Hyperparameter Tunning, Cross-Validation Evaluation Matrix

Lecture 11

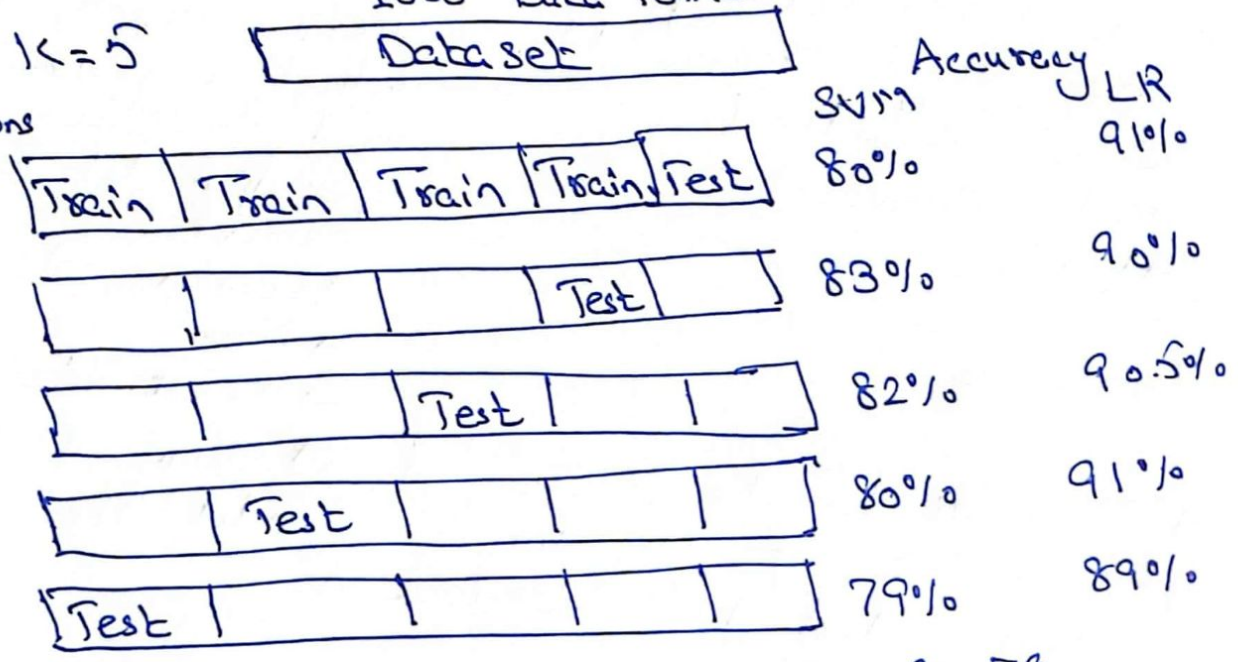
Cross-Validation (CV)

K-Fold CV

→ We split the dataset into "K" numbers of folds (subsets). One fold is used for testing & and remaining are used for training the model. Each time a different fold is used as the test data.

→ Example (Student Exam Evaluation).

→ K=5



SVM Mean Accuracy = $\frac{80 + 83 + 82 + 80 + 79}{5}$

LR " " = $\frac{91 + 90 + 90.5 + 91 + 89}{5}$

→ Better alternative for train-test split when dataset is small.

CamScanner

→ Trade off betw time & model accuracy (large DS)
→ Better for multiclass classification problems.
→ More reliable.
→ Useful for model selection.

Hyperparameter Tunning (HPT)

There are two types of parameters in AI/ML.

Parameters

Model Parameters

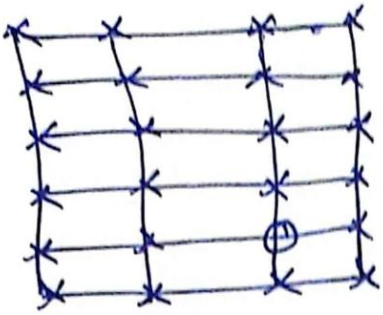
These model's parameters can be determined by training through training data. These are considered as internal parameters.
→ Weight
→ Bias

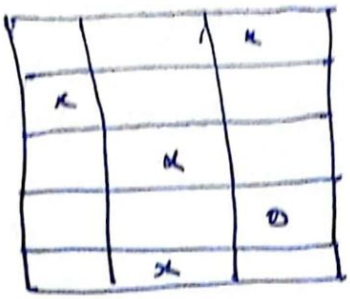
$Y = wx + b$

Hyperparameters

are those whose values control the learning process. These are adjustable parameters used to obtain an optimal model. External parameters
→ Learning rate
→ Number of Epochs
→ n-estimators

→ HPT refers to the process of choosing the optimum set of HP for ML model. The process is also called HP Optimization.
→ HPT has two types





Grid Search CV

Example

Support Vector Classifier

C: [1, 5, 10]

kernel: ('linear, poly, rbf, sigmoid)

C-Value is how much you can tolerate misclassifi higher the value, the lower the misclassification.

→ Objective is for which C-value & kernel value, we get the highest accuracy.

Model Selection (MS)

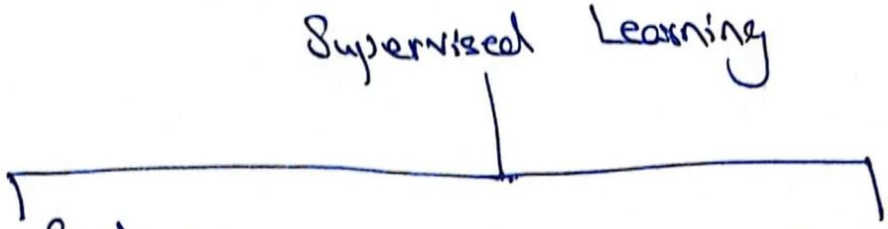
→ MS in ML is the process of choosing the best suited model for particular problem. It depends on several factors such as dataset, task, nature of model etc.

→ Two factor to be considered
1). Logical Reason to select a model.
2). Comparing the performance of the model.

→ Logical reasons to be considered.
1). Types of data
a). images & Videos — CNN
b). Text data or speech data — RNN
c). Numerical data (SVM, LR, DT etc)
2). Based on the task
a). Classification (SVM, LR, DT etc)
b). Regression tasks (LR, RF, MLR)
c). Clustering (k-Means, Hierarchical etc)

Accuracy Score & Confusion Matrix

→ Types of supervised Learning.



Classification
(Predicting a class or distinct value)
eg. Red/Green, T/F

Evaluation Metric: Accuracy

Regression
is about predicting quantity or contains value eg. salary, age, Price.

Evaluation Metric:
Mean Absolute Error.

→ In classification, Accuracy score is the ratio of number of correct predictions to the total number of input data.

$$\text{Accuracy Score} = \frac{\text{No of Correct Predictions}}{\text{Total No. of data points}} \times 100\%$$

$$= \frac{128}{159} = 85.3\%$$

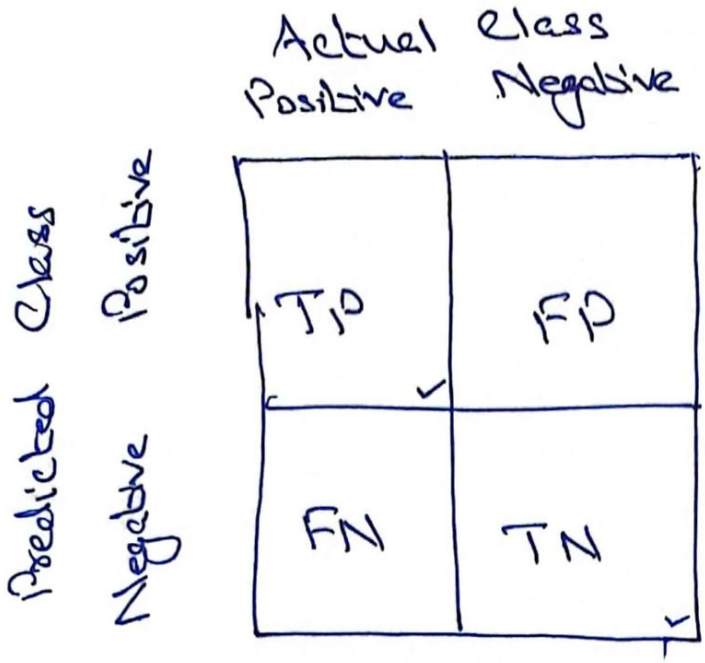
→ Accuracy Score is not reliable when the dataset has an uneven distribution of classes.

→ Eg.
No of Dog images = 800
No of images as dog = 1000
No of cat images = 200
Accuracy = 80%

Now
Dog images = 200
No of images as Dogs = 400
Cat images = 200
Accuracy = 50%

→ Confusion Matrix is used for evaluating the performance of a classification model. It gives more info than the accuracy.

CamScanner



TP + TN = Correct Predictions

FP + FN = Wrong Predictions

→ Heat Map

	1	2	3	4	5
1	30	5	0	2	1
2	2	85	0	3	1
3	0	1	40	0	1
4	1	0	0	30	1
5	3	1	0	1	32

Precision

$$\text{Precision} = \frac{TP}{TP + FP}$$

→ How many are actually positive, out of all predicted positives.

→ Measures the error caused by false positives (Face Authentication).

CamScanner

Recall

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{\text{True Positives}}{\text{Total Actual Positives}}$$

→ It is important where false positive predictions are critical. (Cancer diagnosis).

F1 Score

→ is an important evaluation metric for binary classification that combines Precision & Recall. It is the harmonic mean of Precision & Recall.

→ Very helpful when a dataset has imbalanced classes.

$$= 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

CamScanner