Lecture 1: Introduction to Supervised Learning

Adam Hawley

February 21, 2019

Contents

1 Organisation of this Section of ARIN **1**

2 What is Machine Learning? **2**

3 Types of Machine Learning **2**
3.1 Supervised Learning 2
3.2 Reinforcement Learning 2
3.3 Unsupervised Learning 2

4 Supervised Learning **2**
4.1 Issues 2
4.2 Learning (Generalisation) Bias 3

1 Organisation of this Section of ARIN

- 6 Lectures of this section
- 2 exercise problem sheets (formative):
  - Hand in solutions in groups of 3-4 to receive feedback
  - Solutions will be posted on the VLE
- Two Q&A sessions attendance voluntary.
- Recommended text: Tom Mitchell, "Machine Learning"

2 What is Machine Learning?

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E. [Mitchell 97]

3 Types of Machine Learning

3.1 Supervised Learning

This is where the learning agent receives training examples and corresponding labels provided by a supervisor. The goal is summarised as "Given a new example, what is its label?".

3.2 Reinforcement Learning

Training experience are state-action paris with the correspondingn numerical reward. The goal is summarised here as "Learn a behaviour that maximises cumulative reward.

3.3 Unsupervised Learning

Unstructured set of examples. Goal: discover patterns/structures in the data. Focus: clustering.

4 Supervised Learning

4.1 Issues

- Where to get training examples from? (e.g medical data can be difficult to acquire because of data privacy)
- How to represent examples?
- How to represent classification procedure (hypothesis)?
- Which learning method to use?
- How to evaluate the result?

4.2 Learning (Generalisation) Bias

Definition: Preference relation between legal hypotheses.

Hypothesis with zero error on training data is not necessarily the best (noise!). **Occam's razor**: the simpler hypothesis is the better one. No supervised learning without some generalisation (can be caused by language or learning bias). **Language bias** is bias that comes from your hypothesis representation. **Learning bias** comes from the machine learning algorithm.