EARLY DETECTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING

ABSTRACT:

Chronic kidney disease (CKD) is a global health problem with high morbidity and mortality rate as it induces other diseases. It is a severe lifelong condition caused either by renal disease or by impaired functions of the kidneys. CKD is classified as a condition that results in decreased kidney function over time, as indicated by Glomerular filtration rate(GFR). Since there are no obvious symptoms during the early stages of chronic kidney disease, patients often fail to notice the disease. As people with CKDs are likely to die at early age, early detection of this disease enables patients to receive timely treatment to ameliorate the progression of this disease. In many hospitals and clinics, there is a shortage of nephrologists or general physicians who diagnose the disease. This has resulted in patients waiting longer to get a diagnosis. Therefore, developing a system to classify a patient into classes of 'CKD' or 'Non-CKD' can help the doctors to deal with multiple patients and provide diagnosis faster. In time, hospitals and clinics can implement the proposed machine learning framework that have lower medical expert retention which can provide early diagnosis to patients in regional areas. Machine learning models can effectively aid clinicians achieve this goal due to their fast and accurate recognition performance. Those who are with Chronic Kidney Disease (CKD) are not aware that the medical tests they take for other purposes sometimes contain useful information about CKD disease. This information is sometimes not used effectively to tackle the identification of the disease. Therefore, attributes of different medical tests are investigated to identify what attributes contain useful information about CKD. A database with several attributes of healthy subjects and subjects with CKD are analyzed with different techniques. Common Spatial Pattern (CSP) filter and Linear Discriminant Analysis (LDA) are first used to identify the dominant attributes that could contribute in detecting CKD. Here CSP filter is applied to optimize separation between CKD and non-CKD. Then, classification methods are also used to identify the dominant attributes. The primary aim of the proposed system is to implement and compare the performance of various supervised and unsupervised algorithms and identify the best possible combinations that can provide better accuracy and detection rate.