

COMP20008 2020 SM2 Workshop Week 9: Classification

1. Consider the following data set for a binary class problem and consider building a decision tree using this data

Feature A	Feature B	Class Label
T	F	+
T	T	+
T	T	+
T	F	-
T	T	+
F	F	-
F	F	-
F	F	-
T	T	-
T	F	-

- Write a formula for the information gain when splitting on feature A.
 - Write a formula for the information gain when splitting on feature B.
 - Which feature would the decision tree induction algorithm choose?
2. Load the **2020SM2-workshop-week9-classification.ipynb** jupyter notebook and complete the two practical exercises.