Chapter 6 Chronic Kidney Disease Prediction Using Data Mining Algorithms

Devesh Kumar Srivastava Manipal University Jaipur, India

Pradeep Kumar Tiwari https://orcid.org/0000-0003-0387-9236 Manipal University Jaipur, India

ABSTRACT

In today's contemporary world, it is important to know about the odds of having a disease because of changing living standards of the population overall in the continent. The disease on which the authors are working is chronic kidney disease. Once the person gets chronic kidney disease (CKD), his working capability decreases along with other adverse effects. It is possible to get rid of diseases like CKD with new methodologies that will help us to predict the stage of kidney disease at an early stage. Under big data analytics, data may be structured, unstructured, quasi- or semi-structured. The CKD detected and predicted by applying classification models: support vector machine (SVM), K-nearest neighbor (KNN), and logistic regression algorithm. It helps in predicting the likelihood of occurrence of disease on various different features. The two algorithms KNN and SVM are compared to find the algorithm that gives better accuracy. Further regression technique has been used to detect the disease based on, which the stages are classified by using GFR (glomerular filtration rate) formula.

INTRODUCTION

India is vast country with the population of more than 133 crores out of which there are more than 10 million cases are found approximately of chronic diseases per year as per survey result. Diagnosis of such chronic diseases like cardiac failure, kidney disease, HIV, diabetes mellitus etc. is important at an early stage as it may be controlled to a certain stage but cannot be cured. Some of the most common

DOI: 10.4018/978-1-7998-2742-9.ch006

Chronic Kidney Disease Prediction Using Data Mining Algorithms

diseases are: - Alzheimer's, Arthritis, Asthma, Cancer, Diabetes, Heart diseases, Chronic Kidney (Suprarenal Gland) disease. There is an increase in death rate in developing countries. As digitalization is growing widely from cities to remote villages, it has become important to use these new technologies for prediction, detection and prevention of such diseases (Ameta & Jain 2017). People putting up in small villages may not get access to proper medical facilities. 60% of all deaths worldwide are caused due to chronic kidney disease. Data mining is the process to extract meaningful information from the large amount of data sets in order to attain knowledge. In this work data mining algorithms are used for analyzing the different stages of chronic suprarenal gland disease. Before applying the algorithms, it is checked the feasibility by two-step process: Learning Step (Training Phase) and Construction of Classification Model. Various Algorithms are utilized to assemble a classifier by making the Model to get the hang of utilizing the preparation set accessible. Model is used to foresee class names and testing the developed model on test information and thus gauge the exactness of the arrangement rules. It has been developed to foresee class marks (Example: Label - "ckd" or "Not ckd" for the endorsement of some occasion). Utilizing mining based strategies are savvy and effective. They help foreseeing illness, Helps Banks and Financial Institutions to recognize defaulters with the goal that they may endorse Cards, Loan, and so forth (Dubey 2015)..

Chronic Kidney Disease (ckd) is referred to Chronic Suprarenal Gland Failure. When this disease occurs in human's body what happens is that kidneys get damaged and in turn toxics cannot be filtered out easily from our body. To classify this disease's likelihood, KNN and SVM data mining algorithms are implemented on the dataset. Important features are extracted which is required for the classification and to execute the work (Devi, 2014; Dubey 2015).

Text mining is the route toward investigating and dissecting a lot of unstructured information supported by programming that can distinguish ideas, designs, points, watchwords and different properties in the information. It is the way toward extricating the concealed learning from the content report. Content mining is characterized as learning serious procedure in which a client cooperates with the report gathering utilizing a suite of examination apparatus. It manages changing over unstructured information into organized information. Content mining fuses and incorporates the apparatuses of data recovery, information mining, AI, measurements, and computational etymology, and thus, it is out and out a multidisciplinary field. Content mining manages characteristic language messages either put away in semi-organized or unstructured organizations.5 fundamental steps involved in the text mining are:

- 1. Gathering unstructured data from numerous information sources like plain content, pages, pdf documents, messages, and online journals, to give some examples.
- 2. Detect and expel inconsistencies from information by directing pre-preparing and purifying activities. Information purifying enables you to separate and hold the significant data covered up inside the information and to help recognize the foundations of explicit words.
- 3. Convert all the important data separated from unstructured information into organized arrangements.
- 4. Analyze the examples inside the information through Management Information System (MIS).
- 5. Store all the significant data into a safe database to drive pattern investigation and upgrade the basic leadership procedure of the association (Arasu, & Thirumalaiselvi 2017; Bala, & Kumar (2014).

Content or Text mining is a subfield of NLP (Natural Language Processing) dedicated to empowering computational investigation of content bolted information. The content mining work process by and large includes distinguishing proof of explicit elements in surface content, for example, illnesses, quali18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/chronic-kidney-disease-prediction-using-datamining-algorithms/263316

Related Content

Reinforcement Learning: A Deep Dive Into Techniques and Future Prospects

Arti Saxenaand Falak Bhardwaj (2024). *Methodologies, Frameworks, and Applications of Machine Learning* (pp. 1-18).

www.irma-international.org/chapter/reinforcement-learning/342646

Development of a Charge Estimator for Piezoelectric Actuators: A Radial Basis Function Approach

Morteza Mohammadzaheri, Mohammadreza Emadi, Mojtaba Ghodsi, Issam M. Bahadur, Musaab Zarogand Ashraf Saleem (2020). *International Journal of Artificial Intelligence and Machine Learning (pp. 31-44).*

www.irma-international.org/article/development-of-a-charge-estimator-for-piezoelectric-actuators/249251

Three-Layer Stacked Generalization Architecture With Simulated Annealing for Optimum Results in Data Mining

K. T. Sanvitha Kasthuriarachchiand Sidath R. Liyanage (2021). *International Journal of Artificial Intelligence* and Machine Learning (pp. 1-27).

www.irma-international.org/article/three-layer-stacked-generalization-architecture-with-simulated-annealing-for-optimumresults-in-data-mining/279277

DFC: A Performant Dagging Approach of Classification Based on Formal Concept

Nida Meddouri, Hela Khoufiand Mondher Maddouri (2021). *International Journal of Artificial Intelligence and Machine Learning (pp. 38-62).*

www.irma-international.org/article/dfc/277433

IoT and Machine Learning on Smart Home-Based Data and a Perspective on Fog Computing Implementation

Asha Rajiv, Abhilash Kumar Saxena, Digvijay Singh, Aishwary Awasthi, Dharmesh Dhabliya, R. K. Yadavand Ankur Gupta (2023). *Handbook of Research on Machine Learning-Enabled IoT for Smart Applications Across Industries (pp. 336-349).*

www.irma-international.org/chapter/iot-and-machine-learning-on-smart-home-based-data-and-a-perspective-on-fogcomputing-implementation/326004