Solution of Q2 of the exam paper 2010

Data:

$P(\text{Def})=0.05$
$P(\text{U})=0.10$
$P(\text{Deb})=0.20$
$P(\text{U} \mid \text{Def})=0.40$
$P(\text{Deb} \mid \text{Def})=0.80$

There are two alternative solutions to this problem.

Solution A:

We are asked to find the posterior probability of default given that the borrower is both unemployed and has other debts:

$P(\text{Def} \mid \text{Deb} \wedge \text{U})$

Applying Bayes' theorem, this is:

$$P(\text{Def} \mid \text{Deb} \wedge \text{U}) = P(\text{Deb} \wedge \text{U} \mid \text{Def}) \ \frac{P(\text{Def})}{P(\text{Deb} \wedge \text{U})} \tag{1}$$

We assume that U and Deb are <u>both</u> "independent" and "conditionally-independent given the value of Def". This is a strong assumption because it ignores the fact that a person who is unemployed is also likely to have made other debts in order to pay their bills.

Under these two assumptions, we have (respectively):

$P(\text{Deb} \wedge \text{U}) = P(\text{Deb}) \ P(\text{U}) = 0.2 \cdot 0.1 = 0.02$
$P(\text{Deb} \wedge \text{U} \mid \text{Def}) = P(\text{Deb} \mid \text{Def}) \ P(\text{U} \mid \text{Def}) = 0.8 \cdot 0.4 = 0.32$

If we plug in these figures in eq (1), we obtain:

$$P(\text{Def} \mid \text{Deb} \wedge \text{U}) = P(\text{Deb} \wedge \text{U} \mid \text{Def}) \ \frac{P(\text{Def})}{P(\text{Deb} \wedge \text{U})} = 0.32 \cdot \frac{0.05}{0.02} = 0.8$$

Solution B:

We are asked to find the posterior probability of default given that the borrower is both unemployed and has other debts:

$P(\text{Def} \mid \text{Deb} \wedge \text{U})$.

Applying Bayes' theorem, this probability is:

$$P(\text{Def} \mid \text{Deb} \wedge \text{U}) = P(\text{Deb} \wedge \text{U} \mid \text{Def}) \ \frac{P(\text{Def})}{P(\text{Deb} \wedge \text{U})} \ . \tag{1}$$

We assume that U and Deb are "conditionally-independent given the value of Def". This is a strong assumption but weaker than that made in Solution A. The assumption we make here coincides with a Naive Bayes approach where the input attributes (Deb and U) are assumed to be conditionally

independent given the output class (Def).
Under this assumption, we have:

$$P(\text{Deb} \wedge \text{U} \mid \text{Def}) = P(\text{Deb} \mid \text{Def})\ P(\text{U} \mid \text{Def}) = 0.8 \cdot 0.4 = 0.32$$

and also:

$$P(\text{Deb} \wedge \text{U} \mid \neg\text{Def}) = P(\text{Deb} \mid \neg\text{Def})\ P(\text{U} \mid \neg\text{Def}) \quad . \tag{2}$$

Applying the "law of total probability" (see the class notes of the first afternoon class), we have:

$$P(\text{Deb}) = P(\text{Deb} \mid \text{Def})\ P(\text{Def}) + P(\text{Deb} \mid \neg\text{Def})\ P(\neg\text{Def}) \quad .$$

that we can use to find:

$$P(\text{Deb} \mid \neg\text{Def}) = \frac{P(\text{Deb}) - P(\text{Deb} \mid \text{Def})\ P(\text{Def})}{1 - P(\text{Def})} = \frac{0.2 - 0.8 \cdot 0.05}{1 - 0.05} \simeq 0.168 \quad .$$

Similarly, we have:

$$P(\text{U} \mid \neg\text{Def}) = \frac{P(\text{U}) - P(\text{U} \mid \text{Def})\ P(\text{Def})}{1 - P(\text{Def})} = \frac{0.1 - 0.4 \cdot 0.05}{1 - 0.05} \simeq 0.084 \quad .$$

Replacing these two values in (2), we obtain:

$$P(\text{Deb} \wedge \text{U} \mid \neg\text{Def}) \simeq 0.168 \cdot 0.084 \simeq 0.014 \quad .$$

Applying the "law of total probability" again, we have:

$$\begin{aligned} P(\text{Deb} \wedge \text{U}) &= P(\text{Deb} \wedge \text{U} \mid \neg\text{Def})\ P(\neg\text{Def}) + P(\text{Deb} \wedge \text{U} \mid \text{Def})\ P(\text{Def}) \\ &\simeq 0.014 \cdot 0.95 + 0.32 \cdot 0.05 \simeq 0.029 \end{aligned}$$

Please note that this value is significantly higher than the corresponding value computed in Solution A.
We can finally compute (1):

$$P(\text{Def} \mid \text{Deb} \wedge \text{U}) = P(\text{Deb} \wedge \text{U} \mid \text{Def})\ \frac{P(\text{Def})}{P(\text{Deb} \wedge \text{U})} \simeq 0.32\ \frac{0.05}{0.029} \simeq 0.551 \quad .$$

As Solution B only requires a subset of the assumptions made for Solution A, the former is clearly the preferred solution. If correctly and clearly presented, Solution B (or an equivalent one) will earn you the full mark for the exercise.