Insurance: bias, discrimination & fairness

Arthur Charpentier[1]

[1] Univeristé du Québec à Montréal & AXA JRI

IVADO-OBVIA, June 2022

Brief overview

Arthur Charpentier

Université du Québec à Montréal

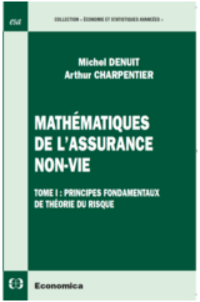
@freakonometrics

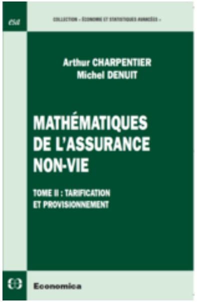
freakonometrics.github.io

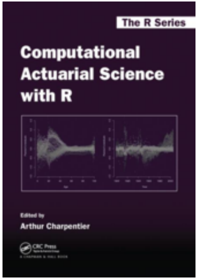
freakonometrics.hypotheses.org

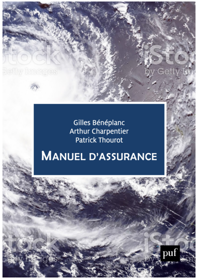
Actuarial Science and Economics of Insurance
Algorithms and Statistical Learning
Dynamic Games and Market Equilibrium

Insurance: Discrimination, Biases and Fairness *ILB*
Assurance: Discrimination, Biais et Équité *ILB*







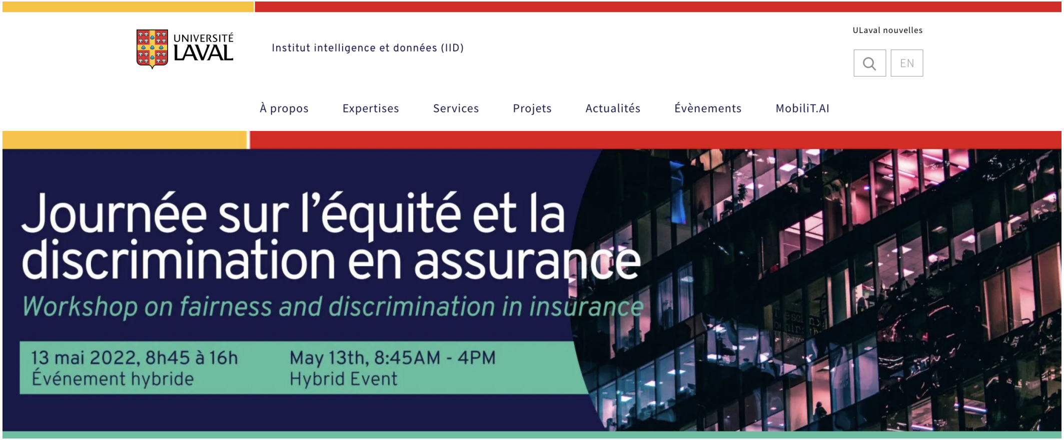


Brief overview

- ILB (Institut Louis Bachelier) publications
- AXA Joint Research Initiative on Unusual Data for Insurance
  see Grari et al. (2022)
- COVEA (Chair 2015-2018 & Fair IA 2023)
- ACPR (Autorité de Contrôle et de Régulation)

- Students MSc & PhD working on fairness, causal inference & interpretability
- Academics invitations for short-term visits 2022-2023
- KPMG joint projects

Brief overview



Workshop on fairness and discrimination in insurance, May 13, Québec City
https://iid.ulaval.ca/jeda2022

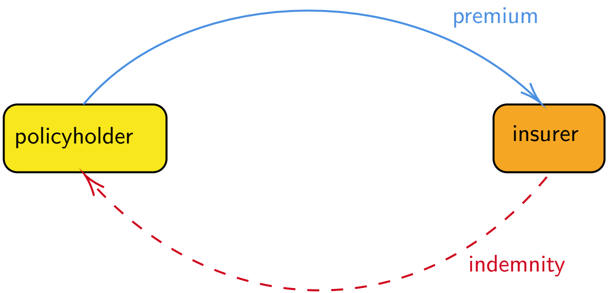
Agenda & keywords

"*Technology is neither good nor bad; nor is it neutral*" , Kranzberg (1986)

- Insurance, mutualization, solidarity vs. individualization, heterogeneity
- Discrimination, *actuarial fairness*, legal aspects, discrimination by proxy
- Biases observation vs. experiment, selection bias, omited variable bias
- Fairness, $\widehat{Y} \perp\!\!\!\perp P$, $\widehat{Y} \perp\!\!\!\perp P \mid Y$ or $Y \perp\!\!\!\perp P \mid \widehat{Y}$, and individual fairness (counterfactual)
- Explainability and interpretability

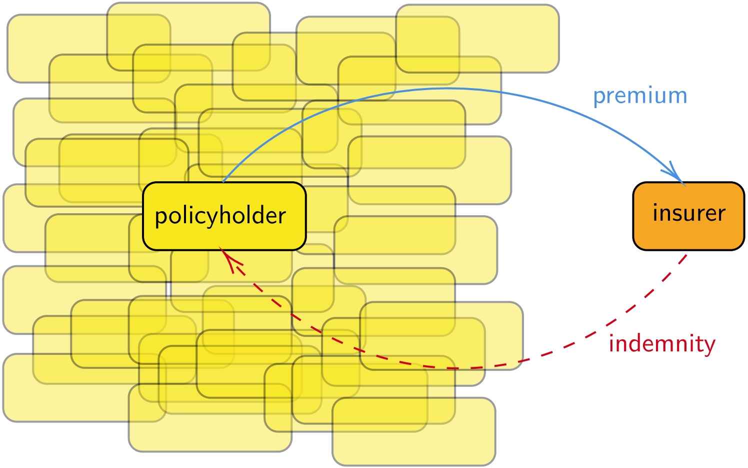
see Charpentier (2022), Barry and Charpentier (2022) and Grari et al. (2022) for further details

Insurance, risk pooling & solidarity



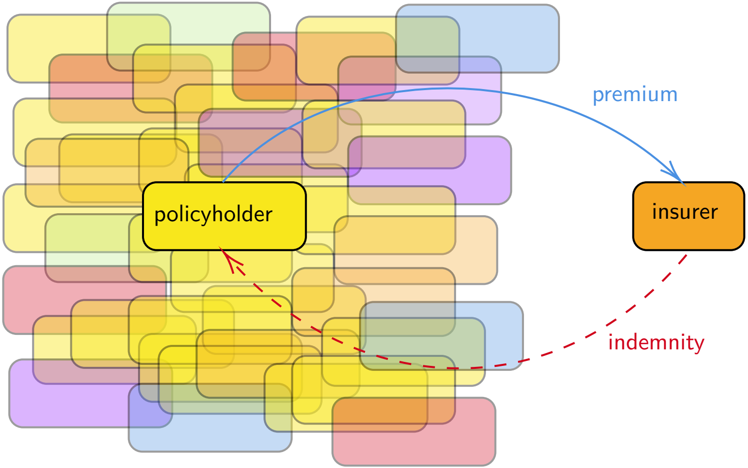
Insurance, risk pooling & solidarity

- Insurance is the contribution of the many to the misfortune of the few



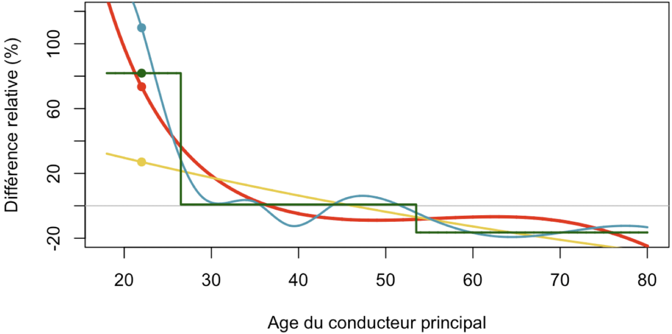
Insurance, risk pooling & solidarity

- Insurance is the contribution of the many to the misfortune of the few



Heterogenous Risks I

- Claim frequency, as a function of the age of the driver (data Charpentier (2014))



- "*actuaries smoothed because smoothing was a 'mathematical and ethical' good*", Bouk (2015)
- interpretability and explanability, "*young drivers are more likely to have an accident*"

Heterogenous Risks II

- Life insurance, life tables function of the age and the genrder
- Men / women table, 1720 (Struyck (1912), page 231)

men						women					
0	1000	29.0%	45	371	16.6%	0	1000	28.9%	45	423	11.8%
5	710	5.6%	50	313	19.2%	5	711	5.2%	50	373	14.7%
10	670	4.2%	55	253	22.9%	10	674	3.3%	55	318	18.2%
15	642	5.5%	60	195	27.2%	15	652	4.3%	60	260	21.2%
20	607	6.6%	65	142	31.7%	20	624	5.8%	65	205	26.8%
25	567	7.9%	70	97	37.1%	25	588	6.8%	70	150	33.3%
30	522	9.2%	75	61	45.9%	30	548	7.3%	75	100	45.0%
35	474	10.5%	80	33	51.5%	35	508	7.9%	80	55	56.4%
40	424	12.5%	85	16		40	468	9.6%	85	24	

Insurance and premium "individualization" I

- "*It is important to distinguish two things when talking about insurance. The first, the insurance operation, is technical and has a collective dimension, the second, the insurance contract, is legal and has an individual dimension*", Bigot and Cayol (2020) (aussi Thiery and Van Schoubroeck (2006), Lehtonen and Liukko (2015))
- **Individualistic approach**
  - The individualistic approach to equality analyses fundamental rights, such as the right to equal treatment, in terms of individuals.
  - An individual cannot be treated differently because of his or her membership in such a group, particularly in a group to which he or she has not chosen to belong.
- **Group approach**
  - The insurance tradition, on the other hand, analyses risks, premiums and benefit schedules in terms of groups
  - Unlike the individualistic approach, insurance classification schemes rely on the assumption that individuals answer to the average (stereotypical) characteristics of a group to which they belong.

Legal Perspective I

	CA	HI	GA	NC	NY	MA	PA	FL	TX	AL	ON	NB	NL	QC
Gender	✗	✗	●	✗	●	✗	✗	●	●	●	●	✗	✗	●
Age	✗	✗	●	✗*	●	✗	●	●	●	●*	●	✗	✗	●
Driving experience	●	✗	●	●	●	●	●	●	●	●	●	●	●	●
Credit history	✗	✗	●	●	●	✗	●*	●	●	✗*	✗	●*	✗	●
Education	✗	✗	✗	✗	✗	✗	●	●	●	●	●	●	●	●
Occupation	✗	✗	✗	●	✗	✗	●	●	●	●	●	●	●	●
Employment status	✗	✗	✗	●	✗	✗	●	●	●	●	●	●	●	●
Marital status	●	✗	●	●	●	✗	●	●	●	●	●	●	●	●
Housing situation	✗	✗	●	●	●	✗	●	●	●	✗	✗	●	●	●
Address/ZIP code	●	●	●	●	●	●	●	●	●	✗	✗	●	●	●
Insurance history	●	●	●	●	●	●	●	●	●	●	●	●	●	●

CA: Californie, HI: Hawaii, GA: Georgia, NC: Caroline du nord, NY: New York, MA: Massachusetts, PA: Pennsylvanie, FL: Floride, TX: Texas

Bureau d'Assurance du Canada (2021)

Legal Perspective II

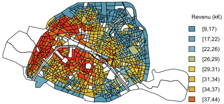
In Québec province

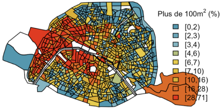
- "*Toute personne a droit à la reconnaissance et à l'exercice, en pleine égalité, des droits et libertés de la personne, sans distinction, exclusion ou préférence fondée sur la race, la couleur, le sexe, l'identité ou l'expression de genre, la grossesse, l'orientation sexuelle, l'état civil, l'âge sauf dans la mesure prévue par la loi, la religion, les convictions politiques, la langue, l'origine ethnique ou nationale, la condition sociale, le handicap ou l'utilisation d'un moyen pour pallier ce handicap.*" (C-12 - Charte des droits et libertés de la personne, art. 10)
- "*la distinction fondée sur l'âge, le sexe ou l'état civil est permise lorsqu'elle repose sur un facteur qui permet de déterminer un risque. Par exemple, une compagnie d'assurance peut vous poser des questions sur votre âge et votre sexe pour fixer votre prime*" (art. 20.1)

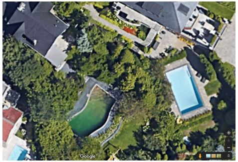
Proxy Based Discrimination (?) I

- location (policyholder home adress)

  Jean et al. (2016), Seresinhe et al. (2017), Gebru et al. (2017), Law et al. (2019), Ilic et al. (2019), Kita and Kidziński (2019), see also redlining



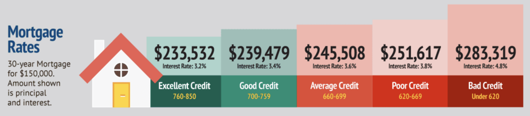


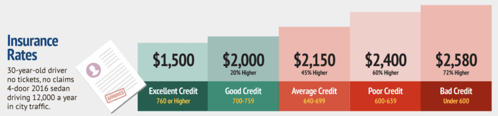


Proxy Based Discrimination (?) II

- credit scoring,

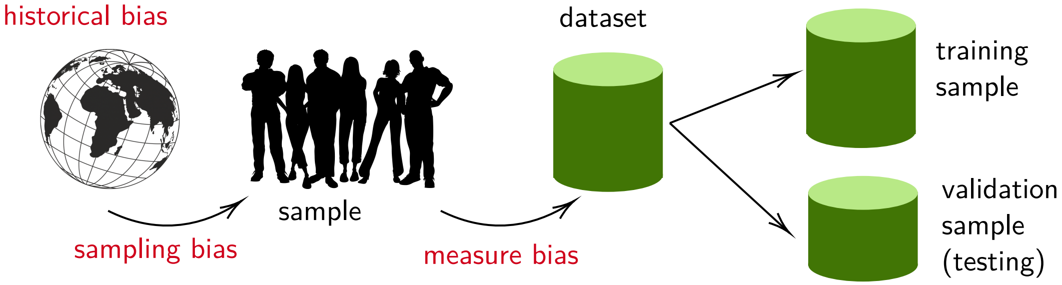
  Kabler (2004), Arya et al. (2013), Miller et al. (2003) Bartik and Nelson (2016), O'Neil (2016), Lauer (2017), Morris et al. (2017), Kiviat (2019)





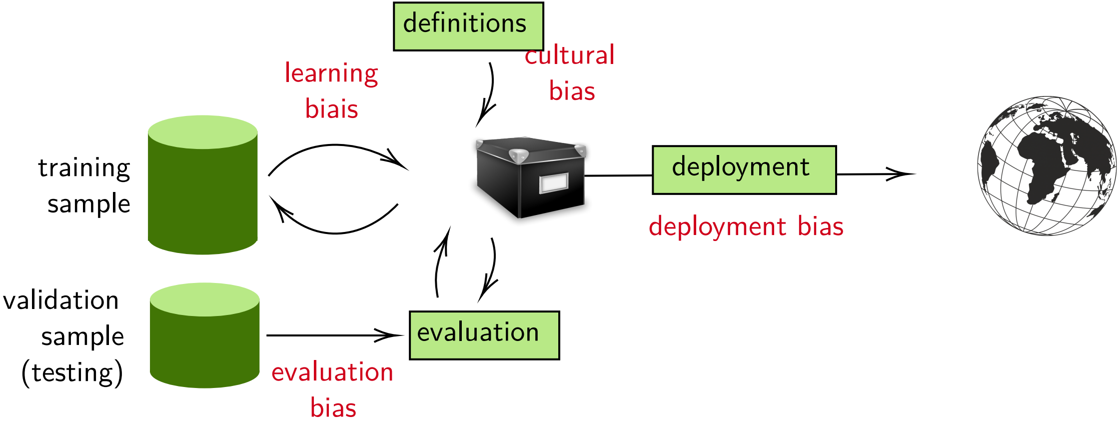
source https://www.incharge.org/debt-relief/credit-counseling/

Biases in Data Generation



(inspired by Suresh and Guttag (2019)).

Biases in Modeling



(inspired by Suresh and Guttag (2019)).

Measuring and quantifying equity I

Notations:

$$\begin{cases} y \in \{0,1\} & \text{variable of interest} \\ p \in \{0,1\} & \text{protected variable (sensitive)} \\ \boldsymbol{x} \in \mathbb{R}^d & \text{`explanatory' variables} \\ s \in [0,1] & \text{score, classically } s = s(\boldsymbol{x}, p) \\ \widehat{y} \in \{0,1\} & \text{classifier, classically } \widehat{y} = \mathbf{1}(s > t) \end{cases}$$

Fairness Through Unawareness, Kusner et al. (2017)
Protected attribute p is not explicitly used in decision function $\widehat{y}$.

Measuring and quantifying equity II

statistical parity	Dwork et al. (2012)	$\mathbb{P}[\widehat{Y}=1\|P=p]=\text{cst},\ \forall p$	independence
conditional statistical parity	Corbett-Davies et al. (2017)	$\mathbb{P}[\widehat{Y}=1\|P=p,X=x]=\text{cst}_x,\ \forall p,y$	$\widehat{Y}\perp\!\!\!\perp P$
equalized odds	Hardt et al. (2016)	$\mathbb{P}[\widehat{Y}=1\|P=p,Y=y]=\text{cst}_y,\ \forall p,y$	separation
equalized opportunity	Hardt et al. (2016)	$\mathbb{P}[\widehat{Y}=1\|P=p,Y=1]=\text{cst},\ \forall p$	
predictive equality	Corbett-Davies et al. (2017)	$\mathbb{P}[\widehat{Y}=1\|P=p,Y=0]=\text{cst},\ \forall p$	$\widehat{Y}\perp\!\!\!\perp P\mid Y$
balance (positive)	Kleinberg et al. (2017)	$\mathbb{E}[S\|P=p,Y=1]=\text{cst},\ \forall p$	$S\perp\!\!\!\perp P\mid Y$
balance (negative)	Kleinberg et al. (2017)	$\mathbb{E}[S\|P=p,Y=0]=\text{cst},\ \forall p$	
conditional accuracy equality	Berk et al. (2017)	$\mathbb{P}[Y=y\|P=p,\widehat{Y}=y]=\text{cst}_y,\ \forall p,y$	sufficiency
predictive parity	Chouldechova (2017)	$\mathbb{P}[Y=1\|P=p,\widehat{Y}=1]=\text{cst},\ \forall p$	
calibration	Chouldechova (2017)	$\mathbb{P}[Y=1\|P=p,S=s]=\text{cst}_s,\ \forall p,s$	$Y\perp\!\!\!\perp P\mid\widehat{Y}$
well-calibration	Chouldechova (2017)	$\mathbb{P}[Y=1\|P=p,S=s]=s,\ \forall p,s$	
accuracy equality	Berk et al. (2017)	$\mathbb{P}[\widehat{Y}=Y\|P=p]=\text{cst},\ \forall p$	
treatment equality	Berk et al. (2017)	$\dfrac{\text{FN}_p}{\text{FP}_p}=\text{cst}_p,\ \forall p$	

Measuring and quantifying equity III

Lipschitz property, Duivesteijn and Feelders (2008)

$$D(\widehat{y}_i, \widehat{y}_j) \text{ ou } D(s_i, s_j) \leq d(\boldsymbol{x}_i, \boldsymbol{x}_j),\ \forall i, j = 1, \cdots, n.$$

Cf formal intervention "$\boldsymbol{X}$ is fixed at $\boldsymbol{x}$", see "$do(\boldsymbol{X} = \boldsymbol{x})$" in Pearl (1998) (or simply $do(\boldsymbol{x})$), (historically, from Wright (1921), Neyman et al. (1923) or Rubin (1974) Holland (1986))

Conterfactual fairness, Kusner et al. (2017) If the prediction in the real world is the same as the prediction in the counterfactual world where the individual would have belonged to a different demographic group, we have counterfactual equity, i.e.

$$\mathbb{P}\big[Y^\star_{P\leftarrow p} = y | \boldsymbol{X} = \boldsymbol{x}, P = p\big] = \mathbb{P}\big[Y^\star_{P\leftarrow p'} = y | \boldsymbol{X} = \boldsymbol{x}, P = p\big],\ \forall p', \boldsymbol{x}, y.$$

To go further on quantifying fairness

LE JOURNAL DE MONTRÉAL FÉVRIER 2022 21H30 -13°c SE CON

LE GUIDE DE L'AUTO

Un homme change de sexe pour faire baisser la facture de son assurance auto

FRÉDÉRIC MERCIER
Vendredi, 27 juillet 2018 13:42
MISE À JOUR Vendredi, 27 juillet 2018 13:42

Désirant faire baisser le montant de sa prime d'assurance, un automobiliste de l'Alberta a fait changer son sexe sur son certificat de naissance.

«J'ai profité d'une faille dans le système», a expliqué l'albertain de 24 ans en entrevue avec CBC News.

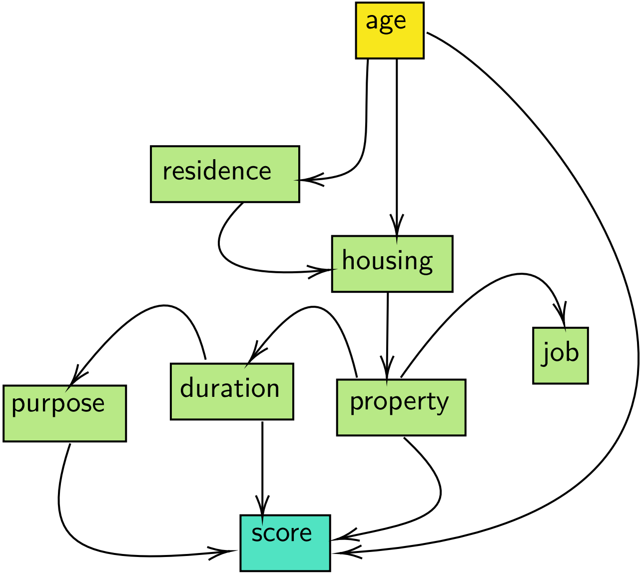
S'il n'a aucunement l'intention d'entreprendre de réelles procédures pour devenir une femme, le jeune homme désirant garder l'anonymat a tout de même fait changer son statut pour devenir «officiellement» une femme auprès du gouvernement albertain. Et il l'a fait uniquement pour économiser sur sa prime d'assurance.

Une différence marquée

L'idée de changement de sexe est venue au jeune homme après avoir appelé une compagnie d'assurance pour une soumission sur une voiture qu'il désirait acheter. Montant de la prime: 4517$ par année.

Curieux, le jeune homme a demandé à l'assureur combien lui coûterait une assurance sur le même véhicule s'il était une femme. On lui aurait alors répondu que la prime chuterait à 3423$ par année.

- P must be collected
- Looking for counterfactual
- DAGs are important



References I

Arya, S., Eckel, C., and Wichman, C. (2013). Anatomy of the credit score. *Journal of Economic Behavior & Organization*, 95:175–185.

Barry, L. and Charpentier, A. (2020). Personalization as a promise: Can big data change the practice of insurance? *Big Data & Society*, 7(1):2053951720935143.

Barry, L. and Charpentier, A. (2022). The Fairness of Machine Learning in Insurance: New Rags for an Old Man? . *ArXiv*.

Bartik, A. and Nelson, S. (2016). Deleting a signal: Evidence from pre-employment credit checks. *SSRN*, 2759560.

Berk, R., Heidari, H., Jabbari, S., Joseph, M., Kearns, M., Morgenstern, J., Neel, S., and Roth, A. (2017). A convex framework for fair regression. *arXiv*, 1706.02409.

Bigot, R. and Cayol, A. (2020). *Le droit des assurances en tableaux*. Ellipses.

Bigot, R. and Charpentier, A. (2019). Repenser la responsabilité, et la causalité. *Risques*, 120:123–128.

Bigot, R. and Charpentier, A. (2020). Quelle responsabilité pour les algorithmes? *Risques*, 121.

Bouk, D. (2015). *How Our Days Became Numbered: Risk and the Rise of the Statistical Individual*. The University of Chicago Press.

References II

Bureau d'Assurance du Canada (2021). Facts of the property and casualty insurance industry in canada.

Charpentier, A. (2014). *Computational Actuarial Science*. The R series. CRC Press.

Charpentier, A. (2022). *Assurance: biais, discrimination et équité*. Institut Louis Bachelier.

Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data*, 5(2):153–163.

Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., and Huq, A. (2017). Algorithmic decision making and the cost of fairness. *arXiv*, 1701.08230.

Duivesteijn, W. and Feelders, A. (2008). Nearest neighbour classification with monotonicity constraints. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 301–316. Springer.

Dwork, C., Hardt, M., Pitassi, T., Reingold, O., and Zemel, R. (2012). Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*, pages 214–226.

Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L., and Fei-Fei, L. (2017). Using deep learning and google street view to estimate the demographic makeup of neighborhoods across the united states. *Proceedings of the National Academy of Sciences*, 114(50):13108–13113.

References III

Grari, V., Charpentier, A., Lamprier, S., and Detyniecki, M. (2022). A fair pricing model via adversarial learning. *ArXiv*, 2202.12008.

Hardt, M., Price, E., and Srebro, N. (2016). Equality of opportunity in supervised learning. *Advances in neural information processing systems*, 29:3315–3323.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American statistical Association*, 81(396):945–960.

Ilic, L., Sawada, M., and Zarzelli, A. (2019). Deep mapping gentrification in a large canadian city using deep learning and google street view. *PloS one*, 14(3):e0212814.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., and Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301):790–794.

Kabler, B. (2004). Insurance-based credit scores: Impact on minority and low income populations in missouri. *State of Missouri Departement of Insurance*.

Kita, K. and Kidziński, Ł. (2019). Google street view image of a house predicts car accident risk of its resident. *arXiv*, 1904.05270.

Kiviat, B. (2019). The moral limits of predictive practices: The case of credit-based insurance scores. *American Sociological Review*, 84(6):1134–1158.

References IV

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., and Mullainathan, S. (2017). Human Decisions and Machine Predictions. *The Quarterly Journal of Economics*, 133(1):237–293.

Kranzberg, M. (1986). Technology and history:" kranzberg's laws". *Technology and culture*, 27(3):544–560.

Kusner, M. J., Loftus, J., Russell, C., and Silva, R. (2017). Counterfactual fairness. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., editors, *Advances in Neural Information Processing Systems 30*, pages 4066–4076. NIPS.

Lauer, J. (2017). *Creditworthy: A History of Consumer Surveillance and Financial Identity in America*. Columbia University Press.

Law, S., Paige, B., and Russell, C. (2019). Take a look around: Using street view and satellite images to estimate house prices. *ACM Transactions on Intelligent Systems and Technology*, 10(5).

Lehtonen, T.-K. and Liukko, J. (2015). Producing solidarity, inequality and exclusion through insurance. *Res Publica*, 21(2):155–169.

Miller, M. J., Smith, R. A., and Southwood, K. N. (2003). The relationship of credit-based insurance scores to private passenger automobile insurance loss propensity. *Actuarial Study, Epic Actuaries*.

Morris, D. S., Schwarcz, D., and Teitelbaum, J. C. (2017). Do credit-based insurance scores proxy for income in predicting auto claim risk? *Journal of Empirical Legal Studies*, 14(2):397–423.

References V

Neyman, J., Dabrowska, D. M., and Speed, T. (1923). On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, pages 465–472.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Pearl, J. (1998). Graphs, causality, and structural equation models. *Sociological Methods & Research*, 27(2):226–284.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688.

Seresinhe, C. I., Preis, T., and Moat, H. S. (2017). Using deep learning to quantify the beauty of outdoor places. *Royal Society open science*, 4(7):170170.

Struyck, N. (1912). *Les oeuvres de Nicolas Struyck (1687-1769): qui se rapportent au calcul des chances, à la statistique général, z la statistique des décès et aux rentes viagèter*. Société générale néerlandaise d'assurances sur la vie et de rentes viagères.

Suresh, H. and Guttag, J. V. (2019). A framework for understanding sources of harm throughout the machine learning life cycle. *arXiv*, 1901.10002.

Thiery, Y. and Van Schoubroeck, C. (2006). Fairness and equality in insurance classification. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 31(2):190–211.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, 20.