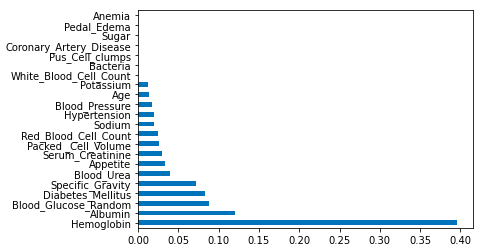
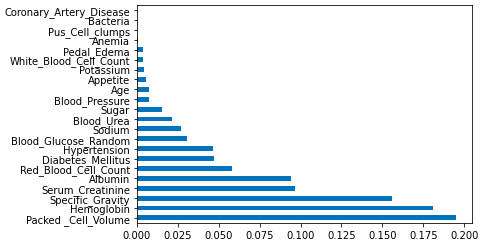
Result for Wemanity exercise on CKD:

The file CKD\_final.ipynb contains the code that can be used with a new dataset in a goal to predict if the patient has the CKD or not. The prediction is done by crossing 5 methods: linear Regression, KNN, SVC, Random Forest and XGB.

According to the results obtains in the file wemanity-exercice.ipynb, I could give some analyses about which data in more important for the prediction of CKD cases.

The first figure is the importance of columns when the algorithm uses a Random Forest method.

The three more variables are Packed Cell Volume, Hemoglobin and Specific Gravity.

The second figure is when it is using the XGB method.

It is clear to see that the Hemoglobin is a really important columns in this dataset. This result is quite normal when we know that people with diabetes are more likely to have CKD.