Performance of econometric and Machine Learning models for the economic study of discrete consumer choices

Thesis abstract

Written by

Nikita Gusarov
GAEL
Univ. Grenoble Alpes
Grenoble, 38000
nikita.gusarov@univ-grenoble-alpes.fr

Under supervision of

Iragaël Joly
GAEL
Univ. Grenoble Alpes
Grenoble, 38000
iragael.joly@grenoble-inp.fr

Pierre Lemaire
G-SCOP
Grenoble INP
Grenoble, 38000
pierre.lemaire@grenoble-inp.fr

Abstract

This works is a cross-disciplinary study of discrete choice modelling, adressing both econometrics and machine learning (ML) techniques applied to individual choice modelling. The problematic arises from the insufficient points of contact among users (economists and engineers) and data scientists, who pursue different objectives, although using similar techniques. To bridge the interdisciplinary gap, the PhD work aims to define a common performance analysis and a comparison framework suitable for different models issued from econometrics and ML. It adresses the performance comparison task from the research procedure perspecive, incorporating all the steps potentially affecting the performance perceptions. To demonstrate the framework's capabilities we propose a serie of applied studies.

***K*eywords** Consumer choice · Preference studies · Econometrics · Data science · Machine learning · Artificial datasets

1 The issues and complexity of model performance assesment

With the development of computational devices and increasing data availability more novel and resource heavy data analysis methods are introduced. In particular, the advances in statistical learning (Hastie, Tibshirani, and Friedman 2009) and data science (Donoho 2017) of the past decades have resulted in propagation of Machine Learning (ML) techniques. The economics discipline is not an exception, the most resource demanding models of previous decades can be executed in several minutes and current research is more and more focused on the big data and analysis automation. The number of available data analysis strategies may make it extremely difficult for the non-experts to select the optimal solution (Athey and Imbens 2019). To address this issue there is an important need for better understanding of the various models' strength and weak points.

In order to limit the scope of our study we will focus our attention on the discrete choice model family in the context of the individual choice related studies, this limitation will establish a baseline for the discussion.

Transportation, **marketing** and **health economics** are among the most impacted fields when it comes to the novel strategies of Choice Modelling (CM). The history of the individual decision making modelling may be traced as far as Luce (1957) works, and the later introduced state of art by McFadden (1974). The other disciplines, such as managerial studies or sociology diverge sometimes from the economics vision of the problematic, accentuating the importance of the domain specific knowledge. The different knowledge acquisition strategies also affect the available toolset and conceptual vision on the data analysis process.

2 The *universal* performance comparison framework

To tackle the issue of inconsistency of theoretical base among the different application fields and knowledge acquisition strategies, we propose a *universal* approach for the model usage exploration and performance analysis. The proposed framework for performance analysis and comparison is based on the standard scientific procedure, with sufficient flexibility to extend it to other fields and disciplines. The adopted procedure may be seen as quite close to many applied and theoretical economic papers. Unfortunately no known to us work, except for the manuscript of Williams and Ortuzar (1982), does not approach the scientific workflow from the same perspective and in such detail as us.

The framework's concept is tightly tied to the performance understanding and definition. We assume that performance should be defined in relation to the particular research question and the indicators allowing to answer such question. But what are the eventual targets the performance can be matched against in the context of the DCM performance analysis? The majority of researchers focus on the predictive accuracy as the main performance metrics for their sample size requirements calculation. However, according to the interdisciplinary works (Japkowicz and Shah 2011) the performance of competing models may be assessed over several criteria: (1) quality of data adjustments; (2) predictive capacity; (3) quality of the field specific (ex: economic and behavioural) indicators derived from estimates; and (4) algorithmic efficiency and computational costs. We are going to explore the available metrics and approaches more in detail.

Another key element of the framework is the data, as we focus on the numerical analysis. In many cases the research is constructed primarily around the available data, which limits the research questions available for exploration. There exist many various problems related to the data collection and usage. Those may be divided into: (1) **theoretical biases** induced through the inconsistency between the chosen theoretical assumptions and actual human behaviour in the CE context; and (2) **statistical biases** associated with experimental design construction and inappropriate modelling strategy choice. In the thesis we address the eventual analysis problems related to data usage, as well as how to use simulation to explore performance differences in a fully controlled and reproducible environment.

Last, but not the least we introduce the *model* concept as an inevitable part of the research procedure. Given the context of the research question some of them might be simply incapable to produce any meaningful answer, while the other will compete producing seemingly identical results. As previously pointed out, in economics, consumer choices data are mainly studied through classification tools from machine learning techniques or regression tools like discrete choice models from econometric techniques. These two practices in particular illustrate two distinct approaches to applying statistical learning. As described by Breiman et al. (2001) and later by Athey and Imbens (2019): the ML focus on the predictive qualities and Econometrics attempts to decipher the underlying properties of the data. Engineering sciences and Computer sciences focuses mainly on ML techniques, whereas in Economics and other applied Social sciences, the scientific community prefers to implement the traditional econometrics techniques to explore hidden patterns (Athey 2018). The understanding of appropriate model families and their place in the framework's context are shown in this part.

3 The framework in action: case studies

In the thesis not only we outline the performance comparison framework, but also we focus on its usage. The reader will encounter several examples of its implementations serving to answer different research needs. Several case studies are proposed, each of them focusing on the different element of the framework: (1) performance criteria selection effects, (2) model comparisons and (3) data processing and relied issues.

3.1 Performance: *Willingness to pay quality estimates in commute mode choice*

ITEA 2023 paper: "*Willingness to pay in commute mode choice: Model performance comparison under sample size and balance impacts*"

In economics studies one of the wide-spread target metrics is the *Willingness to Pay* (WTP) of individuals for particular attributes of transportation mode choices. There already exists a wast literature addressing some major issues of the WTP elicitation task. We propose a performance comparison framework, allowing to systematize the previous research. With its help, in this work we explore models perform in WTP elicitation task under potential misspecifications, sample size and dataset balance changes. The `swissmetro` dataset is used for application purposes. We use simulation to vary sample size and configuration, which are used for model estimation and WTP elicitation. The results illustrate the variability in WTP estimates under different configurations, confirming some of the previously published results.

3.2 Data: *Exploration of model performances in the presence of heterogeneous preferences*

DA2PL 2020 paper: "*Exploration of model performances in the presence of heterogeneous preferences and random effects utilities*"

This work is adresses the issue of data specificities affecting the consumer preference analysis. Through usage of a simulation and theory-testing framework we illustrate the potential spillowers of incorrect modelling strategy choices under erroneous assumptions over the data generation laws. The flexibility of the framework in theory-testing and models comparison over economics and statistical indicators is illustrated based on the work of Michaud, Llerena, and Joly (2012). Two datasets are generated using the predefined utility functions simulating the presence of homogeneous and heterogeneous individual preferences for alternatives' attributes. Three models issued from econometrics and ML disciplines are estimated and compared across multiple criteria.
The highligh the strenghts and weaknesses of the most common models for both the homogeneous or heterogeneous consumer preferences.

3.3 Models: *Issues of model selection*

ICMC 2024 paper / IATBR 2024[1]: "*Issues of model selection: Willingness to pay in commute mode choice*"

This work adresses the popular issue of econometrics and ML models comparison in the context of the commute mode choice modelling. We propose a model performance comparison framework inspired by the work of Williams and Ortuzar (1982) for model selection purposes. The main focus is made on contrasting of the traditional DCM approach with the emerging and not yet fully accepted ML driven methodology (Wang, Wang, and Zhao 2020; Wang et al. 2021). For application purposes we use the popular `swissmetro` dataset.

4 References

Athey, Susan. 2018. "The Impact of Machine Learning on Economics." In *The Economics of Artificial Intelligence: An Agenda*, 507–47. University of Chicago Press. `http://www.nber.org/chapters/c14009`.

Athey, Susan, and Guido W. Imbens. 2019. "Machine Learning Methods That Economists Should Know About." *Annual Review of Economics* 11 (1): 685–725. `https://doi.org/10.1146/annurev-economics-080217-053433`.

Breiman, Leo et al. 2001. "Statistical Modeling: The Two Cultures (with Comments and a Rejoinder by the Author)." *Statistical Science* 16 (3): 199–231.

Donoho, David. 2017. "50 Years of Data Science." *Journal of Computational and Graphical Statistics* 26 (4): 745–66.

Gusarov, Nikita. 2022a. "A Discrete Choice Experiment (DCE) Simulator: 'Dcesimulatr'." Grenoble, France. `https://github.com/nikitagusarov/dcesimulatr`.

———. 2022b. "Performance Metrics for Discrete Choice Analysis (DCA): 'Performancer'." Grenoble, France. `https://github.com/nikitagusarov/performancer/`.

Gusarov, Nikita, Iragaël Joly, and Pierre Lemaire. 2023. "Willingness to Pay Quality Estimates in Commute Mode Choice: Model Performance Comparison Under Sample Size and Balance Impacts." In *ITEA 2023 Proceedings*. Santander, Spain: GAEL, G-SCOP, Univ. Grenoble Alpes, CNRS, INRAE, Grenoble INP.

Gusarov, Nikita, Amirreza Talebijamalabad, and Iragaël Joly. 2020. "Exploration of Model Performances in the Presence of Heterogeneous Preferences and Random Effects Utilities Awareness." In *DA2PL 2020 Proceedings*. Trento, Italy. `https://hal.archives-ouvertes.fr/hal-03019739`.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Science & Business Media.

[1]Submitted for review.

Japkowicz, Nathalie, and Mohak Shah. 2011. *Evaluating Learning Algorithms: A Classification Perspective*. Cambridge University Press. https://doi.org/10.1017/CBO9780511921803.

Luce, R Duncan. 1957. "A Theory of Individual Choice Behavior." COLUMBIA UNIV NEW YORK BUREAU OF APPLIED SOCIAL RESEARCH.

McFadden, Daniel. 1974. "The Measurement of Urban Travel Demand." *Journal of Public Economics* 3 (4): 303–28. https://doi.org/10.1016/0047-2727(74)90003-6.

Michaud, Celine, Daniel Llerena, and Iragael Joly. 2012. "Willingness to Pay for Environmental Attributes of Non-Food Agricultural Products: A Real Choice Experiment." *European Review of Agricultural Economics* 40 (2): 313–29. https://doi.org/10.1093/erae/jbs025.

Wang, Shenhao, Qingyi Wang, Nate Bailey, and Jinhua Zhao. 2021. "Deep Neural Networks for Choice Analysis: A Statistical Learning Theory Perspective." *Transportation Research Part B: Methodological* 148 (June): 60–81. https://doi.org/10.1016/j.trb.2021.03.011.

Wang, Shenhao, Qingyi Wang, and Jinhua Zhao. 2020. "Deep Neural Networks for Choice Analysis: Extracting Complete Economic Information for Interpretation." *Transportation Research Part C: Emerging Technologies* 118 (September): 102701. https://doi.org/10.1016/j.trc.2020.102701.

Williams, H. C. W. L., and J. D. Ortuzar. 1982. "Behavioural Theories of Dispersion and the Mis-Specification of Travel Demand Models." *Transportation Research Part B: Methodological* 16 (3): 167–219. https://doi.org/10.1016/0191-2615(82)90024-8.