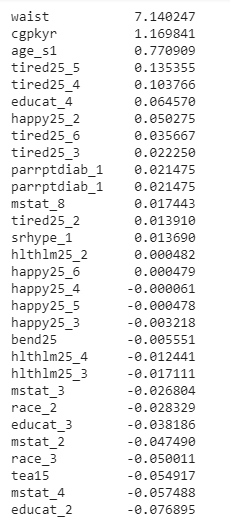
## **CA05 A Logistic Regression**

Logistic Regression

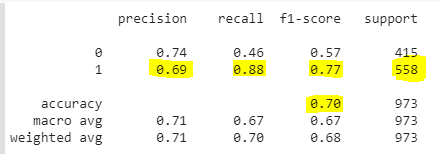
**Part 3:** Evaluate the performance of your model (including ROC Curve), explain the performance and draw a meaningful conclusion.

* I have used RandomSearchCV() module to get the best performing model and below is the performance score of the best model using Logistic Regression model.
* The top 10 parameters in the increasing order of ranking are as follows. Waist, cgpkyr, age\_s1, tired25\_5 are the most important variables in predicting whether a person has cardiovascular disease or no.
* These estimates tell you about the relationship between the independent variables and the dependent variable, where the dependent variable is on the logit scale (0 to 1).  These estimates tell the amount of increase in the predicted log odds of cvd= 1 that would be predicted by a 1 unit increase in the predictor, holding all other predictors constant.



The model’s performance metrics is given below:

* Precision: 0.69, Recall: 0.88, f1-score: 0.77, support: 558
* Accuracy 0.70 – the model has 70% accuracy in determining the relationship between the independent variables and dependent variables



ROC AUC score - 0.67

**ROC** is a probability **curve** and **AUC** represents degree or measure of separability

