

Examen “Conduite de projets en science des données et Predictive Analytics”

GORBATOFF Alexane

1 : IA

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

L'approche réaliste défend le fait que le but de la science est de créer de la connaissance. Dans ce cas la méthodologie de l'IA, basée sur une sorte de biomimétisme, n'est pas forcément adaptée étant donné que ce n'est pas parce que quelque chose fonctionne dans le monde physique que cette même chose fonctionnera dans le monde informatique.

L'approche utilitariste, quant à elle, défend le fait que la science doit servir à quelque chose, peu importe les concepts qu'il y a derrière. Contrairement à l'approche réaliste, ici la méthodologie de l'IA est adaptée puisqu'elle va permettre de se détacher petit à petit du biomimétisme établie au départ afin de faire un objet répondant aux lois physiques plutôt qu'un objet calqué sur ce que l'on voit dans la nature.

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 :

- La variable `input` correspondra à un tableau vide
- La variable `result` correspondra à un tableau contenant les valeurs du tableau `input` initial rangées dans l'ordre croissant

L'algorithme permet de trier un tableau donné en entrée pour obtenir en sortie un tableau trié par ordre croissant. L'algorithme parcourt le tableau *input* pour chercher le minimum qu'il ajoute au tableau trié *result*. Ce minimum est ensuite retiré du tableau *input*. Le nouveau minimum du tableau *input* est alors recherché pour être ajouté à la suite dans le tableau trié. Et ainsi de suite jusqu'à ce que le tableau *input* soit vide.

La complexité de cet algorithme est en $O(n^2)$ puisqu'on recherche successivement le minimum sur un tableau de taille n.

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.

Copilot :

Copilot est une solution GPS et d'optimisation d'itinéraires proposant les services suivants :

- Instructions étape par étape fiables et précises
- Logiciel conforme à la législation spécifique camions
- Dimensions personnalisables (hauteur, longueur, largeur, poids total et poids par essieu) pour un calcul optimal d'itinéraires
- Calcul d'itinéraires prenant en compte le type de charge dont 30 catégories différentes de substances dangereuses
- Avertissements audio et visuels spécifiques camions
- Profils personnalisés de camions Interfaçage aux principaux systèmes de géolocalisation
- Ecran Sécurité Conducteur pour une concentration optimale
- Planification d'itinéraires à arrêts multiples
- Prise en compte des séquençage d'horaires Optimisation des itinéraires
- Connexion à des services Live en temps réel*- ActiveTraffic™
- Options visuelles et sonores personnalisables pour répondre à des exigences métiers strictes
- Calculs rapides et automatiques de changement d'itinéraire si le conducteur prend la mauvaise sortie ou ne suit pas l'itinéraire

<https://copilotpro.com/>

Mapotempo :

Mapotempo est une solution de gestion des tournées (livraisons, collectes, interventions, commerciaux) proposant les services suivants :

- Géocodage d'adresses : Géocoder en passant d'une adresse à ses coordonnées géographiques et inversement. Visualiser ces adresses sur une carte.
- Planification de tournées : Planifier des tournées de livraison, de collecte ou de ramasse en respectant les contraintes liées à l'activité.
- Calcul d'itinéraire : Calculer rapidement des itinéraires, au plus court, au plus rapide, avec différents modes de transport : véhicule léger, poids lourd, vélo etc.
- Optimisation de tournées : Optimiser des tournées en quelques clics et réaliser des gains immédiats en temps et en distance.
- Suivi d'exécution de tournées : Remontée temps réel de la position du véhicule (géolocalisation) et du statut des missions (planifié, en cours, fait, annulé etc.).
- Alertes/infos clients et personnel mobile : Envoi automatique de SMS et de l'heure estimée d'arrivée (avec trafic) à votre base clients et chat live avec votre personnel mobile.

Source : <https://www.mapotempo.com/>

TourSolver :

TourSolver est une solution permettant d'optimiser les tournées et proposant les fonctionnalités suivantes :

- Une interface personnalisable en fonction qui s'adapte au langage du métier exercé
- Une optimisation des tournées en prenant en compte la réalité du terrain (horaires de visite, de travail, de pause, capacité de la flotte, quantités à livrer, compétences, contraintes géographiques, et bien d'autres...)
- Optimisation de la sectorisation géographique des opérations afin de d'équilibrer la charge de travail en fonction des ressources disponibles
- Une application mobile destinée aux ressources présentent sur le terrain (agenda, tournées à réaliser, fiche de visite, état de réalisation des visites, signature, photos, compte-rendu personnalisé, géolocalisation...) et permettant de remonter les informations en temps réel au planificateur

<https://www.toursolver.com/>

4 : Trafic routier

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

Les principales variables mesurées par un détecteur de trafic sont :

- Le débit, c'est-à-dire, le nombre de véhicules qui passent sur une heure
- Le taux d'occupation, c'est-à-dire, le pourcentage de temps pendant lequel il y avait un véhicule sur la chaussée. Il s'agit là normalement d'un assez bon proxy pour la

vitesse. Par exemple, si le taux d'occupation est élevé et donc que les véhiculent sont très proches les uns des autres, normalement ils vont aller moins vite.

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 comporte, en abscisse le taux d'occupation (ou la vitesse) et en ordonnée le débit, c'est-à-dire le nombre de véhicules qui passent. Tous les diagrammes de ce type présentent globalement la même forme, à savoir une forme en cloche. On peut généralement identifier 3 tendances principales :

- Lorsque le taux d'occupation est faible, plus il y a de véhicules qui entrent, plus il y a de véhicules qui sortent et donc le débit augmente.
- Puis on arrive sur un plateau, car plus on ajoute de véhicules, plus ils vont être contraints d'aller lentement. On atteint alors le débit maximal admissible sur le détecteur.
- Enfin, on arrive à un stade où le nombre de véhicules est trop important, on ne peux plus en ajouter pour augmenter le débit. Cela correspond à l'embouteillage et à la congestion.

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

- **en zone urbaine**
- **sur voie rapide ou autoroute**
- **Pourquoi cette différence?**

Le débit typique maximal d'un tronçon à une voie en zone urbaine est de 1200 veh/h, tandis que celui en voie rapide ou autoroute est de 2200 veh/h.

La densité du bâti, les carrefours, l'étroitesse des voies, les rayons de girations ou encore la faible visibilité sont autant de paramètres qui justifient d'une vitesse d'approche plus faible en zone urbaine que sur voie rapide ou autoroute. De ce fait, le débit maximal admissible en zone urbaine est plus faible qu'en zone non-urbaine.

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)

Les principales variables prédictives du temps de parcours d'un camion de livraison en ville sont :

1. La distance à vol d'oiseau
2. L'identifiant de la tournée, qui comprend à la fois les caractéristiques du chauffeur (motivation, expérience, type de véhicule...) et le type de tournée

-
- 3. La durée statique
 - 4. Le taux d'occupation
 - 5. L'angle (coordonnées polaires)

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

L'expérience du chauffeur et sa motivation sont deux facteurs potentiels affectant le temps de parcours et difficiles à mesurer avec les données contenues dans les fichiers de TP.

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. La classe NP est une extension de la classe P, en autorisant des choix non-déterministes.

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

Vrai. Les bonnes heuristiques permettent souvent de formaliser les problèmes industriels.

c- Le problème suivant possède exactement trois solutions:

$u \in \{1,3\}$
 $v \in \{1,2\}$
 $w \in \{3,4\}$
 $x \in \{1,5\}$
 $y \in \{4,5\}$
 $\text{allDifferent}(u,v,w,x,y)$

Faux. Le problème suivant possède exactement 2 solutions.

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.

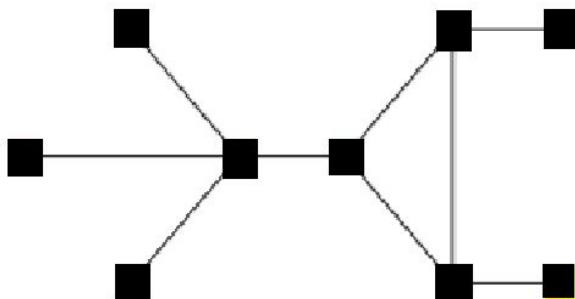
Le problème du graphe k-colorable :

Il s'agit d'un problème d'optimisation consistant à trouver le nombre k minimal permettant de colorier de l'ensemble des sommets du graphe de sorte à ce que deux nœuds reliés entre eux ne portent pas la même couleur.

Source : <https://pagesperso.g-scop.grenoble-inp.fr/~bousquen/Slides/semindoc.pdf>

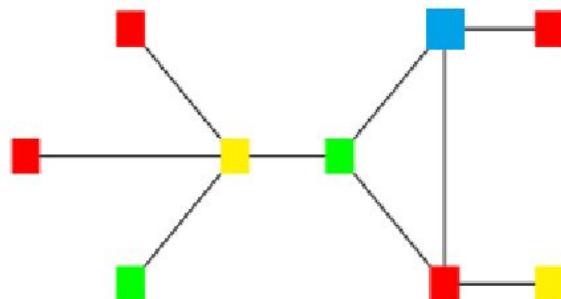
b- Décrivez une petite instance particulière de ce problème d'optimisation (avec des valeurs pour chacune des données).

On considère l'instance ci-dessous, à savoir un graphe à 9 nœuds.



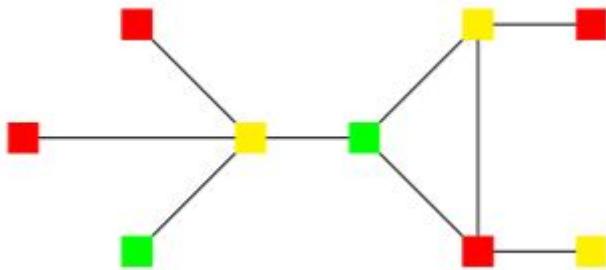
c- Donnez une solution faisable non-optimale et une solution optimale de cette petite instance.

Solution faisable non-optimale :



Tous les nœuds du graphe sont coloriés et les nœuds reliés entre eux ne sont pas de la même couleur. Néanmoins $k=4$ n'est pas optimal.

Solution optimale :



Ici, $k=3$ est la solution optimale.

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 œuvre durant la résolution.

La recherche arborescente :

On part d'une racine correspondante à une représentation de l'ensemble des variables de décision du problème, puis on va explorer un arbre en fixant une par une les variables de décision. Pour un même niveau de profondeur, chaque branche de l'arbre fixera la valeur d'une même variable à une des valeurs possibles pour cette variable.

Le filtrage du domaine des variables :

Pour un nœud donné on va essayer de filtrer le domaine des variables possibles pour les variables de décision et d'enlever toutes les variables pour lesquelles on sait qu'il y a au moins une contrainte qui fait que ces valeurs sont impossibles. Pour cela, on va analyser individuellement chaque contrainte et on va filtrer toutes les variables sur lesquelles portent la contrainte. A chaque nœud de l'arbre on va partir d'une racine du domaine des variables et on va faire un filtrage automatique, puis quand on ne pourra plus filtrer on va choisir un variable et une valeur pour cette variable et en la fixant on obtiendra en dessous le nœud fils. On recommence ensuite cette opération en partant du nœud fils et ainsi de suite.