

Interview

Dan: Okay, so now we will end wrap up with an interview. So you started you touched on this. A bit earlier during our meeting. But could you talk about what your current approach is in your work for using PDP or ICE plots. So I guess the first question would be. Do you use both PDP and ICE plots? I think earlier you said you just you just use PDPs, but if you want to expand then.

Participant: Yeah, yeah, the common approach in our team is just, well, train a suitable model. It's mostly random forest and then analyze the behavior of different features using PDP. Well, after feature importance, so we first we get top features. Mostly top 10 or top five features plot their importance and then go and see what are the hidden meaning of of the features of the top features. Like (unintelligible) if you if you have two most most useful features, we go and analyze them.

Dan: Okay, so is there a reason why you your team uses PDPs, but not ICE plots?

Participant: There is some reason, I guess, but it's I guess cultural thing and influence of my team lead. And I've been like following this approach, not not really thinking about it.

Dan: Okay.

Participant: I guess I guess ICE plots are mostly feel more like cluttered like messy.

Dan: Okay

Participant: In cases of big data set, they tend to be very, very, very messy.

Dan: Right. So you mentioned that you use feature importance to get like the top five or 10 features. How, what do you use to calculate the feature importance? Are you using like what is provided by like the random forest or using some other method?

Participant: We are using. So (unintelligible), one second. Permutation importance.

Dan: Okay, so you use permutation importance to compute the feature importance for.

Participant: For for the we calculate permutation importance for each feature and then plot it on the, on the graph and see it like top 10 of them.

Dan: Okay. And. When you compute the PDPs, do you use the entire data set or do you use a subset of. Like some like sample of instances.

Participant: In my practice, it's mostly all the data set. Because we have a pretty large server server. Uh, which is, which is kind of computationally powerful. And, uh, if my data set is not like millions, millions of. Uh, of cases, I, I can safely and, well

Dan: Right.

Participant: Pretty. It's pretty speedy for me to. To calculate PIMP For all the data set. Uh, well, for validation, but.

Dan: Okay. Um. And so then. You mentioned, so you use the PDPs to. Uh, I guess you said you said see the hidden meaning of the, of the top features, right?

Participant: Yeah. Yeah. So, uh, we look at the PDP graphs and. Uh, try to answer the question. Uh, how this feature influences the output of the model and. Uh, try to explain in in few words.

Dan: Right. Um, so are there any pain points or difficulties in, in your approach?

Participant: Uh, I guess. I guess as you pointed out, uh, the. mmm. The big, big drawback is that, uh, if we have. Uh, pair. Uh, if two features influence our results in pair, we do not. Uh, we do not see that. And also. In cases where. Uh, for most of the. Uh, for most of our validation data set, we have. Uh, one particular influence and we have a, uh, we have other, like. Uh, subset of features where the. Uh, behavior is. Uh, different. Uh, we do not see that because we like, uh, see the. Average picture.

Dan: Right.

Participant: So maybe, maybe we should. Uh, start exploiting ICE. Looking more into ICE.

24 **Dan:** Do you have other approaches besides PDPs that you use for model explainability?

25 **Participant:** Uh, well, this is the most. Uh, most used in our team and the next one should be. Uh, can't remember it now because, uh, haven't used it in a while. Uh, well.

26 **Dan:** Okay, so. Um, but it sounds like. So on your team, you don't look at. Interactions between features.

27 **Participant:** Uh, yeah, yeah. It's, uh, it's mostly we have. Uh, we look at the overall model's behavior. Uh, on this feature on one particular feature. And, uh, we. We tend to go, uh, and deal into. Uh, particular cases that. Explain. Uh, that explain why. Uh, what what's going on in this case? Uh, we, we go and look at the, this, this data point with all the features. Uh, for, for this particular case. And, uh, in some cases, when we have a very interesting predictions, a very interesting behavior. Uh, we might go and use. Uh, and use waterfall charts for. Uh, analyzing. Uh, the influence of the, of the features. Uh, and it's quantity. Uh, in particular cases, not in, uh, in the average for all data set. But for, uh, particular data points.

28 **Dan:** Okay. Um, so for that, um. How do you, how do you determine what cases to, to dive into? Or what, what, like individual data points you look at?

29 **Participant:** Uh, well, mostly to derives from a value. Uh, value is, uh, a value is, uh, where. Uh, we can get to, uh, less errors. Yeah. And so, uh, mostly we go and explore. Uh, top, top five, maybe, top 10. Uh, are worst predictions.

30 **Dan:** I see. Okay. And then for that waterfall chart. Um, is that showing like, SHAP values or something similar?

31 **Participant:** Yeah. Yeah. It should be Shapley values.

32 **Dan:** Okay. Um, and then last question. So for, for PDPs, you just look at one-way PDPs, right? You don't look at two-way PDPs.

33 **Participant:** Yeah.

34 **Dan:** Okay.

35 **Participant:** Uh, yes, it's, it's, uh, well, uh, our approach for now.

36 **Dan:** Okay. Okay. So then I have some questions about, uh, your experience with using PDPilot. Um, so the first question is, how did PDPilot support or not support your model analysis? Um, so like, were there any questions that you were unable to answer or any tasks that you were unable to perform?

37 **Participant:** Uh, well, I guess, I guess no. Um, there are, there were some. Uh, some, uh, some unclear. Uh, some, some issues with navigation. Like, uh, if we need to sort uh, by the difference between clusters and the, uh, mean, uh, PDP plot. Uh, what should we do? But I guess, uh, it, uh, it may be that. Well, I did not, uh, dive deep into into this, uh, into this service. So I guess, uh, I guess it's, it's a matter of, uh, experience. Uh.

38 **Dan:** Okay.

39 **Participant:** Where there any, any.

40 **Dan:** Sorry, go ahead.

41 **Participant:** Any, any tasks. Uh, I wasn't able to perform well, uh, between those you asked me to do no. No. Uh, when talking about a particular, uh, in cases where, uh, I think, uh, there's, can can be, can be some, uh, some things to look into, I think, uh, maybe, uh, business user would like to maybe dip into particular cases. Well, like, uh, we have a client with, uh, strange, uh, behavior. Uh, can you please, uh, Mr., Mr. business user go and look, uh, what what is up with, uh, this client. Uh, but, uh, I think it's, uh, uh, I think it's, it's a great instrument overall.

42 **Dan:** Okay. Um. Okay. So I guess just to summarize that, um, so you felt like you had some issues with navigation and some confusion over like sorting by cluster difference. Um, and so you didn't dive deep into that, but felt it was mostly a matter of experience. Um, and that there were some cases in your exploration that you weren't able to understand. Um, and those would be cases where you would want to involve, uh, like a business stakeholder to get their, um, opinion on it.

43 **Participant:** Yeah. Yeah.

44 **Dan:** Okay. Um, so for the navigation confusion, are there any particular, particular, like, changes to the user interface that you think could be. Helpful to, like, uh, make you more comfortable with the navigation.

45 **Participant:** Oh, uh, I don't know. It's, uh, I think it's, uh, it may be that the system is, uh, looks, uh, at the beginning looks pretty complicated and like, uh, like a Swiss knife. Like, uh, many, many great features. Uh, but, uh, you understand the,

uh, value of them, uh, with some time. So it, it may be, uh, not intuitive. Uh, at the beginning.

46 **Dan:** I see.

47 **Participant:** Um, but nothing comes to mind in terms of, like, particular changes that could help make it more intuitive.

48 **Participant:** I guess. No. No.

49 **Dan:** Okay. Uh, okay. So before we move on, any other things to note in terms of, like, how it supported or did not support your, your your analysis.

50 **Participant:** Uh, well, uh, overall, it's, uh, it's a great experience. Uh, I would like to know that, uh, note that, uh, with, uh, PDPilot. It's, it's a lot easier to explore the, uh, hidden, uh, hidden dependencies. Like, you do not, uh, do not need to write code. Uh, you just, uh, you just click and, uh, like, select, uh, things that you want to explore and, uh, this feels a lot like, uh, AutoML. But, uh, like, auto model interpretation, which is, which is really nice, really nice. And Uh, um, I think, I think if, uh, if, uh, licensing, uh, like, well, I'd like to try this at work. I'd like to try this at work. I don't know if it's allowed by, by the, uh, terms of use, but, uh, this looks, uh, pretty pretty great. And it felt, uh, felt very, uh, very, like, uh, so, so much easier than, uh, what I'm doing. Like, on a weekly basis.

51 **Dan:** Okay. Great. Um. Okay. Yeah. The, so the open source license is, uh, I believe it's MIT or something similar. So it's some permissive open source license. Okay. So next question, uh, were the visualizations useful? Uh, were any of them unclear?

52 **Participant:** Um, yeah. Uh, they were, uh, useful, uh, 100%. Uh, if they were unclear, uh, I think it was okay. Um, I feel like, uh, if I had more time with this instrument, I would be, uh, it would be. Uh, easier for me to exploit all the, uh, possibilities and instruments. So, uh, it has.

53 **Dan:** Okay. Um. So what impact do the different rankings have on your model analysis, uh, which rankings did you find to be the most useful or the least useful? Um, so for this, it seemed like you mostly just used the, um, importance ranking.

54 **Participant:** Yeah. Yeah.

55 **Dan:** Um, so was there a particular reason why you chose to use importance and didn't look at the other rankings or other sorting metrics?

56 **Participant:** Uh, well, for me, it's, uh, I think that's, uh, that, uh, it's, uh, it's not a choice to go into the most important features first and, uh, explore how they influence the model and then go into other sort of features because, uh, well, the more important the feature, the more value I can extract from it.

57 **Dan:** Okay. Um. Okay. And are there any additional ways you think could be helpful for ranking or sorting the plots?

58 **Participant:** Um, well. Um, no, for now, nothing comes to mind.

59 **Dan:** Okay. Um, so how did analyzing subsets or clusters of instances impact your analysis? Was the clustering useful? Was the highlighting useful?

60 **Participant:** Uh, yes. And yes, highlighting feature is, uh, is a great find. Uh, and, uh, the ability to, uh, put some, um, some representatives of well, uh, cluster one to cluster B to move them is, is really great. Uh,

61 **Dan:** Okay.

62 **Participant:** I think it's a great feature.

63 **Dan:** Okay. Um, so in your analysis, um, it looked like you mostly were looking at like the, the clusters and the cluster descriptions. Um, but didn't really use the, like the highlighting or the brushing of the ICE lines. Um, so was there, is there like a particular reason you preferred the cluster descriptions over over brushing or a reason why you didn't use the brushing?

64 **Participant:** Uh, I, uh, I did not find the, the, uh, situations, uh, the particular, uh, cases to use it, like, uh, in the case of, uh, the task you demonstrated. Like, with, bike rentals, it was, uh, it was interesting to go, uh, into like how the rain influences the bike rentals, like, uh, look into cases when, uh, rain negatively influenced, uh, bike rentals. But, uh, in case of, uh, my data set, the data set you provided. Uh, I analyzed, uh, it seemed, seemed like, uh, I did not see, uh, obvious, uh, obvious, uh, features, obvious uh, anomalies to go into to, to, well, like, go and select them by hand. Uh, so, uh, it was more obvious to me to go and look into clusters.

65 **Dan:** Okay. Um, okay, so were the filtering capabilities useful for your analysis? Are there any additional ways to filter the plots you think would be helpful? Um, so I think in your analysis, I don't think you did any filtering. Um, so I guess if you

want to comment on like, uh, I guess your like your thoughts on like, why you didn't filter or are there like any other ways that you think would be helpful to filter the plots.

- 66 **Participant:** Um, I didn't filter because, uh, well, um, I did not, uh, like, uh, I didn't understand how to, uh, add it into, like, to the set of the tricks I would like to try. Uh, it's like, uh, there were, uh, many, many features I've seen and, uh, in a short period of time, 30, 30 minutes, I feel like it's not, uh, enough to try and exploit all of them, all of the features.
- 67 **Dan:** Okay.
- 68 **Participant:** It it was, it wasn't the first, uh, the most obvious feature to go into.
- 69 **Dan:** Right. Okay. Uh, so then I guess just in general, do you think there would be any useful ways to filter the features of the plots, besides by their name or, or shape.
- 70 **Participant:** Uh yes, I guess, uh, I guess the, uh, the possible way is to, uh, filter by, uh, maybe, uh, maybe by, uh, error. Um, like, we, uh, we have our prediction and, uh, uh, where do we, uh, miss the most? Maybe, uh, maybe this, uh, would be, uh, would be something to look into.
- 71 **Dan:** So for that, um, can you explain like how you might like calculate that. In terms of like, or I guess when you say, would this be like filtering the features or filtering the instances?
- 72 **Participant:** Uh, filtering instances.
- 73 **Dan:** Okay.
- 74 **Participant:** Uh, it's like, uh, the instances we have for the biggest error. Uh, like, um, well, we have our, our metrics, like, you know, regression...
- 75 **Dan:** Right. Yeah. Okay. Yeah.
- 76 **Participant:** ...,mean squared error uh, and so we like, uh, go and see where's the biggest errors. Uh, and, uh, I guess they could be, it could be useful for that.
- 77 **Dan:** Okay. So like you would want to see, like, what do the ICE lines look like for the instances that have the biggest errors. Uh, yes, yes. Uh, I would like to, I would like to know, uh, which, uh, which feature. Uh, which feature like misled our prediction. To, to the point where we get, uh, like a big error.
- 78 **Dan:** Okay.
- 79 **Participant:** In in in in general answer to to this to this question.
- 80 **Dan:** Right. Okay. Yeah, I think that makes sense. I'm not sure. Like what, like the. Exactly what it would look like in the context of like partial dependence plots, but I can see the. I can see the benefit of wanting to know, like. For the instances that have the biggest errors, what features. Led to that prediction and getting the big error. Uh, yeah. Okay. I can see that being useful. I'm not exactly sure how I would do that in the tool, but. I can see, I can see the value in that. Um, okay. So this is almost the last question now. So how well did the tool enable you to analyze feature interactions? Um, if you have anything else to mention for this.
- 81 **Participant:** Uh, well, for me, uh, feature interactions were. Uh, less, less obvious. Uh, less obvious feature. Uh, like. It's a, it's pretty interesting, but. Uh, I feel like I. I did not, uh, get it as well as I would like to. Uh, and, uh, and, uh, like, uh, I've had a chance to go and see the interaction between two features and. Uh, well, I've got some, uh, some ideas from that, some interpretation. Uh, but, uh, I'm not sure, uh, in terms of, uh, in terms of the value, uh, which I, I can bring with, uh, bring with using this feature because, uh, maybe, uh, maybe the business insight, uh, business insights are, uh, bigger when when you analyze, uh, like, uh, general influence, not, uh, interactions, uh, between two features. Uh, while it's, it's interesting too. But, uh, uh, in, in particular case of, uh, well, uh, dataset you gave me. Uh, I think, uh, the influence of of feature interactions was, uh, pretty small. If that makes sense.
- 82 **Dan:** Okay.
- 83 **Participant:** Like, like, uh, it feels like, uh, great, uh, additional feature. Uh, but, uh, other features seem to bring, uh, bring bigger value.
- 84 **Dan:** Okay.
- 85 **Participant:** At first. At first. Obviously.
- 86 **Dan:** Right. Um, yeah, I think that's all good points. Um, so I think like one of the plots you did look at, uh, I think was. Uh,

let's see what might have been like one of these plots or something like that. Um, where like, you did notice that there's like quite like a substantial interaction. Um, but yeah. Um, I guess another thing to add to that. And like the type of interact, the type of like analysis you were doing with like the clusters and the cluster descriptions. This is in a way like exploring interactions as well. Um, since you're seeing like, how does the effect of basement finish area. Uh, change like the impact of the above ground living area. So like this view is kind of also looking at feature interactions. Uh, just through like a different lens. Um, okay. Um, so then, uh, what would you say are the tool's biggest weaknesses or limitations? And are there any, uh, additional capabilities or improvements you would want it to have?

87 **Participant:** Uh, I think the biggest limitation is that. Um, it can be complicated for people who do not understand PDP plots and. Um, well, by far, if people are new to machine learning, like maybe business users. Uh, I'm not sure. Uh, how to, how would I. Uh, explain some of the graphs. Uh, we've analyzed today.

88 **Dan:** Mm hmm. Yeah, I think that's a good point. Yeah. So I mean, we we designed this for. For data scientists and machine learning experts. Um, so yeah, you're right in that it's that it's not suited. It's not suited for. Uh, like business people or those that don't have, um, machine learning expertise. So are there any other capabilities or improvements that you'd want PDPilot to have?

89 **Participant:** Uh, well, maybe, uh, to, uh, to repeat what I've said about error analysis maybe, uh, it would be useful to go, uh, go into, uh, go into errors and understand why, uh, we, we have a great performance on, uh, some of the cases and, uh, very, very bad performance on some other subset of the cases.

90 **Dan:** Okay.

91 **Participant:** And, uh, well, uh, I think that's, that's all. That's all.

92 **Dan:** Okay. And now this is the last question. So this is just for if you have any other feedback that you want to give that you didn't haven't already given.

93 **Participant:** Mm. No. Uh, great job. Uh, I think, uh, I would, uh, I would like to use it. And, uh, I think that says it all.

94 **Dan:** All right. Great. Um, all right. So that's it. So thank you very much. I really appreciate your time and, uh, participating in the study. Um, so thank you very much. Uh, I understand this is a big time commitment. So I'm very grateful that you chose to participate. Um, so, uh, after this, I will send the compensation. XX XXX XXX XXXX XXXXX XXXX XX XXXX X
XXXX XXXX XXX XX XXXXXXXX XXX XXX XXXX XXXX XXXX XXX XXXXXXX XXXXXXXXXX XXX XXXX XXXXXXX XXXX
XXXX XX XXXXXXX XXXX XXXXXXXXXX XXX XXXX XXXXXX XXXXXX XX XXXXXXXXXX XXXX XXXX XXX XXX XXXX
XXXXXX X XXXX XXXXXXXXXX XXXX XXX XX XXXX XXXXXXXXXX XXXX XXXX XX XXXXXXXXXX XXXX XXXX XXXXXX
XXX XXXX XXXX XXX XX XXXXX XXXX XXXXXXXX XXXXXXX XXXXXXX XXX XXXXX XXX XXXX XXXXXX XXXX XXXX
XXXX XXX XXXX

95 **Participant:** All right. Yeah. Yeah.

96 **Dan:** Okay. Um, so PDPilot is open source and installable through PIP. Um, so there is a GitHub repository that has, uh, (unintelligible) so I can send that over slack. Um, so it is open source. You can install through PIP. Um, if you're interested in using it for your work or with your own data model, um, then let me know. Um, I'm happy to answer any questions or help you get set up. Um, I'm also happy to make any changes to PDPilot to support your needs. And I'm interested, always interested in hearing any, anything about your experience, your findings and any feedback. Um, so again, thank you very much. I really appreciate it.

97 **Participant:** Uh thanks. Thank you. It was pretty interesting two hours. Thank you.

98 **Dan:** Great. All right, I hope you have a great rest of your day.

99 **Participant:** Yeah. Uh, have a great one.

100 **Dan:** Thank you.