AI for Learning & Optimization
AI CENTER TELKOM UNIVERSITY

Telkom University

Model Evaluation

Artificial Intelligence dan Big Data | AAK2KAB3 | Kur. 2024 | 2024/2025

Content List

- The idea of model evaluation
- Investigating Model Fitness in Machine Learning
- Exploiting Hyperparameter Tuning & Optimizer

The idea of model evaluation

AI for Learning & Optimization
AI CENTER TELKOM UNIVERSITY

What is model evaluation?

Why do we need to evaluate model?

How to evaluate model?

The idea of model evaluation

Preparing dataset for further operation, including data cleansing, correction, imputation, adjustment, etc.

- Web-app
- Mobile-app
- Desktop-app
- etc.

Data Collection → Data Preprocessing → Model Training → Model Evaluation → Model Deployment

- Algorithm selection based on characteristics of datasets
- Classification: Decision Tree, Naive Bayes, Random Forest, K-Nearest Neighbours, SVM, etc.
- Regression: linear regression, decision tree regressor, etc.
- Dataset separation: training and testing
- Fitting model using training dataset

The idea of model evaluation

Data Collection → Data Preprocessing → Model Training → Model Evaluation → Model Deployment

Raw Data

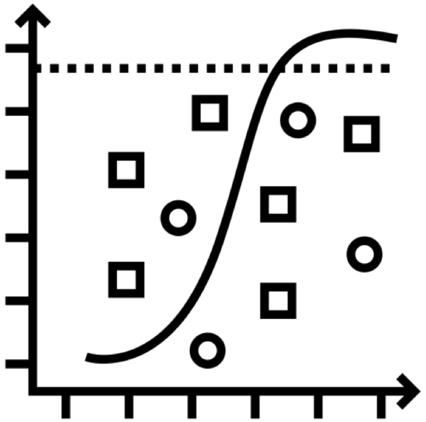
Preprocessing

Model

Is it edible?
Is it delicious?
So forth

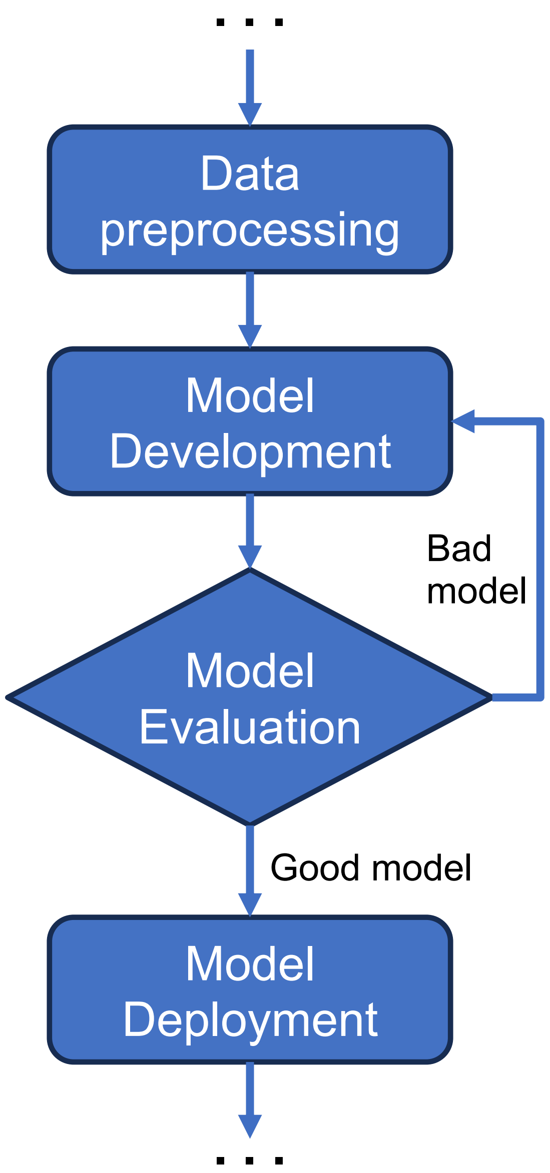
The idea of model evaluation

- Model Evaluation is a way to check our model developments whether they have reached our target (good enough) or even worst compared to other works.
- In real world, we may create several model or run multiple times to get the best model.



Is it good enough?

The idea of model evaluation



What is good/bad model?

The idea of model evaluation

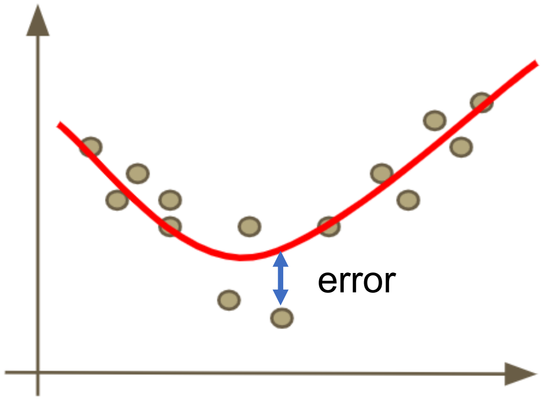
What is a good Machine/Deep Learning model?

- [Acad.] Solving given problem within certain dataset → high accuracy & precision, min. RMSE, etc.
- [Acad.] Low complexity / computational cost → high throughput, low latency, min. memory usage, etc.
- Gaining revenue/profitable
- So forth

Some Evaluation Metrics

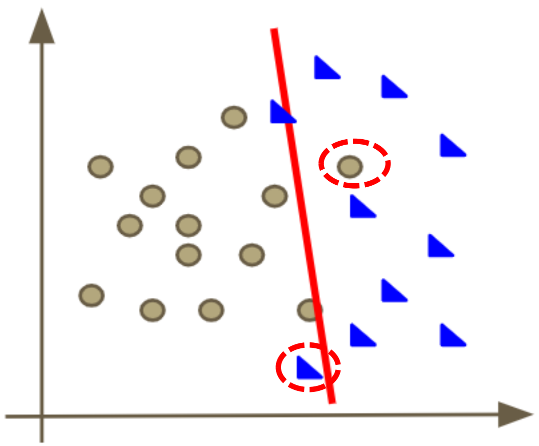
- Regression → predicting trend of data

  RMSE (root mean square error),
  MAE (mean absolute error)
  R^2 (R-squared)
  Etc.



- Classification → assigning object into certain group of classes

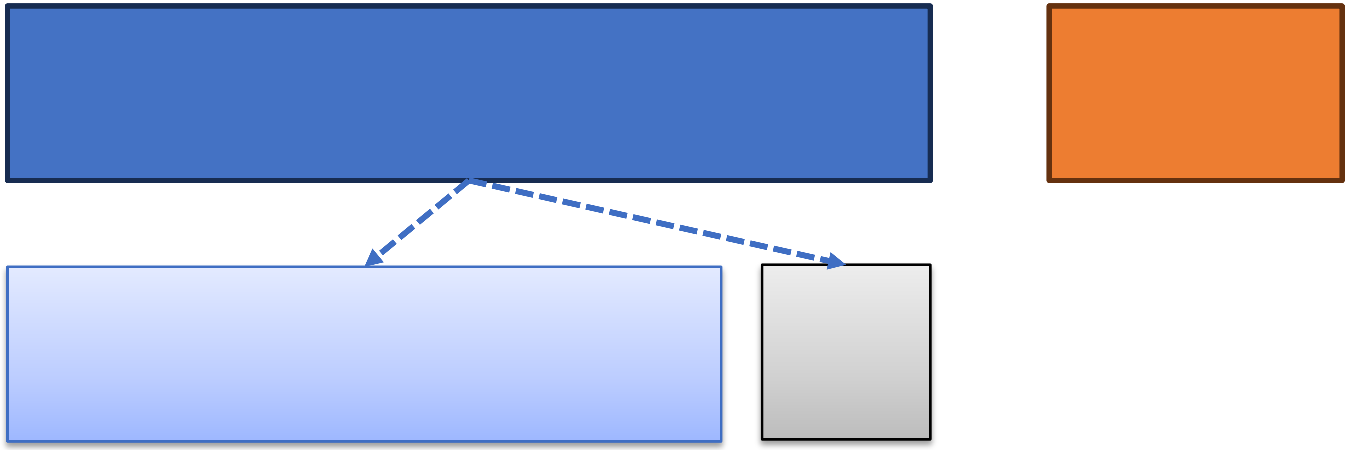
  Accuracy
  Precision
  recall
  F1-score



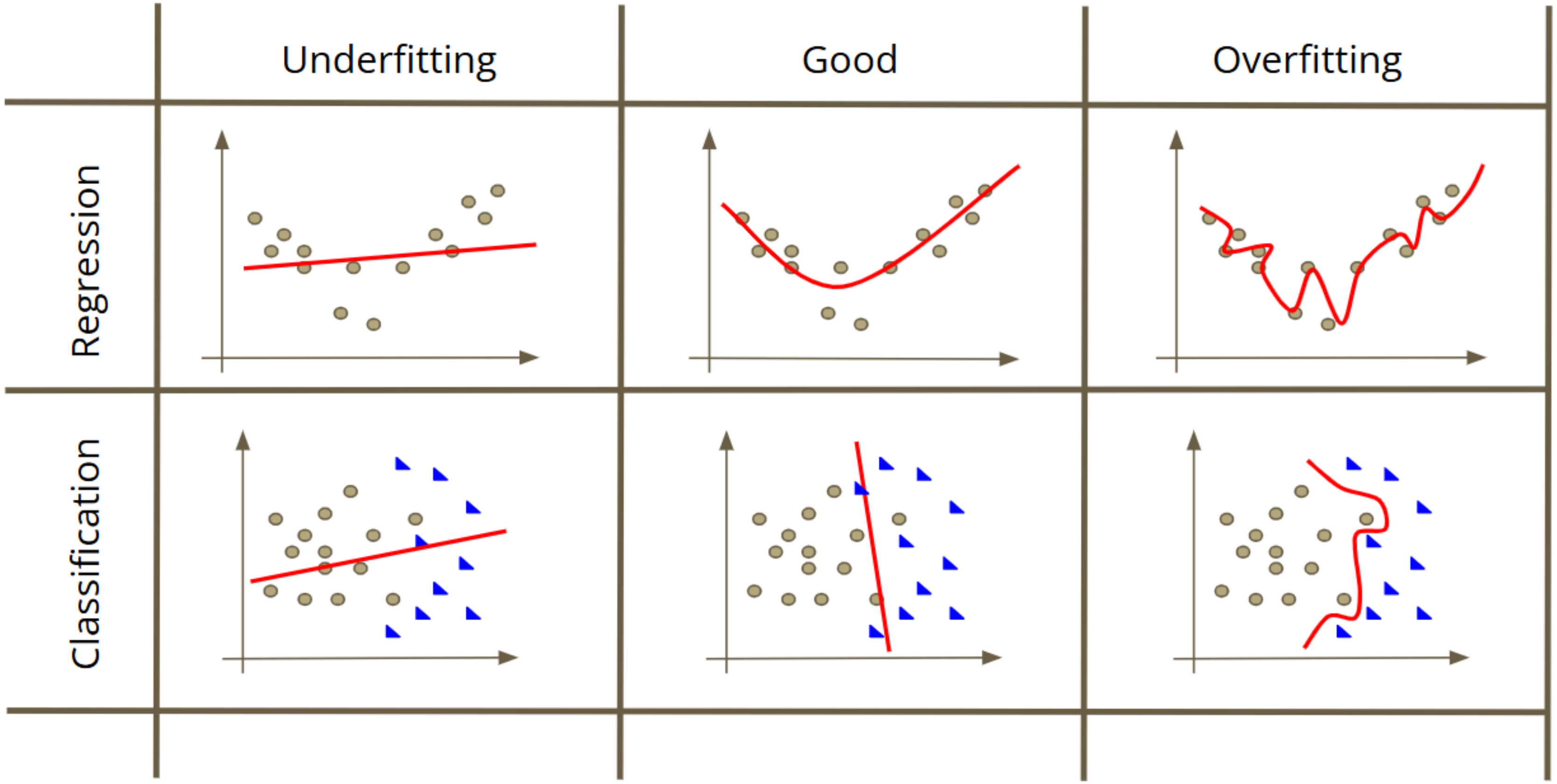
Validation vs Testing

- The validation is a part of model building/ development/ training
- It is utilized to fine tune hyperparameter during model training by dividing training dataset into several partitions of dataset

- The testing is a way to test the model against new set of random data (testing dataset or totally new samples of data)



Validation vs Testing



Underfitting	Good	Overfitting
Bad in validation	Good in validation Good in testing	Good in validation Bad in testing

Investigating Model Fitness in Machine Learning

Cross-Validation is a technique for evaluating a model by dividing the dataset into several parts (folds). This technique helps in reducing the variance that arises due to the selection of an unrepresentative dataset and provides a better picture of the model's performance.

Types of Cross-Validation:

- **K-Fold Cross-Validation**: The dataset is divided into k parts or "folds". The model is trained on k-1 folds and tested on the remaining folds. This process is repeated k times, with each fold used as test data once. The evaluation results are then averaged to reduce bias.

- **Leave-One-Out Cross-Validation (LOOCV)**: Each data point is used as test data once, and the rest are used to train the model. This technique is more computationally intensive, but is very useful for small datasets.

- **Stratified K-Fold Cross-Validation**: Splitting the data in a way that keeps the class distribution balanced in each fold, often used in classification cases with imbalanced data.

Exploiting Hyperparameter Tuning & Optimizer

Hyperparameter tuning is the process of finding the best combination of model parameters that can produce the best performance. Hyperparameters are parameters that are not learned by the model during training, such as the depth of the decision tree, the number of estimators in Random Forest, or the learning rate in Gradient Boosting models.

Metode untuk Hyperparameter Tuning:

1.**Grid Search**: Grid Search tries all combinations of the hyperparameters we specify in a grid. Although this process is very thorough, it can be very slow for large datasets.

2.**Random Search**: Random Search randomly selects a combination of hyperparameters within a specified range. It is faster than Grid Search, but may not find the optimal combination.

1.**Bayesian Optimization**: A more sophisticated approach that uses probabilistic to choose more efficient hyperparameters, reducing over-exploration.

Exploiting Hyperparameter Tuning & Optimizer

An optimizer is an algorithm used to update model parameters during the training process. In the context of deep learning models, optimizers are used to minimize the loss function through incremental changes in model parameters.

Types of Optimizer:

- **Gradient Descent**: Gradient Descent is a basic algorithm used in optimization. It works by calculating the gradient of the loss function and updating the model parameters in the opposite direction to the gradient to minimize the loss.
  - **Stochastic Gradient Descent (SGD)**: Only uses one data sample to update parameters in each iteration.
  - **Batch Gradient Descent**: Using all data for each iteration, more stable but slower.
- **Momentum**: Momentum is a technique used to speed up the convergence process by preserving the "momentum" of previous gradients, preventing them from moving too slowly.
- **Adagrad**: Adagrad sets the update step for each parameter based on how often the feature occurs in the data. The more frequently a feature occurs, the smaller the update step applied.
- **RMSprop**: RMSprop is a modification of Adagrad that addresses the problem of too small update steps for frequently occurring features.
- **Adam (Adaptive Moment Estimation)**: Adam is one of the most popular optimizers in deep learning. Adam combines the advantages of momentum and RMSprop, by optimizing the update step based on estimates of the mean and variance of the gradient.

Summary

- **Cross-Validation** is used to evaluate models more robustly by dividing the data into several parts and testing the model on each part.
- **Hyperparameter Tuning** helps in selecting the best combination of parameters to maximize model performance.
- **Optimizer** is used to update model parameters during training to achieve faster convergence and better results.