

References

- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). There's software used across the country to predict future criminals. *And it's biased against blacks*. ProPublica.
- Barocas, S., Hardt, M., & Narayanan, A. (2023). *Fairness and machine learning: Limitations and opportunities*. MIT press.
- Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *Calif. L. Rev.*, 104, 671.
- Bellamy, R. K., Dey, K., Hind, M., Hoffman, S. C., Houde, S., Kannan, K., Lohia, P., Martino, J., Mehta, S., & Mojsilović, A. (2019). AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM Journal of Research and Development*, 63(4/5), 4: 1-4: 15.
- Binns, R. (2020). On the apparent conflict between individual and group fairness. Proceedings of the 2020 conference on fairness, accountability, and transparency,
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321-357.
- Corbett-Davies, S., Gaebler, J. D., Nilforoshan, H., Shroff, R., & Goel, S. (2023). The measure and mismeasure of fairness. *Journal of Machine Learning Research*, 24(312), 1-117.
- Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Mach. Learn.*, 20(3), 273-297. <https://doi.org/10.1023/a:1022627411411>
- Cover, T., & Hart, P. (2006). Nearest neighbor pattern classification. *IEEE Trans. Inf. Theor.*, 13(1), 21-27. <https://doi.org/10.1109/tit.1967.1053964>
- Cramer, J. S. (2002). The Origins of Logistic Regression. *Tinbergen Institute Working Paper*. <http://dx.doi.org/10.2139/ssrn.360300>
- Domingos, P., & Pazzani, M. (1997). On the Optimality of the Simple Bayesian Classifier under Zero-One Loss. *Machine Learning*, 29(2), 103-130. <https://doi.org/10.1023/A:1007413511361>
- Feldman, M., Friedler, S. A., Moeller, J., Scheidegger, C., & Venkatasubramanian, S. (2015). Certifying and removing disparate impact. proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining,
- Friedler, S. A., Scheidegger, C., Venkatasubramanian, S., Choudhary, S., Hamilton, E. P., & Roth, D. (2019). A comparative study of fairness-enhancing interventions in machine learning. Proceedings of the conference on fairness, accountability, and transparency,
- Hardt, M., Price, E., & Srebro, N. (2016). Equality of opportunity in supervised learning. *Advances in neural information processing systems*, 29.
- In-Vehicle Coupon Recommendation* ((2020). UC Irvine Machine Learning Repository. <https://doi.org/10.24432/C5GS4P>
- Kamiran, F., & Calders, T. (2012). Data preprocessing techniques for classification without discrimination. *Knowledge and Information Systems*, 33(1), 1-33. <https://doi.org/10.1007/s10115-011-0463-8>
- Kearns, M., Neel, S., Roth, A., & Wu, Z. S. (2018). Preventing fairness gerrymandering: Auditing and learning for subgroup fairness. International conference on machine learning,
- Kharoua, R. E. (2023). *Predict Liver Disease: 1700 Records Dataset* (<https://doi.org/10.24432/C55S3H>
- Kohavi, B. B. R. (1996). *Adult* (UC Irvine Machine Learning Repository. <https://doi.org/10.24432/C5XW20>
- Lien, I.-C. Y. C.-h. (2009). *The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients* (UC Irvine Machine Learning Repository. <https://doi.org/10.24432/C55S3H>
- Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1), 50-60. <http://www.jstor.org/stable/2236101>
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM computing surveys (CSUR)*, 54(6), 1-35.
- Naveen, D. (2019). *Logistic regression to predict heart disease* (<https://www.kaggle.com/datasets/dileep070/heart-disease-prediction-using-logistic-regression>
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447-453.
- ProPublica. (2016). *COMPAS Recidivism Racial Bias* (Kaggle. <https://www.kaggle.com/datasets/danofer/compass>
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81-106. <https://doi.org/10.1007/BF00116251>
- Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. Proceedings of the conference on fairness, accountability, and transparency,
- Student. (1908). The Probable Error of a Mean. *Biometrika*, 6(1), 1-25. <https://doi.org/10.2307/2331554>
- Suresh, H., & Gutttag, J. (2021). A framework for understanding sources of harm throughout the machine learning life cycle. Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization,
- Uddin, S., Lu, H., & Hajati, F. (2026). Fairness-preserving framework for machine learning: data bias quantification, model evaluation, and robustness across multiple datasets. *AI and Ethics*, 6(1), 25.
- Uddin, S., Lu, H., Rahman, A., & Gao, J. (2024). A novel approach for assessing fairness in deployed machine learning algorithms. *Scientific Reports*, 14(1), 17753.
- Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *International data privacy law*, 7(2), 76-99.
- Welch, B. L. (1947). The Generalization of 'Student's' Problem when Several Different Population Variances are Involved. *Biometrika*, 34(1/2), 28-35. <https://doi.org/10.2307/2332510>
- Yun, H. (2026). *preprocessing_comparative_study_for_fair_ml*. https://github.com/geun-yun/preprocessing_comparative_study_for_fair_ml
- Zemel, R., Wu, Y., Swersky, K., Pitassi, T., & Dwork, C. (2013). Learning fair representations. International conference on machine learning,