Q1. For the following dataset shown in Figure 1, now let us assume that date is also an input feature. Which attribute will be chosen as the root node for the decision tree? Do you think choosing this attribute is a good choice?

Ans :

An attribute which has the highest information gain will be chosen as a root node of the decision tree so if the date will be chosen as root node then it will perform badly because we know that to predict output we need unseen future data. That is why we will select an attribute which has the second-highest information gain, which is an "Outlook".

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Figure 1: PlayTennis Dataset