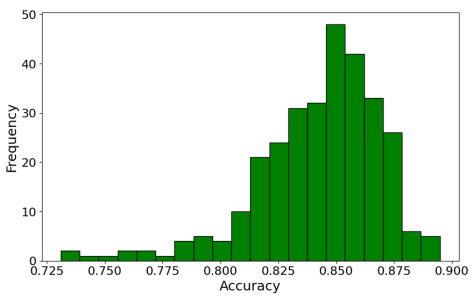
Build an Ensemble of Neural Network Ensembles Using a Predictive Function

The Problem

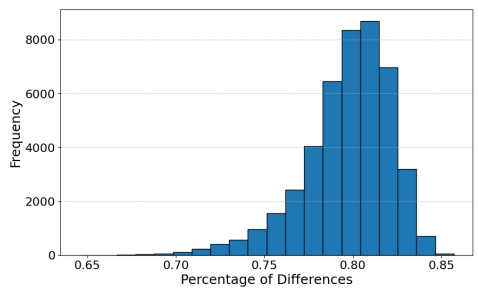
The neural network architecture space $\mathcal{A}$ is exceedingly large (on the order of 10^{24} architectures in our case), raising the question of developing an efficient method for searching for the optimal ensemble $S \subset \mathcal{A}$, i.e., the ensemble that achieves the highest accuracy.

Predictive Function

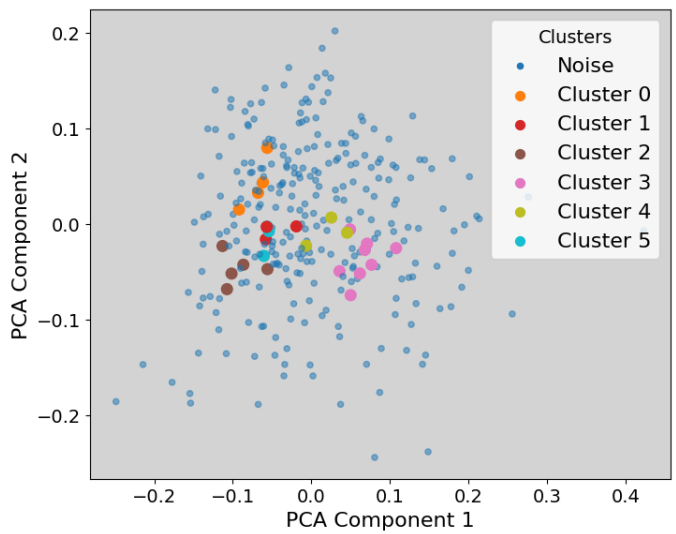
We propose a function to estimate model similarity (e.g., using the Jensen-Shannon distance on model predictions from the test dataset) to guide the selection of an optimal ensemble, where "optimal" refers to maximizing the ensemble's overall predictive accuracy.



Distribution of model accuracies



Distribution of model diversity



Clustering of embeddings