

Exercices

L'analyse pour la prise de décision Instructeur : Patrick Boily

Description : L'analyse des données améliore l'utilisation de solutions fondées sur des preuves pour résoudre des problèmes de gouvernance difficiles. Les organisations se tournent de plus en plus vers les données pour faciliter les changements transformationnels.

Ce cours fournit aux participants une connaissance de base de l'analytique pour soutenir la prise de décision basée sur des preuves. Les participants amélioreront leur compréhension des processus utilisés dans la prise de décision, ainsi que des outils et techniques pour l'application de l'analytique à ces processus. Les discussions portent sur les méthodes, techniques et logiciels d'analyse les plus courants et les plus utiles. Ce cours fournira aux participants une base de connaissances qui leur permettra d'améliorer l'allocation des ressources de la manière la plus efficace possible.

C'est un cours de type survol. Des compétences en programmation et en statistique ne sont pas requises.

Instructeur/Facilitateur : Patrick Boily est diplômé de l'Université d'Ottawa. Il a obtenu son doctorat en mathématiques en 2006. Il a enseigné plus de 65 cours dans des universités de la région d'Ottawa depuis 1999. Il a travaillé sur de nombreux projets en tant que fonctionnaire fédéral de 2008 à 2012 (y compris l'étude primée sur l'utilisation des véhicules canadiens). Il a lancé et géré le Centre pour l'analyse quantitative et l'aide à la décision de l'Université Carleton de 2012 à 2019.

Il est maintenant professeur au département de mathématiques et de statistiques de l'Université d'Ottawa. Au fil des ans, il a fourni des services de consultation à de multiples entités, dont l'Administration canadienne de la sûreté du transport aérien et la Société de gestion des déchets nucléaires. Il possède une vaste expérience en recherche opérationnelle et en science des données. Il dirige divers ateliers sur la science des données et l'analyse statistique.

Le matériel de cours a été préparé en collaboration avec Jen Schellinck, Stephen Davies, John Stroud et Marwan Kashef.

Contenu

Module 1

Prise de décision

L'importance de la prise de décision ; stratégies de prise de décision ; biais de décision ; superprévisions et **crowdsourcing** ; intelligence augmentée et prise de décision ; conseils pour la prise de décision.

Module 2

Raisonnement, preuves, informations, données

Analyse ; stratégies empruntées des sciences ; types de raisonnement.

Module 3

Personnel, éthique des données et lois

Rôles et équipes d'analystes ; l'éthique dans le contexte des données ; les données et la loi.

Module 4

Intelligence économique et analytique

Historique et termes de l'intelligence économique ; description, explication, prédiction, prescription ; exploration de scénarios " que se passe-t-il si " ; valeur de la présentation et de la persuasion.

Module 5

Analyse et appui décisionnel

Décisions concernant l'analytique ; options, méthodes et outils ; attentes raisonnables ; pièges ; conseils pratiques ; communication des résultats par les visualisations et les tableaux de bord.

Module 6

Ingénierie des données et gouvernance des données

Aperçu du pipeline de données ; composants du pipeline de données ; rôles de la gouvernance des données ; pratiques de la gouvernance des données.

Exercices : Module 1 – Principes de base de l'analyse des données

1. Avec 26 secondes à jouer dans le Super Bowl, les Seattle Seahawks sont menés de 4 points par les New England Patriots. Sur un **2nd & Goal**, les Seahawks ont le ballon sur la ligne d'une verge des Patriots. La sagesse commune dans cette situation est de passer le ballon au **running back** et de le laisser essayer de percer la ligne défensive. Les Seahawks avaient deux options :
 - A. Faire courir le ballon (1 jeu). Risque : Ne parvient pas à marquer et le temps s'écoule.
 - B. Lancer la balle à la place, puis courir si nécessaire (2 jeux). Risque : 2% de chances d'interception.Quel jeu l'entraîneur devrait-il choisir ? Pourquoi ?
2. Dans l'article de Vanity Fair "**You Could Fit All the Voters Who Cost Clinton the Election in a Mid-Size Football Stadium**", Tina Nguyen écrit :

Alors que près de 138 millions d'Américains ont voté lors de l'élection présidentielle, l'étonnante victoire électorale de Donald Trump s'est jouée dans une poignée d'États que l'on s'attendait à voir gagner par Hillary Clinton. Le fait que Clinton ait remporté le vote populaire a été une maigre consolation pour les démocrates - au dernier décompte, elle avait environ 2,5 millions de voix d'avance, et ce chiffre ne cesse d'augmenter au fur et à mesure que les bulletins de vote sont comptés. Ce qui est encore plus désolant, c'est l'écart minuscule par lequel Clinton a perdu le Wisconsin, le Michigan et la Pennsylvanie - trois États qui étaient censés être son pare-feu dans la Rust Belt, mais qui ont finalement fait pencher la carte du collège électoral de manière décisive en faveur de Trump.

La marge de victoire de Trump dans ces trois États ? Seulement 79 316 voix. Ce dernier chiffre provient du décompte final des votes de la Pennsylvanie effectué par Decision Desk, où Trump a remporté 2 961 875 voix contre 2 915 440 pour Clinton, soit une différence de 46 435 voix. Ajoutez à cela les résultats officiels du Wisconsin, où Clinton a perdu par 22 177 voix, et du Michigan, où elle a perdu par 10 704 voix, et vous obtenez le résultat suivant : 0,057 % du total des électeurs a coûté la présidence à Clinton.

Il n'est pas tout à fait inhabituel que le collège électoral soit perdu par une marge aussi mince. En 2000, Al Gore a perdu la Floride (et donc l'élection) à 1 754 voix près, ce qui a déclenché un drame de recomptage pénible qui ne s'est terminé que par un jugement de la Cour suprême. Et en 2004, John Kerry a perdu contre George W. Bush en perdant l'Ohio par un peu plus de 118 000 voix. Mais il convient de se demander combien peu d'électeurs ont finalement mis le pays sur la voie actuelle, sans doute terrifiante. Les 79 316 personnes qui ont voté pour Trump dans le Wisconsin, le Michigan et la Pennsylvanie - tous des États que les démocrates ont conquis depuis 1992 - représentent moins que l'ensemble des étudiants de Penn State (97 494 étudiants), ou à peine plus que le nombre de personnes qui ont assisté à Desert Trip, le festival de musique des baby-boomers familièrement appelé "Oldchella". Si l'on mettait tous ces électeurs dans le Rose Bowl, il resterait un peu plus de 13 000 places.

Il y a plus de gens qui vivent à Nampa, Idaho, une ville dont vous n'avez jamais entendu parler. Pour replacer les choses dans une perspective encore plus douloureuse, la candidate du Parti vert Jill Stein a obtenu environ 130 000 voix dans ces trois États. Le candidat libertarien Gary Johnson en a obtenu environ 422 000.

Mais la donnée la plus douloureuse pour Clinton est peut-être la suivante : la candidate démocrate à la présidence a largement négligé la Pennsylvanie et le Michigan, alors que Trump a prospecté ces trois États sans relâche. Son blitz furieux de dernière minute à travers la Rust Belt pour gagner les électeurs blancs de la classe ouvrière, combiné au manque de ressources investies par Clinton, a essentiellement donné leurs 46 votes électoraux combinés à Trump. Au lieu de cela, Clinton a passé les dernières semaines de sa campagne à dépenser des ressources dans des endroits comme l'Arizona et le Texas - des États qui ont été remportés par Trump par une marge énorme.

Alors est-ce que c'était de la malchance, ou une erreur ? Pourquoi ?

3. Revoir les deux dernières questions à la lumière de la diapositive "La chance et l'information".
4. Faites l'exercice de traitement médical (voir la présentation de diapositives).
5. Au cours de la semaine prochaine, prêtez attention à votre processus de décision (ou au processus de décision de votre organisation). Quels sont les biais de décision auxquels vous êtes le plus sensible ? Le moins ?

Exercices : Module 2 – Raisonnement, preuves, informations, données

1. Les arguments suivants sont-ils solides ? S'ils sont faibles, quelles sont leurs faiblesses ? Comment les amélioreriez-vous ?
 - a. Les vaccins COVID entraînent une augmentation des hospitalisations car la moitié des personnes hospitalisées étaient vaccinées.
 - b. La mise en service du grand collisionneur de hadrons était une erreur, car soit il détruit la Terre, soit il ne le fait pas ; une chance sur deux est bien trop risquée.
 - c. Nous savons que la Terre n'est pas plate car aucune des autres planètes connues n'est plate.
 - d. Vous ne devriez pas voter aux prochaines élections car une voix ne fait jamais la différence.
 - e. La solution pour réduire la congestion est de réduire le nombre de voies, car avec moins de voies, les gens chercheront d'autres modes de transport.
 - f. Les mesures de sécurité dans les aéroports sont proportionnelles au risque, car il est acceptable d'attendre quelques heures si cela signifie que mon avion ne sera pas détourné.
2. Considérez les éléments trouvés dans une note d'information relative à la construction d'un pipeline à travers le territoire des caribous :
 - a. Les 7 dernières fois que des pipelines ont été construits dans des territoires de caribous, les populations ont diminué dans le territoire.
 - b. Les biologistes ont une carte montrant les voies de migration du caribou. Sur la base de cette carte, nous concluons que la pose du pipeline sur le territoire n'entravera pas la migration des caribous.
 - c. Les pipelines n'ont pas affecté les populations d'oies ; comme elles et les caribous sont tous deux des animaux sociaux, le pipeline n'affectera pas la population de caribous.
 - d. Les biologistes ont montré que les caribous n'ont pas peur des gros objets. Si les caribous n'ont pas peur, leurs habitudes de reproduction ne seront pas affectées. Comme les pipelines sont de grands objets, la construction de ce pipeline n'affectera pas les habitudes de reproduction des caribous.
3. Identifiez les stratégies de raisonnement (généraliser à partir d'exemples, faire une déduction, raisonner par analogie, raisonner à la meilleure explication) utilisées dans chacun de ces arguments. En appliquant une lentille de raisonnement plausible, quelle serait votre conclusion ? De quelles informations supplémentaires auriez-vous besoin ou souhaiteriez-vous disposer avant de tirer une conclusion ?
4. Dans la pièce **Rosencrantz et Guildenstern sont morts** (1966) de Tom Stoppard, Rosencrantz tire à pile ou face et obtient pile 92 fois de suite. Est-ce possible ? Est-ce plausible ? Si cela vous arrivait, quelle serait votre conclusion ?

Exercices : Module 3 – Personnel, éthique des données et lois

1. Effectuez l'exercice généraliste/spécialiste pour vous-même ou votre équipe.
2. Une banque a l'obligation d'accroître la valeur pour ses actionnaires. Elle envisage un nouveau processus décisionnel piloté par l'IA pour les demandes de prêt, les objectifs étant de libérer le personnel de tâches fastidieuses et de réduire le taux de défaut. La banque a accès à des informations exclusives et publiques sur les demandeurs de prêts. Sans savoir quelle position vous pourriez occuper à l'avenir, quelles questions (le cas échéant) vous posent cette proposition ? Quels principes éthiques voudriez-vous que la banque garde à l'esprit ?

3. Discutez de l'étude de cas d'embauche d'Amazon présentée en classe.
4. Meltwater propose un logiciel qui récupère des informations sur les actualités à partir de mots-clés spécifiques. Les clients commandent des résumés sur des sujets contenant des extraits d'articles d'actualité. L'Associated Press (AP) affirme que son contenu a été volé et que Meltwater avait besoin d'une licence avant de distribuer les informations extraites. Le juge a donné raison à l'AP en faisant valoir que Meltwater est un concurrent. Quel est votre verdict ?
5. En 2014, six géologues italiens ont été blanchis des accusations d'homicide involontaire dans la mort de 29 personnes, qui ont péri lors d'un important tremblement de terre en 2009. Ils avaient initialement été condamnés pour homicide involontaire pour n'avoir pas su prévoir le tremblement de terre qui a ravagé la ville italienne de L'Aquila. Mais après le succès de leur appel, certaines questions demeurent, notamment celle de savoir si les scientifiques seront disposés à continuer à donner des conseils aux profanes, maintenant qu'ils savent qu'ils peuvent être menacés de poursuites pour l'avoir fait. Que vous suggère une telle histoire sur l'utilisation des données/preuves pour prendre des décisions ?

Module 4 – Intelligence économique

1. Effectuez un exercice de prémortem/rétrospective pour une nouvelle initiative en matière de santé mentale des jeunes. Supposons que votre ministère ait créé une application qui vise à améliorer la santé mentale des adolescents canadiens, après la pandémie. Supposons que nous sommes maintenant deux ans dans le futur et que vous faites un retour sur le lancement de l'application.
 - a. Prémortem : Donnez 3 raisons sous votre contrôle et 3 raisons hors de votre contrôle pour lesquelles le lancement a échoué.
 - b. Backcasting : Même chose, mais pour expliquer pourquoi le lancement a réussi.
2. Construisez un récit de base (**story spine**) relatif à une décision réussie qui a été prise par votre organisation (comme l'allocation d'un budget, l'introduction d'un nouveau programme, etc.).

Module 5 – L'analyse pour l'aide à la décision

1. Comment pourriez-vous utiliser des méthodes analytiques pour améliorer la qualité des recommandations dans l'exercice de la note d'information du module 2 ?
2. Quels sont les exemples d'utilisation de l'analyse dans votre département ?
3. Considérez la Marche vers Moscou de Minard. De quelle(s) manière(s) s'agit-il d'une bonne visualisation ? Une mauvaise visualisation ? Quelle histoire raconte-t-elle ? Quels concepts sont représentés dans ce graphique ? Qu'en est-il des autres graphiques de la présentation ?
4. Quelles histoires les affiches de visualisation de données soutiennent-elles ? Quel est le niveau de contexte nécessaire pour leur donner un sens ?

Module 6 – Ingénierie des données et gouvernance des données

1. Rédigez l'équipe de données idéale (ressources finies, cases ETP, coût, priorité).
2. Selon vous, quelles devraient être les 5 principales priorités en matière de gouvernance des données pour votre organisation ? Pour le gouvernement du Canada ?
3. Avez-vous eu des problèmes liés à la disponibilité, la facilité d'utilisation, la cohérence, l'intégrité, la qualité, la sécurité et/ou la fiabilité des données ? Discutez avec la classe.
4. Votre groupe de travail crée-t-il ou génère-t-il des données ? Quel type de données ? Utilisez-vous des données provenant de sources extérieures à votre groupe ? Combien de sources ? Lesquelles ? Publiez-vous l'analyse des données à l'intérieur de votre groupe, à l'extérieur ou les deux ?