Abstract

The traditional data analytic might not have the capacity to handle enormous amount of data. Due to the rapid growth of information, solutions need to be contemplated and provided in order to handle and extract value and knowledge from these data sets. Moreover, decision makers should have the capacity to increase significant bits of knowledge from such fluctuated and quickly evolving information. Such esteem can be given utilizing big data analytic, which is the utilization of advanced analytic techniques on big data using MapReduce approach. This paper examines to develop a high performance platform to efficiently analyse big SEER (Surveillance, Epidemiology, and End Results) breast cancer data set using MapReduce to find the recurrence of breast cancer.

References

21-29.
Intrusion Prediction with Machine Learning. ACM SIGMETRICS Performance Evaluation
Commun ACM 51(1):107–113
machine learning on multicore. Adv neural Info processing systems 19:281
45(6):983–1006
Classification for Big Data Analysis Using Naïve Bayes Classifier. IEEE International
Conference on Big Data, pp. 99-104.
of MapReduce Program Model for Parallel Support Vector Machine. International Journal of
15. Han, J., Liu, Y., and Sun, X., 2013. A Scalable Random Forest Algorithm Based on
MapReduce. 4th IEEE International Conference on Software Engineering and Service Science,
pp. 849-852.
16. Santi Wulan Purnami, S.P. Rahayu and Abdullah Embong, “Feature selection and
17. Farzaneh Keivanfard , Mohammad Teshnehab , Mahdi Aliyari Shoorehdeli , “Feature
Selection and Classification of Breast Cancer on Dynamic Magnetic Resonance Imaging by
Using Artificial Neural Networks”, Proceedings of the 17th Iranian Conference of Biomedical
Engineering (ICBME2010), 3-4 November 2010.
Breast Cancer diagnosis”, Proceedings of the 9th International Conference on Information
Technology and Applications in Biomedicine, ITAB 2009, Larnaca, Cyprus, 5-7 November 2009.
cancer survivability on imbalanced data ”, IEEE 2009.
20. Ankit Agrawal, Sanchit Misra, Ramanathan Narayanan, Lalith Polepeddi, Alok
three data mining methods,” Artificial Intelligence in Medicine, vol. 34, pp. 113-127, 2005
22. A. Bellachia and E. Guvan, “Predicting breast cancer survivability using data mining techniques”, Scientific Data Mining Workshop, in conjunction with the 2006 SIAM Conference on Data Mining, 2006


27. BIOGRAPHY


29. Dr. B. Ramachandra working as Professor and Head in Department of Electrical & Electronics, PES College of Engineering Mandya. He had his Ph.D. From Indian Institute of Science, Bangalore, Master’s from Indian Institute of Technology, Bombay.

**Index Terms**

Computer Science Biomedical

**Keywords**

Breast cancer; Big data, Classification; Data analytics, MapReduce, SEER.