Machine Learning - Week 1

ZVENC

December 2024

1 Introduction

There are many problems that have been solved using rule-based algorithms, such as finding the shortest path between two points, sorting lists of items, etc. There are problems that cannot be solved this way because they are simply too complex; this is a problem area where machine learning excels. In simple terms, machine learning is a field concerned with teaching computers how to learn to do certain tasks. Machine learning is heavily relevant today, being used in various fields such as healthcare, scientific research, social media, manufacturing, etc.

1.1 Types of Machine Learning Approaches

Machine Learning algorithms can be categorized into two main types:

1. Supervised Learning
2. Unsupervised Learning

While there are other types of machine learning, these two are the most relevant at this stage.

1.1.1 Supervised Learning ($x \to y$)

This approach to machine learning involves giving the computer a number of examples to learn from in the form of input-output pairs. The output represents what is expected given a particular input. The computer then learns from these input-output pairs and tries to guess the correct output for an input that it has never seen before.

For example let's say you had a plot of house prices against size for about 15 houses. The goal would be to get the computer to estimate the price of a house based on its size, even for houses not explicity represented in the dataset. This is called a *'regression problem'*, and the relationships could either be linear, polynomial, etc.

Another example is breast cancer detection. A model can learn to predict whether a tumor is cancerous based on its size. It can be provided with a plot of patient age against tumor size with the malignant and benign tumors marked on the plot. The model might the attempt to create a boundary to identify malignant tumors. In this example, the model only has to guess between two possible output classes, 'benign' or 'malignant'. This is what's known as a *'classification problem'*. Unlike regression, classification requires the model to predict the correct value from a finite set of classes, not necessarily just two, like in the example. A model could also be designed to detect different classes of malignant tumors, in which case the classes could be 'benign', 'malignant-type-1', and 'malignant-type-2'.

1.1.2 Unsupervised Learning (x)

In supervised learning, a model learns from data labeled with 'correct answers'. In contrast, unsupervised learning involved providing the model with unlabeled data and tasking it with finding any patterns or structures hidden within the data. Let's say we have another patient age vs tumor size plot except the points aren't labeled. An model might try to group the points into categories or *'clusters'*. This is type of supervised learning algorithm called a *'clustering algorithm'*. Google News employs this algorithm to group related news stories. Achieving this with supervised learning is unfeasible. Other unsupervised learning algorithms include *'anomaly detection'* and *'dimensionality reduction'*.

1.2 Terminology

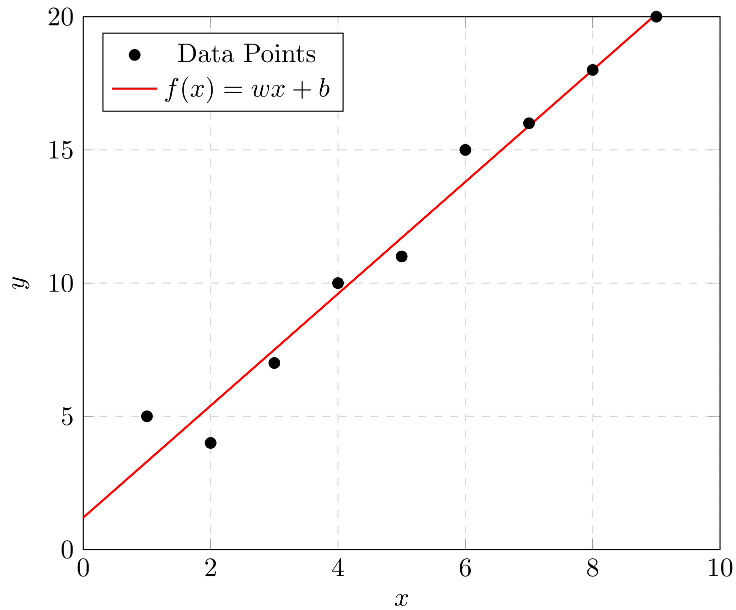
1. Training Set: Data used to train the model
2. Input Variable (x): The input provided to the model in the training set. Also called *'feature'* or *'input feature'*.
3. Output Variable (y): The output correct output provided to the model in the training set. Also called the *'target'* variable.
4. Training Example ($x^{(i)}, y^{(i)}$): This is an input output pair from the training set. The superscript 'i' represents the position of the training example in the training set, for instance, in a table.
5. Total Number of Training Examples (m): I believe this is self-explanatory.

1.3 Linear Regression Model

The training set, which contains the input features and output targets, is fed into the supervised learning algorithm. The algorithm then produces a function (historically called a hypothesis) that it then uses to estimate an output for any given input. That function, also called a model, takes in a feature, x, and then produces and estimate, $\hat{y}$. Note the difference between y and $\hat{y}$; y is the target/true value in the training set, while $\hat{y}$ is an estimate of y. A linear regression model takes on the form:

$$f_{w,b}(x) = wx + b$$

For the sake of convenience, $f_{w,b}(x)$ will be written as $f(x)$. If the features and targets were plot on a graph, the algorithm would generate a best-fit line for the plot which would be represented by that equation.



This is one-variable or *univariate* linear regression, which means there's a single feature for each data point. Using the house example, it is possible to take other factors into consideration rather than just the size of the house such as number of bedrooms, location, etc.