School of Computing and Information Systems
The University of Melbourne
COMP90049 Knowledge Technologies (Semester 2, 2018)
Workshop exercises: Week 11

1. For the following dataset:

ID	*Outl*	*Temp*	*Humi*	*Wind*	PLAY
TRAINING INSTANCES					
A	s	h	h	F	N
B	s	h	h	T	N
C	o	h	h	F	Y
D	r	m	h	F	Y
E	r	c	n	F	Y
F	r	c	n	T	N
TEST INSTANCES					
G	o	c	n	T	?
H	s	m	h	F	?

   Classify the test instances using a Decision Tree:

   (a) Using the Information Gain as a splitting criterion

   (b) Using the Gain Ratio as a splitting criterion

   (c) Using the Gini Index as a splitting criterion

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

   (a) What is Content-based Recommendation?

   (b) What is Collaborative Filtering?

3. 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).