1st Conference on Healthcare Computing Systems and Technologies (CHEST 2019)

A New Machine Learning Technique for Breast Cancer Detection

1st Moein Hajiabadi
Department of Computer Engineering
University of Birjand
Birjand, Iran
m.hajiabadi@birjandut.ac.ir

2nd Hamideh Hajiabadi
Department of Computer Engineering
Birjand University of Technology
Birjand, Iran
hajiabadi@birjandut.ac.ir

Abstract— **Cancer, also called malignancy, is an abnormal growth of cells. it is characterized as a heterogeneous disease containing many various subtypes. The early diagnosis of cancer has a significant effect on the chance that the patient being cured. Automatically classifying cancer disease into two categories including high risk and low risk has been attracted many researchers in the fields of biomedical and bioinformatics. Recently, Machine Learning (ML) methods have been widely utilized in disease diagnosis. A variety of these techniques, including Articial Neural Networks (ANNs), Naive Bayes (NB), Support Vector Machines (SVMs), etc. have been widely applied in cancer detection and have achieved good results. In this paper, we review the most recent ML techniques in breast cancer research area and propose a new classifier based on a new loss function to automatically detect breast cancer and classify it into two categories. We investigate the effectiveness of our method through comprehensive experiments on several benchmark datasets. The experiments show an improvement compared with the state-of-the-art methods.**

Keywords-component; formatting; style; styling; insert (key words)

[1] Liangjun Chen, Hua Qu, and Jihong Zhao. Generalized correntropy based deep learning in presence of non-gaussian noises. Neurocomputing, 2017.

[2] Joann G Elmore, Carolyn K Wells, Carol H Lee, Debra H Howard, and Alvan R Feinstein. Variability in radiologists' interpretations of mammograms. New England Journal of Medicine, 331(22):1493–1499, 1994.

[3] Marc G Genton. Highly robust variogram estimation. Mathematical Geology, 30(2):213–221, 1998.

[4] Hamideh Hajiabadi, Diego Molla-Aliod, and Reza Monsefi. On extending neural networks with loss ensembles for text classification. In Proceedings of the Australasian Language Technology Association Workshop 2017, 2017.

[5] Hamideh Hajiabadi, Reza Monsefi, and Hadi Sadoghi Yazdi. Relf: Robust regression extended with ensemble loss function. arXiv preprint arXiv:1810.11071, 2018.

[6] Douglas Hanahan and Robert A Weinberg. Hallmarks of cancer: the next generation. cell, 144(5):646–674, 2011.

[7] Amruta Hebli and Sudha Gupta. Brain tumor prediction and classification using support vector machine. In Advances in Computing, Communication and Control (ICAC3), 2017 International Conference on, pages 1–6. IEEE, 2017.

[8] Matthew J Holland and Kazushi Ikeda. Minimum proper loss estimators for parametric models. IEEE Transactions on Signal Processing, 64(3):704–713, 2016.

[9] W. Liu, P. P. Pokharel, and J. C. Principe. Correntropy: Properties and applications in non-Gaussian signal processing. IEEE Transactions on Signal Processing, 55(11):5286–5298, Nov 2007.

[10] Weifeng Liu, P. P. Pokharel, and J. C. Principe. Correntropy: A localized similarity measure. In The 2006 IEEE International Joint Conference on Neural Network Proceedings, pages 4919–4924, 2006.

[11] Weifeng Liu, Puskal P Pokharel, and Jose C Principe. Correntropy: A localized similarity measure. In Neural Networks, 2006. IJCNN'06. International Joint Conference on, pages 4919–4924. IEEE, 2006.

[12] Hamed Masnadi-Shirazi and Nuno Vasconcelos. On the design of loss functions for classification: theory, robustness to outliers, and SavageBoost. In Advances in neural information processing systems, pages 1049–1056, 2009.

[13] Robert Moore and John DeNero. L1 and l2 regularization for multiclass hinge loss models. In Symposium on Machine Learning in Speech and Language Processing, 2011.

[14] Yamuna Prasad, K Kanad Biswas, and Chakresh Kumar Jain. Svm classifier based feature selection using ga, aco and pso for sirna design. In International conference in swarm intelligence, pages 307–314. Springer, 2010.

[15] Rebecca Siegel, Deepa Naishadham, and Ahmedin Jemal. Cancer statistics for hispanics/latinos, 2012. CA: a cancer journal for clinicians, 62(5):283–298, 2012.

[16] Ingo Steinwart. Support vector machines are universally consistent. Journal of Complexity, 18(3):768–791, 2002.

[17] Long Tang, Yingjie Tian, Chunyan Yang, and Panos M Pardalos. Ramploss nonparallel support vector regression: robust, sparse and scalable approximation. Knowledge-Based Systems, 2018.

[18] Vladimir Vapnik. Statistical learning theory. 1998. Wiley, New York, 1998.

[19] Yingchao Xiao, Huangang Wang, and Wenli Xu. Ramp loss based robust one-class svm. Pattern Recognition Letters, 85:15–20, 2017.

[20] Kaobi Yan, Zhixin Li, and Canlong Zhang. A new multi-instance multilabel learning approach for image and text classification. Multimedia Tools and Applications, 75(13):7875–7890, 2016.

[21] Bichen Zheng, Sang Won Yoon, and Sarah S Lam. Breast cancer diagnosis based on feature extraction using a hybrid of k-means and support vector machine algorithms. Expert Systems with Applications, 41(4):1476–1482, 2014.