Effective Heart Disease Prediction using IBM

Auto AI Services

Introduction

Cardiovascular diseases (CVDs) are the number 1 cause of death globally, taking an estimated 17.9 million lives each year, which accounts for 31% of all deaths worldwide. Heart failure is a common event caused by CVDs. According to WHO data, heart disease is the leading cause of mortality globally, resulting in 17.9 million deaths annually [1]. The most behavioural risk factors for cardiovascular disease and stroke are unhealthy food, lack of physical activity, smoking, and alcohol drinking [1]. A heart attack occurs when the heart's blood circulation is obstructed by arteries plaque build-up. A thrombus in an artery causes a stroke by impeding blood flow to the brain [2]. The symptoms are common to other illnesses and might be confused with indicators of ageing, making diagnosis difficult for practitioners. Precision prediction and timely identification of cardiac disease are essential for improving patient survival rate. Because of the increased collection of medical data, practitioners now have a great opportunity to promote healthcare diagnosis. ML, which is a part of AI plays a vital role in many applications like text detection and recognition [3], early prediction [4], power quality disturbance detection [5], truck traffic classification [6], and agriculture [7]. ML has now become an essential tool in the healthcare sector to aid with patient diagnosis. The current methods for predicting and diagnosing cardiac disease are mostly dependent on practitioners' evaluation of a patient's medical history, signs, and physical assessment reports. Nowadays, information about patients with clinical reports is widely accessible in databases in the healthcare field, and it is rising rapidly day by day. In this article, the UCI ML repository's Cleveland HD dataset was utilized for developing the prediction model to heart disease. The machine is trained for learning patterns

based on the features that are already present in the dataset. Classification is an effective ML approach for prediction. When properly trained with adequate data, classification is an effective supervised ML method for identifying disease [8].

Literature Survey

Nave Bayes, random forest, PART, C4.5, and multilevel perceptron algorithm-based predictive model accuracy to HD dataset were determined to be in the range of 75.58%– 83.17% [9]. Moreover, Nave Bayes algorithm has the highest accuracy as 83.17%, while other algorithms have less than 80% accuracy [9]. Kumar et al. discovered that the Random Woodland ML classifier had an 85 percent precision for cardiovascular disease [10]. Gudadhe *et al*. [11] described the framework for predicting the heart disease using SVM and obtained the accuracy as 80.41%. Kahramanli and Allahverdi [12] combined fuzzy and crisp values in health data and attained accuracy rates of 84.24% to Pima Indian diabetes dataset and 86.8% for the Cleveland HD dataset, respectively. Various ML classification models [13–17] could be used to improve intelligence. Kahramanli and Allahverdi [12] established the artificial and fuzzy-based model to the Pima Indian diabetes dataset and the Cleveland HD dataset and found 84.24% and 86.8% accuracy, respectively. Olaniyi *et al*. [18] established a prediction model and achieved an accuracy of 85% using feedforward multilayer perceptron (MLP) and 87.5% using SVM on the UCI ML datasets. Polat *et al*. [19] have employed k-nearest neighbour algorithm and an artificial immune recognition framework and achieved 87% accuracy on the Cleveland dataset. On a Cleveland dataset, Detrano et al. [20] achieved 77% using the logistic regression algorithm. Saw *et al*. [21] have implemented the improved logistic regression classification model for heart disease dataset. The fast decision tree and C4.5 tree have been employed for HD prediction [22]. As a result of the proposed model's initial phase, trees and features have been extracted. The genetic and fuzzy logic-based approach has been proposed [23] which is a hybrid model to instantly generate the rules using a fitness function, appropriate genetic operators, and a rule encoding method. Teekaraman *et al*., employed contemporary ML techniques to construct healthcare heart disease predictive model. The Cleveland HD dataset was subjected to SVM with radial basis function (RBF) kernel, Gaussian Naive Bayes, logistic regression, Light GBM, XG Boost, and random forest algorithm, and the best performing prediction model for early diagnosis of heart disease was found by Teekaraman *et al* ., [24]

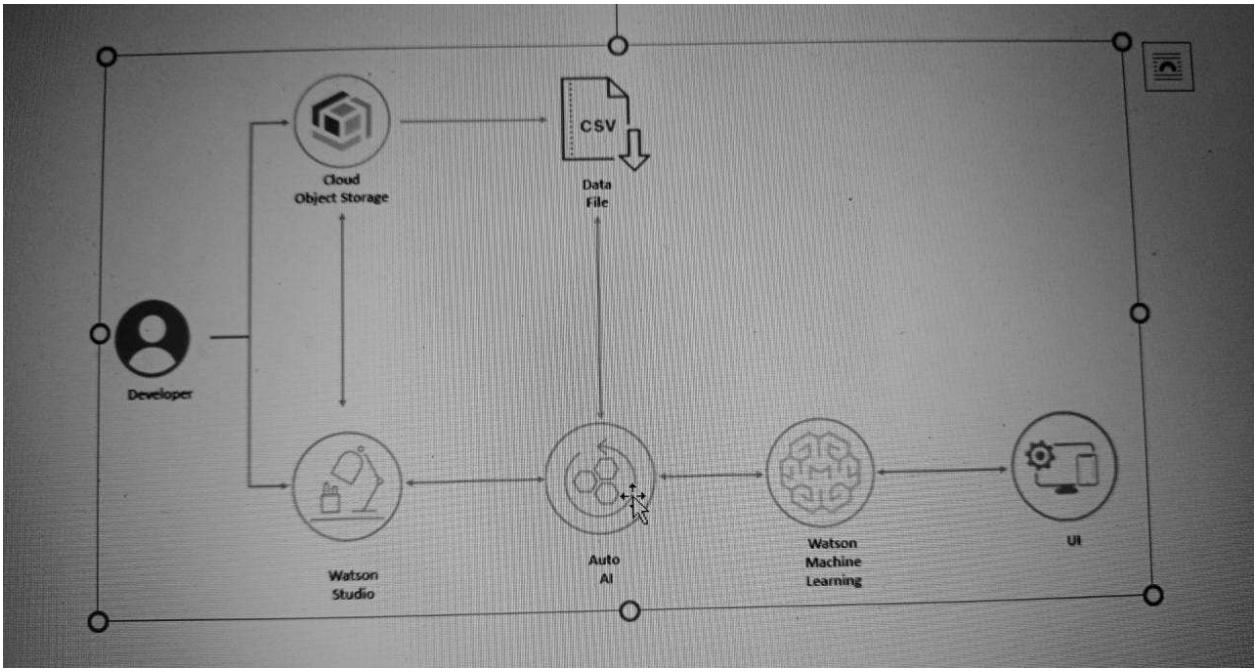
In this Project I have built a model using Auto AI and build a webpage application where we can showcase the prediction of Heart failure.

Services Used:

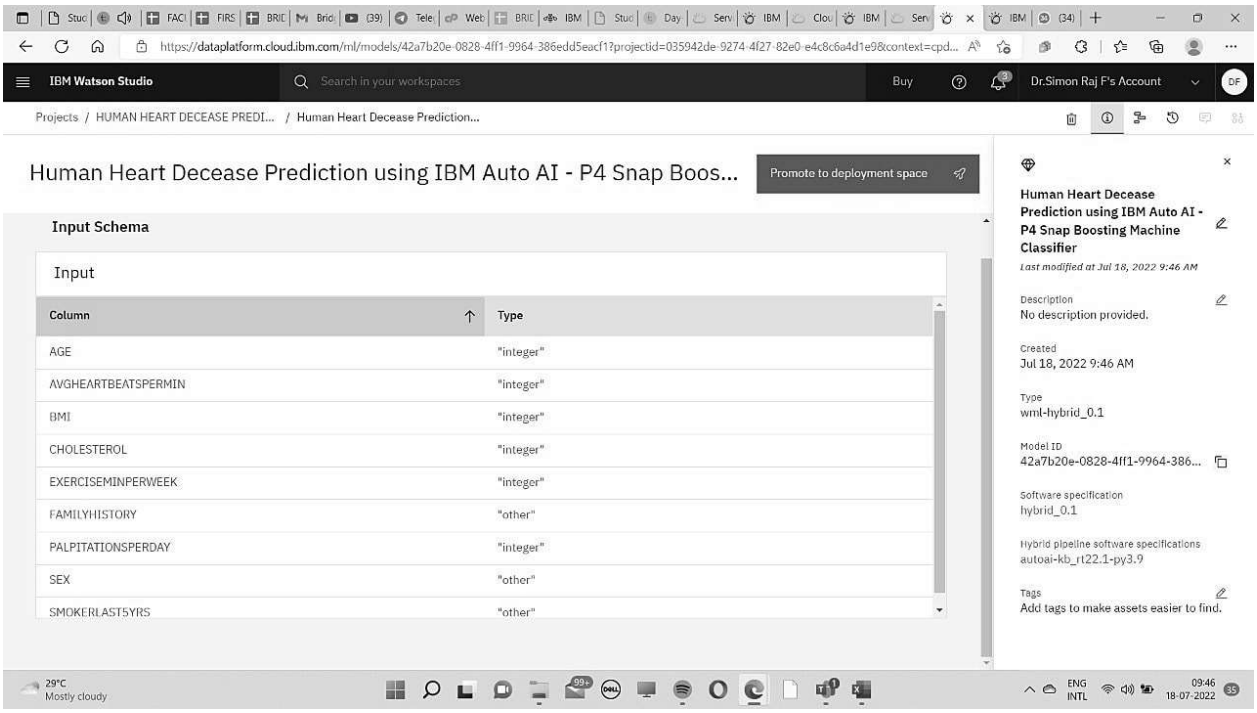
IBM Watson Studio

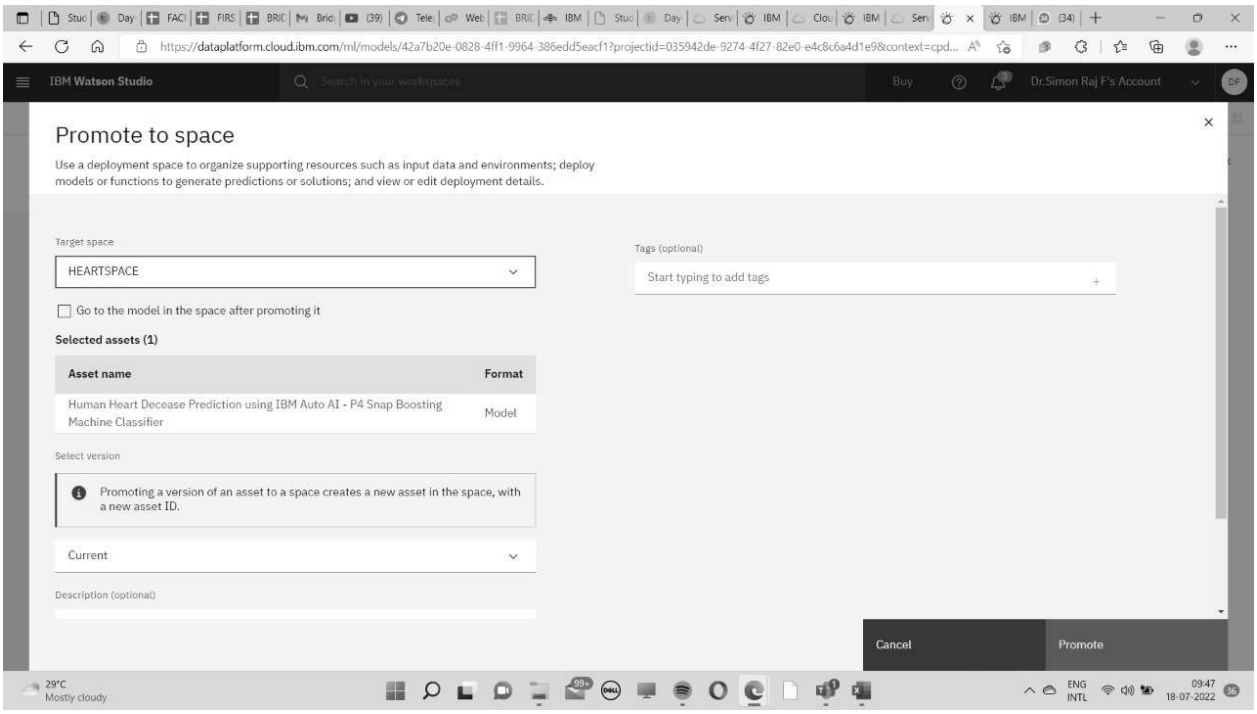
- IBM Watson Machine Learning
- Node-RED
- IBM Cloud Object Storage

Technical Architecture:

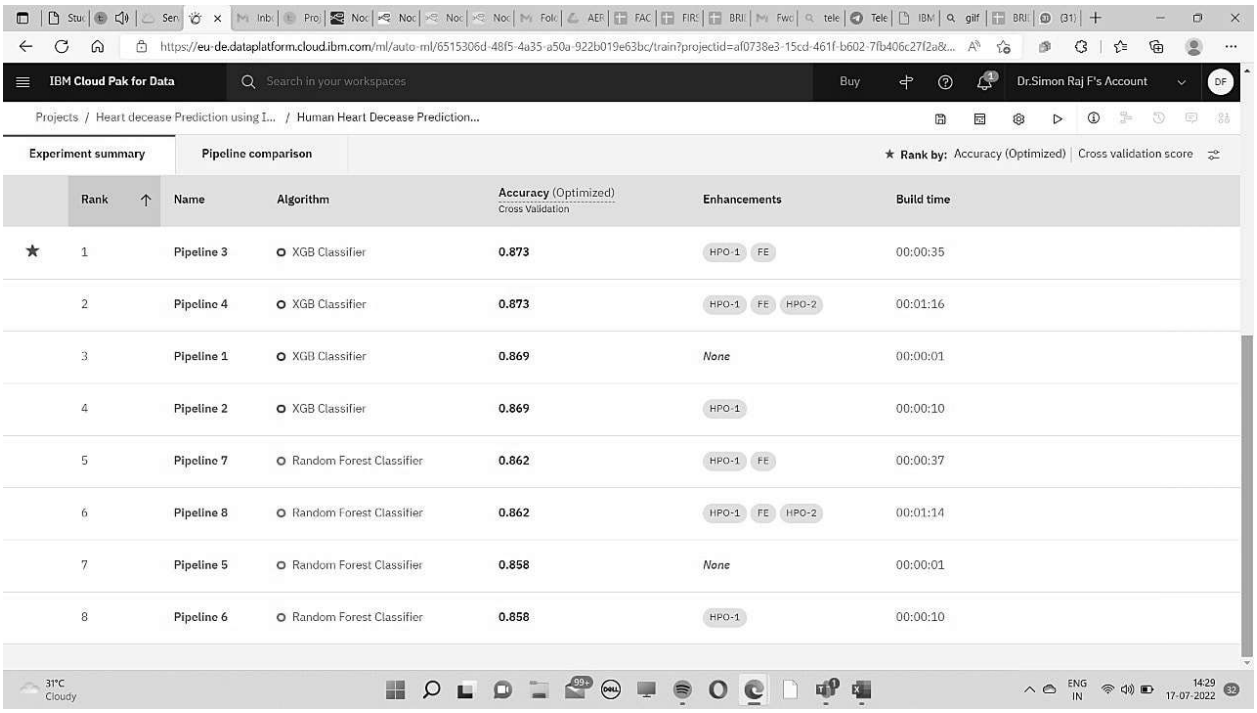


Effective Heart decease Pridiction using Auto AI

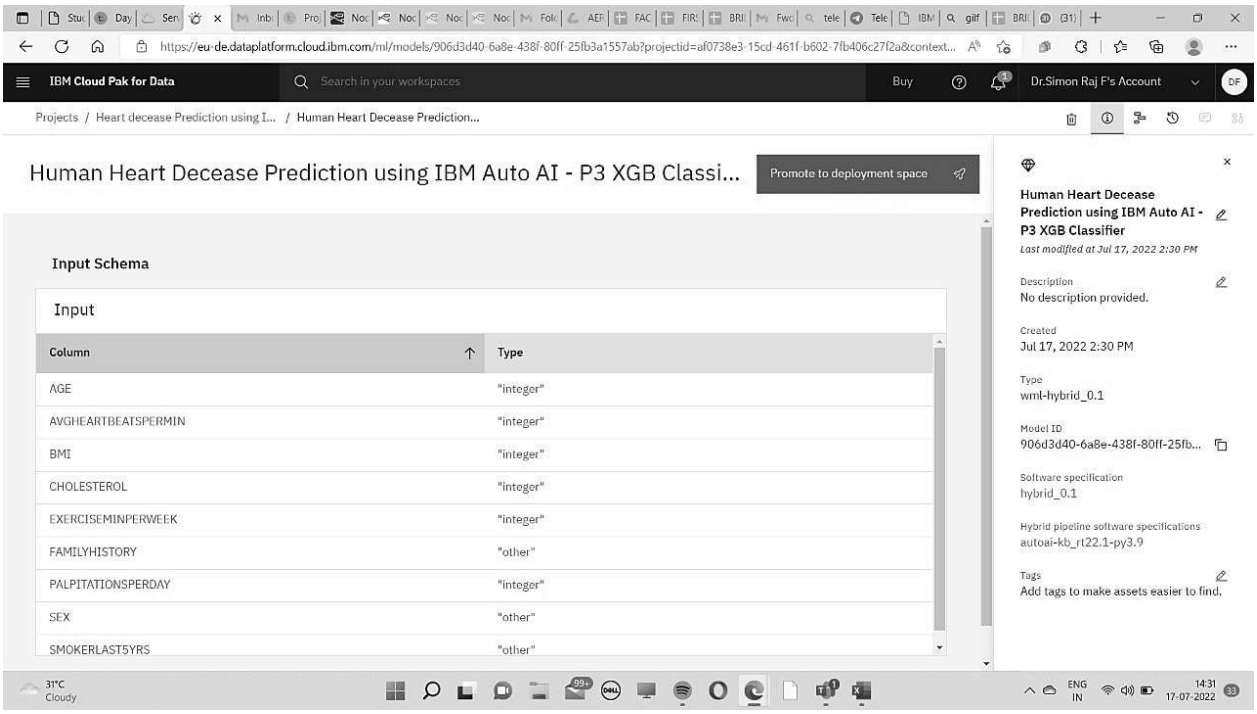




Effective Heart decease Pridiction using Auto AI

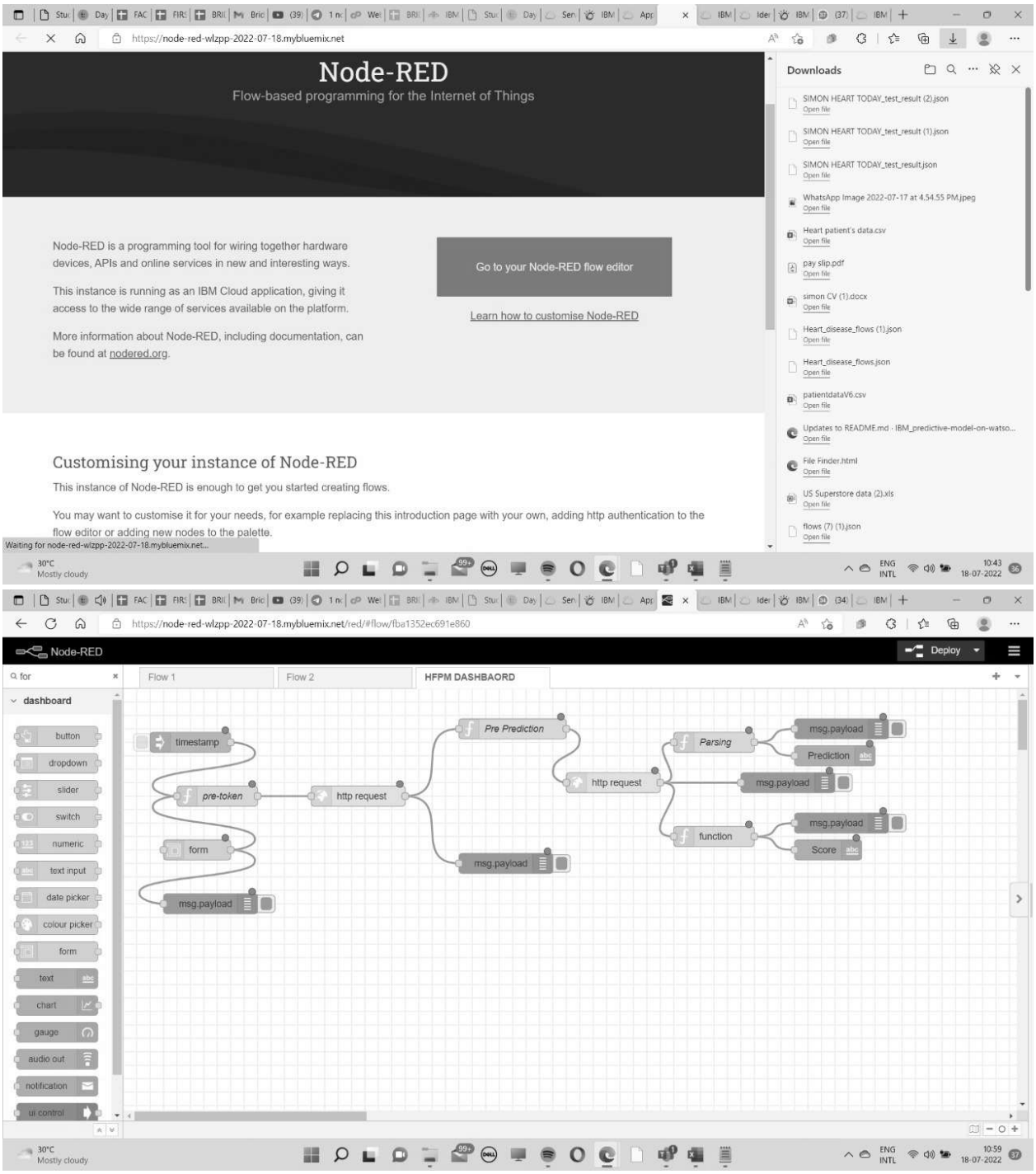


	Rank	Name	Algorithm	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 3	XGB Classifier	0.873	HPO-1 FE	00:00:35
	2	Pipeline 4	XGB Classifier	0.873	HPO-1 FE HPO-2	00:01:16
	3	Pipeline 1	XGB Classifier	0.869	None	00:00:01
	4	Pipeline 2	XGB Classifier	0.869	HPO-1	00:00:10
	5	Pipeline 7	Random Forest Classifier	0.862	HPO-1 FE	00:00:37
	6	Pipeline 8	Random Forest Classifier	0.862	HPO-1 FE HPO-2	00:01:14
	7	Pipeline 5	Random Forest Classifier	0.858	None	00:00:01
	8	Pipeline 6	Random Forest Classifier	0.858	HPO-1	00:00:10

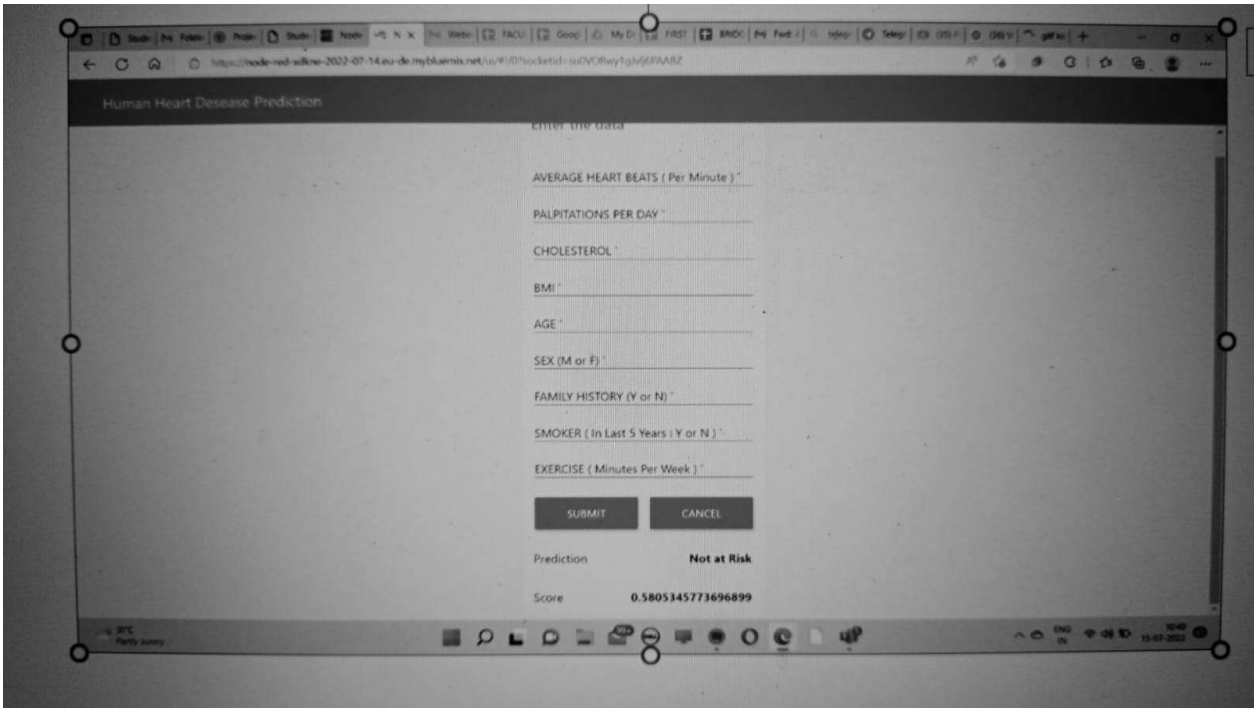


Column	Type
AGE	"integer"
AVGHEARTBEATSPERMIN	"integer"
BMI	"integer"
CHOLESTEROL	"integer"
EXERCISEMINPERWEEK	"integer"
FAMILYHISTORY	"other"
PALPITATIONSPERDAY	"integer"
SEX	"other"
SMOKERLAST5YRS	"other"

Effective Heart decease Pridiction using Auto AI



OUTPUT



Conclusion

In this project we have made webpage application model for predicting Human Heart decease prediction using IBM cloud-based Auto AI Techniques. Using such model and using huge collection of medical data, we can give warning or alert for patient using their medical record.

References:

[1] Who link: https://www.who.int/health-topics/cardiovascular[1]diseases/#tab=tab_1.

[2] J. L. Quah, S. Yap, S. O. Cheah et al., “Knowledge of signs and symptoms of heart attack and stroke among Singapore resi[1]dents,” BioMed Research International, vol. 2014, 2014.

[3] K. Karthick, S. K. Aruna, and R. Manikandan, “Development and evaluation of the bootstrap resampling technique based statistical prediction model for Covid-19 real time data: a data driven approach,” Journal of Interdisciplinary Mathematics, pp. 1–13, 2022.

[4] K. Kanagarathinam and K. Sekar, “Estimation of reproduction number and early prediction of 2019 novel coronavirus disease (COVID-19) outbreak in India using statistical computing approach,” Epidemiology and Health, vol. 42, 2020.

[5] K. Sekar, S. Kumar, and K. K, “Power quality disturbance detection using machine learning algorithm,” in 2020 IEEE International Conference on Advances and Developments in Electrical and Electronics Engineering (ICADEE), pp. 1–5, 2020.

[6] N. Tahaei, J. J. Yang, M. G. Chorzepa, S. S. Kim, and S. A. Dur[1]ham, “Machine learning of truck traffic classification groups from weigh-in-motion data,” Machine Learning with Applica[1]tions, vol. 6, p. 100178, 2021.

[7] V. Meshram, K. Patil, V. Meshram, D. Hanchate, and S. D. Ramkteke, “Machine learning in agriculture domain: a state[1]of-art survey,” Artificial Intelligence in the Life Sciences, vol. 1, p. 100010, 2021.

[8] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, “Compar[1]ing different supervised machine learning algorithms for dis[1]ease prediction,” BMC medical informatics and decision making, vol. 19, no. 1, p. 281, 2019.

[9] C. B. C. Latha and S. C. Jeeva, “Improving the accuracy of pre[1]diction of heart disease risk based on ensemble classification techniques,” Informatics in Medicine Unlocked, vol. 16, p. 100203, 2019. [10] N. K. Kumar, G. S. Sindhu, D. K. Prashanthi, and A. S. Sulthana, “Analysis and prediction of cardio vascular disease using machine learning classifiers,” in 2020 6th International Conference on Advanced Computing and Communication Sys[1]tems (ICACCS), pp. 15–21, 2020.

[11] M. Gudadhe, K. Wankhade, and S. Dongre, “Decision support system for heart disease based on support vector machine and artificial neural network,” in 2010 International Conference on Computer and Communication Technology (ICCCT), pp. 741– 745, 2010. Computational and Mathematical Methods in Medicine 13.

[12] H. Kahramanli and N. Allahverdi, “Design of a hybrid system for the diabetes and heart diseases,” Expert systems with applications, vol. 35, no. 1-2, pp. 82–89, 2008.

[13] A. Gupta, Anjum, S. Gupta, and R. Katarya, “InstaCovNet-19: a deep learning classification model for the detection of COVID-19 patients using chest X-ray,” Applied Soft Comput[1]ing, vol. 99, p. 106859, 2021.

[14] I. H. Sarker, “Machine learning: algorithms, real-world applications and research directions,” SN COMPUT. SCI., vol. 2, no. 3, p. 160, 2021.

[15] Y. Pathak, P. K. Shukla, A. Tiwari, S. Stalin, S. Singh, and P. K. Shukla, “Deep transfer learning based classification model for COVID-19 disease,” Ing Rech Biomed, vol. 22, 2022.

[16] R. Muazu Musa, A. P. A. Majeed, Z. Taha, S. W. Chang, A. F. A. Nasir, and M. R.

Abdullah, “A machine learning approach of predicting high potential archers by means of physical fitness indicators,” PLoS One, vol. 14, no. 1, p. e0209638, 2019.

[17] P. Kota, A. Madenahalli, and R. Guturi, “Heart disease classification comparison among patients and normal subjects using machine learning and artificial neural network techniques,”International Journal of Biosensors & Bioelectronics, vol. 7, no. 3, 2021.

[18] E. O. Olaniyi, O. K. Oyedotun, and K. Adnan, “Heart diseases diagnosis using neural networks arbitration,” International Journal of Intelligent Systems and Applications (IJISA), vol. 7, no. 12, pp. 75–82, 2015.

[19] K. Polat, S. Sahan, and S. Günes, “Automatic detection of heart disease using an artificial immune recognition system (AIRS) with fuzzy resource allocation mechanism and k-nn (nearest neighbour) based weighting preprocessing,” Expert Systems with Applications, vol. 32, no. 2, pp. 625–631, 2007.

[20] R. Detrano, A. Janosi, W. Steinbrunn et al., “International application of a new probability algorithm for the diagnosis of coronary artery disease,” The American journal of cardiology, vol. 64, no. 5, pp. 304–310, 1989.

[21] M. Saw, T. Saxena, S. Kaithwas, R. Yadav, and N. Lal, “Estimation of prediction for getting heart disease using logistic regression model of machine learning,” in 2020 International Conference on Computer Communication and Informatics (ICCCI), pp. 1–6, 2020.

[22] R. El-Bialy, M. A. Salamay, O. H. Karam, and M. E. Khalifa, “Feature analysis of coronary artery heart disease data sets,” Procedia Computer Science, vol. 65, pp. 459–468, 2015.

[23] K. Mankad, P. S. Sajja, and R. Akerkar, “Evolving rules using genetic fuzzy approach - an educational case study”, International Journal on Soft Computing (IJSC), vol. 2, no. 1, pp. 35–46, 2011.

[24] K. Karthick, S. K. Aruna , R. Samikannu , R. Kuppusamy , Y. Teekaraman , and A. Thelkar, “Implementation of a Heart Disease Risk Prediction Model Using Machine Learning”, Computational and Mathematical Methods in Medicine Volume 2022, Article ID 6517716, 14 pages https://doi.org/10.1155/2022/6517716