CSCC11: Machine Learning
Lecture Notes

Joshua Concon

University of Toronto Scarborough – Winter 2020

Instructor is Dr. David Fleet. If you find any problems in these notes, feel free to contact me at conconjoshua@gmail.com.

Contents

1 LEC 1: Tuesday, January 7, 2020

1.1 Machine Learning Definitions

AI View Automatic Learning

Software Engineering View Can be fine-tuned

Statistics View Machine Learning is fast Statistics

Machine Learning methods are broken into 2 phases

Training a model is learned from a collection of training data

Application the model is used to make decisions about some new test data

Types of Machine Learning

- Supervised Learning: The training data is labelled with the correct answers
  - Classification: Outputs are discrete labels
  - Regression: Outputs are real-valued
- Unsupervized Learning: Deriving patterns and structure from unlabelled data
- Reinforcement Learning: An agent seeks to learn optimal actions to take based on state of the world, and learn from the consequences of past actions

1.2 A Simple Problem

Consider a problem where the goal is to fit a 1D curve. There are many curves that may fit.

Using Machine Learning requires us to make certain choices:

1. How do we parameterize the curve? (linear, quadratic, sine...)

2. What criteria (objective function) do we use to judge the quality of the fit?

3. How long are we willing to wait for a solution (or can we use approximations)?

4. How do we prevent overfitting?

Overfitting is when a model fits the training data well, but performs poorly on test data.
Properties of the input that are not relevant to the task at hand are called **Noise**.