

Learning From Data
INTRODUCTORY MACHINE LEARNING COURSE

Yaser Abu-Mostafa
Professor of Electrical Engineering and Computer Science, Caltech

LECTURE 1
The Learning Problem

April 3, 2012
Hameetman Auditorium, Caltech

INFORMATION SCIENCE AND TECHNOLOGY Division of Engineering and Applied Science CALIFORNIA INSTITUTE OF TECHNOLOGY Caltech

Following program is brought to you by cal tech .welcome to machine learning and welcome to our online audience as well . Let me start with an outline of the course and then go into the material of today's lecture as you see from the outline the topics are given colors and that designates their main content with its mathematical or practical .machine learning is a very broad subject it goes from very abstract theory too extreme

practice as in rows of some .and the inclusion of a topic in the course depends on their relevance to much in learning so some mathematics is useful because it gives you the conceptual framework and then some practical aspects are useful because they give you the way to deal with any a learning systems .now if you look at the topics these are not meant to be separate topics for each lecture the just highlight the main content of those lectures but there is a story line that goes through it .and let me tell you what the story line is like

.starts here .ways .what is learning .can we live our .how to do it .how to do it well .and then that take home lessons . there is a logical dependency that go through the course .and there is one exception to that and knows because dependency .one lecture which is the third one doesn't really belong here .it's a practical topic and the reason I included it early on

because I need it to give you some tools to play around with the test they the theoretical and conceptual aspects .if I waited until it belongs normally which is to the second aspect of the learning mothers which is down there .the .the that that a beginning of the course would be just too theoretical for people's taste .okay and as you see if you look at the colors it is mostly a red in the beginning and mostly blue in there and so it's thousand big big building that concept and the fairy and then it goes on to the practical aspects .now .let me start today's lecture . and the subject of the lecture is the learning problem it's an introduction to what learning is .and I would draw your attention to one aspect of this light which is this bought .that's the logo of the course .and believe it or not this is not our to work .this is actually a technical figure out of that we've come up in one of


Outline of the Course	
1. The Learning Problem (April 3)	11. Overfitting (May 8)
2. Is Learning Feasible? (April 5)	12. Regularization (May 10)
3. The Linear Model I (April 10)	13. Validation (May 15)
4. Error and Noise (April 12)	14. Support Vector Machines (May 17)
5. Training versus Testing (April 17)	15. Kernel Methods (May 22)
6. Theory of Generalization (April 19)	16. Radial Basis Functions (May 24)
7. The VC Dimension (April 24)	17. Three Learning Principles (May 29)
8. Bias-Variance Tradeoff (April 26)	18. Epilogue (May 31)
9. The Linear Model II (May 1)	
10. Neural Networks (May 3)	

● theory; mathematical
 ● technique; practical
 ● analysis; conceptual

Learning From Data
Yaser S. Abu-Mostafa
California Institute of Technology

Lecture 1: The Learning Problem

Sponsored by Caltech's Provost Office, E&AS Division, and IST • Tuesday, April 3, 2012



the X. as I'm not going to tell you which one so you can wait in anticipation until it comes up .but this will actually be a scientific figured that was talked about .now when we moved to deter it to today's lecture .I'm going to talk today about the following .machine learning is a very broad subject .and I'm going to start with one example that captures the essence of machine learning it's a fine example about movies that everybody watches .and then after that I'm

going to abstract from their learning problems a practical learning problem .aspects that are


common to all at learning situation that you are going to face .and in extracting them we have the mathematical formalization of the learning problem .and then we will get our first algorithm for machine learning today it's a very simple algorithm but it will fix the idea about what is the role of a group in this case .and we would survey the types of learning so that we know which bought we are for him emphasizing in discourse and which

parts are at a nearby .and I would end up with a positive and very interesting puzzle and it's a puzzle in more ways than one as he was .okay so let's start with the example .they example of machine learning that I'm going to start with is .how a viewer would today it's a movie .day .now that is an interesting problem and it's interesting for us because which movie is a very interesting for a company that the rents out movies .and indeed a company which is Netflix .wanted to improve their in house sisteme by a mere ten percent .okay so they make recommendations when you login vehicle men movies that they think you're like so they think that you ordered them highly .and they had a system and they wanted to improve the system .so how much is a ten percent better improvement in performance worth to the company .it was actually .one million dollars .that was paid out to the first group that actually manage to get the ten percent improvement .so ask yourself up ten percent improvement in something like that why should that be worth a million dollars .it's

Example: Predicting how a viewer will rate a movie

10% improvement = 1 million dollar prize

The essence of machine learning:



2/19

because if they recommendations that the movie company makes .are .spot on .you would pay more attention to the recommendation you are likely to rent the movies at their command and they would make lots of money much more than the million dollar they promised .and this is very typical in machine learning for example might you Lenny has a petition in finance are forecasting .you can imagine that a the minute just

improvement in financial forecasting can make a lot of money .so the fact that you can actually push the system to be better using machine learning is a very attractive aspect of the technique in out wide spectrum of applications .okay so what did these guys do .they give the thing is that the data and people started working on the on the problem using different algorithms until someone managed to to to get the price .now if you look at the problem over eating a movie .it captures the essence of machine learning and the essence has three components .if you find the three components in a problem you have in your field then you know that machine learning is already as an application to a .what are the three .the first one is that a pattern exists .okay if a pattern didn't exist there would be nothing to look for .so what if the pattern here .there is no question that that way appearance on rates a movie

The learning problem - Outline

- Example of machine learning
- Components of Learning
- A simple model
- Types of learning
- Puzzle

Learning From Data - Lecture 1

2/19

Example: Predicting how a viewer will rate a movie

10% improvement = 1 million dollar prize

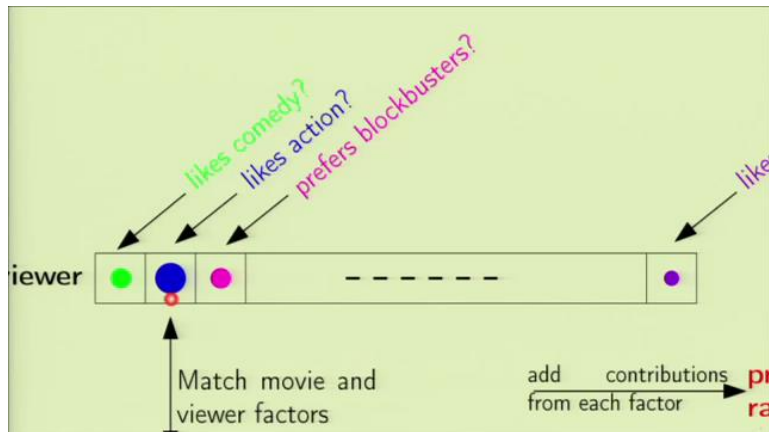
The essence of machine learning:

- A pattern exists.
- We cannot pin it down mathematically.
- We have data on it.

Learning From Data - Lecture 1

3/19

.is it related to how was the of it did other movies .and is also related to how other people but it did that movie .we know that much .so there is a pattern to be discovered .however .we can not to really pin it down mathematic Eddie I cannot ask you to write a seventies or the polynomial that captures how people love it movies .so the fact that there is a pattern and that we cannot pin it down mathematically is the reason why you are going for machine learning fro for learning from data we couldn't write down the system on our own so we are going to depend on data in order to



be able to find this is .there's a missing component which is very important if you don't have that you are out of luck .we have to have data we are learning from data .so if someone knocks on my door was an interesting machine learning application and they didn't how exciting it is and how good the application would be and how much money they would make the first question I asked

.what data do you have okay if you have that that we are in business .if you don't you are out of luck .if you have the three components you're ready to apply machine learning .okay .now .let me give you a solution to the movie rating you know to have to start getting a feel for it okay so here is a system let me start to focus on part of it .okay .we are going to describe a view out as the Victor or factors .okay a profile if you will okay .so if you look here .example the first one would be comedy content .okay is that it does the movie have a lot of comedy .okay .