Unknown Speaker 0:00

We are

P45 0:02

not talking about the topic at hand because you're waiting for somebody to get recording permissions. Right? Yeah, sorry about that. We have a lot of technical issue. Great. Um, so you will have a discussion and then you will have a video and audio after the meeting. So you just need to upload the audio. Okay. Have fun discussion. So do we also need to do the Google Doc now or do we do that after this? On You mean to assay writing? Yeah, yeah, you can do it like simultaneously. Okay, cool. Haha. You, thank you.

Unknown Speaker 0:37

Thank you. Okay.

P45 0:43

All right. Let me let me add a Google Doc to this chat.

P46 0:52

How much time do we have for the discussion?

Unknown Speaker 0:56

30 minutes. Okay.

P45 1:06

workgroup for a right now

I'm just going to send a Google Doc you guys can like I guess get started on which model you like the most.

P46 1:25

It says 20 minutes for discussion and 10 minutes for writing.

Unknown Speaker 1:54

Can I share my screen?

P45 1:56

I'll just put a record on the app to share your screen.

P46 2:01

Right. I mean, I'm just yet this checklist for facilitating a discussion.

Unknown Speaker 2:08

All right, cool.

P45 2:13

So which model Do you guys like the most?

Unknown Speaker 2:19

I,

P47 2:22

I like model two.

Sorry, way, is not model. Model three. Yeah, because I think model three has the highest accuracy. And the disparity is not very high. So the bias between different groups is okay to me.

P45 2:46

Yeah, that makes sense. I mean, I also was a fan of model three again, because the accuracy I think another thing that's kind of underrated about is that when compared to another accurate model, which I believe is model seven models, 769.1% Mall 370 point four. Model three also has significantly slower disparity inaccuracy between racial groups between white Americans and African Americans, which obviously is something that's pretty topical right now. So it has a pretty good ability to actually understand whether or not somebody is likely to reoffend independent of their race, which obviously is a huge issue that we're dealing with right now. And the colonel justice system, having kind of racist undertones at times. I will also say that model three is pretty great because it's dealing with a relatively balanced data set. So if we were to look at our EAM activity and the instructions, it said that the data set had 3000 people 1063 were charged with another crime within two years and 1937 were not charged with a crime within two years. That is relatively balanced as a data set in my mind. And so you don't run into the traditional issues of a data set that has accuracy as its main metric, in which you kind of assign a positive or sign I get for the whole data set and then luck upon being accurate. So, I think accuracy here actually means something.

P48 4:20

A Model T also has a really high false negative rate. So, I mean, I think overall, we need to have like a balanced model, which kind of has good rates. Also false negative would pretty much mean that we are expecting something to be not positive, but it is, right. So I don't think false negative rates can be high.

Unknown Speaker 4:52

Just wondering.

P45 4:54

So false negative rate. If we're interpreting this model would mean that you predict you're not going to refund them to someone that somebody actually does refund.

P48 5:03

Right? Exactly.

P45 5:04

Yeah. So if I were to look at like the societal costs, like a false positive versus false negative rate, I think it's really dependent on the types of crimes that you're diagnosing. So for example, if you're diagnosing crimes like petty theft, or shoplifting, things like that, then a false negative is not a very high cost for society to pay in exchange for overall accuracy. Because what you're basically saying is that we will potentially let some people slip through the cracks in exchange for being overall accurate and overall, allowing fewer criminals on the streets and allowing for your chances for for re offending.

The rate Sorry, go ahead.

P46 5:50

Sorry to interrupt I just like we're just to the point you mentioned, I guess. I mean, letting people slip through the cracks. Yeah. Like depending on the crime, like it could be okay. It could Not be, but I guess like what is worse is in terms of like having people, you know, like, be convicting people for things they have not really done. But just on like, you know, depending on the model or depending on some of the predictions that we're making that someone has a higher tendency to commit a crime we would be stopping, they can only be just putting them in jails, and when facilities, so I think false positive rate is also a major factor. And with model three, even though we have an accuracy, which is really high, and like that's really good, and like absolute terms, it's like it has an accuracy of 70% but the false positive rates would be 10%. So like, of all the people who we are putting in jails or in facilities like 10% of them would have like should not have been there. So think of I think I would like to like champion for model number five because like it has the lowest false positive rate of just 1.9%. And I like I completely agree like it has a high false negative rate as well, of 90.5%. But in terms of like people being falsely, like accused of something, they believe this is the best one. So like, if we compare this, like let's say, we do not have a model, which can predict, like all the people vote have, you know, been lead free? Like, there would be no consideration in terms of if they're going to commit crimes in future or not, but just that like, what they have done so far, they're being convicted for that or challenged for that, but like nothing in terms of future so I think the false positive rate, like a low false positive rate is very important for this. Yeah,

P47 7:48

I just I have a quick question here. Um, so if a person who will not a refund but is predicted to reaffirm or well we put him in to the jail immediately, or we just watch him closely, and depends on what Griffin,

P46 8:07

I think we would be putting them in jail immediately.

Unknown Speaker 8:10

Yes.

P45 8:13

Yeah. So the recidivism score is basically saying that, okay, so this person has offended, we're going to determine their sentence time based off their likelihood to reoffend. So these people, everyone that goes through this model has already committed one crime. They like they've already committed a crime, we're basically giving them a longer sentence based off whether or not we think they're going to reoffend after they get out of jail, right? That's false. positive rate it would be doing giving them a longer sentence without them deserving it.

Unknown Speaker 8:40

Thank you.

P45 8:41

Yeah. So model five is good. is pretty good. What I will say is that to bear significant societal costs and other areas. Yeah, so although you wouldn't be letting people letting innocent sorry putting innocent people in jail for longer than they need to be. At the same time you have trade offs and overall accuracy. You have as large, larger disparities in accuracy between the African American and white American population and you wouldn't model three, right. And at the same time, you are letting 90% of the people that are likely to refund, you are going to predict they're not going to refund which, which is kind of mind boggling, especially white Americans, you're basically saying 96% of them that we predict are not going to refund are actually going to reinvent. So at the expense of letting innocent people go free. You're also putting all of society at an increased risk, which is something that doesn't sit well with me, right.

P46 9:51

I think I think I completely agree with that. But I think we need to question what the role of this like this model prediction system is. It would be like trying to minimize, you know, the risk of more crimes happening. Right, like, like people trying to repeat their crimes. So that's like their primary goal. But we need to be very careful that it does not lead to something that so let's say, even if we have some, like a slight increase of correctly catching the people who might, you know, recommend the crimes, I think that is a win for the model. But like, we do not want to, like, you know, increase this so much that like the other, like the other factors are getting affected. So, I completely agree that 90% is a very high false negative rate. And there is also disparity between the different demographics. But, I mean, I don't know if this is like part of the model or like the part of the system like that we're supposed to discuss, but there could be like alternative measures, which could be in place to just try to stop crime or like prevent crimes separately from the model, I think like it is out of the scope of this model to be able to, you know, completely Stop trying to guess or like, you know, completely control crime, the crime is still going to happen. But there might be other ways to like stop it, rather than just try to increase the accuracy of this module. Okay.

Unknown Speaker 11:11

Go ahead.

P45 11:12

I was gonna say, I think this is a good segue for us to go into part two, which is like, who actually matters, like who the stakeholders are? So sorry, I didn't mean to cut you off. You can totally continue

P48 11:21

our I was just gonna say exactly the same thing. And having said that, I think we just need to decide what metrics are the most important to the society or the stakeholders as well as to have like a equal distribution, whether it comes to false positives or false negatives, while not compromising on the accuracy as well. So yeah, definitely jumping on to step two, but I did want to put out that model number two was also, at least to me looks like a good competitor just because while it has a higher false positive Read, it also takes care of the false negative rate and does not sacrifice too much on the accuracy. So yeah, just putting that out there.

Unknown Speaker 12:09

Okay. Um, so

P45 12:13

when we look at assessing these models, let's try to figure out who exactly the stakeholders are. Right? So I think first and foremost, are the actual people going through this model. So the attendance that the shred the defendants who are being assessed on whether or not they're going to reoffend, there are the judges. There is the prison system and like basically wardens that are one of them, that prison system, and then there's society at large, right. Are there any other stakeholders that you think would be impacted by this model?

P46 12:46

No, I think I think that pretty much covers most of them. Yeah, so then,

P47 12:51

yeah, yeah, I guess I think there might be larger community like white Americans And African American, so that they may think there's a parity or discrimination between the two different ratios.

P45 13:12

Now, that's a good point. So even when we look at these four groups of like judges, prison systems, defendants and their society, it's important for us to look at it through the lens of racism and thinking about it from a white American versus black American perspective, which is totally fair. I think that is a good perspective take, especially when it comes to the defendants. So if I look at it from a case by case basis, from a judge's perspective, I think a judge over anything else would want to make sure that things are accurate, right? I think a judge's job is to get the right sentence more than anything else, so making sure that the number of years that it gives the defendant is a reflection of how many much time they should be served. I think that is the number one metric that they would be using. For the defendants prison systems, a society at large happy to think about what you guys think would be good metrics for those what they would care about.

P46 14:27

I think what they care about, like, along with, you know, like the point to that they mentioned about in terms of different combinations of race, gender, age,

maybe

the model can be a little bit more detailed, and that, what do you think that will bring in more confusion in terms of the decision making? Like how important are these factors in terms of making the right decision?

P45 14:55

So, race definitely matters, right? Because I think there is a lot of institutional racism in the print system in terms of age, if I'm injecting my own personal bias, I think that churches would generally be more likely to give a longer sentence to someone who's younger, because they do see them as more likely of reoffending. Whether or not the facts would back that up, I'm totally unsure. So that is another component that we could definitely test for whether the models a just in some way. And in terms of gender, I think a more, at least a judge or society would also be more likely to predict men as being more likely to reoffend than women does. There's like a perceived danger of it again, not sure whether it's true or not. Yeah, I think these are things that we need to be mindful of, ultimately, these models, and the data that they're drawing from are reflections of society, right? So right, when they look and when they use their data and the sentences that they're recommending the base off sentences that were already given to people, certain races, genders, ages, etc. So that's something that we need to correct for and something that we already see a lot of in this model when it comes to the rates of what should prosecutes African American versus white Americans just very consistently prosecuting African Americans more identical

P46 16:11

just like just to add on one more point is that I think like by definition it says like having a database model or database system is bound to have some biases based on the previous experiences it has had like it does not know like the absolute goals or the absolute relations between things, but just that you know, if it has seen a certain demographic committing lot more crimes, it might be just a you know, more than like, the reason for that could be like, you know, we have more data from a certain time period or something like you know, which is more biased towards certain something, then it would lead to like incorrect prediction. So, I think like, if we are deciding things based on such demographic, demographics or different variables like that, we need to be very careful of like how balanced it is, like, you know, trying to like, do some kind of like augmentations and all There's different kinds of methods to just balance out data sets, because it needs to be from a more uniform perspective, like if it makes sense, like just to try to avoid any biases, which might be based on how much data and what kind of data we have.

Unknown Speaker 17:15

That makes sense.

P48 17:16

Having said that, is it safe to say that as far as survey scores and based on the model data that we are given? Is it safe to say that the accuracy for vite Americans and African Americans needs to be similar or on the same range?

P45 17:36

I would say I would say that a good model would be one that has high overall accuracy and minimal disparity between races.

P46 17:45

And what if, you know, like the disparity is like that, you know, like, let's say we have less data for one of the demographics and the model is just as unable to learn the right prediction for that demographic.

P45 18:01

To overcome institutional biases

P46 18:04

based like is it an institutional bias? Or is it a bias in the data set? You know, like, for example, let's say like for, just as an example, we have 70% white American in the data set and only 30% African American would like data points. So, like, it would be harder for the model to learn proper things for the smaller data set right. And we might see some kind of a higher inaccuracy or lower of accuracy in the in the in the test data. So then we will try to fine tune the model in a way that we're trying to match the disparity between the demographics and like, as a result, we might be having a like a like good our disparity values, but like, you know, low accuracies or even like, trying to learn incorrect or biased variables or features. So really, yeah, I feel like having like, you know, a number along with it, like for example, like we need to know how Much of what demographics are in the data set to make a better prediction problem.

P45 19:05

Yeah, before we start assigning what we'd like to have, it's important to understand what's realistic. Yeah, no, that's totally. So then if I'm just thinking through the actual metrics that we want to use them, so it sounds like we care about accuracy and having relatively similar accuracy, or at least Yeah, in the same ballpark accuracy between races. I'm not sure where we fall in the hole false positive and false negative piece. I mean, if I'm just thinking about it from a societal because there's two elements to it, right? There's what the defendant would care about, and there's what society care about. Like, for example, if I'm a defendant, I only care about false positives. I'm like, if I'm innocent, you cannot lock me up. That is always going to be zero. But if I'm society, All I care about is false negative, because I'm like, Okay, I need to make sure That I am safe, and that there's no criminals in the streets there and that you minimize the chances of a criminal being on the street. Right, that you make sure that you catch someone's I think we offend. Right. And I think realistically, we as a society are so distanced from the prison system and from criminals, that we disassociate ourselves from the impact that it might cause them, right. And so I think what we need to do is we need to make a bit of a stand on who matters more to us, when we're designing this model. Are we doing it in the perspective of society and maximizing the safety of society? Or are we doing it to maximize the accuracy for not accuracy, but minimize the amount of innocent people in jail? Which one is I think that's really what we need to make the call on. Because then we all agree that good accuracy actually high accuracy is important. I personally would really lean towards the false negative rate either That's more important. I think that false positive rate, obviously you don't want it to be ridiculously high. But I would be more likely to tolerate a false positive rate if at least a higher overall accuracy, I would not be willing to tolerate a false negative rate is super high.

Unknown Speaker 21:21

If at least that high accuracy

Unknown Speaker 21:24

I agree with you.

P46 21:25

I think a good way to make a decision is just following the point three, identifying the state stakeholders by which we can just have a pros and cons or you know the potential benefits and the potential harms for like either of these things if you want to go for like high false negative and high false positive. I think that would just help us make a decision also like in the in the final part, which is analyzing the impact Google Doc, or are we so one more question actually Are we are slick? Are we supposed to have separate proposals? Or? It's a group proposal? Right? Okay.

P45 22:06

Yeah. So on the group proposal, we just have to put down the model we choose and the rationale Exactly. If I'm going to these models, I'm gonna start with models that have high accuracy, right? That's from the start, and then I'm going to compare them on the false negative false positive rates. Yeah. Oh, one, three and seven. Have good accuracy rates, right? So model number one, if I look at it,

P46 22:31

good. I can argue even five years and six has great accuracy. How do you say it's just like one or 2%? Right? There you can say I mean,

P45 22:43

honestly, I think there's a one and 2% disparity, which no longer right, which is why I'm like looking at that overview sheet, where it has like the stars, and then seems accurate. You're basically saying 68% is kind of your cutoff for accuracy, which is Yeah, that's fair. So,

P46 23:03

um, yeah, I'm not sure that's completely fair. What would be your argument for that? Because, like, again, I guess, I mean, we said that accuracy is good for, you know, the, like the absolute scale. But I guess we still need to figure out like the pros and cons and the impact of that on the society before we make the final decision. And so we need to have like, at least not just three models, like slightly more. So like, I would recommend, just like from 10, let's get the top five. And on top five, we can have slightly more detailed discussion based on you know, what our metrics and what the impacts are, and then make the final choice. Finally, two choices, I guess, right?

P45 23:42

Yes, I'm fine with one three and seven. And then what would you add? you'd add six and five.

P46 23:48

I would add five instead of six, just because of the insanely low false positive rate. So yeah,

P48 24:00

Ready to go super high?

Unknown Speaker 24:01

Yes, it's Yeah. Well, we we do we did say that we care more about false negatives and false positives.

P46 24:06

I think we just need to balance out everything. Even if it means lower. I'm not sure if he if he agreed on like false negatives are higher, or like, you know, better rated than false.

Unknown Speaker 24:21

Right, welcome back.

P13 24:24

I hope you guys have fun discussing and negotiating with each other. Um, so now we're gonna wrap up.

So a couple to dues for you guys.

The most important thing is to follow the instruction to the end in the in the assignment on campus, and actually click on the Google Drive link. So when you click online, you should see something like what I'm seeing right now, basically a Google Drive. So one of the documents called team proposals So you should click on opinion proposals. The person who is responsible for writing the document or maintain document should put the link here. So I see already two of the rooms already put the link here, which is great. Another thing is, the person who is responsible for the recording should actually upload the recording to your corresponding room. So you already created the photos for each of the room. So it will take a while for you to actually compress the audio. But when when the audio is ready, just uploaded into these folders.

Lastly,

make sure you finish the post survey hopefully in the next couple of hours before the end of today. So we won't be able to actually finish the poster right now because the class is about to be over. Other than miss anything, how are you?

P48 26:05

I'm good. I'm good. So for those of you who have questions, again, feel free to stay on this link. And we will hang around for a few more minutes. And for rescue, it's the end of the class. And remember to fill out the survey post survey. And for each team, make sure that your team upload the audio file and put the Google Doc English in the drive.

P18 26:34

You know, go off to the bathroom, just putting my hand on is good.

P13 26:39

Yeah, got it. So yeah. Thanks, guys, for a great discussion for a great class. See you next Tuesday. Bye.

P47 26:52

I have a question.

P13 26:54

Yeah, go ahead.

P48 26:56

Are we going to discuss these models in the next class

Unknown Speaker 27:00

Yes, we will. Okay.

Unknown Speaker 27:05

This is Michael here got a question as well. Um, I'm having trouble on the canvas homepage for this course determining when quizzes are given work, find out because it'd be great if I can, you know,

Unknown Speaker 27:23

do a little bit of a review beforehand.

P13 27:28

Uh, what do you mean? Like you? You haven't submitted a quiz? No, I did submit the quiz on but I just been trying to figure out when the next quizzes are coming up. So I I have it like, I just can't find that on the Canvas page. Oh, the quizzes every Thursday. Right. Oh, okay. Yeah, we announced it in the first first week. So expect a quiz unless there are exceptions every Thursday, beginning of class. Oh, perfect. Okay, thank you. Yeah.

P18 27:58

Um, I used to have I have another technical question, but multicuisine will it be gone over by a person like, because I notice I got something wrong when I asked my friend who wrote for over 15 years, Grandpa question ever, for over bracket 11 plus four and that's wrong with somebody going in and out, or should I be concerned about?

P13 28:16

Like, most likely No. Yeah, so the short answer, most likely, no, unless there's a exception. We'll actually go over some of the points. Most likely, it's more like just like a check point to see whether you understand something basic.

P18 28:31

I expanded it and it's wrong if I wanted to ask whether I should be concerned about losing my spot. Just

Unknown Speaker 28:41

a thank you.

Unknown Speaker 28:43

Should we start recording, by the way?

P48 28:45

Oh, yes, definitely. Yeah, yeah, definitely. Let's Okay.

Unknown Speaker 28:48

Are you doing the recording? Or I think it's you. Okay, stop recording. Okay.

Unknown Speaker 28:57

Professor, the quiz will always be on On in at the beginning of the class is it the possibility to start it a bit earlier than the class because I found it very distracting to kind of be on the zoom call and then focus on the quiz. Especially because of

Transcribed by https://otter.ai