

IS5152 Data-driven Decision Making
SEMESTER 2 2023-2024
Assignment 3

- Upload a softcopy of your solution as a **pdf** document to Canvas.
- Deadline: 11.59 pm, Friday, 29 March, 2024.

-
1. (10 points) Producto Corporation has experienced a substantial increase in sales during the past few years. In order to keep up with demand it appears necessary to expand production lines. Of course, Producto could decide to do nothing to keep up with the anticipate demand, simply keeping up existing facilities with minimal maintenance. They could also modernize the existing facilities, contract some of the work to a subcontractor, or build a new additional plan.

The four potential strategies will produce different profits in conjunction with any of three perceived states in the future: declining demand (D_1), no change in demand (D_2), and increasing demand (D_3). Profits for each strategy under each state have been estimated, as well as probabilities for occurrences of each state. These estimates are given in the following table:

Strategy	Type of future demand		
	D_1	D_2	D_3
Maintenance	300,000	25,000	-100,000
Modernization	200,000	10,000	-50,000
Subcontracting	-25,000	0	100,000
New plant	-500,000	-200,000	500,000
$P(D_i)$	0.15	0.60	0.25

- (a) What is the best strategy to maximize the potential profit?
- (b) What is the Expected Value With Perfect Information **EVWPI**?
- (c) What is the best strategy using the minimax-regret criterion? (note: ignore the probabilities $P(D_i)$).
2. (10 points) You are planning to publish your first book “My life as a student”. You can either publish the book yourself or through a publisher. If you decide to publish the book yourself, the initial cost for printing and promoting the book is \$25000 and each copy of the book sold will net you a profit of \$10.

The publisher is offering you \$8000 for signing the contract and pay you a royalty of \$1 for each book sold.

Past records show that there is a probability of 30% that an autobiographical book by a first time author will be a success where 8000 copies were sold. When the book is not successful, it would sell only 1000 copies.

- (a) Based on the information provided, what is the best course of action?

- (b) Suppose that you may hire a literary agent to give you a review of your book and a prediction of the potential success of the book. From past experience, you know that when a book is successful, the agent will predict the correct outcome 75% of the time. When the book was not successful, the agent was correct 90% of the time. How would this information affect your decision?
- (c) Compute the Expected Value of Perfect Information and the Expected Value of Sample Information.
3. (10 points) The following table consists of training data from an employee dataset. The attribute *department* has 3 possible values: sales, systems, marketing. The attribute *age* has 4 possible values: [21 to 25], [26 to 30], [31 to 35], [36 to 40]. The attribute *salary* has 3 possible values: < 30K, [30K to 40K], > 40K. The target attribute *status* is binary-valued: junior or senior.

department	age	salary	status
sales	[31 to 35]	> 40K	senior
sales	[26 to 30]	[30K – 40K]	junior
systems	[21 to 25]	< 30K	junior
systems	[36 to 40]	> 40K	senior
systems	[26 to 30]	> 40K	junior
systems	[31 to 35]	[30K – 40K]	junior
marketing	[31 to 35]	[30K – 40K]	senior
marketing	[36 to 40]	[30K – 40K]	senior
marketing	[26 to 30]	> 40K	senior
marketing	[21 to 25]	[30K – 40K]	junior

- (a) What would be the naive Bayes prediction for the status of an employee with the attribute values: (marketing,[31 to 35], > 40K)?
- (b) Laplace smoothing is applied to the data with parameter k set to 3. What is the predicted status for an employee with attribute values: (sales,[36 to 40],< 30K)?