Related Work Research Paper Summary

List of Research Papers is as follows:

1) Paper Title:
"Big Data Analytics to Predict Breast Cancer Recurrence on SEER Dataset using MapReduce Approach"
Paper Authors:
   1. Umesh D. R., Assistant Professor Department of Computer Science & Engineering PESCE, Mandya, Karnataka, India
   2. B. Ramachandra, PhD Professor Department of Electrical & Electronics Engineering PESCE, Mandya, Karnataka, India

Link: https://pdfs.semanticscholar.org/890d/b7e3b41add9e4824f13476de4dba0ae111da.pdf

2) Paper Title:
"Machine Learning Classification Techniques for Breast Cancer Diagnosis"
Paper Authors:
   1. David A. Omondiagbe, Curtin University, Malaysia, CDT 250, Miri 98009, Sarawak, Malaysia
   2. Shanmugam Veeramani , *Corresponding author: s.veeramani@curtin.edu.my
   3. Amandeep S. Sidhu , 2- Curtin University, Kent St, Bentley WA 6102, Australia

Link: https://iopscience.iop.org/article/10.1088/1757-899X/495/1/012033/pdf

3) Paper Title:
"Prediction of Breast Cancer Using Big Data Analytics"
Paper Authors:
   1. K. Shailaja, M. Tech Scholar, Centre for Data Science, Vardhaman College of Engineering, Hyderabad, Telangana
   2. B. Seetharamulu, Professor, Department of CSE, Vardhaman College of Engineering, Hyderabad, Telangana
   3. M.A. Jabbar, Professor, Centre for Data Science, Vardhaman College of Engineering, Hyderabad, Telangana

Link: https://www.researchgate.net/publication/334456222_Prediction_of_Breast_Cancer_Using_Big_Data_Analytics

4) Paper Title:
"On the Scalability of Machine-Learning Algorithms for Breast Cancer Prediction in Big Data Context"
Paper Authors:
   1. Sara Alghunaim National Center for Artificial Intelligence and Big Data Technology, King Abdulaziz City for Science and Technology, Riyadh 11442, Saudi Arabia
   2. Heyam H. Al-Baity, 2 IT Department, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Link: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8755975

5) Paper Title:
"Survey of Breast Cancer Detection Using Machine Learning Techniques in Big Data"
Paper Authors:
   1. Madhuri Gupta, Jaypeee Institute of Information Technology, Noida, India
   2. Bharat Gupta, Jaypee Institute of Information Technology, Noida, India

Link: https://www.igi-global.com/article/survey-of-breast-cancer-detection-using-machine-learning-techniques-in-big-data/227680

Paper abstract:

Cancer is a disease in which cells in body grow and divide beyond the control. Breast cancer is the second most common disease after lung cancer in women. Incredible advances in health sciences and biotechnology have prompted a huge amount of gene expression and clinical data. Machine learning techniques are improving the prior detection of breast cancer from this data. The research work carried out focuses on the application of machine learning methods, data analytic techniques, tools, and frameworks in the field of breast cancer research with respect to cancer survivability, cancer recurrence, cancer prediction and detection. Some of the widely used machine learning techniques used for detection of breast cancer are support vector machine and artificial neural network. Apache Spark data processing engine is found to be compatible with most of the machine learning frameworks.

Summary:
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. In this paper, we address the problem of breast cancer prediction in the big data context. The research work carried out focuses on the application of machine learning methods, data analytic techniques, tools, and frameworks in the field of breast cancer research with respect to cancer survivability, cancer recurrence, cancer prediction and detection. Some of the widely used machine learning techniques used for detection of breast cancer are support vector machine and artificial neural network. The proposed approach obtained an accuracy of 98.82%, sensitivity of 98.41%, specificity of 99.07% and area under the receiver operating characteristic curve of 0.9994.
This paper also proposed a hybrid approach for Breast Cancer diagnosis by reducing the high dimensionality of features using linear discriminant analysis (LDA), and then applying the new reduced feature dataset to Support Vector Machine (SVM). This would help in establishing a standardized approach for performing big data analytics and generating some valuable information which could be shared for various medical predictions. 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 analytics, which is the utilization of advanced analytic techniques on big data using the MapReduce approach. The main usage of referring these papers is to understand how the analysis of Breast Cancer diagnosis takes place in order to build out significant information to aid in detecting Breast Cancer at an early stage. The focus of this reference is to integrate these machine learning techniques with feature selection/feature extraction methods and compare their performances to identify the most suitable approach to perform the analytics.