

Exercices, consignes et annexes

Rédigés par Nida Meddouri (nida.meddouri@epita.fr) et Marie Puren (marie.puren@epita.fr)

Le devoir doit être déposé sur Moodle.

Barème

Exercice 1 : 8 points

- Question 1 : 2 points
- Question 2 : 6 points

Exercice 2 : 12 points

- Question 1 : 4 points
- Question 2 : 4 points
- Question 3 : 4 points

Exercice 1 (...../8) :

Un ambulancier-urgentiste intervient à la suite d'un appel d'urgence. Il dispose d'un outil d'aide à la décision doté d'une intelligence artificielle.

Il arrive sur scène et il fait entrer des données et des informations concernant l'état du patient. L'outil d'aide à la décision analyse ces données et applique un algorithme de Machine Learning dit **CNC (Classifier Nominal Concept)**¹ pour prédire l'évolution de l'état du patient.

L'algorithme du CNC est le suivant :

L'ALGORITHME CLASSIFIER NOMINAL CONCEPT

ENTREE : Séquence de "m" observations $\{(o_1, y_1), \dots, (o_m, y_m)\}$ décrites par "n" caractéristiques nominales et étiquetées par y_i avec $i=\{1, \dots, k\}$

SORTIE : Une règle de classification

(de la forme SI variable¹ ET variable² ET ... ALORS étiquette).

DEBUT :

1. Déterminer le descripteur (caractéristique/attribut) qui a le gain informationnel le plus important.
S'il y a plusieurs descripteurs ayant la même valeur du gain, on choisit le premier.
2. Calculer la fermeture associée à la valeur la plus pertinente de ce descripteur afin de générer un concept pertinent.
3. Déterminer la classe (étiquette) majoritaire associée à ce concept pertinent.

¹ N. Meddouri, M. Maddouri: Efficient closure operators for FCA based classification. International Journal of Artificial Intelligence and Machine Learning (IJAIML), Volume 10, Issue 2, 2020.

4. Dédurre la règle de classification/prédiction :

On associe la classe majoritaire à ce concept pertinent et on obtient la règle de classification associé.

FIN.

Dans un premier temps, un algorithme de Machine Learning doit disposer d'un ensemble initial de données d'apprentissage (des observations antérieures). L'ensemble initial des données d'apprentissage dans notre cas est le suivant :

	Outlook	Temperature	Transpiration	Drug	Label (urgence)
O ₁	sunny	hot	high	false	no
O ₂	sunny	hot	high	true	no
O ₃	overcast	hot	high	false	yes
O ₄	rainy	mild	high	false	yes
O ₅	rainy	cool	normal	false	yes
O ₆	rainy	cool	normal	true	no
O ₇	overcast	cool	normal	true	yes
O ₈	sunny	mild	high	false	no
O ₉	sunny	cool	normal	false	yes
O ₁₀	rainy	mild	normal	false	yes
O ₁₁	sunny	mild	normal	true	yes
O ₁₂	overcast	mild	high	true	yes
O ₁₃	overcast	hot	normal	false	yes
O ₁₄	rainy	mild	high	true	no
<i>GainInfo</i>	<i>0.24</i>	<i>0.02</i>	<i>0.15</i>	<i>0.04</i>	<i>NA</i>

Dans un deuxième temps, l'outil en question applique l'algorithme du CNC pour générer un modèle d'apprentissage (sous forme de règles de classification SI ... ALORS ...).

Enfin, la règle générée, servira à prédire l'état d'un nouveau patient selon ses informations / descripteurs d'état.

Dans notre cas actuel, l'ambulancier-urgentiste constate les descripteurs d'état suivant :

- i. Outlook = sunny
- ii. Temperature = mild
- iii. Transpiration = normal
- iv. Drugs = True

L'ambulancier décide de faire confiance à l'outils doté d'une Intelligence Artificielle pour prendre une décision.

Questions :

- **D'après l'algorithme CNC, quelle décision va être proposée ?** Expliquez rapidement comment vous arrivez à ce résultat. Pour vous aider, référez-vous à l'annexe 1.
- **Est ce responsable et éthique de s'appuyer totalement sur l'outil d'aide à la décision ?** Développez votre réponse en expliquant ce que vous feriez et pourquoi.

Exercice 2 (...../12) :

Une voiture électrique dotée d'un programme d'intelligence artificielle pour une conduite autonome, provoque un accident mortel. Le témoignage des témoins au moment de l'accident indique la responsabilité de la voiture électrique dans le sinistre. On considère qu'il y a 3 origines possibles de la défaillance :

- i. Défaillance matérielle à cause d'un manque d'entretien,
- ii. Défaillance matérielle faute d'un défaut de fabrication,
- iii. Défaillance de conduite autonome à cause de l'algorithme de décision utilisé.

Questions :

- D'abord, en vous appuyant sur le cours et sur les références données en annexe, expliquez, pour chacun des cas, qui est responsable selon vous et pourquoi. Commencez par identifier chacun des acteurs impliqués et développez de manière détaillée votre avis, en fonction des différentes origines de la défaillance.
- Ensuite, expliquez si, selon vous, le dilemme du tramway, peut et/ou doit être utilisé pour déterminer les responsabilités en cas d'accidents causés par une voiture autonome.
- Enfin, donnez votre avis. Par exemple, si vous deviez programmer le logiciel d'une voiture autonome, sur quelle(s) critère(s) éthiques vous baseriez-vous ? Ou pensez-vous qu'il s'agisse d'un problème purement technique ?

Vous développerez et justifierez vos réponses en vous appuyant sur le cours et sur les références listées en annexe 2.

Annexe 1 : Comment prédire en utilisant CNC (Classifier Nominal Concept)

Supposons que nous avons le contexte d'apprentissage (des données/observations) suivant (un petit extrait du précédent):

	Outlook	Temperature	Transpiration	Drug	Label (urgence)
o ₂	sunny	hot	high	true	no
o ₆	rainy	cool	normal	true	no
o ₉	sunny	cool	normal	false	yes
o ₁₀	rainy	mild	normal	false	yes
o ₁₃	overcast	hot	normal	false	yes
<i>Gain Info</i>	<i>0.17</i>	<i>0.17</i>	<i>0.32</i>	<i>0.97</i>	<i>NA</i>

Ce contexte représente 5 observations (o₂, o₆, o₉, o₁₀ et o₁₃). Chaque observation est décrite par 4 descripteurs/caractéristiques/attributs nominaux (Outlook, Température, Transpiration et Drug). Ces

observations sont étiquetées par deux labels (étiquettes) : « yes » et « no » (urgence médicale ou pas).

- **Etape 1 : Déterminer le descripteur (caractéristiques) qui a le gain informationnel le plus important.**

Les valeurs des gains associées à chaque descripteur sont données dans le tableau (dernière ligne). Nous retenons le descripteur Drug qui a 0.97 (le gain le plus important)

	Outlook	Temperature	Transpiration	Drug	Label (urgence)
o ₂	sunny	hot	high	true	no
o ₆	rainy	cool	normal	true	no
o ₉	sunny	cool	normal	false	yes
o ₁₀	rainy	mild	normal	false	yes
o ₁₃	overcast	hot	normal	false	yes
<i>Gain Info</i>	<i>0.17</i>	<i>0.17</i>	<i>0.32</i>	<i>0.97</i>	<i>NA</i>

- **Etape 2 : Calculer la fermeture associée à la valeur la plus pertinente de cet attribut afin de générer un concept pertinent.**

Le descripteur Drug dispose de deux valeurs nominales : « true » et « false ».

- TRUE est présent dans deux observations : o₂ et o₆.
- FALSE est présent dans trois observations : o₉, o₁₀ et o₁₃.

Nous retenons dans ce cas la valeur false, et nous allons calculer sa fermeture. Dans un premier temps, nous cherchons les observations qui partagent FALSE à savoir { o₉, o₁₀ et o₁₃ }

	Outlook	Temperature	Transpiration	Drug	Label (urgence)
o ₂	sunny	hot	high	true	no
o ₆	rainy	cool	normal	true	no
o₉	sunny	cool	normal	false	yes
o₁₀	rainy	mild	normal	false	yes
o₁₃	overcast	hot	normal	false	yes

Dans un deuxième temps, nous cherchons les descripteurs partagés par {o₉, o₁₀ et o₁₃} et nous aurons Transpiration=normal et Drug=FALSE.

	Outlook	Temperature	Transpiration	Drug	Label (urgence)
o ₂	sunny	hot	high	TRUE	no
o ₆	rainy	cool	normal	TRUE	no
o₉	sunny	cool	normal	false	yes
o₁₀	rainy	mild	normal	false	yes
o₁₃	overcast	hot	normal	false	yes

Notre concept pertinent sera ({o₉, o₁₀, o₁₃}, {Transpiration=normal, Drug=FALSE})

- **Etape 3 : Déterminer la classe majoritaire associée à ce concept pertinent généré.**

Dans ce cas, la classe majoritaire est yes

	Outlook	Temperature	Transpiration	Drug	Label (urgence)
O₉	sunny	cool	normal	false	yes
O₁₀	rainy	mild	normal	false	yes
O₁₃	overcast	hot	normal	false	yes

- **Etape 4 : Dédire la règle de classification/prédiction.**

Notre concept pertinent est $\{o_9, o_{10}, o_{13}\}$, $\{Transpiration=normal, Drug=FALSE\}$ et associé à Label(urgence)=yes, donc la règle de prédiction sera :

SI Transpiration=normal **ET** Drug=false **ALORS** Label(urgence)=yes

Annexe 2 : Documents et références à utiliser pour l'exercice 2

- Une vidéo pour commencer car elle explique ce qu'est le dilemme du tramway et donne un aperçu de quelques théories éthiques applicables à la voiture autonome : ARTE, "Les voitures autonomes ont-elles une morale ?", *Gymnastique, la culture en s'amusant*, 19 mars 2022, <https://www.youtube.com/watch?v=icKz8P2Hf2o>
- Pour accéder à la *Moral Machine* évoquée dans la vidéo et tester différents scénarios : <https://www.moralmachine.net/>
- Le chercheur interviewé dans la vidéo, Jean-François Bonnefon, a beaucoup travaillé sur ce sujet :
 - Il a publié un article court dans *Science* dans lequel il examine la théorie utilitariste dans le cadre de la conduite autonome : Jean-François Bonnefon *et al.*, The social dilemma of autonomous vehicles. *Science* 352, 1573-1576 (2016), <https://www.science.org/doi/10.1126/science.aaf2654>. (le PDF est accessible sur Moodle). Ces résultats ont été résumés dans la presse et en ligne :
 - Laure Andrillon, "Jean-François Bonnefon : « la voiture autonome pose un dilemme éthique »", *Libération* 22 août 2016, https://www.liberation.fr/futurs/2016/08/22/jean-francois-bonnefon-la-voiture-autonome-pose-un-dilemme-ethique_1474033/
 - Jean-François Bonnefon, "Voiture autonomes : par quoi serez-vous choqués ?", *CNRS Le journal*, 19 octobre 2017, <https://lejournel.cnrs.fr/billets/voitures-autonomes-par-quoi-serez-vous-choques>
 - Un article portant dans lequel il évoque le dilemme du tramway : J. -F. Bonnefon, A. Shariff and I. Rahwan, The Trolley, The Bull Bar, and Why Engineers Should Care About The Ethics of Autonomous Cars [point of view], in *Proceedings of the IEEE*, vol. 107, no. 3, pp. 502-504, March 2019, <https://ieeexplore.ieee.org/abstract/document/8662742>

- La synthèse des résultats de la *Moral Machine* a été publiée dans cet article : Awad, E., Dsouza, S., Kim, R. *et al.* The Moral Machine experiment. *Nature* **563**, 59–64 (2018). <https://doi.org/10.1038/s41586-018-0637-6> (le PDF est accessible sur Moodle).
Heureusement, les résultats ont été résumés dans diverses publications plus faciles à lire :
 - Marc Lajoie, “Voitures autonomes. Qui sauver, qui sacrifier”, *Radio Canada*, 25 avril 2019, <https://ici.radio-canada.ca/info/2019/voitures-autonomes-dilemme-tramway/>
 - Allez à la **page 32** dans Israel Collier, Sacha. *Quelles règles éthiques pour les véhicules autonomes en cas d'accident ?*. Louvain School of Management, Université catholique de Louvain, 2020. Téléchargement direct du PDF avec ce lien : https://dial.uclouvain.be/downloader/downloader.php?pid=thesis%3A25644&datastream=PDF_01&cover=cover-mem
- D'autres pistes que le dilemme du tramway ?
 - Sarit Mizrahi, “La mise au point de véhicules autonomes : au-delà du dilemme du tramway 3.0”, *Laboratoire de Cyberjustice*, 27 février 2019, <https://www.cyberjustice.ca/2019/02/27/la-mise-au-point-de-vehicules-autonomes-au-dela-du-dilemme-du-tramway-3-0/> (vous pouvez directement consulter la dernière partie intitulée “Le dilemme du tramway 3.0”)
 - Cosmo, “Voiture Autonome : Laissons le Dilemme du Tramway à la Philosophie Morale !”, *Medium*, 11 décembre 2019, <https://medium.com/la-voiture-autonome/voiture-autonome-laissons-le-dilemme-du-tramway-%C3%A0-la-philosophie-morale-698f0e41205e>
 - Laurent Bibard, “Le terme « véhicule autonome », un non-sens philosophique”, *The Conversation*, 16 octobre 2022, <https://theconversation.com/le-terme-vehicule-autonome-un-non-sens-philosophique-191120>