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

LE COMITE DE SUIVI INDIVIDUEL

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

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

August 24, 2021

Abstract

This works is a cross-disciplinary study of econometrics and machine learning (ML) models applied to consumer 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. First of all, as an interdisciplinary research project, the PhD work aims to define a common analysis and a comparison grid of performances of different models issued from econometrics and ML in application to individual decision making process. A second objective is to incorporate the recent developments in the decision making theory into the general model comparison frameworks. A third objective is to evaluate performance of these different modelling techniques in presence of specific behavioural bias in applied fields. To bridge the interdisciplinary gap an integrated simulation and theory-testing framework is proposed. It incorporates all essential steps from hypothetical setting generation to the comparison of various performance metrics.

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

1 Introduction

The advances in statistical learning (Hastie, Tibshirani, and Friedman 2009), data analysis (Zielesny 2011) and data science (Donoho 2017) of the past decades have resulted in propagation of Machine Learning (ML) techniques. Nowadays, it is impossible to imagine a field of science that is not benefiting from the fruits of statistical learning. 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 illustrate two distinct approaches to applying statistical learning. As described by Breiman and others (2001) and later by Athey and Imbens (2019): the Machine Learning (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).

But how the scientists working on the applied problems, such as modelling discrete individual choices, can correctly select the appropriate tools given the confusions between the fields and disciplines? Some pluri-disciplinary scientists make attempts to breach this wall between ML and Econometrics on a high level: Athey and Imbens (2019), Mullainathan and Spiess (2017), Varian (2014). Their advances are mostly focused on the general interdisciplinary question, without entering into the application specific details.

2 Project presentation

The differences among the scientific communities make it difficult to correctly select a model to be used in a given context. There is no known to us work proposing a complete and comprehensive taxonomy of the DCM, studying at the same time the questions related to model performances, applications and reliability.

One of the major breaks, which prevents the appearance of a unique solution is the ambiguity in the performance metrics. Although there exists a expanding literature in the DCM performance evaluation (Flach (2019), Japkowicz and Shah (2011)), the consensus on applying the developed techniques is not reached. The existing performance metrics are far from optimal and rarely offer a complete information about model performances, which leads to a clear need of using many different metrics at once to obtain a clear picture. Many scientists propose some limited overview of statistical tools, although there is still no known work proposing a systematic analysis of performance of all baseline models over datasets of different configuration. The performance of competing models can be studied in terms of the quality of data adjustments, the predictive capacity, the quality of the economic and behavioural indicators derived from estimates, as well as in terms of algorithmic efficiency and resource usage. None of the known to us articles manages all these aspects into their benchmarks, limiting their studies only with several performance criteria.

Another problem is that the overall framework of performance evaluation is not completely established in literature. For example, there is no consensus on which data should be used (Japkowicz and Shah 2011). Many of the existing studies explore the impacts of different specifications on the same observed choice situation, among them we may name Fiebig et al. (2010), McCausland and Marley (2013) or Bouscasse, Joly, and Peyhardi (2019). Other, like Munizaga and Alvarez-Daziano (2005), Chorus (2010) or Pan, Rasouli, and Timmermans (2016) appear to use simulation to control their data and model accuracy.

Finally, the vocabulary and terminology differences existing among the scientific communities prevent the consensus to be reached. Even the *model* term definition itself is rather ambiguous and is usually determined application-wise. Given the economic context, the decision making modelling is directly intertwined to the process of decision making, various theoretical aspects specific to behavioural studies have to be considered. For example, the economic decision theory derives mostly from the random utility theory (RUM) of McFadden (2001). This decision theory was challenged by alternative visions such as random regret minimisation theory (RRM) of Chorus (2010), or the relative advantage maximisation theory (RAM) of Leong and Hensher (2015), or even quantum decision theory (QDT) of Yukalov and Sornette (2017), which offers a wide range of tools for modelling under uncertainty.

Consequently, there exists a particular interest to provide a framework which will allow to select appropriate models given the application and desired outcomes in discrete choice economic studies. In order to achieve this goal it appears sensible to make the focus at least on the state of art discrete choice models (Agresti (2013), Agresti (2007)) as well as their counterparts used in machine learning (Hastie, Tibshirani, and Friedman (2009), Kotsiantis, Zaharakis, and Pintelas (2006)), alimented with the emerging cross-field studies (Molina and Garip 2019).

Consequently, there exists a clear need to propose a toolset facilitating the decision making in the DCM choice, configuration and tuning for the applied practitioners. The work should take into account many different aspects at once, primary focusing on: (1) exploration of the performance metrics and their use cases depending on the models' architectures and scientific objectives; (2) proposition of a common analysis grid suitable for model comparison regardless its specification and background; (3) breaching the interdisciplinary vocabulary gap among econometricians, behavioural scientists and ML community.

2.1 Objectives

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. A work that uses unified knowledge from both disciplines might be highly beneficial for researchers and provide support for future applied studies. Following the logic of Athey (2018) and Mullainathan and Spiess (2017) the project will attempt to merge the essentials of ML and Econometrics paradigms, retaining their key concepts, to propose a toolset for model configuration and selection decision making.

First of all, as an interdisciplinary research project, the PhD work aims to define a common analysis and a comparison grid of performances of different models issued from econometrics and ML in application to individual decision making process. Definition of the performance criteria will have to take into account the specificity of the research objectives (data exploration, prediction, explanatory power, operational effectiveness, etc) and constraints of the two scientific domains.

A second objective is to incorporate the recent developments in the decision making theory into the general model comparison framework. This is a demanding procedure, that will require a thorough study of the discrete choice models in the context of consumer behaviour exploration, taking into account the underlying theoretical decision making theories. An exploration of the eventual changes in models' performances depending on the underlying theory is of particular interest, taking into account the trends in the recommender systems (Abbassi et al. 2009), preference learning (Pigozzi, Tsoukiàs, and Viappiani (2016), Domshlak et al. (2011)) and econometrics choice modelling (Bouscasse, Joly, and Peyhardi (2019)).

A third objective is to evaluate performance of the different modelling techniques in presence of specific behavioural bias in applied fields. This additional step distinguishes the PhD work from a simple statistical study and switches its focus to behavioural economics domain. It is assumed, that behavioural assumptions have a direct impact on the observed performance in DCM modelling, following the results obtained by Chorus (2012) or Pan, Rasouli, and Timmermans (2016).

2.2 Expected contribution

This work is expected to contribute to the interdisciplinary unification of very different domains, such as Econometrics and Machine Learning in the context of decision making modelling. Crossing econometric simulation tools and economic experimentation will rely on the definition of a unified framework to test and validate behavioural assumption. A tool to design simultaneously a choice experiment and its associated simulated choice could benefit the research community to ease and improve the bridge between experimentation method and the corresponding data analysis method.

In terms of immediate results the PhD thesis will produce a methodological support, describing and putting in relations the different models specific to the field of application. The valuable insights into model performance in the context of discrete individual choice, as well as underlying behavioural theories influences on the results are expected to be obtained. A detailed study will allow for future researchers to choose the appropriate method for a specific application in relation with cognitive and behavioural studies.

2.3 Starting point

The doctoral work proposes a theory-testing framework (Figure 1) exploring the performances of different econometric and ML models, while having the control over the data generation procedure and behavioural assumptions. The possibilities offered by such framework have already been demonstrated in the working paper presented during the DA2PL[1] conference.

3 Advances

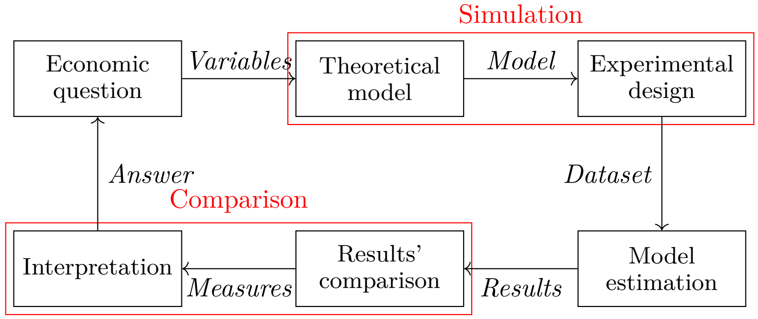
This year several advances have been achieved in the frame of the doctoral work. In this section we will focus on these breakthroughts, reviewing them one by one.

3.1 Understanding of the DCM history

Time has been spent on the literature review and systematisation of the knowledge on the DCM history. Many disciplines and application fields have different references and different reference points as it was observed during the project presentation to the *Preference Studies* community (this occurred during the DA2PL conference).

[1]"From Multiple Criteria Decision Aid to Preference Learning", Trento, Italy, 5-6 november 2020

Figure 1: Proposed framework



In this research we have gone back to the roots of the Logistic regression development and the early works of Bradley and Terry (1952) and Luce (1957). The further development of the concurrent models and the considerations about individual choice modelling was traced through further works of Plackett (1975) and Luce (1977), as well as the more typical and well known in econometrics community works of McFadden (1974) and McFadden (1976).

3.2 Model Taxonomy considerations

To simplify the research procedure associated with discrete choice modelling (DCM) in experimental economics new organisation of the DCM's taxonomy is being devised. The new organisation attempts to eliminate a rather common flaw encountered in contemporary hierarchical taxonomies, which is dictated by their case-specificity.

Many of the existing handbooks and articles in DCM field may be roughly separated into two different types: (1) those proposing an historical view on the modelling techniques development (Hastie, Tibshirani, and Friedman (2009) or Agresti (2013)) and (2) those that offer some sort of taxonomy based on subjective criteria (Kotsiantis, Zaharakis, and Pintelas (2006), Ayodele (2010), Molina and Garip (2019)). Both of the approaches are far from optimal when it comes to making a decision about the model to apply in a particular case. On the one hand, the methodological historical approach lacs sometimes the versatility and makes it difficult for the reader to decide among unrelated model types. On the other hand, a subjective taxonomy may be more instructive, because it offers a clear logical path (or rather a deprogrammed) that relates the various models. However, the subjectivity of any given taxonomy renders the model choice procedure conditional on the authors principles and vision.

We propose an alternative solution, which aims to overcome all the disadvantages of the previously described one. The main idea in this case is based on the recent understanding of ML models, such as NN, which can be found in Wang et al. (2020) or Welleck et al. (2017). In these recent works we discover the *modular* design of the ML models based on the studies of the loss functions. To be more precise, we may assume that every single model comprises a number of attributes, specifications and transforms, that may be combined and recombined as "*modules*" to obtain new models. One of the main disadvantages of the proposed method is the complexity of the resulting structure, which may be inaccessible for the user. The completeness of the resulting product, which should adopt the format of a database, goes without saying, but for someone who is unfamiliar with some of the advanced ML techniques it may be difficult to understand the implications. Consequently such database should inevitably incorporate some elements which will facilitate the comprehension of the *modules'* choice for the researcher.

3.3 Agent Based Simulation (ABS) toolset

To facilitate the simulation procedure it was decided to adopt the Agent Based Simulation (Macal and North 2014) approach to data generation procedure (Chan, Son, and Macal 2010): each agent (member) of the artificial population is associated with a number of choice situation (choice sets) and a decision rule. ABS may be understood as a simulation procedure governed by the exact simulation of multiple agents' behaviour. This approach was created to model "systems comprised of individual, autonomous, interacting agents" (Macal and North 2014), facilitating the modelling of individual behaviour.

A new simulation tool was developed during the spring semester of 2021 by an intern Antoine Dubois, as a result of reconsideration and reorganisation of the previous work, made by Amirreza Talebijamalabad. This tool was developed bearing in mind the ABS paradigm, combined with Object Oriented Programming (OOP) approach. Such tool offers much more flexibility and extensibility compared to previous, rather limited implementations.

3.4 Monte-Carlo implementation

In our research, aiming to explore the advantages and limitations of different DCM models, we rely on the various performance metrics to achieve the goal. The statistical properties of the target metrics are easily obtained through the repeated simulation using the Monte-Carlo method. Usually the model performance metrics are unique values devised to synthesise the information about how good the given statistical model performs in a given scenario. When facing a particular applied problem it is rather complex to provide some generalised performance assessment and compare the models based on a single-valued metrics. Such metrics allow only to compare the models in particular explored case taking into account the limited data available. However, MC technique (dataset reconstitution given the input distributions) as well as Boosting method (one of the possible resampling solutions from the original dataset) both allow to explore statistical properties of the performance metrics in the given context.

This approach offers us the possibility to provide an extensive analysis of the model performances, depending on the varying factors such as population size and/or number of choice sets observed per individual. After deriving the statistical properties of the target metrics it will be possible to inspire more external validity into our conclusions. However, such analysis is very demanding in terms of computation resources. In order to overcome this problem our team works currently with the GRICAD infrastructure[2].

4 Supporting activities

The presentation of the current PhD work cannot bypass the numerous associated activities, among which: (1) various academic courses, conferences and schools; as well as (2) teaching and improvement of the interpersonal skills. This report will not include all of the above activities, but only the most important ones[3].

4.1 Studies

As it can be traced from the main advances of the first year, the studies may be identically divided into several categories. The main focus in the choice of courses for this year was attributed to the development of skills and competences serving to assure the correct development and accomplished of the PhD.

Effectively, first of all come the technical studies aiming at the development of technical skills, as well as many other supporting competences. Among them we may distinguish: (1) scientific calculation and informatics in application to science; (2) data management; (3) workflow organisation.

Into another group of studies we can unite all the academic lectures, which are related to the main objective of this work: (1) providing knowledge on the statistical learning and other ML related techniques; (2) offering the knowledge in behavioural or economic related disciplines. The courses in this group were not among the main objectives for the current year, but rather some opportunistic choices, which will inevitably be useful in future.

4.2 Teaching

In order to support the possibility of the continuation in the research field after the PhD end, as well as to become accustomed to public presentation, it has been decided to incorporate some amount of teaching experience into the PhD curriculum. The main requirements for these activities have been: (1) a close relation to one of the spheres involved in the PhD work or (2) the possibility to impact the work directly.

Both of the intended needs have been successfully satisfied. The first one through the participation in teaching of *Statistics* at Grenoble-INP, while the second objective was attained through management of several Works of Study and Research (Travaux d'Etude et Recherche, TER) projects.

[2]"*Infrastructure de Calcul Intensif et de Données*", 38400, Saint Martin d'Hères, France

[3]For a complete list of the mentioned activities, please refer to the CSI form.

References

Abbassi, Zeinab, Sihem Amer-Yahia, Laks V. S. Lakshmanan, Sergei Vassilvitskii, and Cong Yu. 2009. "Getting Recommender Systems to Think Outside the Box." In *Proceedings of the Third Acm Conference on Recommender Systems*, 285–88. RecSys '09. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/1639714.1639769.

Agresti, Alan. 2007. *An Introduction to Categorical Data Analysis, Second Edition.*

———. 2013. *Categorical Data Analysis, Third Edition.*

Athey, Susan. 2018. "The Impact of Machine Learning on Economics." Book. In *The Economics of Artificial Intelligence: An Agenda*, by Ajay Agrawal, Joshua Gans, and Avi Goldfarb, 507–47. National Bureau of Economic Research; University of Chicago Press. https://doi.org/https://doi.org/10.7208/chicago/9780226613475.001.0001.

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.

Ayodele, Taiwo Oladipupo. 2010. "Types of Machine Learning Algorithms." *New Advances in Machine Learning*, 19–48.

Bouscasse, Hélène, Iragaël Joly, and Jean Peyhardi. 2019. "A new family of qualitative choice models: An application of reference models to travel mode choice." *Transportation Research Part B: Methodological* 121 (C): 74–91. https://doi.org/10.1016/j.trb.2018.12.010.

Bradley, Ralph Allan, and Milton E. Terry. 1952. "Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons." *Biometrika* 39 (3/4): 324–45. http://www.jstor.org/stable/2334029.

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

Chan, Wai Kin Victor, Young-Jun Son, and Charles M Macal. 2010. "Agent-Based Simulation Tutorial-Simulation of Emergent Behavior and Differences Between Agent-Based Simulation and Discrete-Event Simulation." In *Proceedings of the 2010 Winter Simulation Conference*, 135–50. IEEE.

Chorus, Caspar. 2012. "Random Regret Minimization: An Overview of Model Properties and Empirical Evidence." *Transport Reviews* 32 (1): 75–92. https://doi.org/10.1080/01441647.2011.609947.

Chorus, Caspar G. 2010. "A New Model of Random Regret Minimization." *European Journal of Transport and Infrastructure Research* 10 (2).

Domshlak, Carmel, Eyke Hüllermeier, Souhila Kaci, and Henri Prade. 2011. "Preferences in Ai: An Overview." *Artificial Intelligence* 175 (7): 1037–52. https://doi.org/https://doi.org/10.1016/j.artint.2011.03.004.

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

Fiebig, Denzil, Michael Keane, Jordan Louviere, and Nada Wasi. 2010. "The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity." *Marketing Science* 29 (3): 393–421. https://EconPapers.repec.org/RePEc:inm:ormksc:v:29:y:2010:i:3:p:393-421.

Flach, Peter. 2019. "Performance Evaluation in Machine Learning: The Good, the Bad, the Ugly, and the Way Forward." In *Proceedings of the Aaai Conference on Artificial Intelligence*, 33:9808–14.

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

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

Kotsiantis, Sotiris, I. Zaharakis, and P. Pintelas. 2006. "Machine Learning: A Review of Classification and Combining Techniques." *Artificial Intelligence Review* 26 (November): 159–90. https://doi.org/10.1007/s10462-007-9052-3.

Leong, Waiyan, and David A. Hensher. 2015. "Contrasts of Relative Advantage Maximisation with Random Utility Maximisation and Regret Minimisation." *Journal of Transport Economics and Policy (JTEP)* 49 (1): 167–86. https://www.ingentaconnect.com/content/lse/jtep/2015/00000049/00000001/art00010.

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

———. 1977. "The Choice Axiom After Twenty Years." *Journal of Mathematical Psychology* 15 (3): 215–33.

Macal, Charles, and Michael North. 2014. "Introductory Tutorial: Agent-Based Modeling and Simulation." In *Proceedings of the Winter Simulation Conference 2014*, 6–20. IEEE.

McCausland, William J., and A. A. J. Marley. 2013. "Prior Distributions for Random Choice Structures." *Journal of Mathematical Psychology* 57 (3): 78–93. https://doi.org/https://doi.org/10.1016/j.jmp.2013.05.001.

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

———. 2001. "Economic Choices." *The American Economic Review* 91 (3): 351–78. http://www.jstor.org/stable/2677869.

McFadden, Daniel L. 1976. "Quantal Choice Analaysis: A Survey." In *Annals of Economic and Social Measurement, Volume 5, Number 4*, 363–90. NBER.

Molina, Mario, and Filiz Garip. 2019. "Machine Learning for Sociology." *Annual Review of Sociology* 45 (1): 27–45. https://doi.org/10.1146/annurev-soc-073117-041106.

Mullainathan, Sendhil, and Jann Spiess. 2017. "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives* 31 (2): 87–106. https://doi.org/10.1257/jep.31.2.87.

Munizaga, Marcela A., and Ricardo Alvarez-Daziano. 2005. "Testing Mixed Logit and Probit Models by Simulation." *Transportation Research Record* 1921 (1): 53–62. https://doi.org/10.1177/0361198105192100107.

Pan, X, S Rasouli, and HJP Timmermans. 2016. "Modeling Bounded Rationality in Choice Behavior: Relative Utility Vs. Random Regret Models." In *13th International Conference on Design and Decision Support Systems in Architecture and Urban Planning, Ddss 2016.*

Pigozzi, Gabriella, Alexis Tsoukiàs, and Paolo Viappiani. 2016. "Preferences in Artificial Intelligence." *Annals of Mathematics and Artificial Intelligence* 77 (3-4): 361–401. https://doi.org/10.1007/s10472-015-9475-5.

Plackett, Robin L. 1975. "The Analysis of Permutations." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 24 (2): 193–202.

Varian, Hal R. 2014. "Big Data: New Tricks for Econometrics." *Journal of Economic Perspectives* 28 (2): 3–28. https://doi.org/10.1257/jep.28.2.3.

Wang, Qi, Yue Ma, Kun Zhao, and Yingjie Tian. 2020. "A Comprehensive Survey of Loss Functions in Machine Learning." *Annals of Data Science*, 1–26.

Welleck, Sean, Zixin Yao, Yu Gai, Jialin Mao, Zheng Zhang, and Kyunghyun Cho. 2017. "Loss Functions for Multiset Prediction." *arXiv Preprint arXiv:1711.05246*.

Yukalov, Vyacheslav I, and Didier Sornette. 2017. "Quantum Probabilities as Behavioral Probabilities." *Entropy* 19 (3): 112.

Zielesny, Achim. 2011. *From Curve Fitting to Machine Learning*. Vol. 18. Springer.