

# Machine Learning Introduction Tutorial

## COMS3007

Benjamin Rosman

**Due date:** Tuesday, 15 March at 2.00pm.

**Instructions:** Use your notes and any resources you find online to answer the following questions. This would be best done in groups of *up to four*. Submit this as a PDF document on Moodle (one copy per group). Make sure all your names and student numbers appear on the document!

1. State if the following are examples of supervised learning (distinguish between classification or regression), unsupervised learning, or reinforcement learning. Justify your answer in each case.
  - (a) Determining if there is a person wearing glasses in an image.
  - (b) Playing a game of chess.
  - (c) Categorising different types of vehicles on the road.
  - (d) Predicting the judgement of a court case.
  - (e) Predicting student performance in a course.
  - (f) Formulating the best treatment plan for a patient.
  - (g) Developing a computer game AI.
  - (h) Predicting tomorrow's BitCoin price.
2. Find and discuss an example of supervised learning being used in a real-world business application.
3. Find and discuss an example of unsupervised learning being used in a real-world business application.
4. What features would be the most useful in the following learning tasks? Why?
  - (a) Classifying physical animals as dogs or cats.
  - (b) Classifying images of dogs and cats.
  - (c) Predicting the flight speed of any given bird.
  - (d) Predicting student performance in a course.
  - (e) Classifying songs by genre.
  - (f) Predicting which team would win a soccer game.
  - (g) Predicting the number of sales of a chocolate in a supermarket.
5. Learning typically requires a lot of training data, and for supervised learning this includes labelled ground truth data (the target values). For each of the problems in (4), how might you go about collecting the ground truth labels for the training data?