Bulletin of Environment, Pharmacology and Life Sciences
Bull. Env. Pharmacol. Life Sci., Vol 12 [4] March 2023 : 269-274

Online ISSN 2277-1808
Journal's URL:http://www.bepls.com
CODEN: BEPLAD

REVIEW ARTICLE **OPEN ACCESS**

IoT Based Heart Disease Prediction using Machine Learning techniques

P. Jagadeesan[1], M.Vedaraj[2], K.Dhanalakshmi[3], D.Surendiran Muthukumar[4]
[1]Assistant Professor, Department of Computer Science and Engineering, R.M.D. Engineering College, TN
[2]Associate Professor, Department of Computer Science and Engineering, R.M.D. Engineering College, TN
[3] Assistant Professor, Department of Computer Science and Engineering, Veltech Multitech Dr.Rangarajan Dr.Sakunthala Engineering College, TN
[4] Assistant Professor, Department of Computer Science and Engineering , Kalasalingam Academy of Research And Education, TN

ABSTRACT
The goal of our heart disease prediction model is to determine if a patient should be diagnosed with heart disease or not, which is a binary outcome, so: Positive result = 1, the patient will be diagnosed with heart disease. Negative result = 0, the patient will not be diagnosed with heart disease. In this work, IoT based Healthcare system for diabetics disease prediction is proposed using Gaussian Naïve Bayes machine learning algorithm. The proposed Gaussian NB algorithm is compared with existing machine learning algorithm such as SVM, LR, DT, and KNN. From the experimental result analysis, the proposed Gaussian NB algorithm achieves 97.20% accuracy of Heart disease prediction. Health monitoring is the major problem in today's world. Due to lack of proper health monitoring, patient suffer from serious health issues. There are lots of IoT devices now days to monitor the health of patient over internet. Health experts are also taking advantage of these smart devices to keep an eye on their patients. With tons of new healthcare technology start-ups, IoT is rapidly revolutionizing the healthcare industry.
Keywords: IoT; Machine Learning; Prediction; Accuracy; Healthcare

Received 18.01.2023 Revised 16.02.2023 Accepted 21.03.2023

INTRODUCTION

In recent years, the Internet of Things (IoT) has been used as a monitoring and assessment tool to track the current condition of structure or machine or equipment. This tool collects, analyses, and transfers the various data parameters related to the service condition of the equipment, structure, or machine. This ultimately leads to optimization of cost as per its repair and maintenance of the machine is concerned. It even reduces the related manpower as well. In addition, the life of the machinery can be extended by early detection or identification of the issues and their corrective measure[1]. IoT is nothing but linking computers to the internet utilizing sensors and networks [2]. These connected components can be used on devices for health monitoring. The used sensors then forward the information to distant locations like M2M, which are machinery for computers, machines for people, handheld devices, or smartphones [3]. It is a simple, energy-efficient, much smarter, scalable, and interoperable way of tracking and optimizing care to any health problem. Nowadays, modern systems are providing a flexible interface [4], assistant devices , and mental health management lead a smart life for the human being.

IoT-based healthcare applications are utilized to gather essential data, such as real-time changes in health parameters and updates of medical parameter's severity within a standard time interval, so IoT devices incessantly produce huge amounts of health data. The IoT is acknowledged as the most crucial future technology and is gaining much attention from healthcare industries [5],

Heart disease is a grave disease that influences the heart's functionality and gives rise to complications such as infection of the coronary artery and diminished blood vessel function. Heart disease patients do not feel sick until the very last stage of the disease, and then it is too late because the damages have become irretrievable. In the past years, sensor networks for healthcare IoT have advanced quickly, so it is now possible to incorporate instantaneous health data by linking bodies and sensors. It is important to diagnose patients early by means of ECG. IoT-based heart attack detection systems raise privacy and security concerns. As mobile devices are possible targets for malevolent attacks, more studies are needed

on safety countermeasures. A fault-tolerant algorithm should be developed for a dependable IoT system [6].

The structure of this paper is as follows. Section II presents literature survey related to heart disease prediction. Section III briefly discusses the proposed methodology, and Section IV explores the experiment's result and analysis. Section V provides conclusions and future enhancement.

RELATED WORK

Ali et al. [7] introduced an automated diagnostic framework for heart disease diagnosis. First, normalization was done on feature vectors, and then the data were divided into training and test datasets. Subsequently, selection and feature ranking were done on the trained data using a statistical framework. The same subset of features that was chosen by the framework in the training phase was utilized for testing the data. The training data with a reduced number of features were employed by a neural network (NN) for training purposes. The performance of the trained NN was assessed using the test data.

Bo Jin, Chao Che et al. [8] proposed a "Predicting the Risk of Heart Failure With EHR Sequential Data Modeling" model designed by applying neural network. This paper used the electronic health record (EHR) data from real-world datasets related to congestive heart disease to perform the experiment and predict the heart disease before itself. We tend to used one-hot encryption and word vectors to model the diagnosing events and foretold coronary failure events victimization the essential principles of an extended memory network model. By analysing the results, we tend to reveal the importance of respecting the sequential nature of clinical records.

Ashir Javeed, Shijie Zhou et al. [9] designed "An Intelligent Learning System based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection". This paper uses random search algorithm (RSA) for factor selection and random forest model for diagnosing the cardiovascular disease. This model is principally optimized for using grid search algorithmic program. Two forms of experiments are used for cardiovascular disease prediction. In the first form, only random forest model is developed and within the second experiment the proposed Random Search Algorithm based random forest model is developed. This methodology is efficient and less complex than conventional random forest model. Comparing to conventional random forest it produces 3.3% higher accuracy. The proposed learning system can help the physicians to improve the quality of heart failure detection.

Abhay Kishore et al. [10] developed Heart Attack Prediction Using Deep Learning in which This paper proposes a heart attack prediction system using Deep learning procedures, explicitly Recurrent Neural System to predict the probable prospects of heart related infections of the patient. Recurrent Neural Network is a very ground-breaking characterization calculation that utilizes Deep Learning approach in Artificial Neural Network. The paper talks about in detail the significant modules of the framework alongside the related hypothesis. The proposed model deep learning and data mining to give the precise outcomes least blunders. This paper gives a bearing and point of reference for the advancement of another type of heart attack prediction platform. Prediction stage.

Bhuvaneswari *et al.* [11] use the Naive Bayes classifier for medical use. The authors used two well-known algorithms, the Back Propagation Neural Network (BNN) and the Nave Bayesian (NB) data mining classification, to study the previous experience and to calculate the probability of an object among all objects. Bayesian techniques have been developed for probability concepts. The previous backend is computed by bay rules based on the exact nature of the probability model, and the Naive Bayes classifier is used to study very efficiently in the supervised learning environment.

Md. Shahriare Satu *et.al* [12] present that Heart Disease is one of the leading diseases that causes enormous loss of lives all over the world. Some unusual approaches to find out significant factors of heart diseases have been considered by the authors. They have used two heart disease data (Cleveland & Hungarian) and both of them are divided into 33%, 65% and 100% data. Values of different range of individual attributes in these data are determined to find out relevant factors of this disease. Then, different semi supervised learning algorithms such as Collective Wrapper, Filtered Collective and Yet Another Semi Supervised Idea are used to analyze heart disease data. Metrics of these classifiers like accuracy, f-measure and area under ROC have been calculated to justify individual classifiers and specify the best semi supervised learning algorithm. This algorithm is explored significant and irrelevant factors of heart disease by removing attributes one after another sequentially and observing the outcomes of classification. Experimental results on two real data demonstrates the effectiveness and efficiency of the analysis

Sarath Babu *et al.* [13 have suggested that medical data mining will explore the secret trends of medical data sets with great potential. These trends may be used to diagnose clinically. The data must be obtained in a structured manner. From the healthcare profiles, the patient can be predicted by 14 characteristics,

such as age, sex, blood pressure and blood sugar etc. Such attributes are fed in the prediction of heart disease with K-means algorithms, MAFIA algorithms and decision tree classification, using data mining technology to diagnose heart disease; it can achieve as much precision as it can when diagnosing a heart condition. By this medical industries could offer better diagnosis and treatment of the patient to attain a good quality of services. The main advantages pointed in the research are: early detection of heart disease and its diagnosis correctly on time and providing treatment with affordable cost.

PROPOSED MODEL

The proposed model developed for medical data classification and prediction employs artificial intelligence and machine learning techniques. Sensors (wearables) and datasets are essential components of the proposed research. The proposed prediction model for heart disease is shown in Figure 1.

The medical data collected by the sensors was transferred to the system via Bluetooth and stored as binary and comma-separated value (.csv) files. The medical data was kept in the cloud for evaluation by users and doctors. Data pre-processing in Machine Learning refers to the technique of preparing (cleaning and organizing) the raw data to make it suitable for a building and training Machine Learning models. Feature Selection is the method of reducing the input variable to your model by using only relevant data and getting rid of noise in data. It is the process of automatically choosing relevant features for your machine learning model based on the type of problem you are trying to solve. The classification algorithms used in the present work are Gaussian NB, In this prediction model, data stored in the cloud was accessed and processed for medical data classification.

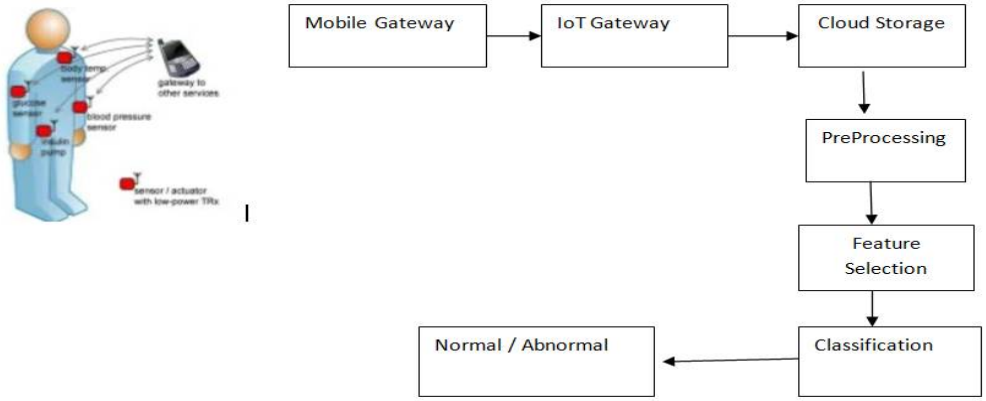

Figure1: Heart Disease Prediction Model

DATASET DESCRIPTION

Cleveland Heart Disease [29] dataset is considered for testing purpose in this study. During the designing of this data set there were 303 instances and 75 attributes, however all published experiments refer to using a subset of 14 of them. In this work, we performed pre-processing on the data set, and 6 samples have been eliminated due to missing values. The remaining samples of 297 and 13 features dataset is left and with 1 output label. The output label has two classes to describe the absence of HD and the presence of HD. The Figure2 shows Cleveland Heart Disease dataset.

S.No	Attribute Name	Description	Range of Values
1	Age	Age of the person in years	29 to 79
2	Sex	Gender of the person [1: Male, 0: Female]	0, 1
3	Cp	Chest pain type [1-Typical Type 1 Angina 2- Atypical Type Angina 3-Non-angina pain 4-Asymptomatic)	1, 2, 3, 4
4	Trestbps	Resting Blood Pressure in mm Hg	94 to 200
5	Chol	Serum cholesterol in mg/dl	126 to 564
6	Fbs	Fasting Blood Sugar in mg/dl	0, 1
7	Restecg	Resting Electrocardiographic Results	0, 1, 2
8	Thalach	Maximum Heart Rate Achieved	71 to 202
9	Exang	Exercise Induced Angina	0, 1
10	OldPeak	ST depression induced by exercise relative to rest	1 to 3
11	Slope	Slope of the Peak Exercise ST segment	1, 2, 3
12	Ca	Number of major vessels colored by fluoroscopy	0 to 3
13	Thal	3 – Normal, 6 – Fixed Defect, 7 – Reversible Defect	3, 6, 7
14	Num	Class Attribute	0 or 1

Figure 2: Cleveland Heart Disease

Performance Evaluation Metrics

Evaluation of the proposed model is performed focusing on some criteria, namely, accuracy, recall, precision, and F-score.

Accuracy is one of the most important performance metrics for classification. It is defined as the proportion between the correct classification and the total sample, as shown in the following equation:

Accuracy = TP+TN/TP+FP+FN+TN ----------(1)

Recall is the small portion of sufficient instances over the overall quantity of applicable instances which have been recovered. The recall equation is shown as follows:

Recall = TP/TP+FN ----------(2)

Precision is identified as follows:

Precision = TP/TP+FP ---------- (3)

The F-measure is often referred to as the F1-score as follows, and it measures the mean value of precision and recall:

F1 Score = 2*(Recall * Precision) / (Recall + Precision) ----------(4)

Experimental Analysis

The proposed Gaussian NB algorithm is compared with existing ML algorithm such as SVM, LR, DT, and KNN. The proposed Gaussian NB archives 97.20% of accuracy, 91% of Recall, 92.22% of precision and 90.11% of F-Score.

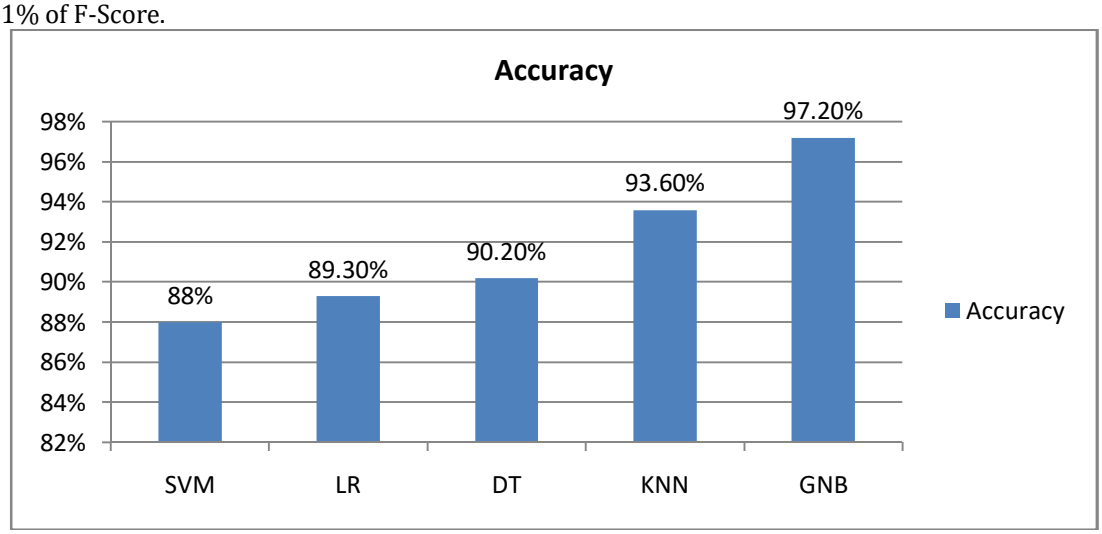

Figure 3: Heart disease prediction Accuracy Comparison

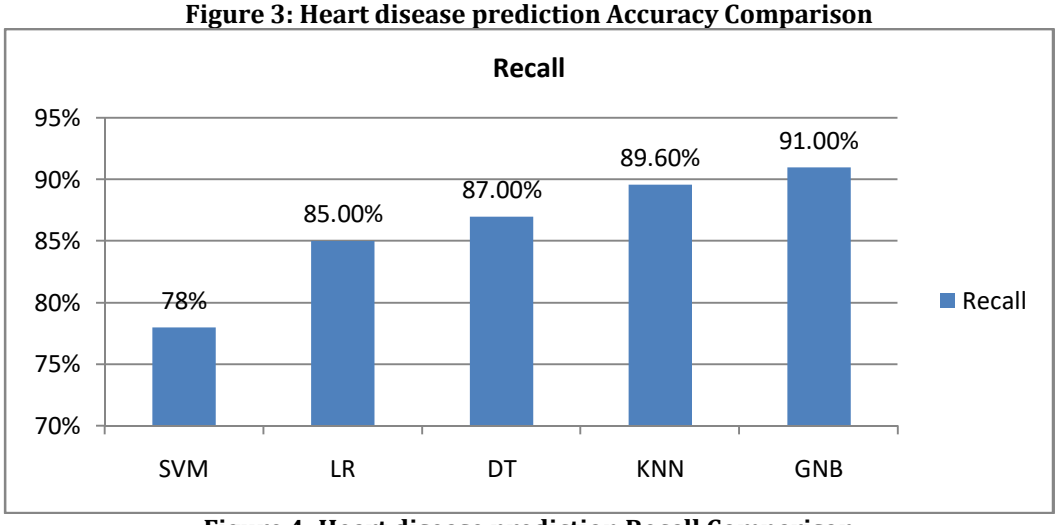

Figure 4: Heart disease prediction Recall Comparison

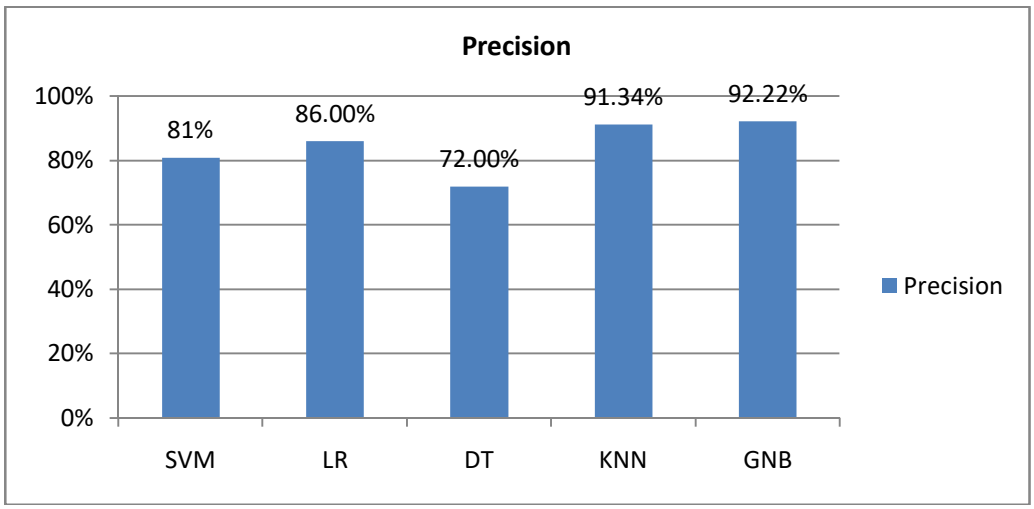

Figure 5: Heart disease prediction Precision Comparison

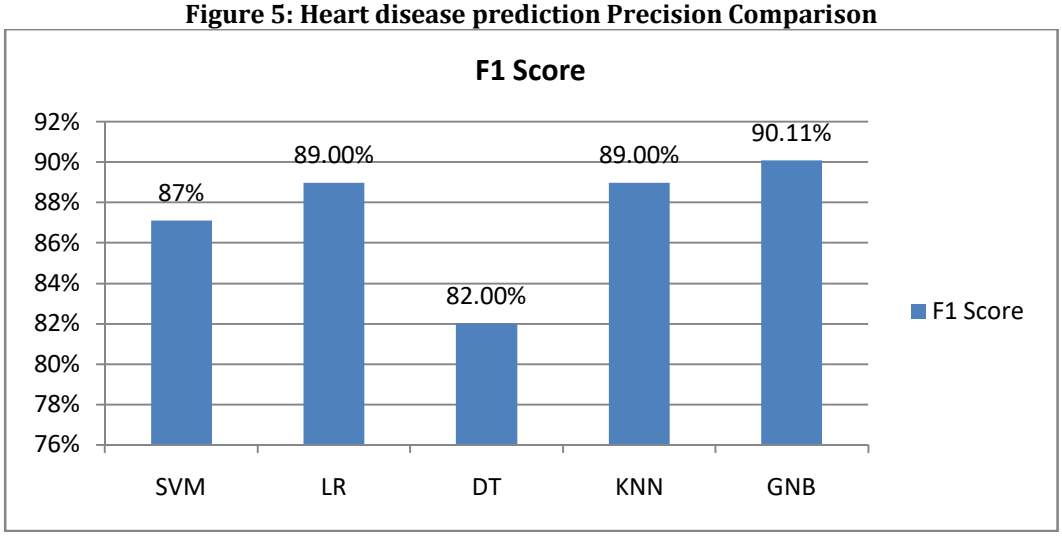

Figure 6: Heart disease prediction Precision Comparison

CONCLUSION

The primary objective of this article is to examine the effect of feature selection approaches on the accuracy of heart disease prediction. This analysis was conducted against a collection of distinctive features extracted from frequently used Cleveland heart disease datasets available at the University of California, Irvine using various feature selection algorithms. In this work, IoT based Healthcare system for diabetics disease prediction is proposed using Gaussian Naïve Bayes machine learning algorithm. The proposed Gaussian NB algorithm is compared with existing machine learning algorithm such as SVM, LR, DT, and KNN. From the experimental result analysis, the proposed Gaussian NB algorithm achieves 98.78% accuracy of Heart disease prediction. In the future, real-time medical datasets gathered from different countries can be used to model development. This could enhance the performance with improved accuracy for heart disease prediction.

REFERENCES

1. Rahaman A, Islam M, Islam M, Sadi M, Nooruddin S. (2019). Developing IoT based smart health monitoring systems: a review. Rev Intell Artif.33:435–40.
2. Hasan M, Islam MM, Zarif MII, Hashem MMA.(2019). Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet Things. 7:100059.
3. Nooruddin S, Milon Islam M, Sharna FA.(2020). An IoT based device-type invariant fall detection system. Internet Things.9:100130.
4. Islam M, Neom N, Imtiaz M, Nooruddin S, Islam M, Islam M. (2019). A review on fall detection systems using data from smartphone sensors. Ingénierie des systems. **DOI:** https://doi.org/10.18280/isi.240602
5. N. Manh Khoi, S. Saguna, K. Mitra, and C. Ahlund, (2015).“IReHMo: An efficient IoT-based remote health monitoring system for smart regions,” in Proc. 7th Int. Conf. E-health Netw., Appl. Services (HealthCom), pp. 563–568, doi: 10.1109/HealthCom.2015.7454565.
6. G. Rathee, A. Sharma, H. Saini, R. Kumar, and R. Iqbal, (2019).“A hybrid framework for multimedia data processing in IoT-healthcare using blockchain technology,” Multimedia Tools Appl., vol. 2, pp. 1–23.

7. Liaqat Ali, Atiqur Rahman, Aurangzeb Khan, Mingyi Zhou, Ashir Javeed, and Javed Ali Khan, (2019). "An Automated Diagnostic System for Heart Disease Prediction Based on Statistical Model and Optimally Configured Deep Neural Network", IEEE Access, vol. 7, pp. 34938-34945, DOI: 10.1109/ACCESS.2019.2904800
8. Bo Jin ,Chao Che, Zhen Liu, Shulong Zhang, Xiaomeng Yin, And Xiaopeng Wei, (2018). "Predicting the Risk of Heart Failure With EHR Sequential Data Modeling" ,IEEE Access.
9. Ashir Javeed, Shijie Zhou, Liao Yongjian, Iqbal Qasim, Adeeb Noor, Redhwan Nour, Samad Wali And Abdul Basit. (2017). "An Intelligent Learning System based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection" , IEEE Access.
10. Abhay Kishore, Ajay Kumar, Karan Singp, Maninder Punia, Yogita Hambir. (2018). Heart Attack Prediction Using Deep Learning, International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 04:90-94
11. R. Bhuvaneswari and K. Kalaiselvi, (2012). "Naïve Bayesian Classification Approach in Healthcare Applications", International Journal of computer Science and Telecommunication", vol. 3, no. 1, pp. 106-112.
12. M. S. Satu, F. Tasnim, T. Akter and S. Halder,(2018). "Exploring Significant Heart Disease Factors based on Semi Supervised Learning Algorithms," International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), Rajshahi, pp. 1- 4.doi: 10.1109/IC4ME2.2018.8465642
13. S. Babu et al., (2017). "Heart disease diagnosis using data mining technique," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, pp. 750-753.doi: 10.1109/ICECA.2017.8203643

CITATION OF THIS ARTICLE

P. Jagadeesan, M.Vedaraj., K.Dhanalakshmi, D.Surendiran Muthukumar. IoT Based Heart Disease Prediction using Machine Learning techniques. Bull. Env.Pharmacol. Life Sci., Vol 12 [4] March 2023: 269-274