

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

Beziau Eva

1: IA

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

Approche utilitariste : s'inspirer du comportement du vivant pour l'appliquer à la recherche ou la conception (ex : voler). Cette approche ne permet pas de modéliser correctement ce comportement, mais permet de l'imiter à notre manière.

Approche réaliste : modéliser des processus biologiques réels du vivant, afin de les comprendre, mais aussi pour les contrôler. Par exemple, tout un pan de la recherche autour du Machine Learning a consisté à étudier le cerveau pour le reproduire numériquement le plus fidèlement possible. Néanmoins, ce n'est pas forcément la meilleure manière de procéder, notamment à cause des inconnus biologiques auxquels font face les scientifiques sur le cerveau.

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.

Après avoir exécuté le programme, les variables `result` et `input` seront :

`result = [0, 2, 4, 5, 9];`

`input = [];`

Le programme trie un tableau d'entiers par ordre croissant. `tmp` marque le plus petit nombre du tableau `input`, tandis que `i` marque la valeur de chaque autre nombre du tableau. On fait des comparaisons entre le nombre marqué par `tmp` et tous les autres. A la fin du tableau `input`, on ajoute à `result` le nombre marqué par `tmp` (le plus petit donc). On fait cela jusqu'à ce que le tableau `input` soit vide.

La complexité de ce programme est de $O(n \log(n))$ ([Algorithme de tri - Wikipédia](#)).

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.

[8 Best Route Planner Apps For Deliveries: 2021 Review](#)

Mapquest route planner : Gère bien les itinéraires avec de nombreux arrêts. Permet d'entrer rapidement de nombreuses adresses (copié-collé ou tableur). Option pour

réordonner les arrêts de la manière la plus optimale possible. Option pour un itinéraire en boucle (pour rejoindre le dépôt). Envoi de l'itinéraire possible par sms.

RouteXL : Permet de répartir les arrêts sur plusieurs itinéraires.

Flightmap : Permet de prendre en compte gratuitement jusqu'à 200 arrêts par mois pour 2 véhicules. Possibilité de télécharger des adresses via tableur. Option "nombre de tâches maximales par véhicule". Option sur la durée limite d'un itinéraire. Intégration directe avec Toukan.

4: trafic routier

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

Le débit et la vitesse / taux d'occupation

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

Le diagramme fondamental d'un détecteur de trafic permet de visualiser le trafic quotidien en fonction du débit, du taux d'occupation et du nombre de véhicules par heure, en fonction de la date.

Cela peut permettre de répartir le trafic si l'on voit qu'à un moment donné un endroit sera coincé par un embouteillage. Le diagramme permet de visualiser les tendances du trafic pendant la journée ou la semaine.

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

UVP = unité de véhicule particulier (voiture = 1, camion = 2, etc.)

Circulation routière

- en zone urbaine
?
- sur voie rapide ou autoroute
5 000 à 7 000 UVP/jour
- Pourquoi cette différence?
La débit dépend de la vitesse, mais aussi de l'attention des conducteurs et de l'état de la chaussée.

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)

La distance entre les points de livraison et l'expérience du chauffeur. Le trafic influe peu.

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

La météo et la fatigue du conducteur

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. VRAI

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.

/

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

/

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

/

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.

/