**Explanation & Conclusion**

**Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.** **When it predicts the positive result, 68% of the results are accurate.**

**Recall is the number of true positives divided by the number of true positives plus the number of false negatives.** **83% of real positive cases are correctly identified (I am using 83% from class 1 here align with the sklearn default setting).** **Recall is also known as “sensitivity” and “true positive rate” (TPR).**

**F1 Score is the weighted average of Precision and Recall. It is useful when you need to take both precision and recall into account.** **A perfect model has an F-score of 1.**

**ROC (Receiver Operating Characteristic) Curve tells us about how good the model can distinguish between two things (here for our CVD Risk is Yes or No).**

**The point of perfect classification is the top-left corner of the ROC plot where the TP rate is 1 and the FP rate is 0 — that is, no 1 are classified as 0 and no 0 are classified as 1.**

**Log-loss is indicative of how close the prediction probability is to the corresponding actual/true value (0 or 1 in case of binary classification). The bolder the probabilities, the better will be your Log Loss (closer to zero). It is a measure of uncertainty (you may call it entropy), so a low Log Loss means a low uncertainty/entropy of your model. In the worst case, let us say you predicted 0.5 for all the observations. So, log-loss will become -log (0.5) = 0.69. Hence, we can say that anything above 0.6 is a very poor model considering the actual probabilities.**

**Since the log loss result here is absurd, there might be an error occurred. Thus, I decide to put this result aside temporarily.**

**To conclude, my logistic regression model regarding the topic of** Cardiovascular Disease (CVD) performs moderately well without the consideration of the log loss.