EARLY RISK PREDICTION OF CERVICAL CANCER: A MACHINE LEARNING APPROACH

Syed Anas Ali[1], Mohammed Lateef Mahmood [2], Mohammed Abdul Zaheer[3], Hajira Sabuhi[4]
[1,2,3] B.E. Student, Department of IT, Lords Institute of Engineering and Technology, Hyderabad
[4] Assistant Professor, Department of IT, Lords Institute of Engineering and Technology, Hyderabad
hajirasabuhi@lords.ac.in

Abstract: *Cervical cancer is a significant public health concern worldwide, necessitating early risk prediction through Machine Learning (ML). This study employs eleven ML algorithms on a UCI ML repository dataset to forecast risks. Initial results show Multi-Layer Perceptron (MLP) achieving 93.33% accuracy with default settings. Further, hyperparameter tuning via Grid Search Cross Validation (GSCV) validates comparable performance across K-Nearest Neighbours (KNN), Decision Tree Classifier (DTC), Support Vector Machine (SVM), Random Forest Classifier (RFC), and MLP, all achieving 93.33% accuracy. This study underscores ML's potential in early cervical cancer risk assessment, benefiting healthcare professionals and at-risk individuals through enhanced predictive capabilities[1].*

I. Introduction

Cervical cancer is a significant global health issue impacting women worldwide, especially in developing nations where it remains a leading cause of mortality according to WHO [1]. In 2012, it accounted for 445,000 new cases, highlighting its pervasive nature [2]. Early detection through methods like the pap smear has proven effective, yet challenges such as limited access to healthcare facilities and variability in diagnostic accuracy persist [3, 4]. Human Papilloma Virus (HPV), a primary risk factor, underscores the importance of lifestyle and prevention strategies [5,6].

This study employs eleven supervised ML algorithms—Decision Tree Classifier (DTC), Multi-Layer Perceptron (MLP), Random Forest Classifier (RFC), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), CatBoost (CatB), Gaussian Naïve Bayes (GNB), Gradient Boosting Classifier (GradB), AdaBoost (AdaB), XG Boost (XGB), and XG Boost with Random Forest (XGBRF)—on a UCI ML repository dataset. Initial results show MLP achieving 93.33% accuracy, subsequently validated through Grid Search Cross Validation (GSCV) across all models.

By leveraging ML's capabilities to analyze complex datasets, this research aims to enhance early prediction and intervention strategies for cervical cancer. Such advancements could revolutionize healthcare by improving screening efficiency and outcomes, ultimately reducing the global burden of cervical cancer [7,8].

II. Literature Survey

1. Automated Cervical Cancer Screening Using Machine Learning Techniques:

Various studies have explored the application of machine learning for automated cervical cancer screening. Techniques such as support vector machines (SVM), artificial neural networks (ANN), and decision trees have been investigated for their efficacy in predicting cervical cancer risks from clinical data and imaging results [Sobar et al., 2021; Kashyap et al., 2019].

2.Integration of HPV Data:

Human Papilloma Virus (HPV) infection is a major risk factor for cervical cancer. Research has focused on integrating HPV genomic data with machine learning algorithms to improve predictive models. This approach aims to enhance accuracy by considering genetic markers associated with HPV strains and their oncogenic potential [Njoroge et al., 2018; Fazal et al., 2020].

3.Comparison of ML Algorithms:

Studies have compared the performance of different machine learning algorithms in cervical cancer risk prediction. Research by Sobar et al. and Kashyap et al. highlighted varying accuracies achieved by SVM, random forest, and neural network models, emphasizing the need for algorithm selection based on dataset characteristics and predictive goals [Sobar et al., 2021; Kashyap et al., 2019].

4.Image-Based Screening Methods:

Machine learning techniques applied to pap smear images or colposcopy findings have shown promising results in automating diagnostic processes. Image classification using convolutional neural networks (CNN) and feature extraction methods like Fourier-Transform Infrared (FTIR) spectroscopy have demonstrated high accuracy in detecting abnormal cervical cells [Hyeon et al., 2017; Wu et al., 2019].

III.System Analysis

In existing system, In 2012, about 445,000 cases were discovered and almost 83% of all were new cases .Symptoms of cervical cancer include irregular periods, unexpected blood, and atypical menstruation. Hence, a pap smear test can diagnose cervical cancer and has been shown to reduce death risk by almost 90% and cervical cancer risk by 60% to 90%. However, the absence of medicinal equipment, inadequate nurturing, simple diagnostic reproducibility, careless maintenance, and ennui on the part of the specialists delivering the exam owing to its droning behavior are main drawbacks of this examination . According to statistics, half of all cervical cancer cases in America occur due to not checking, with another 10% not having been tested in the prior times. The Pap smear is a well-established screening test for cervical cancer. However, the Pap smear has a relatively low accuracy.[9,10]

- The Pap smear is a relatively simple and inexpensive test.
- The Pap smear can only detect cervical cancer cells that are already present in the cervix.
- ThePap smear cannot detect precancerous cells that have the potential to develop into cervical cancer.
- **Algorithm**:

Proposed System:

Many researchers have developed automated and computer-aided diagnosis systems for cervical cancer. These systems can reduce screening time and ease the diagnosis process.[11,12] Sobar et al. achieved 91.67% accuracy with their classifier. Kashyap et al. achieved 95% accuracy with their SVM algorithm. Njoroge et al. achieved 72% accuracy with their FTIR spectroscopy-based classifier. Fazal et al. achieved 99.5% accuracy with their DBSCAN and random forest classifiers. Wu et al. achieved 99.1% accuracy with their SVM-PCA classifier. Hyeon et al. achieved 89.7% accuracy with their convolutional neural network and machine learning classifiers. In this study, eleven supervised machine learning models were employed, achieving accuracies of up to 93.33%. These findings could have a significant impact on computer-assisted diagnosis and e-healthcare systems.[13]

Advantages:

- They can be used in remote areas where access to healthcare providers is limited.
- They can be used to train healthcare providers on how to diagnose cervical cancer.
- They can be used to monitor patients who are at high risk of developing cervical cancer.

Algorithm: DTC, MLP, RFC, KNN, SVM, CatBoost (CatB), Gaussian Naïve Bayes (GNB), Gradient Boosting Classifier (GradB), AdaBoost (AdaB), XG Boost (XGB), XG Boost with Random Forest (XGBRF).

As a result, it is recommended that the world have a trustworthy computer-based system for detecting cervical cancer, as the number of patients is growing day by day. ML techniques may be ideal for implementing such a system, and they can quickly identify cervical cancer in its early stages. Cervical cancer can be detected using ML algorithms in two phases. Firstly it is utilized after meaningful feature extraction, and finally, used after ML methods that have been employed to extract important features. To increase the precision of cancer risk prediction, the ML approach uses complicated relationships between risk factors to automate the process. For early-stage cervical cancer detection, several fundamental ML approaches are being used, including support vector machine (SVM), k-nearest neighbors (KNN), random forest tree (RFT), classification, and regression tree (CART).

HPV contamination is one of the responsible factors for the development of cervical malignancy. HPV is primarily transmitted through sexual contact. Precancerous lesions take approximately 5 to 10 years to develop into malignant cells (MacCosham et al., 2020; Schiffman et al., 2007). Thereby providing an accessible time for women to go for cervical screening at least once in every three years with a Pap smear or visual inspection with acetic acid (VIA) or HPV DNA test, which assists them in diagnosing cervical cancer at an early stage. Cervical cancer is a preventable disease condition because it can be detected early by employing predictive and screening models.

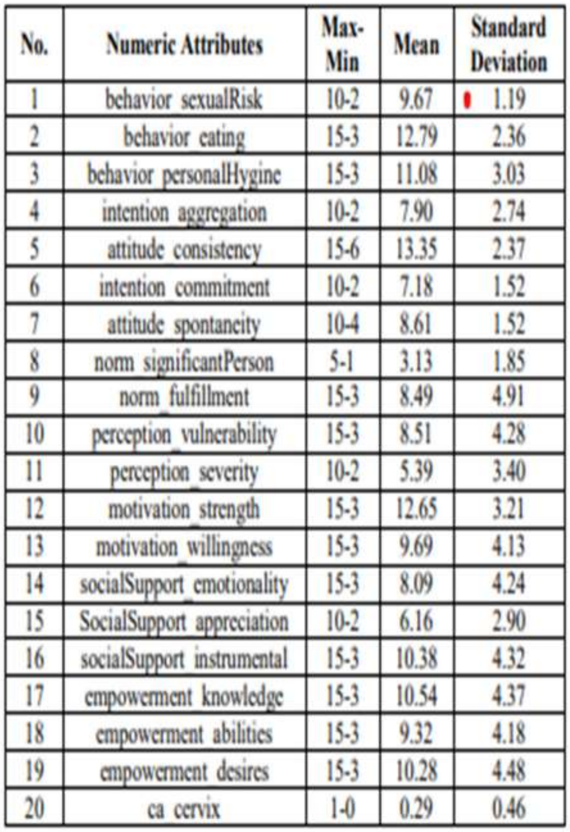
However, in developing nations, only 5% of women participate in cervical screening (Aminisani et al., 2012). Furthermore, cytological factors in Pap smear test are considered as diagnostic- predictive aspects because they identify the structure of the gland cell, squamous epithelial tissue, metaplastic cells, aberrant polymorphic cells, and dysplasia cells, as well as the existence of blood, bacteria, and fungus in the client's samples (Do et al., 2001).

IJESR

Numerous susceptible predictors of cervical carcinoma have been identified, include smoking by the individual or their partner, inadequate nutrition, immunosuppression, use of immunomodulatory drugs, prolonged contraception utilization, racial groups, deficiency of vitamin A, C and folate, having multiple sex partners, subsequent pregnancies, childbearing at a young age, low socioeconomic status, and illiteracy (Latha et al., 2014; Mandelblatt etal., 1991; Randall et al., 2016; Workowski et al., 2015).

Daily, health sector creates huge volumes of data that can be utilized to forecast future sickness based on a patient's treatment history and health information. Furthermore, by incorporating essential healthcare data, these areas can be enhanced. Non-invasive classification technologies, such as supervised machine learning (ML), are critical for cervical carcinoma prediction (Ramondetta, 2013; Workowski et al., 2015). ML in health care allows researchers to process massive amounts of complex medical data and evaluate it for curative ideas. Physicians use this material to extend medical treatment to patients. Patient satisfaction may enhance due to ML in healthcare coverage. Among the methods utilized are the Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), Neural Network (NN), Naive Bayes, Artificial Neural Network Decision tree, and Support Vector Machines (SVM) (Mukhopadhyay et al., 2016;
Nematollahi et al., 2017; Rezaianzadeh et al., 2019).

TABLE I. DIFFERENT ATTRIBUTES



No.	Numeric Attributes	Max-Min	Mean	Standard Deviation
1	behavior sexualRisk	10-2	9.67	1.19
2	behavior eating	15-3	12.79	2.36
3	behavior personalHygine	15-3	11.08	3.03
4	intention aggregation	10-2	7.90	2.74
5	attitude consistency	15-6	13.35	2.37
6	intention commitment	10-2	7.18	1.52
7	attitude spontaneity	10-4	8.61	1.52
8	norm significantPerson	5-1	3.13	1.85
9	norm_fulfillment	15-3	8.49	4.91
10	perception_vulnerability	15-3	8.51	4.28
11	perception_severity	10-2	5.39	3.40
12	motivation_strength	15-3	12.65	3.21
13	motivation_willingness	15-3	9.69	4.13
14	socialSupport_emotionality	15-3	8.09	4.24
15	SocialSupport appreciation	10-2	6.16	2.90
16	socialSupport instrumental	15-3	10.38	4.32
17	empowerment knowledge	15-3	10.54	4.37
18	empowerment abilities	15-3	9.32	4.18
19	empowerment desires	15-3	10.28	4.48
20	ca cervix	1-0	0.29	0.46

IJESR

IV.Methodology

The dataset was attained by dint of UCI ML repository and it contains information about cervical cancer risk behavior [14]. The dataset consists of 72 instances and 19 attributes, including one target column. All of the attributes are numerical type with no missing value. 2022 19th International Conference on Electrical Engineering/Electronics, Computer, Tele Communications and Information Technology (ECTI-CON) At first, exploratory data analysis was performed on the dataset, and a correlation heatmap was created and portrayed, defining the correlation between attributes.

Decision Tree (DT) algorithm

DT trails the rules of divide and conquer. In DT algorithm, the features will take up the different values called as classification trees. To resolve classification and regression issues, the classification and regression tree to be used. DT purports to have a lot of tree branches, which is why it has the Tree in its name. The DT starts with the root nodes, just as a tree originates through its roots. The leaves represent special categories, while the branches represent the mix of characteristics that result in the categorical variables. DT can also accept the continuous variables known as the regression trees. The commonly used DT algorithms in medical field are C4.5 and EC4.5 (Lilhore et al., 2022; Tiwari et al., 2018).

The dataset was divided into 80:20 ratios for training and testing ML algorithms upon exploratory data analysis. This study used two approaches: one used default hyperparameters of ML algorithms, while the other one used Grid Search Cross-Validation with a 10-fold method to tune hyperparameters. Eleven supervised machine learning algorithms, such as DTC, GNB, RFC, KNN, SVM, CatB, MLP, GradB, AdaB, XGB, XGBRF are trained using train dataset. Then the performance metrics are evaluated for both the default hyperparameter method and the tuned hyperparameter approach using the test dataset. Fig. 2 depicts the complete algorithm of the work.[15]

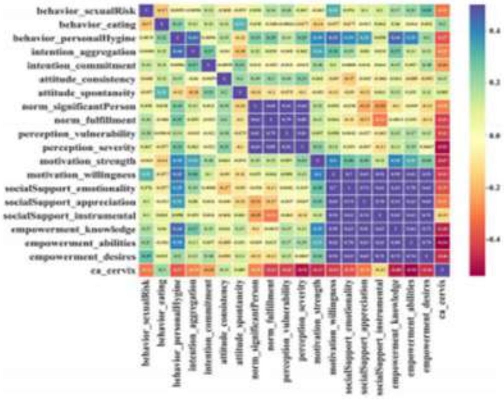


Fig. 1 Correlation heatmap of attributes

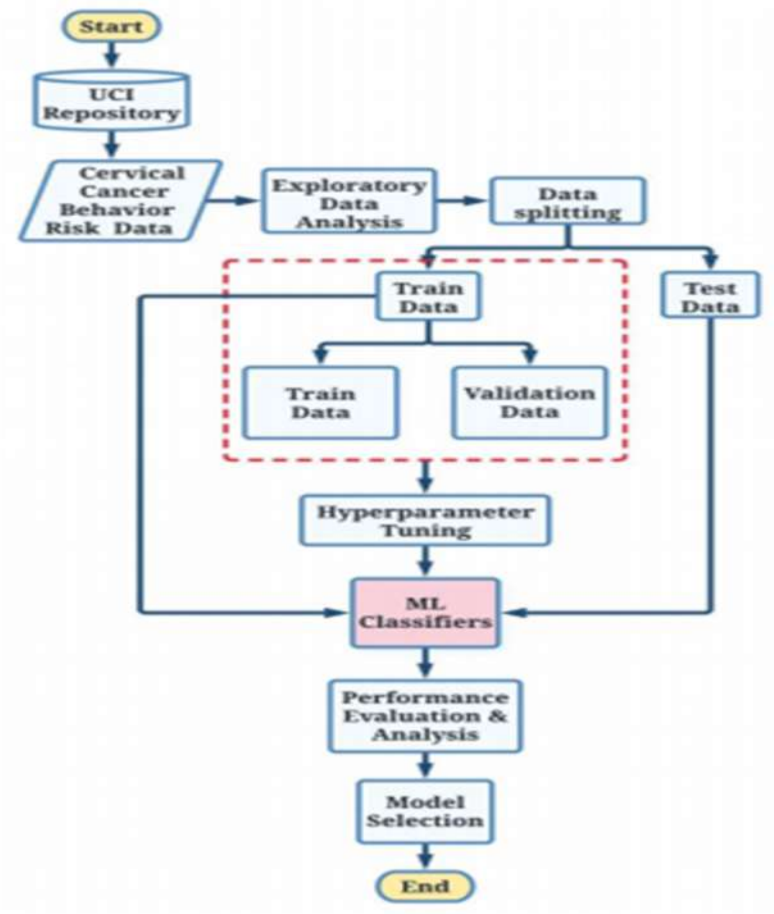


Fig. 2 Algorithm of the work

Process Mechanisms:

Cervical cancer is a significant public health concern among females worldwide. Despite being preventable, it remains a leading cause of mortality. Early detection is crucial for successful treatment and improved survival rates. This study proposes an ensemble Machine Learning (ML) classifier for efficient and accurate identification of cervical cancer using medical data.

The proposed methodology involves preparing two datasets using effective preprocessing techniques, extracting essential features using the scikit-learn package, and developing an ensemble classifier based on Random Forest, Support Vector Machine, Gaussian Naïve Bayes, and Decision Tree classifier traits.

Comparison with other state-of-the-art algorithms using several ML techniques, including support vector machine, decision tree, random forest, Naïve Bayes, logistic regression, CatBoost, and AdaBoost, demonstrates that the proposed ensemble classifier outperforms them significantly, achieving accuracies of 98.06% and 95.45% for

Dataset 1 and Dataset 2, respectively. The proposed ensemble classifier outperforms current state-of-the-art algorithms by 1.50% and 6.67% for Dataset 1 and Dataset 2, respectively, highlighting its superior performance compared to existing methods. The study also utilizes a five-fold cross-validation technique to analyze the benefits and drawbacks of the proposed methodology for predicting cervical cancer using medical data. The Receiver Operating Characteristic (ROC) curves with corresponding Area Under the Curve (AUC) values are 0.95 for Dataset 1 and 0.97 for Dataset 2, indicating the overall performance of the classifiers in distinguishing between the classes.

Additionally, Additive exPlanations (SHAP) as an Explainable Artificial Intelligence (XAI) technique to visualize the classifier's performance, providing insights into the important features contributing to cervical cancer identification. The results demonstrate that the proposed ensemble classifier can efficiently and accurately identify cervical cancer and potentially improve cervical cancer diagnosis and treatment.

V. Conclusion

Cervical cancer is utmost perilous risk for females worldwide. Early detection or risk prediction can help reduce the number of deaths caused by this disease. Numerous data are collected and evaluated to build a reliable prediction model using ML algorithms. This research compares the performance of eleven supervised ML algorithms in envisaging the menace of cervical cancer. To improve the performance of classifiers, hyperparameter tuning was performed using GSCV, and the maximum accuracy found in this study is 93.33 % by DTC, RFC, KNN, SVM, and MLP algorithms. The foremost finding of this work is the better prediction accuracy and consistency which may aid in the development and implementation of computer-aided diagnosis and serve as an efficient tool for healthcare practitioners. However, this should be thoroughly tested before being used in a clinical setting. Many future developments in this area can be accomplished by gathering more data which can contribute significantly to the establishment of an e-healthcare system.

VI. References

[1]"Cervicalcancer." https://www.who.int/news-room/factsheets/detail/cervical-cancer (accessed Feb. 19, 2022).

[2] WHO, "Comprehensive Cervical Cancer Control," Geneva, pp. 366– 378, 2014.

[3] M. Safaeian, et al., "Cervical Cancer Prevention-Cervical Screening: Science in Evolution," Obstet. Gynecol. Clin. North Am., 34(4), pp. 739–760, 2007, doi: 10.1016/j.ogc.2007.09.004.

[4] N. Colombo, et al. "Cervical cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up," Ann. Oncol., vol. 23, no.SUPPL.7,2012,doi: 10.1093/annonc/mds268.

[5] J. Doorbar, “Molecular biology of human papillomavirus infection and cervical cancer,” Clin. Sci., vol. 110, no. 5, pp. 525–541, 2006

[6] I. C. Scarinci et al., “Cervical cancer prevention: New tools and old barriers,” Cancer, vol. 116, no. 11, pp. 2531–2542, 2010

[7] D. Saslow et al., “American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer,” CA. Cancer J. Clin., vol. 62, no. 3, pp. 147–172, 2012, doi: 10.3322/caac.21139.

[8] J. Lu, et al. “Machine learning for assisting cervical cancer diagnosis: An ensemble approach,” Futur.Gener. Comput. Syst., vol. 106,pp.199205,2020,doi:10.1016/j.future.2019.12.033.

[9] B. Nithya and V. Ilango, “Evaluation of machine learning based optimized feature selection approaches and classification methods for cervical cancer prediction,” SN Appl. Sci., vol. 1, no. 6, 2019

[10] R. Weegar and K. Sundström, “Using machine learning for predicting cervical cancer from Swedish electronic health records by mining hierarchical representations,” PLoS One, 15(8), pp. 1–19, 2020

[11] F. Faisal and M. M. Nishat, "An Investigation for Enhancing Registration Performance with Brain Atlas by Novel Image Inpainting Technique using Dice and Jaccard Score on Multiple Sclerosis (MS) Tissue," Biomed. and Pharm. J., vol. 12, no. 3, pp. 1249-1262, 2019

[12] M. M. Nishat, F. Faisal, T. Hasan, M. F. B. Karim, Z. Islam and M. R. Kaysar, "An Investigative Approach to Employ Support Vector Classifier as a Potential Detector of Brain Cancer from MRI Dataset," 2021 Int. Conf. on Electronics, Comm. and Inf. Tech. (ICECIT), 2021, pp. 1-4, doi: 10.1109/ICECIT54077.2021.9641168

[13] M. M. Nishat et al., "Performance Investigation of Different Boosting Algorithms in Predicting Chronic Kidney Disease," 2020 2nd Int. Conf. on STI 4.0, pp. 1-5, 2020,doi: 10.1109/STI50764.2020.9350440

[14] M. A. A. R. Asif et al., "Performance Evaluation and Comparative Analysis of Different Machine Learning Algorithms in Predicting Cardiovascular Disease," Engineering Letters, vol. 29, no. 2, pp. 731- 741, 2021

[15] M. M. Nishat et al., “A Comprehensive Analysis on Detecting Chronic Kidney Disease by Employing Machine Learning Algorithms,” EAI Endorsed Transactions on Pervasive Health and Technology, vol. 18, no. e6, 2021, doi: 10.4108/eai.13-8-2021.170671

IJESR

[16] M. R. Farazi, F. Faisal, Z. Zaman and S. Farhan, "Inpainting multiple sclerosis lesions for improving registration performance with brain atlas," 2016 Int. Conf. on Medical Engg., Health Info. and Tech. (MediTec), 2016, pp. 1-6, doi: 10.1109/MEDITEC.2016.7835363.

[17] M. A. A. R. Asif et al., "Computer Aided Diagnosis of Thyroid Disease Using Machine Learning Algorithms," 2020 11th Int. Conf. on Electrical and Computer Engineering (ICECE), 2020, pp. 222- 225, doi: 10.1109/ICECE51571.2020.9393054

[18] M. M. Nishat and F. Faisal, "An Investigation of Spectroscopic Characterization on Biological Tissue," 2018 4th Int. Conf. on Electrical Engg. and Information & Communication Tech. (iCEEiCT), 2018, pp. 290-295, doi: 10.1109/CEEICT.2018.8628081.

[19] F. Faisal, M. M. Nishat, M. A. Mahbub, M. M. I. Shawon, and M. M. U. H. Alvi, "Covid-19 and its impact on school closures: a predictive analysis using machine learning algorithms," 2021 Int. Conf. on Science & Contemporary Technologies (ICSCT), 2021, pp. 1-6, doi: 10.1109/ICSCT53883.2021.9642617

[20] A. A. Rahman et al., "Detection of Epileptic Seizure from EEG Signal Data by Employing Machine Learning Algorithms with Hyperparameter Optimization," 2021 4th Int. Conf. on BioEngineering for Smart Technologies (BioSMART), 2021, pp. 1-4, doi: 10.1109/BioSMART54244.2021.9677770

[21] A. A. Rahman, L. I. Khalid, M. I. Siraji, M. M. Nishat, F. Faisal and A. Ahmed, "Enhancing the Performance of Machine Learning Classifiers by Hyperparameter Optimization in Detecting Anxiety Levels of Online Gamers," 2021 24th Int. Conf. on Computer and Information Technology (ICCIT), 2021, pp. 1-5, doi: 10.1109/ICCIT54785.2021.9689911

[22] T. Hasan et al., "Exploring the Performances of Stacking Classifier in Predicting Patients Having Stroke," 2021 8th NAFOSTED Conference on Information and Computer Science (NICS), pp. 242-247, 2021, doi: 10.1109/NICS54270.2021.9701526.

[23] M. M. Nishat, T. Hasan, S. M. Nasrullah, F. Faisal, M. A. A. R. Asif, and M. A. Hoque, "Detection of Parkinson's Disease by Employing Boosting Algorithms," 2021 Joint 10th Int. Conf. Informatics, Electronics & Vision (ICIEV) and 2021 5th Int. Conf. on Imaging, Vision & Pattern Recognition (icIVPR), pp. 1-7, 2021, doi: 10.1109/ICIEVicIVPR52578.2021.9564108

IJESR

[24] Nishat, M. M., et al. “Detection of Autism Spectrum Disorder by Discriminant Analysis Algorithm,” BIM, pp. 473-482, Springer, Singapore, 2022 doi: 10.1007/978-981-16-6636-0_36

[25] A. A. Rahman, M. I. Siraji, L. I. Khalid, F. Faisal, M. M. Nishat and M. R. Islam, "Detection of Mental State from EEG Signal Data: An Investigation with Machine Learning Classifiers," 2022 14th International Conference on Knowledge and Smart Technology (KST), pp. 152-156, 2022, doi: 10.1109/KST53302.2022.9729084.

[26] M. R. Rahman, S. Tabassum, E. Haque, M. M. Nishat, F. Faisal and E. Hossain, "CNN-based Deep Learning Approach for Micro-crack Detection of Solar Panels," 2021 3rdInt. Conf. on STI 4.0, 2021, pp. 1- 6, doi: 10.1109/STI53101.2021.9732592.

[27] Nishat, M. M., et al. “A Comprehensive Investigation of the Performances of Different Machine Learning Classifiers with SMOTE-ENN Oversampling Technique and Hyperparameter Optimization for Imbalanced Heart Failure Dataset,” Scientific Programming, https://doi.org/10.1155/2022/3649406

[28] Nishat, M. M, Faisal, F., Mahbub, M. A., Mahbub, M. H., et al., “Performance Assessment of Different Machine Learning Algorithms in Predicting Diabetes Mellitus”. Biosc.Biotech.Res.Comm.,14(1), 2021 doi: http://dx.doi.org/10.21786/bbrc/14.1/10.

[29] Islam, Mahdi, Musarrat Tabassum, Mirza Muntasir Nishat, Fahim Faisal and Muhammad Sayem Hasan, “Real-Time Clinical Gait Analysis and Foot Anomalies Detection Using Pressure Sensor and Convolutional Neural Network” 2022 7th International Conference on Business and Industrial Research (ICBIR), IEEE. accepted, in press.

[30] Rahman, A. A., Siraji, M. I., Khalid, L. I., Faisal, F., Nishat, M. M., Ahmed, A., Mamun, M. A. A., “Perceived Stress Analysis of Undergraduate Students During COVID-19: A Machine Learning Approach,” 2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON), Palermo, Italy, accepted, in press

[31] F. Faisal, M. M. Nishat and M. A. M. Oninda, "Spectroscopic Characterization of Biological Tissue using Quantitative Acoustics Technique," 2018 4th Int. Conf. on Electrical Engineering and Information & Communication Technology (iCEEiCT), pp. 38-43, 2018, doi: 10.1109/CEEICT.2018.8628146.

[32] Sobar, R. Machmud, and A. Wijaya, “Behavior Determinant Based Cervical Cancer Early Detection with Machine Learning Algorithm,” Adv. Sci. Lett., 22,(10), pp. 3120–3123, 2016

[33] K. R. Debashree Kashyap, et al. “Cervical Cancer Detection And Classification Using Independent Level Sets And Multi SVMs,” Biomed. Pharmacol. J., 9(2), pp. 663–671, 2016

IJESR

[34] E. Njoroge, et al. "Classification of cervical cancer cells using FTIR data," Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc., pp. 5338–5341, 2006, doi: 10.1109/IEMBS.2006.260024.

[35] Y. S. Muhammad Fazal Ijaz, Muhammad Attique, "Data-Driven Cervical Cancer Prediction Model with Outlier Detection and OverSampling Methods," Sensors, vol. 20, no. 10, pp. 1424–8220, 2020.

[36] W. Wu and H. Zhou, "Data-driven diagnosis of cervical cancer with support vector machine-based approaches," IEEE Access, vol. 5, pp. 25189–25195, 2017, doi: 10.1109/ACCESS.2017.2763984

[37] J. Hyeon, H. J. Choi, K. N. Lee, and B. D. Lee, "Automating papanicolaou test using deep convolutional activation feature," Proc. - 18th Int. Conf. Mob. Data Manag. MDM 2017, pp. 382–385, 2017

[38] "UCI Machine Learning Repository: Cervical Cancer Behavior Risk Data Set." https://archive.ics.uci.edu/ml/datasets/Cervical+Cancer+Behavior+Ri sk

IJESR