**Assignment 3.1 Short Answer Template**

These questions are based on the Assignment 3.1 Exercises document.

**Written Questions**

(short answer, 2-3 sentences):

* For the animal shelter dataset, the classification results using k-nearest neighbor are not as good as the other datasets. Why do you think this is the case for this data? Explain in terms of inductive bias, and remember that you used OneHotEncoding, making the data more high-dimensional and sparse.

* Which distance metric seemed to perform the best on the test dataset, Euclidean or Cosine distance? How large are the text feature vectors (how many features/words are in the dataset)? Do you think high-dimensional data like this are better suited for the Euclidean or Cosine distance, based on your understanding of the methods and the results you have obtained here?
* On the audio (Drum) dataset, you should have used nearest neighbor, perceptron, and your own handwritten implementation of the perceptron. Which classifier performs best, or are they all giving about the same results? If there is a clear winner, explain why this classifier might be a good match for the audio dataset.