OPEN ACCESS

Kashf Journal of Multidisciplinary Research
Vol: 01 - Issue 12 (2024)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

ENHANCING EARLY BREAST CANCER PREDICTION: A COMPARATIVE STUDY OF MACHINE LEARNING MODELS ON CLINICAL ACCURACY AND INTERPRETABILITY

Shazma Tahseen
Information Technology, Liaquat University of Medical and Health Sciences Jamshoro, Pakistan

Jawad Hussain Awan
Faculty of Engineering, Sciences and Technology, IQRA University, Karachi, Pakistan

Abdul Rahman Baloch
Faculty of Engineering, Sciences and Technology, IQRA University, Karachi, Pakistan

****Corresponding author: Shazma Tahseen (shazma.tehseen@lumhs.edu.pk)***

Article Info

Abstract
The purpose of this research is to evaluate and compare several machine learning models according to their accuracy, considering their generalization capability and their ability to provide clinical explanations. Primary data was acquired from 150 case (cancer patients) and 80 control (cancer-free) respondents from cancer treatment centers. Multilayer perceptron (MLP) and logistic regression algorithms performed better than the traditional methods in predicting breast cancer incidence. The above results are meant to speed up the process of early cancer prediction and show AI experts new ways AI is being used in health care.

Keywords: *AI, Breast Cancer, Machine Learning, ML Algorithms.*

I. BACKGROUND AND RELATED WORK

As one of the fastest-growing NCDs, cancer adds 19.3 million new cases and 10 million fatalities by the year 2020. Cancer affects one in five men and one in six women worldwide, with one in eight men and one in eleven women ultimately succumbing to the disease. In 2020, there will be an estimated 2.3 million new instances of female breast cancer, making it the most commonly diagnosed illness and a significant cause of cancer deaths [1]. Cancer of the breast is an abnormal development of malignant cells. It might spread to other areas of the body if not addressed. Among females, breast cancer has the highest incidence rate (excluding skin cancer) [2].

The process of diagnosis and treatment for breast cancer is time-consuming and costly, both for the individual and for society as a whole. This highlights the need for research into BC prevention strategies and early detection methods [3]. Early identification of cancer reduces the cost and duration of therapy, as well as the likelihood that the disease may spread to other regions of the body [4, 5]. The World Health Organization (WHO) and the International Alliance for Cancer Early Diagnosis (IACED) are just two of the many global health organizations that advocate for early detection as a means of lowering mortality rates [6].

The latest figures show that the incidence and mortality rate of breast cancer are on the rise worldwide, especially in the Americas, Asia-Pacific, and Africa. This emphasizes the need for early detection. Numerous studies have demonstrated the importance of early cancer diagnosis in reducing mortality rates [4]. Computer-aided programs with cancer-related experience were finally developed as a solution to the problem of early cancer detection. The field of early breast cancer diagnosis utilizing ML for AI is growing as its own subject of medicine and computer science [6]. Thanks to the advancement of statistical methods and AI, doctors, scientists, and computer scientists may now work together to improve prognostic tools by applying techniques like factor analysis and regression analysis. These hybrid systems outperform simple empirical predictions in terms of accuracy [7].

One must proceed with caution when addressing the unique characteristics and screening limitations of cancer, and instead rely on a broad, non-specific approach to diagnosis that often involves minimal medical intervention. Generalized linear models, regression, correlation, and other statistical approaches have been used and advocated for in the diagnosis and categorization of malignant malignancies [8].

Using AI, which can be aided by statistical approaches, can improve diagnostic accuracy and precision. Since AI algorithms can examine massive volumes of multi-modal data, they are helping to enhance the screening process by locating signals that would be difficult to locate manually. [9-11]. Through the use of algorithms and learned knowledge, artificial intelligence (AI) teaches computers to comprehend their environment and act accordingly. The data may be used to define criteria and automate the scientific process, which will improve cancer diagnosis. Supervised learning, in which the machine already has some idea of the outcome, and unsupervised learning, in which the machine has no prior knowledge of the outcome, are the two broad categories into which Machine Learning (ML) may be broken down. Data from both methods can be used to predict the presence or absence of cancer, as well as the likelihood of cancer occurrence. Natural Language Processing (NLP), Convolutional Neural Network (CNN), and Multilayer Perceptron (MLP) are the most frequently used algorithms in the health sciences. [12, 13].

Artificial intelligence systems can sort out confusing signals from massive amounts of heterogeneous and multimodal data, gradually learning about the factors involved in breast cancer and their effects [11, 12]. The

capability of the diagnosis system may be increased by using AI, which can begin analysis in screened patients using clinically defined parameters, hence improving cancer detection [7].

Numerous machines and deep learning strategies have been created for the detection and categorization of breast cancer. These methods may be broken down into three distinct subfields: preprocessing, feature extraction, and classification. Preprocessing facilitates the transformation of raw data into a format that can be processed by automated systems and meets other prerequisites. However, the feature extraction process may discriminate between cancerous and noncancerous tumours. [14–17].

A machine learning (ML) technique called Random Forest Classifier (RFC) is being used to find breast cancer in its earliest stages. RFC is a member of the family of algorithms that employ a collection of decision trees [18]. The model's final prediction is determined by the tree with the most votes from among all the trees in the random forest. Each tree (classifier) in RFC works as a member of a team, contributing to the overall prediction. Researchers found that RFC has several benefits, including excellent accuracy and efficiency when working with heterogeneous data, handling dichotomous classification, and addressing datasets with fewer variables than observations [19-26]. Because of its exceptional benefits and advantages, RFC is being used for the prediction and categorization of breast cancer in a variety of ways depending on the area of application.

In addition to the ML algorithms listed above, there is another one known as logistic regression, which is a supervised machine learning approach utilized by AI in statistics. Despite the fact that the name contains the term "regression," it is a classification model rather than a regression. It is a popular categorization method due to its simplicity and excellent results with linearly separable classes. A logistic regression model, like the Adaline and perceptron statistical techniques, classifies binary classes and can be extended to multiclass classification [27]. Logistic regression is used to estimate the likelihood of a binary output, which might contain any of the two values: yes or no; true or false; and so on. The multinomial logistic regression version may handle scenarios with more than two possible outcomes. It assists in identifying the best category for a fresh sample [28].

II. MATERIAL AND METHODS

A. Data Collection

For the purpose of gathering data, two sampling methods were used: convenience sampling was used for data on cancer patients [28, 29] and simple random sampling was used for data on healthy women. Two important factors were rigorously considered while selecting the sample: avoiding sampling bias and selecting a sample that accurately reflects the socioeconomic class and ethnic makeup of communities [30, 31]. Second, the study's findings should be as broadly applicable to the community from which the sample was drawn as possible [32, 33].

B. Application of machine learning based prediction models

Logistic Regression (LR)

Logistic regression is a classification algorithm. The logistic or sigmoid function is used to generate this. The most popular method for determining which of two categories an observation belongs to is multinomial logistic regression, because of the presence of more than two categories. By using the GridSearchCV method, we were able to fine-tune the LogR model's hyper-parameters and get better results from the predictions.

Random Forest Classifier

A random forest is a type of Meta estimator that employs averaging to increase predicted accuracy and control over-fitting by applying several decision tree classifiers to different subsamples of the dataset. A parameter determines the size of the subsample used to construct each tree; otherwise, the entire dataset is applied to each tree. Random Forest is also a form of supervised learning; it is one of the well-known algorithms in the machine learning field. The ML tool may apply to both classification and regression problem types. This is in line with ensemble learning where several classifiers are used in solving a difficult problem to improve the accuracy of the model. It worked well even without fine-tuning of the hyper-parameters.

Multilayer Perceptron (MLP)

The foundation of artificial neural networks is the node structure. Coherent neurons form many of its many levels. Layers are concealed between the input and output layers, which are themselves forced layers. Each prediction is made using a set of variables that are fed into the input layer, and the output layer then shows the result. They specify and update the weight of each neuron in each layer as they progress through the training process, which allows the model to generalize across new sets of data. The weights are used by the trained model to determine which units to activate given a set of inputs. The Multilayer Perceptron (MLP) is a well-established model of neural networks frequently employed in classification and regression (Ahmed et al., 2019). Prediction of breast cancer occurrence risk were accomplished with a Multilayer Perceptron Neural Network (MLP-NN) using Python version 3.9.7 with a configuration (10,30,10) of max 100 eps and a learning rate ("constant," "adaptive").

C. Model calibration

Calibration of a model often entails a reduction in the cost function used by the training process. Model learning from the used training dataset then determines the appropriate weights. Overfitting, a kind of collective deficit, can occur at this stage, causing unwanted background noise and potentially negative training outcomes. The training dataset is inaccessible to hyperparameters, yet they add complexity in order to get the best model structure possible. Therefore, all models had their hyperparameters specified during the calibration process.

Model performance metrics

The mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), root mean square error (RMSE)-standard deviation of observations (RSR), and determination of coefficients are among the methods used to assess the model's correctness during training and testing in this investigation (R2). According to Moriasi et al. (2007), the model's accuracy may be rated as excellent (RSR 0.50), acceptable (RSR 0.50065), satisfactory (RSR 0.60050), or poor (RSR > 0.70). When the MAPE is small, the model is accurate, and vice versa (Lu and Ma, 2020)

III. RESULTS AND DISCUSSION

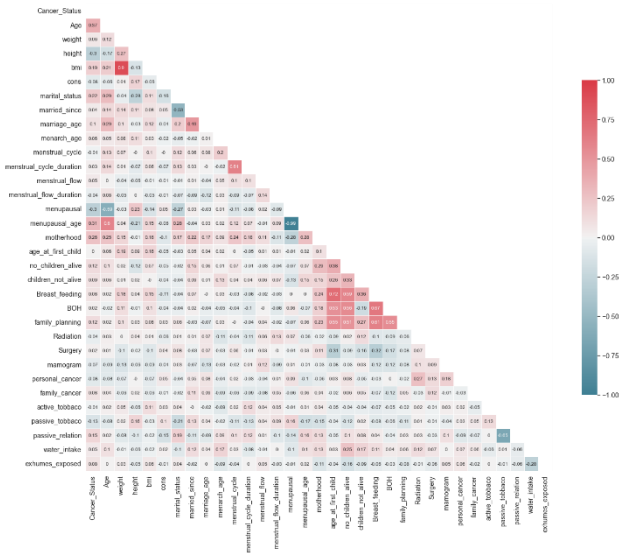

Fig. 1. The Pearson correlation coefficient between variables

Correlation matrix was developed to find the association between different identified factors of breast cancer, and how they are effecting and moving with each other.

A. Performance of developed machine learning models

Numerous studies in a wide variety of fields have found that algorithms developed using machine learning improve prediction accuracy. This research suggests that ML approaches may accurately and precisely predict the probability of breast cancer occurrence in women who have never been diagnosed with breast cancer but who are exposed to the same causes as breast cancer patients. Three machine learning algorithms were examined in this study using both classification and regression techniques. There were two algorithms that significantly improved the accuracy with which breast cancer could be predicted. Both the training and testing data performed well after being separated. ML algorithms can improve the accuracy of classification and regression for the prediction of breast cancer, but current methods of breast cancer screening are either prohibitively expensive or require a physical examination at specialist facilities. The scope of this research can be expanded to include other cancers and other stages of the disease. This study's sample is representative of the general population, encompassing a wide range of racial and socioeconomic groups. The application of machine learning algorithms is required. Specifically, this research made use of two supervised machine learning algorithms: classification and regression. Here, we've taken on the challenge of figuring out how to compare the two kinds of algorithms in a meaningful way. For classification, accuracy, precision, recall, and FI score are some of the ways to measure performance. For regression, MSE, RMSE, MAE, R squared, and RSR are used.

B. Results of the Algorithms

The likelihood of developing breast cancer was predicted with the use of ML algorithms. The training was conducted using the 43 parameters shown in Fig. 1: correlation matrix. A set of all factors were considered for the training of the machine. Table 1, displays the prediction results, which reveal that MLP-NN outperformed other algorithms in training the machine in a specific setting, with an accuracy of 0.96, a precision of 0.96, a recall of 0.96, and an F1 score of 0.96.

Table 1: Performance of ML Algorithms

Algorithm	Accuracy	Precision	Recall	F1 Score
Logistic regression	0.95	0.94	0.95	0.94
Random Forest Classifier	0.95	0.94	0.95	0.94
MLP-NN	0.96	0.96	0.96	0.96

The results of the training and testing datasets for the prediction of breast cancer occurrence in breast cancer-free women using the similar exposure they have as breast cancer patients are presented in Table 2 in terms of Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R-Squared. The results show that MLP performs slightly better than logistic regression and random forest classifiers. Although other models also did well in predicting cancer incidence, the results show that the MLP algorithm appeared asa co most efficient. In this work, three different prediction models are trained and verified using supervised learning-based algorithms (LogR, Random Forest Classifier, and MLP), and their results are compared to determine the most effective model. During model training, it has also been demonstrated that the MLP technique works well with a limited and finely tuned dataset. As a prediction tool, it is more precise and accurate.

Scatter plot was constructed using the best fit line of the best acquired model to help examine the accuracy of the created models (Figures 2, 3 and 4, Plots showing prediction versus training and testing data). Predicted values from the models are shown in relation to the test values in the figure; these values are more closely aligned with the best-fit line for the MLP model than for any of the other models, demonstrating the model's validity. In contrast to the logistic regression and random forest classifier models, the MLP model's predictions were generally more in line with the testing data. In addition, logistic regression and the random forest classifier had a greater coefficient of determination (R^2).

Table 2: Performance metrics for ML model result

Algorithm	MSE	RMSE	MAE	R-Squared
Logistic Regression				
Test Data	0.06	0.25	0.06	0.73
Training Data	0.02	0.15	0.03	0.89
Random Forest Classifier				
Test Data	0.04	0.21	0.04	0.81
Training Data	0.02	0.15	0.02	0.89
MLP				
Test Data	0.04	0.21	0.04	0.82
Training Data	0.03	0.19	0.03	0.85

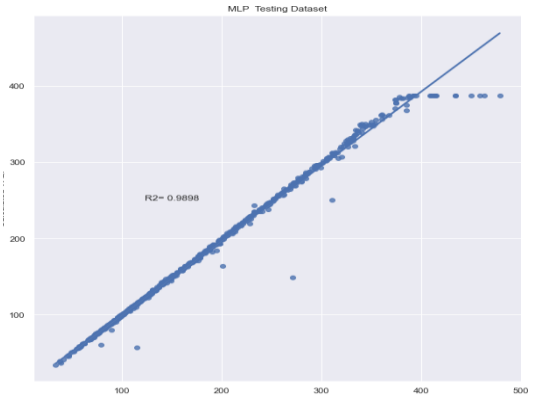

Fig. 2: MLP Testing Dataset

Fig. 2: Logistic Regression Dataset

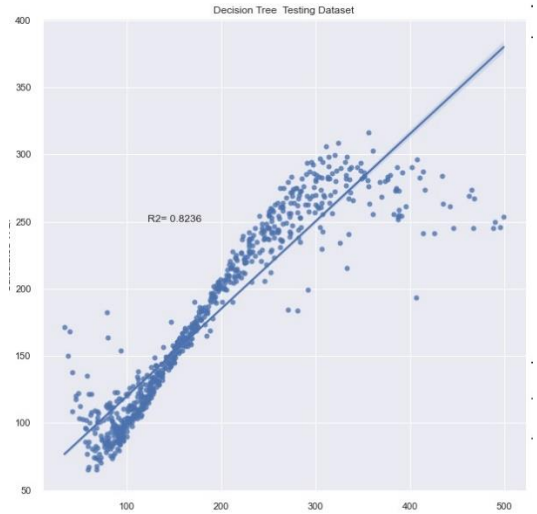

Fig. 3: Random Forest Dataset

IV. CONCLUSION

Before the development and use of AI and ML systems, accurate breast cancer forecasting was essentially unattainable. The process of diagnosis, screening, and detection was laborious, time-consuming, and costly because it included a physical examination and medical treatment. In this study, we explored many different machine learning strategies for making accurate breast cancer risk predictions. Possibly as a result of data tweaking through GridSearchCV, three ML models in this study showed promising results in categorizing and predicting breast cancer. After controlling for balance and normalisation, the findings indicated that MLP performed better with the dataset, whereas LogR was best suited to the smaller dataset. To determine which prediction model performed best, we utilized performance indicators such as mean squared error (MSE) = (0.04-0.03), root mean squared error (RMSE) = (0.21-0.19), mean absolute error (MAE) = (0.04-0.03), and coefficient of accuracy (R2) = (0.82-0.85). The results also confirm the MLP and LogR's ability to accurately identify the importance of input factors in BC prediction calculations. Therefore, the MLP and LogR models built by machine learning may be effectively utilised to assess BC forecasting in the current study region. These models are also applicable to other types of cancer research. Small data from just two cancer centers is a drawback of the study; additional data from a wider variety of sources might provide light on the disease at hand. This factor is expected to greatly increase the utilisation of ML models in the cancer research sector. Improve the current study by comparing the created models' prediction power to that of other ML models, considering a wide range of medical and socioeconomic inputs.

REFERENCES:

[1] H. Sung *et al.*, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA. Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, May 2021, doi: 10.3322/caac.21660.

[2] S. Chaudhury, M. Rakhra, N. Memon, K. Sau, and M. T. Ayana, "Breast Cancer Calcifications: Identification Using a Novel Segmentation Approach," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–13, Oct. 2021, doi: 10.1155/2021/9905808.

[3] A. Büyükavcu, Y. E. Albayrak, and N. Göker, "A fuzzy information-based approach for breast cancer risk factors assessment," *Appl. Soft Comput.*, vol. 38, pp. 437–452, Jan. 2016, doi: 10.1016/j.asoc.2015.09.026.

[4] S. McPhail, S. Johnson, D. Greenberg, M. Peake, and B. Rous, "Stage at diagnosis and early mortality from cancer in England," *Br. J. Cancer*, vol. 112, no. S1, pp. S108–S115, Mar. 2015, doi: 10.1038/bjc.2015.49.

[5] L. Uttley, B. L. Whiteman, H. B. Woods, S. Harnan, S. T. Philips, and I. A. Cree, "Building the Evidence Base of Blood-Based Biomarkers for Early Detection of Cancer: A Rapid Systematic Mapping Review," *EBioMedicine*, vol. 10, pp. 164–173, Aug. 2016, doi: 10.1016/j.ebiom.2016.07.004.

[6] B. Hunter, S. Hindocha, and R. W. Lee, "The Role of Artificial Intelligence in Early Cancer Diagnosis," *Cancers (Basel).*, vol. 14, no. 6, p. 1524, Mar. 2022, doi: 10.3390/cancers14061524.

[7] S. Huang, J. Yang, S. Fong, and Q. Zhao, "Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges," Cancer Lett., vol. 471, pp. 61–71, Feb. 2020, doi: 10.1016/j.canlet.2019.12.007.

[8] Z. Li, S. Wang, and X. Lin, "Variable selection and estimation in generalized linear models with the seamless L0 penalty," *Can. J. Stat.*, vol. 40, no. 4, pp. 745–769, Dec. 2012, doi: 10.1002/cjs.11165.

[9] K. Dembrower *et al.*, "Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study," *Lancet Digit. Heal.*, vol. 2, no. 9, pp. e468–e474, Sep. 2020, doi: 10.1016/S2589-7500(20)30185-0.

[10] Y. Xu, J. Kepner, and C. P. Tsokos, "Identify Attributable Variables and Interactions in Breast Cancer," *J. Appl. Sci.*, vol. 11, no. 6, pp. 1033–1038, Mar. 2011, doi: 10.3923/jas.2011.1033.1038.

[11] J. T. Beck *et al.*, "Artificial Intelligence Tool for Optimizing Eligibility Screening for Clinical Trials in a Large Community Cancer Center," *JCO Clin. Cancer Informatics*, no. 4, pp. 50–59, Sep. 2020, doi: 10.1200/CCI.19.00079.

[12] B. H. Kann, R. Thompson, J. Charles R. Thomas, A. Dicker, and S. Aneja, "Artificial Intelligence in Oncology: Current Applications and Future Directions," *Oncology*, vol. 33, no. 2, pp. 46–53, 2019, [Online]. Available: https://www.cancernetwork.com/view/artificial-intelligence-oncology-current-applications-and-future-directions

[13] W. W. Yim, M. Yetisgen, W. P. Harris, and W. K. Sharon, "Natural Language Processing in Oncology: A Review," *JAMA Oncol.*, vol. 2, no. 6, pp. 797–804, Jun. 2016, doi: 10.1001/JAMAONCOL.2016.0213.

[14] J. A. Cruz and D. S. Wishart, "Applications of Machine Learning in Cancer Prediction and Prognosis," *Cancer Inform.*, vol. 2, p. 117693510600200, Jan. 2006, doi: 10.1177/117693510600200030.

[15] M. F. Akay, "Support vector machines combined with feature selection for breast cancer diagnosis," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 3240–3247, Mar. 2009, doi: 10.1016/j.eswa.2008.01.009.

[16] D. N. Ponraj, M. E. Jenifer, P. Poongodi, and J. S. Manoharan, "A survey of the preprocessing techniques of mammogram for the detection of breast cancer," *J. Emerg. Trends Comput. Inf. Sci.*, vol. 2, no. 12, pp. 656–664, 2011, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.651.592&rep=rep1&type=pdf

[17] G. Valvano *et al.*, "Convolutional Neural Networks for the Segmentation of Microcalcification in Mammography Imaging," *J. Healthc. Eng.*, vol. 2019, pp. 1–9, Apr. 2019, doi: 10.1155/2019/9360941.
[18] L. Kuncheva, T. Christy, I. Pierce, … S. M.-C. on I., and U. 2011, "Multi-modal biometric emotion recognition using classifier ensembles," in *Springer*, 2011, pp. 317–326. Accessed: Apr. 04, 2022. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-21822-4_32
[19] R. Díaz-Uriarte and S. Alvarez de Andrés, "Gene selection and classification of microarray data using random forest," *BMC Bioinformatics*, vol. 7, Jan. 2006, doi: 10.1186/1471-2105-7-3.
[20] T. Khoshgoftaar, … M. G.-19th I. I., and U. 2007, "An empirical study of learning from imbalanced data using random forest," in *19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*, 2007, pp. 310–317. Accessed: Apr. 04, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4410397/?casa_token=2G38dwBUA34AAAAA:9-lusLBducdwWRCDMk1OfH_beajk8lLD3-0okol8KcmGJWdQZQrmpKlGxBr0CwPxnslUaQjhPZc
[21] M. Khalilia, S. Chakraborty, and M. Popescu, "Predicting disease risks from highly imbalanced data using random forest," *BMC Med. Inform. Decis. Mak.*, vol. 11, no. 1, p. 51, Dec. 2011, doi: 10.1186/1472-6947-11-51.
[22] J. S. Evans, M. A. Murphy, Z. A. Holden, and S. A. Cushman, "Modeling Species Distribution and Change Using Random Forest," in *Predictive Species and Habitat Modeling in Landscape Ecology*, New York, NY: Springer New York, 2011, pp. 139–159. doi: 10.1007/978-1-4419-7390-0_8.
[23] Z. Masetic and A. Subasi, "Congestive heart failure detection using random forest classifier," *Comput. Methods Programs Biomed.*, vol. 130, pp. 54–64, Jul. 2016, doi: 10.1016/j.cmpb.2016.03.020.
[24] M. Mohammady, H. R. Pourghasemi, and M. Amiri, "Land subsidence susceptibility assessment using random forest machine learning algorithm," *Environ. Earth Sci.*, vol. 78, no. 16, p. 503, Aug. 2019, doi: 10.1007/s12665-019-8518-3.
[25] A. Subasi, E. Alickovic, and J. Kevric, "Diagnosis of Chronic Kidney Disease by Using Random Forest," in *CMBEBIH 2017. IFMBE Proceedings*, V. 62, Ed. Singapore: Springer Singapore, 2017, pp. 589–594. doi: 10.1007/978-981-10-4166-2_89.
[26] M Belgiu and L Drăguţ, "Random forest in remote sensing: A review of applications and future directions," *ISPRS J. Photogramm. Remote Sens.*, vol. 114, pp. 24–31, Apr. 2016, Accessed: Apr. 04, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271616000265?casa_token=v9qGvhQ48v4AAAAA:xAOeL_Ph-zlaPbCtbPMwmiIRH-LKWEZpIQ6kHLfHPF3kO9fKfHN8EGruOuKS49Xr4YgDpEagv0c
[27] S. Raschka, *Python Machine Learning*. Birmingham: Packt Publishing, 2015. [Online]. Available: https://books.google.com.pk/books/about/Python_Machine_Learning.html?id=GOVOCwAAQBAJ&printsec=frontcover&source=kp_read_button&hl=en&redir_esc=y#v=onepage&q&f=false
[28] T. W. Edgar and D. O. Manz, "Exploratory Study," in *Research Methods for Cyber Security*, Elsevier, 2017, pp. 393–404. doi: 10.1016/B978-0-12-805349-2.00030-3.
[29] U. Boehmer, M. A. Clark, A. Timm, M. Glickman, and M. Sullivan, "Comparing Sexual Minority Cancer Survivors Recruited Through a Cancer Registry to Convenience Methods of Recruitment," *Women's Heal. Issues*, vol. 21, no. 5, pp. 345–352, Sep. 2011, doi: 10.1016/j.whi.2011.03.003.
[30] K. S. Coyne, C. Sexton, R. J. LoCasale, F. R. King, M. K. Margolis, and S. H. Ahmedzai, "Opioid-Induced Constipation among a Convenience Sample of Patients with Cancer Pain," *Front. Oncol.*, vol. 6, Jun. 2016, doi: 10.3389/fonc.2016.00131.
[31] C. Haddad *et al.*, "Correlates of orthorexia nervosa among a representative sample of the Lebanese population," *Eat. Weight Disord. - Stud. Anorexia, Bulim. Obes.*, vol. 24, no. 3, pp. 481–493, Jun. 2019, doi: 10.1007/s40519-018-0631-x.
[32] A. M. Bueno *et al.*, "Injury prevalence across sports: a descriptive analysis on a representative sample

of the Danish population," *Inj. Epidemiol.*, vol. 5, no. 1, p. 6, Dec. 2018, doi: 10.1186/s40621-018-0136-0.

[33] Z. Hu and J. Qin, "Generalizability of causal inference in observational studies under retrospective convenience sampling," *Stat. Med.*, vol. 37, no. 19, pp. 2874–2883, Aug. 2018, doi: 10.1002/sim.7808.

[34] S. J. Kamper, "Generalizability: Linking Evidence to Practice," *J. Orthop. Sport. Phys. Ther.*, vol. 50, no. 1, pp. 45–46, Jan. 2020, doi: 10.2519/jospt.2020.0701.