School of Computing and Information Systems
The University of Melbourne
COMP30027
MACHINE LEARNING (Semester 1, 2019)

Tutorial exercises: Week 2

1. Revise the definitions of **instances** and **attributes** (or **features**). For the following problems, identify what the instances and attributes might consist of:

   (a) Building a system that guesses what the weather (temperature, precipitation, etc.) will be like tomorrow

   (b) Predicting products that a customer would be interested in buying, based on other purchases that customer has previously made

   (c) Automatically identifying the author of a given piece of literature

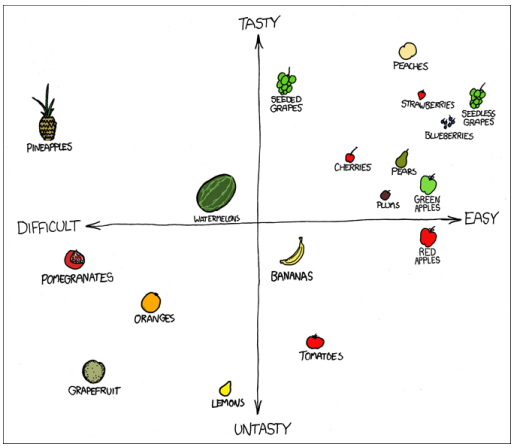
   (d) Finding the best burrito in the United States of America

2. What are the main differences between **supervised** and **unsupervised** machine learning? What kinds of "concepts" do we typically attempt to "learn" in Machine Learning — for each, identify whether they are primarily supervised or unsupervised.

   (a) Based on the problems in the previous question, identify the "concept" for each one, and conjecture whether a typical strategy is likely to use supervised or unsupervised Machine Learning.

   (b) (Extension) Brainstorm some mechanisms for solving supervised problems, and some mechanisms for solving unsupervised problems. In what ways are they the same? In what ways are they different?

3. Based on the following dataset[1] representation, would you consider "bananas" to be more similar to "apples" or "oranges"?



   (a) Identify the attributes in this dataset. What **types** of attribute are they (implicity)? What other attributes might be relevant, and what are their types?

   (b) (Extension) Critically assess the data: do you agree with how the problem is framed; do you consider the instances to be fairly located in the feature space?

[1] By Randall Munroe, `https://xkcd.com/388/`, used under a Creative Commons Attribution-NonCommercial 2.5 License.