

richardwu.ca

CS 485/685 COURSE NOTES

MACHINE LEARNING: STATISTICAL AND COMPUTATIONAL FOUNDATIONS

SHAI BEN-DAVID • WINTER 2019 • UNIVERSITY OF WATERLOO

Last Revision: January 8, 2019

Table of Contents

1	January 8, 2019	1
1.1	What is machine learning?	1
1.2	Why do we need machine learning?	1
1.3	Types of machine learning	1

Abstract

These notes are intended as a resource for myself; past, present, or future students of this course, and anyone interested in the material. The goal is to provide an end-to-end resource that covers all material discussed in the course displayed in an organized manner. These notes are my interpretation and transcription of the content covered in lectures. The instructor has not verified or confirmed the accuracy of these notes, and any discrepancies, misunderstandings, typos, etc. as these notes relate to course's content is not the responsibility of the instructor. If you spot any errors or would like to contribute, please contact me directly.

1 January 8, 2019

1.1 What is machine learning?

In machine learning, we aim to construct a program that takes as input **experiences** and produces as output **expertise**, or what we have learned from the experience.

We can then apply the **expertise** to produce useful programs such as a spam filter.

An example of learning in nature is **bait shyness**: rats who become sick from eating poisoned bait will become more cautious of food of similar characteristic in the future. Since rats will become more cautious of bait in the future, a delayed poison mechanism (rat is poisoned only 2 days after consuming the bait) is necessary for effective bait by de-associating poison from the bait.

Another example is an experiment called **pigeon superstition** by Skinner (1947): pigeons are starved in a cage with various objects. At random intervals, food is dispersed to satiate the pigeons. Eventually, each pigeon develops a "superstition": they each associate one arbitrary behaviour (e.g. a specific object or a specific movement) that results in food being dispersed.

On the contrary, Garcia (1996) tried a similar experiment to bait shyness with rats where poisoned and un-poisoned bait were identical in characteristic. Whenever a rat approached poisoned bait, a stimulus (e.g. bell ringing, electric shock) was applied to the rat. Surprisingly, the rats did not associate the arbitrary stimulus to the poisoning. This is contrary to the pigeon superstition: this can be explained by evolution (future generations are those that can become aware of poisonous bait) and the fact that rats have **prior knowledge** that poisoning comes from the bait itself, not some arbitrary stimulus.

1.2 Why do we need machine learning?

We desire machines to perform learning because machines can **process lots of data** and are (generally) **fast**. We desire machines to *learn* because some tasks are simply too complex to hardcode in (e.g. image recognition). Some tasks we do not fully understand how to solve with hardcoded rules. Furthermore, learning allows adaptivity where the machine can constantly learn from new experiences and inputs.

1.3 Types of machine learning

Supervised and unsupervised Machine learning can be generally classified as either **supervised** or **unsupervised**.

Supervised learning takes labelled examples as experience and tries to re-produce these labels on future examples by learning rules. Spam detection may be supervised learning.

Unsupervised learning does not require labelled training data. Examples of unsupervised learning is outlier detection and clustering.

Semi-supervised learning takes as input both labelled and unlabelled data and sits between supervised and unsupervised.

Reinforcement learning also sits between supervised and unsupervised: the machine knows only the rules of the environment and takes actions until a reward (i.e. label) is produced. The machine then learns to label intermediary actions to the final reward produced in the episode (sequence of actions that resulted in the reward).

Passive and active We can also distinguish between **passive** and **active** learning: the former simply takes observed data whereas the latter involves actively performing experiments and interpreting the consequences of the experiments.

Teacher Machine learning can be guided by a “teacher” i.e. how the random sample used as input is generated. Teachers may be **indifferent**, **helpful** or **adversarial**. Helpful teachers produce hints and try to guide the program in the right direction whereas adversarial tries to fool the program.

Batch and online **Batch** learning is learning from a relatively large corpus of data before producing expertise. In contrast **online** learning requires the program to learn as experience is streamed and may result in more mistakes being made.