

# Introduction to Machine Learning

Alan Zheng

September 2018

## 1 Introduction

Machine Learning is a subfield of computer science focusing on programming computers to learn from data without being explicitly programmed. Machine learning algorithms learn and recognize patterns in seen data (*training data*) and use these patterns to predict characteristics of unseen data.

## 2 Types of Machine Learning

Machine learning algorithms generally fall into one of the three categories: Supervised Learning, Unsupervised Learning, and Reinforcement Learning.

### 2.1 Supervised Learning

Supervised learning algorithms analyze known training data with labels, the characteristics of the data that we want to predict, to predict the labels of unseen data. For example, an e-mail spam filtering algorithm would analyze previously seen e-mails that are already labeled as being spam or non-spam to predict whether new, unseen e-mails are spam or non-spam. A supervised learning problem where the class labels are discrete (i.e. are made up of distinct categories), such as the spam filter example, is called a classification task. Regression is another type of supervised task where the predicted value is continuous (e.g. predicting a student's SAT score based on their GPA).

### 2.2 Unsupervised Learning

Unsupervised learning algorithms analyze unlabeled training data. One common unsupervised learning task is clustering, which creates different groups for data and categorizes similar data into each group. It seeks to determine how data is organized without labels on each data point. An example might be categorizing similar visitors of your website into different groups.

## 2.3 Reinforcement Learning

Reinforcement learning is a different subset of Machine Learning where the learning system (agent) can perform different actions and receives rewards or penalties in return and must learn the correct policy, which dictates which action the agent should take in a given situation, to get the most rewards over time.

## 3 Vocabulary

Here we have listed a few important words that we will use throughout the year.

- **Label** - The thing we're trying to predict. Ex: If we're trying to predict what kind of animal is in a picture, the label would be the type of animal in the picture.
- **Classification** - Taking each instance and assigning it to a particular category. Ex: Determining if tumors are benign or malignant by looking at MRI scans.
- **Regression** - Instead of having discrete classes, like classification, the "class" to be predicted is made up of continuous numerical values. Ex: Predicting house prices based on square footage, number of rooms, etc.
- **Clustering** - For data without pre-labeled classes, clustering is the act of grouping similar data points together. A form of unsupervised learning. Ex: Clustering U.S. households for marketing data.
- **Training Data** - The initial set of data used to discover potentially predictive relationships. It's what your machine learning algorithm "trains" on and learns patterns from.
- **Testing Data** - Set of data used to assess the strength and utility of predictive relationship. Your machine learning algorithm does not see this data during training.
- **Error** - The difference between algorithm's prediction and ground-truth values.
- **Ground Truth** - Data that is known to be correct. A data-set's labels.
- **Features** - The attributes of the data that are used to make a prediction about the labels. Ex: In the house price example in the regression definition, the features would be the square footage, number of rooms, etc.
- **Feature space** - The  $n$ -dimensions in which the features live where  $n$  = the number of features. Typically, the larger the feature space, the more complex your algorithm will be.
- **Model** - The relationship between features and label. Training means "learning" this relationship based on examples.