



# Applied Machine Learning in Engineering

**Exercise 13, July 18, 2023**

Prof. Merten Stender

Cyber-Physical Systems in Mechanical Engineering, Technische Universität Berlin

[www.tu.berlin/cpsme](http://www.tu.berlin/cpsme)

[merten.stender@tu-berlin.de](mailto:merten.stender@tu-berlin.de)

# Teaching Evaluation



Cyber-Physical Systems  
in Mechanical Engineering TU Berlin

- This is our „loss function“ for better teaching. We would love to backpropagate the feedback into our teaching for the next term 😊
- Please give your feedback
  - Things that you did like, things we should continue
  - Things that we should improve
- Link: <https://befragung.tu-berlin.de/evasys/online.php?p=U3MHS>

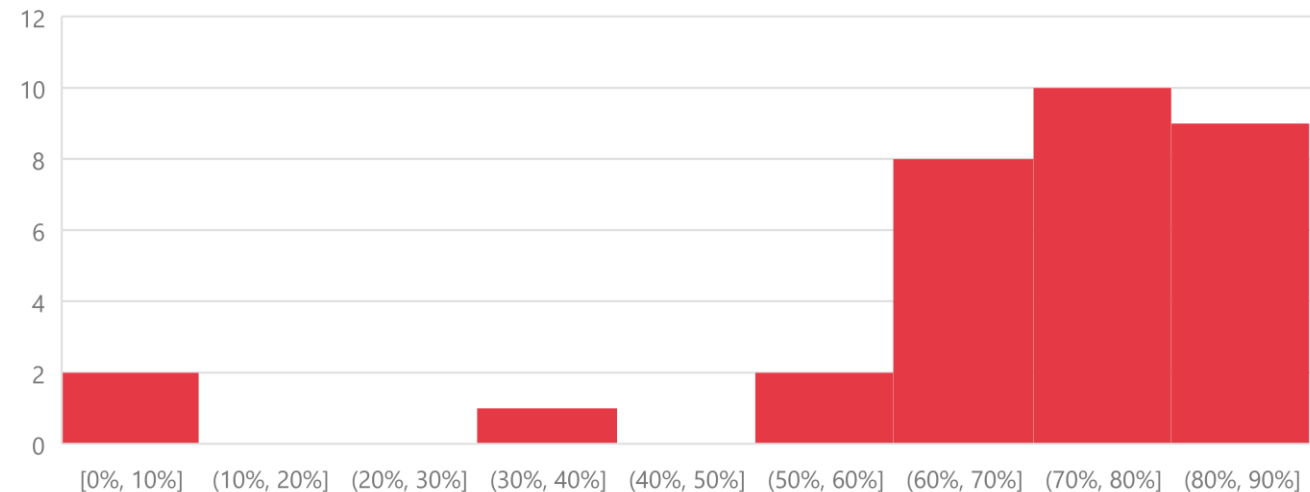


# Homework No. 2



- Results published online (ISIS)
- Task 1: mostly very high points
- Task 2: some points missed
  - Explaining hyperparameters
  - Explaining the basic working mechanism of the models
  - Reporting CV scores
- Task 3:
  - Many good ideas for improvements
  - Sometimes only focused on where the model was **weak**, not where the model was **strong**

Scores Homework No. 2



Main reason for losing points:

**K-fold cross validation for comparing models**

# Comparing Different ML Models



- Individual model performance (can be) highly sensitive to randomly initialized internal parameters (weights for neural nets) or training data set distribution
- We do not want to compare models that are highly sensitive, but reduce that uncertainty
- K-fold cross validation (or k-times bootstrapping): multiple models trained on multiple data splits.
- If a model performs well on average (bias!), it has a low sensitivity to data and initialization, hence we can use it. The larger the spread of the predictions (variance), the more uncertain the model is.

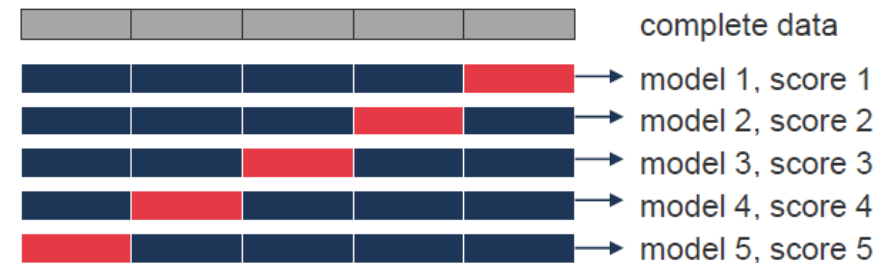


## k-fold Cross Validation



- **Procedure:**

1. Split data set into  $k$  subsets
2. Use  $(k - 1)$  subsets as training set, and remaining subset as validation set
3. Repeat by iterating over all subsets



- Report of evaluation metric:

score = mean  $\pm$  std. dev. (validation set score)

training

validation

- **When to use?**

- Comparatively small data set with strong data spread
- Simple train-test split yields non-repeatable results (within some tolerance)
- Measure bias and variance statistics of your model
- Compare different model architectures against each other

# Student Job



Cyber-Physical Systems  
in Mechanical Engineering TU Berlin

- 40h/month (negotiable)



# Q&A