

Model Evaluation

Why?

- To determine the predictive power of the model
- To analyze the performance of the model

Evaluation Metrics

- Confusion Matrix
- Accuracy
 - Ratio of the number of correct predictions to the total number of predictions
 - $Accuracy = (TP+TN)/(TP+TN+FP+FN)$
 - Drawback: cannot perform well on an imbalanced dataset
- Precision
 - Ratio of true positives to the summation of true positives and false positives
 - $Precision = TP/(TP+FP)$
- Recall
 - ratio of true positives to the summation of true positives and false negatives
 - $Recall = TP/(TP+FN)$
- F1-Score
 - Harmonic mean of precision and recall
 - The goal of the F1 score is to combine precision and recall
 - $F1\ score = (2 \times Precision \times Recall)/(Precision + Recall)$

Cross-validation

- What is it?
 - Technique used in ML to evaluate the performance of a model on unseen data
 - Divide the data into multiple folds
 - Use one of the folds as a validation set & the remaining folds for training
 - This process is repeated multiple times - each time, different fold for validation
 - Purpose: To prevent overfitting
 - Different types of cross validation
 - K-Fold Cross-Validation, Stratified K-Fold Cross-Validation, Leave One Out Cross Validation, Holdout Validation
- K-Fold Cross-Validation
 - The dataset is split into k equally sized folds
 - For each fold
 - The model is trained on k-1 folds
 - The remaining fold is used for validation
 - Compute Metrics
 - Compute the desired metrics (e.g., accuracy) on the validation fold
 - Repeat for all k folds
 - Average the metrics across all folds to get a final evaluation score

Hands-on