IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, May 2023

A Machine Learning Methodology for Diagnosing Chronic Kidney Disease

Saraswathi. P[1], Vidya Shree. CH[2], P. Geethika[3], Shree Latha[4], M Priyanka[5]
Assistant Professor, Department of Computer Science[1]
Students, Department of Computer Science[2,3,4,5]
Rao Bahadur Y Mahabaleswarappa Engineering College, Bellary, Karnataka, India

Abstract: *Chronic kidney disease (CKD) is a global health problem with high morbidity and mortality rate, and it induces other diseases. Since there are no obvious symptoms during the early stages of CKD, patients often fail to notice the disease. Early detection of CKD enables patients to receive timely treatment to ameliorate the progression of this disease. Machine learning models can effectively aid clinicians achieve this goal due to their fast and accurate recognition performance. In this study, we propose a machine learning methodology for diagnosing CKD. The CKD data set was obtained from the University of California Irvine (UCI) machine learning repository, which has a large number of missing values. KNN imputation was used to fill in the missing values, which selects several complete samples with the most similar measurements to process the missing data for each incomplete sample. Missing values are usually seen in real-life medical situations because patients may miss some measurements for various reasons. After effectively filling out the incomplete data set, six machine learning algorithms (logistic regression, random forest, support vector machine, k-nearest neighbor, naive Bayes classifier and feed forward neural network) were used to establish models. Among these machine learning models, random forest achieved the best performance with 99.75% diagnosis accuracy. By analyzing the misjudgments generated by the established models, we proposed an integrated model that combines logistic regression and random forest by using perceptron, which could achieve an average accuracy of 99.83% after ten times of simulation. Hence, we speculated that this methodology could be applicable to more complicated clinical data for disease diagnosis.*

Keywords: Logistic Regression, Random forest, Support Vector Machine, k-nearest neighbour, Naïve Bayes classifier, Neural network

REFERENCES

[1]. Z. Chen et al., "Diagnosis of patients with chronic kidney disease by using two fuzzy classifiers," Chemometr. Intell. Lab., vol. 153, pp. 140-145, Apr. 2016.
[2]. A. Subasi, E. Alickovic, J. Kevric, "Diagnosis of chronic kidney disease by using random forest," in Proc. Int. Conf. Medical and Biological Engineering, Mar. 2017, pp. 589-594.
[3]. L. Zhang et al., "Prevalence of chronic kidney disease in china: a crosssectional survey," Lancet, vol. 379, pp. 815-822, Aug. 2012.
[4]. A Singh et al., "Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration," J. Biomed. Inform., vol. 53, pp. 220-228, Feb. 2015.
[5]. M. Cueto-Manzano et al., "Prevalence of chronic kidney disease in an adult population," Arch. Med. Res., vol. 45, no. 6, pp. 507-513, Aug. 2014.
[6]. H. Polat, H.D. Mehr, A. Cetin, "Diagnosis of chronic kidney disease based on support vector machine by feature selection methods," J. Med. Syst., vol. 41, no. 4, Apr. 2017.
[7]. Barbieri et al., "A new machine learning approach for predicting the response to anemia treatment in a large cohort of end stage renal disease patients undergoing dialysis," Comput. Biol. Med., vol. 61, pp. 56-61, Jun. 2015.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, May 2023

[8]. V. Papademetriou et al., "Chronic kidney disease, basal insulin glargine, and health outcomes in people with dysglycemia: The origin study," Am. J. Med., vol. 130, no. 12, Dec. 2017.

[9]. N. R. Hill et al., "Global prevalence of chronic kidney disease - A systematic review and meta-analysis," Plos One, vol. 11, no. 7, Jul. 2016.

[10]. M. M. Hossain et al., "Mechanical anisotropy assessment in kidney cortex using ARFI peak displacement: Preclinical validation and pilot in vivo clinical results in kidney allografts," IEEE Trans. Ultrason. Ferr., vol. 66, no. 3, pp. 551-562, Mar. 2019.

[11]. M. Alloghani et al., "Applications of machine learning techniques for software engineering learning and early prediction of students' performance," in Proc. Int. Conf. Soft Computing in Data Science, Dec. 2018, pp. 246258.

[12]. Gupta, S. Khare, A. Aggarwal, "A method to predict diagnostic codes for chronic diseases using machine learning techniques," in Proc. Int. Conf. Computing, Communication and Automation, Apr. 2016, pp. 281-287.

[13]. L. Du et al., "A machine learning based approach to identify protected health information in Chinese clinical text," Int.

[14]. J. Med. Inform., vol. 116, pp. 24-32, Aug. 2018.

[15]. R. Abbas et al., "Classification of foetal distress and hypoxia using machine learning approaches," in Proc. Int. Conf. Intelligent Computing, Jul. 2018, pp. 767-776.

[16]. M. Mahyoub, M. Randles, T. Baker and P. Yang, "Comparison analysis of machine learning algorithms to rank alzheimer's disease risk factors by importance," in Proc. 11th Int. Conf. Developments in eSystems Engineering, Sep. 2018.

[17]. Alickovic, A. Subasi, "Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier," J. Med. Syst., vol. 40, no. 4, Apr. 2016.

[18]. Z. Masetic, A. Subasi, "Congestive heart failure detection using random forest classifier," Comput. Meth. Prog. Bio., vol. 130, pp. 56-64, Jul. 2016.

[19]. Q. Zou et al., "Predicting diabetes mellitus with machine learning techniques," Front. Genet., vol. 9, Nov. 2018.

[20]. Z. Gao et al., "Diagnosis of diabetic retinopathy using deep neural networks," IEEE Access, vol. 7, pp. 3360-3370, Dec. 2018.

[21]. R. J. Kate et al., "Prediction and detection models for acute kidney injury in hospitalized older adults," Bmc. Med. Inform. Decis., vol. 16, Mar. 2016.

[22]. N. Park et al., "Predicting acute kidney injury in cancer patients using heterogeneous and irregular data," Plos One, vol. 13, no. 7, Jul. 2018.

[23]. M. Patricio et al., "Using resistin, glucose, age and BMI to predict the presence of breast cancer," BMC CANCER, vol. 18, Jan. 2018.

[24]. X. Wang et al., "A new effective machine learning framework for sepsis diagnosis," IEEE Access, vol. 6, pp. 48300- 48310, Aug. 2018.

[25]. Y. Chen et al., "Machine-learning-based classification of real-time tissue elastography for hepatic fibrosis in patients with chronic hepatitis B," Comput. Biol. Med., vol. 89, pp. 18-23, Oct. 2017.

[26]. Hodneland et al., "In vivo detection of chronic kidney disease using tissue deformation fields from dynamic MR imaging," IEEE Trans. BioMed. Eng., vol. 66, no. 6, pp. 1779-1790, Jun. 2019.