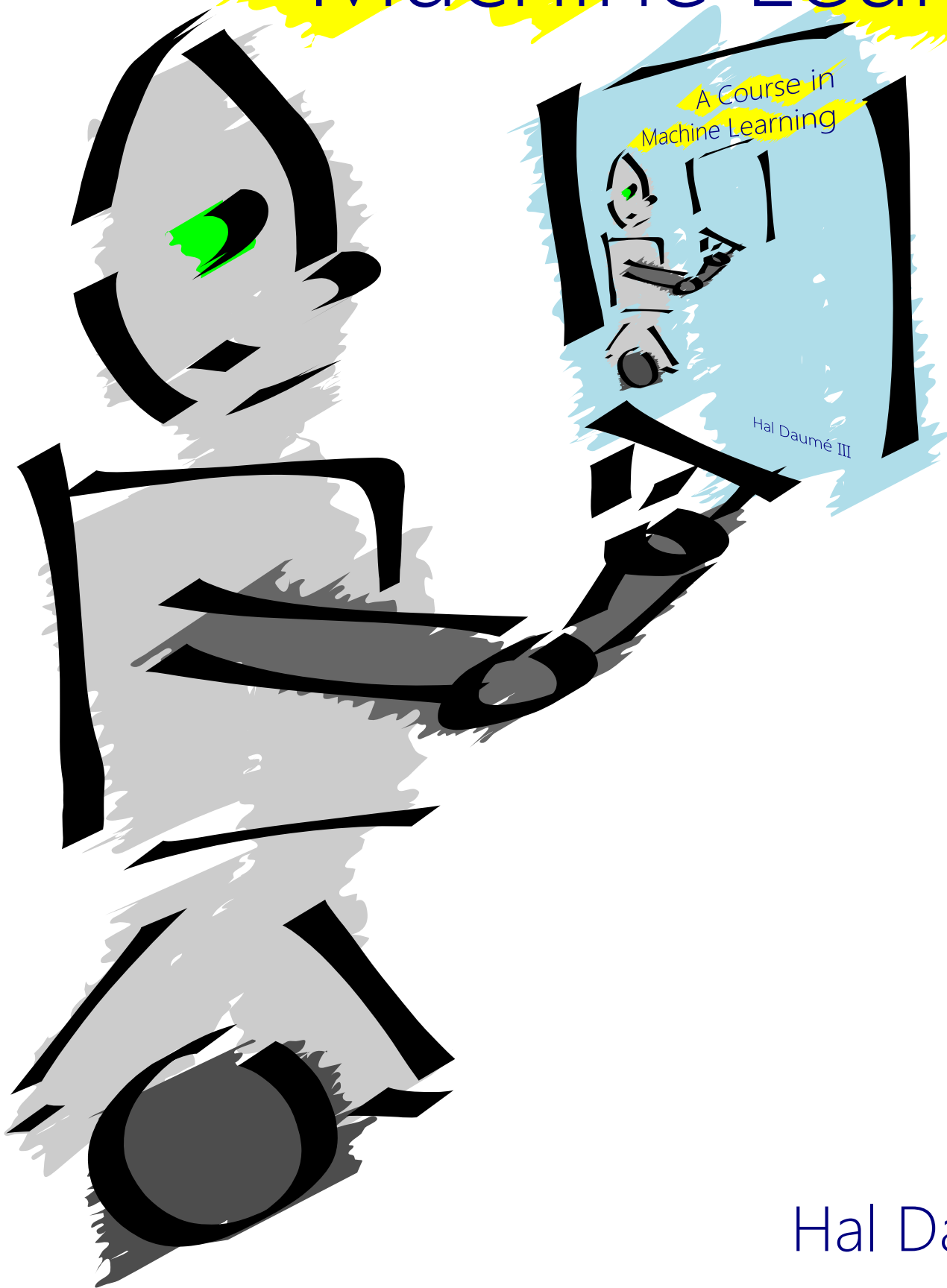


# A Course in Machine Learning



Hal Daumé III

Copyright © 2012 Hal Daumé III

<http://ciml.info>

This book is for the use of anyone anywhere at no cost and with almost no restrictions whatsoever. You may copy it or re-use it under the terms of the CIML License online at [ciml.info/LICENSE](http://ciml.info/LICENSE). You may not redistribute it yourself, but are encouraged to provide a link to the CIML web page for others to download for free. You may not charge a fee for printed versions, though you can print it for your own use.

version 0.8, August 2012

ALL OF THE LEARNING ALGORITHMS that you know about at this point are based on the idea of training a model on some data, and evaluating it on other data. This is the **batch learning** model. However, you may find yourself in a situation where students are constantly rating courses, and also constantly asking for recommendations. **Online learning** focuses on learning over a stream of data, on which you have to make predictions continually.

You have actually already seen an example of an online learning algorithm: the perceptron. However, our use of the perceptron and our analysis of its performance have both been in a batch setting. In this chapter, you will see a formalization of online learning (which differs from the batch learning formalization) and several algorithms for online learning with different properties.

#### Learning Objectives:

- Explain the experts model, and why it is hard even to compete with the single best expert.
- Define what it means for an online learning algorithm to have no regret.
- Implement the follow-the-leader algorithm.
- Categorize online learning algorithms in terms of how they measure changes in parameters, and how they measure error.

Dependencies:

## 17.1 Online Learning Framework

regret

follow the leader

agnostic learning

algorithm versus problem

## 17.2 Learning with Features

change but not too much

littlestone analysis for gd and egd

## 17.3 Passive Aggressive Learning

pa algorithm

online analysis

## 17.4 Learning with Lots of Irrelevant Features

winnow

relationship to egd

## 17.5 Exercises

**Exercise 17.1.** TODO...

Draft:  
Do Not  
Distribute