Haonan Hu

3/18/2021

2863545

In-Class Problem

1. In the "perturb the current feature subset" step of the SA algorithm:
   a) If the random percentage is 4% and the number of features is 200, what is the number of features that will be added or deleted from the current subset of features?
   b) Describe a simple algorithm for determining whether to randomly add or delete the number of features calculated in 2a.
2. Calculate Pr[accept] for iteration 4, where the accuracy for iteration 3 was is 0.820 and the new accuracy is 0.790 (c = 1).
3. Determine the Status (Improved, Accept, or Reject) for the table to the right.

Status:
Improved = accepted because new Acc > old Acc
Accept = accepted because Random Uniform ≤ Pr[accept]
Reject = rejected because Random Uniform > Pr[accept]

Old Acc	New Acc	Pr[accept]	Random Uniform	Status
0.786	0.791	1.019	0.023	
0.791	0.780	0.946	0.980	
0.780	0.814	1.244	0.683	
0.814	0.803	0.922	0.769	
0.803	0.789	0.885	0.000	
0.789	0.786	0.970	0.980	
0.786	0.808	1.287	0.004	
0.808	0.784	0.743	0.975	

Note:
- Treat each row as a separate iteration
- The Old Acc is the previous New Acc regardless of the Status

Simulated Annealing Feature Selection — David O. Johnson EECS 690 (Spring 2021) — 34

1. In the "perturb the current feature subset" step of the SA algorithm:
   a. If the random percentage is 4% and the number of features is 200, what is the number of features that will be added or deleted from the current subset of features?

   200 * 0.04 = 8

b. Describe a simple algorithm for determining whether to randomly add or delete the number of features calculated in 1a.

After calculating current feature subset, fit the model and estimate performance. If the performance is better than previous subset, then add the calculated set of features. Else, an extra step of calculating acceptance probability is needed, if random uniform variable is greater than the calculated acceptance probability, then delete the calculated set of features from the current feature subset, otherwise, accept the new subset, which means add the set from original feature set.

2. Calculate Pr[accept] for iteration 4, where the accuracy for iteration 3 was is 0.820 and the new accuracy is 0.790 (c = 1).

$$\boldsymbol{Pr[accept] = e^{-\frac{4}{1}(\frac{0.82-0.79}{0.82})}}$$ **= 0.864**

3. Determine the Status (Improved, Accept, or Reject) for the table to the right.

Status
Improved
Rejected
Improved
Accepted
Accepted
Rejected
Improved
Rejected