Earlier Detection of Breast Cancer Using Gaussian Mixture Algorithm

Dr.D.J .Samatha Naidu[1], Komerla Fabiya Sulthana[2]

[1, 2] Dept of Master of Computer Application,
[1, 2] Annamacharya PG College Of Computer Studies,Rajampet, Annamayya Dist,AP,516126.

Abstract- *Breast cancer is type of tumor that occurs in the tissues of the breast. It is most common type of cancer found in women around the world and it is among the leading causes of deaths in women. This paper presents the comparative analysis of machine learning, deep learning and data mining techniques being used for the prediction of breast cancer. Many researchers have put their efforts on breast cancer.*

I. INTRODUCTION

A dataset of 938 images has been taken for the classification of skin diseases. They include melanoma, nevus, sebborheickeratosis. By using CNN Algorithms, 70% accuracy is achieved in classification of skin disese. We have also tried with alexnet, which gives 80% accuracy.

The usage of deep learning algorithms has reduced the need for human labor, such as manual feature extraction and data reconstruction for classification purpose

PURPOSE

Cancer is the common problem for all people in the world with all types. Particularly, Breast Cancer is the most frequent disease as a cancer type for women. Therefore, any development for diagnosis and prediction of cancer disease is capital important for a healthy life. Machine learning techniques can make a huge contribute on the process of early diagnosis and prediction of cancer.

SCOPE

In this paper author is using SVM (Support Vector Machine) and ANN (Artificial Neural Network) to predict breast cancer diseases. First this algorithms will be trained using past disease dataset called 'Wisconsin Breast Cancer', this dataset contains 11 integer values and last value contains class label either 0 or 1, 0 means person is normal and 1 means person is infected with disease

II. EXISTING WORK

We have applied SVM and ANN techniques for prediction of the classification of breast cancer to find which machine learning methods performance is better. Support Vector Machines (SVMs) have been first explained by Vladimir Vapnik and the good performances of SVMs have been noticed in many pattern recognition problems. SVMs can indicate better classification performance when it is compared with many other classification techniques . SVM is one of the most popular machine learning classification technique that is used for the prognosis and diagnosis of cancer.

DIS-ADVANTAGES

1. As the diagnosis of this disease manually takes long hours and the lesser availability of systems
2. Statistical methods for medical information are too Heavy
3. This model makes use of all the features without any restrictions while selecting them.
4. Prediction accuracy is less.

Proposed System

Machine learning involves predicting and classifying data and to do so we employ various machine learning algorithms according to the dataset. SVM or Support Vector Machine is a linear model for classification and regression problems. It can solve linear and non-linear problems and work well for many practical problems. The idea of SVM is simple: The algorithm creates a line or a hyperplane which separates the data into classes. In machine learning, the radial basis function kernel, or RBF kernel, is a popular kernel function used in various kernelized learning algorithms.

ADVANTAGES

This is feature learning In terms of performance, CNNs outperform NNs on conventional image recognition tasks and many other tasks.

RELATED WORK

1.Cancer sound analysis for diagnostic information
Authors: Rutuja Mhetre and U.R.Bagal

Abstract: The most important concern in the medical domain is to consider the interpretation of data and perform accurate diagnosis.The bronchitis, pneumonia and many other pulmonary diseases causes Cancer disorders which affects Cancer systems. Diagnosis of these diseases is facilitated by pulmonary auscultation by using stethoscope. This method depends on individuals hearing capability, experience and ability to differentiate the sounds The signal processing techniques may be used for diagnostic information.

2.Automatic Wheezing Detection Based on Signal Processing of Spectrogram and Back - Propagation Neural Network.
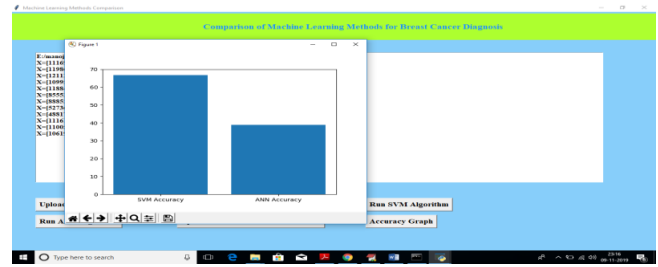Authors: Bor-Shing Lin, Huey-Dong Wu and Sao-Jie Chen.

Sample Screens

Now click on 'Upload Wisconsin Dataset' button to upload dataset

In above screen we can see Accuracy value, precision and recall obtained from SVM. SVM got 66.66% accuracy. Now click on 'Run ANN Algorithm' to get ANN Accuracy.

In above graph x-axis represents algorithm name and y-axis represents accuracy of that algorithm. From above graph we can conclude that SVM is better than ANN

III. CONCLUSION

Breast Cancer is the most frequent disease as a cancer type for women. Therefore, any development for diagnosis and prediction of cancer disease is capital important for a healthy life. In this paper, we have discussed two popular machine learning techniques for Wisconsin Breast Cancer classification.

REFERENCES

[1] K. Graves, Ceh: Official certified ethical hacker review guide: Exam 312-50. John Wiley & Sons, 2007.

[2] R. Christopher, "Port scanning techniques and the defense against them," SANS Institute, 2001.

[3] M. Baykara, R. Das¸ and I. Karado ˘gan, "Bilgi g ¨uvenli ˘gi sistemlerinde kullanilan arac ¸larin incelenmesi," in 1st International Symposium on Digital Forensics and Security (ISDFS13), 2013, pp. 231–239.

[4] RashmiT V. "Predicting the System Failures Using Machine Learning Algorithms".International Journal of Advanced Scientific Innovation, vol. 1, no. 1, Dec. 2020, doi:10.5281/zenodo.4641686.

[5] S. Robertson, E. V. Siegel, M. Miller, and S. J. Stolfo, "Surveillance detection in high bandwidth environments," in DARPA Information Survivability Conference and Exposition, 2003. Proceedings, vol. 1. IEEE, 2003, pp. 130–138.

[6] K. Ibrahimi and M. Ouaddane, "Management of intrusion detection systems based-kdd99: Analysis with lda and pca," in Wireless Networks and Mobile Communications (WINCOM), 2017 International Conference on. IEEE, 2017, pp. 1–6.

[7] Girish L, Rao SKN (2020) "Quantifying sensitivity and performance degradation of virtual machines using machine learning.",Journal of Computational and Theoretical Nanoscience, Volume 17, Numbers 9-10,

September/October 2020, pp.4055-4060(6) https://doi.org/10.1166/jctn.2020.9019

[8] L. Sun, T. Anthony, H. Z. Xia, J. Chen, X. Huang, and Y. Zhang, "Detection and classification of malicious patterns in network traffic using benford's law," in Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2017. IEEE, 2017, pp. 864–872.

[9] S. M. Almansob and S. S. Lomte, "Addressing challenges for intrusion detection system using naive bayes and pca algorithm," in Convergence in Technology (I2CT), 2017 2nd International Conference for. IEEE, 2017, pp. 565–568