

Research Article

Optimizing Medicine Dosage Recognition through Symptom Intensity Assessment with Support Vector Machines

Rekha Gangula^{1,*}, Murali Mohan² and DayakarThalla³

¹Assistant Professor, Department of CSE, Vaagdevi Engineering College, Warangal-506005, Telangana, India.

² Associate Professor, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India-522503

³Assistant Professor, Department of CSE(AI&ML), Vaagdevi Engineering College, Warangal, 506005, Telangana, India.

*Corresponding Author: Rekha Gangula. Email: gangularekha@gmail.com

Received: 28/01/2024; Accepted: 22/02/2024.

DOI: <https://doi.org/10.69996/jcai.2024003>

Abstract: Machine Learning (ML) is a set of contemporary approaches to predicting, recognising, and making decisions without the use of humans. From disease detection to simulation, machine learning is rapidly emerging in the medical industry. The purpose of the proposed study is to examine how supervised machine learning algorithms, such as support vector machines (SVMs), can be used to predict medicine dose. The main advantage of employing SVM-based algorithms is that they can handle an unlimited number of input parameters (patient characteristics, like symptoms), and each parameter is treated identically no matter how differently it looks on the surface. SVM learning, or machine learning with maximisation (support) of separating margin (vector), is a strong classification algorithm that's used for classification or subtyping. Building a machine learning model to determine the appropriate medication dose for and patient is critical to clinical practise and time-consuming for modelling software. Rather than using conventional explicit approaches, we suggest a machine learning approach in this article for predicting disease and determining medicine dosage based on the patient's symptoms.

Keywords: Machine learning; SVM; disease prediction; medicine dosage; EDA.

1 Introduction

The integration of computer technology into almost every aspect of our everyday lives has raised our standard of living. The use of computers has increased dramatically, especially in the field of medicine, in areas such as disease diagnosis and treatment, as well as patient tracking. Since the fields in which computers are used are so complex and unpredictable, intelligent systems like machine learning, artificial neural networks, and genetic algorithms have been developed. We introduced the SVM algorithm in this paper for determining medication dosage based on a patient's symptoms dataset.

SVM: A supervised machine learning approach called support vector machine is used to categorise data into various classifications. From Fig1 Therapeutic can make use of the procedure for drug monitoring.

SVM, in contrast to other algorithms, employs hyperplanes to give a clear and accurate decision border between various classes. By splitting the data into segments (small sets, records, or attributes) with various types of data in each row or attribute, SVM can be used to build several separating hyperplanes. SVM is used on a dataset to train how to process the model and predict the outcome.

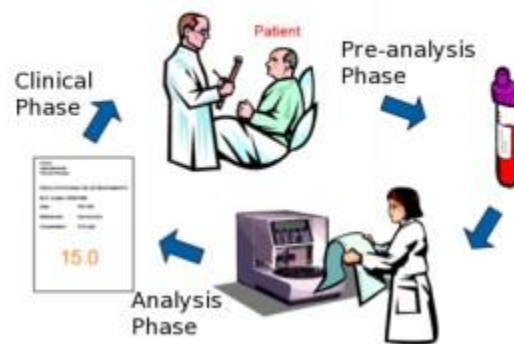


Figure 1: Drug Monitoring

SVM looks for named training data and then produces new input data based on it. The ability to solve classification and regression problems is a key advantage of SVM. However, SVR (classification) focuses on classification while SVM (support vector regression) focuses on regression issues.

SVM can be either Linear or Non-Linear. Linear data is the data with which we can easily separate by just drawing a margin Fig 2, hyper plane we can predict the exact solution. If the data is non-linear, we use kernel trick as mentioned above, it is used to transform or correct data into another dimension. Up to this point we had only 2 dimensions example namely x & y , but now one more dimension is introduced called z Fig 3. Now easily we can visualize the data and separate the two classes.

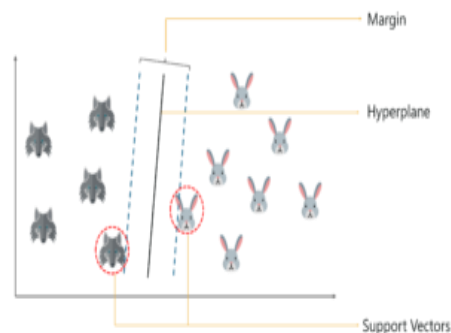
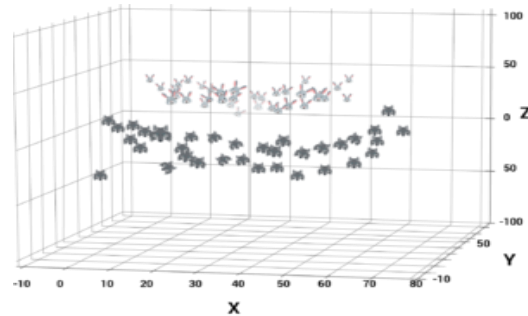


Figure 2: Linear data**Figure 3: Non-linear data**

2 Literature Survey

The classification method is called Support Vector Machine. In the 1990s, cooperation between the machine learning and statistical research fields led to the development of the support vector machine (SVM) by Cortes and Vapnik. SVM finds a dividing line called a hyper plane in an effort to discern between situations. The main benefit of the SVM is its ability to quickly resolve the "high dimensional problem," which arises when there are too many input variables in relation to the total number of observations. Additionally, since the SVM approach is data-driven and can be applied without a theoretical framework, it has a high degree of discriminative power for classification, especially with small sample sizes. It was recently observed that this technology can improve clinical settings' procedures for disease diagnosis. Additionally, SVM has demonstrated efficacy in resolving bioinformatics classification issues. In this investigation, conventional variables that are quite straightforward to obtain were used to predict both the patient's disease and the medicine dose using support vector machines (SVM).

3 Proposed System

The proposed system is used to develop a model for disease prediction and aids users in identifying the medicine dosage for the predicted disease. Here we propose the process of identification of medicine dosage of patients based on symptoms using the SVM algorithm. This process consists of following steps,

1. Import all the crucial libraries
2. Load the training and test data set.
3. Exploratory Data Analysis (EDA).
4. Splitting of training and validation features.
5. Creating the ML model

6. Predict the disease and medicine dosage using the SVM algorithm (Model).

4 Methodology

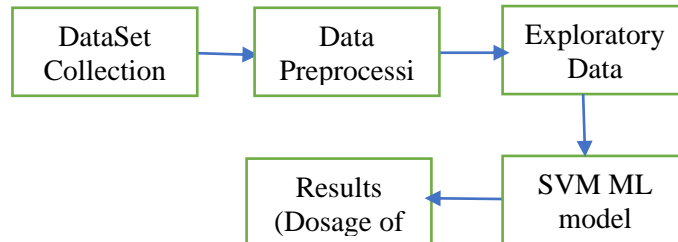


Figure 4: Methodology

The approach A step-by-step procedure for creating the model is shown in Fig 4. The entire procedure is split into two sections, the first of which is created based on the symptoms and the second of which is created based on the anticipated illness.

- **Dataset collection**

A fictitious dataset has been constructed through the examination of earlier research. It has 4921 samples with 132 attributes, each of which indicates whether an attribute is present or absent with a value of 1. Sweating, nausea, headaches, vomiting, skin rashes, lethargy, mild fever, depression, appetite loss, muscle atrophy, diarrhoea, back pain, chills, exhaustion, stomach discomfort, shivering, and so forth are a few of the symptoms.

- **Dataset cleaning**

After loading, the dataset is examined. If there are any null, junk, or missing values, the relevant cleaning techniques are applied.

- **Exploratory Data Analysis (EDA)**

The relationship between the attributes is found for additional processing in EDA Fig 5. Understanding of the correlation between the features Fig 6 along with the disease wise distribution Fig 7 and symptom wise distribution Fig 7 is a part of the EDA process.

- **Importing packages**

All the crucial libraries and packages namely pandas, joblib, sklearn.metrics, numpy, yaml, flask, seaborn, itertools are imported.

- **Splitting Data**

To create the model, the dataset is divided into training and testing sets. A split function that was imported from the sklearn package can be used to do this.

- **Building the model**

The model is constructed by importing the SVM package from the sklearn library, after which functions are implemented to train and test the model.

- **Preprocessing and preparing to train**

To get ready for training the model, the train and test data are pre-processed using the necessary function.

- **Training the model**
Using the prepared data, the support vector machine (SVM) algorithm is used to train the model.
- **Classification report:**
The model's performance is reported in the categorization report. Confusion matrix Fig 9 is built to showcase the accuracy of the model and performs the classifier Fig 10. Here we can render the classification report of validation data Fig.11 and test data Fig. 12.

5 Algorithm

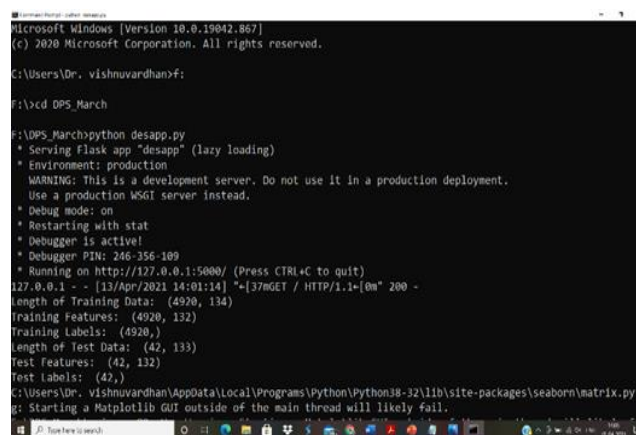
Support Vector Machine (SVM)

SVM is a supervised machine learning algorithm that is uniquely recognized for converting data into extraordinary instructions. SVM is unique in that it uses a hyper plane to act as a selection boundary between the various instructions. SVM can be used to produce a large number of isolating hyperplanes, allowing the data to be divided into parts, each of which contains only one type of data. The SVM algorithm is trained on a set of categorized data. SVM learns the previously categorized training data and classifies new input data based on the learnings made during the training process. SVM has the benefit of being applicable to both form and regression issues. One of the most important advantages of the Support vector machine algorithm is this. Despite the fact that SVM is well-known for type, the Support Vector Regressor can be used to solve regression problems. SVM can also be used for non-linear data using the kernel trick. The kernel trick converts data to any other size with a clean isolating margin between information instructions. After that, we can easily draw a hyper plane among a large number of data instructions.

6 Results and Discussion

The medicine dosage for the patient predicted using the SVM algorithm is the final result of this project. Here we used GUI to render the data entry of patient's information from Fig 13 and disease Prediction shown in Fig 14.

EXPLORATORY DATA ANALYSIS (EDA):



```

Microsoft Windows [Version 10.0.19042.867]
(c) 2020 Microsoft Corporation. All rights reserved.

C:\Users\Dr. vishnuvardhan>

F:\>cd DPS_March

F:\DPS_March>python desapp.py
 * Serving Flask app "desapp" (lazy loading)
 * Environment: production
   WARNING: This is a development server. Do not use it in a production deployment.
   Use a production WSGI server instead.
 * Debug mode: on
 * Restarting with stat
 * Debugger is active!
 * Debugger PIN: 246-356-109
 * Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
127.0.0.1 - - [13/Apr/2021 14:01:14] "GET / HTTP/1.1" 200 -
length of Training Data: (4920, 134)
Training Features: (4920, 132)
Training Labels: (4920,)
length of Test Data: (42, 133)
Test Features: (42, 132)
Test Labels: (42,)
C:\Users\Dr. vishnuvardhan\AppData\Local\Programs\Python\Python39-32\Lib\site-packages\seaborn\matrix.py:
g: Starting a Matplotlib GUI outside of the main thread will likely fail.

```

Figure 5: Loading Training and Test data sets

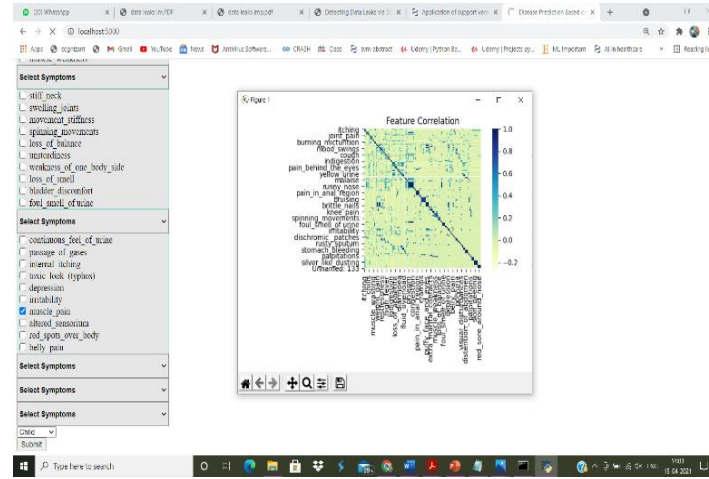


Figure 6: Feature correlation

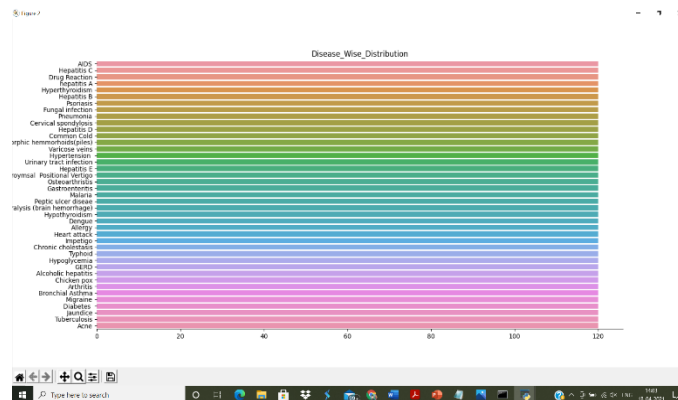


Figure 7: Disease wise distribution

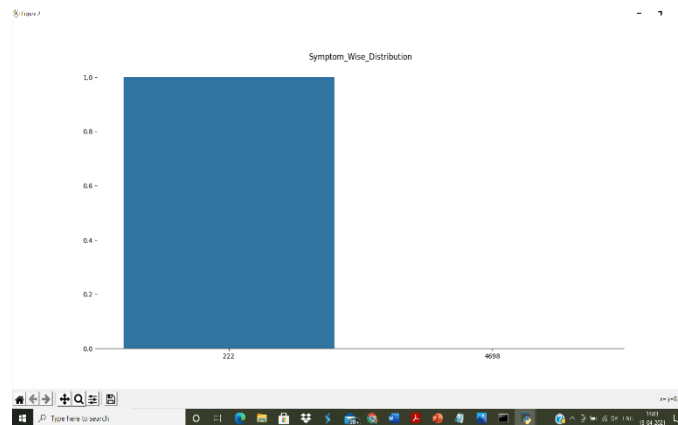


Figure 8: Symptom wise distribution

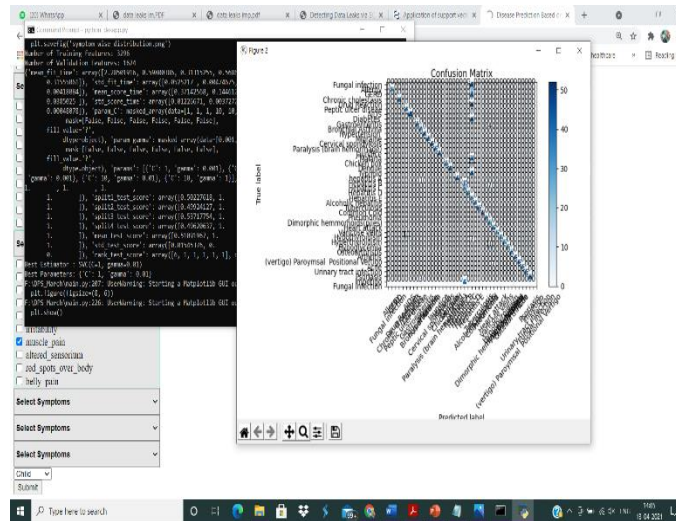


Figure 9: Confusion Matrix

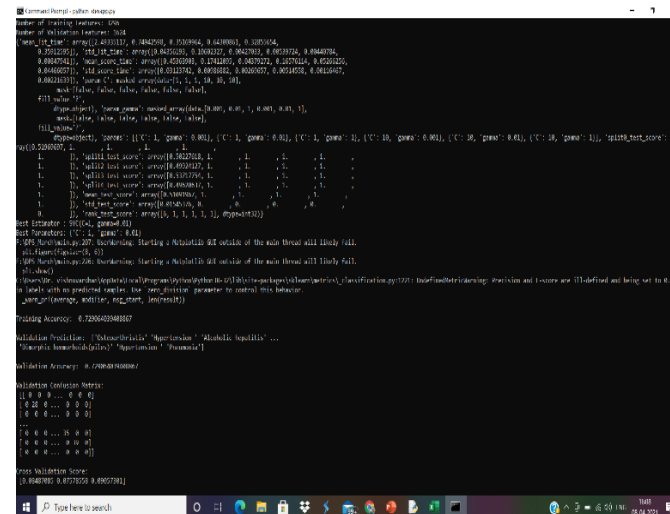


Figure 10: Classifier

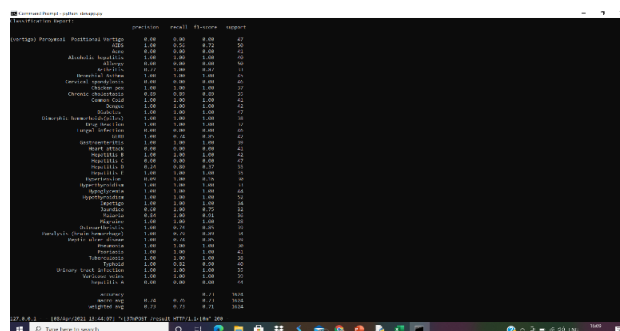


Figure 11: Classification report for validation data

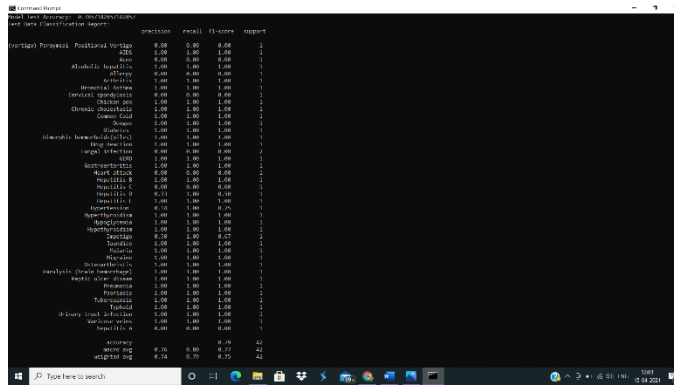


Figure 12: Classification report for test data

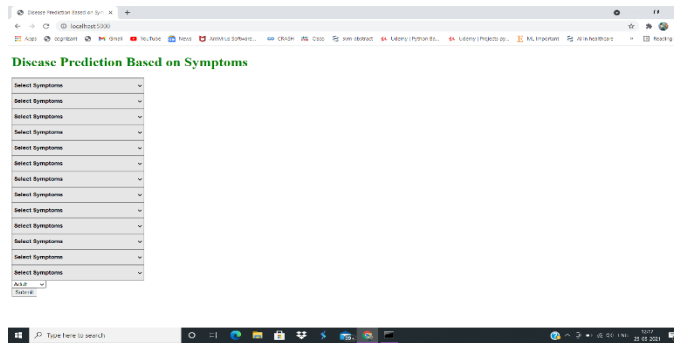


Figure 13: Patient's data entry page

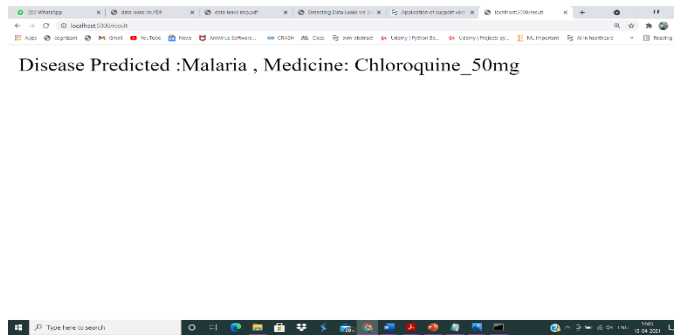


Figure 14: Predicted output

7 Conclusion and Future Work

Assisting the predictions of medication dosage values with inside the medical scenario with SVM has proved to yield a few great outcomes. It can construct personalised version for each affected person in medication dosage evaluation to be able to assist provide the steerage to the docs with inside the dosage prescription. The problem with this approach, despite the previously mentioned advantages, is that it mostly depends on the library's initial patient data for

accurate results. The proposed version on this paper makes use of SVM – Support Vector Machine Algorithm for prediction which has true overall performance while as compared to different models. From the experimental outcomes we will finish that proposed version has stepped forward accuracy in prediction of medication dosage. In destiny this version also can be stepped forward via way of means of the use of the entered records too as take a look at records, which similarly improves the overall performance of this algorithm. Future scope is Disease prediction and medication dosage can be hosted online, making them accessible to anyone with an internet connection. Future research has shown that an SVM-based approach can be used to create a customized model for each patient.

Acknowledgement: Not Applicable.

Funding Statement: The author(s) received no specific funding for this study.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

- [1] W. You, N. Midmer and G.D. Micheli, “Personalized modeling for drug concentration prediction using Support Vector Machine,” *2011 4th International Conference on Biomedical Engineering and Informatics (BMEI)*, Shanghai, China, 2011.
- [2] Y.J Son, H.G. Kim, E.H. Kim, S. Choi and S.K.Lee, “Application of Support Vector Machine for Prediction of Medication Adherence in Heart Failure Patients,” *Healthcare Informatics Research*, vol.16, no.4, pp. 253–259 2010.
- [3] W. A. David Bourne, “Mathematical Modelling of Pharmacokinetic Data”, *Technomic Publishing Company, Inc.*, 1995.
- [4] F. Reginald Brown, “Compartmental System Analysis: State of the Art”, *IEEE Transactions on Biomedical Engineering*, vol. BME-27, No. 1, pp. 1-11, 1980.
- [5] Gary Blau and Seza Orcun, “A Bayesian Pharmacometric Approach for Personalized Medicine - A Proof of Concept Study with Simulated Data,” *Proceedings of the 2009 Winter Simulation Conference*, pp.1969-76, 2009.
- [6] Wenqi You, Nicolas Widmer and Giovanni De Micheli, “Example-based Support Vector Machine for Drug Concentration Analysis”, *33rd IEEE EMBS*, USA, pp. 153-157, 2011.
- [7] Ethem Alpaydin, “Introduction to Machine Learning (Adaptive Computation and Machine Learning),” *MIT Press*, 2004.
- [8] S. Ni Karl and Q. Truong Nguyen, “Image Super resolution Using Support Vector Regress”, *IEEE Transactions on Image Processing*, vol. 16, pp. 1596-1610, 2007.
- [9] C. Cortes and V. Vapnik . “Support-vector networks”. *Mach Learn*, vol. 20, pp. 273-297, 2009.
- [10] I. Maglogiannis, E. Loukis and E. Zafiropoulos, “Support vectors machine-based identification of heart valve diseases using heart sounds,” *Comput Methods Programs Biomed* vol. 95, pp. 47-61, 2009.
- [11] K. Choi, S. Chung and H. Rhee, “Classification and sequential pattern analysis for improving managerial efficiency and providing better medical service in public healthcare centers,” *Healthc Inform Res*, vol. 16, pp. 67-76, 2010.
- [12] I. Guyon, J. Weston and S. Barnhill, “Gene selection for cancer classification using support vector machines,” *Mach Learn*, vol. 46, pp. 389-422, 2002.
- [13] C. C. Chang , C. J. Lin . “LIBSVM: a library for support vector machines [Internet]”. [cited at 2010 Oct 29]. Availablefrom: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.

- [14] D. G. Morrow, M. Weiner and J. Young, "Improving medication knowledge among older adults with heart failure: a patient-centered approach to instruction design," *Gerontologist*, vol.45, pp. 545-552, 2005.
 - [15] R Steve, Gunn, "Support Vector Machines for Classification and Regression," *Technical Report*, University of Southampton.
 - [16] U. Srilakshmi, "Energy-efficient heterogeneous optimization routing protocol for wireless sensor network," *Instrumentation Mesure Métrologie*, vol. 19, no.5, pp. 391-397, 2020.
 - [17] N. Veeraiah and B.T. Krishna, "Trust-aware Fuzzy Clus-Fuzzy NB: intrusion detection scheme based on fuzzy clustering and Bayesian rule," *Wireless Networks*, vol. 25, no.1, pp. 4021-4035, 2019.
 - [18] A. Alsirhani, M. Ezz and A. M. Mostafa, "Advanced authentication mechanisms for identity and access management in cloud computing," *Computer Systems Science and Engineering*, vol. 43, no.3, pp. 967-984, 2022.
 - [19] M. Ragab, H. A. Abdushkour, A. F. Nahhas and W. H. Aljedaibi, "Deer hunting optimization with deep learning model for lung cancer classification," *Computers, Materials & Continua*, vol. 73, no.1, pp. 533-546, 2022.
 - [20] U. Srilakshmi, S. A. Alghamdi, V. A. Vuyyuru, N. Veeraiah and Y. Alotaibi, "A secure optimization routing algorithm for mobile ad hoc networks," in *IEEE Access*, vol. 10, pp. 14260-14269, 2022.
 - [21] U. Srilakshmi, N. Veeraiah, Y. Alotaibi, S. A. Alghamdi, O. I. Khalaf *et al.*, "An improved hybrid secure multipath routing protocol for manet," in *IEEE Access*, vol. 9, no.1, pp. 163043-163053, 2021.
-