Question Answers

- Based on accuracy which dimensionality reduction method, PCA, simulate annealing, or the genetic algorithm worked the best?
    - Genetic Algorithm worked the best with an accuracy of 98.68.
- For each of the two other methods, explain why you think it did not perform as well as the best one.
    - PCA – Most of the variance is concentrated on one Principal component which amounts to 92% accuracy only.
- Did the best dimensionality reduction method produce a better accuracy than using none (i.e. the results of Part 1)? Explain possible reasons why it did or did not.
    - No
    - Genetic algorithm produced 98.67% accuracy which is better than the decision tree without any feature selection or dimensionality reduction.
    - Part 1 used only original features, whereas GA used principal components and original features, that gave better information.
- Did Part 2 produce the same set of best features as Part 3? Explain possible reasons why it did or did not.
    - No
    - Part 2, PCA, gives us only principal components.
    - Whereas for Part 3, we used both original features and Principal components. Using combination of these, we cover more domain knowledge than principal components alone. This resulted in a higher accuracy.
- Did Part 2 produce the same set of best features as Part 4? Explain possible reasons why it did or did not.
    - No
    - Part 2, PCA has less accuracy than Genetic algorithm. PCA uses original features that are interdependent.
    - Genetic algorithm converged to optimal accuracy – 98.67%. The feature set for this is a combination of both original features and principal components which has more domain information.
- Did Part 3 produce the same set of best features as Part 4? Explain possible reasons why it did or did not.
    - No.
    - Part 3, Simulated annealing produced the same (best) accuracy, 98.67, in earlier iterations but did not converge to it in the end after all the iterations.
    - Due to a more stochastic nature in the algorithm, compared to Genetic algorithm, it did not converge in the end.
    - On the other hand, Genetic algorithm and simulated annealing were able to converge to optimal accuracy.