

QA: ML Fairness (theoretical) case study

17-313 Fall 2022

Rohan Padhye, Michael Hilton, Chris Timperley, and Daye Nam



Administrativia

- HW4 implementation and Design Doc due today.
- US Election Day



Learning goals

- Understand different fairness approaches
- Describe strengths and weaknesses of fairness approaches
- Reason about tradeoffs in fairness



Fairness



ML Fairness

• Getting answers is the easy part... Asking the right questions is the hard part.



https://towardsdatascience.com/a-tutorial-on-fairness-in-machine-learning-3ff8ba1040cb



Perception:







Life is often not this simple...







Fairness

- Is a deeply technical topic, but we will discuss it at a higher level of abstraction.
- The formulas are important, but knowing which formula to apply is MUCH more important
- This is a special case of how to to test when the desired outcome is hard to measure.



FIGURE 2. A graph of SPS ([stream] starts per second) over a 24-hour period. This metric varies slowly and predictably throughout a day. The orange line shows the trend for the prior week. The y-axis isn't labeled because the data is proprietary.



What does "fair" mean?



What is Fairness?

- Law
 - fairness includes protecting individuals and groups from discrimination or mistreatment with a focus on prohibiting behaviors, biases and basing decisions on certain protected factors or social group categories.
- Social Science
 - "often considers fairness in light of social relationships, power dynamics, institutions and markets."3 Members of certain groups (or identities) that tend to experience advantages.



What is Fairness? continued

• Quantitative Fields

 (i.e. math, computer science, statistics, economics): questions of fairness are seen as mathematical problems. Fairness tends to match to some sort of criteria, such as equal or equitable allocation, representation, or error rates, for a particular task or problem.

- Philosophy:
 - ideas of fairness "rest on a sense that what is fair is also what is morally right." Political philosophy connects fairness to notions of justice and equity.



Fairness as QA



How can we define "fair"

- For the purposes of creating an oracle
- We must have a better definition than infamous 1964 Supreme Court obscenity test:
 - I shall not today attempt further to define [obscene material], and perhaps I could never succeed in intelligibly doing so. But *I know it* when I see it, and the motion picture involved in this case is not that.



We don't need to start from scratch...



Varieties of fairness (names vary)

Group unaware

- Ignore group data (one group could get excluded)
- Group thresholds
 - Different rules per group (rules differ by group)
- Demographic parity
 - Same percentage in pool as outcomes (might result in random selection)
- Equal opportunity
 - Equal chance out positive outcomes regardless of groups (focus on individual, rules differ per group)
- Equal accuracy
 - Equal chance of both outcomes per group (focus on group, rules differ per group)



Explainability

Simulating loan thresholds Drag the black threshold cars left or right to change the out-offs for loans.

Threshold Decision



Outcome

Correct 84% loans granted to baying applicants and denied to defaulters.



Incorrect 16%

loans denied to caving epplicants and granted to defaulters

True Positive Rate 86% percentage of paying applications getting loans

Positive Rate 52% percentage of all applications getting loans



https://research.google.com/bigpicture/attacking-discrimination-in-ml/



Activity

Consider the different approaches to fairness. Can you come up with different scenarios where each fairness approach might or might not be appropriate?

Remember the fairness approaches are:

- Group unaware
- Group thresholds
- Demographic parity
- Equal opportunity
- Equal accuracy



Resources

- Fairness Textbook:
- <u>https://fairmlbook.org/testing.html</u>



Decision Making/ Tradeoffs



Life is all about tradeoffs

In this course, we have talked a lot about tradeoffs.

Some tradeoffs we have discussed:

Writing Tests vs writing more features

Choosing a familiar tech stack vs a "trendy one"

Other tradeoffs...?



Think about structured ways to make decisions

Do what is ethical, legal, moral, obvious

Sometimes, there are multiple (legitimate) options, that all seem to have positives and negatives.

What to do?



People have studied this before...

Economics

Biology + Microbiology

Sociology

Engineering

Computer Science

Strategy Games

Ethics

Medicine

Politics



Architecture Tradeoff Analysis Method

- 1. Present the ATAM.
- 2. Present business drivers.
- 3. Present architecture.
- 4. Identify architectural approaches.
- 5. Generate quality attribute utility tree.
- 6. Analyze architectural approaches.
- 7. Brainstorm and prioritize scenarios.
- 8. Analyze architectural approaches.
- 9. Present results.

https://concisesoftware.com/architecture-tradeoff-analysis-method-atam/



🔁 S3D

Pros and cons

THE PROS & CONS of MAKING A PROS & CONS LIST PROS CONS LISTS ARE FUN! CLEVER WAY .3 of RATINALIZING P MAKES YOU Appear A BAD DECKION THOUGHT ful & DELIBFRATE GREAT PRACTICE AT DRAWING STRAIGHT LINES -THE OATHERL

https://theoatmeal.com/blog/pros_cons_list

S3D

Moral or Prudential Algebra

1. Make a list of Pros and Cons

2. Take a few days to think, forcing your brain to come up with all of the positive and negative aspects of taking the measure you're considering.

3. Assign weights to each item on your list depending on its importance.

4. When the weight of a item on your "pro" list is equal to the weight of an item on your "con" list, strike both of them out.5. What is left in the balance is the choice you need to make.Wait a few more days and, if nothing new occurs to you, act on the decision you've made.





Even Swaps

Premise:

It is easy to make decisions when there is only one objective.

But having only one objective, as any decision maker knows, is a rare luxury.

Even swaps provides a practical way of making trade-offs among any set of objectives across a range of alternatives.

THE POLITIES OF E- MAIL	Đ	405.12
Harvar	d Business Revie	ew
S	MARCH-AR	11. 1999
	8 A A	2
antoine B. closes and and status I. brane	THE ALCHO	19
taffileer B. Shantar B and Shina I. Index (on Regards)	The Autor The Autor Constraints in Market The Autor State State The Bower of Vittle Barrow Internation Are entering with Disk Constraint Internations	10 12
unteleur II. disecurit and Milai I. belan per magnetic relation focuer in method (THE NAME AND	1+ T2 86
unterber M. Schweizung and Miller I. Leiden om Heigherts schweizung der Kennen im diese, und Kanner im Height diese, und Kanner im Height weiter Neusons Heighteit	THE RECEIPT OF ALL AND ADDRESS OF A LEADER	14 72 86
tarbiten R. disectori and Olice I. below per magnetic relation. Institut method (r. per methods method) and per field tensors pergin t. bedenten, jr.	THE RECEIPTER OF BARRIES CONTACTOR	19 72 86 891
uniciano III, disensa etti add Milard I. Soniani com maganetta sollard I. Soniani mishand J. Sinon, tesi Kamari mishand S. Mani Manopan mishangoli and Juan Manopan mishangoli and Juan Manopan mishangoli peri kandi Kamariani peripeti kandinasiani, jut. antinang Canobai	THE RECEIPTING OF BALANCE CONTACT TO A CONTACT OF A CONTA	19 19 191 194 194
uniciana II. diamanan'i ani Mina I. Lanan one maganita indiari I. Indiari Mathari I. Jine ani Antoni mantani ani Juan India Kanana ani Juan India Kanana ani Juan India Kanana ani Juan India Kanana Ani Dadi Cangga Millin Cangga	THE NAMES THE REACHS IN MARKET THAT VERY TRANSPORT AN INTERIM THE DIS CONFERENCE INCOME. THE REPORT OF INATURE REPORT TO A REPORT MILLION THE REPORT OF A REPORT THE REPORT OF INATURE REPORT THE REPORT OF A REPORT OF THE INTERNET ARTICLE REPORT OF A REPORT OF THE INTERNET MILLION OF THE INCOME.	10 12 13 13 13 13 13 13 13
tantahan K. dawaani ani dawa L balan Internet Internet Internet Jam Makan K. Matha Jam Makan K. Matha Jam Makan K. Matha Matha Jam Makan K. Ja Matha M	THE NAME OF ANALYSIS THAT WORK IS AND SOLD. THE REAL OF MARKET I THAT WORK IS AND SOLD. THE REAL WARKET I THAT WORK IS AND SOLD. THE REAL WARKET IS AN ADDRESS. THE REAL WARKET IS A ADDRESS. THE REAL WARKET IS ADDRESS. THE REAL WARKET IS A ADDRESS. THE REAL WARKET IS ADDRESS. THE REAL WARKET IS ADDRESS. THE REAL WARKET IS A ADDRESS. THE REAL WARKET IS ADDRESS. THE REAL WARKET IS A ADDRESS. THE REAL WARKET IS A ADDRESS. THE REAL WARKET IS A ADDRESS. THE REAL WARKET IS ADDRESS. THE REAL WAR	19 12 19 19 19 19 19 19 19 19 19 19 19 19 19
tartichen R. disectore and ethna I. beiter (een megenetis) entran I. beiter einen, end kanner en erebet genetik bestere eregin I. bedienster, jo unterland darossi somethik bestere eregin I. bedienster, jo unterland darossi somethik bestere eregin I. bedienster, jo unterland darossi somethik bestere eregin I. bedienstere eregin I. bedie bestere eregin I.	THE NAME OF AN ADDRESS OF A DESCRIPTION	10 12 13 14 134 134 134 134 134



Overview - 3 step process

1. Create a Consequences Table

2. Eliminate "Dominated" Alternatives

3. Make Even Swaps



Running Example from research paper

Example: Alan Miller is a computer scientist who started a technical consulting practice three years ago.

For the first year, he worked out of his home, but as his business grew he decided to sign a two-year lease on some space in the Pierpoint office park.

Now that lease is about to expire. He needs to decide whether to renew it or move to a new location.



Define the objectives

Alan defines five overriding objectives that he needs his office to fulfill:

- 1. a short commute from home
- 2. good access to his clients
- 3. good office services (clerical assistance, copiers and fax machines, and mail service)
- 4. sufficient space
- 5. low costs

He finds five viable alternatives: Parkway, Lombard, Baranov, Montana, and his current building, the Pierpoint.



Create Consequences Table

MILLER'S CONSEQUENCES TABLE								
Objectives	Alternatives Parkway	Lombard	Baranov	Montana	Pierpoint			
COMMUTE IN MINUTES	45	25	20	25	30			
CUSTOMER ACCESS (%)	50	80	70	85	75			
OFFICE SERVICES	A	В	c	A	с			
OFFICE SIZE (SQUARE FEET)	800	700	500	950	700			
MONTHLY COST (\$)	1850	1700	1500	1900	1750			

https://hbr.org/1998/03/even-swaps-a-rational-method-for-making-trade-offs



Eliminate "Dominated" Alternatives





Make Even Swaps

- 1. Determine the change necessary to cancel out an objective.
- 2. Assess what change in another objective would compensate for the needed change.
- 3. Make the even swap.
- 4. Cancel out the now-irrelevant objective.
- 5. Select the dominant alternative.



First Even Swaps

MILLER'S EVEN SWAPS 1

Alternatives							
Objectives	Lombard	Baranov	Montana				
COMMUTE IN MINUTES	25	20, 25					
CUSTOMER ACCESS (%)	80	79, 78	85				
OFFICE SERVICES	в	с	A				
OFFICE SIZE (SQUARE FEET)	700	500	950				
MONTHLY COST (\$)	1700	1500	1900				



Second Even Swaps





Final Swap

MILLER'S EVEN SWAPS 3





Advice for swaps

- Make the easier swaps first
- Concentrate on the amount of the swap, not on the apparent importance of the overall objective.
- Remember that the value of an incremental change depends on what you start with.
- Make consistent swaps.
- Seek out solid information.

