Q9

I believe Rashomon set is realistic, though it is still explorable. I think Rashomon set is an idealized concept, if we prove it accurately, we need that large of models. But until now, the number of models is not enough. We need more scholars to explore.

According to Cynthia Rudin (Rudin, 2019), Rashomon set is a large number of models that are highly interpretable and accurate. There are three keywords, they are a large number of models, highly interpretable and accurate respectively. The first keyword is a large number of models. As we know, there are many interpretable models for Machine Learning, such as decision tree, logistic regression, Naïve Bayes, KNN, etc. Therefore, it is possible to find a large number of models to predict the same data set. At present, most machine learning models have little difference in performance on the same problem, so sometimes the adjustment of the same model may have a greater impact. Assuming that a machine learning model has noise, (it can be understood as parameter adjustment or algorithm fine-tuning), the result is better than the original performance, which leads to the possibility of the emergence of Rashomon sets. Then in the case of enough noise, there must be an optimal solution and Rashomon set. Then I will explain the other two keywords. Assume that a model is a collection of simple functions, then the model can be explained by these simple functions. If a Rashomon set contains enough models to solve the same problem, there will have models that contain simple functions. We can change these simple functions to get a close-to-optimal value and they are interpretable. In short, the Rashomon set is realistic and can be used to meaningfully capture the interpretable models.

There are many deep learning models, such as convolutional neural networks, recurrent neural networks, etc. They are all black-box models. We cannot understand the specific steps in them. Many problems can often be completed by simple machine learning, but most scholars have made the model more and more complicated and unexplainable to improve a little bit of accuracy. Even I think that in most cases, the model is too complicated, which will make the result worse. Many published papers will only tell others the best results. And we ignored most of the poor results. According to the Rashomon set, we can use the simplest equation to achieve the best result. I think it is very possible. If scholars can agree on interpretable artificial intelligence, then a large number of interpretable models will appear, and one of them is likely a perfect model.