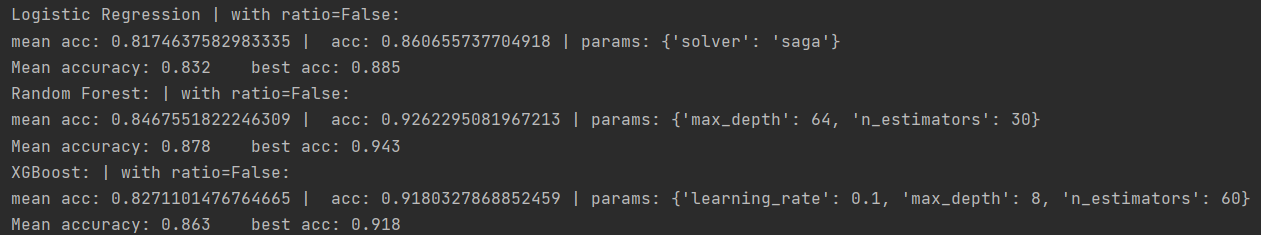
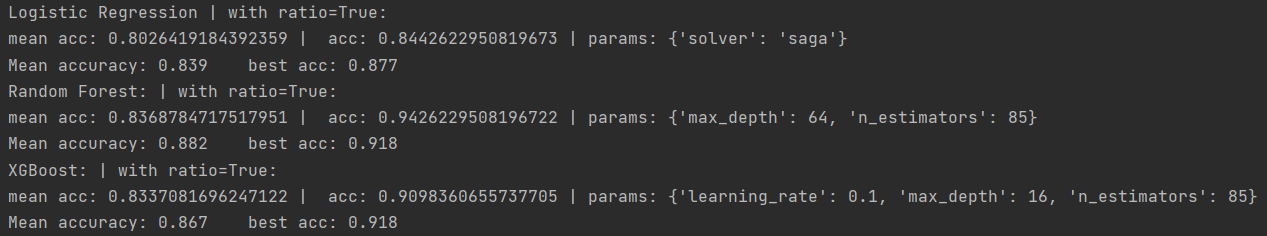
**Task 1 – Predict COVID-19 from routine blood tests:**

3. - The first new feature will be the ratio between Hemoglobin and Red blood cells () because they are strongly related with the amount of oxygen in the blood.  
 - The second new feature will be the ratio between Leukocytes and Red blood cells () which is the ratio between white blood cells (the type that fight against infections) and red blood cells.  
 - Another new feature will be the ratio between Lymphocytes and Red blood cells (), which is the ratio between white and red blood cells.  
 - Another new feature will be the ration between Monocytes and Red blood cells () which is the ratio between immune cells and red blood cells.  
 - Another new feature will be the ration between Neutrophils and Red blood cells () which is the ratio between immune cells and red blood cells.  
**Results:**

Results without new features:  
Results with new features:  
The mean accuracy rate of the models are +-2% deviation from the results without the new features. However, the hyper-parameters changed. Both in RF and XGboost the number of estimators and the max depth increased so we assume that the new models are less general and over-fitting the dataset.

4. Results:  
LightGMB:  
mean accuracy is 83.2%, while the best model with accuracy of 91.8% with the hyper parameters:  
learning rate = 0.1, max depth = 16 and number of estimators is 100.

CatBoost:



mean accuracy is 83.2%, while the best model with accuracy of 90.9% with the hyper parameter:  
learning rate = 0.01.