***Questions for Chapter 1 (based heavily on the book exercises)***

*Submission instructions: Copy this document to your own drive so you can edit it. Answer each question inline (your answer should be immediately after the question). Format your answers in a color other than black so they are easy to see. Share your document with* [*rebeccalevitan@share.brooklyn.edu*](mailto:rebeccalevitan@share.brooklyn.edu) *and* [*denys.katerenchuk@gmail.com*](mailto:denys.katerenchuk@gmail.com)*. Submission closes Thursday 9/3 at 6pm.*

1. What is an example of supervised learning? Unsupervised learning? (In your answer, be sure to make it clear what makes your example supervised or unsupervised.)

An example of supervised learning could be labeling comments as positive or negative, using labeled training data (hence, supervised). On the other hand, we might use unsupervised learning to group different types of computer peripherals shoppers (enterprise, home, gaming, etc). The training data would be unlabeled, and the algorithm would try to find similarities between the datapoints to group them.

2. What is the purpose of a labeled training set? Test set? Validation set?

Labeled training sets are to train your algorithm to recognize similarities in the features of different datapoints. Test sets are to test the performance of your model before putting it into use, minimizing risk of backlash. The validation set is used for fine-tuning your model to prevent overfitting and prevent over-regularizing that would reduce performance.

3. What type of Machine Learning algorithm would you use to allow a robot to walk in various unknown terrains?

Reinforcement Learning: it’s easy to fit into a reward and penalty concept. Falling (and possibly stumbling) will be penalized, while distance and/or difficult-to-reach areas get points.

4. What type of algorithm would you use to segment your customers into multiple groups?

A form of clustering algorithm so that similar customers will be grouped. This goes well with the assumption that there is no single type of customer representative of everyone, and that not every customer is different. (However, for a business with only a few similar products this may not be worth the investment).

5. Make a copy of the life satisfaction collab notebook and do the exercises we started in class. (Put the last cell into a loop to do leave-one-out cross-validation and try out different models and predictors.)