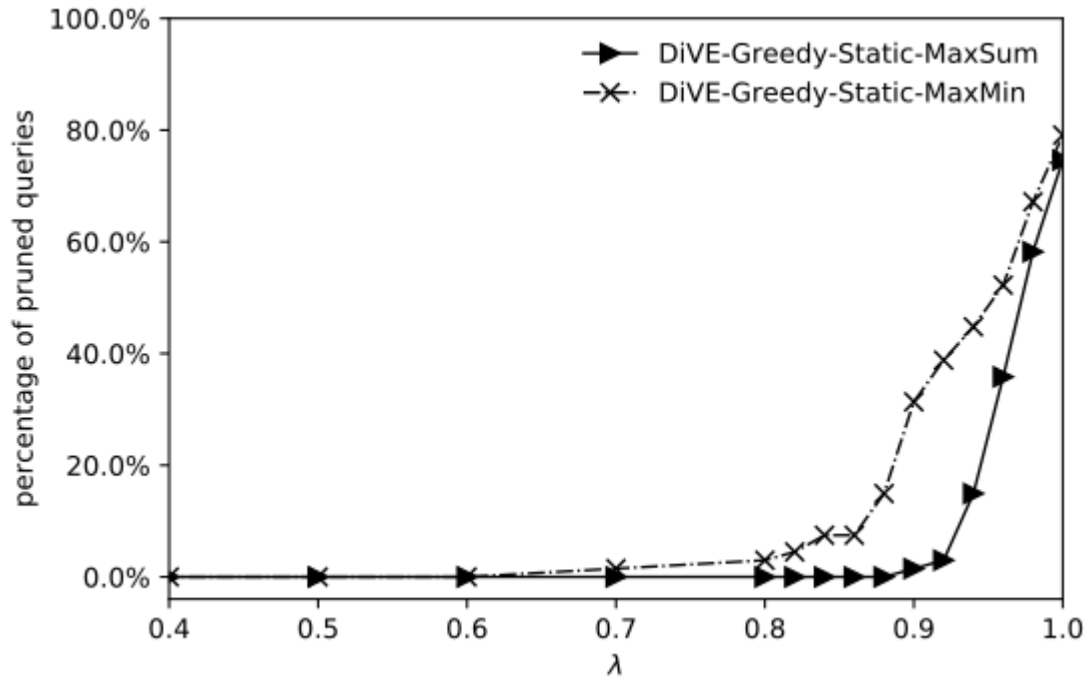


## MaxSum vs. MaxMin Diversification

Current results that we have are using MaxSum diversification. Based on my hypothesis and experiments seem that MaxMin Diversification can improve the pruning performance. Below the results while using MaxSum vs. MaxMin on Greedy Static Top1.



MaxSum uses average score of diversity of the set  $S$  while MaxMin uses the minimum score of distance in the set. Hence, the range diversity score from those both approaches are different.

The variance of *setDist* score while  $\lambda=0.8$  between MaxSum diversification and MaxMin diversification can be seen in the Table below. In this experiment, I selected two most distant views as the initial set  $S$  and then calculate the *setDist* of all views in  $X$ . For instance, the highest score of *setDist* is  $v1$ , where on MaxSum the maximum score is 0.5 and on MaxMin the maximum score of diversity is 1. This Table is just an example, in the real data there are many views have same score. In this Table, I only want to show the distributions of *setDist* score and the different range of *setDist* score between MaxSum and MaxMin.

<b>v</b>	<b>MaxSum setDist</b>	<b>MaxMin setDist</b>
v1	0.5	1
v2	0.491666667	0.833333333
v3	0.483333333	0.666666667
v4	0.475	0.5
v5	0.466666667	0.333333333
v6	0.458333333	0.166666667
v7	0.45	0.166666667

Using MaxSum diversification means we calculate the average of diversity score from all views in the set  $S$  whereas MaxMin uses the minimum score of diversity among views in the set. As an example above, if there are no same views in the set (maximum score of setDist), MaxSum will give 0.5 and MaxMin will give 1 as the diversity score.

Due to this different diversity score, MaxMin diversification can improve the pruning performance as shown in the first Figure. However, this MaxMin makes the value unbalance between the importance score and diversity score. The maximum diversity can be equal to 1 while the value of importance score is low. This thing makes the shape of objective function unbalance.

#### **Update:**

There is a strange behaviour on my Swap algorithm while implementing MaxMin diversification and I am still working on it since Yesterday and can't find the problem. Pruning performance of MaxSum is better than MaxMin in terms of number but in terms of  $\lambda$  MaxMin is better, it can starts prune early compared to MaxSum. I am still inspect what's wrong on my code.