Econometrics and Machine Learning

Kennedy Odongo

1/19/2021

Why has the adoption of ML methods been much slower in Econometrics? Economics journals emphasize the use of methods with formal properties of a type that many of the ML methods do not naturally deliver.The focus in the econometric literature is usually on large sample properties of estimators and tests for example consistency, Normality and efficiency. On the other hand the ML literature focuses on the on the working properties of algorithms.Breiman [2001b],Hastie et al. [2009] and Efron and Hastie [2016],Chamberlain [2000],(Matzkin [1994, 2007]),Varian [2014]

ML techniques often require careful tuning and adaptation to effectively address the specific problems that economists may have: These adaptations include:

- Exploiting the structure of problems: the causal nature of many estimands, the endogeneity of variables, the configuration of data such as panel data, the nature of discrete choice among a set of substitutable products, or the presence of credible restrictions motivated by economic theory, such as monotonicity of demand in prices or other shape restrictions.
- Changing the optimization criteria of machine learning algorithms to prioritize econometric considerations.
- Using techniques such as sample splitting can be used to optimize the performance of machine learning estimators.

ML to Econ

- non-parametric regression-supervised learning for regression problems.
- supervised learning for classification problems-non-parametric regression for discrete response models.
- Unsupervised learning, or clustering analysis and density estimation.
- estimates of heterogeneous treatment effects and optimal policies mapping from individuals' observed characteristics to treatments.
- ML approaches to experimental design.
- Matrix completion problems, including its application to causal panel data models and problems of consumer choice among a discrete set of products.
- Analysis of text data.

Econometrics and Machine Learning:Goals, Methods and Settings

Goals:Wu et al. [2008],Terminology,Validation and Cross validation, **Overfitting,regularization and Tuning Parameters**:(Vapnik[1998]),Morris[1983] **Sparsity**:Hastie et al. [2009, 2015],Belloni et al. [2014]). **Computational Issues and Scalability** :(Bertsimas et al. [2016]),Hastie et al. [2017]),Ruiz et al. [2017]),Friedman [2002], Bottou [1998, 2012]).**Ensemble methods and model averaging, Inference** Wu et al. [2008],Dietterich [2000]),(Bell and Koren [2007]).

The focus in econometric is on the quality of the estimators of the target, traditionally measured through large sample efficiency. In ML however, the literature generally focuses on developing algorithms.

estimating vs. training the model. Regressors, covariates vs. features, regression parameters vs. weights. Unordered discrete response problems are generally called classification problems.

First, the goal is predictive power, rather than estimation of a particular structural or causal parameter.Second, the method uses out-of-sample comparisons, rather than in-sample goodness-of-fit measures.

Supervised Learning for Regression Problems:Mullainathan and Spiess [2017]),Bierens [1987],Chen [2007],(White [1992], Hornik et al. [1989]),Friedberg et al.[2018] **Regularized Linear regression-Lasso**:(Tibshirani [1996]).(Hoerl and Kennard [1970]).,(Miller[2002], Bertsimas et al. [2016]),(Zou and Hastie [2005]),(Efron et al. [2004]),(Meinshausen [2007]),(Efron et al. [2004]),(Candes and Tao [2007]),(Breiman [1993]),Hastie et al.[2017]. **Ridge and elastic nets**:,**Regression Trees and Forests**:(Breiman [2001a]),(Breiman et al. [1984]),Athey and Imbens [2016],(Breiman [1996]),Wager and Athey [2017],Athey et al. [2017d],Tibshirani and Hastie [1987],(Imbens and Lemieux [2008]),(Abadie and Imbens [2011]),Friedberg et al. [2018] **Deep learning and Neural Nets**:Farrell et al.[2018],LeCun et al. [2015],(Krizhevsky et al. [2012]),(Friedman [2002], Bottou [1998, 2012]),(Rumelhart et al. [1986]),**Boosting**:Schapire and Freund [2012].

Supervised Learning for Classification problems: Cortes and Vapnik [1995], (Krizhevsky et al. [2012]) **Classification Trees and Forests**:Breiman et al. [1984], **Support Vector Machines** :Scholkopf and Smola [2001]

Unsupervised Learning: **K-means clustering**:(Hartigan and Wong [1979]), Alpaydin [2009]). , **Generative adversarial Networks**:Arjovsky and Bottou [2017], Goodfellow et al. [2014]).

Machine learning and Causal inference: Abadie and Cattaneo [2018] and Imbens and Wooldridge [2009] **Average Treatment Effects**:(Rosenbaum and Rubin [1983], Imbens and Rubin [2015],Belloni et al. [2014],(Robins and Rotnitzky [1995], Chernozhukov et al. [2016a,b]),(Athey et al. [2016a]),(Zubizarreta [2015]),Athey [2017,2018], **Orthogonolization and Cross-Fitting**:Bickel et al. [1998] and Van der Vaart [2000], Chernozhukov et al. [2017],Chernozhukov et al. [2018a,c],Athey et al. [2017d] **Heterogenous treatment effects**:Athey and Imbens [2017b],(Holland [1986]),Athey and Imbens [2016],[Zeileis et al., 2008],Wager and Athey [2017],„Chernozhukov et al. [2018b],Friedberg et al. [2018],Hartford et al. [2016],[van der Laan and Rubin, 2006],Imai et al. [2013],Chipman et al. [2010],Hill [2011], Green and Kern [2012],Hirano and Porter [2009],Kitagawa and Tetenov [2015],Strehl et al.[2010], Dudik et al. [2011], Li et al. [2012], Dudik et al. [2014], Li et al. [2014], Swaminathan and Joachims [2015], Jiang and Li [2016], Thomas and Brunskill [2016], Kallus [2017]. Zhou et al. [2018],Athey and Wager [2017].

Experimental Design,Reinforcement Learning and Multi-armed Bandits:(Scott [2010], Thompson [1933]),Lai and Robbins [1985],(Sutton et al.[1998]) **A/B Testing Versus Multi-Armed Bandits**:Athey and Imbens [2017a], **Contextual bandits**:Bastani and Bayati [2015],Dimakopoulou et al. [2017],[2018]

Matrix Completion and Recommender systems:**The Netflix Problem**(Bennett et al. [2007]),**Matrix completion Methods for Panel Data**:(Candes and Recht [2009], Mazumder et al. [2010]), Athey et al. [2017a]), **The Econometrics Literature on Panel data and synthetic Control methods**:(Bai and Ng [2017, 2002], Bai [2003],Abadie et al. [2010, 2015],Doudchenko and Imbens [2016],Abadie et al. [2015],Athey et al. [2017a), **Demand Estimation in Panel Data**(Keane et al. (2013), Jacobs et al. [2014],Gopalan et al. [2015]),Ruiz et al.(2017),Athey et al.[2017b],Wan et al. [2017],Semenova et al. [2018])

Text Analysis Gentzkow et.al (2017),Blei and Lafferty (2009), Mnih and Hinton(2007), Mnih and Teh (2012),Mikolov et al. 2103 a,b,c, Mnih and Kavukcuoglu(2013), Pennington et al.2014, Levy and Goldberg, 2014, Vilnis and McCallum, 2015, Arora et al. 2016, Barkan 2016, Bamler and Mandt,2017, Bengio et al.,2003,2006, Harris 1954, Firth 1957