# Grading Rubric for Machine Learning Report

**General (25 points total):**

_____ / **2:** Report includes student's name

_____ / **3:** Report uses proper formatting (12 pt. font, 1 inch margins, etc.)

_____ / **20:** The data was submitted and opens without error in Weka 3.8 or higher.

**Section 1: Data Set (20 points total):**

_____ / **2:** The type of thing being observed or described is clearly identified.

_____ / **6:** A clear explanation of each attribute in the data set is given.

_____ / **2:** A clear explanation of the class label is given. It is important to justify why this particular attribute is the class label; you may not simply choose one at random.

_____ / **3:** Any pre-processing done to the data is clearly described. Examples of pre-processing include removing extreme or impossible values, bootstrapping, and filling in missing values. If no pre-processing was done, this must be explicitly stated.

_____ / **2:** The impact of this classification task is described. Why will it be helpful to classify these kinds of things accurately?

_____ / **5:** Either an explanation of how the data set was collected/generated is presented, or an explanation is provided, with relevant sources cited, if the data was taken from a third party source.

**NOTE**: You may not choose one of the example data sets provided by Weka. These data sets can be found in the "data" directory where you installed Weka. Choosing one of Weka's example data sets will result in receiving no credit for this project.

**NOTE**: If the data set is taken from a third party source and no citation is given, this will be considered plagiarism and the student will receive no credit for this project.

**Section 2: Baseline Classifier (7 points total):**

_____ / **2:** The type of baseline classifier used is named and briefly described.

_____ / **3:** The results of the baseline classifier are given, including accuracy (or other relevant performance metrics).

_____ / **2:** At least one example of an observation that the baseline classifier failed to classify is given, and the impact of this failure on the problem is stated.

**Section 3: "Intelligent" Classifier (14 or 19 points depending on section):**

_____ / **1:** The type of classifier used is identified

_____ / **2:** Some specific reason for using this classifier is identified. Classifiers cannot simply be chosen at random or because they were the most accurate of all the ones tried. Student has identified some reason this particular classifier is appropriate and explained it well.

_____ / **2:** Any parameters used by this classifier are given and described *in the student's own words*, not just copied from existing documentation. An example of a parameter would be the number of neurons used by a neural net and how they are arranged into input / hidden / output layers.

_____ / **5:** CSCI 5525 only: A description of how the classifier constructs its model is given. (About 1 page of detail. At least two concrete examples must be included in the description).

_____ / **2:** The results of this classifier are given, including accuracy (or some other relevant measure given the type of classifier used).

_____ / **5:** The results of this classifier relative to the baseline are visualized using an appropriate graphic (for CSCI 5525, the second classifier is also included).

_____ / **2:** A justification for why this classifier's accuracy is sufficiently high for this task is provided.


**Section 4: Second "Intelligent" Classifier (CSCI 5525 only, 19 points):**

_____ / **1:** The type of classifier used is identified

_____ / **2:** Some specific reason for using this classifier is identified.

_____ / **2:** Any parameters used by this classifier are given and described in the students own words.

_____ / **5:** A description of how the classifier constructs its model is given. (About 1 page of detail. At least two concrete examples must be included in the description).

_____ / **2:** The results of this classifier are given, including accuracy (or some other relevant measure given the type of classifier used).

_____ / **5:** The results of this classifier relative to the baseline and the first classifier are visualized using an appropriate graphic.

_____ / **2:** A justification for why this classifier's accuracy is sufficiently high for this task is provided.



**Total for CSCI 4525: _____ / 66**

**Total for CSCI 5525: _____ / 90**