Interview

Dan: Okay. Now I'll share my screen. And the last step will just be, I have some questions to ask you and then we will be done.

Participant: All right.

Dan: Okay. So you can see the slides again?

Participant: Yes.

Dan: Okay, so my first question is about if you could talk about what your current approach is to using PDP or ICE plots. So I guess the first question is like in your work, do you or have you used PDP and ICE plots?

Participant: Yes, we have used PDP and ICE plots.

Dan: Okay. Do you use both PDP and ICE plots?

Participant: ICE plots, I think we use regularly. PDP, I don't think so.

Dan: Okay. Do you use both one-way and two-way plots?

Participant: One-way plots, definitely. I (unintelligible), I have not, I've never used two-way plots. I don't think they provide that much information.

Dan: Okay. So for, if we could dive into that, so why don't you feel like two-way plots provide much information?

Participant: Well, just looking, just looking at them, you can, well, from what I can see, they do show the interaction between the variables, but they don't show or they do not, you're not able to tell why the model is making those predictions from the two-way plots. Okay. So it doesn't give any more information than you would just plotting the two variables on a scatter plot, if you understand what I'm trying to say, against the ground truth.

Dan: Uh huh. Okay. So for, so for, what are your, so if not using two-way plots, what are your approaches for analyzing interactions?

Participant: Analyzing interactions. Well, most of the time we basically just use scatter plots to try and see what the relationship between two variables are. We could also use, what is it called? I can't remember. The name is escaping my mind at the moment, but we have actually never used two-way ICE plots to judge interactions.

Dan: Okay, so you mainly use scatter plots, where like you just like look at pairs of features and see the distribution.

Participant: See, yes, see distribution and if there're any correlations also.

Dan: Okay and is that like colored against the ground truth as well?

Participant: Yeah, of course the ground truths is what you are judging against. Or you could just try and section or shed the points against the ground truth as it's especially useful when you're looking at predicting, I'm sorry, classification task, where you can easily segment what the ground truth is and which points predict that value.

Dan: Okay, and apart from that, were there any other interaction techniques worth noting?

Participant: Not that I can recall at the moment.

Dan: Okay, so then for for going back to one-way plots, can you talk about what tasks you use them for?

Participant: One-way ICE plots, well, if you're trying to determine what the correlation between the model's prediction and the value of the feature are, the one-way ICE plots kind of show you what the general trend is as to the changing that feature's value or in relation with the prediction of the model.

Dan: Okay, so, so it seems like you could also like look at like a scatter plot that shows like the feature value versus like the ground truth, so like what value do you see in ICE plots over over that?

Participant: That's a good question. Well, I suppose I just kind of see see scatter plots as the go-to for seeing or checking interactions between two plots, sorry two features, whereas ICE plots are just kind of, they're well, they basically are used for just one feature alone or the way we use them, we use them for just one feature alone and of course, I know scatter plots can be used to, you know, analyze a single feature, but, at the moment we do not use them that way. It's just

kind of convention at this point. But in speaking of advantages, if I could just think off the top of my head, I'd say, with ICE plots, you would you would see the relationship between every instance in your datasets in a single glance and that might not be possible on a scatter plot.

Dan: Okay, and so, can you explain what you mean by that because like a scatter plot could also show every instance, right?

Participant: Scatter plots could also show every instance. Yeah, but for every for every for every instance, you only get, well, a single point on the coordinate graph, right? But with ICE plots, you get an entire line that shows how the model changes its prediction based on the the change changes in value of that feature, the of that feature, for every single instance, which you can't really tell that much information on a scatter plot.

Dan: Right. Okay, makes sense. So, for one way plots, how do you determine which plots to look at?

Participant: How do you determine which plots to look at? Well, I suppose the way we use them, we just kind of have already have a feature in mind that we want to investigate and we bring up the one-way plot for that. There are other ways we predetermine what features to investigate for that. For example, if you've trained a model on decision tree, you could use the feature importance value that model give gives out to you or a random forest model, for example, and then you use that for the investigate, so we use the ICE plot as a a tool that we we utilize on towards the end of our investigation or analysis of the model rather than a tool to determine what features to investigate.

Dan: Got it. Are there any pain points in your approach with working with ICE plots?

Participant: Specifically, I can't think of any, but I'd say, I'd flip that question and say I never thought being able to brush a section of the ICE plot itself and see what, first of all, what data points are in that section of the ICE plot and see also how they, how they are represented on the ICE plots of every other feature. I didn't imagine how useful that could be until, well, this demonstration.

Dan: Right.

Participant: So that's a pain point I never had, but I think this will be a useful solution for.

Dan: Okay, so moving on to the next questions are going to focus more about PDPilot, so can you talk about how did PDPilot support or not support your model analysis and like, were there any questions that you were unable to answer or any tasks that you were unable to perform?

Participant: Yeah, so like I've just stated there, I thought that the being able to, he brushing feature itself was particularly useful. It can help you better judge interaction between some features and also see why, what order, sorry, XXX XX X XXXXXX XXXX XX XXXXXXX, and also see for those data points how those particular data points fair on the ICE plots of other features, right? And yeah secondly, I thought the highlighted line similarity feature also helped with that. If you're not just judging it visually and you wanted to see, you know, which, especially with the Ames dataset there that had a lot of features and you just wanted to see which ones had the closest similarity. That was one thing I found particularly useful. I did not exhaustively use the tool because I thought the time allocation was just too little and for me to really get in depth with it. But in general, I think it's a pretty useful tool for analysis.

Dan: Okay, great. Okay, so just to recap that, so you found the highlighted line similarity was feature was useful, but you didn't feel like you were able to exhaustively use the use of tool since I guess like there just wasn't enough time for you to get in depth with it. Is that right?

Participant: Yeah, and also like the analysis of the model, I didn't have a goal in mind. No, I wasn't trying to, usually when I investigate a model you're trying to figure out, you know, answer a question basically. And you kind of tailor your analysis towards that. Just to broadly analyze a model, just felt like an open question that I wasn't sure how to answer. But in general, I try to see if there are correlations between any of the features, if there are any instances where the model was predicting, for example, a higher value. I'm sorry. No, no, no, when the PDP was decreasing as the value of the feature increases. Situations like that to see if there are any interesting things worthy of note in the ICE plots themselves.

Dan: Got it. So when you were doing your analysis, were there any, like tasks that you wanted to do or questions that you wanted to answer, but you weren't sure or were'nt able to to do?

Participant: No, I was able to do everything I wanted to do.

Dan: Okay. So were the visualizations useful and were any of them unclear?

Participant: Yes, I found particularly, I found the, the highlighted Instagrams quite useful. The ones at the top of the plots themselves. I found those quite useful. I think all of them were clear enough. If you follow the, if you read the documentation and also because I had a presentation earlier, which was exhaustive of all the visualizations, so I think they were clear enough.

42 **Dan:** Okay. So you touched on this a bit, but just to if we can expand more, what impact did the different rankings have on your model analysis and which rankings or sortings did you find to be the most and least useful?

43 **Participant:** Most or least useful. I'll have a look at the tool again, I think there was one of them, I just, I had one look at and I just thought, no, I'm not using this or I wouldn't need this. Okay. Oh, I see. So I, I, I'm just going to backtrack a bit on that statement. So it was sorting the two-way plots by variance instead. We did not touch on that during the presentation, so I wasn't exactly sure what it does. But I can imagine it to be useful if you really wanted to get into it. But I think first of all, right off the bat, ranking the plots by importance is the most useful cause usually that's what that's the question you're trying to answer when you are analyzing a model. What's the most important feature? Right. So you kind of see the features arranged according to the importance. That's quite useful. I thought being able to sort by the highlighted line similarity was also quite useful.

44 **Dan:** Okay. Are there any additional rays? Are there any additional ways that you think would be helpful to rank the plots?

45 **Participant:** Not at the moment. I think that felt quite exhaustive.

46 **Dan:** Okay. How did analyzing subsets or clusters of instances impact your analysis? And did you find the clustering to be useful?

47 **Participant:** I found highlighting to be very useful. Clustering itself, I didn't explore enough enough of, but I'd imagine it to be useful to if I had more questions to answer in my analysis.

48 **Dan:** Okay. So next, were the filtering capabilities useful to your analysis? Are there any additional ways that you would like to be able to filter the plots that you think would be helpful?

49 **Participant:** Yeah. The filtering capabilities were useful, especially for the two-way plots. If you want to see the interactions between two features immediately, of course, you'd have to filter. I don't think there are any additional ways that I can think of at the moment.

50 **Dan:** Okay.

51 **Participant:** The options that were available were useful were helpful enough for me.

52 **Dan:** So how well did the tool enable you to analyze feature interactions?

53 **Participant:** I didn't do a lot of feature interactions. I just had a brief look of, a brief look at how the the neighborhood feature was interacting with other features. And the one thing I noticed was that I couldn't get a lot of information. Yes, I was able to see what the interaction between the two features are. But either because I didn't dig deep enough or I just, the tool just didn't have that feature. I wasn't able to tell why that interaction is present and I suppose that might be a question PDP and ICE plots might not be able to answer on their own.

54 **Dan:** Mm-hmm. What type of information would help you answer that question about why the interaction was present?

55 **Participant:** Hmm. Good question. Let me see. So. Correlation. A correlation plot might be able to answer that question if you can see the correlation between the two features, but then in that case you'd be assuming both features are numeric features. If um, in in the case I was referring to in this dataset it was the neighborhood feature which is just a list of names of neighborhoods, so I'm not sure how you could do that. But with two numeric features I'd say the correlation should be should be enough to give you that information.

56 **Dan:** Okay, and so in this case like a correlation would be like like a scatter plot?

57 **Participant:** No, no, no. An actual correlation plot itself.

58 **Dan:** Okay, can you explain what a what a correlation plot is? I'm not sure if I'm familiar with that term.

59 **Participant:** Correlation plot between all of the features. So, yeah. What's it called again? Yeah, so you can have a correlation plot between all of the features that just plots each feature or against the other one.

60 **Dan:** Uh-huh.

61 **Participant:** You see how they correlate with each other. So, for example, if two numeric features are highly correlated they would have the highest score on the correlation plot.

62 **Dan:** Right. Okay, and can you explain like how would that help you? So if you know that there's an interaction between neighborhood and some other feature.

63 **Participant:** Yeah.

64 **Dan:** How would that, like how would the correlations then help you understand that?

65 **Participant:** Well, so if you if you have the, if you know that there's an interaction, you don't actually know what the interaction is, right? With correlation, you get a numeric value for what, obviously it's a rough estimate, but for what the, what the model thinks the correlation between that value. So it could be if for every, I don't know, increase in the, in one numeric value, you get twice as much increase in the other one. So it's just going to show for, for each of the instances in that data set, what the correlation between one feature and the other is.

66 **Dan:** Okay. So if I could just go. And which one are we looking at? I think it was, was it neighborhood and total basement area? Was it this one?

67 **Participant:** Yeah, neighborhood and total basement area.

68 **Dan:** So like we see that when, for the, I don't remember which, so for the Edwards neighborhood, with total basement area is above like 2.5k, we see a big negative interaction. And so I guess so you're saying is that you weren't sure how to understand like why is this interaction happening, right?

69 **Participant:** Yeah, basically.

70 **Dan:** Okay. And you think having some like correlation plot, where you're able to see how the features are correlated with each other would, so in this

71 **Participant:** So in this instance, in this particular instance itself, I don't think that would be able to answer the question. But if there were numeric features, you'd be able to tell the relationship on the correlation plot.

72 **Dan:** So is this what you mean? Like, is this what you mean by like a correlation plot?

73 **Participant:** Yeah, something, something like this. (unintelligible)

74 **Dan:** Okay. Okay, so I guess like, so. Well, our intention is with the tool is that like here, like you see that there's like this big dip and then we can look at like this plot and we can see that there're like these three data points here, which have high total basement areas, but are like very low, low priced homes. So we have one here, one here, one here. And it seems like it's those three data points, which are dragging down the model's prediction for this particular neighborhood.

75 **Participant:** Yeah, I actually did not look at that when I was analyzing the model myself. Oh, well, in that case, I suppose generally there's more information on this plot than you would have on your correlation plot. (unintelligible) I was just saying earlier, what I was just trying to say earlier was that. Well, I assumed that that would give you more information on how the two variables are related to one another, but now that you explained, I can see how you'd, you'd be able to tell directly from the two-way plots.

76 **Dan:** Okay. And then I guess if we could just go to another example, maybe one that has like a, so I think maybe you looked at this one. So, when like trying to understand this interaction, could you just explain once more like what what additional information would help you to understand like why we have this weird interaction for above ground living area above 5k and basement finish area above 1k. It's like, would there be any additional information that would be useful to like show on the screen to help you understand why that interaction is happening.

77 **Participant:** If you look at the chart on the right and you have a look at the the, you know, just the individual points on the graph themselves, you should be able to tell, well you should be able to deduce that those outliers are probably causing the the reduction or that negative interaction between the two features. I don't think there's anything else that I can think of right now (unintelligible).

78 **Dan:** Okay.

79 **Participant:** (unintelligible)

80 **Dan:** Okay. Okay, so I think we can move on from that question. So I guess like now that you like, see like these plots and how they can be used, you think that, like you think that does give you enough information to like start try to understand why interactions are taking place or why they are present.

81 **Participant:** Yes. Yes. Well, there's a lot a lot of information on the screen that I must say.

82 **Dan:** Right.

83 **Participant:** But if you take your time to look at it and try to figure out what exactly is happening, you'd have enough information to to tell why that interaction is happening.

84 **Dan:** Mm hmm. Okay. Yeah, but that's a good point that there's a lot going on on the screen so I can definitely understand that like there's like a bit. it's kind of overwhelming the amount of data that's being presented, so that's a good point. Okay. So what would you say are the tool's biggest weaknesses or limitations and are there any improvements or additional capabilities that you would want the tool to have?

85 **Participant:** Hmm. Limitations. Well, I thought the computation was, for how much information it was computing, I thought it was quite fast. That's one thing I first saw might be a weakness in the library itself.

86 **Dan:** So in this case, that was because, so that JSON file that you downloaded, that had like all the precomputed PDP and ICE plots.

87 **Participant:** No, I, I understand that. Yes. But I also tested the library like 30 minutes before I joined the this call and I followed the documentation on the, was it on the GitHub page on one of the links you sent me and I ran a test myself, just to get to the widget but I didn't, I did not go through any of the plots.

88 **Dan:** Okay.

89 **Participant:** Yeah. And I thought the computations were fast enough, especially on my very slow system.

90 **Dan:** Do you, do you remember which notebook you ran for that?

91 **Participant:** I made my own notebook. I can share it with you if you want to have a look.

92 **Dan:** Yeah. That would be great if you could.

93 **Participant:** Just give me a second to get it up.

94 **Dan:** XXX XXX XXX XXXX XXX XXXXX XXXX XXXXX?

95 **Participant:** XXX X XXXX X XXXXX XX XX XXXXXX XXX XXXX X XXXX XXXXXXX XX XXX XXXXXX XXXXX

96 **Dan:** Got it.

97 **Participant:** XXX X XXXXXXX XX XXX XXXXXXXXXXX. Okay. Yeah. Am I, are you seeing my screen?

98 **Dan:** Uh, right now I am not.

99 **Participant:** Okay. I'll share my screen. All right. Can you see this?

100 **Dan:** Yes.

101 **Participant:** So this is, um, I use the iris flower dataset...

102 **Dan:** Uh huh.

103 **Participant:** to compute the PDPs.

104 **Dan:** I see. Okay. And you ran this in Google Colab. Got it.

105 **Participant:** Yeah. And I ran this in Google Colab. Actually, I should point out I had to do, well, I suppose it's not, um, it's a Jupyter widget thing, but I had to add this extra line of code to enable support for Jupyter widgets in Google Colab.

106 **Dan:** Mm hmm. Yes. Okay. Cool. Um, so now I'll return to the presentation, but thank you for for showing that.

107 **Participant:** No problem.

108 **Dan:** Um, okay. So I guess going back to the question, uh, were there any, let me bring it up. Uh, so were there any weaknesses or limitations that you saw or any improvements or like additional capabilities that you would want the tool to have?

109 **Participant:** Additional capabilities. (unintelligible) To be honest, I think it's quite well rounded, however, the only thing I would say, if it was possible, if you could eject the, the widget into its own window.

110 **Dan:** Mm hmm.

111 **Participant:** So you don't have to manually resize every time in code, you can just like throw it on another screen and

have a full screen version of it where you can play around with the plots.

112   **Dan:**   Got it. Yeah. Okay. Uh, but no other weaknesses that you would want to mention?

113   **Participant:**   No, that I can, that I can think of right now, but, um.

114   **Dan:** Okay.

115   **Participant:** I could send an email if anything comes to mind.

116   **Dan:**   Yeah, absolutely. Um, okay. So this is the last question, just like if there any other feedback that you haven't already mentioned that you want to mention, um, this is a place for it.

117   **Participant:**   No, I think I've said it all.

118   **Dan:**   Okay. Um, so thank you very much. I'm really sorry that this went over time. I'm sorry with the, the slowness in getting this installed. Um, but I'm glad that we were able to complete the study. Um, so either tonight or tomorrow morning, I will send you a $100 Amazon gift card. Um, so, uh, I'll send it to the email that you provided, uh, in the form. Um, so as you know, PDPilot is open source and installable through PIP. Um, so if you're interested in using it for your work or with your own data and model, then I'm happy to answer any questions or help you get set up. Um, also happy to make any changes to PDPilot to support your needs. And I'm always interested in hearing about your experience, your findings and your feedback. Um, but overall, just thank you so much. I know that that two plus hours is a lot of time and I', really grateful for your generosity with your time and helping with this work.

119   **Participant:**   (unintelligible)   Thanks, Daniel.

120   **Dan:**   All right. Have a great night.

121   **Participant:**   You too, bye.

122   **Dan:**   Bye.