

Math 348 Machine Learning in Python

Cameron P. Dart

February 8, 2016

1 Introduction

There are three main types of machine learning algorithms.

- I. Supervised
- II. Unsupervised
- III. Semi-Supervised

In order to gain a more deep understanding of these types of artificial intelligence algorithms, I would like to research some commonly used supervised, unsupervised, and semi-supervised algorithms. While researching these algorithms, I will create my own implementations of each algorithm in Python, describe and prove the mathematical reasoning, and analyze the space-time complexities of my own to open source implementations of these same algorithms in big O, θ , and Ω notation.

1.1 Machine Learning Definition

Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed. These types of algorithms are deeply rooted in mathematical functions and theory. The brute computational powers of a computer greatly exceeds that of a human. Allowing them to process multitudes more data than humanly possible. Machine learning algorithms are being used by a wide variety companies along the likes of from Google, Tesla, Goldman Sach's, and Amazon.

2 Supervised Learning Algorithms

2.1 Definition

Supervised Learning is the task of inferring a function from labeled training data. The training data consist of a set of training examples. In supervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal). A supervised

learning algorithm analyzes the training data and produces an inferred function, which can be used for mapping new examples.

2.2 Example: Linear Regression

2.2.1 Definition

In statistics, a **linear regression** is an approach for modeling the relationship between a scalar dependent variable y and one or more explanatory variables (or independent variables) denoted X . The case of one explanatory variable is called simple linear regression.

2.2.2 Mathematical Description

First, suppose there is a set S that contains n data points such that

$$S = \{(x_i, y_i), i = 1, 2, \dots, n\}$$

The function that describes x_i, y_i is as follows.

let α be the y-intercept and β the slope of the line

$$y_i = \alpha + \beta x_i + \epsilon_i \tag{1}$$

In order to find the more general equation

$$y = \alpha + \beta x \tag{2}$$

2.2.3 Example Use Cases

I. You are a home owner and you are trying to decide when you should best sell your home and for what price.

Let x be the time in years and y be the price of your house

3 Unsupervised Learning Algorithms

3.1 Definition

3.2 Example K-Means Clustering

3.2.1 Defintion

3.2.2 Mathematical Description

4 Semi-Supervised Learning Algorithms

4.1 Definition

4.2 Example: Transductive Support Vector Machine

4.2.1 Definition

4.2.2 Mathematical Description

References