School of Computing and Information Systems
The University of Melbourne
COMP90049 Introduction to Machine Learning (Semester 2, 2020)

Workshop: Week 12

1. For the following set of instances:

a_1	a_2	a_3	c
hot	windy	dry	Yes
mild	windy	rainy	No
hot	windy	rainy	Yes
cool	still	dry	Yes
cool	still	rainy	No
hot	still	dry	No
mild	still	dry	Yes

   Construct all of the **1-itemsets** and calculate their confidences and supports. Discuss how you would continue mining for effective **Association Rules**.

2. What does "correlation does not imply causation" mean? Why is it important to keep this adage in mind, when working in the field of Data Mining?

3. Review the concepts of **Recommendation Systems**:

   (a) What is Content-based Recommendation?

   (b) What is Collaborative Filtering?

4. Consider the following rating table between five users and six items:

ID	Item A	Item B	Item C	Item D	Item E	Item F
User 1	5	6	7	4	3	?
User 2	4	?	3	?	5	4
User 3	?	2	4	1	1	?
User 4	7	4	3	7	?	4
User 5	1	?	3	2	2	7

   (a) Predict the value of the unknown rating for User 4 using User-based Collaborative Filtering. (i.e. Find the Pearson correlation between users and adjust User 4's mean score).

   (b) Predict the value of the unknown rating for User 4 using Item-based Collaborative Filtering. (i.e. Find the correlation between items (using "Adjusted Cosine Similarity") and take a weighted average of User 4's scores).