

École Polytechnique Fédérale de Lausanne



Introduction to Machine Learning CS-233(a)

---

Milestone 1 Report

---

Ann-Kristin Bergmann, 362802

Caspar Martin Schön, 362091

Adrian Zvizdenko, 362739

*November 7, 2022*

## 1 Logistic Regression

The accuracy of classification using Logistic Regression, depends in general on 2 hyper-parameters (**learning rate** and **maximum iteration number**). We decided to fix the maximum iteration parameter to 1000 iterations and apply cross-validation on the learning rate. The results of different search parameters in the range  $10^{-2}$  to  $10^{-5}$ , are plotted below, with the appropriate metric on the x-axis (F1-score for classification). The best hyper-parameter found is 0.0001. Another optimization we applied in the algorithm, is to stop the gradient descend algorithm, whenever the accuracy doesn't further improve. That allowed the model to predict the test labels with up to  $\sim 78\%$  accuracy and 0.76 F1-score.

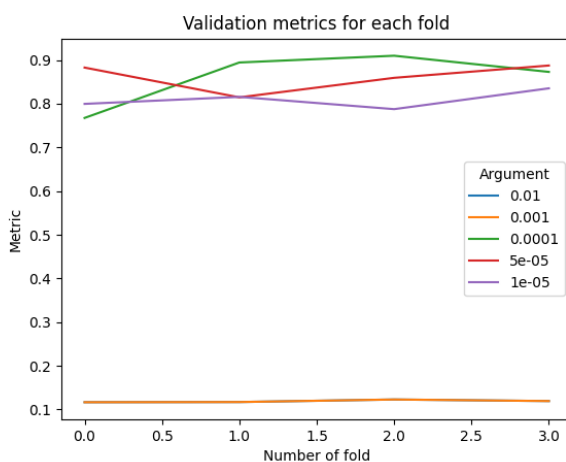


Figure 1: Cross-validation for learning rate

```
Best hyperparameter found is -> 0.0001
Metric for the best hyperparameter -> 0.8618786156492214
Started Logistic Regression training with learning rate 0.0001 and max iterations 1000
Training accuracy at iteration 0 is 0.6512166859791425
Final accuracy after training is 0.9165701842873696
Runtime of Logistic Regression training: 14.186721086502875 sec
Final classification accuracy is 77.831715210356
Final macro F1 score is 0.7644447313911666
```

Figure 3: Console output for cross-validated logistic

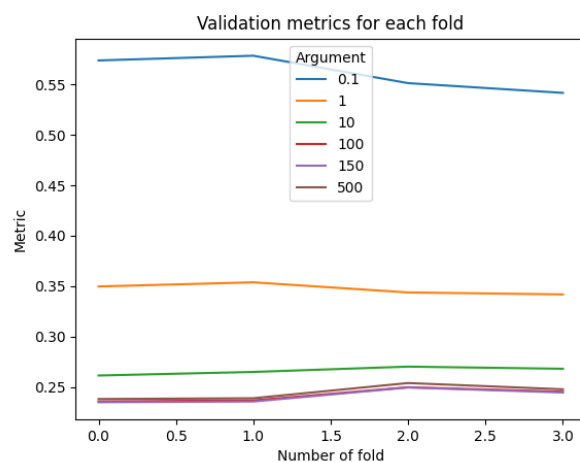


Figure 2: Cross-validation for ridge  $\lambda$

```
Best hyperparameter found is -> 150
Metric for the best hyperparameter -> 0.239901975251897
FIT & PREDICT phase ...
Started Ridge/Linear Regression training with lambda 150
MSE value for ridge/linear regression training 0.16365054275831062
Final loss is 0.36744674733568694
```

Figure 4: Console output for cross-validated ridge

## 2 Ridge/Linear Regression

The Ridge and Linear Regression are combined within the same class, where the hyper-parameter is  $\lambda$ . A value of  $\lambda = 0$ , would correspond to the Linear Regression. In order to find the best hyper-parameter for the regression problem, cross-validation was applied to a range of arguments from 0.1 to 500. The best hyper-parameter found was 150, having the least *MSE* loss. The results of cross-validation, can be visualized in the plot below, having the *MSE* metric on x-axis and the number of the fold as y-axis. Finally, the *MSE* loss of the trained Ridge Regression was lowered to 0.369.