Gradient Boosted Utility Models

Nicolas Salvadé, Tim Hillel

Behaviour and Infrastructure Group, UCL

15th Workshop on Discrete Choice Models

02.06.2023

Gradient Boosted Utility Models

1. Introduction
2. Methodology
3. Results
4. Future work

1. Introduction

Interpretable ML model

- Attempts to make machine learning models interpretable, but **with little success.**

- We propose a machine learning model that is **fully interpretable** based on **gradient boosting decision trees** (GBDT), and inspired by **random utility** models (RUM)

- We are able to derive **non-linear utility** functions

Introductory example – RUM

3 alternatives

Driving	PT	Walking
TT_d $Cost_d$	TT_{PT} $Cost_{PT}$	TT_w

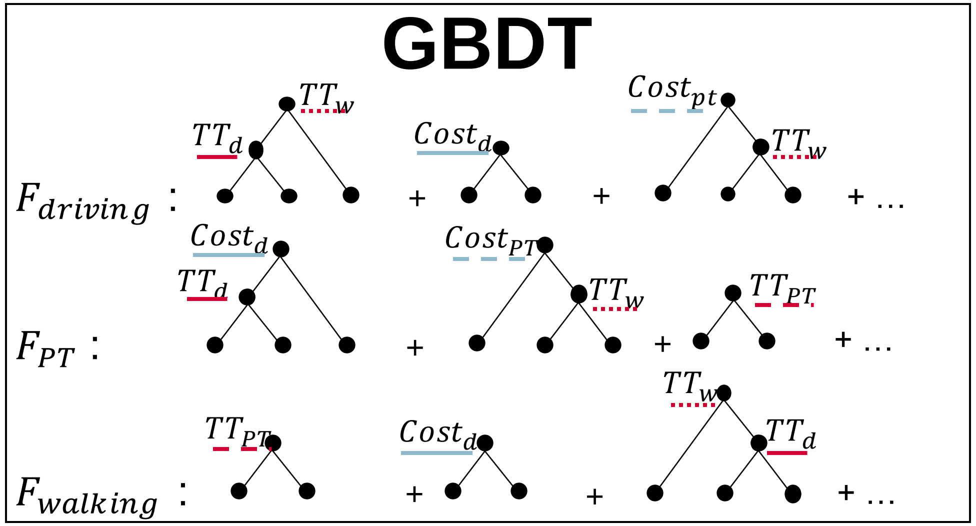
RUM (MNL)

$$V_{driving} = ASC_d + \beta_d TT_d + \beta_d Cost_d$$
$$V_{PT} = ASC_{PT} + \beta_{PT} TT_{PT} + \beta_{PT} Cost_{PT}$$
$$V_{walking} = ASC_w + \beta_w TT_w$$

Introductory example – GBDT

3 alternatives

Driving	PT	Walking
TT_d $Cost_d$	TT_{PT} $Cost_{PT}$	TT_w

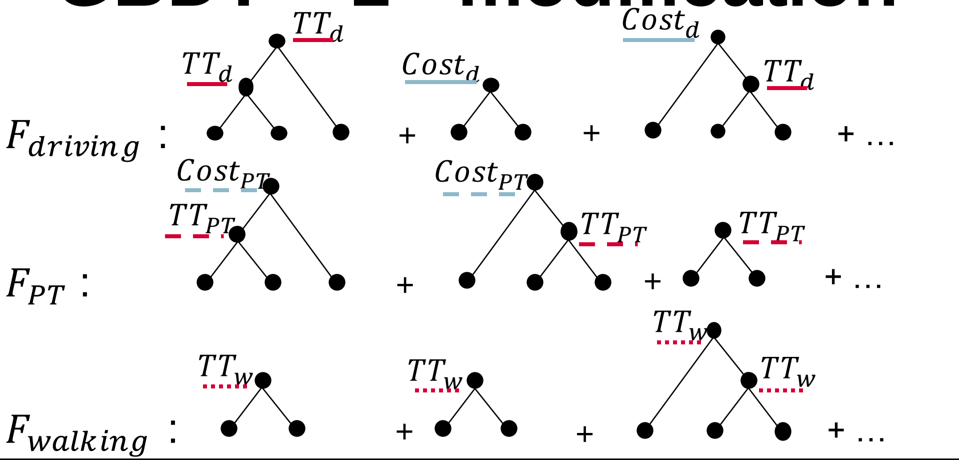


Introductory example – 1st modification

3 alternatives

Driving	PT	Walking
TT_d $Cost_d$	TT_{PT} $Cost_{PT}$	TT_w

GBDT – 1st modification



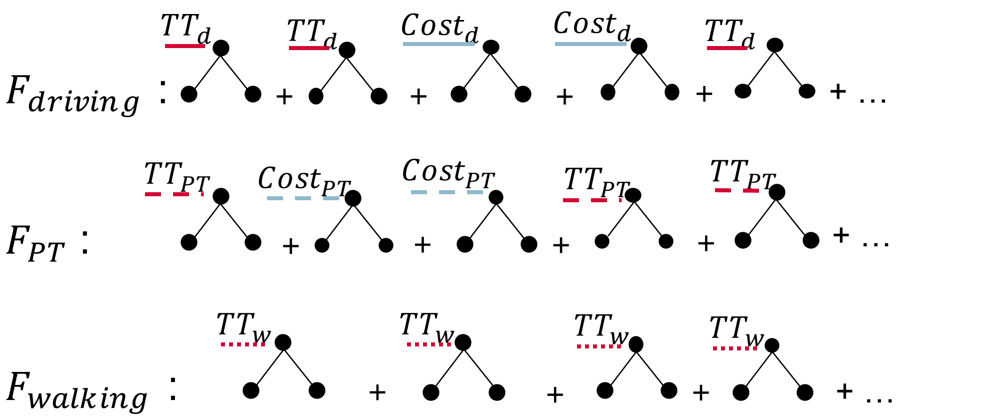
1. Choosing which features compose an ensemble

Introductory example – 2nd modification

3 alternatives

Driving	PT	Walking
TT_d $Cost_d$	TT_{PT} $Cost_{PT}$	TT_w

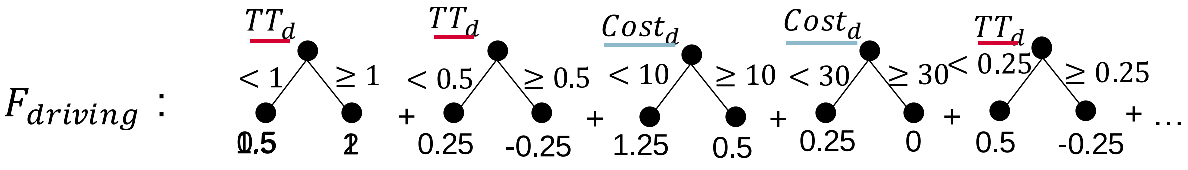
GBDT – 2nd modification



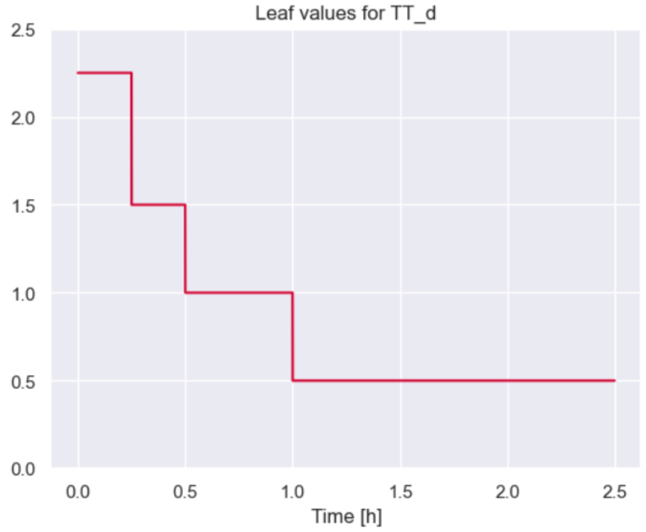
1. Choosing which features compose an ensemble
2. Restricting feature interaction

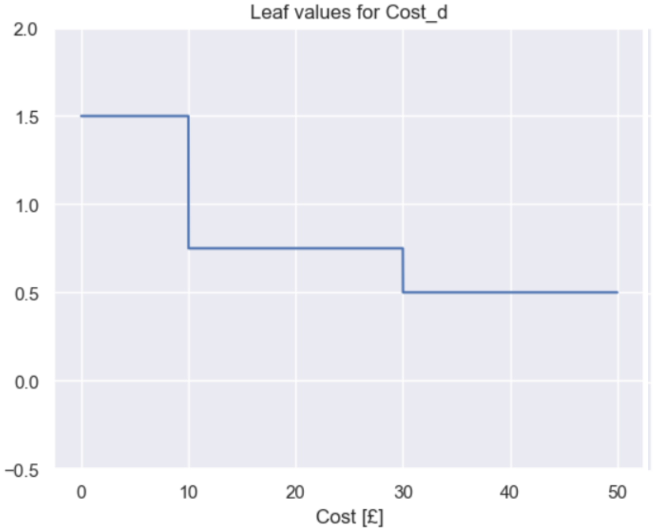
Introductory example – 3rd modification

GBDT – 3rd modification



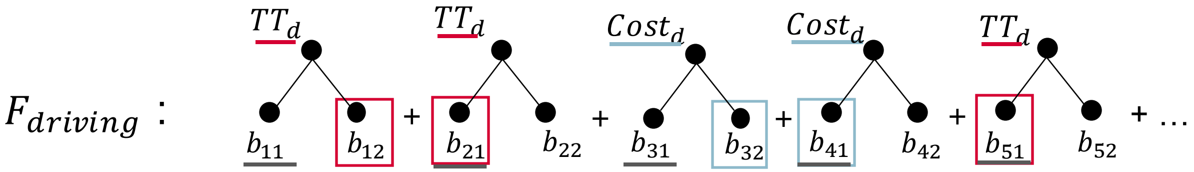
1. Choosing which features compose an ensemble
2. Restricting feature interaction
3. Monotonic constraint





Introductory example – Utility function

RUMBooster



$$V'_{driving} = (b_{12} + b_{21} + b_{51} + \cdots) \ +(b_{32} + b_{41} + \cdots)$$

$$V_{driving} \ = ASC_d \ + \beta_d TT_d \quad + \beta_d Cost_d$$

Motivations

- GBDT have great **predictive power**

  LightGBM and XGBoost are **state-of-the art** libraries and **powerful** tools

- Lack **behaviour interpretability** and exhibit **poor extrapolation** properties (Martin-Baos et al. 2023)

- We propose **gradient boosted utility models** (RUMBooster) were we aim to combine the predictive power of GBDT with RUM interpretability

RUMBoost

1. **Utility** specification: only chosen features can contribute to an alternative utility function

2. **Feature** interaction constraint: only specified features can interact with each other

3. **Monotonic** constraint: left and right leaf values can only be increasing or decreasing for a feature

RUMBooster

Interpretable model where **non-linear utilities** can be derived

2. Methodology

RUM to RUMBoost

- We can estimate **any MNL model** with gradient boosting
- We follow the **same** utility specification
- **Bounds** on Beta parameters are interpreted as **monotonic constraints**
- Since there is no feature interaction in RUMBoost, very **few risk of overfitting**

3. Results

LIVE DEMO

MNL model and dataset

- Used MNL model from Martin-Baos et al. (2023)

  The model has **76 parameters,** and its performance is compared with several ML models

- The model is trained on the **LPMC** dataset (Hillel et al. 2018)

  4 alternatives: walking, cycling, PT and driving, and over 80000 observations

Model specification

- **Ensembles** of RUMBooster are based on the **LightGBM** library

- **No interaction** between features

- Monotonic constraints:

Table 1: Monotonic constraints on features

Negative	Positive
Travel time, cost	Car ownership*, driving license*

For car utility only

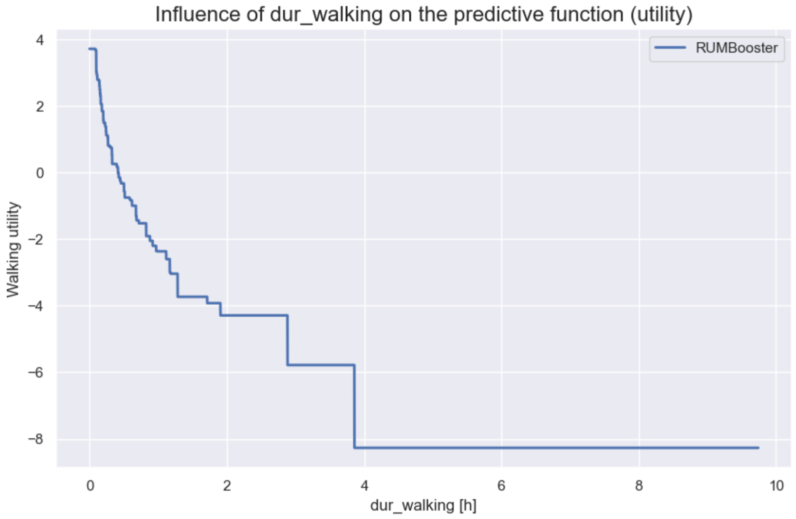
Negative cross-entropy

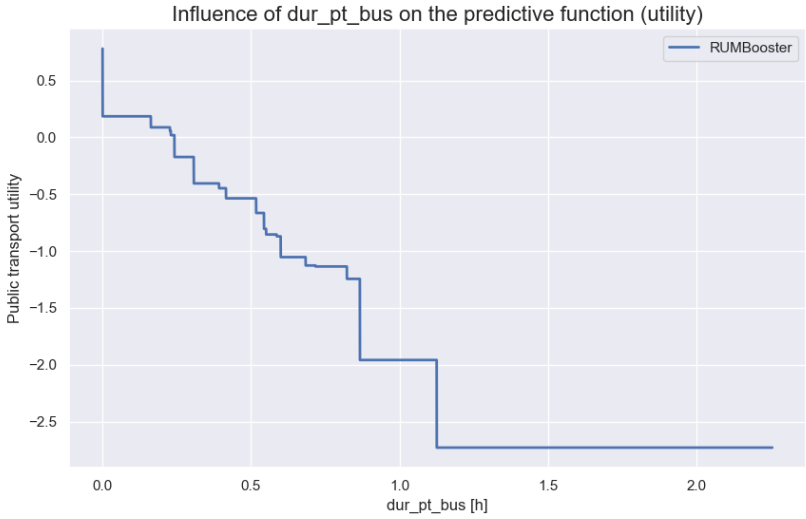
Table 2: Benchmark of classification on LTDS dataset (negative cross-entropy)

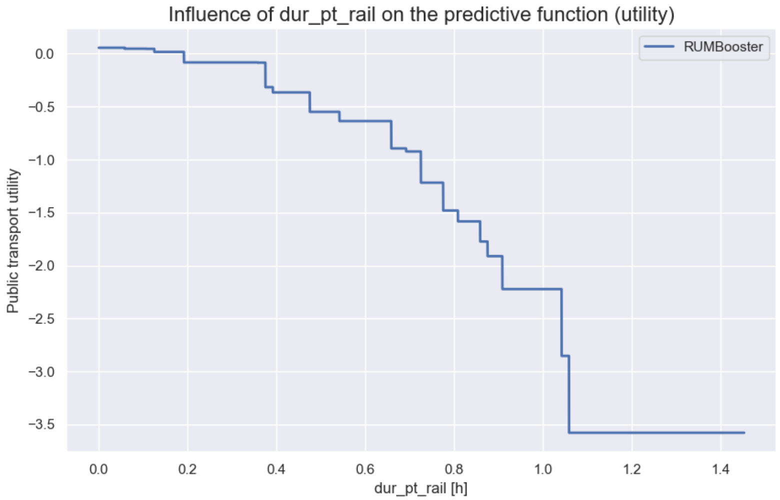
	MNL	NN	DNN	RF	SVM	XGBoost	RUMBooster
Martin-Baos et al. (2023)	0.7164	0.6728	0.6702	0.6900	0.6755	0.6568	0.6791

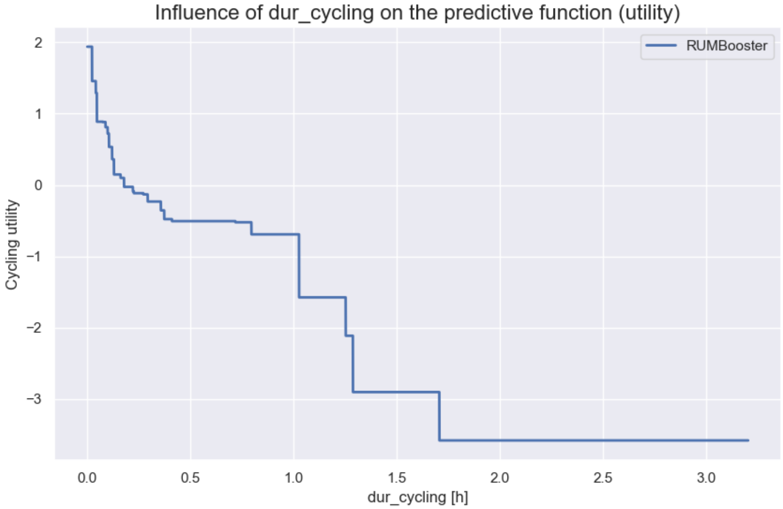
Learning rate = 0.1

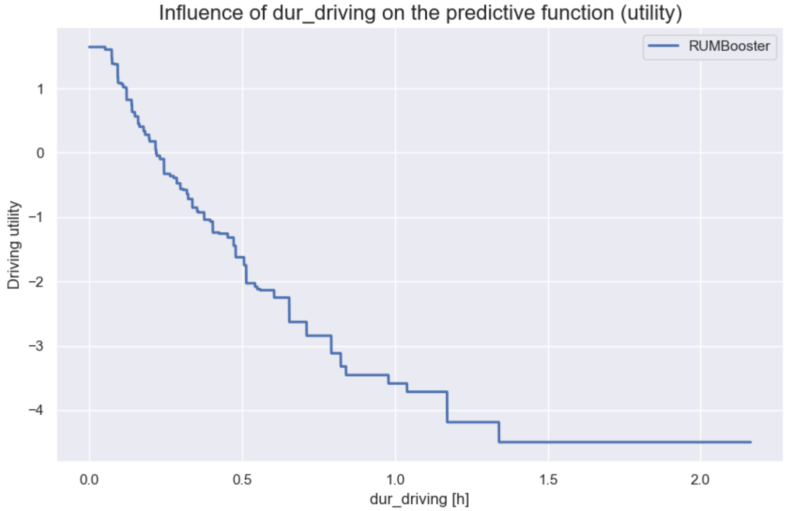
Impact of travel time



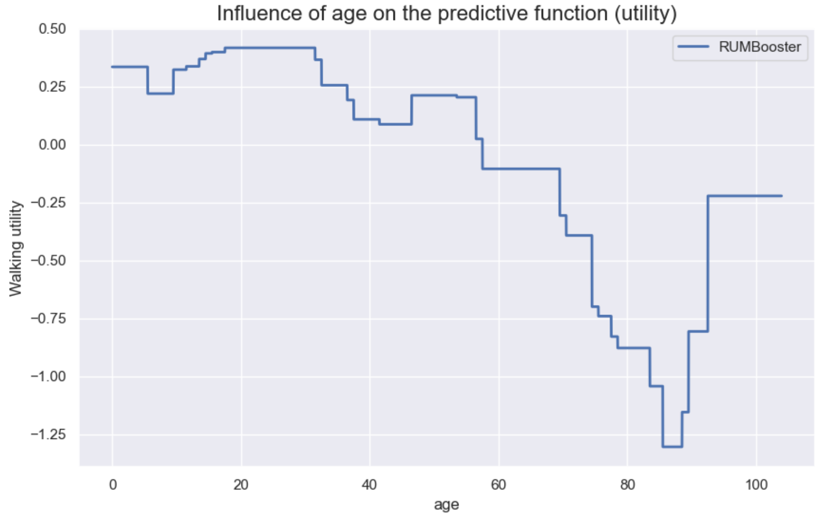


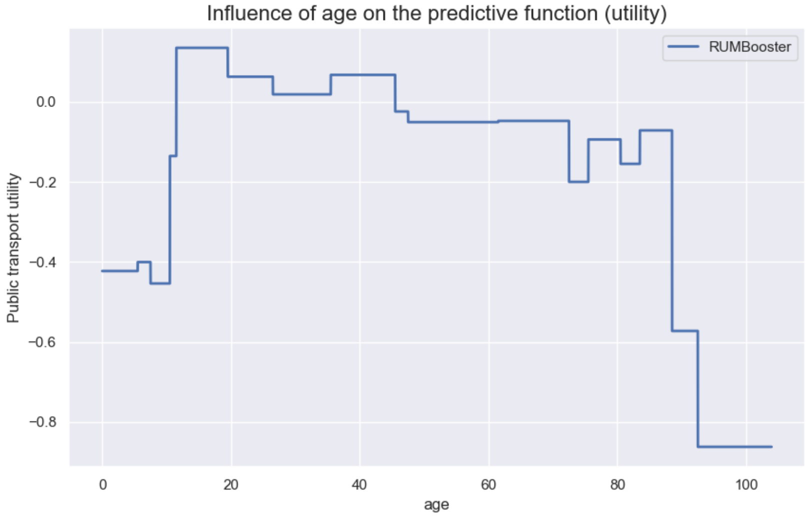


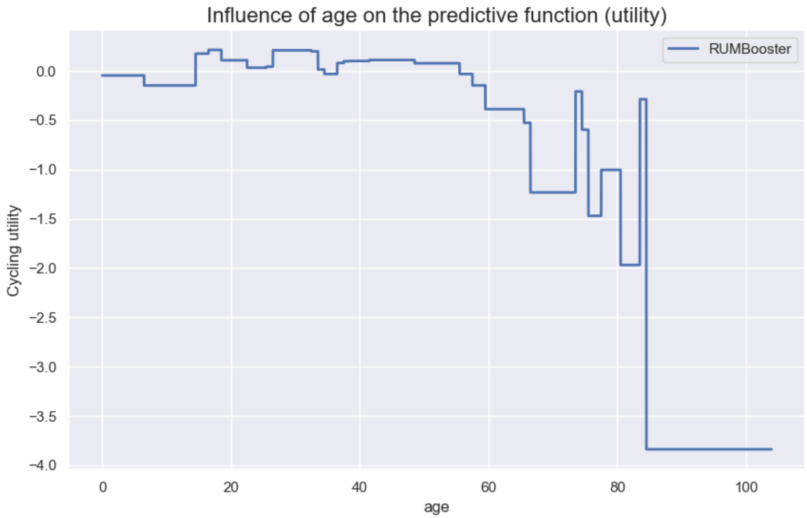


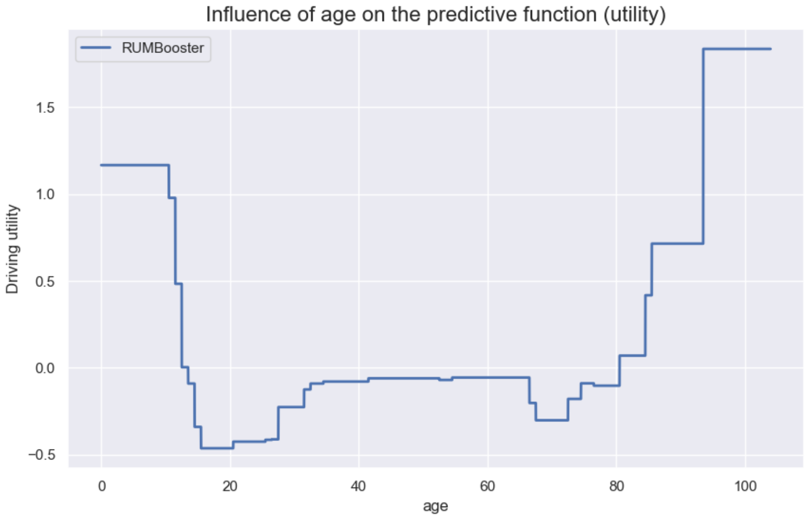


Impact of age









4. Further work

Further work

- Intersecting continuous and categorical variables as **latent variables** to keep nonlinear utility interpretability but increase predictive power

- **Generalise** RUMBooster to MEV models

- Apply the methodology on **different** datasets

Thank you!

References

- Bierlaire, M. 2023. A short introduction to Biogeme. Technical report TRANSP-OR 230620. Transport and Mobility Laboratory, ENAC, EPFL.
- Cano, J.-R., Gutiérrez, P.A., Krawczyk, B., Woźniak, M., García, S., 2019. Monotonic classification: An overview on algorithms, performance measures and data sets. Neurocomputing 341, 168–182. https://doi.org/10.1016/j.neucom.2019.02.024
- Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 785–794. https://doi.org/10.1145/2939672.2939785
- Friedman, J., Hastie, T., Tibshirani, R., 2001. ADDITIVE LOGISTIC REGRESSION: A STATISTICAL VIEW OF BOOSTING.
- Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Statist. 29. https://doi.org/10.1214/aos/1013203451
- Hillel, T., Bierlaire, M., Elshafie, M., Jin, Y., 2019. Weak teachers: Assisted specification of discrete choice models using ensemble learning.
- Hillel, T., Bierlaire, M., Elshafie, M.Z.E.B., Jin, Y., 2021. A systematic review of machine learning classification methodologies for modelling passenger mode choice. Journal of Choice Modelling 38, 100221. https://doi.org/10.1016/j.jocm.2020.100221
- Hillel, T., Elshafie, M.Z.E.B., Jin, Y., 2018. Recreating passenger mode choice-sets for transport simulation: A case study of London, UK. Proceedings of the Institution of Civil Engineers - Smart Infrastructure and Construction 171, 29–42. https://doi.org/10.1680/jsmic.17.00018
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. LightGBM: A Highly Efficient Gradient Boosting Decision Tree.
- Martín-Baos, J.Á., López-Gómez, J.A., Rodriguez-Benitez, L., Hillel, T., García-Ródenas, R., 2023. A prediction and behavioural analysis of machine learning methods for modelling travel mode choice.
- Shi, Y., Li, J., Li, Z., 2019. Gradient Boosting With Piece-Wise Linear Regression Trees.