

# Examen ' Conduite de projets en science des données et Predictive Analytics'

Calice Laetitia

## 1: IA

Décrire la différence entre approche réaliste et approche utilitariste *dans la démarche scientifique*.

**Une approche utilitariste dans une démarche scientifique** consiste à chercher à acquérir des connaissances dans le monde réel pour pouvoir avoir des éléments pour expliquer son raisonnement afin de justifier son but.

**Alors qu'une approche utilitariste, proche de l'Intelligence Artificielle, consiste à copier le monde réel comme il est, il ne s'agit pas de chercher les connaissances mais juste de faire du mimétisme.**

## 2: programmation logique/chainage avant

Dans un langage à base de règles simple en chaînage avant (on appelle cela un système de production) on a le programme:

```
var input[]=[], result[], i=1, tmp=0;  
  
when input.length>0 and i >= input.length then result.append(input[tmp]),  
input.removeAt(tmp), i=1, tmp=0;  
when input[i] < input[tmp] then tmp=i, i=i+1;  
when input[i] >= input[tmp] then i=i+1;
```

Lorsque l'instruction `input=[2,0,5,4,9];` est exécutée, que contiendront les variables `result` et `input` en retour? Expliquer l'algorithme. Donner sa complexité.

**Lorsque l'instruction `input=[2,0,5,4,9]` est exécutée, en retour `result` vaut 0 (`result=[0]`) et l'`input` vaut 2,5,4,9 (`input=[2,5,4,9]`).**

**Il s'agit d'un algorithme de tri. L'algorithme parcourt l'ensemble de la liste et y cherche l'élément qui a la plus petite valeur, il le stocke ensuite dans la variable `tmp`. A la fin de la liste la valeur de `tmp` est enlevée de `input` et mise dans `result`.**

**3: Smart City:** Trouver sur internet 3 logiciels commerciaux **professionnels** destinés à remplir un rôle similaire à celui du projet SmartDeliveries. En vous basant sur la présentation commerciale, identifiez leurs principales caractéristiques démarquantes (quelles fonctionnalités

mettent-ils particulièrement en avant par rapport à la concurrence). Fournir les références utilisées.

**Kardinal** (<https://kardinal.ai/fr/>) permet de créer et gérer des tournées optimisées en temps réel. Ainsi les opérateurs peuvent s'adapter aux aléas (retard, panne, embouteillage ...). De plus, Kardinal est dans une démarche auto-apprenante avec les éléments récoltés sur le terrain avec un algorithme de deep learning.

**Antsroute** (<https://antsroute.com/>) est utile pour optimiser les trajets et permet de pouvoir parcourir le même trajet en moins de kilomètres. Ainsi les livreurs peuvent réaliser plus de tournées.

**Mapotempo** (<https://www.mapotempo.com/>) permet de gérer les tournées des personnes ayant des activités de livraison, de collecte ou de service. L'outil permet de planifier, de calculer les tournées, de suivre les trajets, de géolocaliser et d'alerter le personnel mobile.

#### 4: trafic routier

a- Quelles sont les principales variables mesurées par un détecteur de trafic?

**Un détecteur de trafic permet de compter le nombre de voitures, la vitesse de ces dernières. Il mesure et détecte.**

b- Qu'est ce que le diagramme fondamental d'un détecteur de trafic, pourquoi est-il utile pour mesurer et prévoir la congestion?

**Il s'agit d'un diagramme qui permet de mesurer l'évolution du trafic en fonction de l'horaire. Il permet de comprendre le fonctionnement d'une section comme prévoir la congestion et observer l'état du trafic s'il est fluide ou non.**

c- quel est le débit typique maximal d'un tronçon à une voie

- en zone urbaine → **environ 1800 voitures par heure quand il n'y a pas de file d'attente**
- sur voie rapide ou autoroute → **environ 2700**
- Pourquoi cette différence? **parce que les voitures roulent plus rapidement**

#### 5: temps de parcours

a) Quelles sont les principales variables prédictives du temps de parcours d'un camion de livraison en ville, par ordre d'importance décroissante? (déterminées en cours)

**Comme variables prédictives du temps de parcours (par ordre décroissant d'importance) il y a la durée, l'heure (en seconde), le jour de la semaine, l'origine, la destination, la catégorie de la route, le nombre de voie sur la route, la durée statistique, la distance (en vol d'oiseau), la mission de ce trajet et la durée estimée.**

b) Citer 2 facteurs potentiels affectant les temps de parcours et difficiles à mesurer avec les données fournies dans les fichiers fournis en TP.

**Comme facteurs difficiles à mesurer et qui peuvent affecter le temps de parcours on peut citer l'expertise du chauffeur ou même son état physique actuel (fatigue, prise de médicaments...) mais aussi les accidents.**

## 6. Prescriptive Analytics

Les affirmations suivantes sont-elles vraies ou fausses:

a- Un problème de décision est dans la classe de complexité NP si et seulement si il n'existe pas d'algorithme polynomial pour le résoudre.

**FAUX**

b- Dans l'industrie, la majorité des problèmes d'ordonnancement sont résolus grâce à des heuristiques.

**VRAI**

c- Le problème suivant possède exactement trois solutions:

u in {1,3}

v in {1,2}

w in {3,4}

x in {1,5}

y in {4,5}

allDifferent(u,v,w,x,y)

**FAUX**

d- L'algorithme de résolution de CP-Optimizer est un algorithme exact: si un problème d'optimisation est faisable, il garantit de trouver une solution optimale.

**FAUX**

## 7. optimisation

a- En cherchant sur internet, décrivez un problème d'optimisation combinatoire non vu dans le cours dont la version de décision est un problème NP-Complet.

**Comme problème on peut citer le problème du sac à dos qui est un problème d'optimisation combinatoire. Il s'agit de remplir le sac à dos en maximisant l'espace et les objets sans dépasser un certain poids.**

b- Décrivez une petite instance particulière de ce problème d'optimisation (avec des valeurs pour chacune des données).

Objet	a	b	c
Valeur	2	3	5
Poids	5	6	2
<b>Poids Max = 10</b>			

c- Donnez une solution faisable non-optimale et une solution optimale de cette petite instance.

$$\text{Contrainte : } 5a+6b+2c \leq 10$$

a, b, c = 0 ou 1

solution faisable non-optimale : a = 1 et b = 1 OU a = 0 et b = 1

solution optimale : a = 0 et b = 0 et c = 1

## 8. programmation par contraintes

Deux principes fondamentaux de la Programmation par Contraintes sont (1) la recherche arborescente et (2) le filtrage du domaine des variables. Décrivez brièvement ces principes, leurs rôles et la façon dont ils sont mis en oeuvre durant la résolution.

**La recherche par arborescence est un algorithme utilisé dans la prise de décision qui cherche à trouver les valeurs impossibles en effectuant une sélection parmi un arbre des possibles avec les situations déjà explorées. Il s'agit d'une recherche de type heuristique.**  
**Le filtrage du domaine consiste à enlever les éléments dont on ne sait pas s'ils pourront prendre part à une solution gagnante.**