STAT 517 – Final Project Critique

Fadhil Salih: HAM10000

Reviewed by Ryan Heiderman

**Coverage of 3 Course Areas**

Fadhil did a good job covering the subject areas of Classification, Clustering and Association.

**Dataset size criteria**

The HAM10000 dataset was a set of 10,015 dermatoscopic images, with each image containing 784 pixels and a label of the type of lesion. If each pixel is a column with a value, then this data set meets the 100,000 record x row minimum.

**Relevancy**

This topic was relevant to the course material as well as the larger aspects of medical analysis.

**Difficulty**

Working with a large dataset of images this was an appropriately difficult task Fadhil undertook.

**Interestingness**

This was an interesting analysis. I think anything medical can be interesting and utilizing imagery for prediction of issues is fascinating.

**Clarity**

This project goals and tasks were clear and straightforward. Fadhil had a large dataset of images, each with a label classifying the lesion type. He set out to used different machine learning techniques to classify, cluster and find particular associations between the imagery.

**Originality**

I believe Fadhil was the only student in the class to use medical imagery in his project, it was an original idea.

**Creativity**

He utilized many of techniques we learned in class to develop unique and creative analysis for classifying medical images of lesions.