1 Interview

2 **Dan:** Okay. So now you can stop sharing your screen and then we'll just end up with a few questions that I have for you.

3 **Participant:** Okay.

4 **Dan:** Okay. Okay. So my first question is if you currently use ICE plots in your work?

5 **Participant:** Ah, yeah. Not all the time, so it's not the main thing we do, but in our initial data explorations, we tend to use ICE plots.

6 **Dan:** And do you use both PDP and ICE plots?

7 **Participant:** Yeah, we use PDP and ICE plots.

8 **Dan:** And do you look at both one way and two way plots?

9 **Participant:** Yeah.

10 **Dan:** Okay. Can you describe what tasks you use them for?

11 **Participant:** We typically use them to identify trends and to identify nonlinear interactions. And for this, instead of the color squares (unintelligible) so we plot the average against the average. Or we put all this. Or we do a scatter plot. Or we put the average against from one feature against the average versus the average against the other feature.

12 **Dan:** And that's the average of the average of what?

13 **Participant:** Let me see. So it would be. For if you have feature one at one value and the average and feature two at another value. So you will get the average value at that point. So that would give you one. So what is it that we plot? Let me remember. We tend to use more of the scatter plots for this. But we had lines on top of how it worked out. I forgot how we calculated those. Yeah. No, I forgot. So I think I'm maybe confusing something, but we tend to use more of the scatter plots to identify if we have some non-linear dependence on the interactions.

14 **Dan:** Okay.

15 **Participant:** Yeah.

16 **Dan:** Okay. Yeah. Sure. I mean, if at some point you remember or have an example you can show me, that would be very helpful. How do you determine which plots to look at?

17 **Participant:** Well, precisely what you showed me. So sorting by importance or by difference. In the case of histograms, we also look at earth mover’s distance, for instance, as a metric for the difference between histograms. And then we sort by that or by feature importance or something like that. And then we plot them so because if you have, I don't know, 30 or 40 features, you end up with hundreds of plots. So you sort and then you look at the first 40 or so (unintelligible).

18 **Dan:** And so for that mover’s distance for histograms, what are the histograms that you're comparing?

19 **Participant:** So we bin the features, right? And then we compute histograms. And I'm thinking, yeah, I asked you for the question. I asked you that question about what metric you're using because I think that earth mover’s distance is the same as the way you're computing because if you have just one-dimensional histogram

20 **Dan:** Mhm.

21 **Participant:** I think that your definition is the same as the earth mover’s distance. But we compute a multi-dimensional histogram, so we bin in many dimensions and then you get points in different (unintelligible) a multi-dimensional histogram and for that, you need the earth mover’s distance, the complete definition. So you normalize the histogram and then you compute basically how many points you have to move from one histogram to the other one to change it to make it look like the other one. So from histogram A to histogram B, you move basically. I think in 1D it is the same as the sum of the difference in each bin. So it's.

22 **Dan:** Okay. So you would have like 2D histograms where you would bin the features and then compute the histograms. Now you have two different 2D histograms. And you compute the difference between those two?

23 **Participant:** Yeah. Yeah. So like if you have like, we do by clustering or we do it by, we select one categorical feature and you have two different sets of points and then you do histograms on those on all the other features and then you compute the distance between all these two, these two multi-dimensional histograms.

24 **Dan:** Okay. I see. So you would choose a categorical feature and then you'd have one histogram. So then you would, then for each other feature, you would compute one histogram per category, basically.

25 **Participant:** Yeah. Yeah. But we use it more with clusters, so if we have like 10 clusters or 8 clusters, we compute the distance between the clusters. Or if you select some points, you compute the distance between those points and the rest of the clusters. So each cluster is basically a histogram.

26 **Dan:** Right.

27 **Participant:** And you compute how much energy does it take to change the selected points to the other ones. So this is what the earth mover's distance gives (??) you.

28 **Dan:** Okay. I see. And so is this, so when you, is the purpose of this to like understand the interactions, so like how is this. Or like, I guess what, so what do you, what do you learn by, by comparing those histograms?

29 **Participant:** Yeah, it's not just the interactions, it's basically what's driving the difference. So trying to understand what feature is driving or causing the difference between the predictions.

30 **Dan:** Okay. So like if you have clusters, you want to know like what features are driving these clusters being separate.

31 **Participant:** Exactly. Yeah.

32 **Dan:** Got it. Okay. And what would you say are like, if any of the pain points in your approach to using PDP and ICE plots.

33 **Participant:** Yeah yeah yeah. So the pain point is we usually create all these plots by hand, so we have to copy pasting old notebooks trying to recreate the plots that we used to use and stuff. So we don't have a (unintelligible) tool like PDPilot, so I think it's nice. So the pain point for me is trying to recover all of the code that created the plots.

34 **Dan:** Mm hmm.

35 **Participant:** And then adapting it to the new dataset and so on.

36 **Dan:** I see. Okay. Um. So next, so now I'm going to ask you questions about your experience with using PDPilot. Um, so how would you say that PDPilot supported or did not support your model analysis?

37 **Participant:** I think it totally helped in the, um, like I said, for me, creating all these charts is usually done by hand. So having a tool that already creates these differences and it makes it very easy to change from one type of view to the other one and recreate that one plot and then brushing or highlighting some and computing differences to the highlighted data set. That's very, very helpful because otherwise when I see something interesting in a chart, and I want to select some of the points that are interesting in that chart, it takes a lot of, you know, handmade code, so this makes it very, very easy.

38 **Participant:** Um. I think by the end of the analysis, I realized that I didn't ask many questions that I wanted to answer. So I wanted to sort of explore the data to see if I saw something interesting. I think this data set of the, of the house in practice is not really amenable to that type of analysis. I should have done more questions beforehand. Like, okay, I think this is, you know, I've heard that remodeling the bathroom is a good way to increase the price. So maybe let's check on that. I should have tried to see the little more hypothesis testing. And I think it was, it was, it should have been good for that.

39 **Dan:** I see. Yeah. I was going to ask these questions. So are there any questions you're unable to perform or tasks you were unable to perform?

40 **Participant:** I think I. So I wasn't. I think I got stuck with trying to take out these outliers that were driving the model down (unintelligible) So I think that those outliers should be taken away and would maybe make the analysis easier because some of the importances were sort of biased with that. And I would have like to say, okay, let's take out those samples, but I think that requires the recomputing all of the plots, right?

41 **Dan:** Right.

42 **Participant:** Okay. So it would be nice to identify some. Well okay, maybe if I did the code for this, it would be easy because I could just say, okay, anything with any area above 4000, I don't want it in the samples.

43 **Dan:** Right.

44 **Participant:** Yeah.

45 **Dan:** I see.

46 **Participant:** So, yeah. If I had written the previous code, I would have maybe done it. And then, yeah, I think I would have been not stuck with that question.

47 **Dan:** Mm hmm, Okay. Um. Okay. So for the visualizations, were they useful and were any of them unclear.

48 **Participant:** Yeah, I think they are very useful. I don't know if it's, so when you see too many variables, some of the residual histograms are too small to see, actually. And you have to go back and forth to the detailed plot to actually see what's going on.

49 **Dan:** Mm hmm.

50 **Participant:** But the rest I found very clear and useful.

51 **Dan:** Okay. So how, how would you say the different rankings impacted your model analysis and which ones did you find to be most or least useful.

52 **Participant:** For this model in particular? The uh...

53 **Dan:** Uh. yeah, yeah for for for what you when you were using it.

54 **Participant:** Yeah, yeah, yeah. I think the importance. The first one is the one that I like the most. And then the highlighted line similarity. I found that the most useful. I was trying to think of a way on how to use the histogram difference, trying to, but I couldn't. I thought it was a very natural and useful difference when you showed it in the training. But I couldn't find a way to use it in this analysis. I was trying to think of how should I use it here. I couldn't find it. So I, it's not that I found it not useful because I thought it was going to be useful, but I couldn't make it. I couldn't use it here.

55 **Dan:** Mm hmm. Okay.

56 **Participant:** Yeah

57 **Dan:** Are there any additional ways you'd like, that you think would be useful to rank the plots.

58 **Participant:** Yeah, I was thinking that maybe a couple of filters, like if I want to see these plots, see these importance plots or (unintelligible) plots, but without some ranges of variables or including only or not including some variables.

59 **Dan:** Okay.

60 **Participant:** Yeah, I was, at some point I was thinking that if I selected all of the variables and then unselected some of them, I would have done that filter, but that's very uncomfortable because this dataset has so many variables that it would have taken a lot of time to.

61 **Dan:** Right.

62 **Participant:** Select all would be nice to do that.

63 **Dan:** Okay. So, like, for example, like you could select the range that excluded those outliers for the total area and then.

64 **Participant:** Yeah. Or or.

65 **Dan:** Re-compute the metrics based on that range. Yeah.

66 **Participant:** And then on the name of the variables, if I had a select all button and then I could unselect the variables that I didn't want. Sort of the opposite of selecting just a couple, I wanted to unselect just a couple.

67 **Dan:** Mm hmm. Right. I see. Yeah. So they're there. You can select all for. Say, for one way plots, you actually, well, I guess in this case, I see, I see. Yes, you're right. Because the default is them not being selected. So there's no, no easy way to select all of them. Okay. I see. But then even for the two way plots that wouldn't necessarily help because it would still show up. So I. Yeah. Okay. So a way to exclude a feature would be useful. Okay. Oops.

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

69 **Participant:** Yeah, in the end, yes, because I found this, this cluster of cars with more than three for space with more than three cars. Because at the beginning, I (unintelligible), I didn't see anything. But by clustering at some point, I found that and I thought it was useful. And then I was able to drill down a little bit more by highlighting those and seeing what

was happening. And then when I looked at only the highest prices cars, houses, what they have, so yeah highlighting was very useful.

70 **Dan:** Okay and what about the, like, the, like, the clustering and viewing the cluster descriptions?

71 **Participant:** Yeah, I think I saw a couple of points. I didn't do that much, but for the very expensive houses. That's where I saw some of the features that were, I found it useful because that's when I was. Yeah. That's when I found that some had, you know, more than three cars and the large, the large area. What's it called? The second floor area had to be about the certain ratio to be important.

72 **Dan:** Okay. So when for that part you're talking about, was that for a visualization? Like one of these, or was it for like this one?

73 **Participant:** In that one, in the, in the cluster description.

74 **Dan:** Okay.

75 **Participant:** Yeah

76 **Dan:** Great. Okay. And so were the filtering capabilities useful for your analysis? And are there any additional ways to filter the plots you think will be helpful, beyond making it easier to exclude a feature or a range of a feature.

77 **Participant:** I don't know. I think it. No. I don't know if I have any additional way to do something.

78 **Dan:** Okay. Did you find the current filtering useful?

79 **Participant:** Yeah. Yeah. I thought it was useful for exploring and taking out some ideas especially.

80 **Dan:** Okay. And then so for feature interactions, how well did the tool enable you to analyze the interaction?

81 **Participant:** I thought it was good. I don't know how much the data set helped (unintelligible). So it's not a data set with very complex interactions. So either flat or linear. And I didn't see. I didn't find anything very obvious that wasn't because of a few samples, basically distorting the chart. But I think it's good. I don't know how well. So the. The sorting work. So I went through the options and. I'm not really sure I got something out of that, but again it could be the dataset.

82 **Dan:** Which sorting was this? This was the sorting by interaction?

83 **Participant:** Sorting by by variance or interaction. Like I imagine it can be useful, but maybe it wasn't for this data set. So. Yeah.

84 **Dan:** Okay. And the reason for that would tie into what you said before with the interactions being obvious or simple or just due to like, few samples. Okay. Okay.

85 **Participant:** Especially for for models like XGBoost. We tend to use them, especially when we have non-linear interactions in a data set.

86 **Dan:** Mmm hmm.

87 **Participant:** Because if everything is very linear, you're better off using some linear model. It's very simpler and explainabilty is better.

88 **Dan:** Mm hmm. I see. Okay. So now the last two questions about PDPilot are like what you see as the biggest weaknesses or limitations. And are there any improvements or additional capabilities that you would want PDpilot to have.

89 **Participant:** Hmm. I thought I find it very good. It's complicated to see or to remember where is everything. So there's a lot to unpack in PDPilot. So remembering and with little experience, remembering where was one plot? What could I do here? It's not easy to remember, but it's not that it's wrongly ordered. It's just that. So that, okay, I saw some type of plot. Where was that? So thinking about that is not obvious at the beginning. And. So additional capabilities I would like to see is maybe some export feature. Like if I wanted to export plots directly without screen capturing. It would be nice. Or maybe I was thinking when we were talking maybe some checkpointing or saving something that if I see something interesting. And I want to keep looking for other stuff. Yeah, maybe saving some state, like to recover that state and go back to some analysis to show then to the rest of the team.

90 **Dan:** I see. Yeah, that makes a lot of sense. Okay, so now this just to catch all if there's any other feedback that you didn't get the chance to mention before.

91 **Participant:** No, no, no, I think I think I'm good. I made some notes and I think I said everything.

92 **Dan:** Okay, so that's it. So thank you very much. I know that two hours is a lot of time to. To spend some really appreciative of you. Being willing and so generous is your time. So we'll send you a 40 on Amazon gift card soon. So you know the pilots open source and soluble through PIP. So if you have any interest in like a plan to your work or like needed to like need help and like getting it set up with. Like your own data. I'm happy to answer any questions or to help you get set up. I'm also happy to like make any changes to better support your needs and always interested in hearing about your experience and finding any feedback. So thank you. Thank you so much.

93 **Participant:** Thank you. Okay.

94 **Dan:** All right, have a great night and I will be in touch soon with.

95 **Participant:** Thank you.