Dear Editors,

Please consider our paper titled "Exploring Local Explanations of Nonlinear Models Using Animated Linear Projections" for publication in the Journal of Computational and Graphical Statistics. We'd like to acknowledge that it was previously submitted to the Journal of Machine Learning Research. The primary reason was that the editor felt it had too much visualisation to be a good fit for the journal. However, some useful feedback was provided, and below we have described how we have improved the paper based on their reviews.

We feel that JCGS is a good fit for this work because it describes useful new visual approaches for understanding model fits based on the relatively new local explanations diagnostics. This work can also be used to assess the reliability of the information reported by these diagnostics.

Here is an explanation of how the paper was improved based on the review from JMLR:

(a) The writing of the paper is informal, lacks clarity and depth in many parts that are critical to understanding the method proposed, and requires significant improvement. The method of radial tour, which (apparently?) seems to be the most important contribution of the paper, is practically not explained at all. Up to Section 4, when the description of the interactive visualization tool starts, all we get for the explanation of the radial tour is one sentence, which is drastically not enough. This is observable throughout the text as a whole, and it is very hard to point exactly where and when it happens. In general, the paper's writing is informal, lacking the depth necessary to make it reproducible and auditable. Without clear, detailed explanations of the methods used, it becomes hard to accept it in a deeply technical journal such as JMLR.

(b) The paper does not seem to fit the scope of the journal. Even though the idea behind the radial tour seems interesting (although it is hard to know because the authors have skipped explaining it in details), it seems that one of the most important contributions of the paper is actually an interactive visualization tool. As such, I believe that JMLR is not the right venue for this paper. The authors should've chosen a visualization venue for this, where they'll be able to get a much more interested audience and much more relevant feedback too. I'd recommend IEEE VIS, IEEE TVCG, EuroVis, Computer Graphics Forum, Computers & Graphics, Information Visualization journal, or the PacificVis conference.

- *Abstract: While it is not wrong, not all XAI techniques use local explanations; some definitely do. The wording here could be a bit more accurate.*

- *The introduction seems to end a bit abruptly. Maybe a couple of sentences about the results and outcomes of the research?*

- *Page 3: "Consider a highly nonlinear model." I feel like this (and the reasoning that follows) is a bit of an informal statement. Although intuitively I agree with the authors, is there any objective (or at least semi-formal) way to explain this problem? What is highly non-linear (vs. just linear)? Why, exactly, is it that it is so hard to interpret nonlinear models? That would improve the argumentation here quite a bit.*

- *"The attribution of feature importance depends on the sequence of the included features." This sentence feels out of place.*

- *"modified breakdown plots" – Since this is taking an important place in the paper, with a figure attached to it, I think it needs to be better explained. What are exactly modified breakdown plots?*

- *"The magnitude of the contributions depends on the sequence in which they appear." – This is unclear. The figure has no label for the horizontal axes, which makes it very hard to interpret. And, in case, the sequence alters the magnitude of the contributions, then it should be explained how the sequence is determined. Right now, this explanation does not help to make the visualization or the method presented in Figure 1 easier to understand.*

- *"The attribution space corresponds to the local explanations for each instance; feature importance in the vicinity of the instance." – I think this needs to be better explained. The concept of the attribution space remains obscure to me. It seems to be a very interesting concept and very important to the paper.*

- *"The projection attribution of the primary instance (PI) is examined and typically viewed with an optional comparison instance (CI)."* – The concept of PI and CI remains confusing to me. Needs a clearer explanation.

- *Section 4.2 as a whole* – This explanation is overly confusing and informal/subjective. It needs to be significantly revised to become more fluid and clear.

- *I don't believe that sections 4.7 and 4.8 are needed.*

Regards,

Nick Spyrison