Unknown Speaker 0:17

Hi,

Unknown Speaker 0:21

reading the steps

P53 0:51

m r we'll finish with reading through the models and stuff

Unknown Speaker 1:00

Go around and

Unknown Speaker 1:00

just say like, what our purpose is and

P53 1:05

our case funnel for top two models that we prefer.

Unknown Speaker 1:14

Okay,

P53 1:15

I guess I'll start I'm, I'm a fairness advocate. So I would prefer a model I think, I believe two and four because they have a lower figure figure and discrepancy. I mean, disparity sorry. And that's sort of what I care the most about.

P54 1:41

Okay, I'm, I'm a judge. So I care the most about having a model that has higher accuracy. So I chose model three and seven because their accuracy is the highest Yeah.

P55 2:03

I'm a community member. So my priority is safety. And with that would be minimizing the false negative rate. So I chose model two and eight.

P56 2:17

I'm also a community member, so

Unknown Speaker 2:30

I can't hear you.

Unknown Speaker 2:34

Yeah, I can't hear you either. me. Oh,

Unknown Speaker 2:37

no, no, I mean, in.

P54 2:42

Sorry if I said that wrong. Hello.

Oh, yeah.

Unknown Speaker 2:50

Oh, I'm a defendant. And I'm supposed to find a model that Lower Falls

Unknown Speaker 2:59

by two

P54 3:10

Should we like, I guess like list out which models he saw, like we all chose and then see if there any overlaps,

Unknown Speaker 3:18

I guess.

P53 3:19

I think there was overlap and seven and then there was overlap in two.

Unknown Speaker 3:25

Okay.

P53 3:28

But maybe we can also send our model number to the chat. That would be clear, maybe. Okay.

Unknown Speaker 4:09

They both have overlap. Yeah. Okay.

Unknown Speaker 4:28

Seven is more

Unknown Speaker 4:31

pain, way lower false positive for you.

P56 4:37

Yeah. And has a much larger disparity in accuracy.

P53 4:51

That's a pretty high false negative rate as well. Yeah.

P56 4:59

I guess that's why false positive rate. So low rate if you're just like Yeah.

Okay, let's look at these axes along which are supposed to

Unknown Speaker 5:21

understanding societal values.

P56 5:27

Stakeholders who's at risk of experiencing impacts.

And one of the trade off between expected benefits and potential harms, or groups, not to mention.

P53 5:49

I feel like now the models that we chose are probably like very, I guess extreme in one way or another like it either has like a super high accuracy or Like very low disparity, I think probably to maybe satisfy everyone we need to reach some common ground I guess,

is where I see this task is going to

P56 6:17

between two and eight, at least when it comes to minimizing false negatives between two and eight. It has a lower false negative rate or like a higher false positive rate. So between those two things, two is like a more

balanced of those options for what password.

Unknown Speaker 6:42

I guess we could like, list out what

P54 6:48

aspects that we're all focused more focused on and then see if there's a model where it's more balanced right?

**P56 7:00**

It sounds like there's one person or in our case like two false negatives, but otherwise, there's one person for each of these bolded headlines, right like one case for accuracy for disparities, positives and false negatives.

So, one way to do preferences and I mean, I don't know how the bonus works is one voting role that often works is to see who's like a winner book a series of pairwise elections. So like between every like possible combination is there like a winner and then the winner who like wins the most pairwise elections is considered to be the strongest contender because push comes to shove, most people would prefer that so that's like a way to think about it if you want to just do a series of like to a comparison

P55 8:04

That's a good idea. Do you think we should do the comparisons between models or between like the headlines like accuracy of model two and accuracy of model?

Unknown Speaker 8:14

Seven or something like that idea?

Unknown Speaker 8:18

I guess maybe.

**P56 8:2**1

I don't know. I'm open to suggestions.

I'm here I guess, between headings is that if we do between headings, and we can kind of come up with like a priority, like a elicit an ordering of like, which one is most important to us? And then we can look at models based on our prioritization. Um, if we do it the other way. I guess it's kind of more ad hoc off the call. So maybe we should, but it's hard to because I'm like, is that you pointed out like some of these are very extreme. So Like, I don't know if my priorities should be super clean cut, but it might be worth it. I don't know. I'm open to either. I'm open to anything. I'm just throwing suggestions out there.

P54 9:15

We could also like, guess, I guess, like, throw out some models where they're like some really extreme numbers. Like, I think model number five has like a 90% false negative rate. Like things with

Unknown Speaker 9:29

fives out. Yeah,

Unknown Speaker 9:31

yeah.

**P56 9:32**

Yeah, that's a pretty song if you like this, the ones that really just want to kind of like, not considered because they're too extreme, like someone's definitely going to be unhappy with it. I don't think any more be like is what a great addition to.

P53 9:54

Okay, so we're done to seven models.

P54 10:05

Model one has like a 75 false negative rate. That's pretty high too.

Unknown Speaker 10:14

Yeah, I agree.

P56 10:22

Yeah, I guess model one is probably desirable among the false positive axes. But I don't see anyone using it as like their choice or anything. So I guess there's probably no real incentive to push for this one. Okay.

Okay. And seven, or one, I'm fine,

too. I like to. I voted for two because it has the second lowest false negative rate, and it has like no disparity and accuracy but Then it's overall accuracy is only 60%. I think the highest accuracy model is like 70% or something. Yeah. Or accuracy ones. But it's the false positive rate. Is it a little high?

Unknown Speaker 11:16

So

Unknown Speaker 11:27

first two,

P56 11:29

yeah. For to the false positive.

P54 11:32

Yeah. And for to like the accuracy is low compared to all the other models to

P56 11:38

Yeah, but there's no difference in accuracy between

Unknown Speaker 11:43

That's true.

P53 11:46

I guess I mean,

I am willing to have to out cuz I feel like we it's impossible to have zero disparities disparity, not Not that I don't care about disparity, but I feel like it's it's inevitable that we have some difference. And yeah, if two's not looking great for anyone else, I am willing to help you out.

P56 12:17

So the reason I'm not sure is because so I'm fine getting rid of eight. So model number eight has the same accuracy as model number two. So it up. Yeah, you and I are false positive rate,

lower false negative rate.

I don't think it's like worth, like I think two is better than eight as far as like a community member is concerned.

P53 12:44

I agree. It's definitely better than a

P56 12:48

prison I'm concerned is that if we don't go for two, I guess the one of the next lowest false negative rate is model six, which is like 35.6.

Unknown Speaker 12:58

I think models Four with a 2.1.

P56 13:04

But model four is false negative rate is much higher.

P53 13:06

Oh, sorry, I thought you were talking about this very sorry, sorry, my bad

P56 13:11

examples negative because that's like the community. So,

two and six have the lowest. But when you look at model six of all three is like overall like 35.6. And the false positive rate is like 30. I guess I could be OK with six if other people are okay with it. But I don't want to write to off just yet, in case. Yeah, because I don't I don't want to I don't think it's worth going up to like a 40 or higher. false negative rate, like the ones that are like 70 and 90 and like 60. That's just my current thinking.

P53 14:04

Yeah, i i agree i would definitely prefer um, for more because I guess I care about disparity and for has a okay. Accuracy I guess overall

P54 14:36

I think out of like already looking at two four and six right I would probably prefer six just because the accuracy is like higher than the other few. I would probably rank two is like my lowest as well and the disparity difference for life I guess for these three are all are they pretty? Well, I'm not sure I ended up with

P56 15:09

like 7.9%. So really kind of middling to is really two is really low four is pretty low six is kind of just like standard seeming.

P54 15:28

Yeah, I would be okay with like, four and six. Not so much

Unknown Speaker 15:40

because I look at the false positive

Unknown Speaker 15:42

rate and that's like, I think it's like the middle of

Unknown Speaker 15:49

the range.

P56 15:56

Got between four and six. I definitely think six is preferable. Because

Unknown Speaker 16:04

we go check this.

Unknown Speaker 16:13

Yeah, just like overall.

Unknown Speaker 16:22

I don't know.

P56 16:23

I also understand that the disparities are large, so I don't, I don't know, maybe I should. I guess if I if I just consider myself as a community member who only cares about false negatives, so then I probably would say that I prefer six.

P53 16:42

I think I'm just logically, I think six has a higher accuracy than four as well as a lower, false negative. So I guess I would also be okay with six Cuz 4.7 is sort of also in the middle of the other, the other accuracy, disparity and accuracy figures from the eight models, so I'd be okay with six

Unknown Speaker 17:22

or 2406 has lower false positive rate.

Unknown Speaker 17:29

Yeah.

P54 17:31

Yeah, I also like six

Unknown Speaker 17:33

too.

P54 17:34

So because of the accuracy and

Unknown Speaker 17:38

well, that's

P54 17:40

what my persona cares about. So

Unknown Speaker 17:47

cool. I guess we picked six six. Okay.

P56 17:57

Let's see. So we didn't really Discuss,

Unknown Speaker 18:03

like the.

Unknown Speaker 18:07

So just let you guys know, you probably need to go ahead and write the policy proposal. And I did he is the TA, so I probably need to leave around 11. So I need to recall, stop recording around 11 and IDT will answer any remaining questions you might have regarding this assignment. And also you need to finish a post survey and upload the documents to to Google Drive. And both of us pretend both of us are now here. So you can collectively write write your proposal and we're not listening. I promise. I'm gonna MIT now okay.

Unknown Speaker 19:04

So we decided on

P56 19:23

I guess let me just say like, Oh, this is Yeah, I guess Exactly. We're talking about like, Oh, we

has the best balance. I guess it's kind of like in the middle for everyone.

P53 19:38

It's not possible to have like the perfect, I guess model for every encoder. So

Unknown Speaker 19:48

it's

P53 20:22

Oh, I guess maybe we can each try. Maybe one reason why we pick this one rationale as they

Unknown Speaker 20:37

work sorting out

Unknown Speaker 20:51

the pros.

P53 20:56

Wait, I'm kind of confused. They're saying how make Sure to discuss the pros and cons of each of the models. Does that mean like in the short proposal or?

P54 21:20

Not sure. Is it? Is that only for if you don't agree on a lot of

Unknown Speaker 21:29

misery

Unknown Speaker 26:47

Is there anything else

Unknown Speaker 27:47

So guys, I'm just checking here so how about the writing progress?

Unknown Speaker 27:53

Do you need more time?

Unknown Speaker 28:03

Yeah, okay, cool. So I have to go to another meeting but I'll just leave the proper recording here in interest in case you have other like things to discuss, and Aditi will be here to answer any questions you might have. So after you guys finish the writing, please let editing know then you guys can move on to the post a survey. So, it should I think we are in pretty good shape right now. So everything should be able to be finished before 1120 Okay, so, so let me know if you have any questions and I will leave the recording here but I have to go to another meeting using another device. So good to see everyone. So please let me know if you have any questions regarding this assignment and anything gardening study. Okay, so I'll see you next time. Thank you.

Unknown Speaker 28:57

Thank you.

P53 31:36

Are we done with the writing? anyone needs to make more adjustments.

P54 31:47

So if you guys are done with the proposal we can move on to the post survey. I think it should be at the end of the canvas assignment. Let me know if none of you can see it

P55 31:59

or any of you can see

Unknown Speaker 32:07

Yeah, I see it.

P54 32:12

I see it too. Yeah, this is basically the last task. It's an individual assignment and it should take like 10 to 15 minutes so you guys can start working on that now.

Transcribed by https://otter.ai