

Hope Artificial Intelligence

Assignment-Classification Algorithm

Problem Statement or Requirement:

A requirement from the Hospital, Management asked us to create a predictive model which will predict the chronic kidney disease (CKD) based on the several parameters. The Client has provided the dataset of the same.

- 1.) Identify your problem statement
- 2.) Tell basic info about the dataset (Total number of rows, columns)
- 3.) Mention the pre-processing method if you're doing any (like converting string to number – nominal data)
- 4.) Develop a good model with good evaluation metric. You can use any machine learning algorithm; you can create many models. Finally, you have to come up with final model.
- 5.) All the research values of each algorithm should be documented. (You can make tabulation or screenshot of the results.)
- 6.) Mention your final model, justify why u have chosen the same.

1. Need to predict the chronic kidney disease (CKD) based on given dataset.

Problem Identification stages,

- Stage 1 - Machine Learning
- Stage 2 – Supervised
- Stage 3 – Classification

2. The dataset contains 399 rows and 25 columns

3. The dataset contains string data so converted into number – nominal (one hot encoding)

Models with score

Algorithm	F score	roc_auc_score
SVM	0.9924946382275899	1.0
DC	0.9625928174473452	0.9658058345289334
RF	0.9924946382275899	0.9998804399808704

LR	0.9924946382275899	1.0
KNN	0.9701163285572423	0.9873266379722621
NB	0.7669172932330827	0.9966523194643712

I have created many models using machine learning algorithm and tested with f score and roc_auc_score. Finally, I have selected "SVM" model which gives high score by comparing other models.

So, I saved "SVM" model with in a filename = "finalized_model_random_forest.sav"

Also, deployment for end user.