***Report – HW4***

Stable Coalition:

For built stable coalition, we look on the problem from 2 different side:

1. Clustering – split all examples to two different groups using a distance function using continuous values, when in each group has similar examples. the big group is the coalition and the smaller group is the opposition.
2. generative modeling – we will use the probability function properties to compare the different parties to find their similarities and differences. for example, for gaussian naïve base, we can get from the model the expectation and variance and use them to compare between the different parties and build two different groups, the big group is the coalition and the smaller group is the opposition.

Stable Coalition using clustering:

We use 2 models that we learn in the class for clustering model K-means and Gaussian mixture. first, we found the hyper parameters using k fold cross validation, we define evaluate function to evaluate the best hyper parameters. The evaluate function is to find a group that contains as many voters from a specific party and does not split voters into several different groups. we set a certain threshold value from which we decide that a group belongs to a specific group, for example 45% for k=3 clusters.

The process the we made to reach stable coalition is:

1. we train the train set using k-means for different k clusters and for different threshold.
2. we check accuracy on the validation set and tuning hyper parameters (k and threshold) accordingly.
3. Finally, we check accuracy on the test set.

Accuracy:

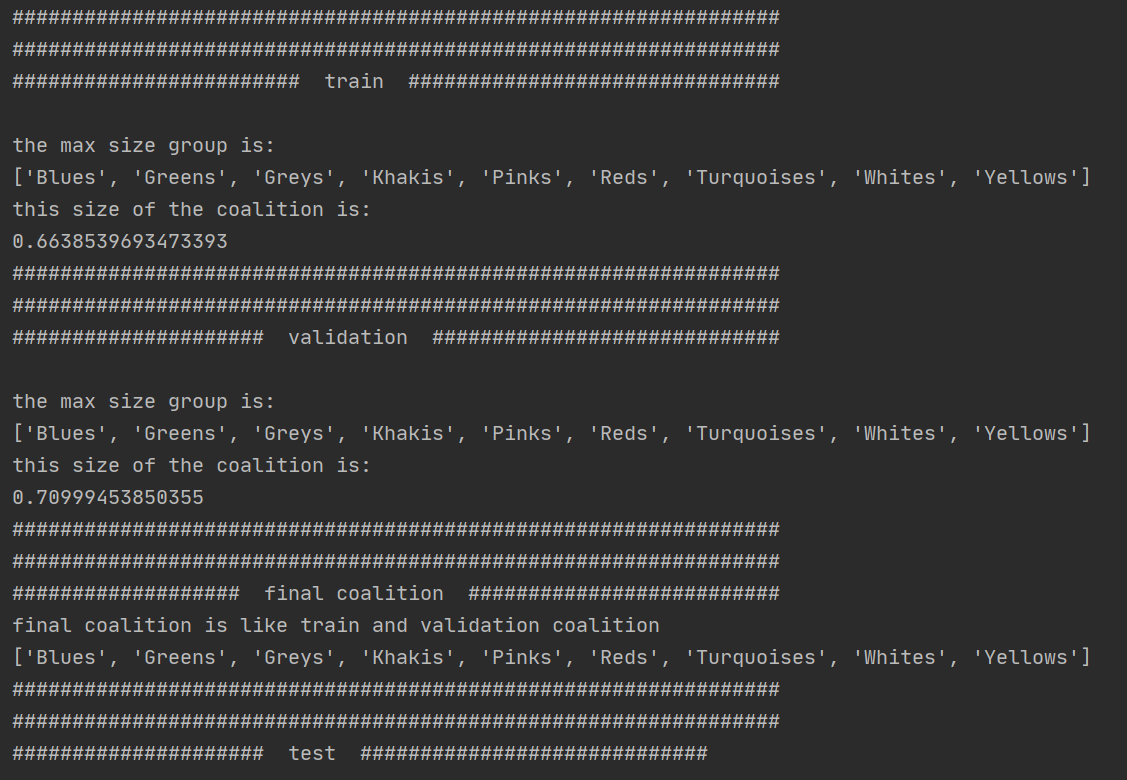
How we check accuracy in the test and validation set?

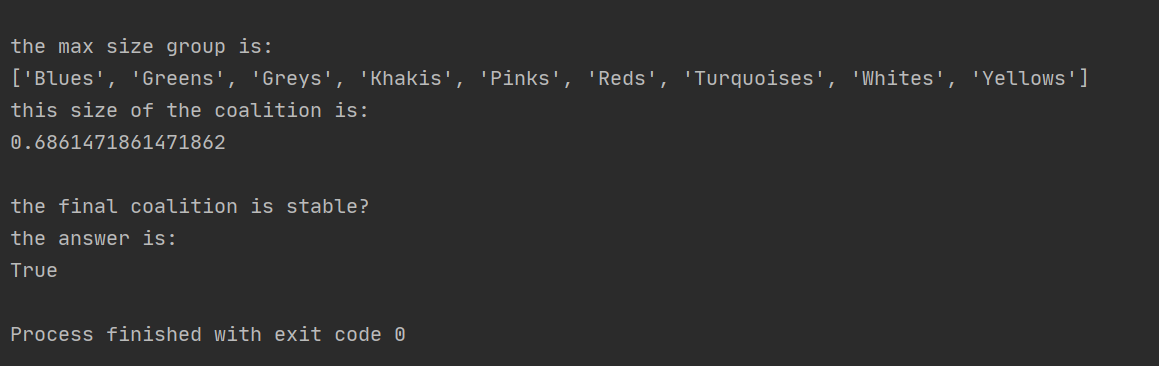
We check if we get equal coalition in all sets, that is:

Train coalition == validation coalition &&

Test coalition == validation coalition == train coalition

Finally, we get these results:





We can see from these results that we get stable coalition for k=3 and threshold=0.45,

The final coalition consists of 9 parties and they:



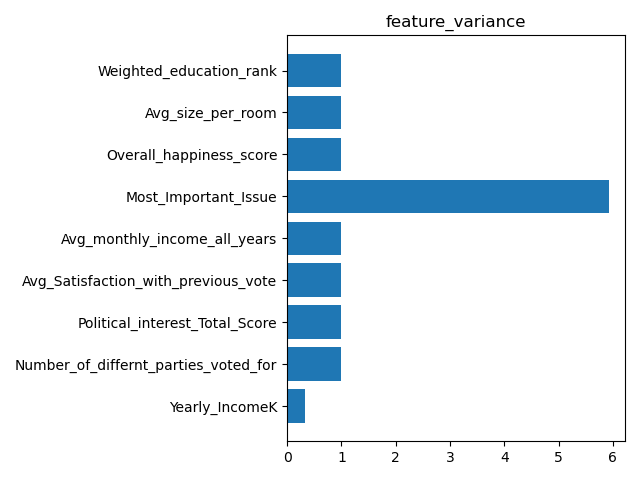
Size: The sizes of the coalitions in different sets is:

Train: 66.3%, validation: 71%, test: 68.6%

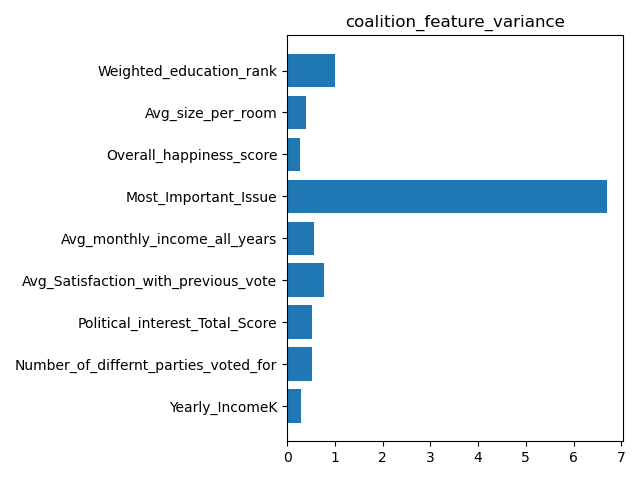
We can see that the sizes are about the same size.

Homogeneous: The homogeneous of the coalition is determined by threshold, the bigger the threshold the bigger the homogeneous. after a lot of experiments, for different k and different threshold, we saw that we get stable coalition for k=3 and threshold=0.45, and if we take threshold that bigger from k=0.45, we did not accept stable accuracy.

Before choosing coalition we get variance like this:



And after choosing coalition we get:



We can see from the graphs that the variance is smaller, hence the coalition is more homogeneous.

***Stable Coalition using Generative Model:***