Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 822-831
Received 15 April 2021; Accepted 05 May 2021.

# A Novel Analysis of Common Heart Disease among Women - A Survey

**Bharathidasan. G[1], Dr. G.V. Sriramakrishnan[2]**
[1]Research Scholar, Department of Information Technology, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai
[2]Department of Information Technology, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai
[1]bharathidasanastro@gmail.com,
[2]greatsri8@gmail.com

**Abstract**
Cancer is the terribly deadliest sickness and hazardous disease. These days' individuals all through world are experiencing different kinds of cancer growth because of heredity impacts, contaminations, poor natural components and helpless way of life. Individuals don't have any mindfulness on Cancer forecast and finding and furthermore it was an unpredictable undertaking for people during the beginning phases. Be that as it may, these days all cancer is reparable utilizing progressed framework innovations. Machine learning approach assumes an essential part into the qualities and cycles associated with cancer recognition and finding. This approach gives a characteristic method to analysts on concentrating singular patient's cancer treatment by giving all the vital information. Likewise Machine learning calculations are utilized to discover the development of disease cells and its connected treatment at level best. This paper examines the different machine learning strategies to anticipate and analyze the regular kind of cancer among ladies. Numerous researchers have analyzed different cancer datasets utilizing a distinctive procedure. Most methods give high exactness results on diagnosing and foreseeing the cancer utilizing various tools. The target of this paper is to figure out all the significant data which are essential for cancer disease prediction and further research.
**Keywords**:- Decision Tree, SVM, KNN, Random Forest

## 1. Introduction

The physical structure of human is comprised of millions of cells. The structure of the body is very much enacted and constrained by these cells by taking energy from food and plays out the particular capacity in human body. Cancer is made by the unusual disease cells in human body [18]. Peoples are experiencing different sorts of cancer disease around the world. Cancer cells can be generated in any piece of the human body and once affected that may spreads into its various parts of the body by making new cancer disease cells. In a human body when the body needs certain energy new vigorous cells might be made. During this cycle when the human body gets affected with cancer this cell production cycle breaks down [17].

"Warburg's demise in 1970 and the disclosure of oncogenes in 1971, most Cancer specialists see disease as a hereditary sickness as opposed to a metabolic infection". The genetic variations that can inherit cancer from our parents can cause cancer [18]. The natural factors like smoking, substance hypersensitivity, bright radiation likewise cause cancer. When the human body is influenced with cancer a few changes will happen naturally and the connected cells gets

transformations in DNA than typical cells. This adjustment in human body is called metastatic cancer growth. For instance spreading breast cancer growth into lung might be treating as metastatic breast cancer. So noticing tissue changes is a basic test in a clinical field. "According to review from cancer institute the event of cancer growth among people is 442.4 per 100,000 every year. The disease death rate among people is 158.3 per 100,000" [18].

Cancer can be of a wide range of types and it named from the tissues where the disease comes from. Among numerous cancer growths there exists some regular cancer disease which affects both men and women. For instance, Breast cancer growth begins from bosom tissue in milk pipes and the lobules. One can understand the breast cancer indications from shape change, a knot on a bosom, shading change in bosom fix skin. Prior determination of bosom cancer prevents hurtful impact[16]. Lung cancer is created in the lungs and can spread to different pieces of the body. It tends to be arrangedinto two types as "small cell and non-small cell lung cancer" [21]. The therapy of these malignant growths is not quite the same as each other. Non-small cell lung cancer is treated with a medical procedure and small cell tumors are treated with radiation and some high level strategies [19]. Smoking is one of the fundamental purposes behind lung cancer. This cancer is a normal sort among people. Thus, Lung disease discovery and prevention in the beginning phase is important to avoid cancer death [20].

Skin disease begins in skin cells because of unusual development. Skin disease is additionally evolved because of abundance daylight or bright beams on a human skin. This disease changes the color of the skin. Yet, compared with different cancers skin cancer can be preventable earlier using some precautionary measures. Cervical cancer created when cells change at the lower part of the uterus. It happens among ladies beyond 30 years old. Cervical cancer spreads gradually and prior therapy is feasible to prevent the disease. This cancer is effectively treatable by Utilizing Pap test. Screening tests are vital to analyze this cancer growth. Brain cancer happens in the actual cerebrum and because of strange change in the brain cells. Developing of affected cells spreads everywhere on the parts of the human body. It causes a leading death in around the world. The manifestations of brain cancer incorporate extreme migraines, vision misfortune, shading variety and disarray.

Any type of cancer disease is hazardous to human existence. So the need emerge to identify prior to analyze the infection. Machine learning is perhaps the most premier applications in the field of clinical application particularly used to diagnose the beginning phases of a wide range of coronary illness. Machine learning algorithms assist with understanding the advancement of cancer cells and what medicines are needed to diagnose the illness. Machine learning approaches applied on available biological cancer data and using sophisticated tools cancer treatment is carried out. Machine learning is a subfield of Artificial Intelligence and incorporates all the most impressive calculations to learn malignant growth information, discovers the pattern, trains the data and predicts the sickness too. By utilizing different learning strategies, these algorithms give a high exactness in foreseeing the disease result with great execution.

In this article we analyzed different types of common cancers like Breast, Lung, Cervical, Skin and Brain which are interrelated among women. The various Machine Learning Algorithms that are suitable for medical application from many researchers has also been noted.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 822-831
Received 15 April 2021; Accepted 05 May 2021.

## 2. Literature Review

Mumine, Kaya Keles [1] proposed a closer review of Machine learning calculations for earlier predictions and the discovery of breast cancer disease.The forecast has been done utilizing antenna database with four features. The features of antenna database were extracted through WEKA tool. The classification algorithms such as random forest, bagging, random committee, simple CART and IBK were utilized to examine the exactness utilizing 10-fold cross-validation. The efficacy of random forest algorithm was higher than others in predicting the cancer more accurately. The algorithm yields 92.2% accuracy using binary classification method.

Sara Alghunaim et al. [2] investigated the cancer dataset to foresee the event of breast cancer utilizing a productive structure of WEKA and Spark. The model was developed using classification algorithms such as the support vector machine (SVM), decision tree, and random forest for breast cancer analyses. Thesetypes of classification algorithms were applied to the huge dataset of Gene expression (GE) and DNA methylation (DM) for breast cancer classification.The comparative study was conducted using the available dataset to predict the best outcome by analyzing accuracy and error rate of all classifiers. Among all algorithms SVM predicts the cancer with an accuracy of 99.68% and lower error rate.

Muhammad Hammad Memon et al. [3] suggested a framework to track down the beginning phases of breast cancer growth among ladies. They utilized Support Vector Machine (SVM) classifier to group the bosom disease as malignant or benign. Then the recursive feature selection algorithm and predictive model were employed inWisconsin Cancer dataset to diagnose and increase classifier performance. The dataset was split into training data and testing data to evaluate classifier performance. The model was performed on subset of 32 features using SVM classifiers like linear, RBF, polynomial, and sigmoid. The SVM linear kernel provides the best accuracy of 99% as an optimal solution.

Priyanka Israni [4] proposed BCD model to analyze the breast cancer in the beginning phase. The model was made utilizing the support vector machine and dissected the Wisconsin Breast Cancer data set with 32 ascribes. A 10-fold cross-validation methodwas utilized to improve performance and overfitting data. Principal component analysis was used to reduce the complexity of feature space indiagnosising the cancer. The model was compared with other models such as Decision trees, Ada Boost, Random Forest, Naïve Bayes using the evaluation parameters F1 score, accuracy, Roc curve. The proposed BCD model accomplished with high accuracy of 98.1% of detecting the breast cancer.

Subrato Bharati et al. [5] were demonstrated the presence of breast cancer among ladies utilizing data mining classifiers. The algorithms like Naïve Bayes, Random Forest, Logistic Regression, Multilayer Perceptron, KNN were employed on UCI Machine learning repository dataset to diagnose the disease. The dataset was analysed using 256 instances with 10 attributes. The performance of the algorithms was analyzed utilizing the boundaries of kappa measurements, TP rate, FP rate and accuracy. K – Nearest Neighbors gives the most noteworthy exactness of 97.90% of diagnosing the disease than different algorithms.

Lilla Boroczky et al. [6] described the performance of false positive reduction framework to discover the cancer disease in the lungs using computer aided design algorithms. The genetic algorithm and future subset selection technique was utilized to choose the required future from the future pool. The proposed strategy has combined with Support vector machine (SVM) classifier and analyzedperformance of the lung nodule database. Among 23 features 10 most related futures were selected for the calculation of optimal solution. The SVM classifier classified the disease with 100% sensitivity and 56.4% specificity.

Yotong Xie et al. [7] proposed a model to recognize the early identification of cellular breakdown in the lungs either malignant or benign. A Multi-View knowledge based collaborative deep model(MV-KBC) was utilized to discover the dangerous nodules from chest CT information. The knowledge based collaborative sub model partitions the data set into nine fixed perspectives for classification and training. The strategy has been used on the LIDC-IDRI database. The model classifies the lung nodule with a precision of 91.60%.

Ozge Gunaydin et al. [8] recommended the significance of nodule examination for an accurate lung cancer treatment. The point of their exploration is the earlier identification of lung disease. The different machine learning techniques, for example, Principal Component Analysis, KNN, Support Vector Machine, Naïve Bayes, Decision Trees and Artificial Neural Networks has been utilized to identify the anomalyon Standard Digital Image Database. The precision has been compared with all strategies and Decision tree technique demonstrates the best strategy in foreseeing the malignant growth with the high accuracy of 93.24%.

Kurnianingosih et al. [9] built up a technique for automatic detection of cervical malignant growth among ladies. A Pap smear cell image was well classified and segmented with the assistance of Herlev Pap smear dataset. A cell image was divided utilizing Mask R-CNN segmentation algorithm. The segmentation process can be measured by precision, recall, and specificity and classification can be measured by F1 score, accuracy and sensitivity.

Yu Peng et al. [10] recommended a productive strategy for cervical cancer diagnosis utilizing Pap tests. The technique uses a productive ellipse fitting algorithm which analyses and fragmented cervical nuclei clusters. Then the related features are extracted from the image and classified accurately. The image classification and clustering process were completed using C4.5 algorithm with related features. The classifier performance wasanalysed in terms of accuracy, sensitivity and eventually 97.8% classification precisionwas achieved.

Anjali Deswal et al. [11] suggested a model for early detection of cervical diseases by analysing at certain attributes inwomen. Screening and diagnostic tests such as Pap smear test and Human papillomavirus were used to detect the infection cells. Random Forest Classifier was used on UGI dataset. The different existing human boundaries like smoking, age, sexual accomplice have been examined and carried out utilizing Eclispse and Weka tool. The proposed framework effectively analyzed and diagnosed by estimating different measurements F-measure, Recall, Precision and Confusion matrix.

Anurak kumar verma et al. [12] applied a novel strategy for skin malignant growth utilizing different machine learning strategies. The algorithms like Classification and Regression Trees,

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 822-831
Received 15 April 2021; Accepted 05 May 2021.

Support Vector Machine, Decision Tree, Random Forest and Gradient Boosting were utilized on dermatology dataset to classify different skin related infections. Finally an ensemble method was applied to predict the skin malignant growth as an individual classifier. The ensemble method outperforms well while examining every one of the classifiers performance. Ensemble method predicts the disease with the high accuracy of 98.64%

Anurag Kumar Verma et al. [13] classified different classes of skin disease to predict skin cancer utilizing machine learning algorithms. The model was proposed with ensemble method using Bagging, Adaboost and Gradient Boosting to classify various types of skin diseases. Finally feature selection strategy was applied to anticipate the skin disease precisely. All the selected features were compared with all classifiers for performance comparison. The better performance has been achieved when ensemble method and feature selection method applied collectively on dermatology dataset than an individual classifier. It predicts the skin sickness with the high precision of 99.68%.

SN Qasem et al. [14] proposed a model to foresee brain malignant growth from a big dataset of clinical images. The proposed strategy includes pre-processing and segmentation processing utilizing watershed segmentation on clinical images. Finally KNN classification algorithm applied to predict the brain disease. The classifier classifies with the precision of 86%.

Mueez Ahmed et al. [15] proposed a new model for detecting tumor in obese and non- obese patients using decision tree classification methods. The proposed model utilized C4.5 classifier to diagnose brain cancer precisely. The database of SKMCH & RC, Lahore was used for analysis. Every one of the clinical records of obese and non- obese patients were broke down by the classifier and concentrate the required data.

## 3. Methodology

In a real-life environment all types of cancer among women require exended treatment and the results should be monitored periodically. These treatment results should be analysed carefully and accordingly the disease should be diagnosed. Machine learning concepts are very useful in analysing such results.In this research, we discuss various machine learning algorithms for predicting and diagnosing different types of cancer.

### *3.1 Decision Tree*

Extricating significant data from enormous measure of information from real world applications is crucial in data analysis process. Classification accomplishes the work proficiently in classifying and predicting the classes precisely. Classification employs two-step measure as learning and classification. In view of exactness of the classification data ought to be classified. It is ought to be deciphered as in equation 1.

$$y = f(X) \quad (1)$$

where y predicts the class label y against a given tuple x. Decision tree is the main method in arranging the data by framing tree like structure. The tree begins with root node and closures with leaf node which holds class forecast. Decision tree handles multidimensional information

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 822-831
Received 15 April 2021; Accepted 05 May 2021.

effectively and yields high precision. The decision tree has number of algorithms and utilized in numerous real world applications to extract the meaningful information. Different attribute selection measures are utilized during decision tree development which is utilized to partition the tuples. The most well-known algorithms are ID3, C4.5 and CART. Measures for selectingattributes such as information gain, Gini index or Gain Ratio can be used depending on the circumstance. These algorithms are used to extract and classify all types of cancer related data productively.

### *3.2 Support Vector Machine (SVM)*

This supervised machine learning algorithm classifies linear and non-linear data efficiently. It is used for classification and numeric prediction problems. SVM transforms the original data into n-dimensional space using a non-linear mapping into classes. This can be accomplished by making a decision limit what isolates one class from another. This decision limit is called hyperplane. This hyperplane is effectively discovered by SVM utilizing preparing tuples. More number of decision boundaries exists in n-dimensional space in which the strategy need to discover the best one to arrange the labels. In view of the number of features the dimension of the hyperplane was characterized. Maximum distance between the data points signifies maximum margin of the hyperplane. SVM is highly accurate and less prone to over fitting.

SVM can be classified in two ways. Inlinear SVM,the information can be classified into two classes by straight line. In non-linear SVM,the information cannot be classified by straight line. The classification of data using the SVMis shown in figure 1.

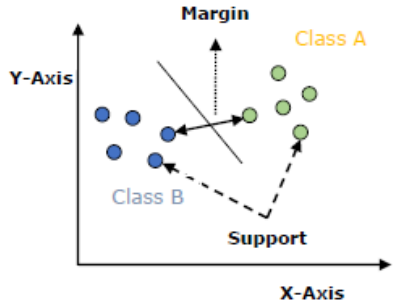


**Figure 1 Data Classification using SVM**

### *3.3 k-Nearest-Neighbor Algorithm*

It is a supervised machine learning classifier. It functions admirably with huge training data sets. Nearest-neighbor classifier works based on similarity that compares a given data set with training data set. All training data sets are put away in n-dimensional space. For a given data set, k-nearest-neighbor searches the pattern for k training data that are similar to it. These training data are called k-nearest-neighbor for the new data set. This distance is estimated by Euclidean distance and it is shown in equation 2as,

$$dist(X1, X2) = \sqrt{\sum(x_i - x_j)^2} \quad (2)$$

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 822-831
Received 15 April 2021; Accepted 05 May 2021.

where x1 and x2 are the points of two datasets. KNN is used for classification as well as a prediction. The algorithm first stores the training dataset and when the new data arrives, classifies them according to similarity. So it is also called as a lazy learner algorithm.

**Naïve Bayes**

This classifier follows the standards of statistical ideas and used to predict class membership in a particular gathering. The classification is based on Bayes hypothesis. Naïve Bayes is a group of algorithms which shares a single common principle. Classifier predicts all the more precisely on object based on conditional probability function. Classifier assumes the occurrence of specific features is an independent of other features. Bayes theorem is interpreted mathematically as shown in equation 3.

$$P(X/Y) = \frac{P(Y/X)P(X)}{P(Y)} \quad (3)$$

WhereX,Y are events and P (X|Y) is a posteriori probability. Naïve Bayes classifier can be applied to huge datasets and its performance can be measured by accuracy, precision and other related measures.

***3.4 Random Forest***

This algorithm is considered the best one for classification, prediction and other learning related concepts and follows an idea of ensemble learning. The classifier contains the number of decision trees by creating a group which is called a forest. More number of trees in a forest gives higher accuracy in result. The algorithm outperforms well all the classification task and numeric prediction. Also algorithm proficiently handles missing qualities and evades the issue of over fitting. The algorithm first selects n features randomly out of m features and the root node can be found in n features by using best splitting criterion. The process can be repeated until random forest creation. The algorithm examines each tree for prediction and the final prediction was taken based on the majority voting. Random forest algorithm functions admirably in clinical application particularly in foreseeing the cancer disease.

***3.5 Gradient Boosting***

It is a powerful classifier on classification and prediction problems. This algorithn is used to predict a model in step by step process and considers the best one than Random forest.It combines weak classifiers to create powerful classifier which produces a powerful classification model. The precision of an algorithm can be gotten either by applying feature engineering or by boosting algorithms. The procedure applies the output of weak classifiers sequentially to minimize the loss function which depends on the problem being solved. The technique uses gradient decent function to diminish the loss. Over fitting of training data carried out efficiently and thereby improves the performance.

**4. Discussion**

This study reviews different types of Classifiers to predict and diagnose cancer. Table 1 provides a summary of various techniques, datasets and methods which are used to predict and diagnose the cancer.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 822-831
Received 15 April 2021; Accepted 05 May 2021.

**Table1. Comparing Machine Learning Techniques in cancer prediction**

| Algorithms Used | Dataset Used | Best Accuracy Method | Accuracy | Reference |
|---|---|---|---|---|
| Random forest, Bagging, Random Committee, CART and IBK | Antenna Database | Random Forest uses 10- fold cross alidation | Accuracy is 92.2% | [1] |
| Support vector machine (SVM), Decision tree, and Random forest | Gene Expression &DNA methylation Repository | Support Vector Machine | Breast Cancer accuracy is 99.68% | [2] |
| Support Vector Machine (SVM) | Wisconsin Diagnostic Cancer dataset | SVM linear kernel | Breast Cancer accuracy is 90% | [3] |
| Support Vector Machine (SVM), Decision trees, Ada Boost, Random Forest, Naïve Bayes | Wisconsin Diagnostic Cancer dataset | Support Vector Machine (SVM) uses 10-fold cross validation | Accuracy is 98.1% | [4] |
| Naïve Bayes, Random Forest, Logistic Regression, Multilayer Perceptron, KNN | UCI Machine learning repository | K – Nearest Neighbors uses kappa measurements | Breast Cancer accuracy is 97.90 | [5] |
| computer aided design algorithms | Lung nodule dataset | SVM & future subset selection used | Lung Cancer sensitivity 100% and specificity 56.4% | [6] |
| multi-view Knowledge based collaborative deep model | LIDC-IDRI database | MV-KBC | Lung Cancer accuracy is 91.60%. | [7] |
| Principal Component Analysis, KNN, Support Vector Machine, Naïve Bayes, Decision Trees and Artificial Neural Networks | Standard Digital Image Database | Decision tree | Lung Cancer accuracy is 93.24% | [8] |
| ellipse fitting algorithm, C4.5 | Cervical nuclei clusters | C4.5 with PAP test | Cervical Cancer accuracy is 97.8% | [10] |
| Classification and Regression Trees, Support Vector Machine, Decision | Dermatology dataset | Ensemble method | Skin Cancer accuracy is 98.64% | [12] |

| | | | | |
|---|---|---|---|---|
| Tree, Random Forest and Gradient Boosting | | | | |
| Bagging, Adaboost and Gradient Boosting | Dermatology dataset | Feature selection | Brain Cancer accuracy is 99.68% | [13] |
| Watershed segmentation Algorithm | Clinical images | KNN | 86% | [14] |

## 5. Conclusion

In this research, we analyzed various Machine learning algorithms for predicting common cancers in females. We found the best algorithm with greatprecision for predicting each kind of cancer. We summarized all the existing Machine learning algorithms and extracted the keynote information to further research for cancer prediction and diagnosis. We reviewed different types of cancersof the Breast, Lung, cervical, Skin and Brain. We analyzed various methods like Decision Trees, Support Vector Machine, KNN, Naïve Bayes, Random Forest and Gradient Boosting to predict cancer. All the methods used different types of dataset, tools and techniques to get accurate results. In the future work the dataset would have to be elaborated from the real world for each type of cancer.

## References

[1] M. K. Keles, "Breast cancer prediction and detection using data mining classificationalgorithms:Acomparativestudy,"*TehničkiVjesnik*,vol.26, no. 1, pp. 149–155, 2019.

[2] S. Alghunaim and H. H. Al-Baity, "On the scalability of machine- learning algorithms for breast cancer prediction in big data context," *IEEE Access*, vol. 7, pp. 91535–91546,2019.

[3] M. H. Memon, J. P. Li, A. U. Haq, M. H. Memon, and W. Zhou, "Breast cancer detection in the IOT health environment using modified recursive feature selection," *Wireless Commun. Mobile Comput.*, vol. 2019, pp. 1–19, Nov. 2019.

[4] P. Israni, ''Breast cancer diagnosis (BCD) model using machine learning,''*Int.J.Innov.Technol.ExploringEng.*,vol.8,no.10,pp.4456–4463, Aug. 2019.

[5] S. Bharati, M. A. Rahman, and P. Podder, "Breast cancer prediction applying different classification algorithm with comparative analysis using WEKA," in *Proc. 4th Int. Conf. Electr. Eng. Inf. Commun. Technol. (iCEEiCT)*, Sep. 2018, pp.581–584.

[6] Lilla Boroczky, Luyin Zhao and K.P.Lee , "Future Subset Selection for Improving the Performance of False Positive Reduction in Lung Nodule CAD," IEEE Transactions on Information Technology in Biomedicine, Vol. 10, No. 3, July 2006.

[7] Yutong Xie, Yong Xia, Jianpeng Zhang, Yang Song, Degan Feng, Michael Fulham, and Weidong Cai, "Knowledge-based Collaborative Deep Learning for Benign-Malignant Lung Nodule Classification on Chest CT," IEEE Transactions on Medical Imaging, Vol. 38, no. 4, pp. 991-1004, April 2019.

[8] Ozge Gunaydin, Melike Gunay, Oznur Snegel, "Comparison of Lung Cancer Detection Algorithms," in Proc. Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), April 2019.

[9] Kurnianingsih, Khalidhameds.allehaibi, Lukitoedinugroho, Widyawan, lutfanlazuardi, Anton satria,Prabuwono,andTeddymantoro,SegmentationandClassificationofCervicalCellsUsingDeepLe

arning," IEEE Access, vol. 7, no. 116925 – 116939, 2019.
[10] Yu Peng,MiraPark,MinXu,Suhuai Luo, Jesse S.Jin,Yue Cui andW.S. Felix WongLeonardoD.Santos, "Clustering Nuclei Using Machine Learning techniques," IEEE/ICME International conference on Complex Medical Engineering", pp. 52-57, 2010
[11] AnjaliDeswal,SanjeevDhawan,KulvinderSingh,"ATechniqueForDeterminingTheEarlyDetection ForCervicalCancer,"5thIEEEInternationalConferenceonSignalProcessing,ComputingandControl( ISPCC2k19), pp. 260-264, 2019.
[12] AnuragKumarVerma,SaurabhPalandSurjeetKumar"ClassificationofSkinDiseaseusingEnsembleD ataMining Techniques," AsianPacificJournalofCancerPrevention, vol20, no. 1887-1894, 2019.
[13] AnuragKumarVerma,SaurabhPalandSurjeetKumar,"Comparisonofskindiseasepredictionbyfeatur eselectionusingensembledataminingtechniques," Informatics in Medicine Unlocked, vol. 16, 2019.
[14] Sultan Noman Qasem, Amar Nazar, Attia Qamar, ShahaboddinShamshirband andAhmadKarim A, "LearningBasedBrainTumorDetectionSystem," CMC,vol.59,no.3,pp.713-727,2019
[15] Mueez Ahmad and Abdul Aziz, "Early Detection of Brain Cancer in Obese and Non-Obese Patients by using Data Mining Techniques," Indian Journal of Science and Technology, vol. 10, no. 24, 2019.
[16] https://en.wikipedia.org/wiki/Breast_cancer
[17] https://www.cancer.gov/
[18] https://www.who.int/health-topics/cancer#tab=tab_1
[19] https://www.wcrf.org/dietandcancer/cancer-trends/worldwide-cancer-data
[20] https://gco.iarc.fr/
[21] https://www.cancer.org/