**Assessment Task – Total Marks: 100 – Weighting: 60%**

The assignment has three tasks: **one individual (Task 1)** and **two tasks to be completed as a group.**

# Task 1

Using the following dataset, create a decision tree using the **entropy**.

Suppose the following dataset is about the properties of 14 people where the attribute “Default” shows whether a person will default on their credit card. The attribute “Default” is the dependent attribute with two values (default = 'yes', default = 'no'). Each person has 4 features, and you want to find out how these features are going to help whether a person will default on their credit card.

1. Which one of these features is the most important feature?
2. What is the best Information Gain (IG)?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Age** | **Income** | **Student** | **Credit\_rate** | **Default** |
| 0 | youth | high | no | fair | no |
| 1 | youth | high | no | excellent | no |
| 2 | middle\_age | high | no | fair | yes |
| 3 | senior | medium | no | fair | yes |
| 4 | senior | low | yes | fair | yes |
| 5 | senior | low | yes | excellent | no |
| 6 | middle\_age | low | yes | excellent | yes |
| 7 | youth | medium | no | fair | no |
| 8 | youth | low | yes | fair | yes |
| 9 | senior | medium | yes | fair | yes |
| 10 | youth | medium | yes | excellent | yes |
| 11 | middle\_age | medium | no | excellent | yes |
| 12 | middle\_age | high | yes | fair | yes |
| 13 | senior | medium | no | excellent | no |

**(20 marks)**

**Task 2**

Using the hand gestures dataset available on the Moodle page, create a hand gesture recognition app in Python. Apply at least three classifiers and compare their performance. Save your Anaconda/Google Colab Notebook in this format: **Your\_Name\_Group(A?,B?,C?)\_StudentNumber\_FaceRECOGNIZE.ipynb**

# Document your code and give full explanation about what each block of code does.

**(30 marks)**

**Task 3**

Use Google API to get reviews for different companies. (***A separate document “Google Places API Key generator.pdf” is uploaded to help you with this task.)***

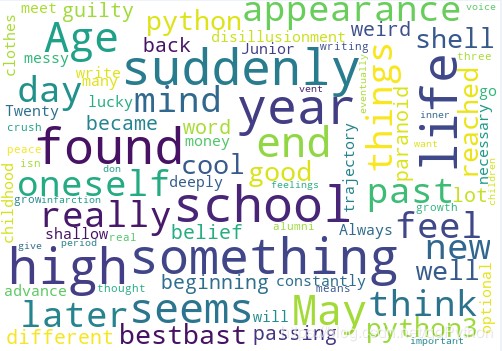
(It used to be in the range of 10,000 reviews that it would let you to download but even if you get anything above 1000 per company, it is

acceptable. NOTE: You can create an API or a Developer account on other platforms like Twitter, Facebook or Google Reviews …. as well)

1. Download the reviews for four companies namely ( 1- Google, Amazon,
2. Linkedin,
3. Tesla)

Using text mining technique, identify the top 10 MOST USED WORDS in each company Reviews that you have downloaded.

Represent your findings in Word Cloud Illustration for each company like the following example. The size of the words gets bigger as they become more frequently used.



Save your Anaconda / Colab Notebook File in **Your\_Name\_Group(A?,B?,C?)\_StudentNumber\_TXTMiner.ipynb format**. **Document your code and give full explanation about what each block of code does.**

# (40 marks)

Individual Contribution (in completing Task2 and Task3) - 400 words

**(10 marks)**

*Each group member must write individually about their contribution and reflection of learning in doing Task 2 and Task3.*

**Assessment Criteria**

Each part will be graded according to the following criteria:

1. Quality of code (correctness and completeness) [Weightage – 40%]
2. Quality of documentation of code ( explanation of each part of code and interpretation of results) [Weightage – 60%]