

# An Analysis on Public Health Data of Chronic Kidney Disease

Akshay Sanjeev

July 11, 2025

Preprocessing

Decision Tree

Random Forest

# Pre-processing

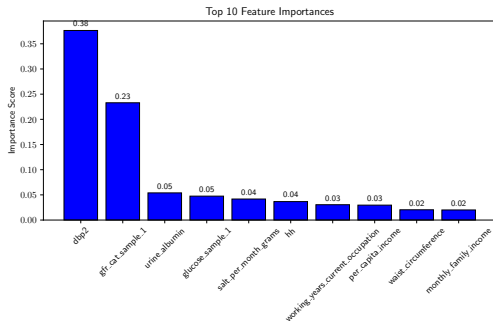
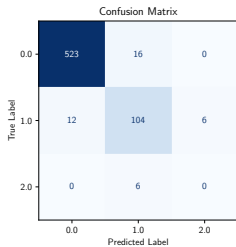
# Informal Description

$$H^i = - \sum_j^n p(x_j^i) \log p(x_j^i) \quad (1)$$

$$IG^i = H^i - \sum_s \frac{|x^{is}|}{|x^i|} H^i_s \quad (2)$$

$$x_{\text{selected}} = \arg \max_i (IG^i)$$

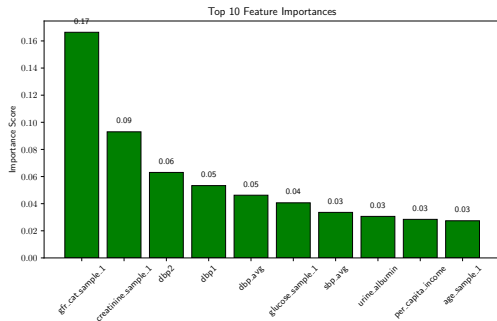
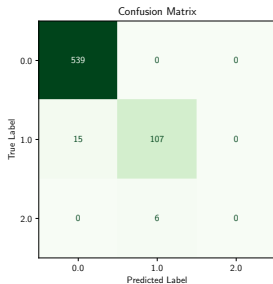
# Results



# Informal Description

$$\hat{f}_{\text{avg}}(x) = \frac{1}{B} \sum_{i=1}^B \hat{f}^i(x)$$

# Results



# Number of Trees

