



MACHINE LEARNING, INTERNET OF THINGS AND THE FUZZY FRONT END OF PRODUCT DEVELOPMENT

Charles Lincoln Kenji Yamamura (charles.yamamura@gmail.com)

ABSTRACT

Product pre-development or “fuzzy” front end is the set of exploratory activities performed before the decision to invest in a project and to develop it. The purpose of this study is to investigate how the Internet of Things and machine learning are applicable to the front end of product development, in the lean approach. A foundation of lean product development is to generate knowledge that creates customer value. Another is to reduce waste of human, material, monetary and time resources. A method to achieve value generation and resource optimization is to bring forward the technical and market investigations (front loading) from the development process, while there is still tolerance for mistakes and significant changes in the product concept. Time and effort dedicated at this stage will be offset by waste avoided during the development phase. The fuzzy front end is a learning activity. It is paramount to investigate and test all possible hypotheses and alternatives. The Internet of Things, through the diffusion of sensors and transmitters, communication, storage, and data processing networks, will allow access to massive volume of data. Machine learning is the set of algorithms capable of potentializing the learning process from that data.

Key words: machine learning; Internet-of-Things; deep learning; product pre-development; fuzzy front end

Area: Potential of IoT and ICT Solutions on Product and Service Development

1. INTRODUCTION

The world is amidst a revolution in the way products and services are created and delivered, comparable in effect to mechanization, mass production, and automation. The agents of that shift are new technologies in data capture, transmission, storage, and processing. Or sensor technology, the Internet-of-Things (IoT), cloud computing, and machine learning (ML).

A sensor is a device that detects inputs from the environment. By getting smaller, cheaper, more accurate and reliable, it becomes ubiquitous. For instance, a modern smartphone have about fifteen sensors – GSM/CDMA cell, touch screen, two cameras, wi-fi, Bluetooth, NFC (near field connection), GPS, light sensor, accelerometer, gyroscope, proximity, thermometer, barometer and magnetometer. Some still add pedometer (foot step counter), humidity, finger print and even radiation sensors (CHOU, 2016). According to Ian Robertson, CEO of BMW, an Airbus A350 aircraft has 4,000 sensors and generates 2.5 terabytes of data per day. However, a single BMW automobile is capable of creating 4 terabytes of data per day (CAR magazine, June 2017).

Cloud computing is a shared pool of computation resources, including storage and processing. The “cloud” potentializes computing time (speed), space (storage), economy, security, and access.

The Internet-of-Things is the third step in the evolution of world-wide-web, from connected computers to connected people (social networks) and connected everything. It is believed the largest impact will happen in business. Significant changes are expected in both production operations and product development.

Sheer amount of data makes it impossible to transform them into useful information and knowledge without proper support of intelligent algorithms, capable of evolving as more data is received. That is the domain of machine learning, computer programs capable of improving prediction, recommendation and decision performance with experience (data).

This paper focuses on the role of machine learning and the IoT in pre-development, or the fuzzy front end of product development. It is “fuzzy” because it is uncertain, ambiguous and probabilistic in nature, in comparison to the clear, specific and deterministic development process. Front loading means moving forward the investigation and solution of market and technical alternatives and issues, optimizing the whole development process (THOMKE; FUJIMOTO, 2000). The central theme of fuzzy front end is concept design, the process of obtaining an optimal concept through generation and selection (HUANG; BO; CHEN, 2005). A clearly defined and detailed product concept allows ealier and easier identification of issues and solutions.

2. LEAN APPROACH

The lean approach was inspired by the Toyota production system and is a set of principles for efficient decision making and process execution. Its foundations are focus on customer value, waste elimination and continuous improvement. Products are understood as bundles of attributes that are valuable to customers, i.e., satisfy their needs and aspirations, and provide an agreeable experience. So, customer value is the measure of product or service relevance. Every activity that does not add customer value is waste and should be precluded, to optimize resource (labor, time, money, energy, materials) utilization. Development activities are learning processes and - by definition - prone to improvement. Lean approach relentlessly seeks learning to enrich customer value and to eliminate waste. In product development, the

trick is to investigate customer requirements and technical solutions as early as possible, avoiding costly corrections later.

3. THE INTERNET OF THINGS

Although details on telematics (long-distance data transmission) are not publicly disclosed, some automobiles collect diagnostic and operation data - such as system performance, environment, usage patterns and styles - and send them to their manufacturers. Companies like Tesla, BMW, Audi, Ford, and others already use telematics generated data to support product development. The most valuable data help to understand customer value and experience. It is important to discern what data to collect and how to use them to create customer value. Rossman (2016) is adamant about the high risk of spending vast amounts of time and resources collecting data and not being able to use them. From his experience at Amazon, the author recommends to go as far as “to think about a terrible customer experience and how IoT could prevent that”.

4. MACHINE LEARNING

An algorithm is a recipe, a sequence of actions to perform a task like calculating a function. Machine learning originated as a sub-field of artificial intelligence and distinguishes itself from conventional algorithms by the capacity to learn - to improve in performance through experience (data) without being explicitly programmed (MITCHELL, 1997).

A machine learning algorithm is trained with sheer volume of data. If sample examples are labeled ($y = f(x)$), learning is of supervised type. If an algorithm learns patterns by itself, without data tags, it does unsupervised learning. In reinforcement learning, there is an interactive process where the algorithm is rewarded, according to the accuracy of its predictions.

Once data is supplied, parameters (or weights) w of dependent variables x in the function $f(x)$ are estimated. The weights w compose a vector θ . Errors ε are calculated in a loss function $L(f(x|\theta), y)$. Learning occurs by back propagation, minimizing the errors by calculating the gradient ∇L along the loss function L and seeking its global minimum (i.e. making the derivative $dL/d\theta$ equal or close to zero, at the lowest downward concavity). The pace of gradient descent is controlled by adjusting a learning rate η , to avoid it being either too slow or to bounce around and diverge from the minimum.

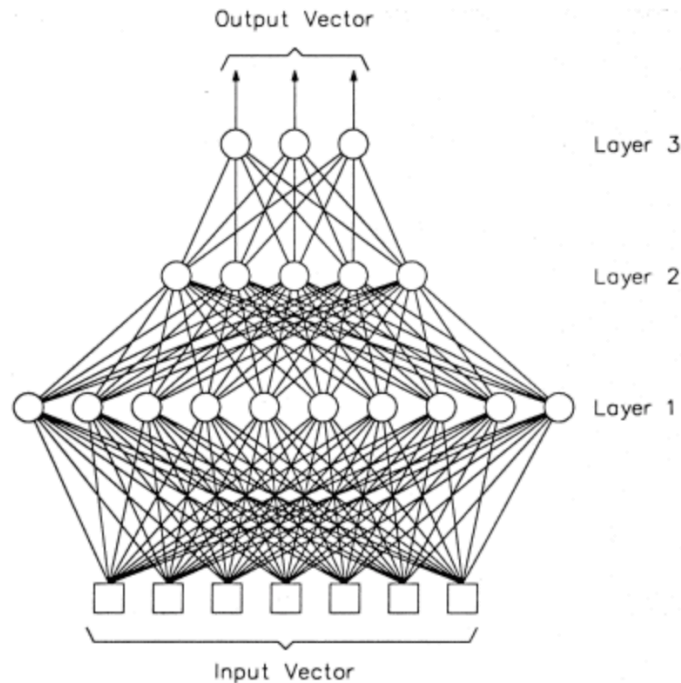
Some popular machine learning techniques are linear regression, logistic regression, decision trees, clustering, Bayes networks, support vector machines (SVMs), and artificial neural networks. Most techniques extract feature representations in up to one latent layer. However, to capture complex relationships among features, it may be necessary to abstract them into higher levels of representation, creating a hierarchy of concepts and purging information that is not relevant for feature variation. Algorithms with more than one level of latent (“hidden”) feature representation belong to the Deep Learning category.

Just to illustrate the evolution from machine learning in general to deep learning, the standard practice in computer vision used to be feeding data to learn feature representations first and then using a learning algorithm for image recognition. Nowadays, convolutional neural networks are capable of learning clusters directly from data, without separate feature extraction (SALAKHUTDINOV, 2015).

The feed-forward neural network, also known as multi-layer perceptron (MLP), is the quintessential DL example (GOODFELLOW; BENGIO; COURVILLE, 2016).

Figure 1. Multi-layer Perceptron (MLP).

Source: http://www.statistics4u.info/fundstat_eng/cc_ann_bp_function.html



Multi-layer perceptrons are graphic algorithms inspired by the human brain. The nodes represent the neurons and the connections are the synapses between them. They are organized in layers, with data (signals) moving from the input “visible” nodes x , through several layers of latent “hidden” nodes h , up to the output nodes y . Data is processed by inputting data into the visible nodes (feature values) and adjusting the weights in the connections. As data moves through higher layers, more abstract representations are processed through the network.

5. CONCEPT DESIGN

The automobile industry is transitioning from an analog hardware based business to a digital data driven model. According to Rossman (2016, quoting a 2014 McKinsey report), an automobile has computing power equivalent to twenty personal computers and features about 100 million lines of programming code. The automobile is one of the most powerful connected devices a person may own (ROSSMAN, 2016). Nikkei Business magazine (May 2017) reports the next step is to reduce tens of ECUs (electronic control units – specialized computers) and hundreds of algorithms down to a single and more powerful GPU (graphic processing unit) and one master algorithm in each car.

Automobiles were chosen to exemplify product concept design in this paper. They are very complex systems, coming in a wide range of segments, prices, sizes, needs, usages, tastes, geographical conditions, legal requirements and cultural flavors. Developing a successful product concept is an ever more complex challenge. A flawed concept may undermine its success from the beginning. According to Clark and Fujimoto (1990), a successful product concept must contemplate four dimensions:

- What the product does, a “dynamic” description of its performance and functionalities;
- What the product is, a “static” description of form, packaging, configuration, and features;
- Whom it serves, or the customer’s personality, profile, life style, psychographics;
- What it means, the product personality, character image, messages, and feelings.

A characteristic trait of a well-designed concept is integrity, in two aspects: a) coherence between concept and its physical elements and functions (internal integrity); and b) between concept and the image and messages it conveys (external integrity) (CLARK; FUJIMOTO, 1990). Most companies do not have well-defined and consistent methods and tools for designing an integrated concept. It is usually handled on an ad hoc and variable basis.

To design a product concept, there are three major key information sources: customers, competitors, and technology (others, such as legislation and environment, will be considered “latent” sources contained in the first three, for simplicity’s sake).

5.1. Customers

Companies usually conduct market research, in the form of interviews, clinics, focus groups, and - increasingly - on-line surveys. Blogs, web page comments, direct inquiries, and CRM reports are also important sources, as well as sales statistics and information from dealers. The difficulty is to extract the voice of the customer in a accurate and complete way, since data come in different formats and nuances, and are subject to varying levels of interpretation.

In the decades of 1990 and 2000, the lion’s share (more than 50%) of the Brazilian light vehicle market belonged to subcompact “B-segment” hatchbacks. In the beginning of 2010’s - in just a couple of years -, their share had dropped to about a third of the market. Compact B SUVs took over the balance (statistics source: fenabrave.org). Despite all the apparatus and market data, most companies did not predict the speed of change and were not prepared for the new market reality.

Natural language processing (NLP) is a field of computer science that investigates how to process human language and is evolving quickly. For instance, some education institutes use NLP to grade student essays. In combination with on-line data streaming, NLP could be deployed to monitor product related web content such as blogs, sites, and comments. Sentiment analysis is a branch of NLP that studies how to capture and measure emotional states. It can be valuable to identify customer product perception in large amounts of data..

A multi-layer perceptron (MLP) can be trained for natural language processing and sentiment analysis, to make sense of information and to predict shifts in customer needs and aspirations. Starting from rough data, they progressively abstract essential information into higher correlated representations. An accomplished tool would probably employ on-line streaming of web data for sentiment analysis and correlate it with market research data and minute sales statistics (with segment, model, version, specification and geographical details). Additionally, as sensors start sending usage data directly to manufacturers, a new source emerges – what customers are doing real-time, not just what they are saying.

There is a Toyota System principle called “genchi genbutsu” (現地現物), which can be roughly translated as “seeing things on the spot”. As much as one collects data and conduct a lot of studies, it is always important to be where the action is and “feel” the customer experience first hand. And it is not enough to understand today’s customer mind. Levitt (1986) used the expression “marketing imagination” - to put oneself in the customer’s shoes today and to understand what they will want tomorrow.

There is space for human judgment. And there is room for artificial intelligence, monitoring the voice and actions of customers in real time and predicting their trends.

5.2. Competitors

Large companies spend significant sums purchasing competitor products. Vehicles are carefully studied for construction, packaging, fit and finish, functionality, styling, and perceived quality. They are dynamically tested on power, acceleration, stability, agility, comfort, safety, silence, economy, braking, steering, and visibility. “Tear down” engineering disassembles vehicles and analyses engineering solutions, processes, and materials.

Based on those inputs, manufacturers establish performance, functionality, and feature benchmarks and targets. Unfortunately, that is not enough. Development takes about three years and - after launch - an automobile must remain competitive for about six or seven years (usually, there is a refreshing in the middle of the life cycle but the essentials remain). Responsibility for environmental impact, parts availability, servicing and recycling remain for at least ten additional years. Benchmarks are set today but products are sold tomorrow. A recently launched product is often matched or surpassed by a new competitor entering the market soon afterwards (and not available for benchmarking at development time), raising the bar and leaving the former behind in both customer value and market performance.

Ford has recently unveiled a high-performance limited production GT supercar. It recalls its historical success in endurance racing and has won the 2016 24 Hours of Le Mans race in its category. Extensive development and benchmarking was conducted and the manufacturer was confident it surpassed direct Italian (Ferrari 488, Lamborghini Huracan) and British (McLaren 675LT) competitors in all performance parameters. But McLaren launched a new model (720S) just afterwards (Ford benchmarked its predecessor, 675 LT) and this is superior in most performance aspects – more powerful, lighter, faster, better accelerating - and costs almost half the price (CAR, June 2017). The Ford GT is both an engineering and commercial success – the manufacturer has sold out a one thousand unit limited production beforehand. But had it been a series product, it would have faced a strategic vulnerability. They are direct competitors, as Ford itself has explicitly acknowledged and it has benchmarked the 675 LT (AUTOCAR magazine, May 17, 2017).

Image 1. Ford GT. Source: ford.com



Image 2. McLaren 720S. Source: 720s.mclaren.com



One way to handle such kind of risk is systematic and predictive competitive intelligence, by carefully collecting current and historical data, sketching competitor life cycle plans, projecting and predicting competitor moves, as war scenery maps in the military.

Game theory and multi-agent reinforcement learning have been applied successfully to model and predict complex systems like international geopolitics (MESQUITA, 2009 / STONE; VELOSO, 2000). Multi-agent reinforcement learning is a Bayesian network, probabilistic graphical model which borrows foundations from decision and game theories (PARSONS; WOOLDRIDGE, 2002). They calculate conditional probabilities of branches in decision trees of self-interested utility maximization seeking agents. To design a multi-agent system, Stone and Veloso (2000) list the following domain issues: number of agents, agent resources, amount of time pressure, dynamically arriving goals, cost of communication, cost of failure, user involvement, and environmental uncertainty.

5.3. Technology

The value of electric/electronic content in a typical automobile is about 40% (WEBER, 2009). For electric and hybrid vehicles it reaches 75%.

A technology road map is a short and long-run plan of technology solutions. It can be split into vehicle sub-systems – powertrain (engine, transmission), body, chassis (floorpan, suspension, brakes, steering, tires), interior (dashboard, seats, interior trim), electronic architecture, controls, displays, climate control, connectivity, entertainment, illumination, safety, and environment. The map may also be segregated into market segments, since technology tends to migrate from higher (more expensive) segments down to mainstream. Concept vehicles (presented in events like auto shows) usually showcase coming technologies and can be monitored and studied. Analogies from different industries are also a source of information. Sectors such as aerospace, military, computers, electronic appliances, and even households (digital virtual assistants come to mind) can provide interesting insights and hints.

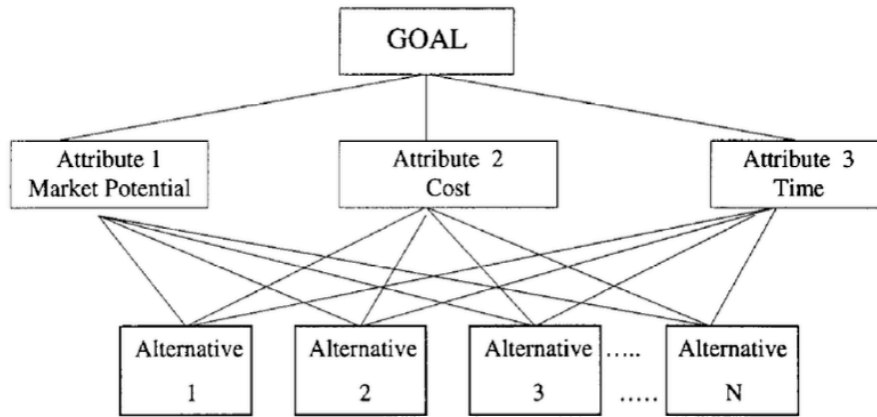
Genetic algorithms draw inspiration from biological evolution. A set of technology solutions in a vehicle sub-system is a population in a genetic algorithm and could be represented by a “chromosome”, a string of binary data. Solutions are individuals, substrings in the chromosome. Solution features or attributes are genes. Alleles are gene values. Recombination happens through crossover and mutation. Crossover is a reproductive process where string segments are swapped between two parent individuals, generating a descendent. Jiao, Zhang, and Wang (2005) adopted a 0.8 crossover rate in a notebook computer product portfolio experiment. Mutation is a random change in a string segment and should be rarer, with a small probability such as 0.01. Selection is carried out by evaluation from a fitness function. For product development, it can be an utility function, ratio between customer value and engineering cost (JIAO et al, 2005). The process continues until termination, when the population reaches equilibrium or a predetermined number of generations is achieved.

5.4. Product Concept

Once customer aspirations, competitor moves, and technology solutions are assessed, a methodology to combine those results is needed. Companies usually rely on heuristic methods for concept design. Huang, Bo, and Chen (2006) point two major flaws in this approach: subjective errors derived from variability in experience, and loss of historical expertise from previous successful cases. They used a multi-layer perceptron (MLP) for designing the concept of an insertion mechanism in a looming machine.

Chi-Hsing, Jiang, and Lee (1999) proposed a neural network product development model using fuzzy logic on market potential, cost, and time attributes to evaluate alternative product concepts (figure 2).

Figure 2. NN Product Development Model. Source: Chi-Hsing; Jiang; Lee (1999)



For automobile development, an MLP can be used to integrate the outputs of the three sources (customers, competitors, technology), condensing and filtering out alternative clusters of concept features. The clusters (alternative concepts) can be ranked using the product portfolio index below.

Equation 1. Product Portfolio Index

$$\text{PPI} = (\%m \text{ Project} / \%m \text{ Brand}) * (\Delta \text{ Segment} / \Delta \text{ Market}) * (\text{ROI} / k)$$

%m : average market share forecast over project life cycle

Δ : average year sales variation forecast over project life cycle

$$\text{ROI} = (\text{Net Income} / \text{Investment}) / \text{Investment}$$

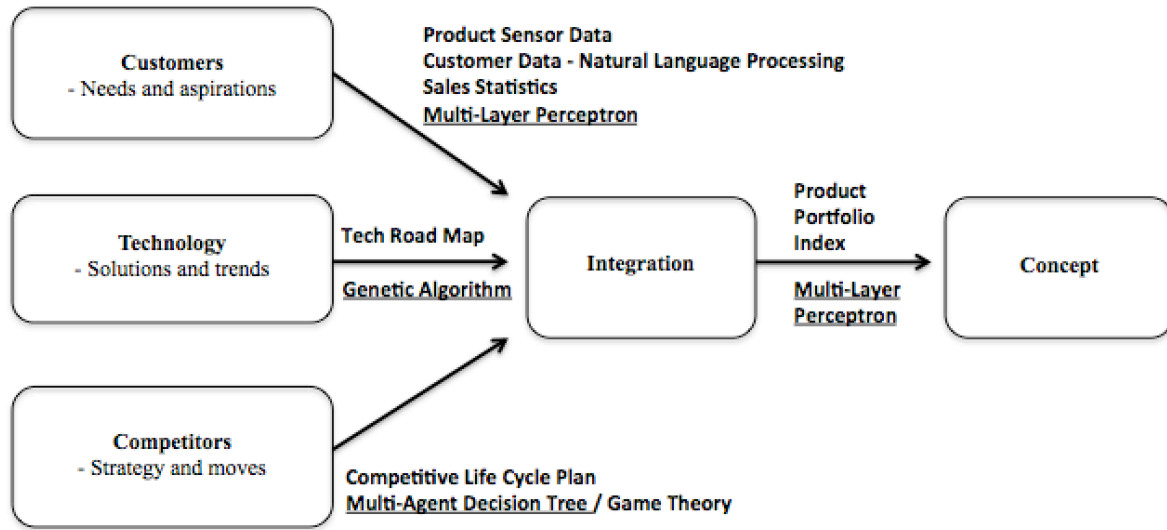
$$\text{Net Income} = \text{Gross Profit} - \text{Expenses}$$

$$k = \text{corporate cost of capital}$$

Segment and market projections are usually available from corporate planning activities. Brand market share and cost of capital are easily available from current operations. ROI is calculated by financial engineering or estimation. Project market share can be estimated by regression or expert estimates (alternatively, historical market share data may be inputted as an additional MLP feature).

Few companies can afford to be present in all segments of a market. Few companies have a clear product portfolio policy either. Product portfolio strategy should start from strategic vision – where the company wants to be, who are its customers, what it wants to mean to them. The product portfolio is the means to achieve that vision. Maybe the company will not be able to reach all its potential customers. Optimum portfolio is not necessarily maximum portfolio. The best product portfolio is one that yields the highest customer value relative to resource allocation (reflecting into the company's bottom line). The product portfolio index ranks product concepts. Ranked concepts may be stacked up until investment budget or engineering capacity ceiling is reached. Or the portfolio can be optimized by an algorithm. Jiao, Zhang, and Wang (2005) proposed using a heuristic genetic algorithm to model the product portfolio. Alternatively, the variables for segment potential, project market strength and return in the product portfolio index may be included as additional features in the multi-layer perceptron, yielding an optimum portfolio.

Figure 3. Concept Design



6. CONCLUSION

Internet-of-Things and machine learning are foundations of a new industrial paradigm. They will be essential elements of both operations and product development, as increasing flexibility will be required to match fleeting customer needs and aspirations.

Current pre-development practice may not be adequate. There is too much variability and loss of experience and knowledge. There is also lack of systems-view and integration. Machine learning may provide the tools to address that pitfall. The model proposed in this paper centers on three intelligence sources - customers, competitors, and technology - and suggests the application of a set of machine learning tools – natural language processing/ multi-layer perceptron, multi-agent reinforcement learning, and genetic algorithm, coordinated by another multi-layer perceptron. Other machine learning tools may prove equally effective. Support vector machines (SVMs) and random forests are very popular and powerful. SVMs carry out linear and nonlinear classification by identifying gaps in the representation space. Random forests rely on multiple decision trees to execute a classification or regression by taking the mode or average of them. To find the best combination, empirical comparative testing is required and may vary according to the type of data available. Inference quality is directly related to data volume and quality.

Machine learning needs large amounts of training data but extracts relationships that are difficult to map mentally or heuristically. Manufacturing turned-digital companies usually do not have access to the same huge volume of data native digital companies do. Less computational power is required. But as much technical skills and probably even more creativity and imagination are needed. And euphonic interaction between human and machine learning.

REFERENCES

- CHI-HSING, H.; JIANG, B.C.; LEE, E.S. Fuzzy Neural Network Modeling for Product Development. Mathematical and Computer Modelling. May 1999, p.71-81.
- CHOU, T. Precision: Principles, Practices and Solutions for the Internet of Things. Crowdstory, 2016.
- CLARK, K.B.; FUJIMOTO, T. The Power of Product Integrity. Harvard Business Review, Nov-Dec 1990.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep Learning. MIT Press, 2016.
- HUANG, H.Z.; BO, R.; CHEN, W. An Integrated Computational Intelligence Approach to Product Concept Generation and Evaluation. Mechanism and Machine Theory, May 2006, p.567-583.
- JIAO, J.; ZHANG, Y.; WANG, Y. A Heuristic Genetic Algorithm for Product Portfolio Planning. Computer and Operations Research, 2007, p.1777-1799.
- LEVITT, T. The Marketing Imagination. Simon and Schuster, New York, 1986.
- LIKER, J.K.; MORGAN, J.M. The Toyota Way in Services: The Case of Lean Product Development. The Academy of Management Perspectives, May 2006, p.5-20.
- MESQUITA, B.B. The Prediction Game. Random House, New York, 2009.
- MITCHELL, T.M. Machine Learning. McGraw-Hill, 1997.
- MORGAN, J.M.; LIKER, J.K. The Toyota Product Development System, Productivity Press, New York, 2006.
- PARSONS, S.; WOOLDRIDGE, M. Game Theory and Decision Theory in Multi-Agent Systems. Autonomous Agents and Multi-Agent Systems, 2002, p.243-254.
- PRIOR, M. Ford GT. Autocar, May 17, 2017, p.24-31.
- ROSSMAN, J. The Amazon Way on IoT. Clyde Hill, 2016.
- ROBERTSON, I. The Next Big Things. Car Magazine, UK, June 2017.
- SALAKHUTDINOV, R. Deep Learning. Lecture at the University of Toronto, Aug 2015.
- SHIMAZU, S.; TAKEDA, K.; OGASAWARA, K.; YAMASAKI, Y. A.I. In Nikkei Business, May 22, 2017.
- STONE, P.; VELOSO, M. Multiagent Systems: a Survey from a Machine Learning Perspective. Autonomous Robots, 2000.
- THOMKE, S.; FUJIMOTO, T. The Effect of “Front-Loading” Problem Solving on Product Development Performance. Journal of Product Innovation Management, Mar 2000, p.128 - 142.
- WEBER, J. Automotive Development Processes – Processes for Successful Customer Oriented Vehicle Development. Springer, Berlin, 2009.