

Reflection

1. What is an Algorithm?

An algorithm is a process or set of rules to follow in order to solve/calculate a problem via a computer

2. What is Machine Learning

Machine learning is a branch of artificial intelligence and computer science that focuses on the use of data and algorithms to imitate the way humans learn and gradually improving its learning/accuracy.

3. Different types of Learning

Learning Problems

- 1. Supervised Learning
- 2. Unsupervised Learning
- 3. Reinforcement Learning

Hybrid Learning Problems

- 4. Semi-Supervised Learning
- 5. Self-Supervised Learning
- 6. Multi-Instance Learning

Statistical Inference

- 7. Inductive Learning
- 8. Deductive Inference
- 9. Transductive Learning

Learning Techniques

- 10. Multi-Task Learning
- 11. Active Learning
- 12. Online Learning
- 13. Transfer Learning
- 14. Ensemble Learning

4. The maths behind machine learning

- Linear Algebra for Machine Learning
- Multivariate Calculus for Machine Learning
- Probability for Machine Learning
- Statistics for Machine Learning

5. Algorithms used in Machine Learning

Model	Applications
Logistic Regression	Price prediction
Linear Regression	evaluate trends and estimates/forecasts
Decision Tree	Fraud Detection
SVM (Support Vector Machine)	Speech Recognition.
Naive Bayes	Spam and noise filtering
KNN (K-Nearest Neighbours)	Recommended systems
K-Means	Segmentation
Random Forest	Fraud detection
Fully connected networks	Classification
Convolutional neural networks	Image processing
Recurrent neural networks	Voice recognition
Reinforcement Learning	Learning by trial and error
Generative Models	Image creation

6. Process of training a Machine Learning Model

1. Collect data
2. Prepare data
3. Choose a model
4. Train your machine model
5. Evaluation
6. Parameter tuning
7. Predictive or Inference

HLT

Linear Regression

Linear Regression, a supervised model is used to estimate real values (cost of houses, number of calls, total sales, etc.) based on continuous variables.

Logistic Regression

Logistic Regression is a classification, not a regression algorithm. A supervised model used to estimate discrete values (Binary values like 0/1, yes/no, true/false) based on a given set of independent variables.

Decision Tree

Decision Tree is a type of supervised learning algorithm that is mostly used for classification problems. Surprisingly, it works for both categorical and continuous dependent variables. In this algorithm, the population is split into two or more homogeneous sets. This is done based on the most significant attributes/ independent variables to make as distinct groups as possible.

SVM (Support Vector Machine)

It is a supervised classification method. In this algorithm, we plot each data item as a point in n-dimensional space (where n is a number of features you have) with the value of each feature being the value of a particular coordinate.

Naive Bayes

It is a classification technique based on Bayes' theorem with an assumption of independence between predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

Naive Bayesian is a form of supervised learning model which is easy to build and particularly useful for very large data sets. Along with simplicity, Naive Bayes is known to outperform even highly sophisticated classification methods.

KNN (K-Nearest Neighbours)

It can be used for both classification and regression problems. However, it is more widely used in classification problems in the industry. K nearest neighbours, a supervised model is a simple algorithm that stores all available cases and classifies new cases by a majority vote of its k neighbours. The case is assigned to the class is most common amongst its K nearest neighbours measured by a distance function.

K-Means

It is a type of unsupervised algorithm which solves the clustering problem. Its procedure follows a simple and easy way to classify a given data set through a certain number of clusters. (Data points inside a cluster are homogeneous and heterogeneous to peer groups).

Random Forest

Random Forest is a supervised model and a term for a set of decision trees. In Random Forest, a collection of decision trees also known as forests are used to classify a new object based on attributes, each tree gives a classification, and we say the tree “votes” for that class. The forest chooses the classification having the most votes (over all the trees in the forest).

