A

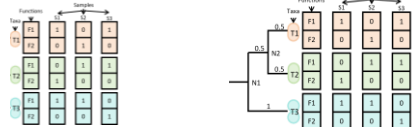## Data Acquisition

Taxon-Function Table

Phylogeny-informed Taxon-Function Table

Distance Metrics Calculation

$$BC_{ab} = \frac{\sum_{i=1}^{N}|a_i - b_i|}{\sum_{i=1}^{N}(a_i + b_i)}$$
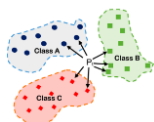
Bray-Curtis Distance

$$d_J(a,b) = 1 - \frac{|a \cap b|}{|a \cup b|}$$

Binary Jaccard Distance

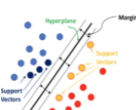$$d_E(a,b) = \sqrt{\sum_{i=1}^{N}(b_i - a_i)^2}$$

Euclidean Distance

$$PiF_{ab} = \sum_{i=1}^{N} l_i d_{ab} p_{ia} p_{ib}$$
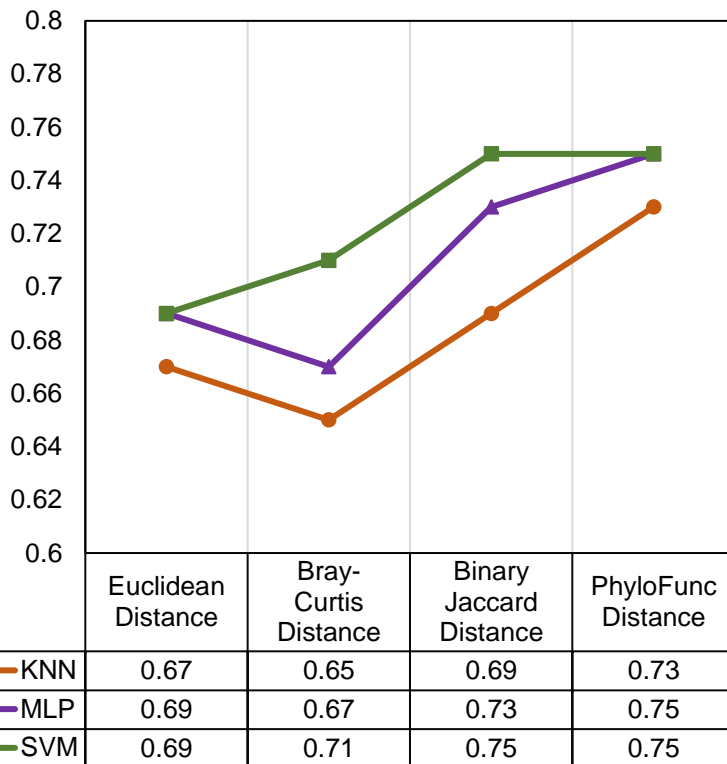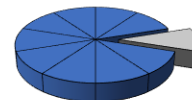
PhyloFunc Distance

Classification Methods

KNN   MLP   SVM

B

	Euclidean Distance	Bray-Curtis Distance	Binary Jaccard Distance	PhyloFunc Distance
KNN	0.67	0.65	0.69	0.73
MLP	0.69	0.67	0.73	0.75
SVM	0.69	0.71	0.75	0.75

Optimize Model Parameters

C、kernel、n_nerghbors、learning_rate、max_depth

Evaluate Models

Leave-one-out cross validation