JAISCR, 2022, Vol. 12, No. 2, pp. 79 – 100
10.2478/jaiscr-2022-0006

MACHINE LEARNING AND TRADITIONAL ECONOMETRIC MODELS: A SYSTEMATIC MAPPING STUDY

María E. Pérez-Pons[1,*], Javier Parra-Dominguez[1,2], Sigeru Omatu[3], Enrique Herrera-Viedma[4], Juan Manuel Corchado[1,2,5,6]

[1] *University of Salamanca, BISITE Research Group Edificio I+D+i, Calle Espejo 2, 37007, Salamanca, Spain*

[2] *Air Institute, IoT Digital Innovation Hub, 37188 Salamanca, Spain*

[3] *Hiroshima University, Digital Manufacturing Education and Research Center Division of Data Driven Smart System 3-10-31 Kagamiyama East-Hiroshima, 739-0046, Japan*

[4] *University of Granada Colegio Máximo de Cartuja, Campus Universitario de Cartuja C.P. 18071 Granada, Spain*

[5] *Department of Electronics, Information and Communication, Faculty of Engineering, Osaka Institute of Technology, 535-8585 Osaka, Japan*

[6] *Pusat Komputeran dan Informatik, Universiti Malaysia Kelantan, Karung Berkunci 36, Pengkaan Chepa, 16100 Kota Bharu, Kelantan, Malaysia.*

[*] *E-mail: eugenia.perez@usal.es*

Submitted: 14th July 2021; Accepted: 15th September 2021

Abstract

Context: Machine Learning (ML) is a disruptive concept that has given rise to and generated interest in different applications in many fields of study. The purpose of Machine Learning is to solve real-life problems by automatically learning and improving from experience without being explicitly programmed for a specific problem, but for a generic type of problem. This article approaches the different applications of ML in a series of econometric methods. *Objective*: The objective of this research is to identify the latest applications and do a comparative study of the performance of econometric and ML models. The study aimed to find empirical evidence for the performance of ML algorithms being superior to traditional econometric models. The Methodology of systematic mapping of literature has been followed to carry out this research, according to the guidelines established by [39], and [58] that facilitate the identification of studies published about this subject. *Results*: The results show, that in most cases ML outperforms econometric models, while in other cases the best performance has been achieved by combining traditional methods and ML applications. *Conclusion*: inclusion and exclusions criteria have been applied and 52 articles closely related articles have been reviewed. The conclusion drawn from this research is that it is a field that is growing, which is something that is well known nowadays and that there is no certainty as to the performance of ML being always superior to that of econometric models.

Keywords: machine learning, econometric models, regression, prediction

1 Introduction

The concept of "Econometry", was first introduced by Pawel Ciompia at the beginning of 20th century. After him, Jan Tinbergen, was one of the first researchers to apply mathematics in the testing of economic hypotheses [71]. Econometrics combine elements of economics, mathematics and statistics. The statistical methods used in econometric models are specifically targeted at this domain and therefore they are non-applicable to other statistical fields [28]. The purpose of machine learning is to solve real-life problems by automatically learning and improving from experience without being explicitly programmed for a specific problem, but for a generic type of problem. Thanks to global economic interdependence, nowadays much more information is available for predictions. These large amounts of data require different types of methods for optimal information processing. As shown in Figure 1, this research comprehends the nature and the growing interdependence of economics, mathematics and computer science.

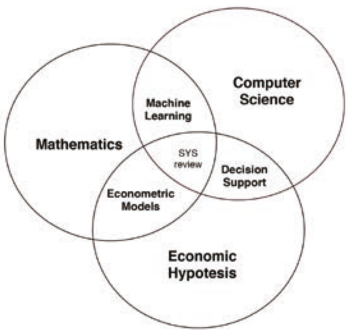


Figure 1. Fields that converge in that SMS.

Therefore, traditional economic models and applications require tools that have a greater computing capacity and new forecasting methodologies that will provide more accurate predictions. In the field of prediction, the ML algorithms based on a previously identified label are classified as supervised ML algorithms [42]. Supervised algorithms look for functions that predict well outside the sample. Economists would denominate this as the dependent variable; the one that will change depending on the changes in other variables. If the labeled attribute does not exist, an unsupervised algorithm is necessary for data exploration rather than the prediction of results. For instance, one could try to predict the value "y" of a company from its observed characteristics "x". Within the field of supervised algorithms there are many different applications that can be employed [12], either for classification or regression. Let's suppose we set out to measure whether having a gym in the workplace would improve the employees' efficiency (where productivity was measured as projects completed per month in a 40 hours work week). Economists would seek for a logical experiment that might entice certain workers to use gym facilities for reasons unrelated to their current productivity (i.e., temporary gym at the workplace). We can estimate a model using a linear regression as shown in the Figure 1:

$$Y_i = \alpha + \beta_0 + \beta_1 X_i + \varepsilon_i \tag{1}$$

or a multiple linear regression as in equation 2:

$$Y_i = \alpha + \beta_0 + \sum \beta_p X_{pi} + \varepsilon_i \tag{2}$$

where Y_i is the outcome (the productivity of the individuals within a year), X_i is the policy of interest (in case the worker has gone to the gym), β is the key parameter of interest (the effect of going to the gym within the working hours), α denotes the other parameters, and ε_i is an error term. Using the same data, a ML approach would involve identifying the variables that are strictly associated with productivity. This is due to the wide range of potential indicators in the data, and the likelihood of building a model that would predict the profitability well, either inside or outside the sample data.

ML models have the capacity to overcome those econometric regression models according to different prediction rules that underlie the systems [48]. The possibilities offered by ML have led several researchers in financial econometrics to carry out comparative studies on the performance of ML applications against traditional models as suggested [45]. However, the two approaches have not always been in conflict with each other. For instance, if just one subset of control variables is predictive, a ML model selection approach could help target the most relevant one. Data mining methods can also be helpful if there are significant interaction effects, so one is focused on predicting effects for certain individuals instead of an overall impact for the whole population [7].

This research work analyzes the state of the art and identifies ML studies that are oriented to econometric models. It is developed under a methodology for the systematic mapping of literature, following the guidelines established by [57, 58], and [41] for carrying out *Systematic Mapping Studies* (SMS). Thanks to this type of SMS it is possible to categorize existing studies in a specific area of research through an outline and structure that indicate how often investigations are conducted in that area. The results of the study are represented visually on a map [51].

The remainder of this study is organized as follows: Section 2 analyzes the state-of-the-art reviews and the surveys performed by other authors in the fields of econometrics and ML. Then, Section 3 details the stages of the systematic mapping study methodology proposed by [58]. In Section 4 the results of the report stage of the conducted systematic mapping study are outlined. In Section 5 results are analyzed in detail. Finally, the conclusions from the conducted research are described in Section 6.

2 Related Work

In recent years, many authors have presented reviews of ML applications in econometric fields. [48], and [3] published theoretical results that emphasised the range of prediction possibilities offered by ML in the field of economics. Since econometrics includes many applications some authors have narrowed down the scope of their studies. For instance, [4] highlights the main machine learning models that are applied in econometric models, including methods based on regression, classification, unsupervised learning methods and matrix completion methods, other authors explore concrete areas of machine learning applications in econometrics, such as methods for demand estimation [6]. Moreover, others conducted an in-depth examination of the major machine learning models for time series forecasting [1, 32] made comparative studies and conclusions derived from traditional and machine learning applications in financial market forecasting. Other comparative studies that have pointed to the good performance of machine learning worked in detecting irregular patterns and making short term forecasts using heterogeneous data [44]. According to [58], SMS studies aim to identify principal studies in the area of interest, more extensive definition of what a SMS is will be given in the following Section 3. Since there are many articles in this field, and some of them mentioned in the previous paragraph, a SMS of the published articles has been carried out. SMS is a methodology that intends to standardize and to objectively review different topics. ML algorithms and their applications are a growing field of study [67, 68] which has many effects in the worldwide economy [23] and therefore, the authors find it interesting to combine all the information with the different applications that have been made, especially at the level of regressions. The authors have found it interesting to create a word cloud with all the titles collected from the searches in order to see quickly which are the topics on which the articles related to this topic mainly focus. The areas covered by the real-world ML applications in traditional econometric models are described in Figure 2.

Figure 2. Word cloud from the titles of all the articles collected in the review.

3 Research Methodology

To complement the definition of SMS given in the previous section, a SMS is defined as a secondary study aimed at identifying if there are scientific evidence and research contributions in a field or specific research area. SMS allows researchers to get an overview of the amount of scientific evidence on the subject under investigation. Several authors have defined procedures or guidelines for conducting a SMS [57, 58, 40]. Some of the main benefits of a SMS are described as follows:

– Uses an objective and validated procedure to identifying papers or relevant information in a specific research field.

– Helps identify future lines of research arising from the information gaps identified in the mapping process.

– The emergence of new research that can be used in future systematic literature reviews.

In the present article, the SMS methodology and phases have been followed as proposed in the review of [58], and [39]. The process has been organized into three phases: planning, development and mapping report. The schema of those three phases is presented in Figure 3 and will be described in detail in the subsections that follow. All the papers included in the review are listed in Table 9.

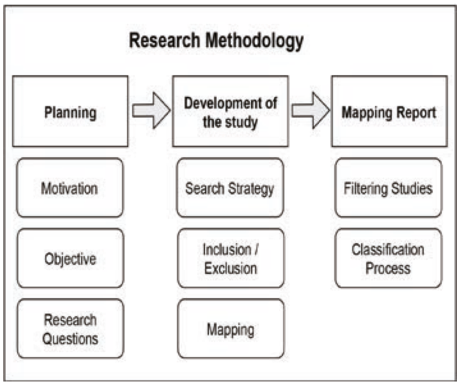


Figure 3. SMS methodology phases that will be described in detail in following subsections.

3.1 Planning

For the planning of our SMS, the following three activities have been considered: Motivation, Objective and Research Questions. For the planning of the activities, the works of [39, 58] have been used as reference. Those activities are described in more detail below;

3.1.1 Motivation

The evolution and adoption of machine learning in the field of traditional econometric models motivated this SMS. The main motivation of this paper is the review of existing papers on the application of machine learning techniques in traditional econometric models, as well as the comparison of their results. The identification of future lines of research in this area is also the motivation behind this study.

3.1.2 Objective

In recent years, it has been possible to develop many models applied to economic problems. All this thanks to the rapid growth of machine learning applications which has been driven by the high demand for real-time applications in real-world scenarios. The objective of this research is to identify the latest applications and do a comparative study of the performance of econometric and machine learning models. The study aimed to find empirical evidence if machine learning algorithms obtain better results than econometric models when compared to the same problem, or on the contrary, in this field so far there is evidence that when machine learning and econometric models are used together is it possible to achieve better results.

3.1.3 Research Questions

Defining research questions helps find the evidence required for the study. These questions will allow to categorize the literature published until June 22, 2020 and to present visual maps in the results section. This SMS is based on the following research questions:

– **RQ1** - How and in which fields machine learning has been implemented as econometric model applications?

– **RQ2** - How does supervised machine learning complement traditional econometric models?

– **RQ3** - Comparing machine learning and econometric models what are the most frequently applied methods, and in what study context?

3.2 Development of the Study

This Section describes the process followed to address the SMS. First the search strategy has been defined to obtain relevant papers. The inclusion / exclusion criteria have been numbered and used to select the studies to be mapped. In this case, a large number of articles were identified as noise and the process has been conducted individually as suggested by the author in [59].

3.2.1 Search Strategy

To conduct a SMS, it is important to build a search strategy and define a search string, even though some authors have concluded that it not possible to prevent bias in systematic reviews [74]. To minimize the possible bias in research, the authors have used PICO (Population, Intervetion, Comparision and Outcomes) guidelines defined by [41, 40] and applied in [58]. PICO was recently referred to as a good set of practical guidelines for conducting SMS [20]. The authors of [59] proposed the PICOC model which has been developed within the PICO framework and its Context has been extended. Nevertheless, in this case, only the PICO framework has been used.

– Population: The identified papers.
– Intervention: The methods implementing machine learning and traditional econometrics.
– Comparison: The different types of results are compared.
– Outcomes: Complementarities and uses of econometric and machine learning models

After applying the PICO approach, our keywords are *econometric* and *ML*. To reduce the search and obtain more accurate results, the keywords considered in the second search were closely related to the research *Regression*, and *Supervised*. The selection of those databases done has been according to the volume of articles as well as the variety of publication topics, being among those selected those recommended as optimal for this type of analysis [10]. The search strings were built for the following databases: Web of Science, Scopus and Springer. The search was restricted to June 22, 2020. Three search strings were designed for each database to reduce the number of articles obtained in the preliminary results, those have been represented in Table 1. In the creation of the search strings, due to the fact that the second condition proved to be very restrictive, as it yielded very few articles, it was decided to extend it a little further by including the possibility of a new word. The number of articles per year found for each of the search engines are those described in the Table 2.

3.2.2 Inclusion/Exclusion Criteria

Regarding the inclusion and exclusion criterias, [58] pointed the importance of establishing the characteristics that the identified studies must meet in order to be included or excluded from a systematic mapping study. Using the [58] guide as a basis, the following inclusion and exclusion criteria were considered when selecting a paper. Those criterias are the ones that have been described in Table 3.

3.3 Mapping Report

The mapping report includes the filtering studies and classification process descriptions. During the filtering process, the relevance of 4 of the papers has been questioned and therefore have been removed from the overall study. To strengthen the criteria for the inclusion-exclusion of articles, the support of three researchers was requested for evaluation by experts. Moreover, this has allowed us to maintain neutrality and objectivity in the process of selection and rejection of articles related to the subject of the review. The full mapping report is shown in Table 9.

3.3.1 Filtering Studies

From the three search strings defined, a total of 356 related studies were found in the first step. To filter all the studies, the previously defined inclusion and exclusion criteria have been applied. The filtering process was conducted by the first author. The papers that have been selected met all the inclusion criteria and none of the exclusion criteria. The first author consulted the rest of the co-authors in cases where it was not clear whether an article should be included or excluded. The authors have requested support from other researchers to resolve disagreements between them when needed.

3.3.2 Classification Process

Once the filtering process has been completed, the selected articles were classified on a spreadsheet into the following categories: Author, Title, Year, Type of Publication, Machine learning vs econometric, Field of application, Research Type, Evolution Activity and contribution Type. The *Type of Publication* is divided into; Journal, Conference and Book Chapter. The *machine learning vs econometric*; describes the output in terms of the performance achieved by machine learning and the econometric models, it is stated if one performed better than the other or if they were used jointly. The

Table 1. Search string designed for database search. The articles have been extracted directly from each Search Engine.

Data Base	Search String
Web of Science (WS1)	((TS=(Econometric AND Machine Learning))) AND LANGUAGE: (English)
Web of Science (WS2)	((TS=(Econometric AND Machine Learning AND supervised))) AND LANGUAGE: (English)
Web of Science (WS3)	((TS=(Econometric AND machine AND learning AND regression OR supervised))) AND LANGUAGE: (English)
Scopus (SC1)	TITLE-ABS-KEY (”Econometric”) OR TITLE-ABS-KEY (”Machine Learning”)) AND TITLE-ABS-KEY (”*supervised learning*”) AND (TITLE-ABS-KEY (”Econometric*”) OR TITLE-ABS-KEY (”regression*”)) AND (LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016)) AND (LIMIT-TO (LANGUAGE , ”English”))
Scopus (SC2)	TITLE-ABS-KEY (”Econometric”) AND TITLE-ABS-KEY (”Machine Learning”)) AND TITLE-ABS-KEY (”supervised*”) AND (LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016)) AND (LIMIT-TO (LANGUAGE , ”English”))
Scopus (SC3)	TITLE-ABS-KEY (”Econometric”) AND TITLE-ABS-KEY (”Machine Learning”)) AND TITLE-ABS-KEY (”*economet*”) OR TITLE-ABS-KEY (”supervised*”) OR TITLE-ABS-KEY (”regression*”) AND (LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016)) AND (LIMIT-TO (LANGUAGE , ”English”))
Springer (SP1)	(”Econometric”) AND (”Machine Learning”)
Springer (SP2)	(”Econometric”) AND (”Machine Learning”) AND (”supervised”)
Springer (SP3)	(”Econometric”) AND (”Machine Learning”) AND (”supervised”) OR (”regression”)

Table 2. Number of articles per year under each search engine criteria. The number of articles in 2020 is only till June 22nd 2020.

	2016	**2017**	**2018**	**2019**	**2020**
SC1	89	129	212	382	162
SC2		2	1	2	
SC3	9	30	28	52	19
WS1	5	12	16	29	11
WS2				2	
WS3	3	6	5	15	4
SP1	129	147	227	315	355
SP2	28	26	46	80	81
SP3	15	17	35	55	63

Table 3. Inclusion and Exclusion Criterias

Inclusion	**Exclusion**
IC1: Peer-reviewed book chapters and papers published in journals or conferences	EC1: Duplicated papers
IC2: Studies published between 2016 to June 2020	EC2: Papers that are not related to machine learning and econometric applications
IC3: Studies in the field of machine learning applications in econometrics	EC3: The papers in which the authors have not identified an econometric and machine learning application (event joint or replaceable methodologies)

Field of application is the field to which the algorithm has been applied (ie: Stock market, Agriculture,...). In the *Research Type category*, studies are classified into: Evaluation research, Validation research, Solution proposal, Philosophical paper, Experience report, and Opinion paper. The *Evolution activity* allows for the classification of the articles determining the following categories; Validate, Implement or Analyse, which means if only an analysis had been carried out in the research or if it implemented and validated a proposed solution. The *Contribution type* consists of the main contribution of the article: Method, model, framework, or platform. There is a set of articles that can contain two or more research types, in those cases, a single record containing the categories is made to avoid the duplicity of the data. The classifications and categories are being presented according to the classification described by [39].

4 Mapping

For each of the databases: Springer, Scopus, and Web of Science, a first simple search string was defined using only machine learning and econometrics as keywords, and 2884 papers appeared in the search results. For this reason, the authors decided to extend the search string by increasing the number of keywords, with a second search string detailed in Table 1. The total number of article results was 268, nevertheless, the results were too strict and not covering the whole expected searches. Therefore, it was decided to expand the search string a little more by extending the concept. The application of the last search chain resulted in 356 papers which were reviewed and filtered by applying the inclusion and exclusion criteria.

That the papers were excluded if they did not meet the inclusion criteria established in the Methodology for the development of the SMS and that have been described in the 3.2.2. The final result is the one shown in Figure 4. As can be seen in the diagram 4, 51 articles were duplicated papers. Then there were 111 articles that contained machine learning and econometrics but not related to both files, and finally, 142 articles were removed because the aim of the study was to identify those papers that were comparing experiments implementing machine learning and econometric models. The final number of articles that have been selected to conduct the research has been 48.

4.1 General Analysis

Due to the variable and extensive nature of research, a graphical analysis of the fields in which this type of research is published is given in Figure 4.1. All the articles from the first search string have been taken so that the volume in terms of the distribution of the publications was wide. The scope of journals that contain most of the publications in the research field of econometric models and machine learning are the ones in computer science and economics. Nevertheless, there are also publications in mathematics and engineering journals, which can be considered essential fields of knowledge when conducting this type of research.

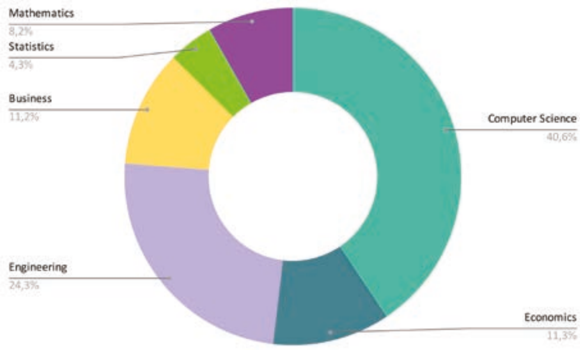


Figure 5. Fields of the journals where the articles have been published. The percentages correspond to the total sample on the first search String created.

5 Results

The results of the systematic mapping are shown in Figure 5. In Figure 5, three dimensions of the previous dimensions described in Section 3.3.2 where the classification process is described have been considered. Those dimensions are: Contribution type, ML vs econometrics, and Research type according to the guidelines for bubble plot graphs, defined by [57]. In the systematic mapping review, the validation of the different models has considered the most important element of the researches, as shown in the table. The different tables are evaluated in Table 4 under Method, 5 Model, and 6 Framework.

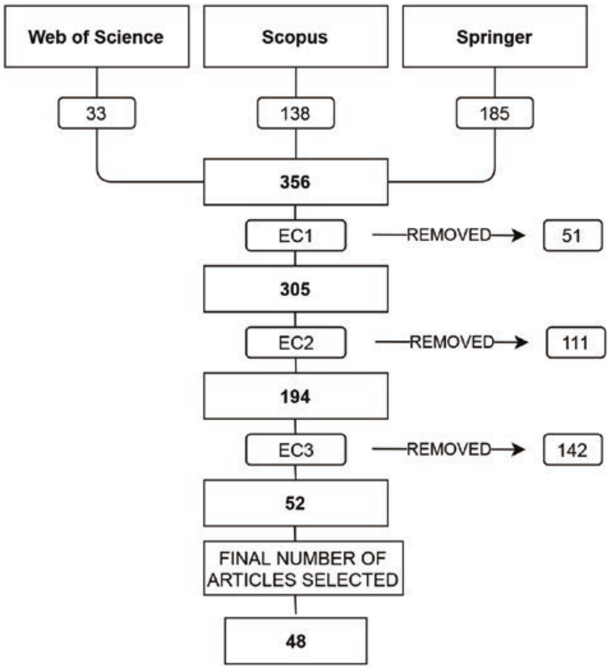


Figure 4. Exclusion Criteria applied to the total articles extraction

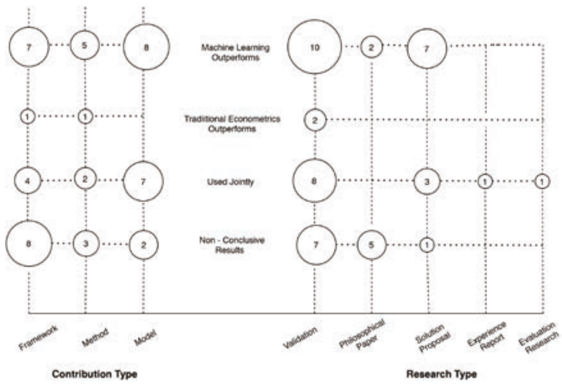


Figure 6. Mapping study

5.1 RQ1 - How and in what fields ML has been implemented as econometric model applications?

The results derived from RQ1 are shown as a stacked bar graph in the Figure 7. As seen in the graph, the articles have been classified into large activity groups into which the research articles could be categorized. A sector of activity has been kept as theoretical due to the fact that some articles were very theoretical and generic in terms of activity leading to an impossibility to adjust it to one sector of activity or another. As can be seen in the graph, the analysis of future performance is oriented to all industries, while the analysis of time series is mostly used in the stock market and investment applications. One of the issues where it is practically a common denominator in all cases is the forecast of future prices.

5.2 RQ2 - How does Supervised ML complement traditional econometric Models?

To answer this question, 4 different hypotheses have been presented as shown in Table 7. In Table 8 the results are shown regarding the differences in the methods, model and framework and in the cases where the machine learning techniques or econometric models outperform one another. In the majority of the cases where machine learning and traditional econometric models were compared, machine learning has performed better in making predictions. There are some situations where the highest accuracy is achieved by using the algorithms jointly, as presented in Table 8. This Table shows that in cases where a new model is proposed, the highest accuracy is achieved when the models are combined, indicating that machine learning models or econometric models do not necessarily work better than each other but that the synergy between both is usually the best option due to the nature of both models. This classification should be placed in the framework of considering that this SMS considers articles from 2016, in which machine learning was already an established field.

Table 4. Articles with Contribution type classified as "Method". ANN: Artificial Neural Network, ARIMA: Autoregressive integrated moving average, D-ML: Diverse ML Methods, LR: Linear Regression, MLR: Multinomial logistic regression, NN: Neural Network, RF: Random Forest, SARIMA: Seasonal Autoregressive Integrated Moving Average, SCL: Supervised Classification algorithms, SVR: Support Vector Regression, SVM: Support Vector Machine, TS: Time Series

Author	ML vs Eco	ML Algorithms	Econometric models	Topic	Output
[17]	Eco Outperforms	RF	MLR	Agriculture	Future Demand
[31]	ML Outperforms	RF/NN	ARIMA	Energy	Future Price
[29]	ML Outperforms	ANN	SARIMA	Energy	Future Price
[56]	ML outperforms	ML	Regression	Tourism	Future performance
[14]	ML Outperforms	ANN	LR	Agriculture	Future Performance
[62]	Used Jointly	ANN	ARIMA	Energy Price	Time Series
[29]	Non Results	ANN	ARIMA	Citizien	Future Performance
[38]	Used Jointly	SCL	LDA	Health	Future Performance
[5]	Non Results	SVR/NN	ARIMA	Theorical	Future Performance
[75]	Non Results	SVM	Theorical	Theorical	Future Performance
[36]	Non Results	NN	TS	Citizienship	Future Performance

Table 5. Articles with Contribution type classified as "Model". ARIMA: Autoregressive integrated moving average, DL: Deep Learning, DNN: Deconvolutional Neural Network, D-ML: Diverse ML Methods, FE -R: Fixed Effect Regression, GA: Genetic Algorithm, GARCH: Generalized AutoRegressive Conditional Heteroskedasticity, KNN: k-nearest neighbors algorithm, LASSO: east absolute shrinkage and selection operator, LR: Linear Regression, NN: Neural Network, SVR: Support Vector Regression, PCA: Principal Component Analysis, PCR: Put-call ratio, RF: Random Forest

Author	ML vs Eco	ML Algorithms	Econometric models	Topic	Output
[66]	ML Outperforms	GA	ARIMA	Citizien	Future Price
[21]	Used Jointly	SVR	ARIMA	Investment	Time Series
[49]	Used Jointly	RNN	GARCH	Agriculture	Future Performance
[33]	ML Outperforms	KNN,..	Regression	Agriculture	Future Demand
[25]	Non results	PCA	LR	Investment	Future Performance
[54]	Non Results	GA	HODRICK-PRESCOTT	Stock Market	Future Price
[60]	Used Jointly	SVR	LASSO	Transport	Future Demand
[26]	Used Jointly	RF	FE-R	Trade	Future Performance
[46]	Used Jointly	DL	Regression	Tourism	Future Performance
[53]	ML outperforms	SVR/KNN	PCR	Tourism	Future Price
[72]	ML outperfoms	AVM ML	Regression	investment	Theorical
[22]	ML outperfoms	ANN	Regression	Tourism	Future Demand
[2]	ML outperfoms	DNN	ARIMA	Investment	Future Price
[27]	Used Jointly	SVR	LASSO	Transport	Future price
[47]	ML outperforms	NN	ARIMA	Investment	Future Price
[69]	Used Jointly	D-ML	Regression	Theorical	Theorical

Table 6. Articles with Contribution type classified as “Framework”. ANN: Artificial Neural Network, ARIMA: Autoregressive integrated moving average, BA:Bootstrap aggregating, GARCH: Generalized AutoRegressive Conditional Heteroskedasticity, DNL: Deep Neural Learning, DT: Decision Tree, D-ML: Diverse ML Methods, LSTM: Long short-term memory, NLP: Natural Language Processing, NN: Neural Network, PCA: Principal Component Analysis, RF: Random Forest, RNN: Recurrent neural network, RT: Regression Tree, SVM: Support Vector Machine, SVR: Support Vector Regression

Author	ML vs Eco	ML Algorithms	Econometric models	Topic	Output
[11]	ML Outperforms	NA	Arima	Investment	Future Performance
[35]	Eco Outperforms	SVR/NN	regression	Stock Market	Future price
[44]	Used jointly	RT,BA	Regression	Film Industry	Future Performance
[64]	Used Jointly	ANN	Barndorff-Nielsen and Shephard	Investment	Future Performance
[24]	Non results	RF	Regression	Tourism	Future Performance
[76]	ML outperforms	NN/RF	ARIMA / Garch	Transport	Future price
[52]	Used Jointly	NLP	Regression	Banking	Future price
[19]	Used Jointly	ML	DF	Theorical	Future performance
[13]	ML outperforms	PCA	Regression	Banking	Future performance
[9]	ML outperforms	D-ML	Multivariate Regresison	Theorical	future performance
[55]	ML outperforms	D-ML	Regresison	Insurance	Future Performance
[48]	Non Results	D-ML	Multivariate Regresison	Theorical	future performance
[37]	ML outperforms	D-ML	Regressison	theorical	Future Performance
[30]	Non Results	D-ML	ARIMA	theorical	Future Performance
[43]	Non Results	D-ML	Bayesian Methods	theorical	Future Performance
[63]	Non Results	DNL	Regresison	Theorical	Future Performance
[16]	Non Results	D-ML	Regresison	Theorical	Future Performance
[8]	Non Results	D-ML	Multivariate Regresison	Theorical	Future performance
[3]	Non Results	ML	Regression	Theorical	Future performance
[18]	Non Results	RNN,LSTM	Regressions	Investment	Future performance

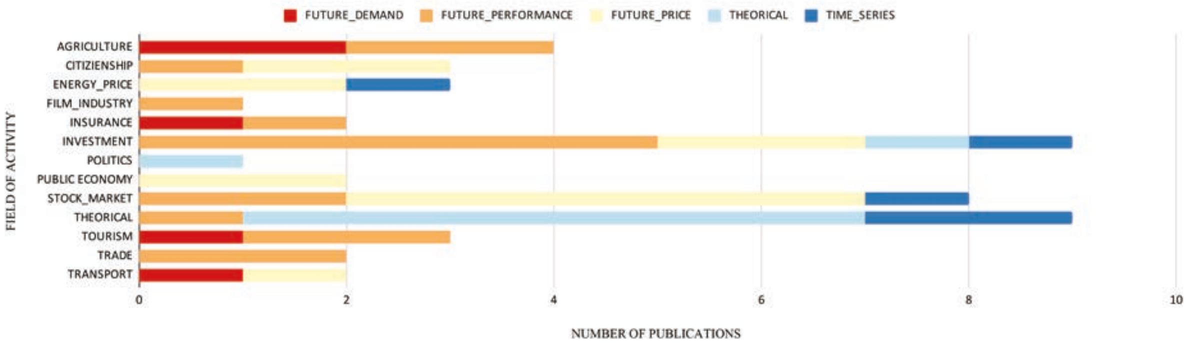


Figure 7. Fields per contribution type

Table 7. Classification according the 4 different hypotheses presented

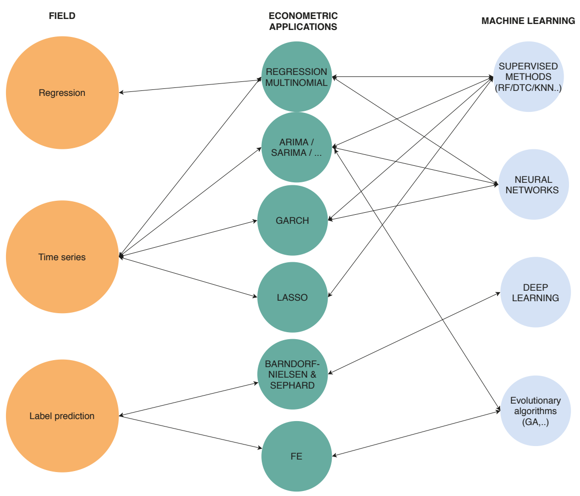
Name	Description
Used Jointly	Combination of Econometric Models and ML algorithms
ML Outperforms	Comparision where ML algorithms achieve better results versus Econometric Models
ECO Outperforms	Comparision where Econometric Models achieve better results versus ML algorithms
No Results	Non Conclusive Results

Table 8. Results in difference

	Used Jointly	ML Outperforms	ECO Outperforms	No results
Method	[38], [62]	[15], [31], [29], [14], [56]	[17]	[36], [5], [75]
Framework	[52], [44], [64], [19]	[11],[55], [37], [48], [9], [76], [13]	[35]	[3], [24], [30], [63] [43] , [16] [18], [8]
Model	[21], [70], [49], [46], [60], [26], [27]	[66], [47],[33], [2], [64], [53], [72], [22]		[25], [54]

5.3 RQ3 - Comparing ML and Econometric Models what are the most frequently applied methods, and in what study context?

Figure 7 compared the algorithms and their applications in econometrics and machine learning. In some cases their performances are contrasted, in others, they are used jointly. The most research methods are ANN and RF vs ARIMA and different types of regression (and the different variants that it can have).



6 Conclusions

Overall, new technologies and applications, such as machine learning, help increase the accuracy of prediction algorithms. A SMS serves to have an overview of the state of the art in the field, this review shows that in terms of the number of publications, this field is growing and more and more innovative and joint models are beginning to emerge, offering better prediction capabilities. In the case of this SMS, it has been observed that there are more than 300 articles in which the two topics are related, however, when it comes to directly compare the two, the number of articles is even smML applications and econometric modelsaller. Taking into account the search criteria of this article, the trend and the highest degree of innovation, as well as the number of articles published, the stock market is where there are more applications of machine learning in econometrics. Taking into account the results of this SMS, there are many possibilities for future lines of research, where econometric models could be combined with supervised machine learning models, and open a new paradigm for the creation and implementation of hybrid models. In addition, other possible future lines of research would be to analyze in more detail some specific sectors of activity in which most of the publications are growing, in which various types of machine learning applications and econometric models converge. An example of these types of activity could be the investment sector (both private equity and securities). Another interesting approach would be to identify

Table 9. ML vs ECO: Machine Learning Algorithm Vs Econometric Model, Publication = Type of Publication, Topic = Field of application, Research = Type of Research, Contribution = Contribution Type, Activity = Evolution Activity

Authors	Title	Year	Publication	ML vs ECO	Topic	Research	Activity	Contribution
[17]	Determinants of credit demand of farmers in Lam Dong, Vietnam: A comparison of Machine Learning and multinomial logit	2019	Article	ECO	Agriculture	Validation	Validate	Method
[15]	Alternative prediction Methods in the stock exchanges of Thailand	2019	Conference	ML	Stock Market	Validation / Solution Proposal	Validate	Method
[31]	Long-term forecast of energy commodities price using Machine Learning	2019	Article	ML	Energy Price	Validation	Validate	Method
[66]	Learning from man or machine: Spatial fixed effects in Urban Econometrics	2019	Article	ML	Citizienship	Validation	Conceptual	Model
[55]	A case study on reducing auto Insurance attrition with Econometrics, Machine Learning, and A/B testing	2019	Conference Paper	ML	Insurance	Experience	Validate	Framework
[11]	The effect of information asymmetries on serial crowdfunding and campaign success	2018	Article	ML	Investment	Validation / Solution Proposal	Analyse	Framework
[33]	Can Machine Learning improve prediction - An application with farm survey data	2018	Article	ML	Agriculture	Validation	Analyse	Model
[8]	The great regression Machine Learning, Econometrics, and the future of quantitative social sciences	2018	Article	Non Results	Theorical	Philosophical	Analyse	Framework
[48]	Machine Learning: An applied Econometric approach	2017	Article	ML	Urban	Philosophical	Analyse	Framework
[37]	Some comments on the current state of Econometrics	2016	Article	ML	Theorical	Philosophical	Analyse	Framework

[61]	Advances in Time Series Analysis and Forecasting	2017	Chapter	ML	Stock Market	Solution	Implement/ Validate	Model
[25]	Empirical Economics	2019	Article	Non Results	Investment	Validation	Implement/ Validate	Model
[16]	Illustrating Statistical Procedures: Finding Meaning in Quantitative Data	2020	Chapter	Non Results	Theorical	Philosophical	Analyse	Framework
[29]	Neural Computing and Applications	2019	Article	ML	Energy Price	Validation / Solution Proposal	Implement/ Validate	Method
[9]	Mathematical and Statistical Methods for Actuarial Sciences and Finance	2018	Chapter	ML	TRADE	Validation	Implement, Validate	Framework
[18]	A Multivariate and Multi-step Ahead Machine Learning Approach to Traditional and Cryptocurrencies Volatility Forecasting	2019	Chapter	Non Results	Stock Market	Validation	Implement, Validate, Analyze	Framework
[63]	Regional Policy Analysis in the Era of Spatial Big Data	2020	Chapter	Non Results	Politics	Philosophical	Analyse	Framework
[36]	Comparative analysis of time series model and machine testing systems for crime forecasting	2019	Article	Non Results	Citizienship	Validation	Implement / Validate	Method
[54]	A novel hybrid model based on Hodrick–Prescott filter and support vector regression algorithm for optimizing stock market price prediction	2017	Article	Non Results	Stock Market	Validation / Solution Proposal	Implement, Validate, Analyze	Model
[62]	A wavelet-based hybrid neural network for short-term electricity prices forecasting	2020	Article	Used Jointly	Energy Price	Validation / Solution Proposal	Implemenr / Validate / Analyze	Method
[50]	Beyond Traditional Probabilistic Methods in Econometrics	2019	Chapter	Non Results	Theorical	Philosophical	Analyse	Framework
[30]	Online workload forecasting	2017	Chapter	Non Results	Theorical	Philosophical	Analyse	Framework
[52]	Computational Data Sciences and the Regulation of Banking and Financial Servicesk	2017	Chapter	Used Jointly	Investment	Validation / Solution Proposal		Framework

[2]	Deep learning and wavelets for high-frequency price forecasting	2018	Chapter	ML	Stock Market	Validation	Implement, Validate, Analyze	Model
[5]	Multistep-ahead prediction: A comparison of analytical and algorithmic approaches	2018	Chapter	Non Results	Theorical	Validation	Implement, Validate, Analyze	Method
[21]	A Machine Learning approach to univariate time series forecasting of quarterly earnings	2020	Article	Used Jointly	Investment	Validation	Implement, Validate, Analyze	Model
[49]	Integration of RNN with GARCH refined by whale optimization algorithm for yield forecasting: a hybrid Machine Learning approach	2020	Article	Used Jointly	Agriculture	Validation / Solution Proposal	Implement, Validate, Analyze	Model
[3]	Machine Learning Methods That Economists Should Know about	2019	Article	Non Results	Theorical	Validation		Framework
[35]	Machine Learning versus Econometric jump models in predictability and domain adaptability of index options	2019	Article	Eco	Stock Market	Validation	Analyse	Framework
[14]	Comparative decision models for anticipating shortage of food grain production in India	2018	Article	ML	Agriculture	Validation	Analyse	Method
[75]	Spatial Choice Modeling Using the Support Vector Machine (SVM): Characterization and Prediction	2018	Chapter	Non Results	Investment	Validation	Analyse	Method
[44]	Machine Learning versus Econometrics: prediction of box office	2019	Article	Used Jointly	Film Industry	Validation	Analyse	Framework
[65]	Refinements of Barndorff-Nielsen and Shephard model: an analysis of crude oil price with Machine Learning	2020	Article	ML	Investment	Validation		Model
[73]	Data science: the impact of statistics	2019	Article	Used Jointly	Theorical	Experience	Analyse	Framework

[69]	Causal tree with instrumental variable: an extension of the causal tree framework to irregular assignment mechanisms	2019	Article	Used Jointly	Investment	Evaluation	Implement, Validate, Analyze	Model
[38]	Data Science	2017	Chapter	Used Jointly	health	Validation	Implement, Validate	Method
[46]	Stacked autoencoder with echo-state regression for tourism demand forecasting using search query data	2018	Article	Used Jointly	Toruism	Validation	Validate	Model
[24]	Real-time inflation forecasting with high-dimensional models: The case of Brazil	2017	Article	Non Results	Public Economy	Validation	Analyse	Framework
[53]	Hedonic Housing Theory - A Machine Learning Investigation	2017	Article	ML	Investment	Validation / Solution Proposal		Model
[60]	Forecasting transportation demand for the US market	2019	Article	Used Jointly	Transport	Validation		Model
[26]	Optimal retail location: Empirical Methodology and application to practice	2019	Article	Used Jointly	Trade	Validation		Model
[76]	Comparison of Econometric Models and Artificial Neural Networks Algorithms for the Prediction of Baltic Dry Index	2019	Article	ML	Transport	Validation	Analyse	Framework
[27]	The Informational Content of the Term Spread in Forecasting the US Inflation Rate: A Nonlinear Approach	2017	Article	Used Jointly	Public Economy	Validation	Validate	Model
[56]	Economic crises and market performance-A machine learning approach	2017	Article	ML	Tourism	Validation	Analyse	Method
[72]	Who performs better? AVMs vs hedonic models	2020	Article	ML	Investment	Validation / Solution Proposal	Proposal	Model

[22]	Artificial Neural networks-based Econometric models for tourism demand forecasting	2017	Conference	ML	Tourism	Validation / Solution Proposal	Proposal	Model
[34]	An Introduction to the NMPC-Graph as General Schema for Causal Modeling of Nonlinear, Multivariate, Dynamic, and Recursive Systems with Focus on Time-Series Prediction	2016	Conference	Non Results	Theorical	Philosophical		
[19]	Batch and incremental dynamic factor Machine Learning for multivariate and multi-step-ahead forecasting	2019	Article	Used Jointly	Theorical	Validation	Implement	Framework
[13]	Nowcasting and forecasting GDP in emerging markets using global financial and macroeconomic diffusion indexes	2019	Article	ML	Public Economy	Validation	Analyse	Framework
[47]	Advances in Time Series Analysis and Forecasting	2017	Chapter	ML	Stock Market	Solution Proposal	Implement	Model

which methods from other scientific fields, for example focusing research on applications of genetic algorithms or deep learning.

7 Funding

This research has been supported by the project ”INTELFIN: Artificial Intelligence for investment and value creation in SMEs through competitive analysis and business environment”, Reference: RTC-2017-6536-7, funded by the Ministry of Science, Innovation and Universities (Challenges-Collaboration 2017), the State Agency for Research (AEI) and the European Regional Development Fund (ERDF).

8 Conflicts of interest

The authors declare no conflict of interest.

Author contributions: M.E.P.-P. and J.P.-D. have conducted the analysis of the articles. E.H-v, S.O and J.M.C. formalized the problem, designed the methodology and reviewed the work. All the authors contributed to the redaction of the paper. All authors have read and agreed to the published version of the manuscript.

References

[1] Nesreen K Ahmed, Amir F Atiya, Neamat El Gayar, and Hisham El-Shishiny. An empirical comparison of machine learning models for time series forecasting. Econometric Reviews, 29(5-6): 594–621, 2010.

[2] Andrés Arévalo, Jaime Nino, Diego León, German Hernandez, and Javier Sandoval. Deep learning and wavelets for high-frequency price forecasting. In International Conference on Computational Science, pages 385–399. Springer, 2018.

[3] Susan Athey. The impact of machine learning on economics. In The economics of artificial intelligence: An agenda, pages 507–547. University of Chicago Press, 2018.

[4] Susan Athey and Guido W Imbens. Machine learning methods that economists should know about. Annual Review of Economics, 11: 685–725, 2019.

[5] Fouad Bahrpeyma, Mark Roantree, and Andrew McCarren. Multistep-ahead prediction: A comparison of analytical and algorithmic approaches. In International Conference on Big Data Analytics and Knowledge Discovery, pages 345–354. Springer, 2018.

[6] Patrick Bajari, Denis Nekipelov, Stephen P Ryan, and Miaoyu Yang. Machine learning methods for demand estimation. American Economic Review, 105(5): 481–85, 2015.

[7] Alexandre Belloni, Daniel Chen, Victor Chernozhukov, and Christian Hansen. Sparse models and methods for optimal instruments with an application to eminent domain. Econometrica, 80(6): 2369–2429, 2012.

[8] Julien Boelaert and Étienne Ollion. The great regression. Revue française de sociologie, 59(3): 475–506, 2018.

[9] Stefano Bonini and Giuliana Caivano. Probability of default modeling: A machine learning approach. In Mathematical and Statistical Methods for Actuarial Sciences and Finance, pages 173–177. Springer, 2018.

[10] Wichor M Bramer, Melissa L Rethlefsen, Jos Kleijnen, and Oscar H Franco. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. Systematic reviews, 6(1): 1–12, 2017.

[11] Vincenzo Butticè, Carlotta Orsenigo, and Mike Wright. The effect of information asymmetries on serial crowdfunding and campaign success. Economia e Politica Industriale, 45(2): 143–173, 2018.

[12] Rich Caruana and Alexandru Niculescu-Mizil. An empirical comparison of supervised learning algorithms. In Proceedings of the 23rd international conference on Machine learning, pages 161–168, 2006.

[13] Oguzhan Cepni, I Ethem Güney, and Norman R Swanson. Nowcasting and forecasting gdp in emerging markets using global financial and macroeconomic diffusion indexes. International Journal of Forecasting, 35(2): 555–572, 2019.

[14] Manojit Chattopadhyay and Subrata Kumar Mitra. Comparative decision models for anticipating shortage of food grain production in india. Theoretical and applied climatology, 131(1-2): 523–530, 2018.

[15] Kanchana Chokethaworn, Chukiat Chaiboonsri, and Satawat Wannapan. Alternative prediction methods in the stock exchanges of thailand. In Journal of Physics: Conference Series, volume 1324, page 012086. IOP Publishing, 2019.

[16] Ray W Cooksey. Illustrating statistical procedures: Finding meaning in quantitative data. Springer Nature, 2014.

[17] Huy Duc Dang, Au Hai Thi Dam, Thuyen Thi Pham, and Tra My Thi Nguyen. Determinants of credit demand of farmers in lam dong, vietnam. Agricultural Finance Review, 2019.

[18] Jacopo De Stefani, Olivier Caelen, Dalila Hattab, Yann-Aël Le Borgne, and Gianluca Bontempi. A multivariate and multi-step ahead machine learning approach to traditional and cryptocurrencies volatility forecasting. In ECML PKDD 2018 Workshops, pages 7–22. Springer, 2018.

[19] Jacopo De Stefani, Yann-Aël Le Borgne, Olivier Caelen, Dalila Hattab, and Gianluca Bontempi. Batch and incremental dynamic factor machine learning for multivariate and multi-step-ahead forecasting. International Journal of Data Science and Analytics, 7(4): 311–329, 2019.

[20] Mette Brandt Eriksen and Tove Faber Frandsen. The impact of patient, intervention, comparison, outcome (pico) as a search strategy tool on literature search quality: a systematic review. Journal of the Medical Library Association: JMLA, 106(4): 420, 2018.

[21] Jan Alexander Fischer, Philipp Pohl, and Dietmar Ratz. A machine learning approach to univariate time series forecasting of quarterly earnings. Review of Quantitative Finance and Accounting, pages 1–17, 2020.

[22] Raffaella Folgieri, Tea Baldigara, and Maja Mamula. Artificial neural networks-based econometric models for tourism demand forecasting. Tourism in South East Europe..., 4: 169–182, 2017.

[23] Jason Furman and Robert Seamans. Ai and the economy. Innovation Policy and the Economy, 19(1): 161–191, 2019.

[24] Márcio GP Garcia, Marcelo C Medeiros, and Gabriel FR Vasconcelos. Real-time inflation forecasting with high-dimensional models: The case of brazil. International Journal of Forecasting, 33(3): 679–693, 2017.

[25] Jose Ignacio Gimenez-Nadal, Miguel Lafuente, Jose Alberto Molina, and Jorge Velilla. Resampling and bootstrap algorithforecastingms to assess the relevance of variables: applications to cross section entrepreneurship data. Empirical Economics, 56(1): 233–267, 2019.

[26] Chloe Kim Glaeser, Marshall Fisher, and Xuanming Su. Optimal retail location: Empirical methodology and application to practice: Finalist–2017 m&som practice-based research competition. Manufacturing & Service Operations Management, 21(1): 86–102, 2019.

[27] Periklis Gogas, Theophilos Papadimitriou, Vasilios Plakandaras, and Rangan Gupta. The informational content of the term-spread in forecasting the us inflation rate: A nonlinear approach. Available at SSRN 2990336, 2017.

[28] Christian Gourieroux and Alain Monfort. Statistics and econometric models, volume 1. Cambridge University Press, 1995.

[29] Coşkun Hamzaçebi, Hüseyin Avni Es, and Recep Çakmak. Forecasting of turkey's monthly electricity demand by seasonal artificial neural network. Neural Computing and Applications, pages 1–15, 2019.

[30] Nikolas Herbst, Ayman Amin, Artur Andrzejak, Lars Grunske, Samuel Kounev, Ole J Mengshoel, and Priya Sundararajan. Online workload forecasting. In Self-Aware Computing Systems, pages 529–553. Springer, 2017.

[31] Gabriel Paes Herrera, Michel Constantino, Benjamin Miranda Tabak, Hemerson Pistori, Jen-Je Su, and Athula Naranpanawa. Long-term forecast of energy commodities price using machine learning. Energy, 179: 214–221, 2019.

[32] Ming-Wei Hsu, Stefan Lessmann, Ming-Chien Sung, Tiejun Ma, and Johnnie EV Johnson. Bridging the divide in financial market forecasting: machine learners vs. financial economists. Expert Systems with Applications, 61: 215–234, 2016.

[33] Jennifer Ifft, Ryan Kuhns, and Kevin Patrick. Can machine learning improve prediction–an application with farm survey data. International Food and Agribusiness Management Review, 21(1030-2019-611): 1083–1098, 2018.

[34] Christoph Jahnz. An introduction to the nmpc-graph as general schema for causal modeling of nonlinear, multivariate, dynamic, and recursive systems with focus on time-series prediction. In Proceedings of SAI Intelligent Systems Conference, pages 825–852. Springer, 2016.

[35] H Jang and J Lee. Machine learning versus econometric jump models in predictability and domain adaptability of index options. Physica A: Statistical Mechanics and its Applications, 513: 74–86, 2019.

[36] Sudan Jha, Eunmok Yang, Alaa Omran Almagrabi, Ali Kashif Bashir, and Gyanendra Prasad Joshi. Comparative analysis of time series model and machine testing systems for crime forecasting. NEURAL COMPUTING & APPLICATIONS, 2020.

[37] George Judge. Some comments on the current state of econometrics. Annual Review of Resource Economics, 8: 1–6, 2016.

[38] Jan Kalina and Jaroslav Hlinka. On coupling robust estimation with regularization for high-dimensional data. In Data Science, pages 15–27. Springer, 2017.

[39] Barbara Kitchenham and Pearl Brereton. A systematic review of systematic review process research in software engineering. Information and software technology, 55(12): 2049–2075, 2013.

[40] Barbara Kitchenham and Stuart Charters. Guidelines for performing systematic literature reviews in software engineering. Citeseer, 2007.

[41] Barbara A Kitchenham, David Budgen, and O Pearl Brereton. Using mapping studies as the basis for further research–a participant-observer case study. Information and Software Technology, 53(6): 638–651, 2011.

[42] Sotiris B Kotsiantis, I Zaharakis, and P Pintelas. Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160(1): 3–24, 2007.

[43] Vladik Kreinovich, Nguyen Ngoc Thach, Nguyen Duc Trung, and Dang Van Thanh. Beyond Traditional Probabilistic Methods in Economics, volume 809. Springer, 2018.

[44] Yan Liu and Tian Xie. Machine learning versus econometrics: prediction of box office. Applied Economics Letters, 26(2): 124–130, 2019.

[45] Marcos Lopez de Prado. Beyond econometrics: A roadmap towards financial machine learning. Available at SSRN 3365282, 2019.

[46] Sheng-Xiang Lv, Lu Peng, and Lin Wang. Stacked autoencoder with echo-state regression for tourism demand forecasting using search query data. Applied Soft Computing, 73: 119–133, 2018.

[47] Dusan Marcek. Statistical models and granular soft rbf neural network for malaysia klci price index prediction. In International Work-Conference on Time Series Analysis, pages 401–412. Springer, 2016.

[48] Sendhil Mullainathan and Jann Spiess. Machine learning: an applied econometric approach. Journal of Economic Perspectives, 31(2): 87–106, 2017.

[49] P Murali, R Revathy, S Balamurali, and AS Tayade. Integration of rnn with garch refined by whale optimization algorithm for yield forecasting: a hybrid machine learning approach. JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING, 2020.

[50] Hung T Nguyen, Nguyen Duc Trung, and Nguyen Ngoc Thach. Beyond traditional probabilistic methods in econometrics. In International Econometric Conference of Vietnam, pages 3–21. Springer, 2019.

[51] Isaac Odun-Ayo, Olasupo Ajayi, Rowland Goddy-Worlu, and Jamaiah Yahaya. A systematic mapping study of cloud resources management and scalability in brokering, scheduling, capacity planning and elasticity. Asian Journal of Scientific Research, 2019.

[52] Sharyn O'Halloran, Marion Dumas, Sameer Maskey, Geraldine McAllister, and David K Park. Computational data sciences and the regulation of banking and financial services. In From Social Data Mining and Analysis to Prediction and Community Detection, pages 179–209. Springer, 2017.

[53] Timothy Oladunni and Sharad Sharma. Hedonic housing theory—a machine learning investigation. In 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), pages 522–527. IEEE, 2016.

[54] Meryem Ouahilal, Mohammed El Mohajir, Mohamed Chahhou, and Badr Eddine El Mohajir. A novel hybrid model based on hodrick–prescott filter and support vector regression algorithm for optimizing stock market price prediction. Journal of Big Data, 4(1): 31, 2017.

[55] Miguel Paredes. A case study on reducing auto insurance attrition with econometrics, machine learning, and a/b testing. In 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), pages 410–414. IEEE, 2018.

[56] José Francisco Perles-Ribes, Ana Belén Ramón-Rodríguez, Luis Moreno-Izquierdo, and Martín Sevilla-Jiménez. Economic crises and market performance—a machine learning approach. Tourism Economics, 23(3): 692–696, 2017.

[57] Kai Petersen, Robert Feldt, Shahid Mujtaba, and Michael Mattsson. Systematic mapping studies in software engineering. In 12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12, pages 1–10, 2008.

[58] Kai Petersen, Sairam Vakkalanka, and Ludwik Kuzniarz. Guidelines for conducting systematic mapping studies in software engineering: An update. Information and Software Technology, 64: 1–18, 2015.

[59] Mark Petticrew and Helen Roberts. Systematic reviews in the social sciences: A practical guide. John Wiley & Sons, 2008.

[60] Vasilios Plakandaras, Theophilos Papadimitriou, and Periklis Gogas. Forecasting transportation demand for the us market. Transportation Research Part A: Policy and Practice, 126: 195–214, 2019.

[61] Ignacio Rojas, Héctor Pomares, and Olga Valenzuela. Advances in Time Series Analysis and Forecasting: Selected Contributions from ITISE 2016. Springer, 2017.

[62] Foued Saâdaoui and Hana Rabbouch. A wavelet-based hybrid neural network for short-term electricity prices forecasting. Artificial Intelligence Review, 52(1): 649–669, 2019.

[63] Laurie A Schintler. Regional policy analysis in the era of spatial big data. In Development Studies in Regional Science, pages 93–109. Springer, 2020.

[64] Indranil SenGupta, William Nganje, and Erik Hanson. Refinements of barndorff-nielsen and shephard model: an analysis of crude oil price with machine learning. Annals of Data Science, pages 1–17, 2019.

[65] Indranil SenGupta, William Nganje, and Erik Hanson. Refinements of barndorff-nielsen and shephard model: an analysis of crude oil price with machine learning. Annals of Data Science, pages 1–17, 2020.

[66] Åvald Sommervoll and Dag Einar Sommervoll. Learning from man or machine: Spatial fixed effects in urban econometrics. Regional Science and Urban Economics, 77: 239–252, 2019.

[67] Standford, Index 2018, https://hai.stanford.edu/ai - index-2018, 2018.

[68] Standford, Index 2019, https://hai.stanford.edu/ research/ai-index -2019, 2019.

[69] Falco J Bargagli Stoffi and Giorgio Gnecco. Causal tree with instrumental variable: an extension of the causal tree framework to irregular assignment mechanisms. International Journal of Data Science and Analytics, pages 1–23, 2019.

[70] Falco J Bargagli Stoffi and Giorgio Gnecco. Causal tree with instrumental variable: an extension of the causal tree framework to irregular assignment mechanisms. International Journal of Data Science and Analytics, 9(3): 315–337, 2020.

[71] Jan Tinbergen. Shaping the world economy; suggestions for an international economic policy. NA, 1962.

[72] Agostino Valier. Who performs better? avms vs hedonic models. Journal of Property Investment & Finance, 2020.

[73] Claus Weihs and Katja Ickstadt. Data science: the impact of statistics. International Journal of Data Science and Analytics, 6(3): 189–194, 2018.

[74] Claes Wohlin, Per Runeson, Paulo Anselmo da Mota Silveira Neto, Emelie Engström, Ivan do Carmo Machado, and Eduardo Santana De Almeida. On the reliability of mapping studies in software engineering. Journal of Systems and Software, 86(10): 2594–2610, 2013.

[75] Yong Yoon. Spatial choice modeling using the support vector machine (svm): Characterization and prediction. In International Conference of the Thailand Econometrics Society, pages 767–778. Springer, 2018.

[76] Xin Zhang, Tianyuan Xue, and H Eugene Stanley. Comparison of econometric models and artificial neural networks algorithms for the prediction of baltic dry index. IEEE Access, 7: 1647–1657, 2018.

Eugenia Pérez-Pons is a Ph.D. candidate in Computer Engineering and Intelligent Systems at University of Salamanca. She holds a degree in International Business from Pompeu Fabra University, an Executive Master in Digital Business from ESADE and a Master in Data Science from the Open University of Catalonia. She has 8 years of professional experience in multinational companies in Spain, Russia and Germany in companies such as Mango, Seat, Ikea or Lidl. Her research interests include the application of artificial intelligence algorithms in different fields. She has participated in national and international conferences and has several publications in international journals and book chapters.

Javier Parra is Ph.D. Assistant Professor in the Department of Business and Economics Administration at the University of Salamanca. Since 2009, Javier has been complimenting his work as a professor with the management of different technology departments in private companies. He is computer technician from the University of Wales and Master in Macroeconometrics and Finance from the Menéndez Pelayo International University, the Institute of Fiscal Studies and the Centre for Economic and Commercial Studies of Spain. Javier is a member of the BISITE Research Group where he develops his research interests related to economics, technological finance, econometrics and digital intelligence. Currently, Javier combines research at BISITE with his work at IoT Digital Innovation Hub, which has been founded with the objective of developing innovation in the Internet of the Things of Small and Medium Enterprises.

Sigeru Omatu received his Ph.D. in Electronic Engineering from Osaka Prefecture University and joined the faculty at University of Tokushima in 1969. He was Professor of University of Tokushima in 1988 and Professor of Osaka Prefecture University in 1995. He has been Professor of Osaka Institute of Technology since 2010. His honors and awards include the Best Paper Awards for Distributed Parameter System Theory, IEE of Japan, 1991, for Intelligent Classification, JSME, 1995, for Coin and Bill Classification, SICE, Japan, 1995, for Intelligent Smell Classification, IARIA, 2008, for Neuro-Control, IARIA, 2009. Furthermore, he received Ichimura Distinguished Award for Intelligent Classification, New Technology Development Foundation, 1996, SICE Fellow and IEEJ Outstanding Achievement Award, 2005. In 2011 he received Commendation for Science and Technology in Research Sector by the Minister of Education, Culture, Sport, Science and Technology. He used to be the President of ISCIA (Institute of Systems, Control and Information Association, Japan) in 2007 and President of Electronics, Information and Systems Society, IEEJ, 2004-2006, Regional Editor of EAAI(IFAC Journal) for 1998-2001, Associate Editor of IEEE Transactions on Neural Networks for 1993-2003, and IMA Journal of Mathematical Control and Information (Oxford University Press) since 1991. His research area covers intelligent signal processing, pattern recognition, intelligent control, and adaptive control.

Enrique Herrera-Viedma received the Ph.D. degree in computer science from the University of Granada, Granada, Spain, in 1996.,He is currently a Professor with the Department of Computer Science and Artificial Intelligence, University of Granada, where he is currently the Vice-Rector for Research and Knowledge Transfer. He has also been a Guest Lecturer in plenary lectures and tutorials in multiple national and international conferences related to artificial intelligence (AI). He has authored or coauthored more than 300 papers in JCR journals. In 2013, he published in the prestigious journal Science about the new role of digital libraries in the era of the information society.,Prof. Herrera-Viedma was the Vice-President for Publications with the IEEE System Man and Cybernetics Society during 2019–2020. He is currently the VP for Cybernetics, founder of the IEEE Transactions in Artificial Intelligence, and a Highly Cited Researcher by Clarivate Analytics in Computer Science and Engineering in 2014–2020. His H-index is 101 in Google Scholar (more than 33 000 citations) and 85 in WoS (more than 23 000 citations). He is an Associate Editor of several AI journals like IEEE Transactions Fuzzy Systems, IEEE Transactions Intelligent Transportation Systems, IEEE Transactions Systems, Man, and Cybernetics: Systems, Knosys, Applied Soft Computing, Fuzzy Optimization and Decision Making, Information Sciences, Soft Computing.

Juan M. Corchado was born in Salamanca, Spain, in 1971. He received the Ph.D. degree in computer sciences from the University of Salamanca and the Ph.D. degree in artificial intelligence from the University of the West of Scotland. He was the Vice President for Research and Technology Transfer, from 2013 to 2017, and the Director of the Science Park with the University of Salamanca, where he was also the Director of the Doctoral School, until 2017. He has been elected twice as the Dean of the Faculty of Science with the University of Salamanca. He has been a Visiting Professor with the Osaka Institute of Technology, since 2015, and a Visiting Professor with University Teknologi Malaysia, since 2017. He is the Director of the Bioinformatics, Intelligent Systems, and Educational Technology (BISITE) Research Group, which he created, in 2000. He is the President of the IEEE Systems, Man and Cybernetics Spanish Chapter and the Academic Director of the Institute of Digital Art and Animation, University of Salamanca, where he is currently a Full Professor. He also oversees the master's programs in digital animation, security, mobile technology, community management, and management for TIC Enterprises with the University of Salamanca. He is a member of the Advisory Group on Online Terrorist Propaganda of the European Counter Terrorism Centre (EUROPOL). He is also an Editor and the Editor-in-Chief of specialized journals such as the Advances in Distributed Computing and Artificial Intelligence Journal, the International Journal of Digital Contents and Applications, and the Oriental Journal of Computer Science and Technology.