ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

DOI:10.48047/IJFANS/V11/I12/193

Machine Learning Algorithm for Brain Stroke Detection

M. Kishore Babu[1], Assistant Professor, Department of CSE,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.
Sk. Reshma[2], **V. Balakota Reddy**[3], **U. Jayanth**[4], **Sk. Naheda**[5]
[2,3,4,5] UG Students, Department of CSE,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.
[1,2,3,4,5]kislatha@gmail.com,shaikreshma9505@gmail.com,
vemireddy77802gmail.com, jayanth14962@gmail.com,nahedashaik2000@gmail.com

Abstract

Insufficient blood flow to the brain results in a condition known as a stroke, which results in cell death. In worldwide it is currently the leading cause of death. Many risk factors that are suspected to be related to the stroke's origin have been identified through examination of the affected individuals. Using these risk factors, numerous research has been done to predict the disorders linked to stroke. Most models are built using machine learning techniques and data mining. In this study, we used data from medical reports and a person's physical condition to use five machine learning algorithms to identify strokes. We use a substantial number of hospital entries that we have collected. The classification outcome demonstrates that the result is satisfactory and can be applied to real-time medical records in order to address the issues. We think machine learning algorithms can aid in better understanding illnesses and make a useful healthcare partner.

Keywords: Brain stroke, Cat boost, machine learning algorithms, MLP, Random Forest.

Introduction

People's health is regarded as being important to their lives, it is imperative to have a system in place for recording data about illnesses and how they are related. Stroke is a serious medical condition that requires urgent attention to prevent permanent d amage. Early detection and treatment are crucial for the successful recovery of stroke patients. Machine learning has the potential to improve stroke diagnosis by analysing medical images and identifying subtle signs of stroke that may be missed by human radiologists. Most of the information about diseases can be found in patient case summaries, clinic records, and other manually preserved data. The sentences in them may be deciphered using a variety of machine learning techniques (ML) and text mining. Information retrieval uses a process called machine learning to distribute content, giving the semantic and syntactic aspects of the content precedence. For feature extraction and implementation, a variety of ML and text mining algorithms are reviewed and put into practise. Medical practitioners often refer to it as a stroke when blood flow to the brain or spinal cord is impaired, causing damage to the tissues or cells. Many viewpoints affect how

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

a stroke is seen, yet a stroke clearly causes a visceral reaction. In fields like surveillance and medicine, machine learning may be used as a major tracker with the use of correctly trained algorithms. One suggestion calls for gathering patient symptoms from their medical data and using them to educate the system. The suggested stemmed extracted the common and distinctive collection of features to detect stroke disease after mining the case sheets using maximum entropy and tagging approaches. The information was then used to train several machine learning algorithms, such as Support vector machines, Decision trees, Logistic Regression, Random Forest,K-Nearest Neighbors, and Logistic Regression using Logistic Regression. One of these techniques, the support vector machine, achieves excellent accuracy.

Literature Survey

1] Badriyah, Tessy et al. Data can be analyzed and used for decision making. This can be done in several ways like using the deep learning method which is becoming more and more popular these days as it has proven to be very effective in solving various problems. The pioneers of deep learning itself started in 1980 when Kuniko Fukushima proposed Neo Cognition, the first model of convolutional neural networks, followed by Yann LeCun, Leon Bottou, Joshua Bengio, and Patrick Haffner improved it. The basis of deep learning itself is ANN. The image, video, and image text data warehouse are very large and uses a lot of computing power. Like other neural networks, deep learning architectures usually consist of visible and hidden layers, where the weights of each perceptron unit are optimized using the backpropagation algorithm. Examples of algorithms applying deep learning include Deep Convolutional Neural Networks (DCNN) for image classification, Deep Belief Networks - Deep Neural Networks (DBN - DNN) for speech recognition, recurrent neural networks (RNN) for language translation, query-oriented deep mining (QODE) which is based on restricted Boltzmann machine (RBM) for multi-document synthesis, conditional restricted Boltzmann machine for drug-target interaction (DTI) prediction (RBM) and deep belief networks (DBN) for prediction of time series data.

[2] N. Venketasubramanian, B. W. Yoon In many nations, stroke is a major cause of mortality and disability. According to statistics, there were 10.3 million new stroke cases, 113 million disability-adjusted life years (DALYs) lost as a result of stroke, 6.5 million stroke deaths, and roughly 25.7 million stroke survivors worldwide in 2013. The bulk of stroke deaths and lost DALYs—75.2% of all stroke-related deaths and 81.0% of all associated DALYs—were seen in developing nations. Stroke is a significant issue in Asia, which is home to more than 60% of the world's population and many of its "emerging" nations. Except for a few nations like Japan, Asia has a higher stroke death rate than the Americas,Western Europe, or Australasia.

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

A review of stroke epidemiology in East and South Asia has already been done. Three of the 11 Southeast Asian countries with predominantly South Asian and East Asian populations were recently included in a stroke study of 12 Asian countries. However, statistics are not available for other countries in Southeast Asia. Based on data from the Global Burden of Disease (GBD) study, World Health Organization (WHO) and recent key publications from Asian countries, this study was conducted to review the recent epidemiology of stroke in South, East, and South-East Asia, including mortality, morbidity, prevalence, lost DALYs, stroke subtypes and risk factors. A better understanding of stroke prevalence in these regions can aid in health planning and resource allocation, as well as understanding the extent and variability of stroke.

[3] G. A. P. Singh and P. K. Gupta Among all cancer-related disorders, One of the main causes of death is lung cancer (Cancer Research UK, Cancer Mortality for Common Cancers). It scans a patient's lungs and the results are most often used to diagnose a condition. An X-ray, CT or MRI may be the subject of this scan analysis. Due to the variety of imaging modalities used to visualize a patient's lungs, one of the challenges is the automatic classification of lung cancer. Detection and classification of lung cancer have been shown to have great potential using image processing and machine learning techniques. In this study, we present an efficient method to identify CT images associated with lung cancer and classify them into benign and malignant categories. The proposed method first applies image processing techniques before classifying photos with supervised learning algorithms. Here, we combined the statistical features with the extracted texture features and passed the various extracted features to the classifier. We have a k-nearest neighbor classifier, a support vector classifier, a decision tree classifier, a multinomial naive Bayes classifier, a stochastic gradient descent classifier, a random forest classifier, and a multilayer perceptron (MLP) classifier. classifier used. To train and evaluate these classifiers, we used a dataset of 15,750 clinical pictures containing both 6,910 benign images and 8,840 malignant images related to lung cancer. The collected results show that the MLP classifier has higher accuracy than other classifiers with a score of 88.55%.

[4] C. L. Chin et al. Over the past few years, stroke has consistently ranked among the top 10 most common deaths in Taiwan. Symptoms of the stroke require emergency treatment. The faster the patient is treated, the better the patient's recovers or chances of recovery. The diagnosis must, however, rely on physicians to evaluate the imaging because the site of the ischemic stroke in the CT scan is not evident. This study aims to create a CNN deep learning algorithm-based early ischemic stroke detection system. The system will start picture pre-processing after receiving the CT scan of the brain to eliminate any areas that cannot possibly be a stroke region. Finally, in order to expand the number of patch photos, we'll choose a few and apply the Data Augmentation

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

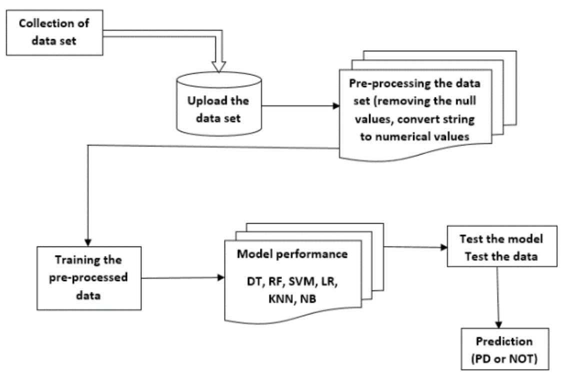
approach In order to train and test the convolutional neural network, we will finally input the patch photos. [15-22] In this study, we trained and evaluated a CNN module's capacity to recognize ischemic strokes using 256 patch images. According to the experimental findings, the accuracy of the suggested method is greater than 90%.

Problem Identification

With the increasing popularity of machine learning, computer approaches are classified into two categories: traditional methods and machine learning methods. The related efforts of brain stroke detection classification are discussed in this part. Machine Learning Model Detection and how machine learning methods outperform older methods. For model development, the present procedure in this project has a specific flow. In the existing system, methods such as logistic regression and naive bias are applied. However, it necessitates a huge memory and produces inaccurate results.

Proposed Methodology

In proposed system, Support Vector Machine(SVM), Random Forest, XG Boost, and Decision Tree are a portion of the AI calculations utilized in the proposed framework. In our proposed framework, we analyse these few strategies and select the model with the most elevated exactness. The ideal model for our dataset will be chosen by usdependent on the precision score and estimate assuming that the patient subtleties are causing stroke or not.



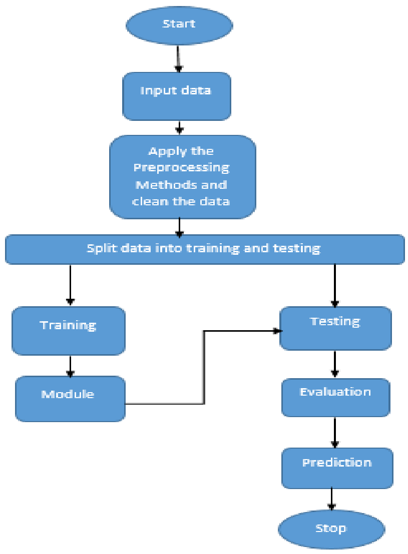


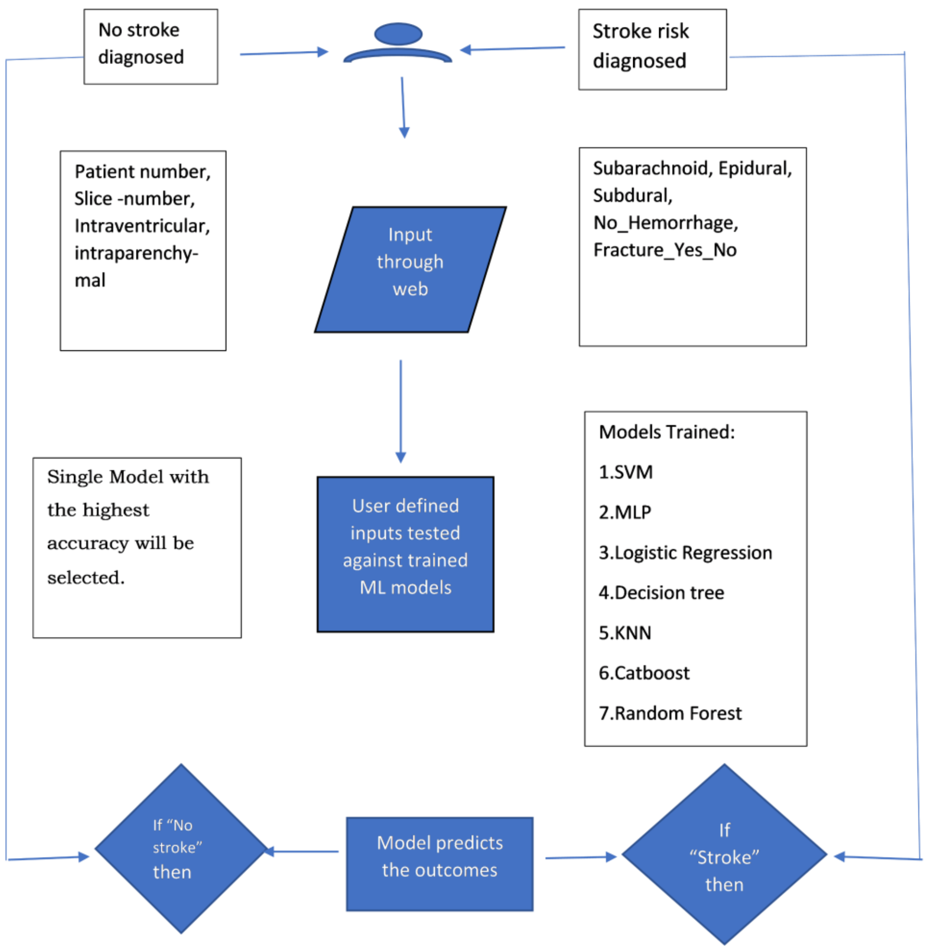
Fig: Flow chart of the project

Fig:Proposed Method

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Implementation



Then, we must register on our website. If it is already registered, we can log in immediately. We then need to load the data. The processed data will be made available after the data has been checked for any nan values. On this, data will be viewed. The provided inputs are being used to train the ML models. The system will then determine if the person is likely to experience a brain stroke based on the provided data.

In the implementation, we are using eight Machine Learning algorithms. And the different values will be detected as outcome.

The accuracy, precision, recall and f1-score values are the outputs. Based on the accuracy of the trained models, the best model will be taken and the further process will be continued.

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

The accuracy given by the trained models are shown:

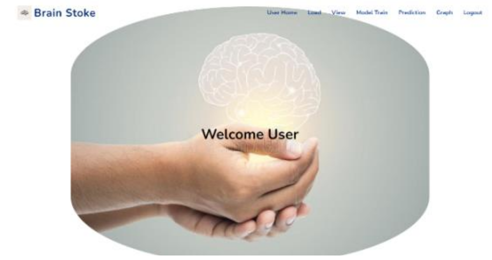
Model	Accura-cy	Precisi-on	Rec-all	F1_sco-re
KNN	0.9407	0.8946	1.0	0.9443
Logistic Regression	0.8259	0.8862	0.75	0.8125
Decision Tree	.9855	.9842	.987	.9857
Naïve Bayes	.8215	.8838	.743	.8072
Random Forest	.9725	.9687	.977	.9729
MLP Classifier	.8215	.8918	.734	.8054
SVC	.7558	.6941	.9195	.7911
Cat Boost	.9863	.9857	.987	.9864

Results and Discussions

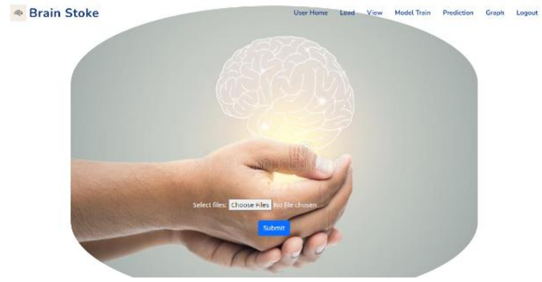
Home Page: Home page of This Project

About: It's a Small information about this Project

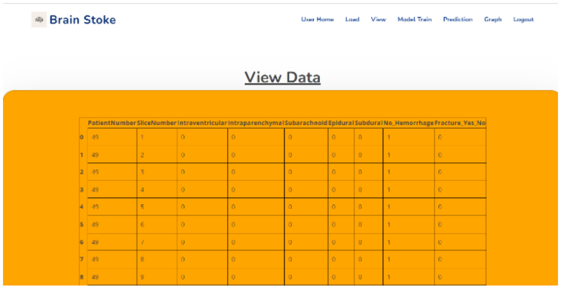
User Home page: User can view the home page after successful login.



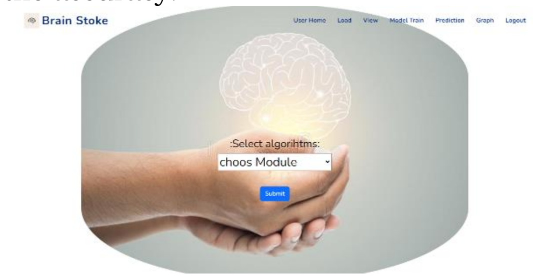
Upload data: User can upload the dataset.



View Data: User can View the Data.



Model: User can select the model and view the accuracy.

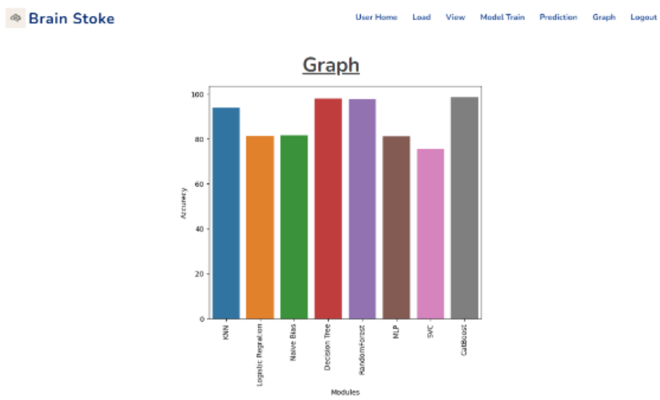


ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Prediction: User can give the input view the predicted result.

Graph: Graph of accuracy of KNN, Logistic Regression, Naïve Bias, Decision Tree, Multi-layer perc.. eptron (MLP),Random Forest, SVC, Cat Boost.



Conclusion

The investigation includes the utility of collection estimations for coordinated substances, for instance, patient case sheets in requesting strokes reliant upon communicated limits (indications) and components. Considering portrayal moves close, this survey predicts the kind of stroke a patient will have. This audit shows that stroke occurs more frequently in men than in women, and in those between the ages of 40 and 60. Individuals who experienced an ischemic stroke differed more from those who experienced a haemorrhagicstroke. The patient's exposure to changeable and immutable risk factors, as well as the particular patient's indications, are taken into account when determining the kind of stroke.

Limitations and Future scope

The Proposed method doesn't specify the type of stroke, and it can use different datasets for further scope to improve the performance of the system.

References

[1] Roger VL, Go AS, Lloyd-Jones DM, Benjamin EJ, Berry JD, Borden WB, Bravata DM, Dai S, Ford ES, Fox CS, Fullerton HJ, Gillespie C, Hailpern SM, Heit JA, Howard VJ, Kissela BM, Kittner SJ, Lackland DT, Lichtman JH, Lisabeth LD, Makuc DM, Marcus

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

GM, Marelli A, Matchar DB, Moy CS, Mozaffarian D, Mussolino ME, Nichol G, Paynter NP, Soliman EZ, Sorlie PD, Sotoodehnia N, Turan TN, Virani SS, Wong ND, Woo D, Turner MB (2012), “Executive summary: heart disease and stroke statistics—2012 update: a report”,Circulation 125(1):188–197,Volume 125 Issue 1,January 2012.

[2] Pahus SH, Hansen AT, Hvas AM (2016),“Thrombophilia testing in young patients with Ischemic stroke”,Thromb Res 137:108–112

[3] DuPont SA, Wijdicks EF, Lanzino G, Rabinstein AA (2010),“Aneurysmal subarachnoid haemorrhage: an overview for the practicing neurologist”, Semin Neurol 30(5):45–54

[4] Santos EMM, Yoo AJ, Beenen LF, Majoie CB, Marquering HA (2016),“Observer variability of absolute and relative thrombus density measurements in patients with acute ischemic stroke”, Neuroradiology 58(2):133–139

[5] Rebouc¸as ES, Marques RCP, Braga AM, Oliveira SAF, de Albuquerque VHC, Filho PPR (2018) ,“New level set approach based on Parzen estimation for stroke segmentation in skull CT images”,Soft Computing,https://doi.org/10.1007/s00500-018-3491-4,Issue 19, 2019.

[6] Shinohara Y, Yanagihara T, Abe K, Yoshimine T, Fujinaka T, Chuma T, Ochi F, Nagayama M, Ogawa A, Suzuki N, Katayama Y, Kimura A, Minematsu K (2011),“Cerebral infarction/transient ischemic attack (TIA)”, J Stroke Cerebrovasc Dis 20(4):S71–S73

[7] 7.B.Mamatha,R.Sreelatha,DrM.Saravanamuth-u,“Brain Stroke detection using machine learning”,IJCRT,volume 10,Issue 8 August 2022.

[8] Dr.Harish B.G and Noorul Huda Khanum “Stroke prediction using machine learning”,IJCRT, volume 10,Issue 9 september 2022.

[9] 9K.Pallavi,B.Saradha and Dr.Saravanan Thirunavakarasu, “Classification of stroke diseases using machine learning algorithms”,IJCRT ,volume 10,Issue 8 August 2022.

[10] 10. Amini L, Azarpazhouh R, Farzadfar MT, Mousavi SA, Jazaieri F, Khorvash F, Norouzi R, Toghianfar N (2013),“Prediction and control of stroke by data mining”,Int J Prev Med 4(2):S245

[11] 11. Bentley P, Ganesalingam J, Jones AL, Mahady K, Epton S, Rinne P, Sharma P, Halse O, Mehta A, Rueckert D (2014),“Prediction of stroke thrombolysis outcome using CT brain machine learning”, NeuroImage Clin 4:635–640

[12] 12. Cheng CA, Lin YC, Chiu HW (2014),“Prediction of the prognosis of ischemic stroke patients after intravenous thrombolysis using artificial neural networks”, Stud Health Technol Inform 202:115–118

[13] 13. Colak C, Karaman E, Turtay MG (2015),“Application of knowledge discovery process on the prediction of stroke”, Comput Methods Programs Biomed 119(3):181–185

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

[14] 14. Maier O, Schro¨der C, Forkert ND, Martinetz T, Handels H (2015),“ Classifiers for ischemic stroke lesion segmentation: a comparisonstudy”,PLoS ONE 10(12):e0145118

[15] Sri Hari Nallamala, et al., “A Literature Survey on Data Mining Approach to Effectively Handle Cancer Treatment”, (IJET) (UAE), ISSN: 2227 – 524X, Vol. 7, No 2.7, SI 7, Page No: 729 – 732, March 2018.

[16] Sri Hari Nallamala, et.al., “An Appraisal on Recurrent Pattern Analysis Algorithm from the Net Monitor Records”, (IJET) (UAE), ISSN: 2227 – 524X, Vol. 7, No 2.7, SI 7, Page No: 542 – 545, March 2018.

[17] Sri Hari Nallamala, et.al, “Qualitative Metrics on Breast Cancer Diagnosis with Neuro Fuzzy Inference Systems”, International Journal of Advanced Trends in Computer Science and Engineering, (IJATCSE), ISSN (ONLINE): 2278 – 3091, Vol. 8 No. 2, Page No: 259 – 264, March / April 2019.

[18] Sri Hari Nallamala, et.al, “Breast Cancer Detection using Machine Learning Way”, International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8, Issue-2S3, Page No: 1402 – 1405, July 2019.

[19] Sri Hari Nallamala, et.al, “Pedagogy and Reduction of K-nn Algorithm for Filtering Samples in the Breast Cancer Treatment”, International Journal of Scientific and Technology Research, (IJSTR), ISSN: 2277-8616, Vol. 8, Issue 11, Page No: 2168 – 2173, November 2019.

[20] Kolla Bhanu Prakash, Sri Hari Nallamala, et al., “Accurate Hand Gesture Recognition using CNN and RNN Approaches” International Journal of Advanced Trends in Computer Science and Engineering, 9(3), May – June 2020, 3216 – 3222.

[21] Sri Hari Nallamala, et al., “A Review on ‘Applications, Early Successes & Challenges of Big Data in Modern Healthcare Management’”, Vol.83, May - June 2020 ISSN: 0193-4120 Page No. 11117 – 11121.

[22] Nallamala, S.H., et al., “A Brief Analysis of Collaborative and Content Based Filtering Algorithms used in Recommender Systems”, IOP Conference Series: Materials Science and Engineering, 2020, 981(2), 022008.

[23] Nallamala, S.H., Mishra, P., Koneru, S.V., “Breast cancer detection using machine learning approaches”, International Journal of Recent Technology and Engineering, 2019, 7(5), pp. 478–481.