Aim

- Learn the differences between how humans' reason versus how the state-of-the-art foundation model chat bots do
- Model search problems and understand how uninformed search strategies work
- Learn how to choose a supervised model which suits your problem

Assessment

To pass this lab write down your answers and explain them to Elmira at the lab.

- For task 4, implement the code in Python and demonstrate it at the lab.
- For tasks 3 and 5, feel free to solve the problem either by hand or implementing a code in Python and demonstrate it at the lab.

Requirements	Grade
Tasks 1, 2, 3	3
Tasks 1, 2, 3, and 4	4
All tasks	5

Task 1

Try the following prompts with Perplexity without logging in (so you get the default models).

A. how many “i”s are there in “Incomprehensibilities”?
B. I have four cats. Three of them are named North, East, West. What is the name of the fourth cat. Can you name all of my cats?
C. What is the heaviest moon of Saturn?
D. I drive one km towards North, one km towards West and then one km towards South. If I end up at the same position where I started, where am I?

1- For which prompts did you get a correct answer?
2- Compare the type of tasks and how the human mind would approach finding an answer to each task versus how LLM-based chat bots are designed to respond to them and explain why the chatbot could not answer one of the questions correctly?

Task 2

Consider the following state space graph to move from starting node, S to the goal node, G.

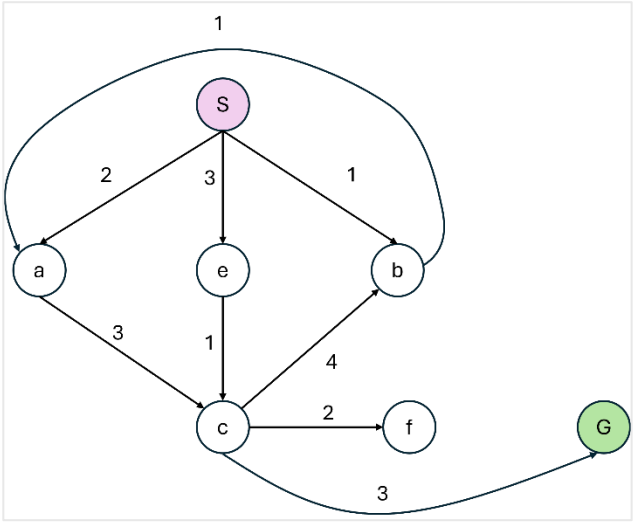


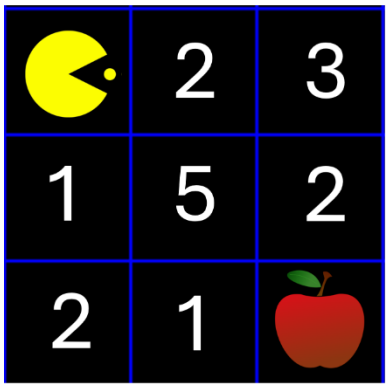
Figure 1: The state space graph to move from S to G

Draw the search tree and specify the order of the nodes that each of the following search strategies will follow to reach the goal.

Search Strategy	Order of explored nodes
BFS	
DFS	
IDS	
UCS	

Task 3

In the following grid maze, Pacman needs to reach an apple in the bottom-right corner. Each cell in the grid has a different cost associated with moving into it, representing the risk of encountering ghosts. The cost of the starting cell is also 1.



1. Create the search tree.
2. Find the path from start to finish with the lowest total cost.
3. Explain what type of uninformed search technique (DFS, BFS, IDS, and UCS) would you find most appropriate to solve this problem and why?

Task 4

Download the Python code for task 4. Make sure you have understood what each part of the code does. Comment on different parts of the code to specify which parts perform each of the following steps.

1. Load the data
2. Visualize the data
3. Fitting a model to the training data
4. Transforming the test data to find decision boundary
5. Visualizing the decision boundary

Add your own script to fit a random forest classifier to the training data. Create two subplots, one at the top and the other at the bottom. Compare the decision boundary specified by SVM and random forest on the two plots. Which one do you find to be a better classifier for this data?

Task 5

A video streaming service that offers monthly subscriptions wants to predict which customers are likely to cancel their subscriptions in the next month. The company has data on 20 customers as, and whether they canceled their subscription, retained or upgraded.

Customer	Subscription Period	Monthly Usage	Support Calls	Price Tier	Subscription Decision
1	> 2 years	Low	No	Basic	Active
2	< 1 year	Low	Yes	Standard	Canceled
3	> 2 years	Medium	No	Premium	Active
4	< 1 year	Low	No	Basic	Canceled
5	> 2 years	Low	No	Standard	Upgraded
6	1 – 2 years	Low	Yes	Basic	Canceled
7	> 2 years	Medium	No	Premium	Active
8	1 – 2 years	Low	No	Standard	Active
9	> 2 years	Low	No	Standard	Upgraded
10	< 1 year	Low	Yes	Basic	Canceled
11	> 2 years	High	No	Premium	Active
12	1 – 2 years	Low	No	Basic	Canceled
13	> 2 years	High	No	Premium	Active
14	1 – 2 years	Low	No	Standard	Active
15	< 1 year	Low	Yes	Basic	Canceled
16	> 2 years	Medium	Yes	Standard	Upgraded
17	1 – 2 years	Low	No	Standard	Active
18	> 2 years	High	No	Premium	Active
19	1 – 2 years	Low	Yes	Basic	Canceled
20	> 2 years	Medium	No	Standard	Active

1- Create a decision tree classifier to predict the outcome for a new customer with the following data:

```
new_customer = {
        "Subscription length": "1-2 years",
        "Monthly usage": "High",
        "Support calls": "No",
        "Price tier": "Standard"
    }
```

2- Specify the information gain of the data split at each node of your tree.