International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No. 39, April 2022 – www.ijais.org

Determining the Impact of ICT Devices on our Environment and Health using Machine Learning Techniques

Ugochi Adaku Okengwu
Department of Computer Science
University of Port-Harcourt Choba, Rivers State, Nigeria

Stanley Ziweritin
Estate Management & Valuation
Akanu Ibiam Federal Polytechnic, Unwana

ABSTRACT
This paper explores the known but silent effects of ICT devices on our environment and health. The use of ICT devices is very crucial and paramount to our daily activities but the negative and positive effects especially with the new idea of Internet Of things (IOT) that will make the sensors part of our daily routine till we go to bed. The negative effect ranges from exposure to toxic compounds, high energy consumption due to high reliance on ICT devices, exposures us to non-thermal radio frequency radiation from Wifi, cellular and many more. The positive effect of ICT and its associated devices are too numerous to list which have been adopted in every spheres of human endeavour. The aim is to build a K-nearest neighbor and random forest technique to access the impact of ICT devices in detecting human heart diseases caused by ICT radiations. This will help reduce the stress of searching and waiting with hope for specialists to look at results of images when diagnosis are performed by lab scientists in determining whether the patient is fine or have heart disease. This contributes positively to the healthcare delivery system and promotes our next level of digital economy in the society at large because of the limited number of medical doctors. We adopted the k-fold cross validation test to have a better classification report. The KNN produced 90% cross validation test accuracy which was observed to be higher than the random forest with 85.71% cross validated accuracy.

Keywords
ICT devices, health, Internet of Things(IOT), KNN, Decision tree

1. INTRODUCTION
Information communication technology(ICT) employs wireless communications for mobile users with wireless fidelity(WIFI) and cellular network supported devices[1],[2]. The fifth generation(5G) is a virtual reality considered as the base technology for Internet of Things(IoT) where machines communicates and help human interpret results of diagnosis[3]. The electromagnetic field(EMF) radiations are harmful to human health based on biological stress response[4],[5],[6]. The concern has been the risk of cancer, cataracts, and DNA degeneration that may also lead to heart problem among people living near based stations or ICT devices([7],[8],[9],[10],[11]. Recently, Prasad *et al.*[12], Stein *et al.*[13] and Szmigielski[14] recorded some proliferative and tissue injury effects of mobile users exposure to 2450 MHz radiation in HL-60 which are liable to have cancer related problem. We are exposed to EMF's of radiations in every sphere of human life around ICT within the environment[15]. The sources of EMF's have increased in modern days with the advancement of technology in adopting 5G technology. The spectrum of EMF's source extends from magnetic to static fields, radio frequency, visible light to gamma rays and infrared[16]. The non-ionizing radio signals cannot directly cause changes in chemical bonds but the ionizing magnetic spectrum/signals can cause changes that may lead to tissue damage and cancer related disease([17],[18]. There are several types of supervised, semi-supervised, reinforcement and unsupervised learning techniques employed to train and detect human disease[19],[20],[21]. The evolution of AI has changed the entire 21th century in terms of technological advancement. The exponential growth of machine learning under AI is the most trending field of study that drives an innovative system for digital economy[22],[23]. Supervised learning algorithms are used to solve regression and classification problem, unsupervised can be employed to solve association rule mining and clustering problems[24] . Supervised meaning to oversee or derive a particular activity and make sure is done correctly. In this type of learning the machine learns and to be guided through a given training data input(s) labeled to produce output. The Unsupervised means to act without anybody supervision or direction and trained on unlabeled data without any guidance[25]. Here the data is not labeled, there is no guidance and the machine have to figure out the dataset given and find out all the hidden patterns about the target(output). Reinforcement means to encourage a pattern of behavour with no predefined data. The agent interacts with its environment by producing actions and discovers errors or rewards. The agent learns through trial and error with punishment and adopts reward system of feedback[26]. The Semi-Supervised learning employs the idea both sizable(small) amount of labeled data and large amount of unlabeled data for training purpose[27]. The semi-supervised learning is the combination of supervised that has labeled data and unsupervised learning with unlabeled data[28]. The reinforcement learning is also known to be learning by critics. The algorithm will be given the task of making different possibilities to find the correct answer from the task of giving a wrong answer[29]. We need specialists (doctors) when diagnosis are performed to look at the images in determining whether the patient is fine or have heart disease. We are automating the initial screening process of doctors detecting human heart disease from results of diagnosis because of the limited number of specialists(doctors). The RF and KNN algorithms are used to determine whether a particular patient is having heart disease or not without waiting for a specialist.

The aim is to build a K-nearest neighbor(KNN) and random forest(RF) model as an ICT tool for detecting human heart diseases that occurs in nature. We are proposing to train and test the model using Kaggle online heart disease dataset. It contributes significantly to the healthcare delivery system and promotes our next level digital economy in the society at large. The k-fold cross validation test will be adopted to produce a better classification report. Python in ANACONDA will be used for simulation with some diagnostic tools to measure the performance metrics of both models.

This paper is divided into sections as stated: section one contained the introduction; Section two presents a brief review of about previous methods to the study area and the gap in exploring the proposed model; Section three, introduces the materials and methods with the different techniques adopted and materials used for developing the model; Section four, focuses on the results and detailed discussion of results; Section five presents the conclusion.

2. RELATED WORKS

Omotosho *et al.*[30]; carried out a research to ascertain the growth rate of ICT users with e-health strategies of wireless technology. From findings advised users to invest on ICT for Internet of Things(IoT) without considering its health related challenges where machines can communicate and expose human to radiations. They also proposed to increase the number of high frequency for powered-base stations of ICT device as a base technology for 5G wireless communications. The result reveals that human exposure to such higher frequencies of WIFI and cellular networks that range from 6-100Hz millimeter waves can cause 58% negative effect on human health. Breckenkamp *et al.*[31]; proposed a study on "the effect of microwave or EMF radiations" emitted by ICT devices(cell phones) on human chromosomes. The study recorded some dose dependent effects such as the risk of having cancer, heart disease and brain damage for mobile users for those living close to base stations[32]. Lacohe, *et al.*[33]**;** developed a multi-layered model to interpret possible health challenges involved in the use of ICT devices. The emission of EMFs causes adverse effect on human health. They stated that the technology behind 5G is to use high radio frequencies in accommodating and transporting large amount of data at high speed across the globe. The research shows that pulse EMF radiations are more dangerous to human health than pulse of non EMF's and signals of 5G radiations which can cause heart disease, cancer and damage in human DNA.

Prasad *et al.*[34]; developed an expert system using context sensitive auto-associative memory neural network(CSAMM), Bayesian nets, DT and PSO to carry out diagnosis on human disease called asthma. The dataset was collected through questionnaires with 25-patients clinical data for analysis. The CSAMM produced 84.32%, followed by PSO(84.16%), BN(81.17%) and C4.5 type of DT with (83.83%) accuracy. The general performance was poor in terms of accuracy and speed which required more clinical data and training time to perform well as required. Ba-Alwi and Hintay [35]; developed a model using data mining classifiers using Naive Bayes(NB), NN, j48 and DT to diagnose and detect the disease called hepatitis. The NB was recorded to have the best classification with 96.52% accuracy rate. A combination CART, ID3 and C4.5 types of DTs are employed in diagnosing the disease known as hepatitis[36]. The CARD gave 83.2%, ID3(64.8%) and C4.5(71.4%) metrics of accuracy which was not encouraging in terms of accuracy and training time. Vijayarani and Dhayanand[37]; proposed Naive Bayes and SVM classification techniques to diagnose and detect live disease. The NB performed better and produced 79.66% metrics of accuracy compared to the SVM with 61.28% rate which was below average. Rajeswari [38]; carried out an analysis on human liver disorderliness using FT-tree, Naive Bayes(NB) and K-star techniques in classifying data instances. From the results; NB produced 96.52%, FT-tree(97.10%) and k-star gave 83.47% accuracy.

3. MATERIALS AND METHODS

In this research; we are considering the popularly known and widely accepted KNN and RF classifiers in diagnosing heart disease. How the KNN and RF can be trained to carryout cross-validation test to detect heart disease with the symptoms in predicting the target. The reason is that, KNN and RF training time complexity are very fast and makes no assumptions about training dataset as very crucial in non-linear data cases with better accuracy rate. RF is the highest in metrics of accuracy among other classification techniques and can systematically balance datasets when a class occurs more frequent than other classes in the data.

3.1 Disease dataset

The disease dataset was obtained from Kaggle online website containing heart disease which can also be caused by EMF radiations or emissions. The disease dataset was divided into (70% or 212) training and (30% or 91) validation set to form a total of 303 items as show in table I.

Table 1: Heart disease dataset(Source: Kaggle online site)

index	age	sex	cp	---	ca	thal	target
1	63	1	3	---	0	1	1
2	37	1	2	---	0	2	0
3	41	0	1	---	0	2	0
4	56	1	1	---	0	2	1
3	41	0	1	---	0	2	0
---	---	---	---	---	---	---	---
298	57	0	0	---	0	3	0
299	45	1	3	---	0	3	0
300	68	1	0	---	2	3	0
301	57	1	0	---	1	3	0
302	57	0	1	---	1	2	0

3.2 Preprocessing

The splitting of dataset was done using train_test_split in Python with testing size = 0.3, random states= 0 and shuffle set to False. We employed the standard scaler in the preprocessing stage because; some features have a lot of variations in scale. The standard scaling is used to rescale

International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No. 39, April 2022 – www.ijais.org

down values within the distribution because of the variation.

3.3 Data exploration

There is need to convert some categorical variables into dummy variables and transformed before training the learning models. The get_dummies method is used to create dummy columns for the categorical variables.

3.4 Random forest classifier

The RF works with an assembly of DT's in the forest space[39]. The gini index term was required to determine the branching for each DT nodes with class probabilities in RF classification problem given as:

$$\text{Gini}=1-\sum_{i=1}^{C}(P_i)^2 \quad (1)$$

Where p_i is the relative frequency class under observation within the dataset while C = the number of classes. The RF classifier is employed to create objects for range of parameters such as estimators, paramter_name, cross validation, scoring and n_jobs. The estimator is used to pass the model or metrics in optimizing the hyper-parameter values. An integer value of 12 was passed to the cross validation(cv) in signifying the number of splits required. The n_jobs specifying the number of jobs to run in parallel and -1 is assigned to n_jobs to use all processor.

3.5 K-nearest neighbor

The KNN is a non-parametric and supervised technique used by professionals to solve classification and regression problem[39],[40]. When tested with new born(unseen) data after training to find k-training samples that are closest to the new born data samples. The model takes and assigns the most common class label among those k-training samples to the testing samples. The K-nearest neighbor was assigned a small value to avoid large variations in KNN predictions and model biasness. A dummy variable was created from the original categorical variable with unique class features. A cross validation test was carried out in finding the optimal k values within the parameter range(1 to 10). A subset of the training dataset was used to validation the training error in the process of building the model. A cross-validation test of 5-folds was adopted to randomly divide the training dataset into approximately 5 equal groups(sizes), parameter name as n_neighbors, and scoring set to be accuracy

The Euclidean distance(ED) was computed by measuring the k-nearest neighbor points between new born data and training data points given as:

$$d(x,y)=\sqrt{\sum_{i=1}^{m}(x_i-y_i)^2} \quad (2)$$

Algorithm 1: Random forest(RF)

Step	**Processes involved**
1	Start
2	Select samples randomly from dataset
3	Build, train and obtain result from each DT predictions
4	Carryout voting for each predicted result
5	Output predicted predictions with the highest votes
6	Stop

Algorithm 2: K-nearest Neighbor(KNN)

Step	**Processes involved**
1	Start
2	Classify training set and target labels with sample data
3	For j=1 in range(K): # loop through all votes of kNN
4	Calculate Euclidean distance d(train and x_samples)
5	Mark the end of For...loop
6	Compute set of indices for the K smallest distance
7	Return majority for Y_training set
8	Stop

3.6 Performance evaluation

This is the last and final step involved in measuring the success metrics of both models. We proposed the use of the following diagnostic metric, namely: accuracy, training score and cross-validation score as tools employed to measure the performance of KNN and RF classifiers.

4. RESULTS AND DISCUSSION

The RF and KNN classifiers are discussed powered by ICT innovations in detecting heart disease. The implementation was done with some fine-tuned hyper-parameter values for better performance using the training and validation accuracy scores.

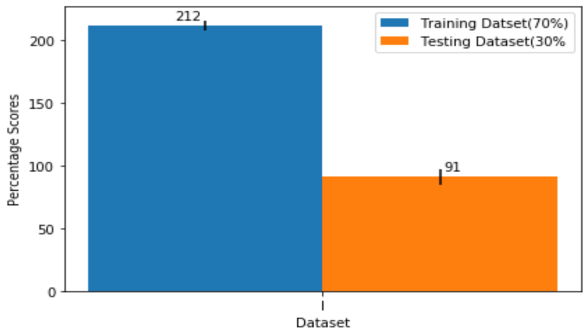

Figure 2: Train and test data split

Figure 2 is shows the total number of training and testing dataset. The total dataset was divide into 70%(212) training and 30%(91) testing set. The model was trained with training dataset and the performance was measured with the testing dataset.

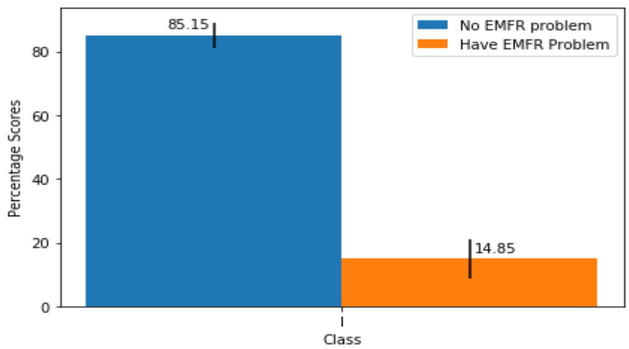

Figure 3: Effect of radio frequency radiation on human health

Figure 3 shows the negative impact of non-thermal radio frequency radiation on human health from Wifi and cellular networks of IOTs. About 14.85% of the total set under study was detected to have EMFR health problem while 85.15% are free after training with the testing dataset. The total numbers of infected persons are 90 and not infected 212 to form a total of 303 items.

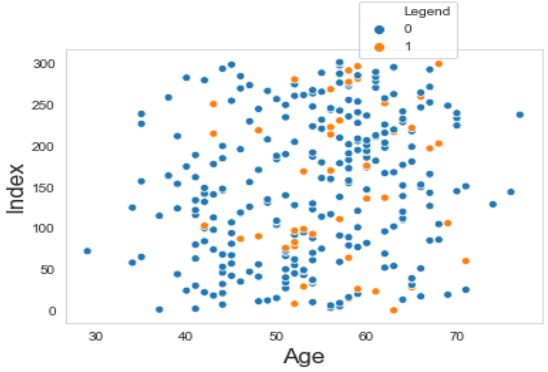

Figure 4: Distribution of people affected by the radiation from ICT devices

Figure 4 is the distribution of patients detected to be fine represented as yellow and those having the diseases with blue colored points. In the target attribute; label "0" is used to denote patients that are free from heart disease and label "1" for those having the disease. The affected number of patients fails within the age range of 40 and above, while from 40 below was not affected.

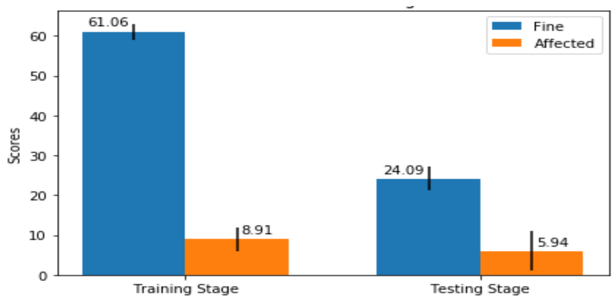

Figure 5: The detected percentage of heart disease

Figure 5 shows the percentage of patients that are fine and those having human heart disease obtained from the training and testing stages. Also from the results; 8.91% was detected to have heart disease and 61.06% are fine/free in the training dataset while 5.94% have heart disease and 24.09% are fine as recorded in the testing dataset.

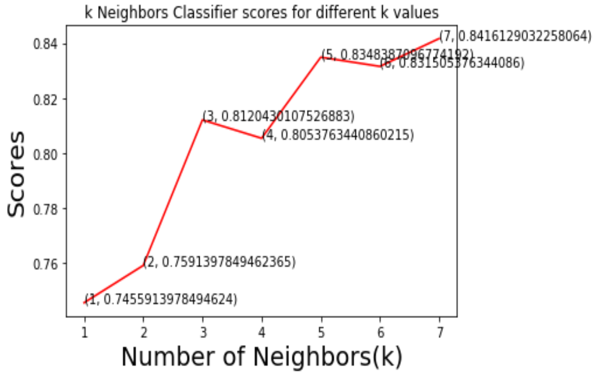

Figure 6: Graph of KNN classifier for different k values

Figure 6 depicts the performance of KNN classifier for different k-values ranging from 0 to 30. When k-neighbor(k) is set to 1 produced 74% validation score, k=2 gave 75%, k=3 produced 81% and k=4 with 80%, k=5 with 83% and so on with cross validation score for 30 iterations. The python script given bellow was used in computing the scores of the Knn classifier for different k values.

```
from sklearn.model_selection import cross_val_score
knn_scores=[]
for k in range(1,8):
    knn_classifier= KNeighborsClassifier(n_neighbors=k)
    score=cross_val_score(knn_classifier, X,y, cv=10)
    knn_scores.append(score.mean())
figure=figsize=(20, 15)
plt.plot([k for k in range(1,8)], knn_scores, color ='red')
for i in range(1,8):
    plt.text(i, knn_scores[i-1], (i, knn_scores[i-1]))
plt.xticks([i for i in range(1, 8)])
plt.xlabel('Number of Neighbors(k)', fontsize=20)
plt.ylabel('Scores', fontsize=20)
plt.title('k Neighbors Classifier scores for different k values')
```

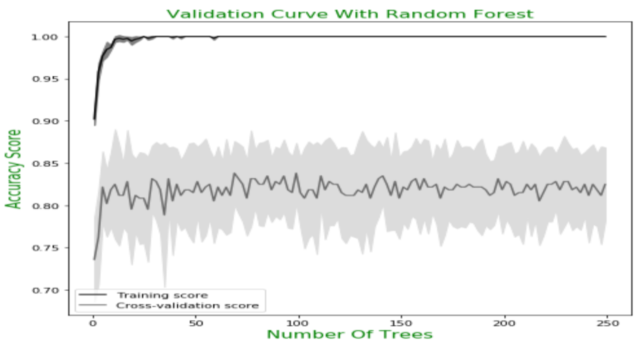

Figure 7: The validation curve graph of RF

Figure 7 is the validation curve graph showing the literal behaviour over range of values for some hyper-parameter values. The validation accuracy score(VSA) of the RF model is measured to be higher and far above the training accuracy score(TAS) with the addition of more trees. The VSA falls with 100% to 90% while TAS recorded within 73% to 84% inclusively. The code snippet was used to validate the random forest model.

```
import matplotlib.pyplot as plt
import numpy as np
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import validation_curve
param_range = np.arange(1, 250, 2)
train_scores, test_scores = validation_curve(RandomForestClassifier(),
                                X, y, param_name="n_estimators", param_range=param_range,
                                cv=4, scoring="accuracy", n_jobs=-1)
train_mean = np.mean(train_scores, axis=1)
train_std = np.std(train_scores, axis=1)
test_mean = np.mean(test_scores, axis=1)
test_std = np.std(test_scores, axis=1)
plt.subplots(1, figsize=(7,7))
plt.plot(param_range, train_mean, label="Training score", color="black")
plt.plot(param_range, test_mean, label="Cross-validation score", color="dimgrey")
plt.fill_between(param_range, train_mean - train_std, train_mean + train_std, color="gray")
plt.fill_between(param_range, test_mean - test_std, test_mean + test_std, color="gainsboro")
plt.title("Validation Curve With Random Forest", fontsize=15, color='Green')
plt.xlabel("Number Of Trees", fontsize=15, color='Green')
plt.ylabel("Accuracy Score", fontsize=15, color='Green')
plt.tight_layout()
plt.legend(loc="best")
plt.show()
```

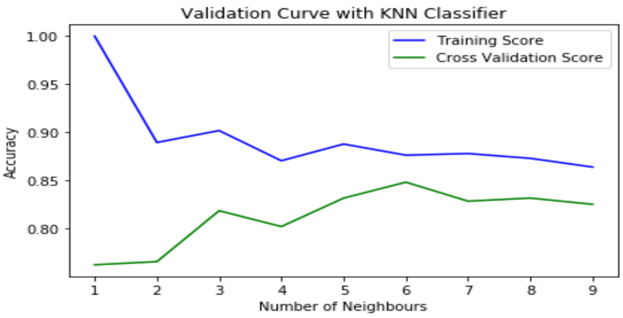

Figure 8: The validation graph of KNN classifier

The training accuracy of KNN is shown to be very high at the beginning and decreases as we increases the number of k-nearest neighbor values. The KNN cross_validation score was measured to be lower at the beginning and increases along with the increasing number of k-neighbor values as shown in figure 8. We can state from the graph that $k = 2$ is the ideal value of kNN value. As the number of neighbors(k) increases, the accuracy of Training score decreases and the cross-validation accuracy scores increases.

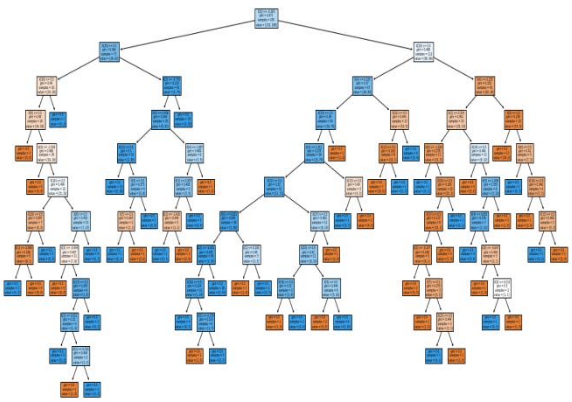

Figure 9: The RF generated from the proposed system dataset

Figure 9 depicts the RF that combined different voting DT's grown to form a forest tree of height 10 with high metrics of accuracy. This was done by choosing samples randomly from the original set with replacement in growing trees with random subset at each step. The random forest is constructed with the python script given bellow:

```
from sklearn import tree
plt.figure(figsize=(15,10))
for i in range(len(model.estimators_)):
    tree.plot_tree(model.estimators_[i], filled=True)
plt.title('Random Forest Feature classification', fontsize=14, color='Green')
plt.show()
```

Table 2: Training and validation test accuracy

Model	Training score	Validation score
RF	81.81%	85.71%
KNN	84.48%	90.00%

Table I shows the recorded training and validation accuracy score of RF and KNN classifiers measured in percentage after training. The KNN classifier produced 84.48% training and 90.00% validation accuracy compared to the RF(81.81%) training and validation(85.71%) accuracy in detection. The KNN performed better than the RF classifier in terms of accuracy.

5. CONCLUSION

The developed model can help doctors diagnose heart related disease caused my EMF emissions with the idea of machine learning driven by ICT innovation to improve decision making. The experimental results of both model proved to be highly accurate in prediction but the KNN is recommended to be more and highly efficient in terms of training and validation accuracy. We therefore; evidently conclude that the KNN performed better than the RF model in predicting the target. The general public should be enlightened about the dangers of user exposure to emissions caused by ICT devices.

The proposed system is recommended and the wired methods of internet connections which would help reduce the rate of exposure to EMF's. The general public should take note of the guidelines stipulated by the world health organization(WHO) for limiting users to the exposure of EMF, optical radiations, ultrasounds and infrasound's.

6. REFERENCES

[1] S. S. Daniel, D. Haluk, G. Michael, and L. Robert, L. An Empirical Examination of the Impact of ICT Investments on Feature Levels of Institutionalized Democracy and Foreign Direct Investment in Emerging Societies, Journal of the Association for Information System(JAIS), 13(3):117-149. 2012.

[2] L. M. Hilty, P. Arnfalk, L. Erdmann, J. Goodman, M. Lehmann, and P. A. Wager, The Relevance of Information and Communication Technologies for Environmental Sustainability - A Prospective Simulation Study, Environmental Modeling and Software, 21(11):1618-1629, 2006.

[3] S. Simko, and M. Mattsson, 5G Wireless Communication and Health Effects-A Pragmatic Review Based on Available Studies Regarding 6 to 100 GHz, International Journal of Environmental Research and Public Health(IJERPH), 16(18):1-23, 2019.

[4] V. S. Benson, K. Pirie, J. Schüz, G. K. Reeves, V. Beral,

J. Green, Mobile Phone Use and Risk of Brain Neoplasms and Other Cancers: Prospective Study, International journal of Epidemiology(IJE), 42(3):792-802, 2013.

[5] J. Wang, H. Ding, F. A, Bidgoli, B. Zhou, C. Iribarren, S. Molloi, and P. Baldi, Detecting Cardiovascular Disease from Mammograms with Deep Learning, IEEE Trans Med Imaging, 26(5):1172-1181, 2017.

[6] L. Zheng, and A. K. Chan, An Artificial Intelligence Algorithm for Tumor Detection in Screening Mammogram, IEE Trans Med Imaging, 20(7):67-559, 2001.

[7] M. Carlberg, T. Koppel, M. Ahonen, and L. Hardell, Case-control Study on Occupational Exposure to Extremely Low - Frequency Electromagnetic Fields and Glioma Risk, American Journal of Industrial Medicine, 60(5):494-503, 2018.

[8] D. R. Kohli, A. Sachdev, and H. S. Vats, Cell Phones and Tumor: Still in No Man's Land, Indian Journal of Cancer(IJC), 46(1):5-12, 2009.

[9] M. J. Adams, P. H. Hardenbergh, L. S. Constine, and S. E. Lipshultz, Radiation-Associated Cardiovascular Disease, PubMed, 45(1):55-7, 2003.

[10] J. K. Meena, A. Verma, C. Kohli, and G. K. Ingle, Mobile Phone Use and Possible Canser Risk: Current Perspectives in India, Idian Journal of Evirn. Med, 20(1):5-6, 2016.

[11] E. Senkus-Konefka, and J. Jassem, Cadiovascular Effects of Breast Cancer Radiotherapy, PubMed, 33(6):578-93, 2007.

[12] D. C. Prasad, P. E. S. N. Krishna, and Y. Sagar, An Approach to Develop Expert Systems in Medical Diagnosis Using Machine Learning Algorithms(ASTHMA and A Performance Study, International Journal of Soft Computing(IJSC), 2(1):26-33, 2011.

[13] Y. Stein, O. Levy-Nativ, and E. D. Richter, A Sentinel Case Series of Cancer Patients with Occupational Exposures to Electromagnetic Non-ionizing Radiation and Other Agents, European Journal of Oncology(EJO), 16(1):21-54, 2011.

[14] S. Szmigielski Cancer Risks Related to Low-Level RF/MW Exposures, Including Cell Phones, Electromagnetic Biology and Medicine(EBM), vol. 32(3):273-280, 2013.

[15] P. A. Valberg, T. E. Van-Deventer, and M. H. Repacholi, Base Stations and Wireless Networks– Radiofrequency (RF) Exposures and Health Consequences, Environmental Health Perspectives, 115(3):416–424, 2007.

[16] B. Zarikoff, and D. Malone, A Comparison of RF Exposure in Macro- and Femtocells, Health Physics, 105(1):39-48, 2013.

[17] C. Bornkessel, M. Schubert, M. Wuschek, and P. Schmidt, Determination of the General Public Exposure Around GSM and UMTS Base Stations, Radiation Protection Dosimetry, 124(1):40-47, 2007.

[18] H. Yinghe, K. L. Vincken, F. P. Lafeber, and M. A. Viergever, Automatic Quantification of Radiographic Finger Joint Space Width of Patients with Early Rheumatoid Arthritis, IEEE Journals and Magazines, vol. 63(10):2177-2186, 2016.

[19] C. K. Shie, C. H. Chuang, C. N. Chou, M. H. Wu, and E. Y. Chang, Transfer Representation Learning for Medical Image Analysis, In IEEE Conf Eng Med Bio:711-714, 2015.

[20] M. Fatima, and M. Pasha, Survey of Machine Learning Algorithms for Disease Diagnostic, Journal of Intelligent Learning Systems and Applications: 1-16, 2017.

[21] D. Yu, and L. Deng, Deep Learning and It's Applications to Signal Information Processing, IEEE Signal Process, Mag., 28(1):145-154, 2011.

[22] K. H. Rao, G. Srinivas, A. Damodhar, and M. V. Krishna, Implementation of anomaly detection technique using machine learning algorithms, International Journal of Computer Science and Telecommunication(IJCST), 2(3):25-31, 2011.

[23] S. Ruggieri, Efficient C4.5 Classification Algorithm , Knowledge and Data Engineering, IEEE Transactions, 14(2):438-444, 2002.

[24] A. Kumar,J., Kim, W. Cai, and D. Feng, Content-based Medical Image Retrieval: A Survey of Applications to Multidimensional and Multimodality Data, Journal of Digit Imaging, 26(6):1025-1039, 2013.

[25] B. Lie, P. Yang, T. Wang, and S. N. D. Chen, Regularized Discriminative Sparse Learning for Alzheimer's Disease Diagnosis," IEEE Transactions on Cybernetics, 47(4):1102-1113, 2017.

[26] Shi, J. and Malik, J. 2000 Normalized Cuts and Image Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp.888-905.

[27] Y. LeCun, Y. Bengio, and G. Hinto, Deep Learning, Nature, 521(7553):436-444, May. 2015

[28] K. C. Tan, W. J. Teoh, EQ. Yu, and K. C. Goh, A Hybrid Evolutionary Algorithm for Attribute Selection in Data Mining, Journal of expert System with Applications, 36:8616-8630, 2009.

[29] M. Abedini, N. C. F. Codella, R. Connell, R. Garnavi, M. Merler, S. Pankanti, I. R. Smith, and T. Yeda-Mahmood, A Generalized Framework for Medical Image Classification and Recognition, IBM J Res Dev. 59(2/3):1-18, 2015.

[30] A. Omotosho, P. Ayegba, J. Emuoyibofarhe, and C. Meinel, Current State of ICT in Healthcare Delivery in Developing Countries, International Journal of Online Engineering (iJOE), 15(8):91-107, 2009.

[31] J. Breckenkamp, G. Berg, and M. Blettner, Biological Effects on Human Health Due to Radiofrequency/Microwave Exposure: A Synopsis of Cohort Studies, Radiat Environ Biophys, 42(3):141-54, 2003.

[32] M. Hours, M. Bernard, L. Montestrucq, M. Arslan, A. Bergeret, L. Deltour, and E. Cardis, Cell Phones and Risk of Brain and Acoustic Nerve Tumours: The French

INTERPHONE Case-Control Study, Rev Epidemiol Sante Publique,. 55(5):321-32, 2007.

[33] H. Lacohe, N. Wakeford, L. Pearson, A Social History of the Mobile Telephone with a View of its Future, BT Technology Journal, vol. 21(3):203 – 211, 2007.

[34] M. Prasad, P. Kathuria, P. Nair, A. Kumar, and K. Prasad, Mobile Phone Use and Risk of Brain Tumours: A Systematic Review of Association Between Study Quality, Source of Funding, and Research Outcomes, Neurological Sciences(NS), 38(5):797-810, 2017.

[35] F. M. Ba-Alwi, and H. M. Hintaya, Comparative Study for Analysis: The Prognostic in Hepatitis Data mining Approach, International Journal of Scientific Engineering Research, 4:680-685, 2013.

[36] G. Sathyadevi, Application of CART Algorithm in Hepatitis Disease Diagnosis, IEEE International Conference on Recent Trends in Information Technology(ICRTIT), MIT, Anna University, Chennai, 3-5:1283-1287, 2011.

[37] S. Vijayarani, and S. Dhayanand, Liver Disease Prediction Using SVM and Naive Bayes Algorithms, International Journal of Science Engineering and Technology Research(IJSETR), 4:816-82, 2015.

[38] P. Rajeswari, and G. S. Reena, Analysis of Liver Disorder Using Data Mining Algorithm," Journal of Computer Science and Technology, 10:48-52, 2010.

[39] M. Fatima,. and P. Maruf, Survey of Machine Learning Algorithms for Disease Diagnosis, Journal of Intelligent Learning Systems and Applications(JILSA), 9(1):1-16, 2017.

[40] F. Burba, F. Ferraty, and P. Vieu, k-Nearest Neighbour Method in Functional Nonparametric Regression, International Journal of Intelligent Engineering Informatics(IJIE), 6(6):453-469, 2018.