Introduction of Machine Learning

Shuo Yang Shijie Geng

Computer Department of Hangzhou Dianzi University

First Group Meeting, April 2022

Table of Contents

Table of Contents

What is Machine Learning?

Machine learning is a subfield of computer science that evolved from the study of pattern recognition and computational learning theory in artificial intelligence. Machine learning explores the study and construction of algorithms that can learn from and make predictions on data. Such algorithms operate by building a model from example inputs in order to make data-driven predictions or decisions, rather than following strictly static program instructions.

Definition of machine learning

- The term machine learning was coined in 1959 by Arthur Samuel

Machine Learning:

Field of study that gives computers the ability to learn without being explicitly programmed.

Definition of machine learning

- The term machine learning was coined in 1959 by Arthur Samuel

Machine Learning:

Field of study that gives computers the ability to learn without being explicitly programmed.

- A widely quoted, more formal definition provided by Tom M. Mitchell

Machine Learning:

A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

An example of machine learning

Examples

Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is the task T, experience E, and performance P in this setting?

- T : Classifying emails as spam or not spam.

An example of machine learning

Examples

Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is the task T, experience E, and performance P in this setting?

- T : Classifying emails as spam or not spam.
- E : Watching you label emails as spam or not spam.

An example of machine learning

Examples

Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is the task T, experience E, and performance P in this setting?

- T : Classifying emails as spam or not spam.
- E : Watching you label emails as spam or not spam.
- P : The number (or fraction) of emails correctly classified as spam/not spam.

Table of Contents

Terminology related to machine learning

- example / instance
- feature / attribute
- attribute value , example/attribute space
- feature vector
- label
- training example , training set
- validation example
- test example
- hypothesis
- loss function

Some notations of machine learning

- x Scalar
- $\boldsymbol{x}$ Vector
- $\mathbf{x}$ Variable set
- $\mathbf{A}$ Matrix
- $\mathcal{X}$ Example Space
- $\mathbf{D}$ Data Set / Data Examples
- $\mathcal{D}$ Probability Distribution
- $\mathbf{H}$ Hypothesis Set
- $\mathcal{H}$ Hypothesis Space
- ...

Table of Contents

A Learning Comic

- : Experience: Training Data
- : What is to be learned: Target Function
- : Learning Algorithm: how to infer the target function from the experience
- : Evaluation: Test Data

Table of Contents

Supervised learning

Supervised learning

Given a set of data points$\{x^{(1)}, \cdots, x^{(m)}\}$ associated to a set of outcomes $\{y^{(1)}, \cdots, y^{(m)}\}$. we want to build a classifier that learns how to predict y from x.

The different types of predictive models are summed up in the table below:

	Regression	Classification
Outcome	Continuous	Class
Examples	Linear regression	Logistic regression, SVM, Naive Bayes

For instance,

- Regression: predict house price;
- Classification: classify Breast cancer;

Unsupervised Learing

The goal of unsupervised learning is to find hidden patterns in unlabeled data $\{x^{(1)}, \ldots, x^{(m)}\}$

- Google News Category
- Organize computing clusters
- Social network analysis
- Market segmentation
- Astronomical data analysis
- Cocktail party problem

Table of Contents

Relations with Other Fields

- Data Mining
- Artificial Intelligence
- Pattern Recognition
- Neurocomputing
- KDD
- ...

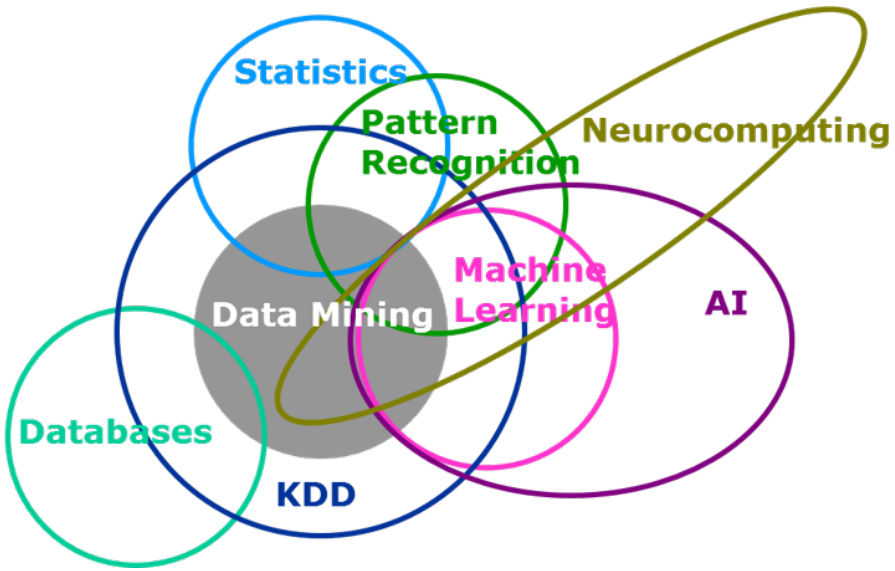

Figure 1: Relations with Other Fields