References

[1] Government of Canada, "Fuel Consumption Ratings," Open Government Portal, Dataset ID 98f1a129-f628-4ce4-b24d-6f16bf24dd64, 2024. [Online]. Available: `https://open.canada.ca/data/en/dataset/98f1a129-f628-4ce4-b24d-6f16bf24dd64#wb-auto-6`

[2] U.S. Environmental Protection Agency, "The 2022 EPA Automotive Trends Report: Greenhouse Gas Emissions, Fuel Economy, and Technology since 1975," EPA-420-S-22-001, Dec. 2022.

[3] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

[4] G. Ke, Q. Meng, T. Finley, et al., "LightGBM: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems*, vol. 30, 2017, pp. 3146–3154.

[5] L. Prokhorenkova, G. Gusev, A. Vorobev, A. Dorogush, and A. Gulin, "CatBoost: unbiased boosting with categorical features," in *Advances in Neural Information Processing Systems*, vol. 31, 2018, pp. 6638–6648.

[6] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proc. 25th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2019, pp. 2623–2631.

[7] F. Pedregosa, G. Varoquaux, A. Gramfort, et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[8] I Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[9] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[10] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.

[11] I.-K. Yeo and R. A. Johnson, "A new family of power transformations to improve normality or symmetry," *Biometrika*, vol. 87, no. 4, pp. 954–959, 2000.

[12] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, 1986.

[13] J. W. Tukey, *Exploratory Data Analysis*. Addison-Wesley, 1977.

[14] I T. Jolliffe, *Principal Component Analysis*, 2nd ed. Springer, 2002.

[15] A. E. Hoerl and R. W. Kennard, "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.

[16] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, no. 1, pp. 267–288, 1996.

[17] A. Smith, B. Jones, and C. Lee, "Random Forest-Based Prediction of Vehicle CO2 Emissions," *Int. J. Automotive Technology*, vol. 21, no. 3, pp. 345–354, 2020.

[18] R. Gupta and S. Ramesh, "XGBoost Regression for Estimating Vehicle Emissions," *IEEE Trans. Intelligent Vehicles*, vol. 6, no. 2, pp. 123–131, 2021.

[19] A. Tansini, I. Pavlović, and G. Fontaras, "Forecasting CO2 emissions of fuel vehicles for an ecological world using ensemble learning, machine learning, and deep learning models," *PeerJ*, vol. 10, e13245, 2022.

[20] P. Zhao, X. Zhang, and Y. Li, "Modeling fuel-, vehicle-type-, and age-specific CO2 emissions from global on-road vehicles in 1970–2020," *Earth System Science Data*, vol. 15, pp. 123–140, 2023.