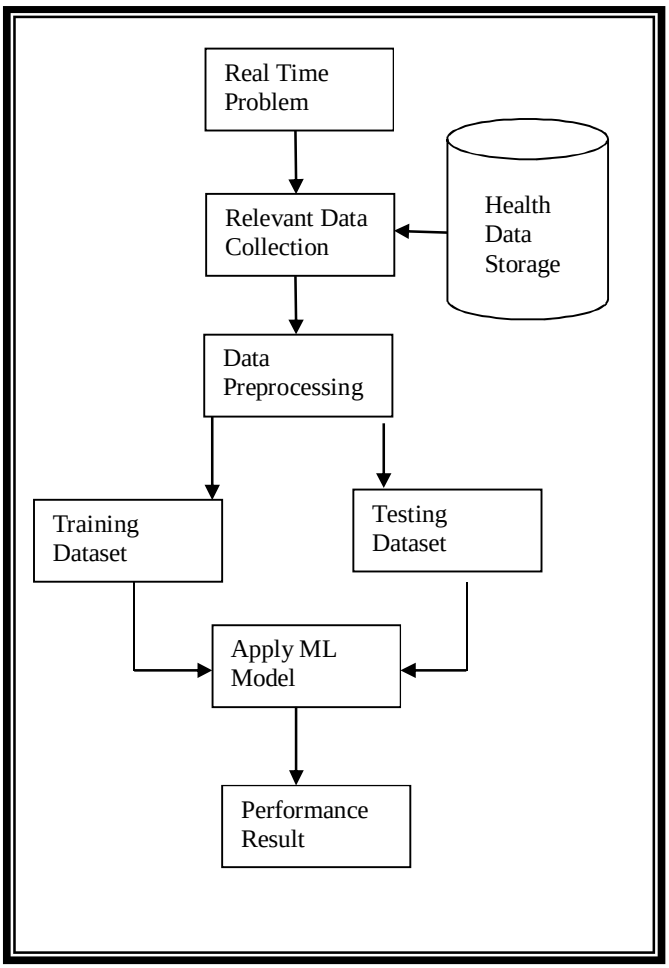
Here Authors explained their results based upon reflective analysis of various patients from Wuhan itself. The name of Hospital is Jinyintan Hospital. Here Author's explained the epidemiological data set. Here patients having different features or symptom gathered and taken all lab results as a features or attributes of the given dataset [16]. By the above research they came to conclusion the pattern of clinical research which will support for further analyses. Here Researcher's go into deep dive for analysing the demographic impact by these diseases. We all knows its behaviour varies from one continent to another continent. Firstly, all data taken by Wuhan only. but analysing more we need data from different Region & Continent [30]. Here Authors took different lab analysis like CT images. Here Machine learning play vital roles to detect any deep analysis of the given Image. By the research Author's want to contribute their research information to the society or doctors' community which will help to our people by developing medicine accordingly. In this Papers Author's Explained that how anyone create a such framework which will detect the Corona Virus using different mechanical devices like smartphone enabled sensors. Here AI comes into existence to find the result from various sensor to predict the symptom of Corona will detect automatically without touch of any human being. so spreading ratio may be goes deep down. and finally, we will get our result to further analysis [42].
In this section Author's describe very informative network-based model which is very informative for analysing the infected people. The name of different Network is (ResNet50, InceptionV3) [32].

III. Aim

The aim of this dissertation is to predict whether a person has COVID-19 or not, using machine learning techniques. The prediction is performed using the clinical information of the patients. The goal is to identify whether a patient can potentially be diagnosed with COVID-19.

IV. Objective

The main objective of our Work is, identifying the most suitable machine learning technique for prediction, to perform on clinical reports of patients. In this paper we are preparing a machine learning model that could make accurate predictions of COVID-19 in patients.



V. ALGORITHMS

Step 01: Store Data from Kaggle Repository
Step 02: Import Prior Libraries:
Step03: Now Import our Required Dataset

```
# Let's read the data set
data = pd.read_csv('covid-19.csv')
data. shape
```

Step04: Apply Feature Extraction
a) Bivariate Analysis
b) Multivariate Analysis
Step 05: Visualize Data for better understanding
Step06: Applying Machine Learning Algorithms
Step07: Apply Different Model
Step08: Repeat Step07 for many times with different Algorithms
Step09: Finally Compare Results with performance parameters like RMSE Score & R2 Score.

VI. FLOW DIAGRAM

In figure 3 At first step, we need to fetched Data from any external source or we can collect Data from Local Market but for better Analysis we are Fetching our Data from Kaggle. That is very reliable Data Source through Word Wide. In Next Step we need to Fetched Different Libraries for processing our Data. At very next Step that

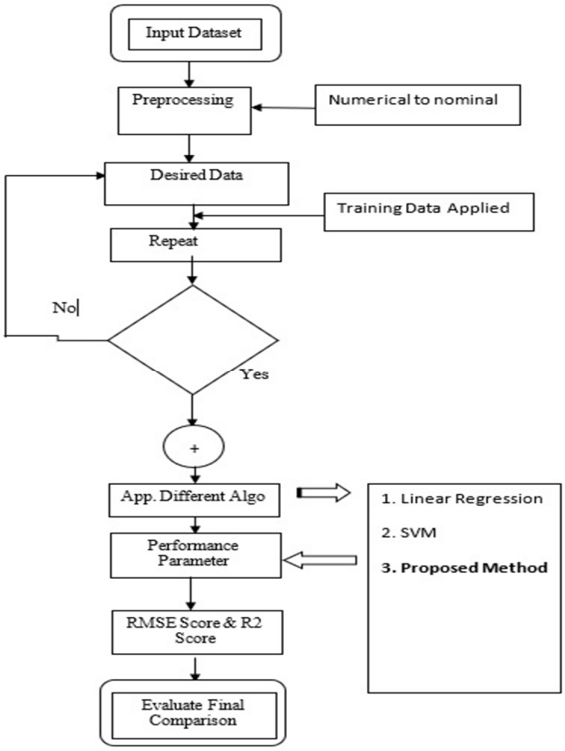


Figure 3: Flow Diagram

is Third Step we need to process our Data for next step Here we have many processing Mechanism. We are using Numerical to Nominal Data Conversion or also using Uni-Variant and Multi-Variant Data Processing. At Next Step i.e., Fourth Step we need to repeat it for different Data split. Then after we will reach at step 5 where we will apply Different Machine Learning Algorithms and Finally, we will apply our own Proposed Methods i.e., Gradient Boosting with Weighted Average values. Here we will adding average values of previous implemented mechanism. At Final Step i.e., Sith step we will have to find Performance Measures i.e., RMSE Score & R2 Score will give clear views of proposed method and Existing one. At Final Step we will compare these given Results. We can say that Our Proposed Methods gives better Result.

VII. Procedure for Implementation

The model employs filters for faster evaluation and lesser overall time. The pre-processing methods and application of filters affect a lot in final evaluation results of classifiers (ML based models). The feature extraction methods, conversion of nominal to binary and cleaning are few of those filters.

```
In [1]: # Lets import the required libraries

        # for mathematical operations
        import numpy as np
        # for dataframe manipulations
        import pandas as pd

        # for data visualizations
        import seaborn as sns
        import matplotlib.pyplot as plt
        import plotly.express as px

        # setting parameters for visualization
        plt.rcParams['figure.figsize'] = (16, 5)
        plt.style.use('fivethirtyeight')
```

Figure 4: Calling Libraries

In the figure 4 we called Libraries which will help you to call all functionality which required.
In the Figure 5 we try to show number of major columns available in our DataSet.we have 8 columns in our data set.
In the figure 6 we try to show Disease Spread ,world wide ,country wise and many more.

	SNo	ObservationDate	Province/State	Country/Region	Last Update	Confirmed	Deaths	Recovered
0	1	01/22/2020	Anhui	Mainland China	1/22/2020 17:00	1.0	0.0	0.0
1	2	01/22/2020	Beijing	Mainland China	1/22/2020 17:00	14.0	0.0	0.0
2	3	01/22/2020	Chongqing	Mainland China	1/22/2020 17:00	6.0	0.0	0.0
3	4	01/22/2020	Fujian	Mainland China	1/22/2020 17:00	1.0	0.0	0.0
4	5	01/22/2020	Gansu	Mainland China	1/22/2020 17:00	0.0	0.0	0.0

Figure 5: Major columns

```
Basic Information
Totol number of countries with Disease Spread:  226
Total number of Confirmed Cases around the World:  96158735.0
Total number of Recovered Cases around the World:  53035240.0
Total number of Deaths Cases around the World:  2056996.0
Total number of Active Cases around the World:  41066499.0
Total number of Closed Cases around the World:  55092236.0
Approximate number of Confirmed Cases per Day around the World:  264172.0
Approximate number of Recovered Cases per Day around the World:  145701.0
Approximate number of Death Cases per Day around the World:  5651.0
Approximate number of Confirmed Cases per hour around the World:  11007.0
Approximate number of Recovered Cases per hour around the World:  6071.0
Approximate number of Death Cases per hour around the World:  235.0
Number of Confirmed Cases in last 24 hours:  599088.0
Number of Recovered Cases in last 24 hours:  400355.0
Number of Death Cases in last 24 hours:  16889.0
```

Figure 6: columns Description

Figure 7: Growth of different types of cases

In the figure 7 we try to show Growth of different types of cases according to confirmed cases,Recovered cases and Death cases.

VIII. Comparison of Result

In figure 8 By the analysis of above graph we can say that we find that Liner Regression ,Polynomial Regression , SVM Prediction values gives Different Results respectively. Here We are going to proposed some tuning mechanism so that Results get improved.

TABLE I: COMPARISON OF ALGORITHMS

Dates	LR Pred.	PR Pred.	SVM Pred.
2021-01-20	64534315.76411	80161880.46839	124201136.41425
2021-01-21	64755138.31053	79272179.85058	126099178.56141
2021-01-22	64975960.85694	78281099.64779	128023400.13859
2021-01-23	65196783.40336	77183957.12743	129974089.22374
2021-01-24	65417605.94978	75975918.78977	131951536.26582

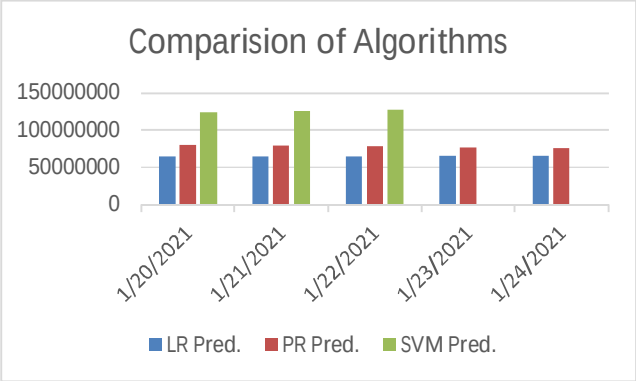


Figure 8: Comparison Analysis

IX. CONCLUSION

This We conclude that when we Implemented Number of Machine Learning Algorithms for finding best results in terms of performance. We Finds the major Features of given Data set are state, country, confirm cases & death cases. Most Important Factors to Predict the Case Distribution of a Patient. Finally, we Implemented Different Machine Learning Algorithms Like Linear Regression, Polynomial Regression & Finally our Proposed Method gives better result.

FUTURE SCOPE

In The future works focus on applying some other techniques to improving the performances of these methods for up to maximum extent. Another concept that can be implemented Deep learning in place of machine learning technology. The reason behind this is best and efficient techniques using nowadays. Deep learning is also introduced nowadays which is becoming more popular for classification purpose. So, we can also implement deep learning in future work also.

REFERENCES

[1] Bhardwaj, S., Bhardwaj, H., Bhardwaj, J. and Gupta, P., 2021, November. Global Prediction of COVID-19 Cases and Deaths using Machine Learning. In 2021 Sixth International Conference on Image Information Processing (ICIIP) (Vol. 6, pp. 422-426). IEEE.

[2] Sen, A., Kala, U. and Manchanda, A., 2021, February. Analysis and prognosis of COVID-19 pandemic in India-A machine learning approach. In 2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT) (pp. 1-6). IEEE.

[3] Gambhir, E., Jain, R., Gupta, A. and Tomer, U., 2020, September. Regression analysis of COVID-19 using machine learning algorithms. In 2020 International conference on smart electronics and communication (ICOSEC) (pp. 65-71). IEEE.

[4] Mary, L.W. and Raj, S.A.A., 2021, October. Machine Learning Algorithms for Predicting SARS-CoV-2 (COVID-19)–A Comparative Analysis. In 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC) (pp. 1607-1611). IEEE.

[5] Lalmuanawma, S., Hussain, J. and Chhakchhuak, L., 2020. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. Chaos, Solitons & Fractals, 139, p.110059.
[6] Kumari, P. and Toshniwal, D., 2020, November. Real-time estimation of COVID-19 cases using machine learning and mathematical models-The case of India. In 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS) (pp. 369-374). IEEE.
[7] Patel, D., Kher, V., Desai, B., Lei, X., Cen, S., Nanda, N., Gholamrezanezhad, A., Duddalwar, V., Varghese, B. and Oberai, A.A., 2021. Machine learning based predictors for COVID-19 disease severity. Scientific reports, 11(1)pp1-7 .
[8] Subudhi, S., Verma, A., Patel, A.B., Hardin, C.C., Khandekar, M.J., Lee, H., McEvoy, D., Stylianopoulos, T., Munn, L.L., Dutta, S. and Jain, R.K., 2021. Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. NPJ digital medicine, 4(1), pp.1-7.
[9] Fernandes, F.T., de Oliveira, T.A., Teixeira, C.E., Batista, A.F.D.M., Dalla Costa, G. and Chiavegatto Filho, A.D.P., 2021. A multipurpose machine learning approach to predict COVID-19 negative prognosis in São Paulo, Brazil. Scientific reports, 11(1), pp.1-7.
[10] Coronavirus Disease (COVID-19) - events as they happen. Library Catalog:www.who.int.
[11] Countries where Coronavirus has spread - Worldometer. Library Catalog:www.worldometers.info.
[12] COVID-19 situation reports. Library Catalog: www.who.int.
[13] Diagnosis of covid-19 and its clinical spectrum dataset.url=https://kaggle.com/einsteindata4u/covid19.
[14] WHO Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020. Library Catalog: www.who.int.
[15] WHO EMRO | Questions and answers | COVID-19 | Health topics.
[16] Support Vector Machine Machine learning algorithm with example and code,January 2019. Library Catalog: www.codershood.info Section: Machine learning.
[17] Ali Al-Hazmi. Challenges presented by MERS corona virus, and SARS corona virus to global health. Saudi journal of biological sciences, 23(4):507–511, 2016.Publisher: Elsevier.
[18] Sina F Ardabili, Amir Mosavi, Pedram Ghamisi, Filip Ferdinand, Annamaria R Varkonyi-Koczy, Uwe Reuter, Timon Rabczuk, and Peter M Atkinson. Covid-19 outbreak prediction with machine learning. Available at SSRN 3580188, 2020.
[19] Hiba Asri, Hajar Mousannif, Hassan Al Moatassime, and Thomas Noel. Using machine learning algorithms for breast cancer risk prediction and diagnosis. Procedia Computer Science, 83:1064–1069, 2016.
[20] Taiwo Oladipupo Ayodele. Types of machine learning algorithms. New advances in machine learning, pages 19–48, 2010.
[21] Taiwo Oladipupo Ayodele. Types of machine learning algorithms. New advances in machine learning, pages 19–48, 2010. Publisher: InTech.
[22] David W Bates, Suchi Saria, Lucila Ohno-Machado, Anand Shah, and Gabriel Escobar. Big data in health care: using analytics to identify and manage highrisk and high-cost patients. Health Affairs, 33(7):1123–1131, 2014.
[23] Hetal Bhavsar and Amit Ganatra. A comparative study of training algorithms for supervised machine learning. International Journal of Soft Computing and Engineering (IJSCE), 2(4):2231–2307, 2012.
[24] Rich Caruana and Alexandru Niculescu-Mizil. An empirical comparison of supervised learning algorithms. In Proceedings of the 23rd international conference on Machine learning, pages 161–168, 2006.
[25] Nanshan Chen, Min Zhou, Xuan Dong, Jieming Qu, Fengyun Gong, Yang Han,Yang Qiu, Jingli Wang, Ying Liu, Yuan Wei, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in wuhan, china:a descriptive study. The Lancet, 395(10223):507–513, 2020.
[26] Dursun Delen. Analysis of cancer data: a data mining approach. Expert Systems,26(1):100–112, 2009.
[27] Manoj Durairaj and Veera Ranjani. Data mining applications in healthcare sector: a study. International journal of scientific & technology research, 2(10):29–35, 2013.
[28] Ashok Kumar Dwivedi. Performance evaluation of different machine learning techniques for prediction of heart disease. Neural Computing and Applications,29(10):685–693,
[29] Arihito Endo, Takeo Shibata, and Hiroshi Tanaka. Comparison of seven algorithms to predict breast cancer survival (< special issue> contribution to 21 century intelligent technologies and bioinformatics). International Journal of Biomedical Soft Computing and Human 13(2):11–16, 2008.
[30] Leslie Pack Kaelbling, Michael L Littman, ali,narin and Andrew W Moore. Reinforcement learning: A survey. Journal of artificial intelligence research, 4:237–285,1996.
[31] Ya-Han Hu, Yi-Lien Lee, Ming-Feng Kang, and Pei-Ju Lee. Constructing inpatient pressure injury prediction models using machine learning techniques. Computers, Informatics, Nursing: CIN, 2020.
[32] G. Kesavaraj and S. Sukumaran. A study on classification techniques in datamining. In 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), pages 1–7, 2013.
[33] Barbara Kitchenham. Procedures for performing systematic reviews. Keele,UK, Keele University, 33(2004):1–26, 2004.
[34] Halgurd S. Maghdid, Kayhan Zrar Ghafoor, Ali Safaa Sadiq, Kevin Curran,and Khaled Rabie. A novel ai-enabled framework to diagnose coronavirus covid 19 using smartphone embedded sensors: Design study. arXiv preprint arXiv:2003.07434, 2020.

[35] R. Alizadehsani, M.J. Hosseini, Z.A. Sani, A. Ghandeharioun, R. Boghrati, Diagnosis of coronary artery disease using cost-sensitive algorithms, in: Proceedings of the IEEE 12th International Conference on Data Mining Workshops, IEEE, 2012, pp. 9–16, https://doi.org/10.1109/ICDMW.2012.29.
[36] C. Mahesh, K. Kiruthika, M. Dhilsathfathima, Diagnosing hepatitis B using netural network based expert system, in: Proceedings of the International Conference on Information Communication and Embedded Systems (ICICES2014), IEEE, 2014, pp. 1–7.
[37] M.E.B. Menai, Random forests for automatic differential diagnosis of erythemato–squamous diseases, Int. J. Med. Eng. Inform. 7 (2) (2015) 124–141.
[38] S. Ansari, I. Sha□, A. Ansari, J. Ahmad, S.I. Shah, Diagnosis of liver disease induced by hepatitis virus using arti□cial neural networks, in: Proceedings of the IEEE 14th International Multitopic Conference, IEEE, 2011, pp. 8–12, https://doi. org/10.1109/INMIC.2011.6151515.
[39] J.S. Sartakhti, M.H. Zangooei, K. Mozafari, Hepatitis disease diagnosis using a novel hybrid method based on support vector machine and simulated annealing (SVM-SA), Comput. Methods Programs Biomed. 108 (2) (2012) 570–579.
[40] T. Mazzocco, A. Hussain, Novel logistic regression models to aid the diagnosis of dementia, Expert Syst. Appl. 39 (3) (2012) 3356–3361.
[41] L. Parisi, N. RaviChandran, M.L. Manaog, Feature-driven machine learning to improve early diagnosis of Parkinson's disease, Expert Syst. Appl. 110 (2018) 182–190.
[42] K.M. Eggers, J. Ellenius, M. Dellborg, T. Groth, J. Oldgren, E. Swahn, B. Lindahl, Arti□cial neural network algorithms for early diagnosis of acute myocardial infarction and prediction of infarct size in chest pain patients, Int. J. Cardiol.114(3)(2007)366–374,
[43] Mahesh, Generalized regression neural network based expert system for hepatitis B diagnosis, J. Comput. Sci. 10 (4) (2014) 563–569, https://doi.org/10.3844/ jcssp.2014.563.569.
[44] JuH Lee, YNa Hwang, S.Y. Park, J.H. Jeong, S.M. Kim, Diagnosis of osteoporosis by quanti□cation of trabecular microarchitectures from hip radiographs using arti□cial neural networks, J. Computat. Theor. Nanosci. 12 (7) (2015) 1115–1120, https://doi.org/10.1166/jctn.2015.3859.
[45] Y.M. Chae, S.H. Ho, C.S. Hong, C.W. Kim, Comparison of alternative knowledge model for the diagnosis of asthma, Expert Syst. Appl. 11 (4) (1996) 423–429, https://doi.org/10.1016/S0957-4174(96)00057-7.
[46] L. Liu, M. Deng, An evolutionary arti□cial neural network approach for breast cancer diagnosis, in: Proceedings of the Third International Conference on Knowledge Discovery and Data Mining, IEEE, 2010, pp. 593–596, https://doi. org/10.1109/WKDD.2010.148.
[47] Kamil, Anas. (2022). Analysis of CT images of the COVID-19 patients. 10.1109/MICEST54286.2022.9790201.
[48] K. Chakrabortya, A. Chattopadhyayb, A. Chakrabarti, T. Acharyad, A. K. Dasguptae, A combined algorithm for malaria detection from thick smear blood slides, J. Health Med. Inf. 6 (1) (2015) 645–652, https://doi.org/10.4172/ 2157-7420.1000179.
[49] L. Shi, Z. Guan, C. Liang, H. You, Automatic classi□cation of Plasmodium for malaria diagnosis based on ensemble neural network, in: Proceedings of the 2nd International Conference on Intelligent Medicine and Image Processing, 2020, pp. 80–85