

School of Computing and Information Systems  
The University of Melbourne  
COMP90049 Knowledge Technologies (Semester 1, 2017)  
Workshop exercises: Week 9

1. Revise the definition of **data mining**. Name the main modes of data mining, and give an example of a task for each one.
2. What is the **Naive Bayes** classifier? How does it classify data? What assumptions do we need to make about the data?
3. For the following dataset:

<i>apple</i>	<i>ibm</i>	<i>lemon</i>	<i>sun</i>	CLASS
TRAINING INSTANCES				
Y	N	Y	Y	FRUIT
Y	N	Y	Y	FRUIT
Y	Y	N	N	COMPUTER
Y	Y	Y	Y	COMPUTER
TEST INSTANCES				
Y	N	Y	Y	?
Y	N	Y	N	?

Classify the test instances according to the method of Naive Bayes, as discussed in this subject.