Information Theory for Complex Systems
Homework 1

Vincent Udén
udenv@student.chalmers.se

Januari - 2023

Below follows my solution for determining the optimal order of attributes using the ID3 algorithm implemented in Python. Before running any code I santised the csv file manually to remove trailing and leading whitespace. The code is attached at the end of the report and is availible on https://github.com/vincent-uden/tif150-information-theory-for-complex-systems.

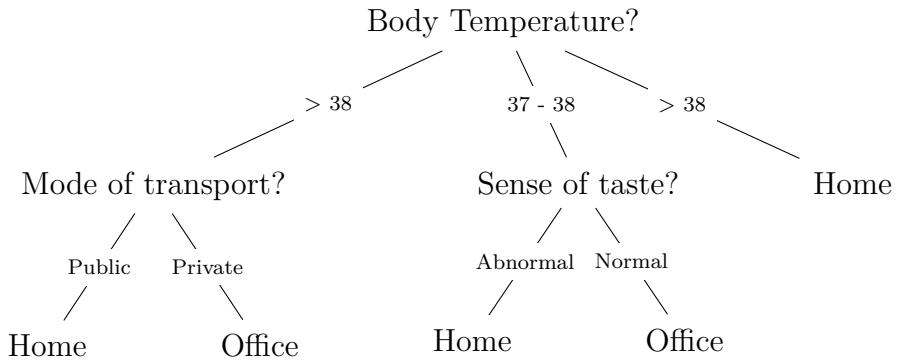

Figure 1: The sequence of optimal attributes for separating individuals working from home or at the office.

Figure 1 shows the optimal attribute order which maximises the information gained at each observation. It also shows the values of the chosen attributes along the edges of the tree. Entropy values are presented below in table 1.

Table 1: Shannon entropy at progressing stages of the ID3 algorithm.

Attribute choice	$H(S\|A)$ at the step	$H(t)$ after step
Start (before any choice)	-	4.39
Body Temperature ($>$ 38)	1.18	1.65
Body Temperature (37 - 38)	1.18	1.65
Body Temperature ($<$ 37)	1.18	0.0
Mode of tranport (Public)	0.0	0.0
Mode of tranport (Private)	0.0	0.0
Sense of taste (Abnormal)	0.0	0.0
Sense of taste (Normal)	0.0	0.0