

Logistic Regression CA 05 – Part A

Part 3: Evaluate the performance of your model (including ROC Curve), explain the performance and draw a meaningful conclusion. (deliverable: Performance outputs in Notebook, explanation and conclusion in Word/PDF document)

Overall, I believe that the model has a good fit because the performance values are decent according to “standards”. The accuracy score is 0.6856, the f1-score is 0.66 and the AUC is 0.6621. In contrast, if we take a look at the log loss score is pretty high, 10.85. The cutoff point, as we saw in class, is 0.69 or lower. In addition, there are some values that, according to coefficients, are not making a high impact on predicting the CVD Risk. However, these are the scores that gave me the highest performance of the model. As we can see in the ROC curve and AUC, our prediction, in fact, is much higher than a “worst prediction” which is AUC = 0.5 or random, but at the same time, it is not close to the perfect prediction or 1.

