RMIT University
School of Science
COSC2110/COSC2111 Data Mining
Tutorial Problems Week 1

The following data was collected from the Melbourne Maternity Hospital on 1-Jan-2019, 1-Feb-2019 and 1-Mar-2019.

1-Jan-2019		
Height	Weight	Sex
59.43	7.38	boy
58.03	7.14	boy
55.66	8.63	boy
57.81	7.71	boy
58.89	8.82	boy
55.64	7.22	boy
55.53	7.93	boy
58.53	7.87	boy
56.25	8.27	boy
55.88	8.04	boy
51.44	4.70	girl
53.59	4.67	girl
51.16	4.31	girl
51.85	4.89	girl
54.03	5.30	girl
51.86	4.95	girl
52.03	5.12	girl
51.45	5.59	girl
50.54	5.62	girl
50.56	5.78	girl

1-Feb-2019		
Height	Weight	Sex
58.00	5.26	boy
55.13	7.31	boy
58.96	8.93	boy
53.02	5.86	boy
57.07	6.39	boy
57.04	5.64	boy
55.43	7.32	boy
58.81	6.20	boy
57.65	8.51	boy
53.61	6.85	boy
51.70	5.52	girl
52.93	5.60	girl
53.23	5.46	girl
55.53	6.04	girl
51.92	6.20	girl
56.11	6.10	girl
53.75	6.25	girl
50.75	6.61	girl
55.67	6.76	girl
54.19	4.66	girl

1-Mar-2019		
Height	Weight	Sex
56.52	6.25	boy
51.91	5.96	boy
58.84	8.12	boy
51.18	7.31	boy
52.72	8.60	boy
51.61	6.36	boy
54.89	7.33	boy
51.22	6.42	boy
54.32	3.02	boy
58.98	8.30	boy
56.39	6.45	girl
52.36	7.54	girl
59.14	3.27	girl
58.61	8.22	girl
52.11	3.42	girl
57.31	4.24	girl
59.95	5.42	girl
50.20	3.17	girl
59.75	7.41	girl
53.69	6.75	girl

1. Consider the data for 1-Jan-2019.
   (a) Would you expect a baby with height 57 and weight 5 to be a boy or a girl? Why?
   (b) Would you expect a baby with height 58 and weight of 5 to be a boy or a girl? Why?
   (c) Can you construct a rule from this data to make the decision?
2. Repeat the previous question for the data of 1-Feb-2019.
3. Repeat the previous question for the data of 1-Mar-2019.

4. Consider the following data about animals.

   (a) How would you expect the gila monster and platypus to be labelled? Why?

   (b) Can you construct a rule from this data to make the decision?

Name	Body Temperature	Skin Cover	Gives Birth	Aquatic Creature	Aerial Creature	Has Legs	Hiber-nates	Class Label
human	warm-blooded	hair	yes	no	no	yes	no	mammal
python	cold-blooded	scales	no	no	no	no	yes	reptile
salmon	cold-blooded	scales	no	yes	no	no	no	fish
whale	warm-blooded	hair	yes	yes	no	no	no	mammal
frog	cold-blooded	none	no	semi	no	yes	yes	amphibian
komodo dragon	cold-blooded	scales	no	no	no	yes	no	reptile
bat	warm-blooded	hair	yes	no	yes	yes	yes	mammal
pigeon	warm-blooded	feathers	no	no	yes	yes	no	bird
cat	warm-blooded	fur	yes	no	no	yes	no	mammal
leopard shark	cold-blooded	scales	yes	yes	no	no	no	fish
turtle	cold-blooded	scales	no	semi	no	yes	no	reptile
penguin	warm-blooded	feathers	no	semi	no	yes	no	bird
porcupine	warm-blooded	quills	yes	no	no	yes	yes	mammal
eel	cold-blooded	scales	no	yes	no	no	no	fish
salamander	cold-blooded	none	no	semi	no	yes	yes	amphibian
gila monster	cold-blooded	scales	no	no	no	yes	yes	?
platypus	warm-blooded	hair	no	yes	no	yes	no	?

5. Here are some examples from the iris data.

```
A1, A2, A3, A4, Class
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
4.6,3.1,1.5,0.2,Iris-setosa
5.0,3.6,1.4,0.2,Iris-setosa
7.0,3.2,4.7,1.4,Iris-versicolor
6.4,3.2,4.5,1.5,Iris-versicolor
6.9,3.1,4.9,1.5,Iris-versicolor
5.5,2.3,4.0,1.3,Iris-versicolor
6.5,2.8,4.6,1.5,Iris-versicolor
5.7,2.8,4.5,1.3,Iris-versicolor
6.3,3.3,6.0,2.5,Iris-virginica
5.8,2.7,5.1,1.9,Iris-virginica
7.1,3.0,5.9,2.1,Iris-virginica
6.3,2.9,5.6,1.8,Iris-virginica
6.5,3.0,5.8,2.2,Iris-virginica
```

   How would the following two examples be classified?

```
6.2,3.4,5.4,2.3,?
5.7,2.8,4.1,1.3,?
```