

### **DSC 630 Assignment 4 Summary**

For this assignment, three methods were used to predict which hotel cluster would be the best to recommend users of Expedia based on their search specifications on the site: Two standard machine learning models, and one model generated from scratch by following the Expedia challenge rules (one base model and one improved model will be compared, both built from scratch). It was found that the improved model generated from scratch output a 19% precision, whereas the two standard machine learning models (Naïve Bayes and AdaBoost) output 3.3% and 1.9% precision, respectively. Because two machine learning models output much lower precision scores, this means that machine learning alone is not enough to generate a good performing classification model. While the Expedia code did not output good precision either, it was still 6 times higher than the highest percentage that could be achieved from machine learning. Why is machine learning not as good at recommending hotels as the code written from scratch? Perhaps it is because there are much too many factors that go into whether a user ultimately chooses one hotel over the other.