|  |  |  |
| --- | --- | --- |
|  |  | **Année : 2021**  **cours : 420 – A52-BB** |

Examen intra – Partie Pratique

Ricardo Vallejo R.

*Les notes de cours sont autorisées dans cette partie*

**Exercice 2 : [20 pts]**

Note : l’exercice sera fait sur papier.

Soit le tableau donnant la qualité d’une réponse d’un candidat correspondant aux deux variables temps de réflexion et complexité de la question :

|  |  |  |
| --- | --- | --- |
| Temps de réflexion(min) | Complexité de la question | Qualité de la réponse |
| 10 | simple | Excellente |
| 10 | élevé | Moyenne |
| 20 | moyenne | Bonne |
| 30 | simple | Bonne |
| 40 | moyenne | Excellente |
| 40 | moyenne | Bonne |
| 50 | élevé | Bonne |

En utilisant l’approche Naive Bayes,

Q.1 : Trouver la fréquence de chaque qualité dans les trois catégories**[5 pts]**

**P(y=Excelente) = 2/7**

**P(y=Bonne) = 4/7**

**P(y=Moyenne) = 1/7**

**P(Temps 0-10) = 2/7**

**P(Temps 11-20) =1/7**

**P(Temps 21-30) =1/7**

**P(Temps 31-40) =2/7**

**P(Temps 41-50) =1/7**

**P(Q Simple) = 2/7**

**P(Q Eleve) = 2/7**

**P(Q Moyene) = 3/7**

**Temps :**

**P(Temps 0-10 | Excellent) = 1/2**

**P(Temps 11-20 | Excellent) =0**

**P(Temps 21-30 | Excellent) =0**

**P(Temps 31-40 | Excellent) =1/2**

**P(Temps 41-50 | Excellent) =0**

**P(Temps 0-10 |** Moyenne**) = 1**

**P(Temps 11-20 |** Moyenne**) =0**

**P(Temps 21-30 |** Moyenne**) =0**

**P(Temps 31-40 |** Moyenne**) =0**

**P(Temps 41-50 |** Moyenne**) =0**

**P(Temps 0-10 |** Bonne**) =0**

**P(Temps 11-20 |** Bonne**) =1/4**

**P(Temps 21-30 |** Bonne**) =1/4**

**P(Temps 31-40 |** Bonne**) =1/4**

**P(Temps 41-50 |** Bonne**) =1/4**

**Complexité :**

**P(**simple **| Excellent) = 1/2**

**P(**élevé **| Excellent) = 0**

**P(**moyenne **| Excellent) = 1/2**

**P(**simple **|** Moyenne**) = 0**

**P(**élevé **|** Moyenne**) = 1**

**P(**moyenne **|** Moyenne**) =0**

**P(**simple **|** Bonne**) = 1/4**

**P(**élevé **|** Bonne**) =1/4**

**P(**moyenne **|** Bonne**) =2/4**

Q.2 : Évaluer les probabilités conditionnelles de chaque prédicteur et à partir de votre résultat, indiquer la qualité de la réponse à prédire selon **un temps de réflexion de 20 min** et **une complexité = élevé [15 pts]**

Temps = 20 min

Complexite = Eleve

Qualite Reponse = ????

P(Excellente | 20min, eleve) = 0

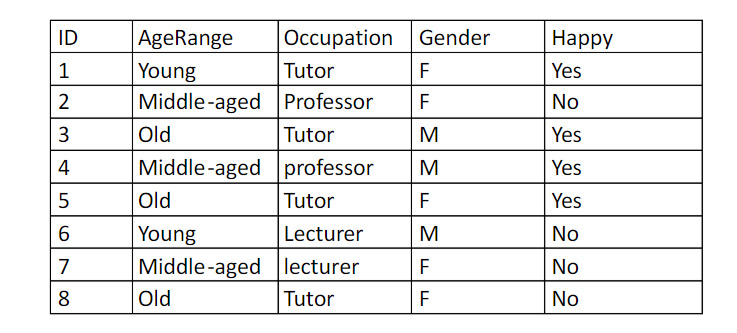
P(Moyenne | 20min, eleve) = 0

P(Bonne | 20min, eleve) = ( 1/4 \* 1/4) \* (2/7) / (1/7 \* 2/7) = 0,43

La plus probable cest Un response de Qualite Bonne.

**Exercice 3 : Classification arbre de décision [30 pts]**

Soit le dataset suivant :



**Question 1 : 20 points**

Développer le modèle arbre de décision en utilisant le langage python

Calculer l'accuracy sur le dataset de training.

**Question 2 : 10 points**

En utilisant l’approche arbre de décision, determiner l’état de la personne ayant les attributs suivants :

AgeRange : young

Occupation : professor

Gender : F

**Noter que les étapes de développement du modèle et son utilisation doivent être indiquées dans un Notebook python.**