An Empirical Study of Machine Learning Algorithms forCancer Identification

Dr.M.Rajaiah,Dean Academics & HOD, Dept of CSE, Audisankara College of Engineering and Technology, Gudur.

Mr.Venkataradhakrishnamurty,Associate Professor ,Dept of CSE, Audisankara College of Engineering and Technology, Gudur.

Ms.K.Lavanya,UG Scholar, Dept of CSE, Audisankara College of Engineering and Technology, Gudur.

Mr.A.Mohansai,UG Scholar, Dept of CSE, Audisankara College of Engineering and Technology, Gudur.

Ms.G.V.Kumari,UG Scholar, Dept of CSE, Audisankara College of Engineering and Technology, Gudur.

Ms.K.Amitha,UG Scholar, Dept of CSE, Audisankara College of Engineering and Technology, Gudur.

ABSTRACT:

A major challenge in the search for a cure for cancer is predicting the illness status of the disease. For instance, distinguishing between benign and malignant tumours helps doctors diagnose cancer more accurately. Although technology advancements produced data on patients with various illness stages, it would be crucial to assess how well machine learning algorithms accomplish predictions. In this article, we suggest employing machine learning algorithms such a variation of AdaBoost, deepboost, xgboost, and support vector machines. We then analyse them using area under curve and accuracyon actual clinical data linked to thyroid cancer, colon cancer, and liver cancer. Results from experiments demonstrate the SVM's strong performance.

*Keywords:*Boosting, Machine Learning, Cancer Genomics,Medical Diagnosis

1.INTRODUCTION:

By 2020, there may be 15 million more cancer cases [1]. A crucial step in the medical diagnosis of cancer patients would be to give healthcare workers efficient tools to spot cancer situations. Researchers from several fields created computational methods to help with medical diagnostics in order to accomplish this objective. For instance, to enhance the pre-surgical diagnoses of thyroid nodules, Stokowy et al. [2]

suggested a computational strategy based on gene expression data to distinguish malignant from benign thyroid nodules. Zhang *et al*. [3] proposed usingextremelearningmachine(ELM)appliedtoseveralmicroarraydataincludinglungdata,lymphomadataandother predictionperformance.Asvariousmachinelearningresearchers have recently proposed computational methods toimprove the performance, there is a need to apply these recentmachine learning algorithms for clinical data generated viavarioustechnologies.

Inthispaper,wepresentusingDeepBoost,whichisanewensemblelearningalgorithm[7]; xgboost, which is scalable end-to-end tree boosting system[8]; a variant of Adaboost [9];and support vectormachines (SVM) [10, 11] . We apply these machine learningalgorithms for cancer identification. Experimental results onreal data pertaining to thyroid cancer, colon cancer, and livercancer show that SVM outperforms the previously mentionedalgorithms.

The rest of the paper is organized as follows. Section II reportsexperimental results of machine learning algorithms on realdatapertainingtothyroidcancer,coloncancerandlivercancer. Section III concludes the paper and points out futurework

2.PROPOSED SYSTEM:

The proposed systeminthispaper,wepresentusingDeepBoost,whichisanewensemblelearningalgorithm xgboost, which is scalable end-to-end tree boosting system,a variant of Adaboostand support vectormachines (SVM).We apply these machine learningalgorithms for cancer identification. Experimental results onreal data pertaining to thyroid cancer, colon cancer, and livercancer show that SVM outperforms the previously mentionedalgorithms.

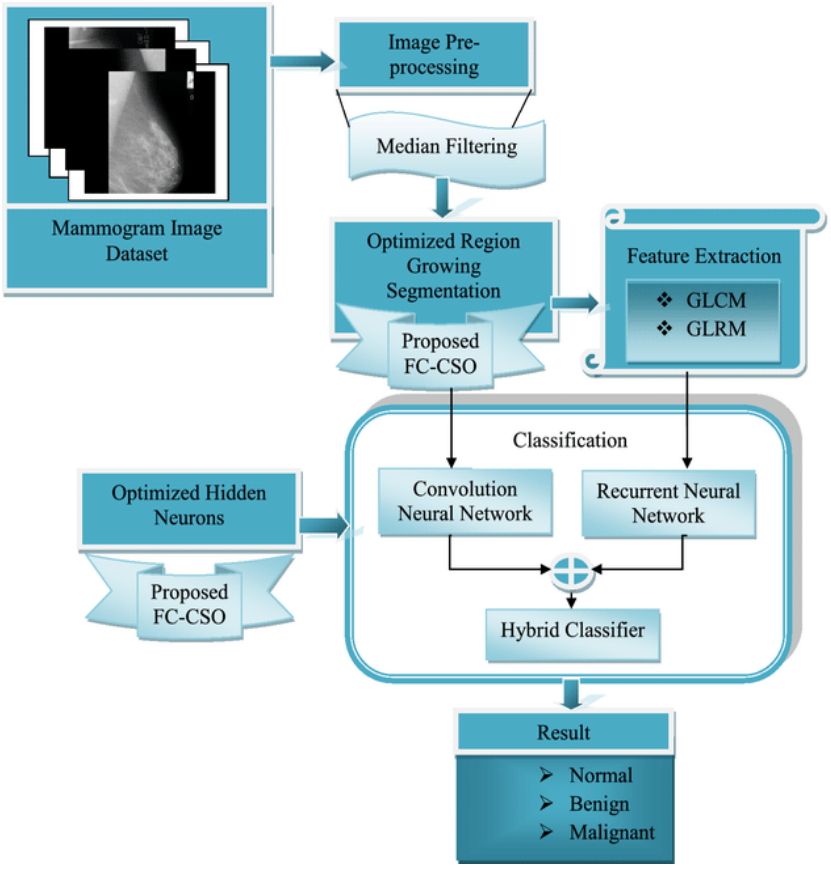
3.LITERARTURE SURVEY:

T. Chen, and C. Guestrin, "Xgboost: A scalable treeboostingsystem,"*Proceedingsofthe22ndacmsigkddinternationalconferenceonknowledgediscoveryanddatamining*.ACM,2016.pp.785-794.T. Turki, and J. T. L. Wang, "Reverse EngineeringGeneRegulatoryNetworksUsingSamplingandBoosting Techniques," *Machine Learning and DataMininginPatternRecognition:13thInternationalConference, MLDM 2017, New York, NY, USA, July15-20, 2017, Proceedings*, P. Perner, ed., pp. 63-77,Cham:Springer InternationalPublishing,2017.C.-C. Chang, and C.-J. Lin, "LIBSVM: A library forsupportvectormachines,"*ACMTransactionsonIntelligentSystemsandTechnology (TIST)*,vol.2,no. 3,pp.27,2011.J. Zhang, S. O. Williams, and H. Wang,

"Intelligentcomputing system based on pattern recognition anddataminingalgorithms,"*SustainableComputing:InformaticsandSystems*,2017.D.Marcous,andY.Sandbank,"deepboost:DeepBoostingEnsembleModeling,"*Rpackageversion0.1.6*.T. Chen, T. He, and M. Benesty, "xgboost: extremegradientboosting(2015),"*Rpackageversion0.3-3*.T. Turki, Z. Wei, and J. T. Wang, "Transfer LearningApproachesto Improve Drug Sensitivity PredictioninMultipleMyelomaPatients,"*IEEE Access*,2017.

Block Diagram:

The below Figure shows the Block diagram of the proposed system which has been



Hardware requirements:

- Adaboost algorithm
- Deepboost algorithm
- Xg boost algorithm

4.CONCLUSION:

For the goal of cancer identification, we suggest employing machine learning algorithms such as DeepBoost, xgboost, a variation of AdaBoost, and SVM. Support vector machines work well, as evidenced by experimental results on actual clinical data relating to thyroid cancer, colon cancer, and liver cancer. Future work will consist of three main objectives: (1) enhancing the performance of machine learning algorithms through the use of transfer learning techniques as in [14], (2) enhancing the performance of the machinelearning used in this study, and (3) combining the machine learning used in this study with feature learning techniques.

FUTURE WORK:

The EASY FARMING SYSTEM USING IOT FOR EFFICIENT PADDY GROWTH adding components make feasible. By adding automation process the results for the crop field management will give better results.

References:

[1] W.H.Organization,"Globalcancerratescouldincreaseby50%to15millionby2020, "*Globalcancer rates could increase by 50% to 15 million by2020*,2003.

[2] T.Stokowy,B.Wojtaś,J.Krajewska,E.Stobiecka,

H. Dralle, T. Musholt, S. Hauptmann, D. Lange, L.Hegedüs, and B. Jarząb, "A two miRNA classifierdifferentiatesfollicularthyroidcarcinomasfromfollicular thyroid adenomas," *Molecular and cellularendocrinology*,vol. 399,pp. 43-49,2015.

[3] R.Zhang,G.-B.Huang,N.Sundararajan,andP.Saratchandran, "Multicategory classification using anextremelearningmachineformicroarraygeneexpressioncancerdiagnosis,"*IEEE/ACM TransactionsonComputationalBiologyandBioinformatics (TCBB),* vol. 4, no. 3, pp. 485-495,2007.

[4] Y.Wang,I.V.Tetko,M.A.Hall,E.Frank,A.Facius,K.F.Mayer,andH.W.Mewes,"Geneselectionfrommicroarraydataforcancerclassification—amachinelearningapproach,"

Computational biology and chemistry, vol. 29, no. 1,pp. 37-46,2005.

[5] Y.Xiao,J.Wu,Z.Lin,andX.Zhao,"Adeeplearning-basedmulti-modelensemblemethodforcancer prediction," *Computer methods and programsin biomedicine,* vol.153,pp.1-9,2018.

[6] Y. Yamamoto, A. Saito, A. Tateishi, H. Shimojo, H.Kanno, S. Tsuchiya, K.-i. Ito, E. Cosatto, H. P. Graf,and R. R. Moraleda, "Quantitative diagnosis of breasttumorsbymorphometricclassificationofmicroenvironmentalmyoepithelialcellsusingamachine learning approach," *Scientific Reports,* vol.7,2017.

[7] C. Cortes, M. Mohri, and U. Syed, "Deep boosting,"*Proceedings of the 31st International Conference onMachine Learning*,2014. pp.1179-1187.

[8] T. Chen, and C. Guestrin, "Xgboost: A scalable treeboostingsystem,"*Proceedingsofthe22ndacmsigkddinternationalconferenceonknowledgediscoveryanddatamining*.ACM,2016.pp.785-794.

[9] T. Turki, and J. T. L. Wang, "Reverse EngineeringGeneRegulatoryNetworksUsingSamplingandBoosting Techniques," *Machine Learning and DataMininginPatternRecognition:13thInternationalConference, MLDM 2017, New York, NY, USA, July15-20, 2017, Proceedings*, P. Perner, ed., pp. 63-77,Cham:Springer InternationalPublishing,2017.

[10] C.-C. Chang, and C.-J. Lin, "LIBSVM: A library forsupportvectormachines,"*ACMTransactionsonIntelligentSystemsandTechnology (TIST)*,vol.2,no. 3,pp.27,2011.

[11] J. Zhang, S. O. Williams, and H. Wang, "Intelligentcomputing system based on pattern recognition anddataminingalgorithms,"*SustainableComputing:InformaticsandSystems*,2017.

[12] D.Marcous,andY.Sandbank,"deepboost:DeepBoostingEnsembleModeling,"*Rpackageversion0.1.6*.

[13] T. Chen, T. He, and M. Benesty, "xgboost: extremegradientboosting(2015),"*Rpackageversion0.3-3*.

[14] T. Turki, Z. Wei, and J. T. Wang, "Transfer LearningApproachesto Improve Drug Sensitivity PredictioninMultipleMyelomaPatients,"

Author Profiles:

Dr.M.Rajaiah,Currently working as an Dean Academics & HOD in the department of CSE at ASCET (Autonomous), Gudur, Tirupathi(DT).He has published more than 35 papers in Web of Science,Scopus,UGC Journals.

Mr .Venkataradhakrishnamurty completed his Master of Technology in Computer Science and Engineering. Pursing Ph.D in Saveetha School Of Engineering. Currently working as an Associate professor in the department of CSE at ASCET (Autonomous), Gudur, Tirupati(DT).

Ms.K.Lavanya,B.Tech student in the department of CSE at Audisankara College of Engineering and Technology, Gudur. She has pursuing in computer science and engineering.

Mr.AMohansai,B.Tech student in the department of CSE at Audisankara College of Engineering and Technology, Gudur. He has pursuing in computer science and engineering.

Ms.G.V. Kumari,B.Tech student in the department of CSE at Audisankara College of Engineering and Technology, Gudur. She has pursuing in computer science and engineering.

Ms.K.Amitha Reema Rathan,B.Tech student in the department of CSE at Audisankara College of Engineering and Technology, Gudur. She has pursuing in computer science and engineering.