Administrative Notes
January 23, 2018

- Reminder: In the News group component due today
- Reminder: Reading quiz due tomorrow
- Reminder: Reading quizzes due next Monday
- Reminder: Project groups due February 7 on Canvas
- Reminder: Midterm #1 February 1 (in class)
- In the News individual marks released
Last class..

- May have been confusing
- Let’s try again
What is classification?

- The idea behind classification is that we want to use patterns and/or correlations to make decisions.

- Classification happens all the time in real life.
  - The doctor uses your symptoms and other measurements like weight/blood pressure/etc. to help make a diagnosis.
  - Google uses classification to determine what an image is.

- Classification is a general class of algorithms.
Classifiers

- Classifiers are algorithms that perform classification

- They are specific
  - E.g., we don’t give loans to anyone an income of less than $50,000 per year

- The algorithm you come up with is no different than the other algorithms you’ve come up with so far
  - You still need to state the steps you need to take to come up with the solution
How do I write a classifier?

- Start with the data!
- Start with a full set of data that you have

<table>
<thead>
<tr>
<th>Applicant</th>
<th>Annual Income</th>
<th>Loan Approved?</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>26 000</td>
<td>No</td>
</tr>
<tr>
<td>#2</td>
<td>60 000</td>
<td>Yes</td>
</tr>
<tr>
<td>#3</td>
<td>50 000</td>
<td>Yes</td>
</tr>
<tr>
<td>#4</td>
<td>47 000</td>
<td>No</td>
</tr>
<tr>
<td>#5</td>
<td>12 000</td>
<td>No</td>
</tr>
<tr>
<td>#6</td>
<td>108 000</td>
<td>Yes</td>
</tr>
</tbody>
</table>
How do I write a classifier?

- Divide the data arbitrarily into the test data and the training data
- We just chose a 50/50 split for our demo but you could do other splits like 60/40, 70/30

<table>
<thead>
<tr>
<th>Applicant</th>
<th>Annual Income</th>
<th>Loan Approved?</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>26 000</td>
<td>No</td>
</tr>
<tr>
<td>#2</td>
<td>60 000</td>
<td>Yes</td>
</tr>
<tr>
<td>#3</td>
<td>50 000</td>
<td>Yes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Applicant</th>
<th>Annual Income</th>
<th>Loan Approved?</th>
</tr>
</thead>
<tbody>
<tr>
<td>#4</td>
<td>47 000</td>
<td>No</td>
</tr>
<tr>
<td>#5</td>
<td>12 000</td>
<td>No</td>
</tr>
<tr>
<td>#6</td>
<td>108 000</td>
<td>Yes</td>
</tr>
</tbody>
</table>
How do I write a classifier?

Given your training data, can you find a pattern that can tell you when to approve a loan?

Last class, we decided an annual income of $50 000 seemed like a good cut off point. \(\rightarrow\) That was a classifier!
How do I write a classifier?

<table>
<thead>
<tr>
<th>Applicant</th>
<th>Annual Income</th>
<th>Loan Approved?</th>
<th>Classifier said to…</th>
</tr>
</thead>
<tbody>
<tr>
<td>#4</td>
<td>47 000</td>
<td>?</td>
<td>No</td>
</tr>
<tr>
<td>#5</td>
<td>12 000</td>
<td>?</td>
<td>No</td>
</tr>
<tr>
<td>#6</td>
<td>108 000</td>
<td>?</td>
<td>Yes</td>
</tr>
</tbody>
</table>

After you come up with a classifier that seems to do okay with your training data, you use it on your test data to see what kinds of decisions it makes.
How do I write a classifier?

<table>
<thead>
<tr>
<th>Applicant</th>
<th>Annual Income</th>
<th>Loan Approved?</th>
<th>Classifier said to…</th>
</tr>
</thead>
<tbody>
<tr>
<td>#4</td>
<td>47 000</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>#5</td>
<td>12 000</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>#6</td>
<td>108 000</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

If the results of your classifier match up with the decisions you’ve made in your test data, it’s looking good. You can start trying to use it on data that you haven’t made any decisions on yet.
What were we trying to do with the cards?

- We took out all the red cards that were 8 and above
- We divided what was left into training data and test data
What were we trying to do with the cards?

• What were we trying to do?
  • We tried to come up with a classifier to predict what the next card coming up would be
  • We weren’t looking for the exact card it would be (i.e., we weren’t trying to predict if the next card is the 7 of spades or the 5 of hearts). We wanted to know if the next card was of high value (8 and above) or low value (7 and below plus aces)
  • Come up with an algorithm that would help predict the next card. What would you look at? What pieces of information would you use to help make a decision?
What were we trying to do with the cards?

• We could have counted the cards that we’ve seen and determine if there are more cards that are 7 or lower or 8 and higher
  • You can try to predict that would come afterwards based on the ratio

• That’s where we ended last class
Let’s look at a more complex example: loan applications (from Hardt et al. at Google)

- The bank makes $300 on a successful loan, but loses $700 on a default
- Training data of historical applicants describes the applicant’s credit rating and are labeled as either successful or defaulters
- Light blue are the defaulters, dark blue are successful
Loan application example

Classification task: approve or deny a loan application, based on credit threshold

Group exercise: choose a threshold (credit rating) at which to approve/deny loans and define why you chose that threshold

Light blue are the defaulters, dark blue are successful
Group Exercise

Choose a threshold (credit rating) at which to approve/deny loans and define why you chose that threshold

- **50**: is at a point where more than 2/3 of the people pay you back

- **55**: counted pale dots as -700 and dark blue dots as +300; found a point where you started to make money

- **70**: 100% return
Loan application threshold #1: 50

Threshold Decision

credit rating

Outcome

Correct 84% loans granted to paying applicants and denied to defaulters

Incorrect 16% loans denied to paying applicants and granted to defaulters

True Positive Rate 83% percentage of paying applications getting loans

Positive Rate 48% percentage of all applications getting loans

Profit: 14800
Loan application threshold #2: 54

**Threshold Decision**

credit rating

loan threshold: 54

denied loan / would default

granted loan / defaults

denied loan / would pay back

granted loan / pays back

**Outcome**

Correct 83%
loans granted to paying applicants and denied to defaulters

Incorrect 17%
loans denied to paying applicants and granted to defaulters

True Positive Rate 75%
percentage of paying applications getting loans

Positive Rate 41%
percentage of all applications getting loans

Profit: 16600
Does this seem fair?

If we maximize our profits, then 17% of people who would likely succeed in paying back their loans will be denied loans. Is this fair?

A. Yes
   Yes: because you want to maximize profits

B. No
   No: If you increase the number of loans overall, then the % difference makes a big deal
Changing the problem: there are two groups of people – blue and orange

- Each group has the same # of dots
- Each group has half defaulters/half successful
- Only the distributions are different
Loan application example: Consider both populations together

Classification task: approve or deny a loan application, based on credit threshold and/or colour
Loan application example: Consider both populations together

The website this example came from has 4 different types of classifiers for this example:

https://research.google.com/bigpicture/attacking-discrimination-in-ml/ (shortened: goo.gl/o5Szyu) – URL is also on website under Extra Resources

Discuss:

• What do you think are the pros and cons of each?
• Which do you think is most fair?

Note: The goal is not to have “the” right answer; just think through the options. In the real world, consider involving someone with ethics expertise.
Which classifier do you think is fairest?
Clicker question

A. Max Profit
B. Group Unaware
C. Demographic Parity
D. Equal Opportunity
Group discussion summary:
Which classifier do you think is fairest?

Max profit: drop the least amount of people; make the most money

Demographic parity and equal opportunity: relatively higher profit while kind of decreasing the gap between the loan threshold of the two different demographics

Group unaware: it’s unethical to discriminate based on something racial or similar types of information to that
“Sometimes, in order to be fair, it is important to make use of sensitive information... This may be a little counterintuitive: The instinct might be to hide information that could be the basis of discrimination.”

- Cynthia Dwork
Using Sensitive Information Discussion

Should people designing programs to assist loan officers in deciding on who to make loans to consider sensitive information (such as race) in designing their programs?

A. Yes
Yes: We don’t want algorithms to increase the amount of discrimination that already exist. Certain groups may have lower credit scores and incomes already because of how society worked out; don’t want to contribute to this problem.

B. No
No: Payback of loans is a character issue and it’s not tied into race. Algorithm should be blind to this info.
Let's talk about bias. There are two main ones involved.

- **Conscious bias** is when you're biased and you know it (and you're generally not sorry)
- **Unconscious bias** is when you're biased... and you may not know it (and if you do, you're sorry)... and you may even be biased against what you believe!

An example of unconscious bias

- [http://wwest.mech.ubc.ca/diversity/unconscious-bias/](http://wwest.mech.ubc.ca/diversity/unconscious-bias/)
Test this on yourself

http://www.understandingprejudice.org/iat/

Seriously, test yourself at some point.
Unconscious bias on gender and work

<table>
<thead>
<tr>
<th>Test Result</th>
<th>% of Test Takers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong association between male and career</td>
<td>40%</td>
</tr>
<tr>
<td>Moderate association between male and career</td>
<td>15%</td>
</tr>
<tr>
<td>Slight association between male and career</td>
<td>12%</td>
</tr>
<tr>
<td>Little or no gender association with career or family</td>
<td>17%</td>
</tr>
<tr>
<td>Slight association between female and career</td>
<td>6%</td>
</tr>
<tr>
<td>Moderate association between female and career</td>
<td>5%</td>
</tr>
<tr>
<td>Strong association between female and career</td>
<td>5%</td>
</tr>
</tbody>
</table>
## Unconscious bias on race

<table>
<thead>
<tr>
<th>Test Result</th>
<th>% of Test Takers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong automatic preference for White people</td>
<td>48%</td>
</tr>
<tr>
<td>Moderate automatic preference for White people</td>
<td>13%</td>
</tr>
<tr>
<td>Slight automatic preference for White people</td>
<td>12%</td>
</tr>
<tr>
<td>Little or no automatic preference</td>
<td>12%</td>
</tr>
<tr>
<td>Slight automatic preference for Black people</td>
<td>6%</td>
</tr>
<tr>
<td>Moderate automatic preference for Black people</td>
<td>4%</td>
</tr>
<tr>
<td>Strong automatic preference for Black people</td>
<td>6%</td>
</tr>
</tbody>
</table>

If your test results showed a preference for a certain group, you may have a hidden, or unconscious, bias in favor of that group. The results of more than one million tests suggest that most people have unconscious biases. For example, nearly two out of three white Americans show a moderate or strong bias toward, or preference for, whites, as do nearly half of all black Americans.
Bias case study

In a group of 3-4 people, discuss the Facebook reading article where workers were accused to be biased against conservative articles. In particular,

• List a decision that could be made that would be the result of *conscious* bias and say why.

• List a decision that could be made that would be the result of *unconscious* bias and say why.
Facebook discussion

Conscious Bias

- Omit conservative articles because of their political beliefs
- Omit things relating to Facebook

Unconscious Bias

- They hired mostly graduates from liberal arts colleges so those graduates may talk about liberal news more often which cause them to subliminally prefer that type of news
- Might be more inclined to trust certain types of news
Speaking of Facebook and Bias…
Facebook and the need for humans in the news!

[Facebook] “enabled advertisers to direct their pitches to the news feeds of almost 2,300 people who expressed interest in the topics of ‘Jew hater,’ ‘How to burn jews,’ or, ‘History of ‘why jews ruin the world.’”
https://www.propublica.org/article/facebook-enabled-advertisers-to-reach-jew-haters

Solution: human reviewers

Fun game: how long until someone complains about bias?
Other sites also grapple with fake news (yes, there is real fake news)

“Google's search algorithm has been changed over the last year to increasingly reward search results based on how likely you are to click on them…”

As a result, fake news now often outranks accurate reports on higher quality websites.

The changes at [have] real world consequences — such as the Comet Ping Pong shooting, done by a man who was convinced from his internet searches that Clinton was using the restaurant as a front for a child abuse ring.”
Google search and fake news
Business Insider

**FAKE NEWS**
Obama Signs Executive Order Banning The National Anthem At All ...
cnn.com.de › News
11 Nov 2016 - President Obama has signed an Executive Order banning the National Anthem with fines and jail time beginning December 1st of this year.

**FAKE NEWS**
Obama Signs Executive Order Banning The Pledge Of ... - ABC News
abcnews.com.co › News
President Obama signs an Executive Order banning the Pledge of Allegiance in schools nationwide
President Obama, seen here signing an Executive Order ...

**REAL NEWS**
Obama Signs Executive Order Banning the Pledge of ... - Snopes
www.snopes.com/pledge-of-allegiance-ban/
16 Aug 2016 - Unimaginative fake news publishers have recycled an old hoax about President Obama's banning the Pledge of Allegiance. ... Claim: President Obama has issued an executive order banning the Pledge of Allegiance in U.S. schools. ... Early this morning, President Obama made what could very ...

**REAL NEWS**
Obama Did Not Ban the Pledge - FactCheck.org
www.factcheck.org/2016/09/obama-did-not-ban-the-pledge/ 
2 Sep 2016 - Q: Did President Obama sign an executive order banning the Pledge of ... Order Banning The Pledge Of Allegiance In Schools Nationwide.
Google search and fake news
Clicker question

“There should just be no situation where fake news gets distributed, so we are all for doing better here,” Google CEO Sundar Pichai said

Should Google search rankings penalize fake news?
A. Yes
B. No
What is the responsibility of a company when their work is shown to be unfair?

Consider when Google Photos labeled black people as gorillas.

What responsibility did Google have once this was discovered? What about before labelling pictures at all?

- Apologize and fix the problem
- Make sure the algorithm actually works before letting the world use it
What is a company’s responsibility when users intentionally introduce inappropriate behaviour?
Clicker question

Should Google have changed their algorithm in response to this?

A. Yes
B. No
Generalizing the examples

Computer algorithms don’t just end at what we’ve discussed and what we’ve read about in the readings.

It’s engrained in many of the things we use every day! Try identifying some that affect your every day life and examine it to see if it has a conscious or unconscious bias.

E.g., If you search for something on Google, it shows up on Amazon as a recommended product or on Facebook/Instagram/Snapchat as an ad
Learning Goals

• CT Building Block: Students will be able to explain examples of how computers do what they are programmed to do, rather than what their designers want them to do.

• CT Impact: Students will be able to list reasons that an algorithm might be biased and what its impact will be.

• CT Impact: students will be able to list arguments why a company should or should not change its algorithms based on “fairness”
In the News Call #1 - Group

Anyone still looking for a group?

If you are, head to the front to see if you can find anyone else also looking for a group.