

# Machine Learning

Raj Kumar Mishra

21111041

rajkumar96369427@gmail.com

August 6, 2023

## Contents

<b>1</b>	<b>Supervised learning</b>	<b>1</b>
1.1	Algorithms used in supervised learning . . . . .	2
1.1.1	Linear Regression . . . . .	2
1.1.2	Logistic Regression . . . . .	2
1.1.3	Decision Trees . . . . .	2
1.1.4	Random Forests . . . . .	2
1.1.5	Support Vector Machines . . . . .	3
<b>2</b>	<b>Unsupervised learning</b>	<b>3</b>
2.1	Algorithms used in unsupervised learning . . . . .	3
2.1.1	K-Means Clustering . . . . .	3
2.1.2	Hierarchical Clustering . . . . .	3
2.1.3	Principal Component Analysis . . . . .	3
2.1.4	Gaussian Mixture Models . . . . .	3
2.1.5	Autoencoders . . . . .	4
<b>3</b>	<b>Reinforcement learning</b>	<b>4</b>
3.1	Algorithms used in Reinforcement learning . . . . .	4
3.1.1	Q-Learning . . . . .	4
3.1.2	Deep Q Networks . . . . .	4
3.1.3	Policy Gradient Methods . . . . .	4
3.1.4	Policy Optimization . . . . .	4
3.1.5	Actor-Critic Methods . . . . .	4

# 1 Supervised learning

Supervised learning is a machine learning technique that is widely used in various fields such as finance, healthcare, marketing, and more. It is a form of machine learning in which the algorithm is trained on labeled data to make predictions or decisions based on the data inputs. In supervised learning, the algorithm learns a mapping between the input and output data. This mapping is learned from a labeled dataset, which consists of pairs of input and output data. The algorithm tries to learn the relationships between the input and output data so that it can make accurate predictions or decisions based on data inputs. In supervised learning, the algorithm learns a mapping between the input and output data. This mapping is learned from a labeled dataset, which consists of pairs of input and output data so that it can make accurate predictions on new unseen data. The labeled dataset used in supervised learning consists of input features and corresponding output labels. The input features are the attributes or characteristics of data that are used to make predictions, while the output labels are the desired outcomes or targets that the algorithm tries to predict.

## 1.1 Algorithms used in supervised learning

### 1.1.1 Linear Regression

Linear regression is a type of regression algorithm that is used to predict a continuous output value. It is one of the simplest and most widely used algorithms in supervised learning. In linear regression, the algorithm tries to find a linear relationship between the input features and the output value. The output value is predicted based on the weighted sum of the input features.

### 1.1.2 Logistic Regression

Logistic regression is a type of classification algorithm that is used to predict a binary output variable. It is commonly used in machine learning applications where the output variable is either true or false, such as in fraud detection or spam filtering. In logistic regression, the algorithm tries to find a linear relationship between the input features and the output variable. The output variable is then transformed using a logistic function to produce a probability value between 0 and 1.

### **1.1.3 Decision Trees**

A decision tree is a type of algorithm that is used for both classification and regression tasks. It is a tree-like structure that is used to model decisions and their possible consequences. Each internal node in the tree represents a decision, while each leaf node represents a possible outcome. Decision trees can be used to model complex relationships between input features and output variables.

### **1.1.4 Random Forests**

Random forests are an ensemble learning technique that is used for both classification and regression tasks. They are made up of multiple decision trees that work together to make predictions. Each tree in the forest is trained on a different subset of the input features and data. The final prediction is made by aggregating the predictions of all the trees in the forest.

### **1.1.5 Support Vector Machines**

This algorithm finds a hyperplane that best separates different classes while maximizing the margin between them.

## **2 Unsupervised learning**

Unsupervised machine learning analyzes and clusters unlabeled datasets using machine learning algorithm. The training model has only input parameter values and discovers the groups or patterns on its own. Once subscribed they are provided a membership card and the mall has complete information about the customer. Now using this data and unsupervised learning techniques, the mall can easily group clients based on the parameters we are feeding in.

### **2.1 Algorithms used in unsupervised learning**

#### **2.1.1 K-Means Clustering**

Divides data points into clusters based on similarity, with each cluster represented by a centroid.

### **2.1.2 Hierarchical Clustering**

Builds a tree-like structure of clusters, where similar data points are grouped together based on distance metrics.

### **2.1.3 Principal Component Analysis**

Reduces the dimension of data while preserving as much variability as possible by projecting it onto a lower-dimensional subspace.

### **2.1.4 Gaussian Mixture Models**

A probabilistic model that represents data as a mixture of several Gaussian distributions, allowing it to capture complex data distributions.

### **2.1.5 Autoencoders**

Neural network architectures used for dimensionality reduction and feature learning by training to reconstruct the input data from a compressed representation.

## **3 Reinforcement learning**

Reinforcement learning involves an agent learning to make decisions by interacting with an environment to maximize a reward signal.

### **3.1 Algorithms used in Reinforcement learning**

#### **3.1.1 Q-Learning**

An off-policy algorithm where an agent learns to take actions to maximize a cumulative reward by updating Q-values based on experience.

#### **3.1.2 Deep Q Networks**

Combines Q-learning with deep neural networks, enabling it to handle high-dimensional state spaces in complex environments.

### **3.1.3 Policy Gradient Methods**

Directly learn a policy that maps states to actions by optimizing the expected cumulative reward.

### **3.1.4 Policy Optimization**

An on-policy algorithm that optimizes the policy in a more stable manner by updating it.

### **3.1.5 Actor-Critic Methods**

Combines value-based and policy-based methods by using two networks an actor network and a critic network.