Learning

A process that leads to change, which occurs as a result of experience and increases the potential for performance and future learning

Learning is a process of acquiring new understanding, knowledge, behavior, skills, values, attitudes, and preferences

If AA=5, BB=6, AAA=50, BBB=60, AAAA=5000, BBBB=6000, what will be the value of AAAAA?

AAAAA = 5000

Training Data: AA=5, BB=6, AAA=50, BBB=60, AAAA=5000, BBBB=6000 Testing data: AAAAA

What about AABB?

Can I find it? (Yes/No)?

No. As my machine does not know how to solve it.

Training:
Cat (‘*ae*’ pronunciation)
Pot (‘*aw*’ pronunciation)
Pat (‘*ae*’ pronunciation)
Tat (‘*ae*’ pronunciation)
Cot (‘*aw*’ pronunciation)
Testing: Not?

Testing: Not? (‘*aw*’ pronunciation)

Testing: Check?

Testing: Check? (Not yet trained)

Can a human recognize a person’s face to whom he has met several times?

Can a human recognize a person’s voice with whom he earlier had conversations?

A computer program is said to learn from **experience** E with respect to some class of **tasks** T and **performance** measure P, if its performance at tasks in T, as measured by P, improved with experience E [*Tom Mitchell 1998*]

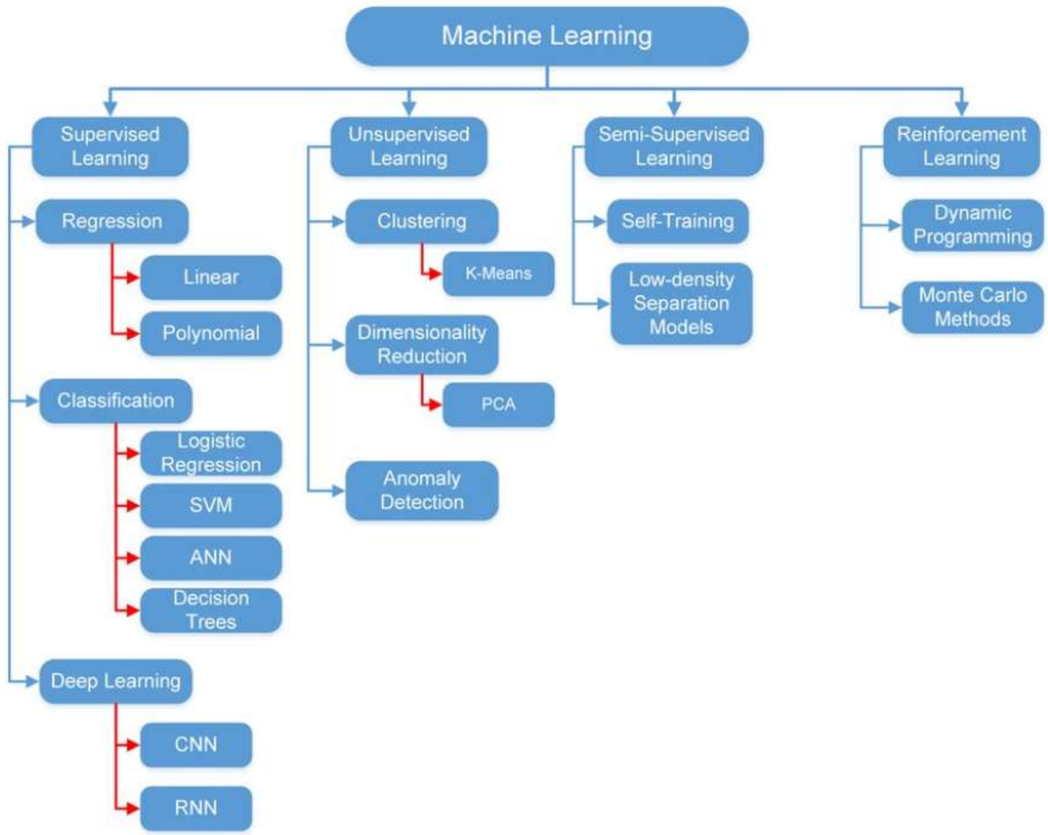
Question: Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is T, P and E?

(a) Classifying emails as spam or not.
(b) Watching you label emails as spam or not.
(c) The number (or fraction) of emails correctly classified as spam / not-spam.
(d) None

(a) T
(b) E
(c) P

Machine Learning

www.hbpatel.in



Source: https://www.researchgate.net/figure/Different-Machine-Learning-Categories-and-Algorithms_fig6_347096492

Machine Learning

www.hbpatel.in

Supervised

It maps an Input (X) to an Output (Y) based on example I/O pairs
Y = F(X)

```
X: 2, 3, 4, 5
Y: 4, 5, 6, 7
```

```
Y = F(X)
Y = X + 2
```

```
For X=102, Y=?
```

- Training Dataset
- Mapping Function
- Testing

Classification

Learning a ***discrete*** valued function
E.g. Various sections in D-Mart

A Predictive modeling problem where a class label is predicted for a given example of input data.

Binary

Spam or not
Buy or not

Multi-class

Face Recognition
Fruit Recognition

Multi-label

Document Classification
(E.g. Keywords)

Imbalanced

- Fraud
- Outlier Detection
- Medical diagnosis test

Regression

Learning a ***continuous*** valued function
E.g. Continuous mathematical series

Unsupervised

Reinforcement

Machine Learning

www.hbpatel.in

Supervised

Unsupervised

Reinforcement

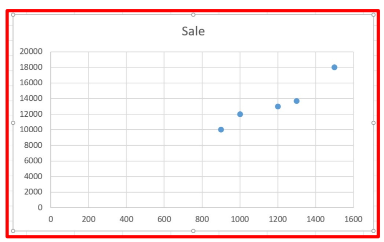
Classification

Regression

Learning a continuous valued function
E.g. Continuous mathematical series

- Determining market trend
- House price prediction

Advertisement Budget	Sale
900	10000
1200	13000
1500	18000
1000	12000
1300	13700
2000	?



Exercise

Exercise: You are given temperature, humidity, pollution (etc.) data for last 2 years.
Now, you are supposed to specify the category under which the answer falls?

Question 1: Is it going to be a hot day or cold day today?

Answer 1: Binary Classification

Question 2: In which zone (Green, Yellow, Red), today's pollution would fall?

Answer 2: Multiclass Classification

Question 3: What would be the temperature today?

Answer 3: Regression

Machine Learning

www.hbpatel.in

Supervised

Unsupervised

Reinforcement

Supervise model is not needed. Instead, it allows the model to work on its own to discover patterns and information that was previously undetected. It mainly deals with unlabelled data trends.

Clustering

A way of grouping the data points into different clusters, consisting of similar data points. The objects with the possible similarities remain in a group that has less or no similarities with another group. (E.g. Grouping coins based on their size, appearance, weight, color etc. without their prior knowledge)

- Partitioning
- Density-based
- Distribution Model-based
- Hierarchical
- Fuzzy

Machine Learning

www.hbpatel.in

Supervised

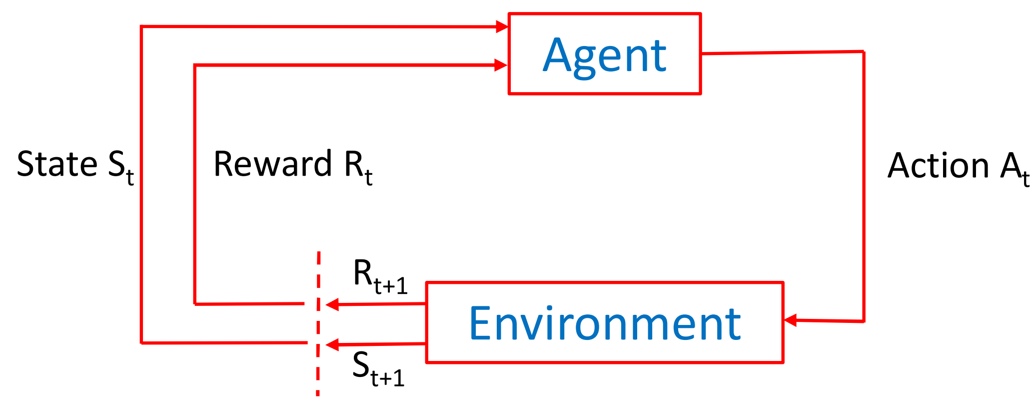
(**Input**, **Correct Output**)

Unsupervised

(**Input**, ?)

Reinforcement

(**Input**, *Some* **Output**, We have to grade output)



Evaluating Estimator Performance

Learning the parameter of a prediction function and testing it on the same data is a methodological mistake.

For the input values of X = [2,3,4,5], output Y=[3, 5, 7, 9], what if I try to find the value of Y for X=2? (Obviously it would give 100% accuracy as it belongs to the training dataset)

A model that would just repeat the labels of the samples, that it has just seen, would have a perfect score but would fail to predict anything useful on yet-unseen data. This situation is called overfitting.

For the input values of X = [2,3,4,5], output Y=[3, 5, 7, 9], what if I try to find the value of Y for X=2? (It should give the value of Y=3, but think what if it does not?)
To avoid this problem, we would not use the complete data as training data (may be 90% is used for training) and other remaining (10%) data can be used for testing.

Train-Test Split

www.hbpatel.in

Training Set	Testing Set

Total Dataset

x	y
2	5
3	8
4	11
5	14
6	17
7	?
8	?

Training Set `Train_x, Train_y`

Testing Set `Test_x, Test_y`

Train-Test Split: Cross Validation

www.hbpatel.in

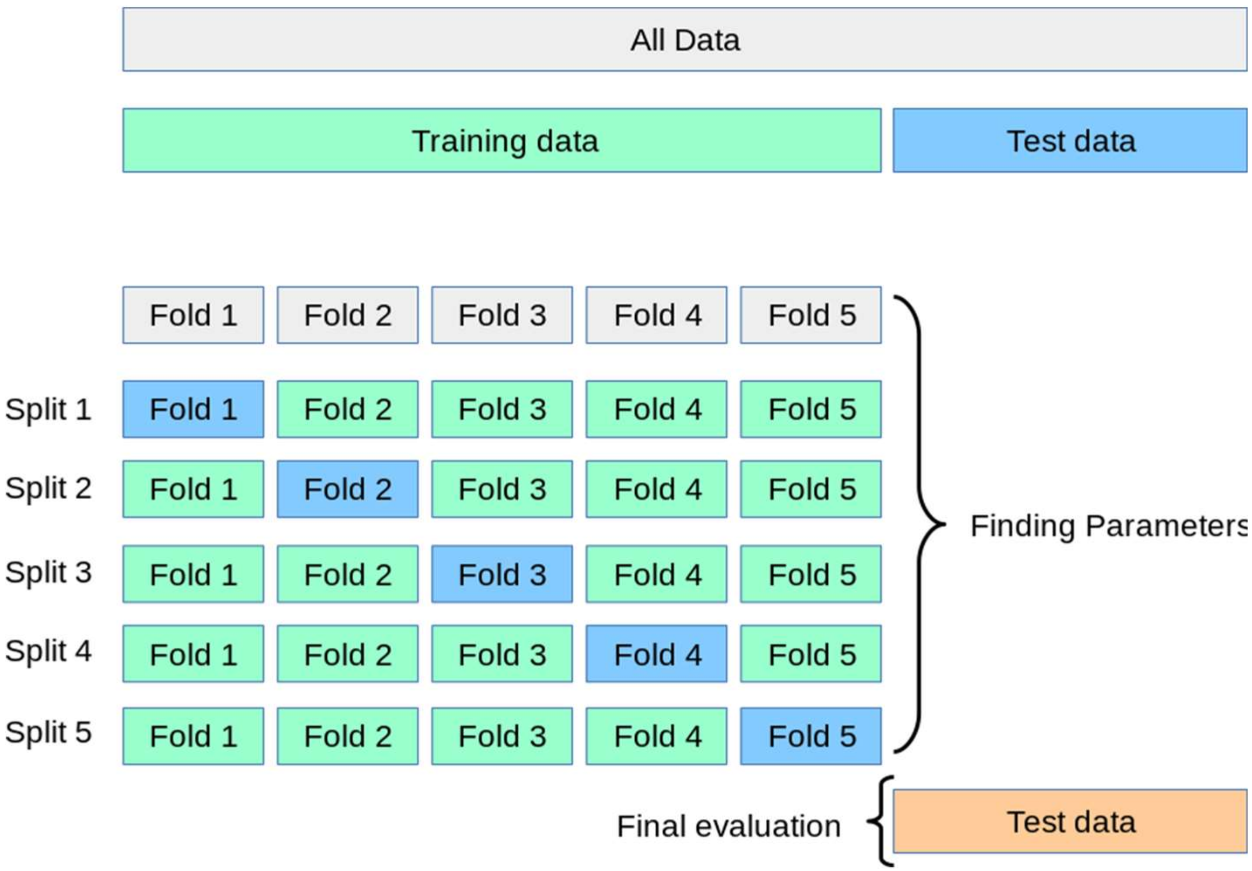


Image Source: https://scikit-learn.org/stable/modules/cross_validation.html

Model Evaluation: Classification

www.hbpatel.in

- Classification accuracy
- Confusion matrix
- Precision recall
- F1 score
- Sensitivity and Specificity

Recall = $\frac{TP}{TP+FN}$

Recall measures how good our model is at correctly predicting positive classes.

Precision = $\frac{TP}{TP+F}$

Precision measures how good our model is when the prediction is positive.

Accuracy = $\frac{Number\ of\ currect\ predictions}{Total\ number\ of\ predictions}$

Confusion Matrix

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

1. True Positive (TP): Predicting positive class as positive (OK)
2. False Positive (FP): Predicting negative class as positive (Not OK) | Type I Error
3. False Negative (FN): Predicting positive class as negative (Not OK) | Type II Error
4. True Negative (TN): Predicting negative class as negative (OK)

Model Evaluation: Classification

www.hbpatel.in

- Classification accuracy
- Confusion matrix
- Precision recall
- F1 score
- Sensitivity and Specificity

F1 Score $=2\text{x}\ \frac{precision\ x\ recall}{precision+recall} = \frac{TP}{TP+\frac{1}{2}(FP+F\quad)}$

(Weighted average of precision and recall)

F1 Score is a more useful measure than accuracy for problems with uneven class distribution because it takes into account both false positive and false negative. The best value of F1 score is 1 and worst is 0.

Sensitivity: It is also known as True Positive Rate (TPR) and it is same as recall. It measures the proportion of positive class that is correctly predicted as positive.

Specificity: It is similar to sensitivity but focused on negative class. It measures the proportion of negative class that is correctly predicted as negative.

Sensitivity and Specificity

www.hbpatel.in