

## AI Academy: Introduction to Data Mining

### Week 7 Workshop

---

Workshop 7 contains 2 questions.

---

# 1 BN Inference (12 points)

The following dataset presents 3 categorical attributes: Gender (M, F), Car Type (Sports, Luxury) and Age Group (G1, G2) with one Class Variable: Class (C0, C1). For each question, please show how you arrived at your answer.

Gender	Car Type	Age Group	Class
M	Luxury	G2	C0
M	Sports	G1	C0
M	Sports	G1	C1
M	Luxury	G1	C1
M	Luxury	G2	C0
F	Sports	G1	C1
F	Luxury	G2	C1
F	Luxury	G1	C0
F	Sports	G1	C0
F	Luxury	G1	C1

Table 1: Dataset for BN Inference

For the following problem, you may find it useful to fill in the following table (optional).

$P(Class = C0) =$	$P(Class = C1) =$
$P(Gender = M \mid Class = C0) =$	$P(Gender = M \mid Class = C1) =$
$P(Gender = F \mid Class = C0) =$	$P(Gender = F \mid Class = C1) =$
$P(CarType = Luxury \mid Class = C0) =$	$P(CarType = Luxury \mid Class = C1) =$
$P(CarType = Sports \mid Class = C0) =$	$P(CarType = Sports \mid Class = C1) =$
$P(AgeGroup = G1 \mid Class = C0) =$	$P(AgeGroup = G1 \mid Class = C1) =$
$P(AgeGroup = G2 \mid Class = C0) =$	$P(AgeGroup = G2 \mid Class = C1) =$

Using the training dataset above, how would a Naive Bayes classifier classify the following data points? Show your work.

1.  $\{Gender = M, Car Type = Luxury, Age Group = G1\}$
2.  $\{Gender = M, Car Type = Sports, Age Group = G2\}$
3.  $\{Gender = F, Car Type = Sports, Age Group = G1\}$
4.  $\{Gender = F, Car Type = Luxury, Age Group = G2\}$

## 2 BN Inference (12 points)

Compute the following probabilities according to the Bayesian net shown in Figure 1. **Note:**  $P(A)$  means  $P(A = \text{true})$ ;  $P(\sim A)$  means  $P(A = \text{false})$ .

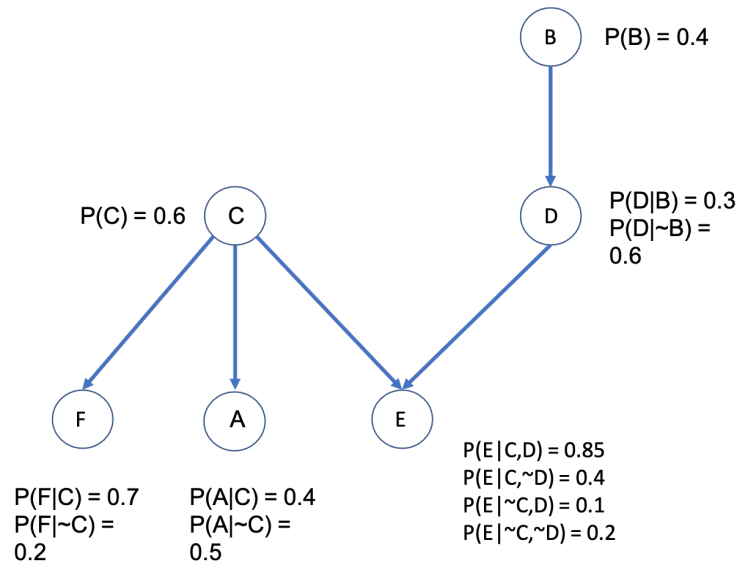


Figure 1: BN Inference

1. Compute  $P(A)$ . Show your work.
2. Compute  $P(D|B, \sim A)$ . Show your work.
3. Compute  $P(A, B, \sim C, D, E, F)$ . Show your work.
4. Are  $E$  and  $F$  conditionally independent given  $C$ ? Justify your answer in 1 sentence.
5. Are  $A$  and  $B$  marginally independent? Justify your answer in 1 sentence.
6. Given evidence that  $A = \text{true}$ ,  $C = \text{true}$ ,  $D = \text{false}$ , and  $F = \text{true}$ , use the Bayes Net to predict whether  $E$  is more likely to be *true* or *false*, or whether both are equally likely.