Journal of Business Research 182 (2024) 114805

Contents lists available at ScienceDirect

Journal of Business Research

journal homepage: www.elsevier.com/locate/jbusres

Prediction-led prescription: Optimal Decision-Making in times of turbulence and business performance improvement

A. Schäfers [a], V. Bougioukos [b,c], G. Karamatzanis [c,d], K. Nikolopoulos [c,*]

[a] Alliance Manchester Business School, United Kingdom
[b] Notre Dame, United Kingdom
[c] IHRR Forecasting Lab, Durham University & Durham University Business School, United Kingdom
[d] Norwich Business School, University of East anglia, United Kingdom

ARTICLE INFO

Keywords:
Prediction-led Prescription
Forecasting
Special events
AI
Turbulence

ABSTRACT

Can you have prescription without prediction? Most scholars and practitioners would argue that a good forecast drives an optimal decision, thus promoting the concept of *prediction-led prescription.* In times of turbulence, Special events like promotions and supply chain disruptions are impacting businesses severely. Nevertheless, limited research has been carried out to date to accurately forecast the impact of, and consequentially prescribe in the presence of special events. Nowadays Artificial Intelligence (AI) predictive analytics methods and heuristics imitate and even improve human intelligence, progressively leading towards innovative cognitive analytics solutions. This research aims to contribute to applying advancements in AI-based predictive analytics to improve business performance. We provide empirical evidence that these AI solutions outperform the popular (especially among practitioners) linear regression models. We corroborate the stream of literature arguing that AI predictive analytics could – via a natural path-dependent process – enhance prescriptive analytics solutions, and thus improve business performance.

1. Introduction & motivation

Can we have prescription without prediction? Although academics' and practitioners' views may vary (Siemsen and Spiliotopoulou, 2023), the dominant school of thought sees them entangled (Nikolopoulos, 2021). Especially in times of turbulence where special events are impacting critical sectors (Nikolopoulos et al., 2021; Nikolopoulos et al., 2015), optimal decisions better be taken once we have accurate estimates of the current as well as the future state of systems (Petropoulos et al., 2022). Thus, despite prescriptive and predictive analytics being perceived as different sets of tools, the truth is that the latter regularly leads the former. In essence, forecasts lead to optimal decisions, and thus, predictive analyses drive prescriptive solutions; in a nutshell: prediction-led prescription.

Nowadays, we live in a world massively dominated by data, as well as an unprecedented increase in data collection and increasing computational power that has led to the phenomenon (and opportunity) of big data analytics (BDA) in recent years. Moreover, Artificial Intelligence (AI) elementsin decision systems have become important parts (Duan et al., 2019). Similarly, BDA and AI created a loop of co-development, which offered clear improvements in decision-making in times of turbulence (Zhang et al., 2021; Gunasekaran et al. 2017; George et al. 2014; Nikolopoulos, 2010). The reason is that these techniques seem promising for prediction but also for prescribing optimal decision thereafter (Kim & Swanson, 2018). Also, external factors like the financial meltdown in 2008, supply chain disruptions, and stock-outs (Nikolopoulos et al., 2021) raised the interest for more refined BDA (Huang et al., 2014).

Despite the breadth and depth of forecasting research (Petropoulos et al., 2022), special events and their respective impact on baseline time series forecasting remain at large an under-researched topic (Nikolopoulos, 2021). This is not a trivial technical challenge, due to mostly limited available and often non-parametric and nonlinear past data (Webby & O'Connor, 1996).

Nikolopoulos (2010; 2021) provided systematic reviews and theoretical propositions to conceptualise and simulate special events in time series. Both articles – despite being a decade apart – highlight the need for further research into this area. Although being tested on simulated and real-life data, the results of these studies with a clear and immediate need for more advanced AI-based predictive, prescriptive, and cognitive

* Corresponding author at: Durham University Business School, Durahm University, Durham DH1 3LE.
E-mail address: kostas.nikolopoulos@durham.ac.uk (K. Nikolopoulos).

https://doi.org/10.1016/j.jbusres.2024.114805
Received 20 December 2022; Received in revised form 20 June 2024; Accepted 25 June 2024
Available online 3 July 2024
0148-2963/© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

analytics models, and if back then (Nikolopoulos, 2010), computational power was not empowering such solutions, nowadays – with a super-computer and the cloud at hand, it really is about time. Furthermore, it is well documented to date that nonlinear methods like neural networks have not been able to consistently outperform statistical methods (Makridakis et al., 2018) – thus, AI-based solutions have a long way to go. Despite AI-based methods' great flexibility and ability to approximate non-parametric relationships, their performance has not always been as expected. Nevertheless, when forecasting baseline time series, Zhang and Haghani (2015) could observe a strong performance of these algorithms in grasping non-regular patterns and special events in the context of forecasting traffic time series. Furthermore, Kraus et al., (2017) found strong performance of these algorithms in the context of complex financial time series data sets.

The primary aim of this research is to assess the performance of predictive analytics methods when forecasting the impact of special events (we employ the real-world "Rossmann store sales" dataset; Kaggle, 2016), leading to the prescription of optimal decisions (ordering, inventory positions, trading strategies etc) and therefore improved *business performance:*

> *"As demand forecasting accuracy increases, and the standard deviation associated with the forecast decreases, the need to hold "just in case" inventory also goes down. This leads to lower inventory carrying costs and thus better cash flow"*[1]

To the extent that.

> *"1% forecast improvement leads to a 2.5 % reduction in the amount of inventory that needs to be held."*[2]

If that is not a manifestation of the potential benefits of prediction-led prescription, then what is?

Secondly, we corroborate,the stream of literature on applying machine learning methods – and AI in general, to achieve better quality prescriptive analytics (Bertsimas & Van Parys, 2022; Bertsimas and Kallus, 2020; Ban and Rudin, 2019; Lee et al., 2018; Van der Vlist, 2016).

Essentially, we argue that our results contribute to the scientific premises of a framework which supports the effective transition from predictive to prescriptive analytics by offering additional credibility on the optimal strategic choice, leading to superior operational performance (Elmachtoub & Grigas, 2022; (Bertsimas & Kallus, 2020). AI tools, such as the ones researched in this article, enhance further the value dimension of analytics in the stage of 'prediction', leading to more value to the next level of 'prescription'; that improved decision-making offering stronger competitiveness which could foster improved performance. This is also the natural path towards 'cognitive' analytics solutions (following the diagonal in the data analytics framework in Charles, Emrouznejad,Gherman,and Cochran, 2022), as such solutions are not created in a void and they follow a natural evolutionary process from simpler solutions to more advanced ones as per the conceptual framework of Charles, Emrouznejad, Gherman, and Cochran (2022).

In our study, the transition from 'predictive' to 'prescriptive' is facilitated by the identification and forecasting of special events. Such an identification will not only offer greater accuracy when anticipating the future, but it will also allow the decision maker to optimize their behaviour by selecting from a finite strategy-set; if the firm has to select a pure or mixed strategy for reaching the greater pay-off, the choice or the probabilities assigned will be much more accurate thus making easier for the firm realizing/answering the question of 'how can we make it happen'?

AI methods can capture information more effectively and create a set of knowledge, making special events more predictable and providing solutions to tackle the issues presented in the prescriptive stage of analytics. Real-world analytics applications often include elements of both predictive and optimization effects (Cohen et al., 2017; Angalakudati et al., 2014) that help the transition from predictive to prescriptive analytics and therefore, our results contribute further into this valuable path dependence.

We will even dare claim that given the combinatory nature of our approach, that our theoretical proposition can be classified as an early-stage cognitive analytics tool, also given the autonomous way the machine learning methods do learn and extrapolate from past data and knowledge.

The remaining of this paper is structured as follows: section 2 offers a comprehensive yet targeted visit in the relevant literature, followed by section 3 on our theoretical foundations and framework. Section 4 discussed our research methodology and section 5 offers our empirical results followed by discussion (section 6). The last section offers conclusion, limitations, implications for theory and practice, and suggestions for future work on the topic.

2. Background literature

We follow a targeted literature review methodological approach focus on three areas of the literature: a) forecasting trends, b) forecasting special events, and c) using AI for a) and b).

2.1. Forecasting trends

Forecasting and prediction methods and models have been successfully applied for smoothing and estimating baseline trends for decades now (Syntetos et al., 2016; Assimakopoulos & Nikolopoulos, 2000; Brown, 1963). Besides the traditional linear statistical models, machine learning algorithms are nowadays often used for prediction and forecasting tasks (Wauters & Vanhoucke, 2017; Chen & Guestrin, 2016; Hendry, 1987).

The strong focus on quantitative methods can be attributed to the fact that very often, we get very mixed results from qualitative forecasting methods – most notably judgmental forecasting methods (Petropoulos et al., 2022). Human judgment is so often based on simple mental strategies and heuristics driven by past experiences (Goodwin & Wright, 2010). This latter argument explains to a large extent the limitations and often very mixed results of qualitative methods (Werner, et al., 2017; Nikolopoulos, et al., 2015; Genrea, et al., 2013; Lawrence, et al., 2000; Webby & O'Connor, 1996) (Lawrence, et al., 2000).

One other dimension of predictive analytics methods is computational cost (Leitch & Tanner, 1991). Newly proposed methods need to be evaluated both on accuracy as well as computational cost – i.e time to compute. Especially in the last two decades nonlinear methods, like Artificial Neural Networks (ANN – Haerdle, 1992) and tree-based methods have attracted wider attention because of their increasing accuracy but also in parallel increasing computational power needed and respective cost (Haykin, 1998).

2.2. Forecasting special events

Special events are as old as time, as is the need to forecast the timing and impact of them (Petropoulos et al., 2022); yet the formal introduction of the term, and a first set of proposition on approaches for the identification and forecasting of such events came from Wilpen (as late as) in 1986 in the context of relational database management systems: [any] "out of the ordinary events". These irregular but (temporally) expected out of ordinary events (Nikolopoulos, 2021; Nikolopoulos, 2010; Armstrong, 2001) are difficult to identify and even more difficult to forecast in a time series context. Limited data (more often than not), or at best non-parametric and nonlinear past data of previous events, makes a forecast difficult (Webby & O'Connor, 1996). Special events are

[1] The Forecasting Accuracy Bugaboo (forbes.com).

[2] Demand Planning Solutions Improve Forecasting By Consuming More And More Data (forbes.com).

also mentioned in the literature as *shocks*, *rare events* or grey swans (or even black swans but these are meant to be unforecastable – Taleb, 2007). Black swans have a major impact that is even more difficult to expect or forecast (Aven, 2015) for example terror attacks or tsunamis (Aven, 2016; Nafday, 2011). Special events are more similar to rare events or shocks (Kesavan and Kushwaha, 2013) like for examples jumps in financial markets or even more regular events like promotions or strikes (Nikolopoulos, 2010). Research also remains inconclusive regarding the value of qualitative forecating – i.e expert adjustements for adjusting the impact of special events (Goodwin , 2002; King & Zeng, 2001; Goodwin & Fildes, 1999).

Therefore, more focus have been given to quantitative approaches – under the caveat that past data are avaialbe, like parameter estimation and model intervention (Webby & O'Connor, 1996). In another proposition Lee and Yum combined two separate ANNs to forecast special events (Lee & Yum, 1998). Nunes and Pimentel developed the statistical jump-diffusion processes (Nunes & Pimentel, 2017), and Martzoukos and Trigeorgis proposed a Markov-chain methodology to forecast these jumps and abnormal peaks (Martzoukos & Trigeorgis, 2002). Other approaches can be also found in the context of maintenance modelling in Cha *et al.* (2018), or Ren et al. (2016), in the context of traffic prediction. Theofilatos et al. (2017) indicate that the small ratio between special events and regular periods causes problems for statistical forecasting methods. This observation was also described by Goodwin and Wright (2010), when predicting the occurrence of terror attacks or kidnapping (Goodwin & Wright, 2010), where that ANNs can outperform linear statistical forecasting techniques and autoregressive methods in the context of special events.

Nikolopoulos (2010) applied a wider range of forecasting tools, including ANNs, multi linear regression (MLR) and nearest neighbour techniques. The linear techniques proved to be sufficient under certain circumstances, especially when forecasting simulated data with linear relationships; nevertheless, ANNs outperformed the other techniques, in data with stronger nonlinearity. Combinations, or sophisticated selection between ANN and MLR depending on data characteristics gave also promising results Another interesting approach was forecasting with rare events logistic regression model (RELR), applied by Ren *et al.* in 2016. These models are based on a further development of conventional logistic regression. The challenge with their research was that satisfactory results could be only obtained by massively shrinking the dataset to a ratio of 10:1 between rare events and regular periods. Furthermore, RELR only performed slightly better than regular regression and the prediction rate was still not very satisfactory (Barrow & Kourentzes, 2018). Barrow and Kourentzes applied a range of forecasting algorithms to capture the impact of holidays and promotional effects. Their paper is of high interest due to the large number of forecasting methods empirically competing seasonal naïve, seasonal moving average, exponential smoothing, seasonal exponential smoothing and ANNs. The main finding from this study was that both linear and nonlinear models, exhibit limited performance without the additional information provided in form of dummy variables (Barrow & Kourentzes, 2018).

2.3. AI-based/ Machine learning approaches

Classical statistical forecasting techniques, that work well in the presence of a few variables and large volume of data, cannot cope with problems with a high degree of dimensionality. (Bansal, et al., 1993). On the contrary, Machine Learning (ML) approaches, with their ability to learn from patterns and input–output relationships (Samuel, 1969), have recently shown very promising performance in such challenging contexts (Makridakis, et al., 2018; Smyl, et al., 2018; Keung, Zhang & Xu, 2017; Ma, et al., 2016; Makridakis, Hogarth, & Gaba, 2009; Carbonneau, et al., 2008).

In recent years ML, due to the unprecedented access to computing power and storage via the cloud, have seen a strong revival due to been capable of handling large data sets with high complexity and non-parametric distributions (Sanders,& Ganeshan, 2018; Fisher & Raman, 2018; Cohen, 2017; Feng & Shanthikumar, 2017). Linear statistical models on the other hand, still work very well in the presence of strong linearity (Finlay, 2011). Ensembles are used to enhance the results of ML algorithms further (Fitzpatrick & Muesa, 2016; Zhang & Haghani, 2015; Dietterich, 2000).

Random Forests (RF – Breiman, 2001) have been very popular and successful in empirical investigations (Lessmann, Sungb, & Johnson, 2011). When comparing to Support Vector Machines (SVMs) and Artificial Neural Networks ANNs (Petropoulos et al., 2022), as well as other non-parametric models, RFs are not perceived as full black box models, and that is valued by both academic and practitioners, as the importance of the individual attributes can be measured visualised too (Liang & Lin, 2014; Lessmann, Sungb, & Johnson, 2010). RFs are faster than ANNs and easier to train, giving them a huge advantage in terms of computational cost and complexity. Successful applications of RFs include customer relationship management, medical science, and bioinformatics (Krauss, et al., 2017; Nagya, et al., 2016; Fitzpatrick & Muesa, 2016). Similar results can be found in Baboota and Kaur‘s paper in 2018, where RF (compared to SVMs) were able to capture complex relationships when dealing with special events. With the help of feature engineering both RF and SVM algorithms could produce quite balanced results, nevertheless RF outperformed SVMs. In 2013 Liu et al. authors also compared RFs with ANNs, as well as SVMs. They also found superior performance of RF (Liu, et al., 2013). Also Ahmad et al., observed that RF were able to discover and forecast sudden nonlinear fluctuations in their dataset. The RF algorithms performed again more robustly, even under significant noise and missing data (Ahmad, et al., 2017).

One very promising ML approach, the Gradient Boosting (GB), unfortunately has architecture that is more challenging to fine-tune (Krauss, et al., 2017; Hastie, Tibshirani, & Friedman, 2009). Despite generating very competitive results in general, in comparison to RF the GB algorithms tend to over-fit (Hastie, Tibshirani, & Friedman, 2009; Mason, et al., 1999). GB algorithms are also non-parametric like RF, thus able to handle complex interactions among attributes and capture complex nonlinear trends. Furthermore, they are able to handle small samples with nonlinear distributions very well.

In a 2007 paper, GB algorithms were applied in a very noisy data set (Death, 2007). The GB algorithms were able to outperform ANNs, especially in the small data sets he used. Similar results were found by Yang et al. (2007). The second main advantage of GB versus ANNs is that like RF models are not fully black box models (Martinez, et al., 2018; Zhang and Haghani, 2015). Also, in a financial analytics context, Finlay (2011) could make use of the GB architecture for a credit scoring system. In this paper standard benchmarks could not perform well due to the complex underlying patterns with many individual attributes. All nonlinear models, ANNs, SVMs and boosted trees outperformed the linear models, with boosting offering the best compromise between accuracy and computational cost (Finlay, 2011).

The aforementioned targeted literature review concludes with the highlight of the gap in having a plethora of a) accurate and relatively computationally cheap forecasting algorithms for modelling and forecasting special events and b) consequently optimally prescribing decisions based on these forecasts.

3. Prediction-led prescription: Theoretical foundations and methodological considerations

The current work essentially links two sub-fields of analytics: predictive and prescriptive; and inevitably paves the way towards cognitive analytics solutions. The theoretical foundations of our proposition are ceased from the Charles et al. (2022) data analytics framework, where the framework is expanded (see Fig. 1) to allow for:

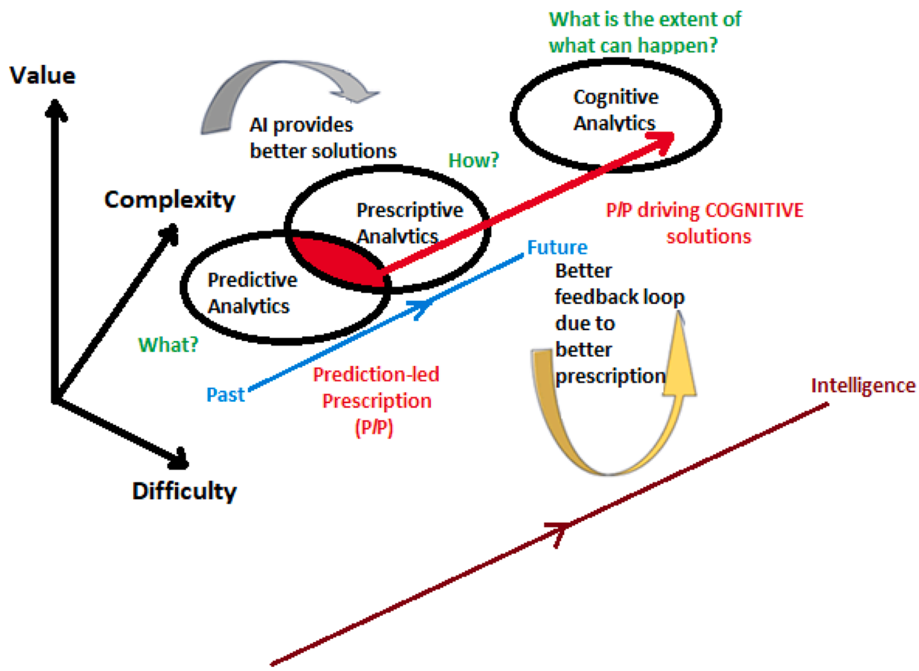

Fig. 1. Expansion of Charles et al., 2022 Data Analytics Framework to include Prediction-led Prescription (PlP) and respective driving of cognitive solutions.

a) the Predictive and Prescriptive analytics to have an intersection, the area exactly where predictive analytics are informing and driving the prescriptive ones, introducing Prediction-led Prescription (P*l*P), and
b) allowing P*l*P[3] to pave the way and drive cognitive analytics solutions too. Cognitive analytics while pooling information and adjusting/optimising model parameters on present time (Gudivada et al., 2016), could be influenced and driven to some extent from the data set until the optimal prescribed action (that has been already influenced by earlier predictions).

Predictive analytics for special events can drive solutions and equally provide guidance as to what steps can be taken to improve business performance by identifying and handling the occurrence of special events, action which can be detrimental to the timely intervention needed to secure and enhance the performance of the business (Leicht-Deobald et al., 2019; Burton et al., 2019; Schafheitle et al., 2019; Isson & Harriott, 2016). Previous research indicates that 'black swan' events such as the COVID-19 pandemic can lead to very severe consequences for the performance of businesses (Donthu & Gustafsson, 2020). A study by Baumgartner (et al., 2020) reveals that up to 45 % of businesses' annual revenue can vanish due to the impact of various adverse events globally. A recent study (Wamba et al., 2019) highlighted the need for further research on the potential impact of Artificial Intelligence in crises. Thus, we strongly believe our research highlights potential ways to navigate through uncertainty and improve decision-making in changing economic conditions and looming disasters, which can clearly influence the performance of an organisation.

Methodologically this study (as illustrated in the Fig. 2 below), attempts to expand our ability to process and handle ***both high volume*** and ***high complexity*** of information, thus creating a better prescriptive set of actions improving our ability to move forward, and thus improve business performance (Schafheitle et al., 2019).

It is worth noting that very often, prescriptive analytics uses solution-oriented simulation and scenario calculations as well as machine learning algorithms with the aim of aiding and implementing decisions (Sivathanu & Pillai, 2019; Lunsford & Phillips, 2018). At that stage, decision-making becomes a joint human-algorithm decision-making process (Burton et al., 2019). As it is argued that AI has improved making predictions in a more efficient manner (Agrawal et al., 2018), the usage of such techniques can help to deliver more accurate optimization of decisions. This study attempts to go further than an accurate forecast, paving the way for better implementation of future actions such as a choice of strategy out of set of alternative strategies. One key conceptualization, which drives a big part of the argument of our endeavour, is connected with the improvement of the reaction time, but as well as with the quality of the decision, while entering the phase of prescriptive actions as given in Fig. 4 and given that these algorithms (machine learning/deep learning) in principle emulate human cognitive processes, we can claim, decisively, that the proposed solution, can drive towards cognitive analytics ones (as depicted in Fig. 3 as well).

The notion that prescriptive analytics offer the ground for proactive actions, which are based on results generated from predictive analytics (Lepenioti et al., 2020), is on the core of our argument which supports the ability for more accurate forecasts in the domain of occurrence of special events that can lead to better business performance. Similar rational is also apparent to the framework offered by Charles at al. (2022) (see Fig. 1).

4. Empirical investigation

We employ in our empirical investigation the publicly available "Rossmann Store Sales" dataset (Kaggle, 2016). The dataset consists of 1115 longitudinal time series for 1115 distinct drug stores in Germany, covering the period of January 2013 to the end of July 2015. The dataset contains a mix of continuous and categorical variables; therefore, we used a series of dummy variables to prevent the problem of spurious ordering (Brown, 2015; Berry & Linoff, 2004, p. 554). Missing values were treated accordingly, too, via either using median values of this attribute or assigned randomly in order to abide by the original distribution (Castro et al., 2017). The dataset was split into an initial fixed 70/30 ratio for fitting versus out-of-sample forecasting evaluation (Petropoulos et al., 2022), and then a rolling origin evaluation took place (Tashman, 2000), thus starting with a six-month training and three-month evaluation period, the dataset was step by step prolonged by

[3] We will be using the terms 'Prediction-led Prescription' and 'P*l*P' interchangeable onwards in the manuscript.

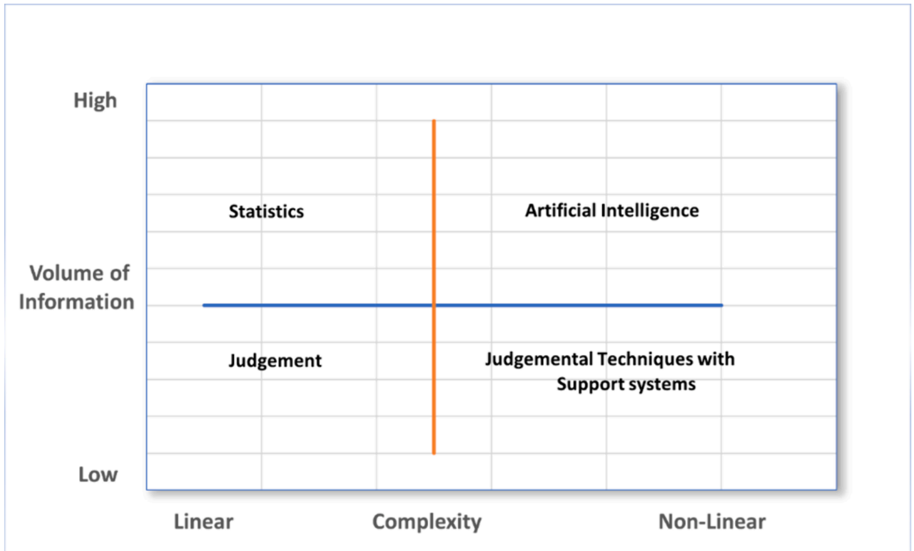

Fig. 2. Methodological choices of forecasting methods when considering **Volume** of Information versus **Complexity** of Information.

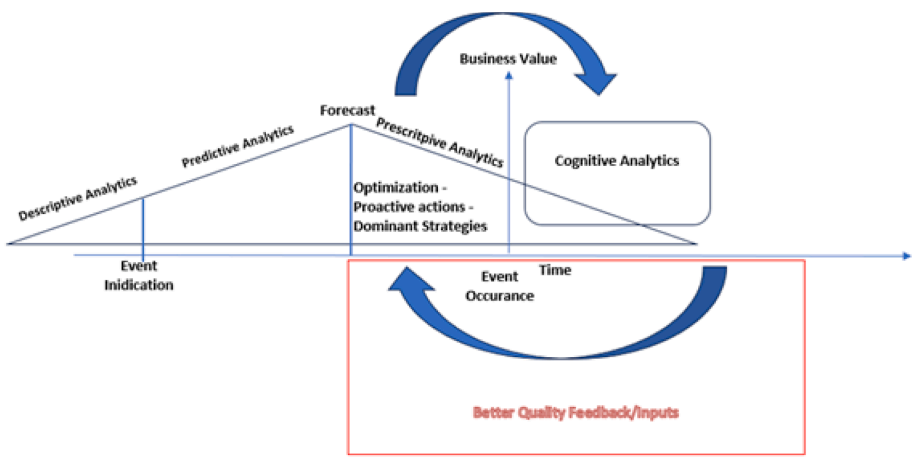

Fig. 3. The theoretical model –
Adopted from Lepenioti et al., 2020.

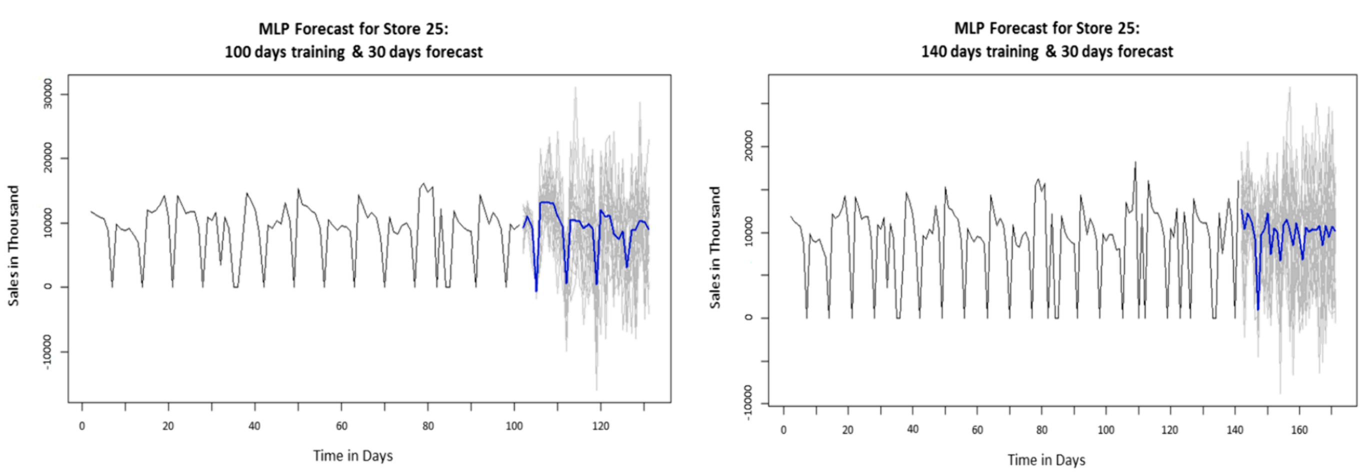

Fig. 4. *Store* #25. **Left** part: 100 days for training and 30 days out-of-sample for forecasting evaluation, **Right** part: 140 days for training and 30 days out-of-sample for forecasting evaluation. Forecasting performed with a Multilayer Perceptron (Artificial Neural Network).

another 6 months period.

An example of the 1115 time series is illustrated in Fig. 4, *store* #25, with a series of Special Events in the training part of the left Fig. 4 (15 data points with zero sales). On the left part of Fig. 4, we set 100 days for training and 30 days out-of-sample for forecasting evaluation. In contrast, in the right part, the rolling origin has been applied, with (an extended) 140 days for training and (still) 30 days out-of-sample for forecasting evaluation. Forecasting in this example is performed with the standard benchmark of Artificial Neural Networks (ANN), the Multiple Layer Perceptron (MLP) (Petropoulos et al., 2022).

From the illustration in Fig. 5, a time series dominated by special events, the importance of forecasting the timing and the extent (i.e. the amplitude) of these events is becoming evident as these rapid changes in sales (and respective demand) affect severely inventory levels, production schedules, staffing, sourcing and whatnot, and thus make clear the link of predictive, predictive, and cognitive analytics and the importance of the P*l*P paradigm shift.

Following the methodological paradigm of Nikolopoulos, Babai and Bozos (2016), we employ multiple measures of forecasting performance: a) Runtime to track the computational cost of the forecasting method, b) the Root Means Squared Error (RMSE) to track the uncertainty of the forecasts, and c) the Mean Absolute Error (MAE) to trach forecasting accuracy. As the multiple metrics may indicate different winners, we also calculate a ranking metric so as to pick the ultimate winner of our empirical investigation (Petropoulos et al., 2022).

Although we are in a time series context, the interest lies in modelling and forecasting the special events that are modelled by cues of information, and as such, we need multivariate methods in order to perform the task. We use Multiple Linear Regression (MLR) as the natural benchmark, the standard Machine Learning (ML) benchmark in the like of an MLP ANN (Haykin, 1998), and employ three state-of-the-art ML methods – Random Forest (RF), Gradient Boosting (GB), and Extreme Boosting(xgBoosT) (Petropoulos et al., 2022; Chen et al., 2018) in order to model and forecast our special events.

From Table 1 and our holistic empirical comparison, MLR as expected dis the fastest approach, while xgBoost the fastest of all the ML approaches. The ML benchmark (MLP ANN) does worse than the standard linear benchmark (MLR) that perform overall quite strongly, and when all is said and done, RF is the overall winner, tha tis consistent with earlier studies on the forecasting field.

We also take one step further our analysis and we plot the *MAE* over *Runtime*. If the objective is to provide a method that is both accurate and fast – as is very often the case in industry, for example the case of UBER[4] and the Theta method (Assimakopoulos and Nikolopoulos, 2000), the efficiency frontier in can help to make a decision (Fig. 5). It can be observed, that for a ‘fast and frugal’ ML estimation, the xgBoost architecture offers the most promising results. Although the xgBoost was the fastest nonlinear algorithm, the RF architecture outperformed the xgBoost architecture in terms of forecasting accuracy, with a slightly longer runtime.

This final result contradicts the findings of he paper of Krauss et al. (2017) and Zhang and Haghani’s (2 0 5) papers, both observing a slightly better performance for GB architectures versus RF.

5. Discussion: Form prediction to prescription to cognition

Prescriptive analytics has been helping businesses to achieve improved performance outcomes for quite some time (den Hertog & Postek, 2016), via answering questions relevant to ‘what should I do?’ (Lepenioti,et al., 2020,p. 57) and ‘how can we make it happen?’ (Charles et al. p.44). As Šikšnys & Pedersen, 2016 argue the aim is prescribing the best choice in order to gain the optimal results from a predicted future via a big data set. This can be achieved by incorporating in the process various predictive techniques such as ML and AI (Syntetos et al. 2026; Basu 2013). AI has improved making prediction in a more efficient manner (Agrawal et al., 2018), and the usage of such techniques can help to deliver more accurate prediction and respective decisions. The notion that prescriptive analytics offer the ground for proactive actions, which are based on results generated from predictive analytics (Lepenioti et al., 2020), is on the core of our argument (on P*l*P) which supports the ability for more accurate forecasts in the presence of special events that can lead to better business performance. Similar rational is also apparent to the framework offered by Charles at al.

Corroborating recent studies have addressed the importance of optimal solutions on newsvendor problems (Ban and Rudin, 2019) or generally optimise an unknown optimisation objective using Machine Learning models (Bertsimas and Kallus, 2020), our study also further contributes to the literature trend of applying AI methods to achieve better quality of prescriptive analytics which are mainly currently developed in a conceptual way (Lee et al., 2018; Bertsimas & Van Parys 2022).

We argue that our results contribute to the scientific premises of the P/P framework, which supports the effective transition from predictive to prescriptive analytics by offering additional credibility on the optimal strategic choice for superior operational performance (Elmachtoub & Grigas, 2022; Bertsimas & Kallus, 2020). The predictive algorithmic decision-making method promoted in our study is practically leading to a prescriptive analytics approach by offering the more insightful alternatives for optimal decision-making similar to other studies (Van der Vlist, 2016). As per the framework offered by Charles et al. (2022), this study contributes to the advising on all possible scenarios to transit from foresight (predictive) to wide sight (prescriptive).

The transition from predictive to prescriptive (Fig. 2) is facilitated with the identification of special events and, in turn, such an identification will not only add greater accuracy on forecasting performance results but it will also allow the decision maker to optimize their behaviour by selecting the dominant strategy from a finite strategy-set (or infinite set if the problem is theoretical) or allowing to achieve optimal scenario building and running the appropriate simulations hence achieving better prescriptive analytics results. Usually, prescriptive analytics set cost minimization objectives (Achenbach & Spinler, 2018); however, our contribution via realizing the special event allows us to optimize any relevant objective. Our study shows that by achieving a better result via AI, we were able to utilize it with a plausible prescriptive technique in order to get the best prescription possible.

Therefore, the incorporation of AI predictive techniques allows the best implementation of the intersecting prescriptive techniques, as in the figure above, leading to the best strategic decisions, optimization tasks or even simulated scenario-based outcomes. The bridging between predictive and prescriptive is based on better filtering of the various options that should be available after an accurate prediction.

On the other hand, cognitive process are more elegant and sophisticated and attempt to develop an insight on deep sight to reveal patterns from data usually given in an unstructured form (Charles et al., 2022). There is a cognitive computing environment that generates some inferences for via feedback which are stored and can then be employed as actions in the future using cognitive models similar to the human brain (Gudivada et al., 2016). Even though the field is still emerging from the very beginning it was clear that cognitive analytics refers to the use of self-learning algorithms simulating on human cognitive processes, which can adapt on differences in data generating immediate responses (Hurwitz et al., 2015).

The cognitive aspect is not just a big-data analysis but “draws upon the cognitive computing environment to generate actionable insights by analysing diverse heterogeneous data sources using cognitive models that the human brain employs” (Gudivada et al., 2016, pp. 169–170). Machine learning and AI on training and learning is of highest importance in cognitive analytics and our analysis and structure and unstructured data can be part of that process (Phillips-Wren et al., 2015).

[4] Forecasting at Uber: An Introduction | Uber Blog.

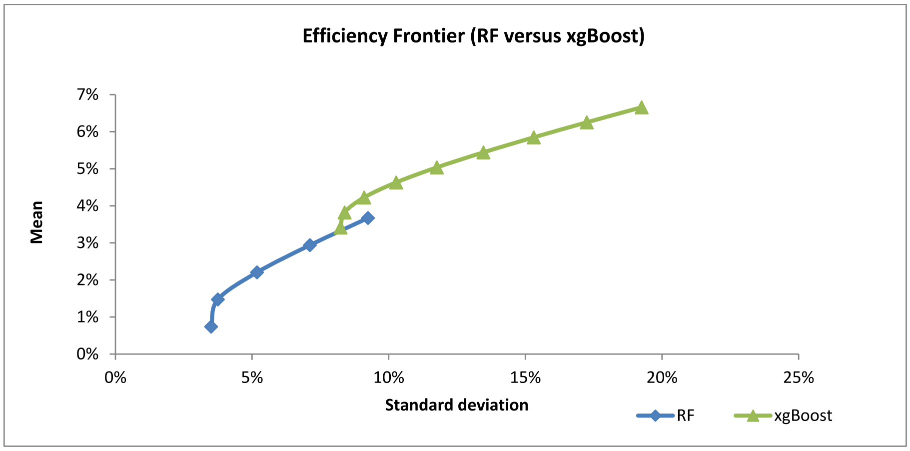

Fig. 5. Efficiency Frontier, RF vs GB.

Table 1
Empirical Forecasting Competition of LR, MLP, RF, GB and xgBoost vs MLR and ANN.

	Computational cost	Uncertainty	Accuracy	Combining all metrics	
Method	Runtime[in min]	RMSE	MAE	RANKING	
				Score	**RANK**
Artificial Neural Network (ANN)	2.570	4485.670	57.150	4 + 5 + 5 = 14	**5**
Random Forest (RF)	**1.800**	**1270.000**	**27.600**	**3 + 1 + 1 = 5**	**1**
Gradient Boosting (GB)	2.690	1325.000	29.300	5 + 2 + 2 = 9	**3**
Extreme Boosting (xgBoost)	0.110	1321.000	29.100	2 + 3 + 4 = 9	**3**
Multiple Linear Regression MLR	*0.006*	*2039.000*	*38.280*	*1 + 4 + 3 = 8*	***2***

With **bold** the best-performing method; With *Italics* the standard benchmark.

Our study has shown that using AI and capturing easier the best response on strategic decisions will foster stronger and more robust training on the cognitive process that is mimicked by the computer capabilities. Any novel system could take advantage of findings which can give better optimal choices on the prescriptive side. Then these choices can feed in the loop via more advance AI methods to promote further intelligence of the system.

This embeddedness of AI can become a very important component into those processes, and their potential of developing data-driven algorithms; algorithms that perform prescriptive analytics, based on cognitive data (Gunasekaran et al 2017), thus our study is an affirmative case showing how greater credibility can be added on the various stages of analytics.

6. Conclusion, Limitations, and the future

Following the methodological contribution discussed in section 5, on the empirical end of this research, the currently most promising nonlinear ML algorithms – RF, GB and xgBoost (Petropoulos et al., 2022), have been successfully applied in the context of forecasting special events, in a systematic way aiming to lead to improved performance during periods of higher uncertainty, and the winner came out to be the well-celebrated RF (Breiman, 2001,1996).

The research could synthesize the initial findings made by Nikolopoulos in 2010 in the field of forecasting special events and the latest results in the field of ML. Furthermore, the research could strengthen Huang *et al.* research. In their paper it was argued that competitive information, like promotions are important factors for forecasting sales for retailers (Huang, et al., 2014) It could be observed that information about promotions and competitors increases the achievable accuracy.

An average runtime of fewer than fifteen minutes on a regular desktop office computer and a dataset with more than 70,000 observations is very practicable, even for decision-making under time pressure. Nevertheless, as stated above already, satisfactory results can only be achieved with initial data pre-processing and feature engineering, like for any other quantitative approach. This study paper also lays the foundation for further research in the field of ML and special events.

For future research on the topic, based on historical successful performance of ensemble methods (Dietterich, 2000: Finlay, 2011), a combination of the best performing methods in form of ensembles, in a similar approach to Nikolopoulos (2010) 'expert' methods, could potentially further increase the accuracy and robustness of our proposition.

Regarding Makridakis et al. observation, that a combination of ML techniques, linear statistical methods and forecasting models achieved the highest results on the M4 time series data set, suggests conducting further research in this direction (Makridakis, et al., 2018). The application of seasonal artificial networks (SANN) could be another approach for further research, based on the poor performance of the pure autoregressive time series MLPs (Adhikari & Agrawal, 2013). AI could feed and support complex special event processing techniques, which will be utilized to prescribe in a way that it improves proactive rather than reactive business decisions and thus, achieving better operational performance. Cognition comes as the last natural step, since these algorithms emulate the way humans operate, assuming large datasets are in place for the necessary training; and this is a very inspirational avenue for further research.

CRediT authorship contribution statement

A. Schaefers: Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **V. Bougioukos:** Writing – review & editing, Methodology, Conceptualization. **G. Karamatzanis:** Writing – review & editing, Project administration. **K. Nikolopoulos:** Writing – review & editing, Validation, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Adhikari, R. & Agrawal, R., (2013). '*An Introductory Study on Time Series Modeling and Forecasting*'. Available at: https://arxiv.org/ftp/arxiv/papers/1302/1302.6613.pdf (Accessed: 20 08 2018).

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction Machines: The simple economics of artificial intelligence*. Harvard Business Press.

Ahmad, M., Mourshed, M., & Rezgui, Y. (2017). Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy and Buildings, 147*, 77–89.

Angalakudati, M., Balwani, S., Calzada, J., Chatterjee, B., Perakis, G., Raad, N., & Uichanco, J. (2014). Business analytics for flexible resource allocation under random emergencies. *Management Sci., 60*(6), 1552–1573.

Armstrong, J. (2001). *Principles of forecasting: A handbook for researchers and practitioners*. Dordrecht: Kluwer Academic Publishers.

Assimakopoulos, A., & Nikolopoulos, K. (2000). The theta model: A decomposition approach to forecasting. *International Journal of Forecasting, 16*(4), 521–530.

Aven, T. (2015). Implications of black swans to the foundations and practice of risk assessment and management. *Reliability Engineering and System Safety, 134*, 83–91.

Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research, 253*, 1–13.

Ban, G. Y., & Rudin, C. (2019). The big data newsvendor: Practical insights from machine learning. *Operations Research, 67*(1), 90–108.

Bansal, A., Kaufmann, R., & Weitz, R. (1993). Comparing the Modeling Performance of Regression and Neural Networks as Data Quality Varies: A Business Value Approach. *Journal of Management Information Systems, 10*, 11–32.

Barrow, D., & Kourentzes, N. (2018). The impact of special days in call arrivals forecasting: A neural network approach to modelling special days. *European Journal of Operational Research, 264*, 967–977.

Baumgartner, T., Malik, Y., & Padhi, A. (2020). 'Reimagining industrial supply chains'. McKinsey Article.

Berry, M., & Linoff, G. (2004). *Data mining techniques*. Indianapolis, Indiana: Wiley Publishing Inc.

Bertsimas, D., & Kallus, N. (2020). From Predictive to prescriptive analytics. *Management Science, 66*(3), 1025–1044.

Bertsimas, D., & Van Prays, B. (2022). Bootstrap robust, prescriptive analytics. *Mathematical Progamming, 195*, 39–78.

Breiman, L. (1996). Bagging Predictors. *Machine Learning, 24*, 123–140.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5–32.

Brown, C., 2015. *Package 'dummies'*. Available at: https://cran.r-project.org/web/packages/dummies/dummies.pdf (Accessed: 09 07 2018).

Brown, R.G. (1963). Smoothing, Forecasting and Prediction of Discrete Time Series, Englewood Cliffs, New Jersey: Prentice-Hall.

Burton, J. W., Stein, M. K., & Jensen, T. B. (2019). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making, 32*(1), 15–29.

Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research, 184*, 1140–1154.

Castro, J., Gomez, D., Molina, E., & Tejada, J. (2017). Improving polynomial estimation of the Shapley value by stratified random sampling with optimum allocation. *Computers and Operations Research, 82*, 180–188.

Charles, V., Emrouznejad, A., Gherman, T., & Cochran, J. (2022). Why data analytics is an art. *Significance, 19*(6), 42–45. https://doi.org/10.1111/1740-9713.01707

Chen, T. & Guestrin, C., 2016. '*XGBoost: A Scalable Tree Boosting System*' Available at: http://www.kdd.org/kdd2016/papers/files/rfp0697-chenAemb.pdf (Accessed: 25 07 2018).

Chen, T., He, T. & Benesty, M., 2018. *Package 'xgboost'* Available at: https://cran.r-project.org/web/packages/xgboost/xgboost.pdf (Accessed 18 07 2018).

Cohen, M. (2017). Big Data and Service Operations. *Production and Operations Management, 27*(9), 1709–1723.

Cohen, M. C., Leung, N.-H.-Z., Panchamgam, K., Perakis, G., & Smith, A. (2017). The impact of linear optimization on promotion planning. *Operation Research., 65*(2), 446–1444.

Death, G. (2007). Boosted trees for ecological modeling and prediction. *Ecological Society of America, 88*(1), 243–251.

den Hertog, D., & Postek, K. (2016). *Bridging the gap between predictive and prescriptive analytics-new optimization methodology needed*. Tilburg, The Netherlands: Tilburg Univ.

Dietterich, T. (2000). Ensemble Methods in Machine Learning. In *Multiple Classifier Systems* (pp. 1–15). Berlin: Springer-Verlag, Berlin Heidelberg.

Donthu, N., & Gustafsson, A. (2020). Effects of COVID-19 on business and research. *Journal of Business Research, 117*, 284.

Duan, Y., Edwards, S. J., & Dwivdei, K. Y. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63–71.

Elmachtoub, A., & Grigas, P. (2022). Smart "Predict, then Optimize''. *Management Science, 68*(1), 9–26.

Feng, Q., & Shanthikumar, J. (2017). How Research in Production and Operations Management May Evolve in the Era of Big Data. *Production and Operations Management, 27*(09), 1670–1684.

Finlay, S. (2011). Multiple classifier architectures and their application to credit risk assessment. *European Journal of Operational Research, 210*, 368–378.

Fisher, M., & Raman, A. (2018). Using Data and Big Data in Retailing. *Production and Operations Management, 27*(09), 1665–1669.

Fitzpatrick, T., & Muesa, C. (2016). An empirical comparison of classification algorithms for mortgage default prediction: Evidence from a distressed mortgage market. *European Journal of Operational Research, 249*, 427–439.

Genrea, V., Kennya, G., Meyler, A., & Timmermann, A. (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting, 29*, 108–121.

George, G., Haas, M. & Pentland A. (2014). 'Big data and management', Academy of Management, 57 (2), pp. 321-326 Briarcliff Manor, NY.

Goodwin, P. (2002). Integrating management judgment and statistical methods to improve short-term forecasts. *Omega, 30*(2), 127–135.

Goodwin, P., & Fildes, R. (1999). Judgmental forecasts of time series affected by special events: Does providing a statistical forecast improve accuracy? *Journal of Behavioral Decision Making, 12*(1), 37–53.

Goodwin, P., & Wright, G. (2010). The limits of forecasting methods in anticipating rare events. *Technological Forecasting & Social Change, 77*, 355–368.

Gudivada, V. N., Irfan, M. T., Fathi, E., & Rao, D. L. R. (2016). "Cognitive Analytics: Going Beyond Big Data Analytics and Machine Learning," in Handbook of Statistics volume 35. *Elsevier*, 169–205.

Gunasekaran, A., Papadopoulos, T., Rameshwar, D., Wamba, S., F., Childe, S., Hazen, B., & Akter, S., (2017)'Big data predictive analytics for supply chian and organizational performance', *Journal of Business Research*, 70, pp. 208-317.

Haerdle, W. (1992). *Applied nonparametric regression*. Cambridge: Cambridge University Press.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *Boosting and Additive Trees, in The Elements of Statistical Learning*. New York: New York Springer.

Haykin, S. (1998). *Neural Networks: A comprehensive foundation*. New York: Pearson US Imports & PHIPEs.

Hurwitz, J., Kaufman, M., & Bowles, A. (2015). *Cognitive Computing and Big Data Analytics*. Indianapolis: John Wiley and Sons.

Hendry, L. (1987). The potential impact of artificial intelligence on the practice of OR. *European Journnl of Operational Research, 28*, 218–225.

Huang, T., Fildes, R., & Soopramanien, D. (2014). The value of competitive information in forecasting FMCG retail product sales and the variable selection problem. *European Journal of Operational Research, 237*, 738–748.

Isson, J. P., & Harriott, J. S. (2016). 'People analytics in the era of big data: Changing the way you attract, acquire, develop, and retain talent'. Wiley.

Kaggle, 2016. *Rossmann Store Sales.* Available at: https://www.kaggle.com/c/rossmann-store-sales (Accessed: 19 03 2018).

Kesavan, S., & Kushwaha, T. (2013). Differences in retail inventory investment behavior during macroeconomic shocks: role of service level. *Production and Operations Management, 23*(12), 2118–2136.

Keung L., R., Zhang, W. & Xu, W. (2017). 'Parallel Aspect-Oriented Sentiment Analysis for Sales Forecasting with Big Data', *Production and Operations Management,* 27 (10) , S. 1775-1794.

Kim, H., & Swanson, N. (2018). Mining big data using parsimonious factor, machine learning, variable selection and shrinkage methods. *International Journal of Forecasting, 34*, 339–354.

King, G., & Zeng, L. (2001). Explaining Rare Events in International Relations. *International Organization, 55*(3), 693–715.

Krauss, C., Do, X., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research, 259*, 689–702.

Lawrence, M., O'Connor, M., & Edmundson, B. (2000). A field study of sales forecasting accuracy and processes. *European Journal of Operational Research, 122*, 151–160.

Lee, J., & Yum, C. (1998). 'Judgmental adjustment in time series forecasting using neural networks. *Decision Support Systems', 22*, 135–154.

Lee, J., Davari, H., Singh, J., & Pandhare, V. (2018). 'Industrial artificial intelligence for industry 4.0 -based manufacturing systems. *Manufacturing Letters, 18*, 20–23.

Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I., & Kasper, G. (2019). The challenges of algorithm-based HR decision-making for personal integrity. *Journal of Business Ethics, 160*(2), 377–392.

Leitch, G., & Tanner, J. (1991). Economic forecast evaluation: Profit versus the conventional error measure. *American Economic Review, 81*(3), 580–590.

Lepenioti, K., Bousdekis, A., Apostolou, D., & Mentzas, G. (2020). 'Prescriptive Analytics. *Literature Review and research challenges', International Journal of Information Management, 50*, 57–70.

Lessmann, S., Sung, M., & Johnson, J. (2010). Alternative methods of predicting competitive events: An application in horserace betting markets. *International Journal of Forecasting, 26*, 518–536.

Liang, T., & Lin, J. (2014). A two-stage segment and prediction model for mortgage prepayment prediction and management. *International Journal of Forecasting, 30*, 328–343.

Liu, M., Wang, M., Wang, J., & Li, D. (2013). Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators B, 177*, 970–980.
Lunsford, D. L., & Phillips, P. P. (2018). 'Tools used by organizations to support human capital analytics. *Performance Improvement', 57*(3), 6–15.
Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). The M4 Competition: Results, findings, conclusion and way forward. *International Journal of Forecasting, 34*(4), 802–808.
Makridakis, S., Hogarth, R., & Gaba, A. (2009). Forecasting and uncertainty in the economic and business world. *International Journal of Forecasting, 25*, 794–812.
Martinez, A., et al. (2018). A machine learning framework for customer purchase prediction in the non-contractual setting. *European Journal of Operational Research*. https://doi.org/10.1016/j.ejor.2018.04.034 [Preprint]. (Accessed: 12.07.2018.
Martzoukos, S., & Trigeorgis, L. (2002). Real investment options with multiple sources of rare events. *European Journal of Operational Research, 136*, 696–706.
Ma, S., Fildes, R., & Huang, T. (2016). Demand forecasting with high dimensional data: The case of SKU retail sales forecasting with intra- and intercategory promotional information. *European Journa lof Operationa lResearch, 249*, 245–257.
Mason, L., Baxter, J., Bartlett, P. & Frean, M., (1999). Boosting Algorithms as Gradient Descent, in: 12, ed. *Advances in Neural Information Processing Systems.* Boston: MIT Press, p. 512–518.
Nafday, A. (2011). Consequence-based structural design approach for black swan events. *Structural Safety, 33*, 108–114.
Nagya, G., et al. (2016). GEFCom2014: Probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach. *International Journal of Forecasting, 32*, 1087–1093.
Nikolopoulos, K. I., Babai, M. Z., & Bozos, K. (2016). Forecasting supply chain sporadic demand with nearest neighbor approaches. *International Journal of Production Economics, 177*, 139–148.
Nikolopoulos, K. (2010). Forecasting with quantitative methods the impact of special events in time series. *Applied Economics, 42*(8), 947–955.
Nikolopoulos, K., et al. (2015). Relative performance of methods for forecasting special events. *Journal of Business Research, 68*, 1785–1791.
Nikolopoulos, K., Punia, S., Schäfers, A., Tsinopoulos, C., & Vasilakis, C. (2021). Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. *European Journal of operational research, 290*(1), 99–115.
Nikolopoulos, K. (2021). We need to talk about intermittent demand forecasting. *European Journal of Operational Research, 291*(2), 549–559.
Nunes, C., & Pimentel, R. (2017). Analytical solution for an investment problem under uncertainties with shocks. *European Journal of Operational Research, 259*, 1054–1063.
Petropoulos, F.; Apiletti, D.; Assimakopoulos, V.; Babai, M.Z.; Barrow, D.K.; Bergmeir, C.; Bessa, R.J.; Boylan, J.E.; Browell, J.; Carnevale, C.; Castle, J.L.; Cirillo, P.; Clements, M.P.; Cordeiro, C.; Cyrino Oliveira, F.L.; De Baets, S.; Dokumentov, A.; Fiszeder, P.; Franses, P.H.; Gilliland, M.; Gönül, M.S.; Goodwin, P.; Grossi, L.; Grushka-Cockayne, Y.; Guidolin, M.; Guidolin, M.; Gunter, U.; Guo, X.; Guseo, R.; Harvey, N.; Hendry, D. F.; Hollyman, R.; Januschowski, T.; Jeon, J.; Jose, V.R.R.; Kang, Y.; Koehler, A.B.; Kolassa, S.; Kourentzes, N.; Leva, S.; Li, F.; Litsiou, K.; Makridakis, S.; Martinez, A.B.; Meeran, S.; Modis, T.; Nikolopoulos, K.; Önkal, D.; Paccagnini, A.; Panapakidis, I.; Pavía, J.M.; Pedio, M.; Pedregal Tercero, D.J.; Pinson, P.; Ramos, P.; Rapach, D.; Reade, J.J.; Rostami-Tabar, B.; Rubaszek, M.; Sermpinis, G.; Shang, H.L.; Spiliotis, E.; Syntetos, A.A.; Talagala, P.D.; Talagala, T.S.; Tashman, L.; Thomakos, D.; Thorarinsdottir, T.; Todini, E.; Trapero Arenas, J.R.; Wang, X.; Winkler, R.L.; Yusupova, A.; and Ziel, Z. Forecasting: theory and practice. International Journal of Forecasting, 38, 3 (2022), 705-871.
Phillips-Wren, G. E., Iyer, L. S., Kulkarni, U. R., & Ariyachandra, T. (2015). Business analytics in the context of big data: A roadmap for research. *Communications of the Association for Information Systems, 37*(23), 449–472.
Ren, Y., et al. (2016). Influential factors of red-light running at signalized intersection and prediction using a rare events logistic regression model. *Accident Analysis and Prevention, 95*, 266–273.
Sanders, N., & Ganeshan, R. (2018). Big Data in Supply Chain Management. *Production and Operations Management, 27*(10), 1745–1748.
Samuel, A. (1969). Some Studies in Machine Learning Using the Game of Checkers. *Annual Review in Automatic Programming, 6*, 1–36.
Schafheitle, S. D., Weibel, A., Ebert, I. L., Kasper, G., Schank, C., & Leicht-Deobald, U. (2019). No stone left unturned? Towards a framework for the impact of datafication technologies on organizational control. *Academy of Management Discoveries, 6*(3), 455–487.
Siemsen, E. and Spiliotopoulou, E. (2023). Forecast vs. Order Adjustments in Inventory Systems, Informs 2023, Phoenix, Arizona, USA, 14-18 Oct 2023.
Sivathanu, B., & Pillai, R. (2019). Technology and talent analytics for talent management – A game changer for organizational performance. *International Journal of Organizational Analysis, 28*(2), 457–473.
Šikšnys, L., & Pedersen, T. B. (2016). Prescriptive analytics. In L. Liu, & M. Özsu (Eds.), *Encyclopedia of database systems*. New York, NY: Springer.
Smyl, S., Ranganathan, J. & Pasqua, A., (2018). *M4 Forecasting Competition: Introducing a New Hybrid ES-RNN Model.* Available at: https://eng.uber.com/m4-forecasting-competition/ (Accessed: 13 07 2018).
Syntetos, A. A., Babai, M. Z., Boylan, J. E., Kolassa, S., & Nikolopoulos, K. (2016). *European Journal of Operational Research, 252*(1), 1–26.
Tashman, L. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. *International Journal of Forecasting, 6*, 437–450.
Theofilatos, A., Yannis, G., Kopelias, P. & Papadimitriou, F., (2017). 'Impact of real-time traffic characteristics on crash occurrence: Preliminary results of the case of rare events'.
Van der Vlist, F. N. (2016). Accounting for the Social: Investigating Commensuration and Big Data Practices at Facebook. *Big Data & Society, 3*(1), 1–16.
Wamba, S. F., Dubey, R., Gunasekaran, A., & Akter, S. (2019). The performance effects of big data analytics and supply chain ambidexterity: The moderating effect of environmental dynamism. International Journal of Production Economics.
Wauters, M., & Vanhoucke, M. (2017). A Nearest Neighbour extension to project duration forecasting with Artificial Intelligence. *European Journal of Operational Research, 259*, 1097–1111.
Webby, R., & O'Connor, M. (1996). Judgemental and statistical time series forecasting: A review of the literature. *International Journal of Forecasting, 12*, 91–118.
Werner, C., Bedford, T., Cooke, R., Hanea, A., & Napole, A. (2017). Expert judgement for dependence in probabilistic modelling: A systematic literature review and future research directions. *European Journal of Operational Research, 258*, 801–819.
Yang, H., Roe, B., & Zhu, J. (2007). Studies of stability and robustness for artificial neural networks and boosted decision trees. *Nuclear Instruments and Methods in Physics Research A, 574*, 342–349.
Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. *Transportation Research Part C, 58*, 308–324.
Zhang, Y., Zhang, M., Li, J., Liu, G., Yang, M., & Liu, S. (2021). A bibliometric review of a decade of research: Big data in business research – Setting a research agenda. *'Journal of Business Research', 131*, 374–390.

A. Schaefers Andreas was a Master student in the programme of Business Analytics in Alliance Manchester Business School and working now full time in Switzerland.

V. Bougioukos Vasileios is an Adjunct Associate Professor at Notre Dame, London and a Research Associate in Durham Business School.

G. Karamatzanis Georgios is an Assistant Professor in Strategy and Entrepreneurship in Norwich Business School in the University of East Anglia.

K. Nikolopoulos Kostas is the Professor and Chair of Business Information Systems & Analytics in Durham University Business School and the Director of the IHRR Forecasting Lab in Durham University.