

Laser classification

Ekaterina Voloshinova

3 сентября 2024 г.

Data

- ▶ Measurements of lasers
- ▶ 200 samples, each assigned a class label
- ▶ -1 for faulty, 1 for functioning
- ▶ 100 functioning, 100 faulty (balanced classes)
- ▶ Classification problem: given a sample of measurements of a laser, say if it is functioning or not

Evaluation and Hyper-parameter tuning

- ▶ Evaluation: Hold-out testing (20% for test dataset)
Measuring accuracy and F1-score
- ▶ Hyper-parameter tuning:
K-fold cross-validation + F1-macro score

Models

- ▶ Logistic regression

$$\operatorname{argmin}_{\theta} \sum_{i=1}^n \log 1 + e^{-y_i x_i^T \theta}$$

- ▶ Accuracy: 0.95
 - ▶ F1 score: 0.9565

- ▶ Naive Bayes

- Accuracy: 0.875
 - F1 score: 0.878

Neural network

- ▶ activation function: relu
- ▶ hyper-parameters: number of neurons in hidden layer, alpha
- ▶ best model: hidden_layer_size = 252, alpha = 0.01
- ▶ accuracy: 0.925
- ▶ F1-score: 0.933

Trees and forests

- ▶ Decision tree

 - criterion: entropy

 - accuracy: 0.95

 - F1-score: 0.9565

- ▶ Random forest

 - Hyper-parameters: n_estimators, criterion, max depth

 - best model: 100, Gini, None

 - accuracy: 0.95

 - F1 score: 0.95