

FMST

Model Performance Measures, ML
Pipeline and Hyperparameter Tuning
(Week 2)

Topics covered in Week 3

- Performance measures
- ROC-AUC
- Concept of Pipeline
- Building a Pipeline
- Performance on train vs test data
- Hyperparameter tuning
- Grid Search and Random Search
- Hands-on Exercises

Session Agenda

- Quick recap of classification metrics
- Use of AUC and ROC
- Use of a pipeline object
- The Train, Validation and Test sets
- Hyperparameter tuning – GridSearchCV and RandomSearchCV
- Case Study
- Questions

Classification Metrics

- Sensitivity/Recall/True positive rate:

$$\frac{tp}{tp + fn}$$

- Specificity/True negative rate:

$$\frac{tn}{tn + fp}$$

- Precision:

$$\frac{tp}{tp + fp}$$

- F1 Score:

$$\frac{2 \times precision \times recall}{precision + recall}$$

- Accuracy:

$$\frac{tp + tn}{tp + tn + fp + fn}$$

Classification metrics

For the given Confusion matrix, what is the F1 score?

65.4%

Confusion Matrix - Test Data

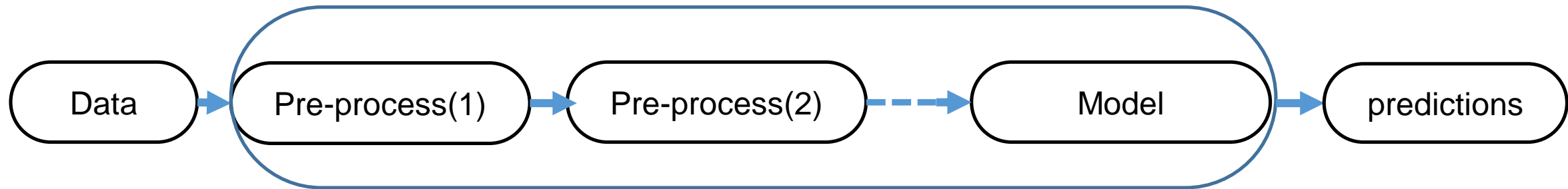
Actual label	Non_diabetic	Diabetic
	TN = 85	FP = 43
Diabetic	FN = 12	TP = 52
Predicted label		

ROC and AUC

- ROC (Receiver Operating Characteristics) is a probability curve with True positive rate in the vertical axis and False positive rate on the horizontal axis for different threshold values
- AUC is the area under the ROC curve

Need for a Pipeline

- Streamlines the process of transforming data, training an estimator and using it for prediction



Train, Validation and Test sets

- It is a general practice to split our data into three sets
- The train set
 - The data that we use to train the model
- The validation set
 - The data that we use to 'validate' a model
 - Any hyper-parameter tuning that is done, is based on the performance of the model on the validation set
- The test set
 - The data that is used to simulate real unseen data

- Always tune the model based on the performance on the validation set, once the model is trained on the Train set
- Never fine-tune a model based on its performance on the test set
- Test set is meant to aid in assessing a model's performance in production before the model hits production

Hyper-parameter tuning

- As opposed to parameters (like the ones in linear regression slope and constant term) which change based on the data for a given parametric model, hyper-parameters are preset values even before a non-parametric model gets trained on the data
- Parameters change during the training process
- Hyper-parameters are preset and do not change while training
- The process of setting the right hyper-parameters to get max performance out of a given model, is called Hyper-parameter Tuning

Grid Search and Random Search

- Both are the two most common methods of choosing the right hyper-parameters
- In Grid search, each and every combination of hyper-parameters tested before selecting the 'best' combination of hyper-parameters
- In Random search, only a subset of combinations can be tested before selecting the 'best' combination of hyper-parameters
- We use Random Search when the parameter grid is fairly large and we want to save on processing time
- [GridSearchCV](#) and [RandomizedSearchCV](#) are included in the sklearn library to perform the same over a parameter grid, that is passed as an argument to the functions along with the estimator

Case study

Q and A



Happy Learning