Characterization of Elevated Tumor Markers in Diagnosis of HCC Using Data Mining Methods

6

Vyshali J. Gogi

 https://orcid.org/0000-0002-2814-567X

Rashtreeya Vidyalaya College of Engineering, India

Vijayalakshmi M. N.

Rashtreeya Vidyalaya College of Engineering, India

INTRODUCTION

The process of identifying patterns in huge datasets comprising methods such as machine learning, statistics, and database system can be considered data mining. It is a multidisciplinary field in computer science, and it excerpts knowledge from the massive data set and converts into comprehensible format. The Medical environment is rich in information but weak in knowledge. Medical systems contain wealth of data which require a dominant analysis tool for determining concealed association and drift in data. The health care condition that comprehends to liver disorder is termed as Liver disease. Liver disorder leads to abrupt health status that precisely governs the working of liver and intern affecting other organs in the body. Data mining classification techniques like Decision Tree, Linear Discriminant, and SVM Fine Gaussian and Logistic Regression algorithms are applied.

BACKGROUND

Hepatocellular Carcinoma (HCC) is an abnormal condition of human liver which is diagnosed at a very advanced stage. The disease is liver disorder which can be predicted after series of clinical and laboratory and imaging studies. Hepatocellular carcinoma is the most malignant tumors which is the major cause of death and requires to be treated at early stage. Benign tumors do not lead to cancer but malignant tumors lead to cancer. The lab investigation accords to austerity of the disease in liver together with physiology of the condition. The laboratory studies incorporate persistence of electrolytes, Complete Blood Count, Liver Function Tests, International Normalized Ratio (INR), partial thromboplastic time (PTT) and AFP (Alpha-Feto Protein) tests. Alpha-fetoprotein (AFP) is considered as one of the tumor markers of HCC. AFP is produced by tumors or by regenerating hepatocytes, but not all tumors produce AFP. The elevations in AFP level greater than 400ng /DL in proper clinical context and appropriate radiologic findings marks the diagnosis of HCC. Increased AFP along with proper imaging studies is considered diagnostic of HCC. The medical imaging techniques like USG (Ultrasonography), CT (Compute Tomography) scan, and MRI (Magnetic Resonance Imaging) creates an interior visualization of the body. Surgical resection or ablation cannot be planned with the anatomic detail obtained from ultrasonography. The lesions are not detected significantly. Although CT scan is highly accurate compared to ultrasonography,

it may miss some of small lesions. MRI is an excellent method to characterize HCC. HCC demonstrates various features which is clearly differentiated in MRI scan. Patients with lesions larger than 1 to 2 cm are recommended for biopsy since they are predicted to be under risk of malignancy. The decision on resection, transplantation and ablative therapy is taken based on the biopsy result.

According to the literature survey advanced hepatocellular carcinoma was treated with Sorafenib molecular target drug which increased the survival time of the patient with 3 to 4 months (Gomaa W., 2017). Presence of Hepatitis C virus and non-alcoholic fatty liver also contributed to liver disorder. Researchers developed a framework considering the imaging studies and found that imaging studies alone was not sufficient for efficient diagnosis of the disease. Artificial Immune Systems and Artificial Neural Network algorithms were applied on laboratory parameters. Artificial Neural Network gave the better result (Shane Dixon & Xiao-Hua Yu, 2016). Data mining algorithms support vector machine and back propagation algorithm was applied on the liver function test attributes. Support vector machine gave best results. Micro array analysis was used as molecular biology approach where the result was affected by diet, age, and ethnicity (Sumedh Sontakke, 2017). Clinical decision support systems support physicians to carry out meticulous prediction of the disease. The Volume of the dataset, and attributes present in the dataset along with the competence and deficiency of the classifiers contributes to deal with medical issues (Ahmed J. Aljaaf, 2015). A cancer prediction system provides warning to the patients thereby reducing the cost and saving time. Validation of the system is done by comparing the predicted results with the previously available medical record of the patient (K.Arutchelvan, 2015). The study of hepatitis dataset and its contribution to Hepatocellular carcinoma. Machine learning algorithms along with feature selection algorithms were applied on the hepatitis dataset. K-fold cross validation was implemented where diverse number of fold were applied and performance was analyzed (Tri Astuti, 2015). The use of logistic regression and artificial neural network models for biomedical data classification was presented by the researchers. The study gave the conclusion that no single algorithm can outperforms all other algorithms. The result depends on the kind of data and cases (Stephan Dreiseitla, 2003). Eleven data mining classification algorithms were applied to the datasets and the performance of all classifiers are compared against each other in terms of accuracy, precision, and recall. Several investigations have also been carried out to improve performance of the classification models. Finally, the results shown promising methodology in diagnosing liver disease during the earlier stages (Sina Bahramirad, 2013). Abnormal LFT result not only mean that the problem is pertaining to liver always, it may also indicate abnormalities in other organs. To overcome this, researchers came out with a prediction model called Algorithm for Liver Function Investigations (ALFI), which gave a decision support tool incorporating a clinical prediction rule that facilitated for the management of patients in primary care. The C-Statistic technique was used for the study (David J. McLernon, 2012). A framework was proposed based on ultrasound images for classifying liver fibrosis. Deep features were learnt via transfer learning and Visual Geometry Group neural network (VGGNet). Fully connected network (FCNet) classifier was trained in order to predict the normal, early stage and late stage fibrosis liver status (Dan Meng L. Z., 2017). Each stage of liver cancer is diagnosed depending on extent to which liver is damaged. To determine the degree of liver damage at four different stages, a KNN based intelligent computational method was presented. The accuracy of the classification was compared with linear discriminant analysis, diagonal linear discriminant analysis, quadratic discriminant analysis, diagonal quadratic discriminant analysis classifiers and classification and regression trees (Singh & Pandey, 2016).

Based on the literature survey, the authors have used both statistical and computational techniques for diagnosis of Hepatocellular Carcinoma. The authors in the computational methods used laboratory parameters and imaging studies, and the authors in statistical techniques considered tumor markers.

Related Content

Manufacturing vs. Services and the Role of Information Technology
Arnab Adhikariand Shromona Ganguly (2018). *Encyclopedia of Information Science and Technology, Fourth Edition (pp. 7234-7247).*
www.irma-international.org/chapter/manufacturing-vs-services-and-the-role-of-information-technology/184420

Research and Implementation of Pedestrian Attribute Recognition Algorithm Based on Deep Learning
Weilan Fang, Zhengqing LU, ChaoWei Wang, Zhihong Zhou, Guoliang Shiand Ying Yin (2024). *International Journal of Information Technologies and Systems Approach (pp. 1-18).*
www.irma-international.org/article/research-and-implementation-of-pedestrian-attribute-recognition-algorithm-based-on-deep-learning/344019

Research on Power Load Forecasting Using Deep Neural Network and Wavelet Transform
Xiangyu Tan, Gang Ao, Guochao Qian, Fangrong Zhou, Wenyun Liand Chuanbin Liu (2023). *International Journal of Information Technologies and Systems Approach (pp. 1-13).*
www.irma-international.org/article/research-on-power-load-forecasting-using-deep-neural-network-and-wavelet-transform/322411

E-Commerce in Logistics and Supply Chain Management
Yasanur Kayikci (2018). *Encyclopedia of Information Science and Technology, Fourth Edition (pp. 5367-5377).*
www.irma-international.org/chapter/e-commerce-in-logistics-and-supply-chain-management/184240

Time Series Invariants under Boundedness: Information Symbols and Early Warning Signs
 (2013). *Boundedness and Self-Organized Semantics: Theory and Applications (pp. 29-47).*
www.irma-international.org/chapter/time-series-invariants-under-boundedness/70272