Figure 19. (**Evidence of MD tree can generalize from small-scale models to large-scale models**). Each colored circle represents one sample (which is one pre-trained model configuration), and the color represents the ground-truth label: blue means the hyperparameter is too large, while red means small. The black dashed line indicates the decision boundary learned by MD tree from the training set. *First row* shows the training set (sub-figure a) and test set (sub-figure b) of transferring diagnosis from small size model to large size model. *Second row* shows the training set (sub-figure c) and test set (sub-figure d) of transferring diagnosis from models trained with a small amount of data to those trained with a large amount of data.

We provide evidence that the MD tree can generalize from small-scale models to large-scale models in Figure 19. From the training sets of Figure 19(a) and 19(c), we can see that the multi-regime pattern of NNs on the space of loss landscape metrics appears even if only small-scale models are present. The boundaries between these regimes are linearly separable, and it is captured by our MD tree method. The decision boundary is indicated by the black lines and can transfer well to the test set (shown in Figure 19(b) and 19(d)) that includes large-scale models. Figure 19 is replicated from Figure 11 and 12 from the submitted paper.