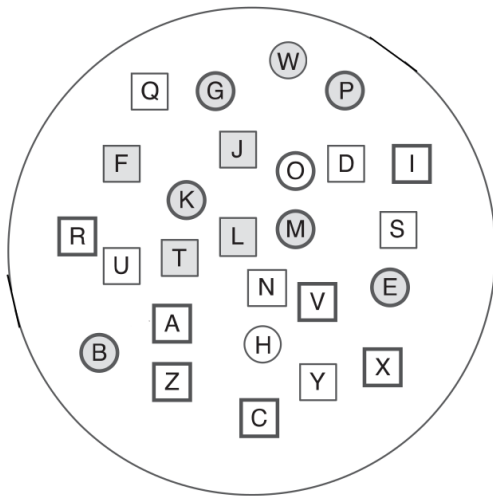# Exercise 1 - Information Gain

The following table shows data relating weather attributes with whether soccer games were played or not. Using information gain (IG) as criterion, determine which of the four weather attributes (Outlook, Temp, Humidity, Windy) is most strongly related to the variable Play.You can use the entropy table at the end of this lab sheet, which lists entropy values for some probability combinations.

Outlook	Temp	Humidity	Windy	Play
Sunny	Hot	High	False	No
Sunny	Hot	High	True	No
Overcast	Hot	High	False	Yes
Rainy	Mild	High	False	Yes
Rainy	Cool	Normal	False	Yes
Rainy	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Sunny	Mild	High	False	No
Sunny	Cool	Normal	False	Yes
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Rainy	Mild	High	True	No

Exercise 2 - Evidence Lift

For the data set in the picture, represented by shapes, colours and thickness of border of the elements in the picture, calculate the evidence lift:

1. of squares for thick borders

2. of thick borders for squares

3. explain the results you got in 1. and 2.

4. of thick borders for circles

Source: [MSD]

Technological University Dublin
Tallaght Campus
Department of Computing

Data Analysis
Pen and Paper Lab Assignment
Relationships (ML)

Formulae and Tables

Entropes for different value distributions of a dichotomous variable

Number of items	Breakdown		Entropy
1	0	1	0.00
2	1	1	1.00
3	1	2	0.92
4	1	3	0.81
5	1	4	0.72
5	2	3	0.97
6	1	5	0.65
7	1	6	0.59
7	2	5	0.86
7	3	4	0.99
8	1	7	0.54
8	3	5	0.95
9	1	8	0.50
9	2	7	0.76
9	4	5	0.99
10	1	9	0.47
10	3	7	0.88
11	1	10	0.44
11	2	9	0.68
11	3	8	0.85
11	4	7	0.95
11	5	6	0.99
12	1	11	0.41
12	5	7	0.98
13	1	12	0.39
13	2	11	0.62
13	3	10	0.78
13	4	9	0.89
13	5	8	0.96
13	6	7	1.00
14	1	13	0.37
14	3	11	0.75
14	5	9	0.94
15	1	14	0.35
15	2	13	0.57
15	4	11	0.84
15	6	9	0.97
15	7	8	1.00

References

[**MSD**] *Making Sense of Data I: A Practical Guide to Exploratory Data Analysis and Data Mining*, by Glenn J. Myatt and Wayne P. Johnson, John Wiley & Sons, 2014.