

As far as I'm concerned, I believe that Rashomon set is realistic and can be used to meaningfully capture explainable models. [1]

Rashomon effect means that there exist a large number of models (the empirical Rashomon set) that perform approximately-equally-well on the training data. Because of that, if the model set is large enough, it's likely that models with desirable properties can exist inside it.

However, it could be a challenge to proving whether an interpretable model exists in the Rashomon set. There's already a few research on the properties of Rashomon sets and types of model it could contain. For example, the Rashomon ratio, a new measure related to simplicity of model classes, which is the ratio of the volume of the set of accurate models to the volume of the hypothesis space. This ratio can help modelers to navigate the trade-off between simplicity and accuracy, then find a good model which is explainable and with good accuracy.[2]

In conclusion, the research on Rashomon ratio give us a way to explain this effect, and then search in the set and find the desirable model, which makes Rashoman set realistic in capture models.[1]

- [1] Rudin C . Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead[J]. Nature Machine Intelligence.
- [2] Semenova L, Parr R, Rudin C. A study in Rashomon curves and volumes: A new perspective on generalization and model simplicity in machine learning; 2018. In progress.