UQÀM Université du Québec à Montréal
Faculté des sciences
Département de mathématiques

Hiver 2024
Groupe : 030

MAT998P
ÉQUITÉ ET DISCRIMINATION DES MODÈLES PRÉDICTIFS
(3 crédits)

Professeur : Arthur Charpentier
Courriel : charpentier.arthur@uqam.ca
Bureau : PK-5330
Disponibilités : me contacter par courriel [1].
Horaire du cours : Mercredi 09h00 - 12h00 (local PK-2605)

› Évaluation

— participation : 50%
— exposé sur un sujet lié au cours : 50%

› Contenu du cours

Ce cours présentera un état de l'art sur l'équité et la discrimination dans le contexte des modèles prédictifs, en mettant l'accent sur le cas de l'assurance. En effet, comme l'expliquent Avraham *et al.* (2014) "*insurance companies are in the business of discrimination*". Si la discrimination basée sur le risque semble légitime, il existe des attributs sensibles sur la base desquels un assureur ne doit pas discriminer. Ce cours reviendra sur la base des modèles prédictifs, et sur les principales variables sensibles. Dans un deuxième temps, nous verrons comment quantifier une éventuelle discrimination, en insistant sur les principales mesures de "group-fairness", avant de discuter de l'approche individuelle, notamment en relation avec les approches causales. En effet, la question centrale de la discrimination est "*le prix aurait-il été différent si cette personne avait été un homme au lieu d'une femme*". Enfin, nous verrons comment corriger une potentielle discrimination.
Ce cours présentera le contexte général, et les outils mathématiques nécessaires pour quantifier les discriminations de modèles prédictifs (modèles linéaires mais aussi des modèles plus opaques d'apprentissage machine) et les développements récents. D'autres sujets pourront être inclus selon temps et l'interet des étudiants. Ce cours sera largement basé sur Charpentier (2024), dont une version sera mise à disposition des étudiant·e·s. Des connaissances de base en modèles linéaires et en apprentissage machine sont un prérequis pour ce cours. Une connaissance de l'optimisation convexe est souhaitable, mais pas essentielle.

1. Recevant beaucoup de courriels, je répondrai aussi attentivement et précisément que possible à vos questions, à condition que le signe du cours [MAT998P] soit mentionné dans l'objet.

› Références

Ronen Avraham. Discrimination and insurance. In Kasper Lippert-Rasmussen, editor, *Handbook of the Ethics of Discrimination*, pages 335–347. Routledge, 2017.

Ronen Avraham, Kyle D Logue, and Daniel Schwarcz. Towards a universal framework for insurance anti-discrimination laws. *Connecticut Insurance Law Journal*, 21 :1, 2014.

Solon Barocas, Moritz Hardt, and Arvind Narayanan. Fairness in machine learning. *Nips tutorial*, 1 :2017, 2017.

Toon Calders and Sicco Verwer. Three naive Bayes approaches for discrimination-free classification. *Data mining and knowledge discovery*, 21(2) :277–292, 2010.

Arthur Charpentier. *Computational actuarial science with R.* CRC press, 2014.

Arthur Charpentier. *Insurance : biases, discrimination and fairness.* Springer Verlag, 2024.

Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. Algorithmic decision making and the cost of fairness. *arXiv*, 1701.08230, 2017.

Sam Corbett-Davies, Johann Gaebler, Hamed Nilforoshan, Ravi Shroff, and Sharad Goel. The measure and mismeasure of fairness. *arXiv*, 1808.00023, 2023.

Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*, pages 214–226, 2012.

Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al. *The elements of statistical learning*, volume 1. Springer series in statistics New York, 2001.

Niki Kilbertus, Mateo Rojas-Carulla, Giambattista Parascandolo, Moritz Hardt, Dominik Janzing, and Bernhard Schölkopf. Avoiding discrimination through causal reasoning. *arXiv*, 1706.02744, 2017.

Drago Plečko, Nicolas Bennett, and Nicolai Meinshausen. fairadapt : Causal reasoning for fair data pre-processing. *arXiv*, 2110.10200, 2021.

Filippo Santambrogio. *Optimal transport for applied mathematicians*, volume 55. Springer, 2015.

ENGAGEMENT ET RESPONSABILITÉS

Par leur présence en classe au moment convenu par le professeur, les étudiants deviennent responsables de leur formation en assumant pleinement les tâches exigées dans leur cheminement académique. De plus, par des comportements éthiques et une attitude professionnelle, ils assurent le maintien d'un environnement de travail et d'étude sain et riche, et ce, dans le respect des autres étudiants du groupe, des responsables de cours et de l'ensemble de la communauté universitaire.

Veuillez consulter, la Charte des droits et responsabilités des étudiantes et des étudiants (instances.uqam.ca/).

POLITIQUE N° 44 D'ACCUEIL ET DE SOUTIEN DES ÉTUDIANTES, ÉTUDIANTS EN SITUATION DE HANDICAP

Par sa politique, l'Université reconnait, en toute égalité des chances, sans discrimination ni privilège, aux étudiantes, étudiants en situation de handicap, le droit de bénéficier de l'ensemble des ressources du campus et de la communauté universitaire, afin d'assurer la réussite de leurs projets d'études, et ce, dans les meilleures conditions possibles. L'exercice de ce droit est, par ailleurs, tributaire du cadre réglementaire régissant l'ensemble des activités de l'Université.

Il incombe aux étudiantes, étudiants en situation de handicap de rencontrer les intervenantes, intervenants (conseillères, conseillers à l'accueil et à l'intégration du Service d'accueil et de soutien des étudiantes, étudiants en situation de handicap, professeures, professeurs, chargées de cours, chargés de cours, direction de programmes, associations étudiantes concernées, etc.) qui pourront faciliter leur intégration à la communauté universitaire ou les assister et les soutenir dans la résolution de problèmes particuliers en lien avec les limitations entraînées par leur déficience.

Le Service d'accueil et de soutien aux étudiantes, étudiants en situation de handicap (SASESH) offre des mesures d'aménagement dont peuvent bénéficier certains étudiants. Nous vous recommandons fortement de vous prévaloir des services auxquels vous pourriez avoir droit afin de réussir vos études, sans discrimination. Pour plus d'information, visitez le site de ce service à l'adresse suivante : http ://vie-etudiante.uqam.ca/etudiant-situation-handicap/nouvelles-ressources.html et celui de la politique institutionnelle d'accueil et de soutien aux étudiantes, étudiants en situation de handicap : https://instances.uqam.ca/wp-content/uploads/sites/47/2018/05/Politique_no_44.pdf

Vous devez faire connaitre votre situation au SASESH le plus tôt possible :
En personne : 1290, rue Saint-Denis, Pavillon Saint-Denis, local AB-2300
Par téléphone : 514 987-3148
Courriel : situation.handicap@uqam.ca
En ligne : http ://vie-etudiante.uqam.ca/

POLITIQUE NO 16 VISANT À PRÉVENIR ET À COMBATTRE LE SEXISME ET LES VIOLENCES À CARACTÈRE SEXUEL

Le harcèlement sexuel se définit comme étant un comportement à connotation sexuelle unilatéral et non désiré ayant pour effet de compromettre le droit à des conditions de travail et d'études justes et raisonnables ou le droit à la dignité.

Les violences à caractère sexuel se définissent comme étant des comportements, propos et attitudes à caractère sexuel non consentis ou non désirés, avec ou sans contact physique, incluant ceux exercés ou exprimés par un moyen technologique, tels les médias sociaux ou autres médias numériques. Les violences à caractère sexuel peuvent se manifester par un geste unique ou s'inscrire dans un continuum de manifestations et peuvent comprendre la manipulation, l'intimidation, le chantage, la menace implicite ou explicite, la contrainte ou l'usage de force. Les violences à caractère sexuel incluent, notamment :

- la production ou la diffusion d'images ou de vidéos sexuelles explicites et dégradantes, sans motif pédagogique, de recherche, de création ou d'autres fins publiques légitimes ;
- les avances verbales ou propositions insistantes à caractère sexuel non désirées ;
- la manifestation abusive et non désirée d'intérêt amoureux ou sexuel ;
- les commentaires, les allusions, les plaisanteries, les interpellations ou les insultes à caractère sexuel, devant ou en l'absence de la personne visée ;
- les actes de voyeurisme ou d'exhibitionnisme ;
- le (cyber) harcèlement sexuel ;
- la production, la possession ou la diffusion d'images ou de vidéos sexuelles d'une personne sans son consentement ;
- les avances non verbales, telles que les avances physiques, les attouchements, les frôlements, les pincements, les baisers non désirés ;
- l'agression sexuelle ou la menace d'agression sexuelle ;
- l'imposition d'une intimité sexuelle non voulue ;
- les promesses de récompense ou les menaces de représailles, implicites ou explicites, liées à la satisfaction ou à la non-satisfaction d'une demande à caractère sexuel.

Toute personne membre de la communauté universitaire contribue à maintenir une culture du respect et du consentement, notamment, en participant aux activités de formations obligatoires sur le sexisme et les violences à caractère sexuel.

Pour plus d'information : :
https://instances.uqam.ca/wp-content/uploads/sites/47/2019/04/Politique_no_16_2.pdf

Les personnes victimes, témoins ou informées d'une situation de sexisme, de violence à caractère sexuel, ou pour en apprendre plus sur ces enjeux, peuvent consulter le
Bureau d'intervention et de prévention en matière de harcèlement (BIPH)
514 987-3000, poste 0886 ; harcelement@uqam.ca ; harcelement.uqam.ca

Soutien psychologique (Services à la vie étudiante)
514 987-3185 ; vie-etudiante.uqam.ca

Service de la prévention et de la sécurité :
514 987-3131

LES INFRACTIONS DE NATURE ACADÉMIQUE



Tout acte de plagiat, fraude, copiage, tricherie ou falsification de document commis par une étudiante, un étudiant, de même que toute participation à ces actes ou tentative de les commettre, à l'occasion d'un examen ou d'un travail faisant l'objet d'une évaluation ou dans toute autre circonstance, constituent une infraction au sens de ce règlement.

La liste non limitative des infractions est définie comme suit :

- la substitution de personnes ;
- l'utilisation totale ou partielle du texte d'autrui en le faisant passer pour sien ou sans indication de référence ;
- la transmission d'un travail aux fins d'évaluation alors qu'il constitue essentiellement un travail qui a déjà été transmis aux fins d'évaluation académique à l'Université ou dans une autre institution d'enseignement, sauf avec l'accord préalable de l'enseignante, l'enseignant ;
- l'obtention par vol, manœuvre ou corruption de questions ou de réponses d'examen ou de tout autre document ou matériel non autorisés, ou encore d'une évaluation non méritée ;
- la possession ou l'utilisation, avant ou pendant un examen, de tout document non autorisé ;
- l'utilisation pendant un examen de la copie d'examen d'une autre personne ;
- l'obtention de toute aide non autorisée, qu'elle soit collective ou individuelle ;
- la falsification d'un document, notamment d'un document transmis par l'Université ou d'un document de l'Université transmis ou non à une tierce personne, quelles que soient les circonstances ;
- la falsification de données de recherche dans un travail, notamment une thèse, un mémoire, un mémoire-création, un rapport de stage ou un rapport de recherche.

Les sanctions reliées à ces infractions sont précisées à l'article 3 du Règlement no 18 sur les infractions de nature académique.

Vous pouvez consulter sur le site r18.uqam.ca des capsules vidéos qui vous en apprendront davantage sur l'intégrité académique et le R18, tout en vous orientant vers les ressources mises à votre disposition par l'UQAM pour vous aider à éliminer le plagiat de vos travaux.

INFOSPHÈRE

Infosphère est l'un de ces outils indispensables : un guide méthodologique visant à promouvoir les bonnes pratiques en matière de recherche documentaire et de rédaction de travaux. Cet outil vous accompagnera tout au long de vos études et vous permettra d'éviter les pièges du plagiat. Cliquez sur le logo à gauche pour être redirigé vers le site.