ECE/CS/ME 532:

Matrix Methods in Machine Learning

(formerly Theory and Applications of Pattern Recognition)

Prof. Rebecca Willett (willett@discovery.wisc.edu)

NOT Rebekah Willett (willett@wisc.edu)

Machine Learning is about learning by example.

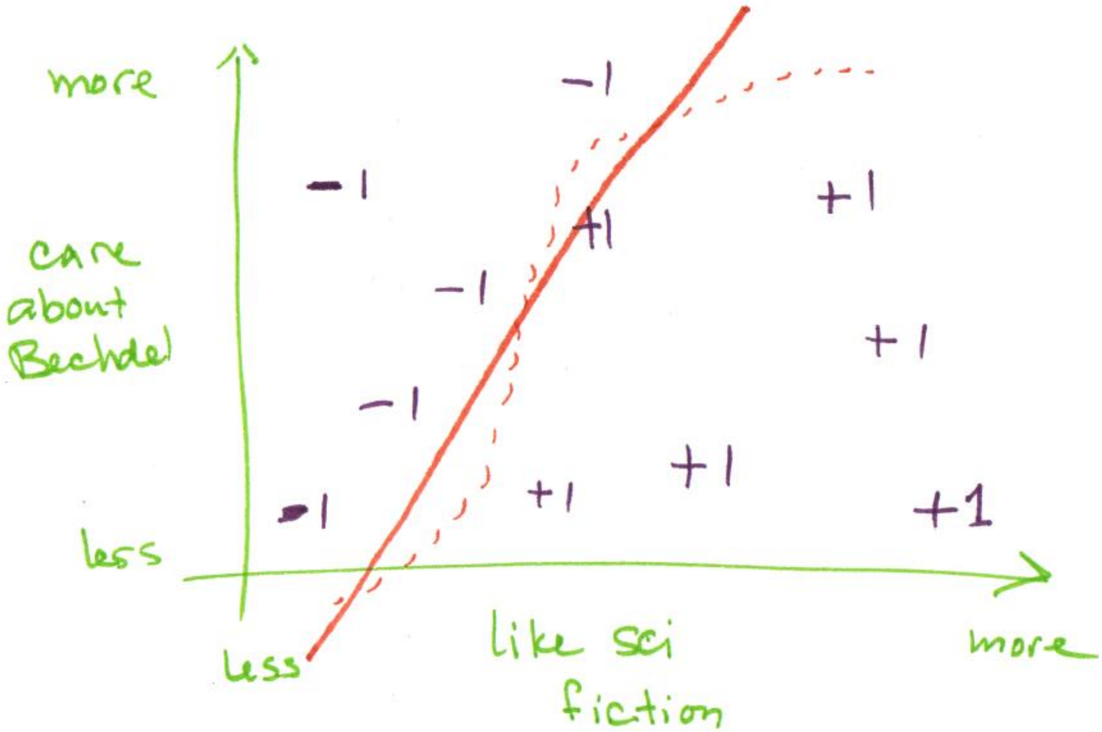
Ex. Imaging we want predict ~~which~~ whether someone will enjoy the orig. Star Wars Trilogy.

First we take a survey of people who watched it; and ask

1. did you like it?
2. how much do you ushally like Sci fi ?
3. how much do you care about the Bechdel test ?

Bechdel test : movie must have
≥ 2 female characters (with names)
talk to each other
about something besides men.

Then we plot the results



Can we now predict whether someone will like Star Wars based on how much they like Sci fi and the Bechdel test?

One approach: predict label y using a weighted combination of features x_1 = sci fi pref.

x_2 = Bechdel pref.

$$\hat{y} = \text{weight}_1 \cdot x_1 + \text{weight}_2 \cdot x_2$$

How do we choose weight_1 and weight_2 based on the survey results?

Let's index the surveys $i = 1, 2, \ldots, n$

for i^{th} survey, have label y_i

and features x_{i1} (sci fi) and x_{i2} (Bechdel)

goal: make sure $\hat{y}_i = \text{weight}_1 \, x_{i1} + \text{weight}_2 \, x_{i2} \approx y_i$

for all i

Matrix Methods help us find good weights.

let $\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}$ be a vector of labels (~~SWF~~ Star wars preference)

let $\underline{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix}$

$x_{i,j}$ = survey response of i^{th} person on j^{th} question

each row of $\underline{X}$ is responseof one person

each column of $\underline{X}$ is all responses to one question vector of weights $\underline{w} = \begin{bmatrix} weight_1 \\ weight_2 \\ \vdots \\ weight_p \end{bmatrix}$

before $\hat{y} = \sum_{j=1}^{p} weight_j \, x_j$

now: $\underline{\hat{y}} = \underline{X} \, \underline{w} \Longrightarrow$ now we use matrix methods to find $\underline{w}$ so that $\underline{y} \approx \underline{\hat{y}}$

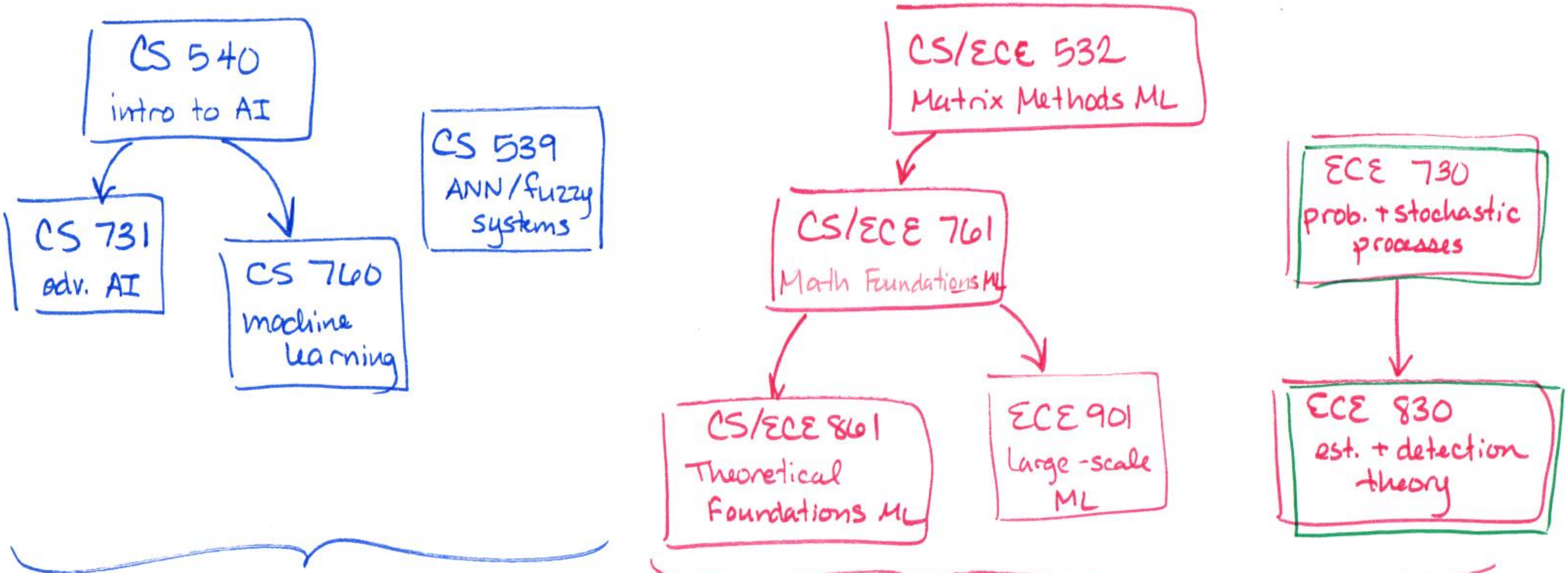
Matrix methods are at the heart of

- machine learning ← our focus
  - predicting labels / classification
  - recommender system
  - image recognition / analysis
- robotics
- finance
- mech / aero engineering
- signal processing
- optimization / operations research.

Beyond toolboxes

- put algorithms in context / understand <u>why</u> they work
- design new algorithms

Maching Learning + Signal Processing Class @ Wisconsin



more applied survey style courses (lots of coding)

more mathematical focus (less coding)

Signal processing courses