

Disease Prediction by Machine Learning over Big Data from Healthcare Communities

ABSTRACT

With big data growth in biomedical and healthcare communities, accurate analysis of medical data benefits early disease detection, patient care and community services. In this paper, we streamline machine learning algorithms for effective prediction of chronic disease outbreak in disease-frequent communities.

EXISTING SYSTEM

With the growth in medical data ,collecting electronic health records (EHR) is increasingly convenient . Besides, first presented a bio inspired high-performance heterogeneous vehicular telematics paradigm, such that the collection of mobile users' health related real-time big data can be achieved with the deployment of advanced heterogeneous vehicular networks. Chen et.al proposed a healthcare system using smart clothing for sustainable health monitoring. Qiu et had thoroughly studied the heterogeneous systems and achieved the best results for cost minimization on tree and simple path cases for heterogeneous systems. Patients' statistical information, test results and disease history are recorded in the EHR, enabling us to identify potential data-centric solutions to reduce the costs of medical case studies. Qiu et al. proposed an efficient flow estimating algorithm for the tele health cloud system and designed a data coherence protocol for the PHR(Personal Health Record)-based distributed system. Bates et al. proposed six applications of big data in the field of healthcare. Qiu et al. proposed an optimal big data sharing algorithm to handle the complicate data set in tele health with cloud techniques. One of the applications is to identify high-risk patients which can be utilized to reduce medical cost since high-risk patients often require expensive healthcare. Moreover, in the first paper proposing healthcare cyber-physical system it innovatively brought forward the concept of prediction-based healthcare applications, including health risk assessment.

DRAWBACKS

- It includes health risk assessment.
- It maybe not satisfy the changes in the disease and its influencing factors.

PROPOSED SYSTEM

In this paper, we streamline machine learning algorithms for effective prediction of chronic disease outbreak in disease-frequent communities. We experiment the modified prediction models over real-life hospital data collected from central China in 2013-2015. To overcome the difficulty of incomplete data, we use a latent factor model to reconstruct the missing data. We experiment on a regional chronic disease of cerebral infarction. We propose a new convolutional neural network based multimodal disease risk prediction (CNN-MDRP) algorithm using structured and unstructured data from hospital.

ADVANTAGES

- It reduce the costs of medical case studies.
- It can improve the accuracy of risk

SYSTEM REQUIREMENTS

H/W System Configuration:-

- Processor - Pentium –IV
- RAM - 4 GB (min)
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

S/W System Configuration:-

- Operating System : Linux
- Application Server : Tomcat5.0/6.X
- Backend coding : Java
- Tool : Virtual Box
- Environment : Ubuntu
- Technology : Hadoop