**Artificial Intelligence Project 1: Design Document**

**Alex Abrahamson, Wilson Harris, Matthew Nitschke**

1. **Description of the Problem**
2. **ARFF Converter**
   1. **Design Decisions**
   2. **UML Diagram with Text Descriptions**
3. **Experimental Design**
   1. **Data Sets**

Perms-SF Data Set: Low number of instances, high number of attributes The data describes the occupancy   
rate, between 0 and 1, of different car lanes of San Francisco bay area freeways. The measurements cover the period from Jan. 1st 2008 to Mar. 30th 2009 and are sampled every 10 minutes.

HEMPMASS Data Set: High number of instances, low number of attributes These signatures are learned from Monte Carlo simulations of the collisions that produce these particles and the resulting decay products. In each of the three data sets here, the goal is to separate particle-producing collisions from a background source.

Parkinson Disease Spiral Drawings Using Digitized Graphics Tablet Data Set: Low number of instances, low number of attributes

* 1. **Machine Learning Algorithms**
  2. **Evaluation Measures**

**References**

Alex – Part 1

Matthew – Part 2

Wilson – Part 3