

Steel Plates Dataset

Richard Szarka

Marko Stahovec

Dataset

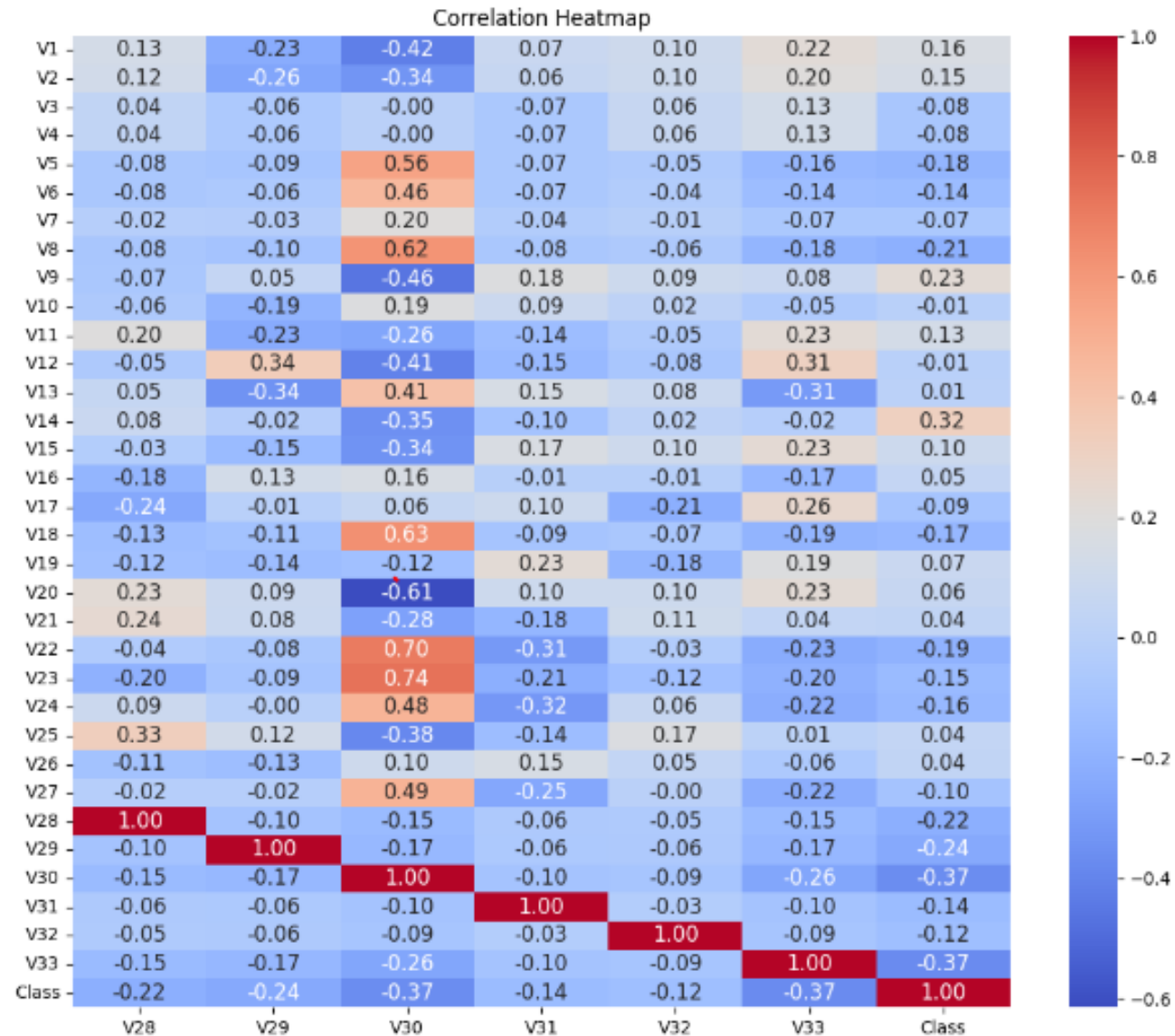
- No missing values
- 34 columns (7 of them for multiclass)
- Last class can be used for binary classification
- shape: (1941, 34)

Variable Information

27 independent variables:

X_Minimum
X_Maximum
Y_Minimum
Y_Maximum
Pixels_Areas
X_Perimeter
Y_Perimeter
Sum_of_Luminosity
Minimum_of_Luminosity
Maximum_of_Luminosity
Length_of_Conveyer
TypeOfSteel_A300
TypeOfSteel_A400
Steel_Plate_Thickness
Edges_Index
Empty_Index
Square_Index
Outside_X_Index
Edges_X_Index
Edges_Y_Index
Outside_Global_Index
LogOfAreas
Log_X_Index
Log_Y_Index
Orientation_Index
Luminosity_Index
SigmoidOfAreas

Dataset



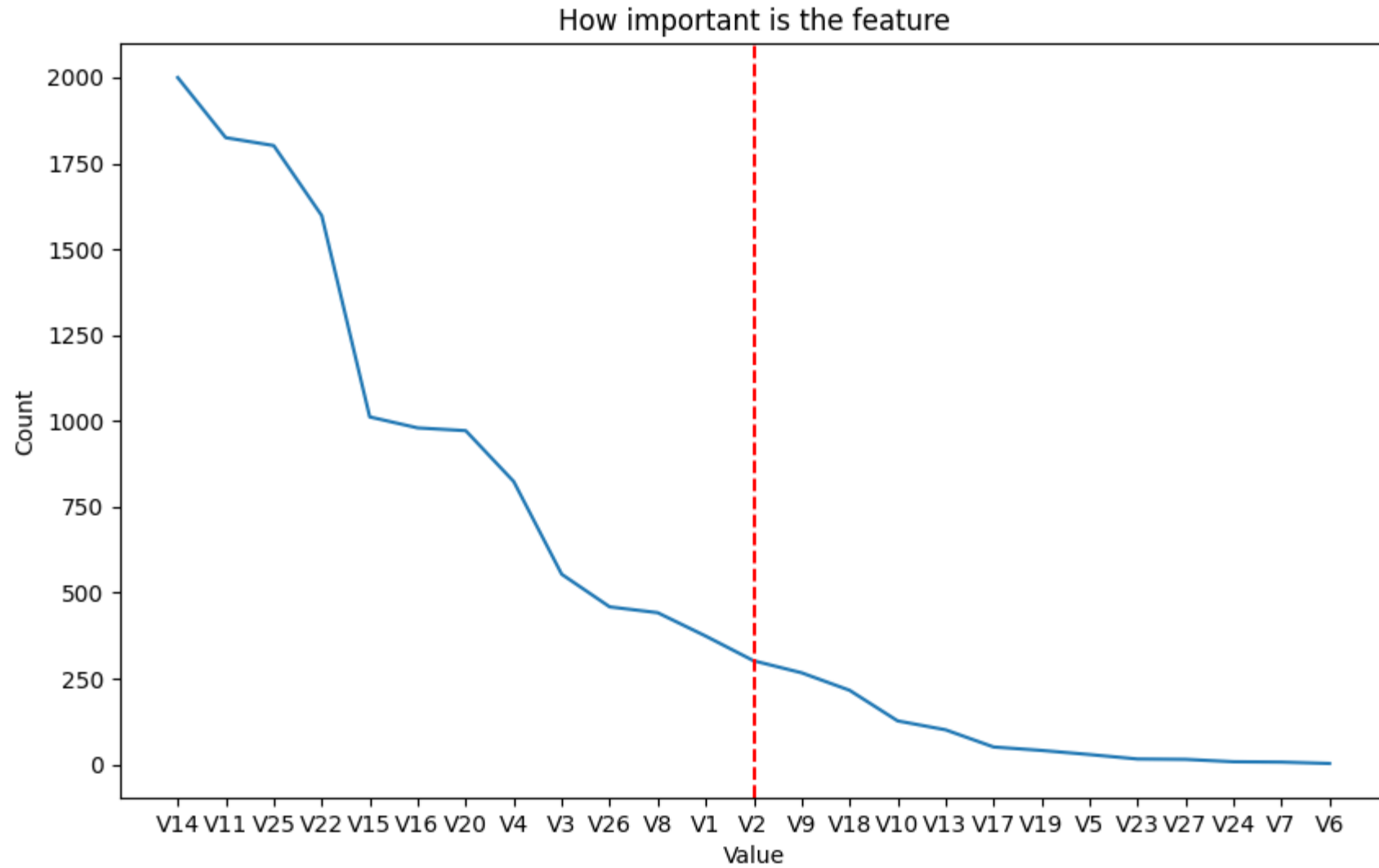
Dataset



Preprocess

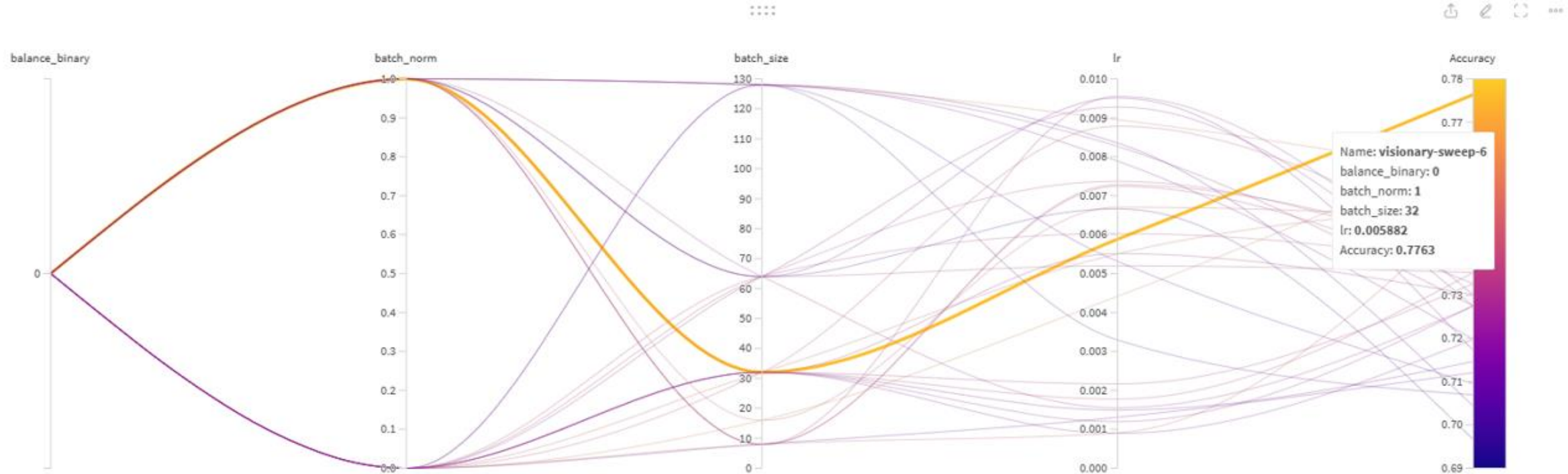
- Z – score normalization
- Train valid split 80:20
- Categorical parameters were only binary (transformation to 0 and 1)
- Feature selection based on DecisionTree importance (≤ 0.05)

Feature Selection



Hyperparameters

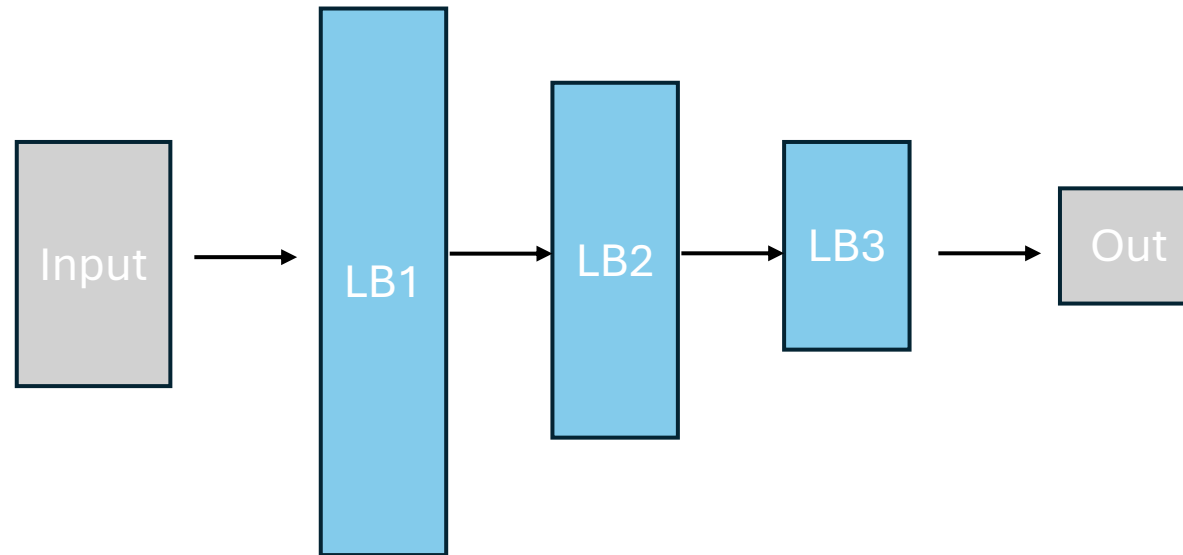
- Batch size: 32
- Learning rate: 0.006
- Patience: 150 epochs
- Optimizer: SGD - (so we can see the whole impact of batch norm)
- Loss -> BCE for binary and CrossEntropy for multiclass



Architecture

Architecture:

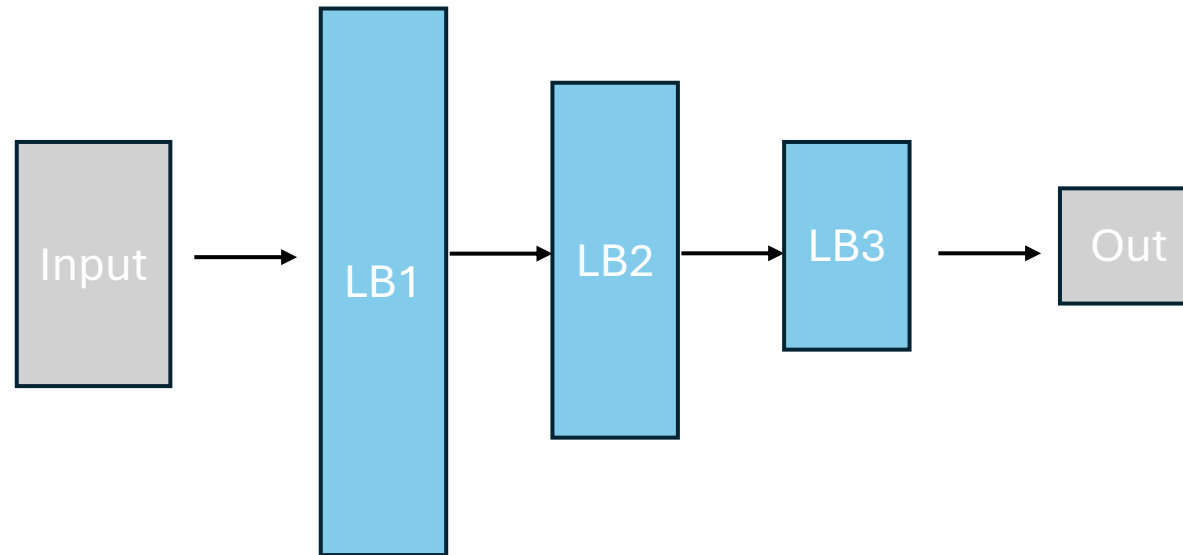
- Linear block - LB (linear -> batch norm -> ReLU)
- Sigmoid (only for binary classification)



Architecture

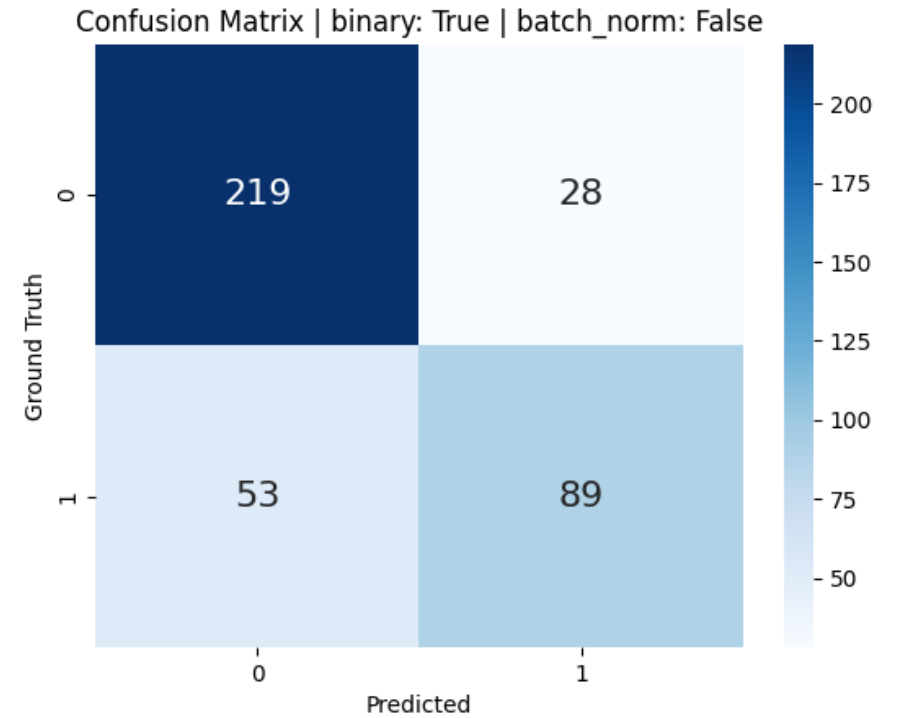
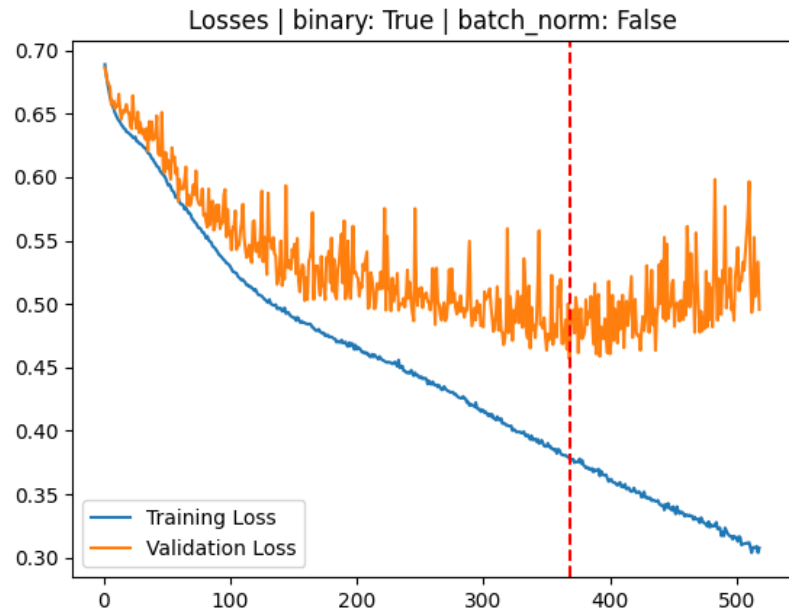
Architecture:

- Linear block - LB (linear -> batch norm -> ReLU)
- Sigmoid (only for binary classification)

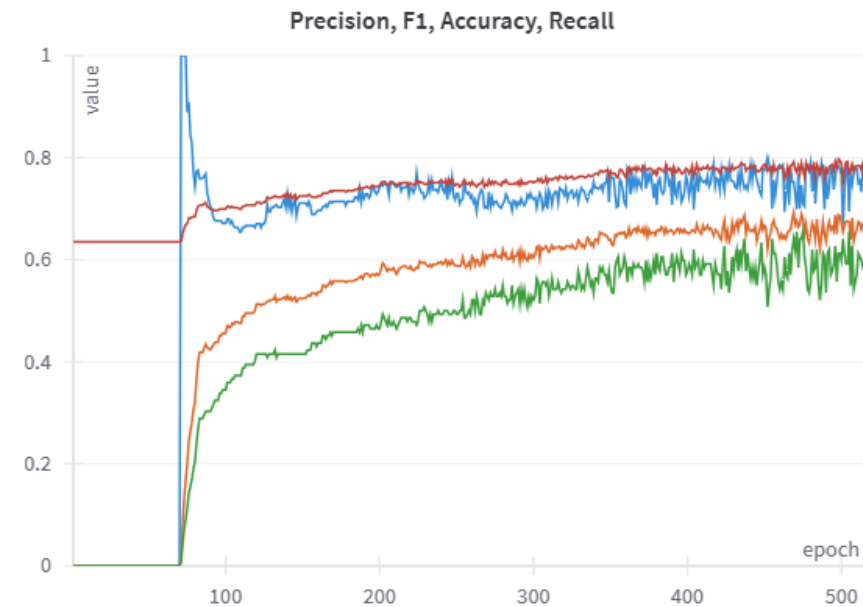


Experiments

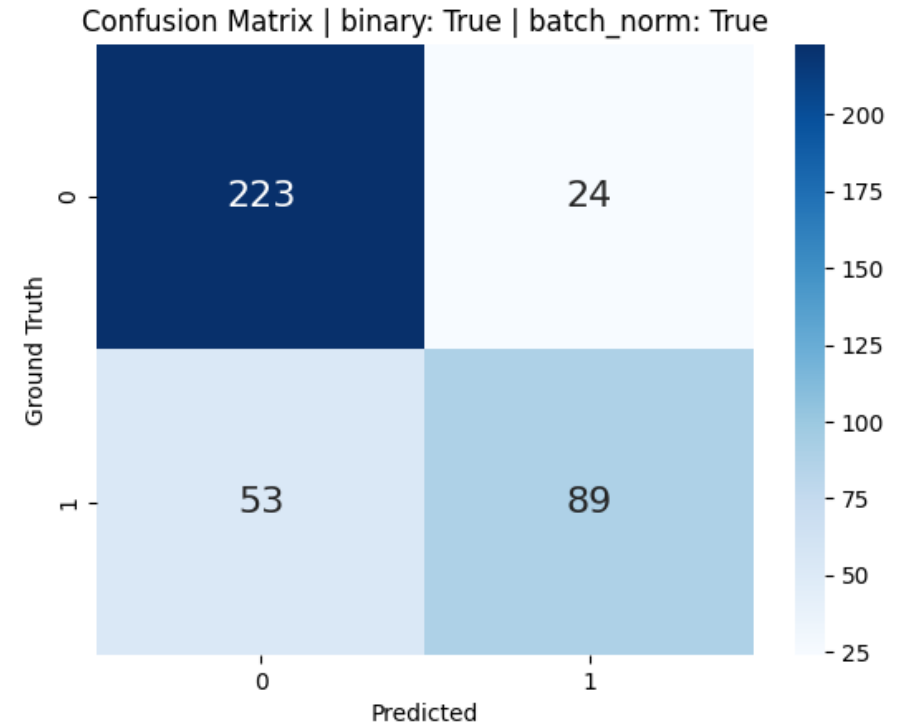
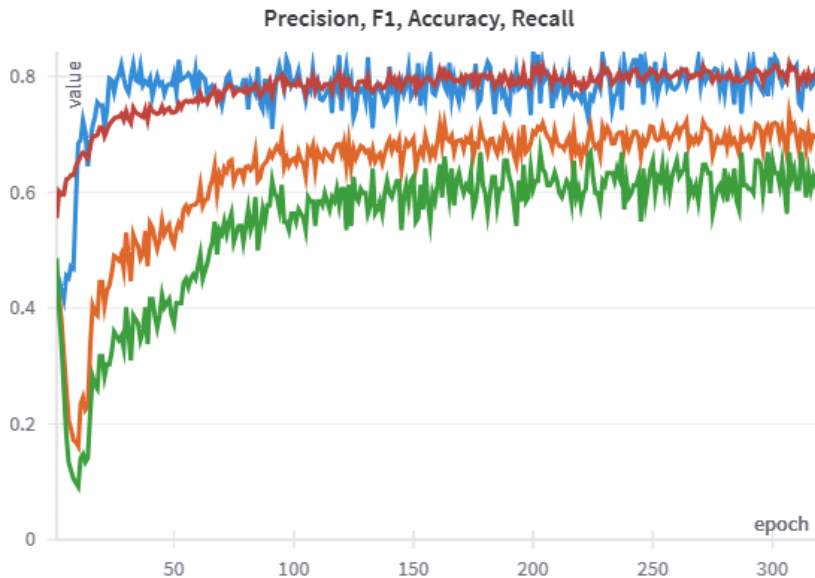
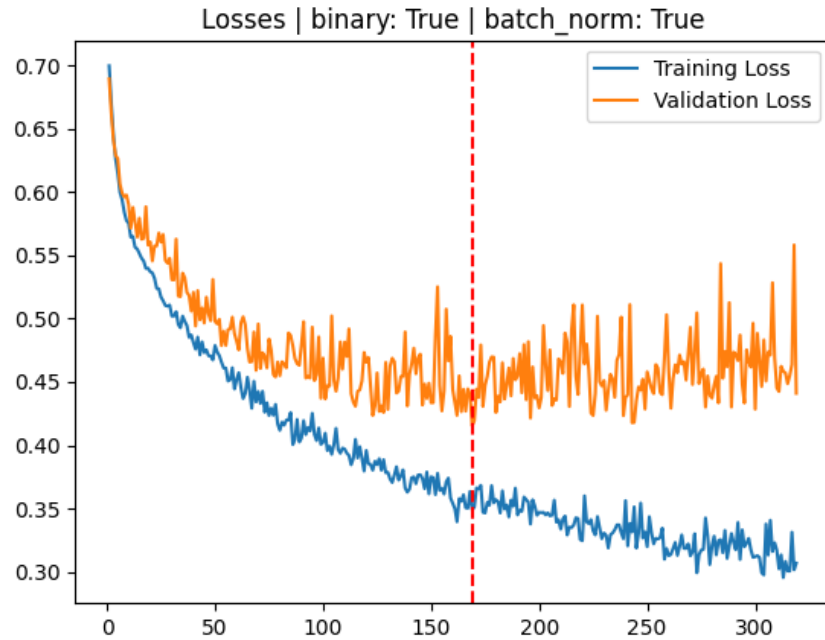
Binary classification



Accuracy: 0.7917737789203085
Precision: 0.7606837606837606
Recall: 0.6267605633802817



Binary classification – batch norm

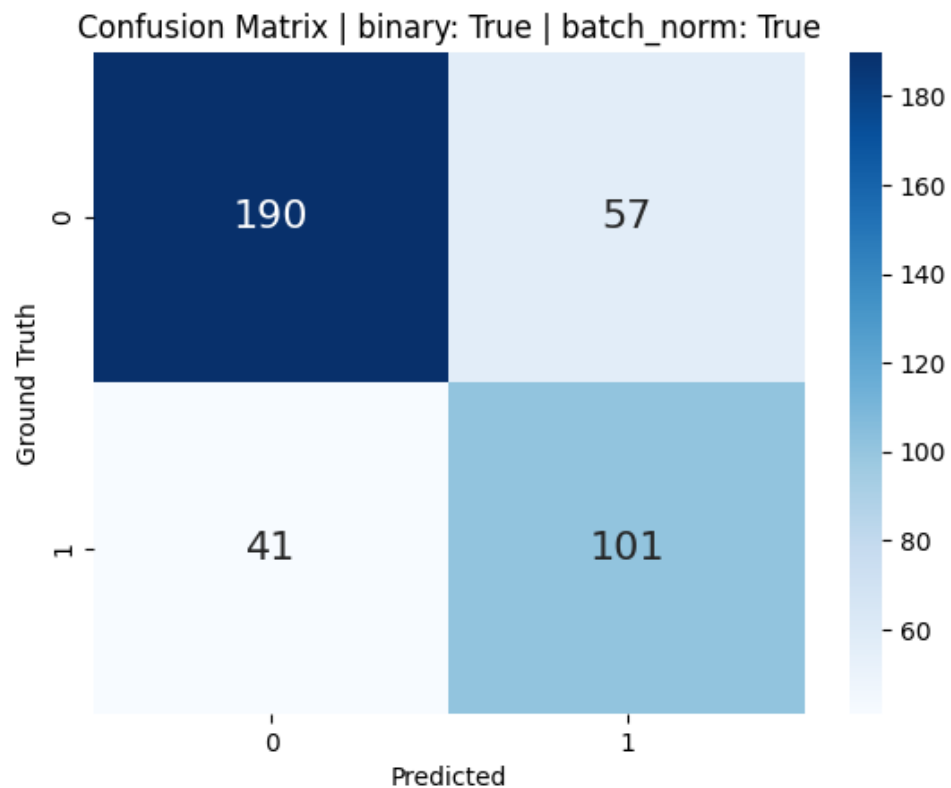
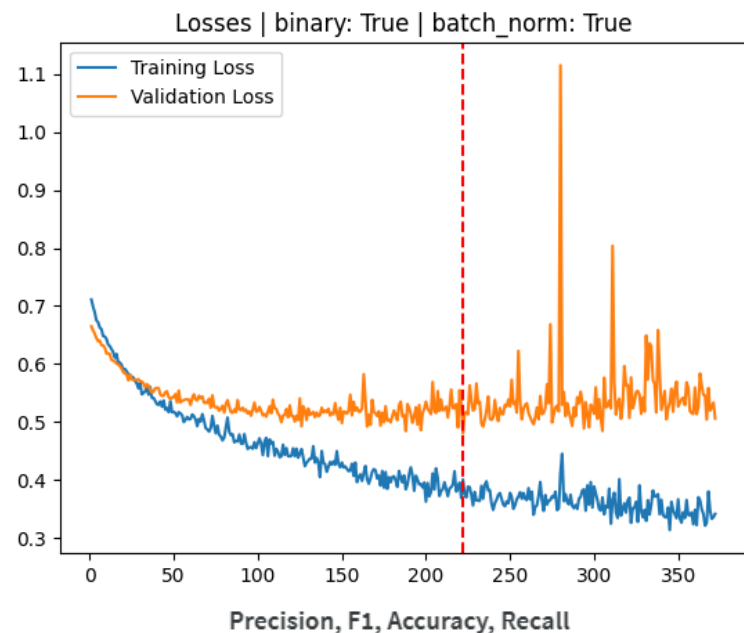


Accuracy: 0.8020565552699229

Precision: 0.7876106194690266

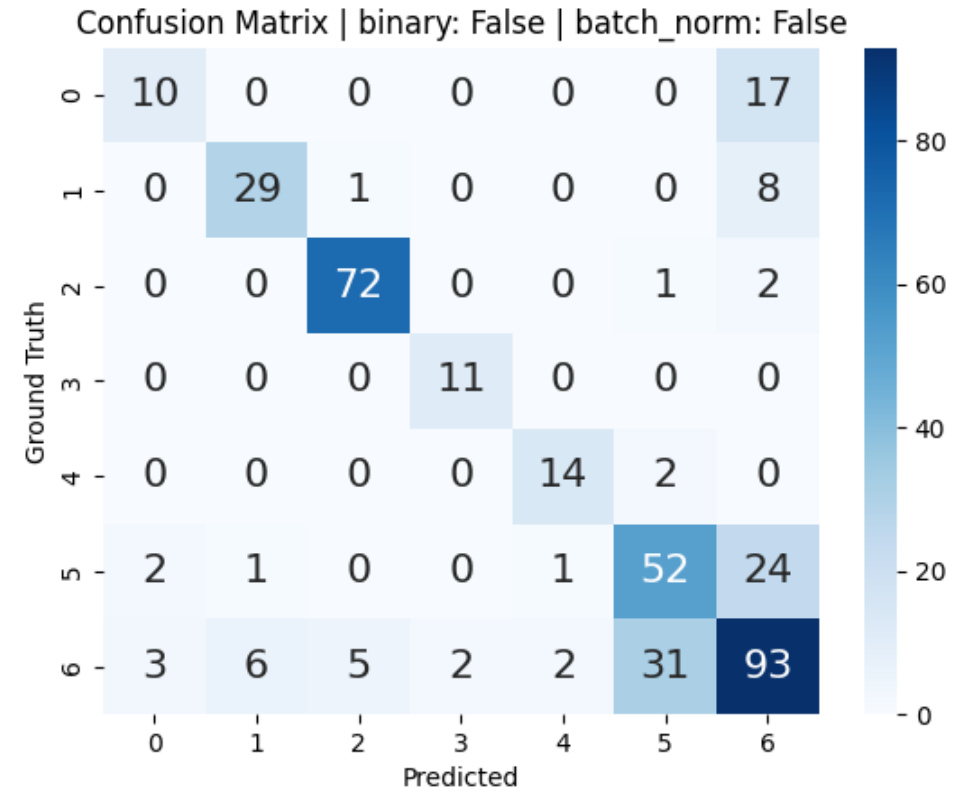
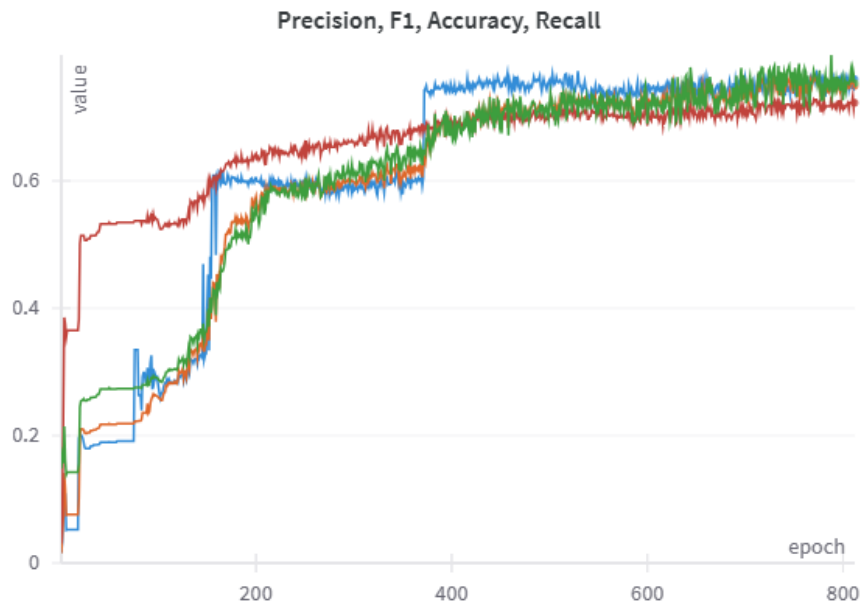
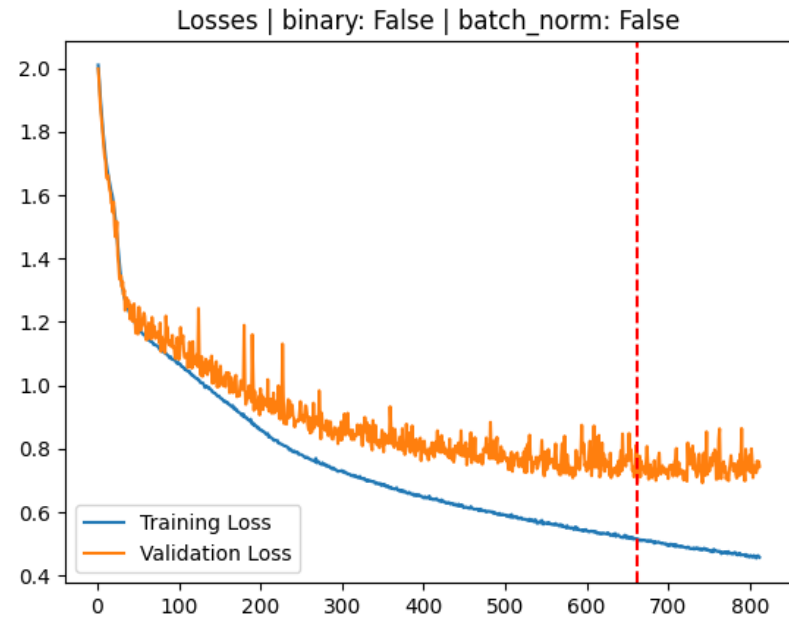
Recall: 0.6267605633802817

Balanced Binary classification – batch norm



Accuracy: 0.7480719794344473
Precision: 0.6392405063291139
Recall: 0.7112676056338029

Multiclass classification

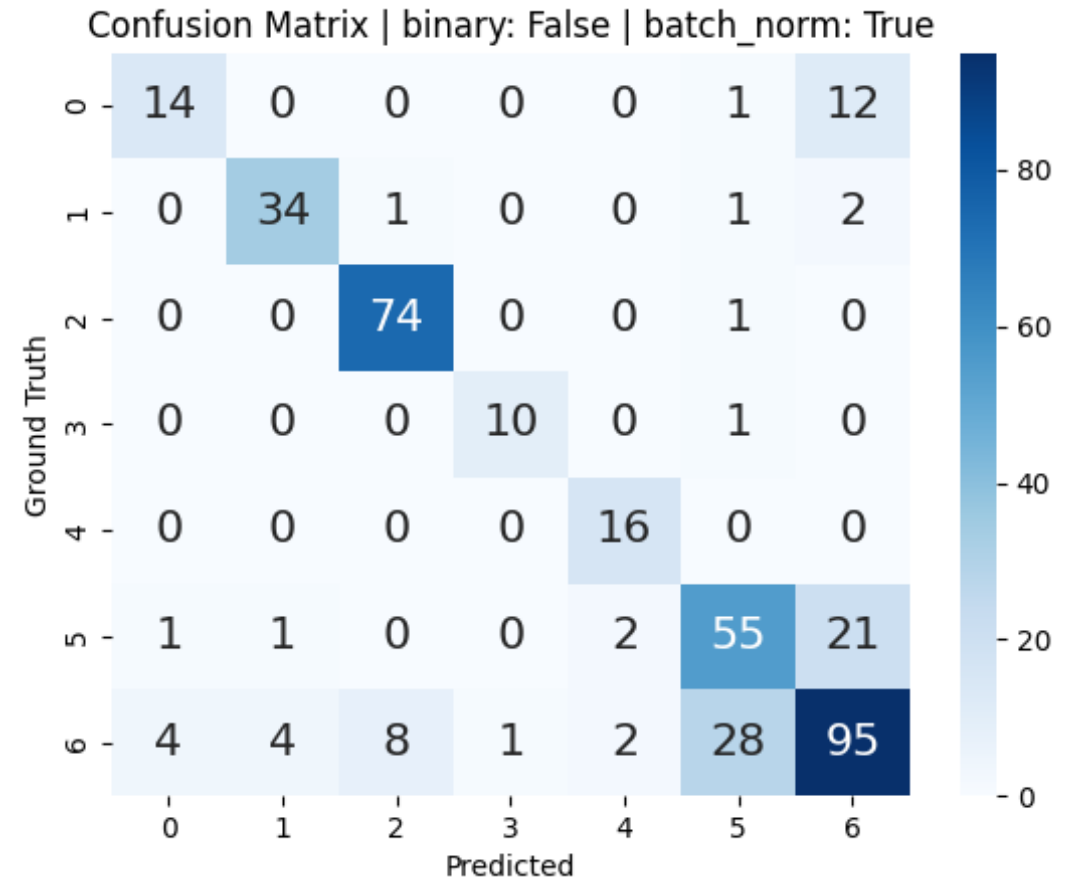
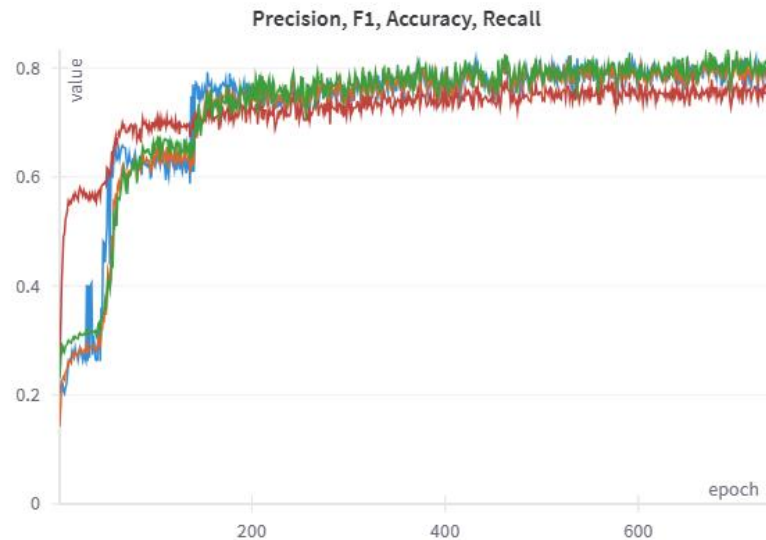
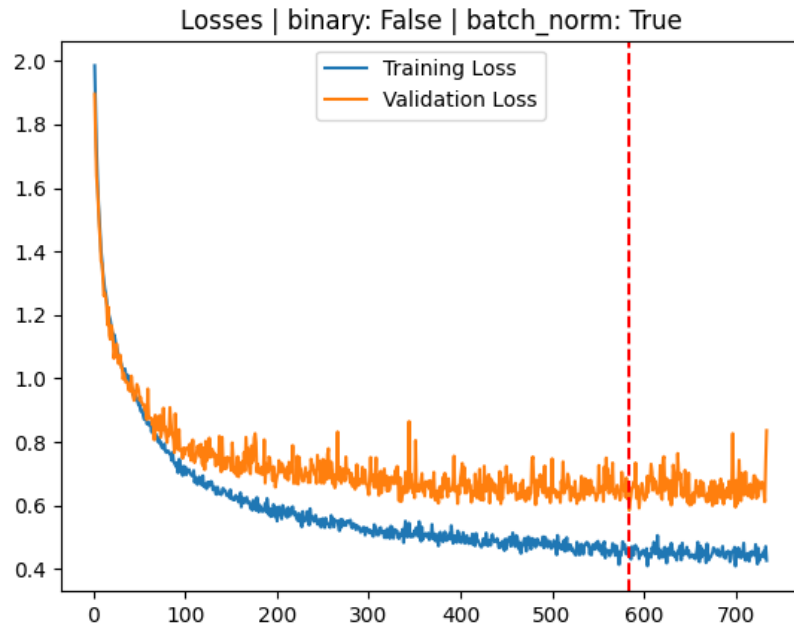


Accuracy: 0.7223650385604113

Precision: 0.7593524141916755

Recall: 0.7533511203674289

Multiclass classification – batch norm



Accuracy: 0.7660668380462725

Precision: 0.7960353271463412

Recall: 0.8093610029840571

Useful links to project

<https://api.wandb.ai/links/r-szarka/q5gr5cww> - sweep

<https://api.wandb.ai/links/r-szarka/ou5z1ct2> - multiclass

<https://api.wandb.ai/links/r-szarka/uoiclrgf> - binary

https://github.com/RichardSzarka/NN_project_1 - github