COMP9417: Machine Learning and Data Mining

Update on aims and learning objectives

Session 1, 2003

1 Introduction

Each lecture in the course had a stated aim and a list of learning objectives. The purpose of this is to provide structure and direction for learning, study and revision. However, I have received feedback that in some cases these aims and objectives are either missing, or are not covered in the lecture or textbook, or are in some way not very helpful (e.g., are too vague). In this note I will try to rectify some of these problems by giving revised or expanded aims and objectives with reference to lecture slides and textbook chapters in Mitchell [1].

2 Week 1: Fundamentals of Concept Learning

The single most important idea introduced in this lecture has to with representing hypotheses so that there is a well-defined generality ordering over the hypothesis space. This enables definitions of minimal generalisation and specialisation of hypotheses within a hypothesis space. See the basic definitions in Mitchell pp. 24–25. Note that these definitions can be extended to hypothesis spaces other than those in concept learning, such as those in Chapter 7 on computational learning theory. Take care with some of the other definitions:

satisfies an instance x satisfies a hypothesis h if and only if $h(x) = 1$, i.e., if the instance is an element of the concept denoted by the hypothesis (see p. 24)

consistent a hypothesis h is consistent with a set of training examples D if and only if $h(x) = c(x)$ for each example $\langle x, c(x)\rangle$ in D (see p. 29)

3 Week 3: Rule Learning

The objective "define association rules" should be expanded to include the definitions of support and confidence and how they relate to coverage and accuracy which are used in, say, sequential covering.

Also, the topics of cross-validation were introduced in this lecture with no corresponding learning objective. The learning objective should be to be able to define leave-one-out cross-validation (LOOCV) and k-fold cross-validation, and the concept of stratified cross-validation. Although these concepts fit more naturally into the lecture from week 9 on Evaluating Hypotheses they were introduced here since they are used by some rule learning methods.

4 Week 4: Machine Learning for Numeric Prediction

There has been feedback to the effect that it is difficult to answer the problems of setting up perceptrons to learn Boolean functions (e.g. problem 4.2 in Mitchell). The best suggestion to tackle such problems is from pp. 86–87 in Mitchell. For a Boolean function such as "and" or "or" consider first setting up the corresponding logical function (e.g. for "or" at least one input variable must be true or set to "1"). For other functions, such as "exor", you can take the same approach essentially by defining a nested logical function using "and", "or" and "not", e.g. $a \,\text{exor}\, b \equiv (a \vee b) \wedge \neg(a \wedge b)$.

5 Week 6: Computational Learning Theory

To address the feedback on this lecture we expand on some of the learning objectives as follows.

- define training error and true error in the setting of computational learning theory, and know the related definitions from statistics given in the lecture on Evaluating Hypotheses; this is in Chapter 7, Sections 1–2 of Mitchell.
- describe a basic theoretical framework for sample complexity of learning is covered by the material on ϵ-exhausting the Version Space with hypotheses of conjunctions of Boolean literals; this is in Chapter 7, Sections 3.1–3.2 (and 3.3 for those interested in extra study) of Mitchell. This also covers the objective of defining Probably Approximately Correct (PAC) learning.
- define the Vapnik-Chervonenkis (VC) dimension should also include the definition of shattering instances by hypotheses; this objective is covered by the material on Slides 21–25 and in Chapter 7, Sections 4.1–4.2 (and 4.3–4.4 for those interested in extra study) of Mitchell.
- define optimal mistake bounds should be expanded to cover being able to define mistake bounds for Find-S, the Halving Algorithm, as well as optimal mistake bounds (but not Weighted Majority); this objective is covered by the material on Slides 26–29 and in Chapter 7, Section 5 of Mitchell.

The objective of being able to reproduce the basic results in each area is covered by the above guidance.

6 Weeks 7 and 10: Bayesian Learning

The set of lecture notes for Week 7 titled "Bayesian Learning" was actually covered over two weeks. Part 1, in Week 7, contained learning objectives up to being able to reproduce Naive Bayes for text classification. However, Part 2 in Week 10, which covered the remainder of the material, did not have explicit learning objectives. Here are those objectives.

- define the structure of Bayes Nets in terms of conditional probability; covered on Slides 51–57 and in the textbook.

- describe the basic outline of inference and learning in Bayes Nets; covered on Slides 58–64 and in the textbook.

To link Parts 1 and 2 would be added the objective of being able to use the basic definitions of useful probabilities for inference, which relates to inference in Bayes Nets.

The material from the review paper by Rabiner on Hidden Markov Models will *not* be included in learning objectives. It will remain solely as background for those interested in extra study. This is because we did not have time to cover the material during lectures.

7 Week 9: Reinforcement Learning

This was a guest lecture, but the material *is* part of the course and is examinable. The learning objectives are as follows.

- describe the setting for reinforcement learning in terms if the agent and its environment
- define the learning task in terms of learning a policy for action selection
- define the optimal policy and the concept of discounted cumulative rewards
- describe the method of Q-learning, including the Q function and an algorithm for Q-learning
- describe the framework for convergence of Q-learning
- describe the method of temporal difference learning

This material is covered in the lecture notes and Chapter 13 of Mitchell.

8 Week 11: No Free Lunch, Ensembles and Kernels

This is the *final* week of lectures containing examinable material. Owing to the limited coverage in the lecture of the last two learning objectives, namely

- describe the concept of kernel methods
- describe the method of support vector machines

these will be *removed* as objectives in the sense of being examinable. Instead, they will remain solely as background for those interested in extra study. For this refer to the review articles posted on the Web page.

The material from the remaining weeks on the course has the same status, i.e., as background for those interested in extra study.

References

[1] T. Mitchell. *Machine Learning*. McGraw-Hill, New York, 1997.