Research-Technology Management

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/urtm20

Forecasting New Product Demand Using Domain Knowledge and Machine Learning

A proposed method uses machine learning and an expert's domain knowledge to enhance the accuracy of new product predictions.

Charles Lincoln Kenji Yamamura, José Carlos Curvelo Santana, Bruno Sanches Masiero, José Alberto Quintanilha & Fernando Tobal Berssaneti

To cite this article: Charles Lincoln Kenji Yamamura, José Carlos Curvelo Santana, Bruno Sanches Masiero, José Alberto Quintanilha & Fernando Tobal Berssaneti (2022) Forecasting New Product Demand Using Domain Knowledge and Machine Learning, Research-Technology Management, 65:4, 27-36, DOI: 10.1080/08956308.2022.2062553

To link to this article: https://doi.org/10.1080/08956308.2022.2062553

Published online: 27 Jun 2022.

Forecasting New Product Demand Using Domain Knowledge and Machine Learning

A proposed method uses machine learning and an expert's domain knowledge to enhance the accuracy of new product predictions.

Charles Lincoln Kenji Yamamura, José Carlos Curvelo Santana, Bruno Sanches Masiero, José Alberto Quintanilha, and Fernando Tobal Berssaneti

OVERVIEW: Forecasting demand for new products is a challenging task, as it involves capturing relations of complex variables in markets where little or no historical data exist. Managers usually rely on surveys, intuition, and heuristics to forecast new products. Linear statistical tools used to predict demand for existing products are not suitable because there are not enough data to capture complex nonlinear relations in yet-to-be launched products. Other tools are appropriate for aggregate new categories but not for incremental company-specific products. Machine learning can capture complex nonlinear relations, but it usually requires significant amounts of data. Using an expert's domain knowledge can circumvent the need for vast training datasets. To support product development activities, we propose a method that combines domain knowledge and machine learning to forecast market share of complex incremental new products. An experiment from the automobile industry shows the approach yields expressive results (82 percent forecast accuracy).

KEYWORDS: New product forecasting; Domain knowledge; Machine learning; Product concept; Artificial neural network

For the last 50 years, automobiles have experienced rapid incremental innovations—in energy efficiency, costs, safety, performance, comfort, convenience—that have kept them from experiencing an industry life-cycle decline (Fujimoto 2014). It takes three to five years to develop a new automobile (Ulrich, Eppinger, and Yang 2020), and it must remain competitive in the market for six to eight years after it is launched. Rather than reacting to customer demands,

Charles Yamamura has over 20 years' experience in strategic planning and product development in the automobile industry. He is a former chief portfolio strategist at Fiat Chrysler, technical director at JAC Motors, product planning manager at Nissan and Yamaha, marketing manager at Volkswagen, and product planner at Toyota. He is currently completing a PhD in industrial engineering at the Polytechnic School of the University of São Paulo, Brazil. He holds an MBA from Yale University and an MSc from the Getulio Vargas Foundation. His research focuses on vehicle electrification, machine learning applied to product strategy, innovation, and small data. charles.yamamura@usp.br

José Santana is a professor of industrial engineering at the Federal University of ABC in Brazil and a researcher at the Polytechnic School of the University of Sao Paulo, Brazil. He holds MSc and PhD degrees in chemical engineering from the State University of Campinas, and a bachelor's degree in industrial chemistry at the Federal University of Sergipe. He has experience in the fields of industrial and chemical engineering, focusing on biochemical and chemical processes, effluent treatments, sustainability development process simulation and optimization, statistical control of quality, factorial planning, accreditation of methods, quality assurance, product planning, and development. jccurvelo@yahoo.com.br

Bruno Masiero is a professor of electrical and computing engineering at the University of Campinas, Brazil. He has a bachelor's degree in electrical engineering and an MSc in electrical engineering from the University of São Paulo, and a PhD in electrical engineering from the RWTH Aachen University, Germany. His research focuses on the application of modern digital signal processing techniques in several audio and acoustic applications, such as spatial sound acquisition and reproduction, acoustic imaging, characterization of acoustic materials, and development of audiological assessment tools. He is a member of the board of the International Commission for Acoustics. masiero@unicamp.br

José Quintanilha is a senior associate professor at the Energy and Environment Institute of the University of São Paulo, Brazil. He has a bachelor's degree in statistics from the Institute of Mathematics and Statistics of the University of São Paulo, an MSc in remote sensing and geoprocessing from the National Institute of Space Research, and PhD in transportation engineering from the University of São Paulo. He has experience in geostatistical analysis, with emphasis on environment, city, and transport applications. jaquinta@usp.br

Fernando Berssaneti is a professor of industrial engineering at the Polytechnic School of the University of São Paulo, Brazil, and a counseling member of the Vanzolini Foundation. He holds a bachelor's degree, a MSc, and a PhD in industrial engineering from the Polytechnic School of the University of São Paulo. He is also a business consultant in industrial engineering, with emphasis on product and innovation management, mobility, and sustainability. He has several articles and books published in those fields. fernando.berssaneti@usp.br

DOI: 10.1080/08956308.2022.2062553

automotive companies must anticipate what customers want, usually with little or no historical data, since the new product is not in the market yet (Kahn 2002).

Given the scarcity of both data and adequate analytical tools to forecast demand for yet-to-be developed products, managers usually rely on surveys, intuition, and heuristics to make decisions. The decision-maker's knowledge and perception of the business environment (customers, competitors, technology) and expected changes over time (Dane and Pratt 2007) drive intuition. Intuitive decision-making strives to extract meaning from available information, but that process is usually vague and subjective (Spangler 1991).

Among the existing analytical methods, linear models are easy to interpret and explain, but they rarely capture the complexity of real-world problems with accuracy (Shrestha et al. 2020). Some analytical approaches rely on diffusion theory (Linton 2002; Ching-Chin et al. 2010; Lee et al. 2014; Yin et al. 2020) from microeconomics. For instance, a Bass model contemplates the initial stages of new product adoption by a population (Linton 2002). Although diffusion models are applicable to aggregate product categories, they are not adequate for incremental company-specific products (Kahn 2002), which are the scope of this study.

Machine learning algorithms can capture patterns that are difficult for humans to delineate mentally or heuristically. But those algorithms are data driven (Yu 2007; Yu, Simoff, and Jan 2010) and require significantly large amounts of training data. Lack of sufficient data is a major challenge in data analytics (Ng 2016; Yang et al. 2019). This study proposes using the manager's domain knowledge to partially mitigate the lack of data for machine learning simulations (Vapnik and Izmailov 2019, 2020). Domain knowledge can accelerate the development of solutions.

Our study focuses on demand forecasting of complex mass-produced products, like automobiles, mobile phones, and microcomputers. Complex products have multiple feature variables with nonlinear relations (Hobday 1998). We present a method practitioners can use to predict market share of new product concepts (sets of features), and thereby support and enhance new product development (NPD). Our method combines domain knowledge and machine learning, precluding the need for vast amounts of training data. The domain expert (manager or market specialist) creates a database using their knowledge and some competitor sales data. An artificial neural network processes the database, simulating the market by capturing the nonlinear relations among the variables, which are expressed in the weights of the connections in the stable machine learning algorithm (artificial neural network models of the market).

Our proposed method uses machine learning and domain knowledge to enhance the accuracy of new product predictions.

To demonstrate our method, we conducted an experiment to predict the market share of small sport utility vehicles (SUVs) in the Brazilian market. The experiment shows how a machine learning algorithm is useful to support NPD. The trained artificial neural network captured the complex nonlinear relations in the variables, producing meaningful results: a forecast accuracy of 82 percent, compared to the standard accuracy of 58 percent achieved by most firms (Kahn 2002, 2010). The SUV dataset contains 1,563 observations and we achieved a correlation factor (R) of 0.96 between predictions and real market share. We conducted an additional smaller scale experiment on medium-size pickup trucks. Machine learning usually requires tens of thousands of observations to capture patterns in data (Allen 2019).

In our study, the combination of machine learning and domain knowledge was able to reduce data size requirements to a much smaller number of observations (Shrestha et al. 2020) while getting superior results.

New Product Forecasting

A product is a bundle of features (attributes and characteristics) that customers perceive as having "value"—that is, which satisfy their needs and aspirations (Woodruff 1997; Ulrich, Eppinger, and Yang 2020). A product concept is a set of the most relevant customer features as interpreted by a company. Effective product decision-making implies good forecasting (Weber 2009). Successful new product forecasting predicts the most likely outcome of a product concept given a set of assumptions (Kahn 2006).

Companies use statistical techniques, like time series and regression methods, to forecast existing products. By contrast, because little or no historical data exist (Kahn 2014), new product forecasting is predominantly judgmental and seeks meaningfulness as the performance criterion, instead of statistical accuracy. The forecasting process for new products relies on surveys, managerial judgment, and sometimes on extrapolations made by assessing similarities with competitor products (Sharma et al. 2020). A meaningful forecast can be used for decision-making because it provides deeper understanding of the problem (Kahn 2002). Due to the lack of data, new product forecasting usually aims to produce realistic and operationally useful approximations in predictions rather than numerical precision. According to Kahn (2002), new product forecasts do not seek numerical precision because usually it is not obtainable due to the scarcity of data. The forecasts are operational approximations that permit understanding and judgment of a problem. In other words, managers know predictions are numerically "wrong," but they can be good enough to support understanding and decision-making, with a reasonable approximation. Our proposed method aims to improve the accuracy of new product forecasting. It uses machine learning and domain knowledge to enhance the accuracy of new product predictions.

Kahn (2014) identifies seven possible new product categories: cost reduction, product improvement (replaces an existing product), product line extension (addition to the existing product line), new usage, new (geographical) market, "new-to-the-company" products, and "new-to-the-world" products. Products that are new to the world are also called radical products, as they disrupt existing industries, technologies, and customer perceptions of the market (Henderson and Clark 1990; Rice et al. 1998; Kahn 2018; Veryzer 1998; Leifer, O'Connor, and Rice 2001). We do not consider radical products in this analysis. The other six categories fall within incremental or continuous product innovation (Veryzer 1998). Note that product innovation requires the convergence of both technical capabilities and customer perceptions to succeed (Veryzer 1998; Leifer, O'Connor, and Rice 2001), and that modest technical innovations may sometimes produce significant competitive advantage (Henderson and Clark 1990).

Companies usually need to balance a mix of incremental and radical new products for a healthy business (Rice et al. 1998; Leifer, O'Connor, and Rice 2001; Veryzer 1998; Kahn 2018). Forecasting becomes harder and less accurate as products move from the incremental to radical scale, and applied techniques tend to shift from statistical to intuitive (Kahn 2014). Machine learning performs best when simulating relatively stable phenomena (Shrestha et al. 2020). Our proposed method is not suitable to predict radical products, as no data exist to extract patterns and trends. But "new-to-the-world" products are rare (Ching-Chin et al. 2010) and forecasting them is still an exercise of vision and brainstorming (Linton 2002).

Kahn's (2002, 2014) survey of 168 industrial companies found that average prediction accuracies are as follows: cost reduction (72 percent), improvement (65 percent), line extension (63 percent), market extension (54 percent), new to the company (47 percent), and new to the world (40 percent). The average accuracy was 58 percent (Kahn 2002, 2010; Armstrong 2002), but the benchmark for an acceptable level should be around 76 percent (Kahn 2010; Ching-Chin et al. 2010).

Machine Learning

A machine learning algorithm is a computer program that improves its performance with data inputs, without being explicitly programmed (Mitchell 1997; Goodfellow, Bengio, and Courville 2016; Yu 2007; Shrestha et al. 2020). It distinguishes itself from conventional algorithms by its capacity to learn or to improve with new data. Most current achievements in machine learning are of the *supervised learning* type, whereby algorithms improve by processing *labeled* data (Ng 2018; Goodfellow, Bengio, and Courville 2016).

An artificial neural network is a machine learning algorithm that maps functions of complex feature variables to target values in nonlinear hierarchical relations (Goodfellow, Bengio, and Courville 2016; Aggarwal 2018). Machine learning methods with multiple levels of representation are also called *deep learning* methods (LeCun, Bengio, and Hinton 2015). Data are processed by inputting feature data and adjusting the weights in the connections through layers of hidden nodes. The artificial neural network calculates errors by successive forward and backward iterations, with the goal of minimizing the derivatives of errors (Figure 1). As data move into higher layers of representation, the algorithm processes more abstract relations through the network (Aggarwal 2018).

Domain Knowledge

People possess knowledge in the forms of facts, ideas, perceptions, and processes that enable them to take effective action (Alavi and Leidner 2001). Domain knowledge is the knowledge pertaining to a particular specialized field, as opposed to general knowledge (Yu 2007). In this study, domain knowledge is the sum of the manager's experience and perceptions about customers, technologies, and competitors related to certain markets and products. Domain

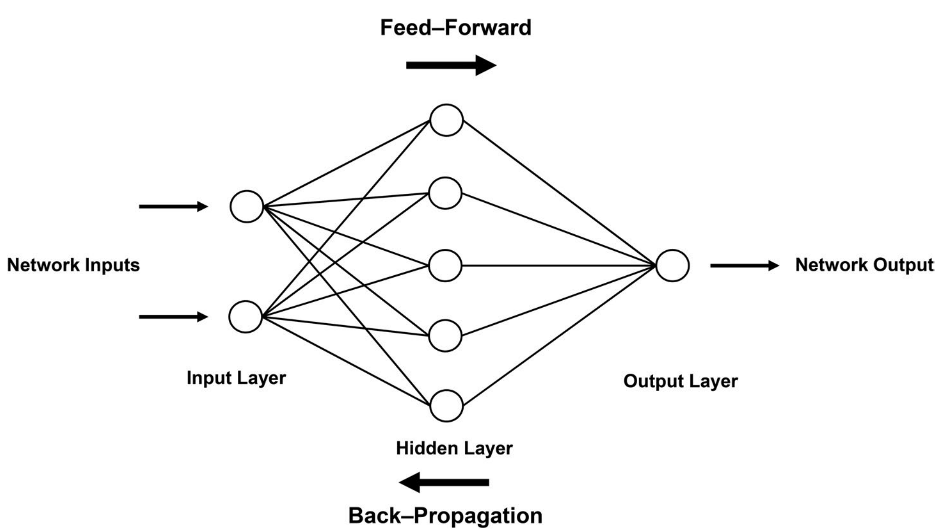

FIGURE 1. An artificial neural network

knowledge is usually informal and poorly structured (Yu 2007; Yu, Simoff, and Jan 2010).

The initial conditions, assumptions, and context constitute the invariants of a problem (Gil et al. 2019). Invariants do not change when an object suffers a transformation (Aleksandrov, Kolmogorov, and Lavrent'ev 1999). In the product design context, invariants are the variables the domain expert takes for granted in a problem—they assume those variables have fixed values in a given context—and does not revisit them every time they handle that problem. Isolating invariants is a difficult task and requires an expert's skill and ingenuity (Wigner 1949, 1960). Due to their use of intelligence, human learning requires far fewer examples compared to machine learning, which uses brute force in data analytics (Vapnik and Izmailov 2019, 2020). Many potential variables in a problem are irrelevant to a task at hand. Based on their knowledge and experience, experts identify and put aside the invariants latent in the problem, and focus on the relevant variables. Vapnik and Izmailov (2019) exemplify jocosely with the "duck test": "If it looks like a duck, swims like a duck and quacks like a duck, it probably is a duck"—just three features to qualify a duck, instead of many possible variables that characterize birds in general.

As an industry matures, innovation becomes predominantly incremental and experts accumulate knowledge on solving specific problems that emerge from established designs and architectures, which are not revisited every time a new task arises. Innovation occurs at the component level, not in the overall design (the way components are coordinated as a system) (Henderson and Clark 1990). Fixed features in a product design are examples of invariants, although they are not the whole set of invariants; the domain expert skims the invariants further by crafting the partial product hypothesis.

An increase in invariants diminishes the size of training data and leads to more accurate predictions, while an increase in variables requires more training data (Vapnik and Izmailov 2019, 2020).

Proposed Forecasting Method

Conventional data analytics is data driven, and an algorithm discovers patterns in that data (Yu 2007; Yu, Simoff, and Jan 2010). The domain expert only provides project guidelines and interprets the results from the trained model. The domain expert formulates a market hypothesis only after reviewing simulation results (Figure 2a). In the proposed method, the manager is involved in the entire process, including the formulation of an initial hypothesis and synthesizing data, not only in inference and judgment (Shrestha et al. 2020). A partial hypothesis using the manager's domain knowledge precedes the simulation (Figure 2b).

The proposed forecasting tool is based on algorithmic machine learning simulation guided by an expert's domain knowledge, which they use to formulate the initial hypothesis and to preprocess the training examples (Yu 2007; Yu, Simoff, and Jan 2010). The initial hypothesis comes from the domain expert's knowledge and experience related to a specific market. Preprocessing involves selecting, cleaning, formatting, and synthesizing the training data. The simulation refines the hypothesis and converges to an optimal model by capturing the relations (weights) among multiple nonlinear variables. The stable set of variables and weights in the algorithm is used to predict the market share of new product concepts (feature sets). Domain knowledge guides the algorithmic training process, improving its performance, and making it more explainable (Yang et al. 2019).

The quality of new product forecasting depends on the quality of assumptions—coming from managerial judgment—and the quality of data (Kahn 2014). The initial hypothesis is a vector (set) of the relevant product features (Simon 1996) from the customer point-of-view (Woodruff 1997); the vector contains the variables (customer features) of the machine learning model. A company usually relies on staff expertise, market research, static and dynamic testing, competitor benchmarking, distributor feedback, literature, and other sources to elaborate the hypothesis.

A manufacturer of complex products handles and defines hundreds (sometimes thousands) of feature variables, as they are elements of the overall product design and development (Weber 2009). But fewer variables comprise the key determinants of customer value (Weber 2009), and too many redundant features introduce noise and can harm the machine learning algorithm's performance (Yu, Simoff, and

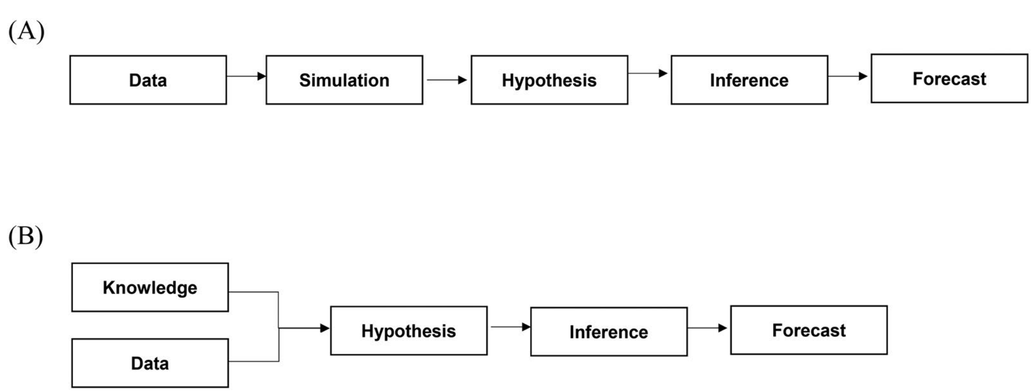

FIGURE 2. New product forecasting: a) conventional data analytics; b) proposal integrating domain knowledge and machine learning

Jan 2010). Identifying a compact set of relevant features to model the simulation is a challenge and prior domain knowledge is required to set the partial hypothesis (initial hypothesis space). Features that get left out are invariants in the model. The set of feature variables is a "partial" hypothesis, but it still misses important information—namely, weights estimated by the machine learning algorithm. The simulation converges the hypothesis space (potentially infinite sets of possible weights) to an optimum hypothesis (a stable set of variables and weights in the trained model).

Database construction starts with the hypothesis of product features. The domain expert collects data on company and competitor product sales and converts them to market share. A relatively modest amount (in the order of a few hundred observations) of sales data are sufficient. The expert synthesizes the sample by evaluating the features of each product in the dataset. A machine learning algorithm (an artificial neural network) processes the features and target values (market share calculated from sales), capturing patterns and estimating the weights of the variables. Then, the domain expert can use the trained model to forecast the market share of potential new products.

The domain expert can use the trained model to forecast the market share of potential new products.

Experiment Using Data from the Brazilian SUV Market

To demonstrate the method, we present an experiment on small SUVs in the Brazilian market. Although no consensus exists globally regarding the definition of an SUV, it refers to a closed passenger vehicle with a "two-box" styling bodywork, which is taller than a conventional automobile. By "small," we mean vehicles shorter than 4.5 meters (177 in) in overall length.

In this small-scale experiment, we formulated a hypothesis of key product features using our industry knowledge and experience and information compiled from specialized magazines and websites (Quatro Rodas 2020a; UOL Carros 2020). After we formulated our hypothesis, we compared it to Weber (2009), who provides a list of primary customer-relevant characteristics, which we aligned with our hypothesis. We identified 16 primary product attributes: novelty, brand, style, robustness, comfort, space, trunk, convenience, finish, equipment, infotainment, economy, performance, agility, safety, and price. Although Weber's (2009) manual is generic and not specific to the Brazilian market, 14 out of 16 of the variables (87 percent) align with our experimental hypothesis. In reality, a large manufacturer would rely on multiple resources, including the experience of its executives and staff, market research, product testing, distributor surveys, and competitor benchmarking. The effective accuracy test comes from the algorithmic test (correlations) and the experiment results.

We collected monthly sales data of 22 small SUV products—from March 2003 to December 2020—from the National Federation of Brazilian Vehicle Distributors website (FENABRAVE 2021) and converted the data to market share. The dataset covers the entire small SUV history in Brazil, and it contains 1,563 observations (Yamamura 2020). We used market share to mitigate seasonal effects. As companies regularly forecast overall market volumes, it is easy to convert market share forecasts back into sales figures.

We used reviews from an automobile magazine (Quatro Rodas 2020a) and customer evaluations from a specialized website (UOL Carros 2020) to measure product features. Some vectors (sets of features) are repetitive; their feature values are constant over a certain period of time (like a few months). However, other vectors differ for the same product at different times in its product life. For example, *novelty* decreases as time passes: a new product scores five for *novelty* in its launch year, but one point is deducted every year after. We based *brand strength* on a survey from a Brazilian automobile magazine (Quatro Rodas 2020b). Although some variables are quantitative—like price, fuel economy, and performance—others are subjective, like brand and style. To normalize all features, we evaluated them using a one-to-five point Likert scale, a level of granularity that was sufficient to capture relations in the database (Table 1).

We split the whole dataset in 70:15:15 ratios among training, test, and validation sets—shuffled, drawn stochastically, and set apart from the beginning. The artificial neural network input layer contains 16 "neurons," corresponding to the feature variables. The output layer has just one node, predicting the target value. The model in the experiment presents a single hidden layer, adequate for a relatively small dataset (Goodfellow, Bengio, and Courville 2016).

No consensus exists regarding an appropriate number of units in a hidden layer. A common heuristic is between half and twice the input layer size (between n/2 and 2n), from 8 to 32 in the experiment. The neural network achieved the highest correlation between real and predicted data ($R = 0.96$) with 10 hidden units. The algorithm estimates the weights of hidden and output layer neurons using *activation* functions. The model optimizes the weights using a function that minimizes the output (market share estimation) errors (Goodfellow, Bengio, and Courville 2016; Aggarwal 2018). We ran the algorithm on MATLAB R2019b, in a MacBook with 1.2 GHz Intel Core M processor.

We tested the method using real data, predicting market share of the six most significant (largest sales volume) new small SUVs launched in Brazil between 2016 and 2020. We used mean absolute percent error (MAPE) to measure the accuracy of those predictions. MAPE is an indicator commonly used in companies to measure the accuracy of predictions (Ching-Chin et al. 2010). MAPE is calculated by subtracting actual demand from the forecast, dividing by

TABLE 1. A sample section of the database

Novel	Brand	Style	Robust	Comf	Space	Trunk	Conv	Finish	Equip	Info	Econ	Perf	Agile	Safe	Price	Share
3	5	5	5	5	4	3	4	5	5	5	1	3	4	4	1	3.16
1	5	5	5	4	2	1	4	5	4	5	1	1	3	4	3	3.00
3	2	5	4	5	5	5	5	5	3	1	1	4	3	5	2	0.08
4	3	5	3	5	4	5	5	5	2	3	1	5	3	4	1	0.13
2	3	3	2	1	2	4	2	2	2	1	4	4	4	3	4	1.85
2	1	1	1	3	1	2	1	1	3	3	3	1	4	3	4	0.31
2	2	4	5	1	5	4	2	1	2	3	1	1	2	3	4	0.44
3	2	2	5	2	5	5	2	2	2	4	1	2	2	2	5	0.69
5	4	4	4	4	4	2	4	4	5	5	4	5	5	5	3	3.72
5	4	5	3	3	2	3	3	4	5	5	4	5	5	5	4	1.24

Note: Novel = novelty; Robust = robustness; Comf = comfort; Space = cabin space; Trunk = trunk space; Conv = convenience; Equip = equipment; Info = infotainment; Econ = fuel economy; Perf = performance; Agile = agility; Safe = safety; Share = market share.

actual demand, taking its absolute value, and multiplying by one hundred:

$$MAPE = \left| \frac{fcst - actl}{actl} \right| * 100$$

Overfitting tends to be a major issue in machine learning algorithms (Shrestha et al. 2020). It consists of low bias but high variance in errors. In other words, an algorithm could perform well on the dataset used for training but struggle to generalize to new data (Shrestha et al. 2020). Increasing the size of the sample or decreasing the size of the model (number of variables) are procedures used to handle overfitting. Machine learning algorithms have built-in procedures like regularization (penalization of variables with small effect) and cross-validation (splitting data in random subsamples for test and validation) to mitigate overfitting (Yu 2007; Yu, Simoff, and Jan 2010). Since overfitting is not an absolute issue, but relative to comparison (Hawkins 2004), it can also be handled by testing, calibrating, and comparing new dataset results, and by interpreting and understanding the circumstances, limitations, and contingencies of the predictions (Hawkins 2004; Ng 2018; Shrestha et al. 2020). We used the aforementioned measures to ensure the number of feature variables and the size of the original dataset were adequate to capture relations in new data. The six new products experiment shows that the model generalizes to new data (not present in the original dataset), and therefore, the model is not overfitting.

The six new products experiment shows that the model generalizes to new data (not present in the original dataset), and therefore, the model is not overfitting.

In our experiment, we limited the number of feature variables (16), using domain knowledge, and made sure the sample size of the original dataset (1,563 observations) was sufficient to generalize to new data. We conducted the MAPE analysis of six real products using data prior to each product's launch time to assess the capacity of the algorithm to generalize to new data.

Results

In the six real products case, we made the predictions with neural networks using the portion of the dataset prior to a new product launch. The results come from artificial neural networks trained only once (not after several trials), which is a reason why the errors exhibit a relatively wide range for new data not in the training set. For instance, the neural network did a remarkable job at predicting product improvement for the Renault Duster, which was likely easier to forecast because it was an improvement and not strictly new. In other words, the product improvement was an enhanced version of an existing product, and it had some feature values that were common to the previous version—that is, present in the training dataset. The neural network did a relatively poor job of predicting the Renault Captur, a completely new SUV introduced into the Brazilian market. The Captur had more new feature values.

Still the model presented meaningful predictions overall, being able to fit new data not present in the training dataset. The estimated MAPE ranged from 2 percent to 48 percent—with an average of 18 percent or accuracy of 82 percent (Table 2). As mentioned previously, Kahn (2010, 2002) suggested the average forecasting accuracy commonly achieved by firms is 58 percent. Ching-Chin et al. (2010) suggest an adequate performance should be about 76 percent. Also as mentioned, overfitting is relative to understanding the circumstances of the prediction. In this case, we used a benchmark from the literature to reference and to confirm the model generalizes to new data.

Second Experiment: Brazil's Medium-Size Pickup Market

For confirmation, we conducted an additional experiment on Brazil's medium-size pickup market. The model is less elaborate as there were some limitations. For example, there is a smaller number of players (seven) and product life-cycles tend to be longer than the small SUV segment—more than 10 years in some cases, like the Volkswagen Amarok, which was launched in 2010 and is still in the market.

The dataset contains the history of medium-size pickup trucks between January 2015 and December 2020, with 491 observations. We identified seven key feature variables—brand, robustness, style, engine power, cabin, price, and size—once again based on industry experience from the authors and validation from specialized press documents and Weber's textbook (2009). We present a database extract (Table 3).

There were no recent product launches in the segment, but historical data include product facelifts, content changes, and engine improvements. Since there were no "new-to-the-company" products in the given time period, we used data from three consecutive years to predict the market share in the following year. For instance, we predicted products in 2020 with a neural network trained with data from 2017 to 2019. The average forecast accuracy yielded was 78 percent (Table 4), confirming useful predictive power.

Discussion

By applying our method, practitioners can realize several benefits.

> **Our method demonstrates it is possible to forecast market share of incremental but complex new products using expert domain knowledge and small datasets.**

New Product Forecasting with Small Data—Our method demonstrates it is possible to forecast market share of incremental but complex new products using expert domain knowledge and small datasets. The manager uses domain knowledge to set the initial conditions for the product hypothesis, by pre-selecting the relevant product feature variables (and isolating the invariants). The machine learning algorithm simulates the market and estimates the weights of feature variables. The proposed method builds on strengths of the managerial profession—market knowledge and experience—to enhance NPD. It sharpens decision-making, as good decisions imply good forecasting.

By providing understanding and reducing uncertainty (Ganguly and Euchner 2018), our method is also a learning tool. As new knowledge becomes available or market conditions change, the model needs to be updated and improved. For instance, market, technology, or legislation may dictate the emergence of new relevant customer features, like

TABLE 2. Six new product market share predictions using artificial neural networks trained with data prior to launch date

Make	Model	Launch Date	Mkt Share (%)	Prediction	MAPE	1 – MAPE
Nissan	Kicks	Aug '16	1.254	1.312	4.6%	95.4%
Hyundai	Creta	Feb '17	1.952	1.879	3.7%	96.3%
Renault	Captur	Apr '17	0.732	1.080	47.5%	52.5%
Volkswagen	T Cross	Aug '18	2.806	3.526	25.7%	74.3%
Chevrolet	New Tracker	Apr '20	3.175	3.928	23.7%	76.3%
Renault	Duster facelift	Apr '20	1.377	1.350	2.0%	98.0%
Average					**17.9%**	**82.1%**

TABLE 3. A section of the pickup database

Brand	Robust	Style	Power	Cabin	Price	Size	Mkt Share
4	4	4	4	3	3	5	1.38
2	1	3	2	2	5	3	3.00
4	3	4	3	4	2	5	1.03
4	4	4	3	3	2	5	0.30
3	4	3	3	2	2	5	0.41
5	5	4	4	3	1	5	1.42
3	2	3	5	3	1	5	0.39

Note: Robust = vehicle robustness; Power = engine power; Cabin = cabin design.

TABLE 4. Prediction of market share using artificial neural networks trained with data from three previous years

		2020				2019				2018			
Make	Model	Mkt %	Pred	MAPE	1–MAPE	Mkt %	Pred	MAPE	1–MAPE	Mkt %	Pred	MAPE	1–MAPE
Chevrolet	S10	1.02	1.26	23.6%	76.4%	1.21	1.88	55.8%	44.2%	1.29	1.33	3.3%	96.7%
Fiat	Toro	2.77	2.40	13.3%	86.7%	2.47	2.28	7.5%	92.5%	2.37	2.48	4.6%	95.4%
Ford	Ranger	1.02	0.81	20.2%	79.8%	0.84	0.83	1.3%	98.7%	0.83	0.66	21.1%	78.9%
Mitsubishi	L200	0.49	0.42	13.5%	86.5%	0.38	0.46	18.8%	81.2%	0.44	0.39	9.9%	90.1%
Nissan	Frontier	0.41	0.31	25.2%	74.8%	0.30	0.06	80.4%	19.6%	0.26	0.18	28.3%	71.7%
Toyota	Hilux	1.66	1.55	6.6%	93.4%	1.52	1.67	9.7%	90.3%	1.59	1.57	1.1%	98.9%
Volkswagen	Amarok	0.54	0.78	44.2%	55.8%	0.71	0.76	7.3%	92.7%	0.76	1.35	77.5%	22.5%
Average				20.9%	79.1%			25.8%	74.2%			20.8%	79.2%
										Overall Average		22.5%	77.5%

powertrain electrification (hybridization), barely present in current small SUVs in Brazil. The diffusion of that technology will require the introduction of a new feature variable in the database and hence in future algorithmic models. On the other hand, "infotainment" technology may become so common and undifferentiated that customers simply take it for granted and it is no longer a relevant variable for the method. In that case, it becomes an invariant and it is no longer explicit in the model. Those changes will demand both redesigning the database and retraining the algorithm, which simultaneously generates new learning for decision makers and raises product development performance.

Constructing the Hypothesis with Domain Knowledge—Constructing an appropriate set of features to frame the new product concept is a major challenge. Regrettably, there is no established formula to do it. As Wigner (1949) suggests, setting the initial conditions of an experiment is an art. Vapnik (2019) suggests it is the intelligence part of learning. But experienced managers have competence, understanding, and insight regarding their businesses. They have knowledge and intuition of what is relevant in those markets, and their skills are routinely used in their work, albeit not always in a structured way.

The initial hypothesis is a set of the most relevant feature variables in the market under consideration. The domain expert (manager) creates a hypothesis by separating key feature variables from invariants (unchanging features) in the pool of potentially infinite variables. The compact set of feature variables left is the product concept to be modeled and refined. The model is complete when the algorithm approximates the weights of the variables and is stable. To avoid potential bias, it is important that the domain expert validates the partial hypothesis using surveys, press information, literature, and maybe a second expert's opinion.

Forecasting New Products Designs—The trained algorithm, with a fixed set of variables and weights, models the market to be investigated for NPD. New product concepts—vectors with evaluated features—can be inputted to predict market share. Our method can test both alternative company product proposals and potential competitor responses. Market experts can predict a new product demand by introducing a new feature vector in the trained algorithm. They may feed alternative variable sets (product concepts), testing their market sensitivity. Or they can input a set of features from a hypothetical competitor product, based on competitive intelligence, predicting its potential market strength and guiding the company's future preemptive actions.

Our proposed method is robust and reproducible for incremental new products in complex but relatively stable markets. However, we do not recommend it to forecast "new-to-the-world" products because as the machine learning algorithm identifies nonlinear patterns in data, it assumes probability distributions that are constant over time, which is not the case for radical products.

Domain and Data Experts—Usually, domain and data experts are different people (Crews 2019; Amershi et al. 2014). Pure managerial judgment relies on intuition and experience, and a data expert is usually not involved. In data analytics, however, the algorithm is responsible for discovering patterns. The domain expert (manager) participates initially to provide general requirements for the project, and returns later in the process, when the model is ready. The domain expert's involvement is collateral and subsidiary (Amershi et al. 2014).

Our proposed approach requires that the domain expert and data scientist work closely throughout the process; otherwise, the domain specialist must possess some systems knowledge to handle both database and algorithm development. The domain expert is responsible for creating the initial product hypothesis (set of feature variables) and guiding data synthesis. Once the model is trained, the domain expert will interpret the results and explain them for NPD action.

With recent commercially available software, the manager does not need to be a full-fledged data expert (Gil et al. 2019; Amershi et al. 2014). In the next few years, using machine learning to support business decisions may become as common as manipulating spreadsheets and databases is today (Allen 2019; Euchner 2019; Daugherty and Euchner 2020).

Limitations

One limitation is that this study is a single-industry one. The logic of our proposed method—that is, machine learning algorithms just read numbers and capture relations among them—is extensible to other sectors and contexts. Still, studies exploring the application in different domains could be valuable contributions. Also, further research exploring alternative machine learning techniques, such as regression support vector machine (SVM) or extreme learning machine (ELM), to new product strategy is a promising venue.

Conclusion

New product forecasting supported by machine learning is an emerging and fertile field of investigation. Our proposed forecasting method, which uses a machine learning algorithm combined with domain knowledge, can yield meaningful predictions. Our method precludes the need for huge datasets, which is important for NPD where little or no historical data exist. Narrowing the range of potential research variables increases efficiency and speed. Our method also reduces training data size and trades data quantity for quality. Our model explains and predicts, thereby supporting NPD and creating knowledge that can be abstracted and transferred to other projects.

This work was supported by the Coordination of Superior Level Staff Improvement of the Brazilian Ministry of Education (CAPES) and the Brazilian National Council for Scientific and Technological Development (CNPq).

References

Aggarwal, C. C. 2018. *Neural Networks and Deep Learning: A Textbook*. New York: Springer.

Alavi, M., and Leidner, D. E. 2001. Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly* 25(1): 07–136. doi:10.2307/3250961

Aleksandrov, A. D., Kolmogorov, A. N., and Lavrent'ev, M. A. 1999. *Mathematics: Its Contents, Methods and Meaning*. New York: Dover Publications.

Allen, R. 2019. Five lessons for applying machine learning. *Research-Technology Management* 62(3): 38–44. doi:10.1080/08956308.2019.1587330

Amershi, S., Cakmak, M., Knox, W. B., and Kulesza, T. 2014. Power to the people: the role of humans in interactive machine learning. *AI Magazine* 35(4): 105–120. doi:10.1609/aimag.v35i4.2513

Armstrong, J. S. 2002. Benchmarks for new product forecast errors. *Journal of Product Innovation Management* 19:133–143. https://www.researchgate.net/profile/J-Armstrong/publication/237409816_Benchmarks_for_New_Product_Forecast_Errors/links/02e7e52814b26c3423000000/Benchmarks-for-New-Product-Forecast-Errors.pdf

Ching-Chin, C., Ieng, A. I. K., Ling-Ling, W., and Ling-Chieh, K. 2010. Designing a decision-support system for new product sales forecasting. *Expert Systems with Applications* 37(2): 1654–1665. doi:10.1016/j.eswa.2009.06.087

Crews, C. 2019. What machine learning can learn from foresight: A human-centered approach: for machine learning–based forecast efforts to succeed, they must embrace lessons from corporate foresight to address human and organizational challenges. *Research-Technology Management* 62(1): 30–33. doi:10.1080/08956308.2019.1541725

Dane, E., and Pratt, M. G. 2007. Exploring intuition and its role in managerial decision making. *Academy of Management Review* 32(1): 33–54. doi:10.5465/amr.2007.23463682

Daugherty, P., and Euchner, J. 2020. Human + machine: collaboration in the age of AI. *Research-Technology Management* 63(2): 12–17. doi:10.1080/08956308.2020.1707001

Euchner, J. 2019. Little ai, big AI–good AI, bad AI. *Research-Technology Management* 62(3): 10–12. doi:10.1080/08956308.2019.1587280

FENABRAVE (National Federation of Brazilian Vehicle Distributors) 2021. New Registrations. http://www.fenabrave.org.br/Portal/conteudo/emplacamentos.

Fujimoto, T. 2014. The long tail of the auto industry life cycle. *Journal of Product Innovation Management* 31(1): 8–16. doi:10.1111/jpim.12076

Ganguly, A., and Euchner, J. 2018. Conducting business experiments: Validating new business models well-designed business experiments can help validate assumptions and reduce risk associated with new business models. *Research-Technology Management* 61(2): 27–36. doi:10.1080/08956308.2018.1421381

Gil, Y., Honaker, J., Gupta, S., Ma, Y., D'Orazio, V., Garijo, D., Gadewar, S., Yang, Q., and Jahanshad, N. 2019. Towards human-guided machine learning. In Proceedings of the 24th International Conference on Intelligent User Interfaces, 614–624. doi:10.1145/3301275.3302324

Goodfellow, I., Bengio, Y., and Courville, A. 2016. *Deep Learning*. Cambridge: MIT Press.

Hawkins, D. M. 2004. The problem of overfitting. *Journal of Chemical Information and Computer Sciences* 44(1): 1–12. doi:10.1021/ci0342472

Henderson, R. M., and Clark, K. B. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* 35(1): 9–30. doi:10.2307/2393549

Hobday, M. 1998. Product complexity, innovation and industrial organisation. *Research Policy* 26(6): 689–710. doi:10.1016/S0048-7333(97)00044-9

Kahn, K. B. 2002. An exploratory investigation of new product forecasting practices. *Journal of Product Innovation Management* 19(2): 133–143. doi:10.1111/1540-5885.1920133

Kahn, K. B. 2006. *New Product Forecasting*. New York: M.E. Sharpe.

Kahn, K. B. 2010. New product forecasting. *Wiley International Encyclopedia of Marketing*. doi:10.1002/9781444316568.wiem05011

Kahn, K. B. 2014. Solving the problems of new product forecasting. *Business Horizons* 57(5): 607–615. doi:10.1016/j.bushor.2014.05.003

Kahn, K. B. 2018. Understanding innovation. *Business Horizons* 61(3): 453–460. doi:10.1016/j.bushor.2018.01.011

LeCun, Y., Bengio, Y., and Hinton, G. 2015. Deep learning. *Nature* 521:436–444. doi:10.1038/nature14539

Lee, H., Kim, S. G., Park, H. W., and Kang, P. 2014. Pre-launch new product demand forecasting using the Bass model: A statistical and machine learning-based approach. *Technological Forecasting and Social Change* 86:49–64. doi:10.1016/j.techfore.2013.08.020

Leifer, R., O'Connor, G. C., and Rice, M. 2001. Implementing radical innovation in mature firms: The role of hubs. *Academy of Management Perspectives* 15(3): 102–113. doi:10.5465/ame.2001.5229646

Linton, J. D. 2002. Forecasting the market diffusion of disruptive and discontinuous innovation. *IEEE Transactions on Engineering Management* 49(4): 365–374. doi:10.1109/TEM.2002.806723

Mitchell, T. M. 1997. *Machine Learning*. New York: McGraw-Hill.

Ng, A. 2016. What artificial intelligence can and can't do right now. Harvard Business Review, November 9. https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now

Ng, A. 2018. Machine learning yearning. https://www.deep-learning.ai/programs/

Quatro Rodas. 2020a. Vehicle tests. https://quatrorodas.abril.com.br/testes

Quatro Rodas. 2020b. Top brands. *Quatro Rodas*, January 2020. 729:69–78.

Rice, M. P., O'Connor, G. C., Peters, L. S., and Morone, J. G. 1998. Managing discontinuous innovation. *Research-Technology Management* 41(3): 52–58. doi:10.1080/08956308.1998.11671210

Sharma, A. K., Goel, N., Rajput, J., and Bilal, M. 2020. An intelligent model for predicting the sales of a product. IEEE. https://ieeexplore.ieee.org/abstract/document/9057880

Shrestha, Y. R., He, V. F., Puranam, P., and von Krogh, G. 2020. Algorithm supported induction for building theory: How can we use prediction models to theorize? *Organization Science* 32(3): 856–880. doi:10.1287/orsc.2020.1382

Simon, H. A. 1996. *The Sciences of the Artificial*. 3rd ed. Cambridge: MIT Press.

Spangler, W. E. 1991. The role of artificial intelligence in understanding the strategic decision-making process. *IEEE Transactions on Knowledge and Data Engineering* 3(2): 149–159. doi:10.1109/69.87995

Ulrich, K. T., Eppinger, S. D., and Yang, M. C. 2020. *Product Design and Development*. New York: Mc Graw-Hill.

UOL Carros 2020. Tests and evaluations. https://autolivraria.com.br/bc/teste-do-leitor/

Vapnik, V. 2019. Complete statistical theory of learning. *Automation and Remote Control* 80(11): 1949–1975. doi:10.1134/S000511791911002X

Vapnik, V., and Izmailov, R. 2019. Rethinking statistical learning theory: learning using statistical invariants. *Machine Learning* 108:381–423. doi:10.1007/s10994-018-5742-0

Vapnik, V., and Izmailov, R. 2020. Complete statistical theory of learning: learning using statistical invariants. *Conformal and Probabilistic Prediction and Applications* 2020 Aug 16: 4–40. http://proceedings.mlr.press/v128/vapnik20a.html

Veryzer, R. W. 1998. Discontinuous innovation and the new product development process. *Journal of Product Innovation Management*, 15(4): 304–321. doi:10.1111/1540-5885.1540304

Weber, J. 2009. *Automotive Development Process*. Berlin: Springer-Verlag.

Wigner, E. P. 1949. Invariance in physical theory. *Proceedings of the American Philosophical Society* 93(7): 521–526. http://www.jstor.org/stable/3143140

Wigner, E. P. 1960. The unreasonable effectiveness of mathematics in the natural sciences. *Communications on Pure and Applied Mathematics* 13:1–14. doi:10.1142/9789814503488_0018

Woodruff, R. B. 1997. Customer value: the next source for competitive advantage. *Journal of the Academy of Marketing Science* 25(2): 139–153. doi:10.1007/BF02894350

Yamamura, C. 2020. Small SUVs Brazil 2003–2020. *Harvard Dataverse*. doi:10.7910/DVN/QTCOJB

Yang, Z., Al-Bahrani, R., Reid, A. C., Papanikolaou, S., Kalidindi, S. R., Liao, W. K., Choudhary, A. and Agrawal, A. 2019. Deep learning based domain knowledge integration for small datasets: Illustrative applications in materials informatics. *2019 International Joint Conference on Neural Networks (IJCNN)*: 1–8. doi:10.1109/IJCNN.2019.8852162

Yin, P., Dou, G., Lin, X., and Liu, L. 2020. A hybrid method for forecasting new product sales based on fuzzy clustering and deep learning. *Kybernetes* 49(120): 3099–3118. doi:10.1108/K-10-2019-0688

Yu, T. 2007. Incorporating prior domain knowledge into inductive machine learning. PhD diss. University of Technology Sidney. https://opus.lib.uts.edu.au/handle/2100/385

Yu, T., Simoff, S., and Jan, T. 2010. VQSVM: a case study for incorporating prior domain knowledge into inductive machine learning. *Neurocomputing* 73(13-15): 2614–2623. doi:10.1016/j.neucom.2010.05.007