Literature survey

Team id:	PNT2022TMID17480
Project Name:	Early detection of chronic kidney disease using machine learning
Team Leader:	Krishna veni.B
Team members:	Harinikrishnaveni.M Lavanya.S Pavithra.V

Paper I

Detection of chronic Kidney Diseases using machine learning Algorithms with Least Number of predictors[Marwa Almasoud , Tomas E Ward]

Abstract :

CKD is one of the most critical health problems due to increasing prevalence. Several statics tests have been done to remove redundant features Such as the ANOVA test, the percent correlation, and the Cramer's V test .We achieve an accuracy of 99.1% according to FI-measure hemoglobin has higher important for both random forest and gradient boosting in detecting CKD. Hence, we can detect CKD at only $26.65 by performing there simple tests.

Advantages:

- The classifiers have been trained ,tested and validated using 10-fold cross-validation.
- Higher performance was achieved with gradient boosting algorithm by FI-measure 99.1%, sensitivity 98.8% and specificity 93.3%.
- This result is the highest among previous studies with less number of feature and hence less cost.
- The hemoglobin has contribution in detecting CKD

Drawbacks:

- The plan to predict if a person with CKD risk factors such as diabetes, hypertension and family history of kidney failure will have CKD in the future on not using appropriate datasets.

Paper II

Early Detection of Chronic Kidney Disease Using Machine Learning Models[Dr.Dhananjaya]

Abstract:

This study developed an algorithm for predicting CKD at early stage. To predict early onset of CKD, 3 machine learning techniques are used:-

- Decision tree,
- Random forest,
- Artificial neural network,
- Support vector machines …

are used for accuracy prediction of models and we choose the best performance by assessing their accuracy.The study reveals the random forest classifier model outperformed, decision tree and support vector machines in predicting CKD and each effectiveness is evaluated.

Advantages:

- "*Support Vector Machine*" has good capability in classification band prediction problems.It increase performance by handling non linear classifications with mapping.It can discover the optimum disjunctive hyperplane.
- "*Random forest*" can accurately predict development of early stage renal disease in immunoglobulin a nephropathy patients.

Drawbacks:

- SVM sometimes ignores interaction with classifier,risk of overfitting,computability intensive,classify with dependent selection .
- "*disadvantage of random forest*": Comparative analysis of results of decision tree algorithm revealed that predicted CKD in all stages better than random forest with an accuracy of 85.5%.