I think Rashomon set is realistic only for well-trained data scientists who are familiar with different types of models and when to use them. The process of generating the set of possible models for a finite dataset requires the understanding of those models at first. Considering the hardness to trust explainable models and understand black box models, according to C. Rudin’s descriptions that the former one may generate misconceptions and the latter one may be too complicated and secretive, I think the selection process of which accurate models should be added into the Rashomon set may generate another sophisticated problem (Rudin 2019). I understand that interpretable models are more beneficial for adjustable applications, but I’m not sure whether Rashomon set will add another layer of subjection or complexity in mining those possible models for a dataset.

If the sets of models are generated by capturing various models, I think it’s difficult to find out the most meaningful one. If the size of the sets is large enough, it will take tremendously long time for computer to test which model will be the most suitable one for the datasets, especially if some models are sophisticated. In the worst scenario, all the models, either interpretable or explainable, may just fit into the datasets equally, but the system keeps looking for the most interpretable model until all of them are tested.

Reference

Rudin, Cynthia. "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead." *Nature Machine Intelligence* 1.5 (2019): 206-215.