1. Genetic algorithm ended up performing the best
2. PCA is a linear dimentionality reduction so the data might not have been linear enough to perform as good as the others. There is also always some info lost so a lot might have been lost
Simulated annealing ended up finding the same exact featureset so it is likely that its fold split was less favorable than the one for GA
3. Yes, Genetic algorithm ended up performing better than no dim reduction, probably because it compared the no dim reduction to others and found that there was some other combination that did better
There is also probably some underlying information that was not captured with the basic data.
4. No it did not, PCA creates 4 new features z1-z4 while Simulated annealing choses the best out of the 4 original and 4 new. Since the best set was not just the 4 that was created by PCA,
the featureset will definetly be different. If the best features were z1-z4, then yes, however that is unlikely
5. Same as the reason above
6. Yes they created the same. This is because the goal of both is to find the best subset of a given feature set so it is possible and even likely that they converge on the same feature set.