Self-organizing tool for smart design with predictive customer needs and wants to realize Industry 4.0

Alfredo Alan Flores Saldivar
School of Engineering
University of Glasgow
Glasgow G12 8LT, U.K.
a.flores-saldivar.1@research.gla.ac.uk

Cindy Goh
School of Engineering
University of Glasgow
Glasgow G12 8LT, U.K.
Cindy.Goh@glasgow.ac.uk

Wei-neng Chen
School of Computer Science and Engineering
South China University of Technology
Guangzhou 510006, China
cwnraul634@aliyun.com

Yun Li*
School of Engineering
University of Glasgow
Glasgow G12 8LT, U.K.
* Corresponding author: Yun.Li@glasgow.ac.uk

***Abstract*— Following the first three industrial revolutions, Industry 4.0 (i4) aims at realizing mass customization at a mass production cost. Currently, however, there is a lack of smart analytics tools for achieving such a goal. This paper investigates this issues and then develops a predictive analytics framework integrating cloud computing, big data analysis, business informatics, communication technologies, and digital industrial production systems. Computational intelligence in the form of a self-organizing map (SOM) is used to manage relevant big data for feeding potential customer needs and wants to smart designs for targeted productivity and customized mass production. The selection of patterns from big data with SOM helps with clustering and with the selection of optimal attributes. A car customization case study shows that the SOM is able to assign new clusters when growing knowledge of customer needs and wants. The self-organizing tool offers a number of features suitable to smart design that is required in realizing Industry 4.0.**

Keywords—Smart manufacturing, Industry 4.0, smart design, big data analytics, self-organizing map.

I. INTRODUCTION

Recently, the U.S. has been driving Cyber Physical Systems (CPS), Industrial Internet, and Advanced Manufacturing Partnership (AMP) research for future manufacturing [1]. Germany is leading a transformation of the entire value chain of the manufacturing industry toward a “fourth Industrial Revolution”, coined “Industrie 4.0” or “Industry 4.0” [2]. In China, a “Manufacturing 2025” plan and Internet Plus have also been announced to advance manufacturing and accelerate service innovation [1].

Following the first three industrial revolutions as characterized by the steam power created industrial production, the electricity powered mass production, and the electronics enabled automation, Industry 4.0 (i4) aims at realizing mass customization at the cost of mass production. It introduces Internet of Things (IoT) to manufacturing by connecting embedded system technologies and smart production processes to pave the way for a new technological era expected to transform industry and business models radically [3]. In i4, a collection of technologies and concepts such as CPS, Smart Factory, IoT, and Internet of Services (IoS) interact with one another to form a closed-loop production value chain. Communications between these components can take place at several levels. For example, within a modular structured smart factory, CPS monitor the physical processes, create a virtual copy of the physical world and make decentralized decisions. Over the IoT, CPS communicate and cooperate with each other and humans in real time. Via the IoS, both internal and cross-domain organizational services are offered and utilized by participants of the value chain [4].

Recognizing that intelligent ways of manufacturing through i4 can transform the way customers acquire goods and services, several German multinational companies have started developing a well-integrated methodology to optimize the communication pathways between different systems in their manufacturing chain. Through the use of the Internet and smart devices, they aim to pursue a service-oriented strategy and strong customization of products to achieve higher throughput and production flexibility. With the introduction of methods that are adaptive, self-learning, self-aware, self-predictive, self-optimized, self-configurable and self-maintained, the automation technology aims to change the way products are manufactured and how services are provided [5].

With the power of IoT and the proliferation of data, companies are facing significant challenges in turning key data trends into informed decisions. This includes predicting customers’ needs and wants, both conscious and underlying. In order to present more reliable customization and production plans, intelligent strategies are currently being developed. The ability to predict key consumer trends based their needs and wants is an advantage that would put companies ahead of their competitors. Nonetheless, despite the benefits to businesses, most manufacturing systems are not ready to leverage on the rich information in these datasets. This is primarily due to a lack of smart analytics tools for design and manufacture.

A. A. Flores Saldivar is thankful to CONACYT for a Mexican Government scholarship. W.-N. Chen and Y. Li are thankful to the National Science Foundation of China and the Royal Society of the UK for the support via a Newton Fund grant (RS NI140008 - NSFC 61511130078).

To address such an issue, this paper investigates a predictive analytics framework and how it can be integrated with cloud computing, big data analysis, business informatics, communication technologies, and CPS in industrial production systems. Predictive technologies can further interlace or join intelligent algorithms with electronic devices and tether-free intelligence to predict product performance degradation and autonomously manage and optimize product service needs and wants [2]. The aim of this work is to develop such a framework that is built on the i4 ideology, i.e. a system that is capable of automatically predicting customer needs and wants for optimal product design and customization with added value for manufacturing.

In Section II of this paper, challenges and trends of i4 are discussed, together with the issues surrounding mass customization. In Section III, we tackle the smart design issue for mass customization and present a self-organizing tool for predicting customer needs and wants. We demonstrate the effectiveness of the proposed methodology through a case study in Section IV. Lastly, Section V draws conclusions with discussions on future work.

II. Challenges and trends of Industry 4.0

In order to be responsive and adaptive to market changes and demands, manufacture tends to integrate technologies and along with industry's research and developments conceived the adoption of mechanical systems, to support production process, to what today is common to find in manufacturing process as highly automated assembly lines. It is not difficult to find that information technology and social media networks have increasingly influenced consumer's perceptions on product innovation, quality, variety and speed of delivery.

A. Customization for smart manufacturing

Smart factories allow meeting singular customer necessities, while maintaining the effectiveness acquired in automated production. Meaning that even one-off items can be produced beneficially. Manufacturing companies generally oppose to various trends: growing global competition, more individualized customer demands, new technologies and rapid technological progress, as well as strict environmental regulations. Nevertheless in i4, dynamic business and engineering processes enable last-minute changes to production and deliver the ability to respond flexibly to disruptions and failures [6]. These trends lead to an increase in product variety, shorter product life cycles, uncertain and fluctuating demands, as well as higher cost pressure. Fig. 1 shows how over time manufacturing has shifted from mass production to mass customization [7].

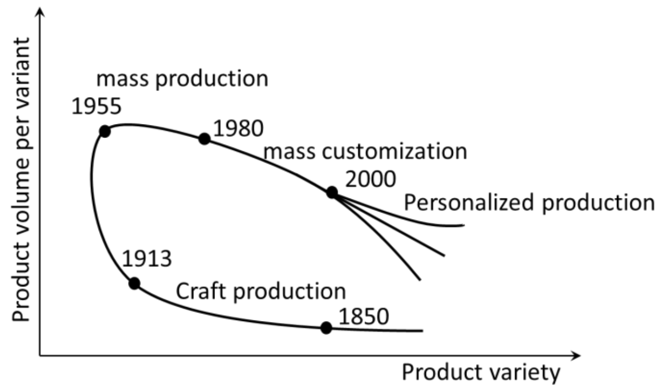

Fig. 1. Trend of mass customization vs mass production [7]

Optimized decisions and design are facilitated through end-to-end transparency given over the manufacturing process. Moving forward, i4 will bring new ways of creating novel business models and value. Especially, it will give start-ups and small businesses with the chance to create and provide downstream services. Additionally, i4 will diminish factory-floor necessities and facilitate progress humanity.

i4 is still in a nascent phase, yet there are issues and unmet necessities for reaching a mature state, describing as a learning curve that quickly has to reach a mature state due to the pressing needs of manufacturing.

B. Main issues in achieving mass customization

So far, implementation of self-learning processes for smart manufacturing have not been successfully achieved. The transformation from the existing state into more intelligent machines will require further advancement in science. In particular, the issues below need to be addressed [2]:

- **Manager and Operator Interaction**: operators control machines, managers design logistic schedules and machines are only performing the assigned tasks. Although these tasks are usually optimized by expert operators and managers, a significantly important factor is missing in these decisions: the health condition of the machine components.
- **Machine Fleet**: It is very common that similar or identical machines (machine fleet) are being exposed to completely different working conditions for different tasks. In contrast, most predictive and prognostic methods are designed to support a single or limited number of machines and working conditions. Currently, available prognostic and health management methods are not taking advantage of considering these identical machines as a fleet by gathering worthwhile knowledge from different instances.
- **Product and Process Quality**: As the final outcome of the manufacturing process, product quality can provide much insight on machine condition via backward reasoning algorithms. Product quality can provide feedback for system management, which can be used to improve production scheduling. Currently, such feedback loop does not exist and needs further research.
- **Big Data and Cloud**: Data management and distribution in Big Data environment is critical for achieving self-aware

and self-learning machines. The importance of leveraging additional flexibility and capabilities offered by cloud computing is inevitable, but adapting prognostics and health management algorithms to efficiently implement current data management technologies requires further research and development.

- **Sensor and Controller Network**: Sensors are the machine's gateway to sense its surrounding physical environment. However, sensor failure and degradation may pass wrong and inaccurate readings to decision-making algorithms, which will result in an incorrect outcome.

It is underlined for the purposes of this work what is discussed in [2] of the aforementioned issues, how data is managed. CPS will only implement mass production, but mass customization needs to be designed beforehand, and it is often found that customer is not clear what their needs and wants are [5], as stated once by world's innovative company Apple Inc. CEO 'people don't know what they want, until you show it to them' [8]. Suddenly how data is managed will lead to evolution for the innovation floor by this constant communication and linkage that IoT enables and with this data aim to move from manually innovation like many successful company leaders (Steve Jobs, Henry Ford, etc…) to automation of this process.

Subsequently it is necessary to discuss about what big data is for i4 and Smart Factory environments. The next section investigates how big data and CPS may be integrated in order to achieve self-prediction among other features.

C. *Big data helping mass customization in Industry 4.0*

Big data and cloud computing for i4 are viewed as data services that utilize the data generated in i4 implementations, but are not independent as i4 components [4]. For Industry 4.0 and Smart Manufacturing processes dealing with large data storage, sharing data, processing and analysing are becoming key challenges to computer science research. Some examples of these include efficient data management, additional complexity arising from analysis of semi-structured or unstructured data and quick time critical processing requirements. In order to resolve these issues, understanding of this massive amount of data, advanced visualization and data exploration techniques are critical [9].

A good understanding of the dataset is crucial to the choice and the eventual outcome of the analysis. Within the context of i4.0, there are two main sources of data: human-generated data and machine-generated data, both present huge challenges for data processing. Furthermore, Big Data cannot be defined by data volume alone. The speed of data production, need for short-time or real-time data storage and processing and incomplete or noisy data all adds to the complexity which render the analysis and interpretation of data a highly non-trivial task. It becomes even more challenging when data analysis and decision making needs to be carried out in real time.

Big data for the i4 era, intelligent analytics and CPS needs to become one (team together) to achieve new conceptualization of production factory transformation and management. By utilizing sensor installations, signals like vibration, pressure, and other measurements can be extracted. Apart from that, verifiable data can be collected for further data-mining. The protocol proposed by [10] developed by MTConnect® and Open Platform Communications (OPC) can help users record controller signals. Big data refers to the accumulation of these records, weather be signals, measures, historic data; all that combination is called "Big Data" [2]. The transforming agent consists of many components: visualization tools, predictive analytics, and an integrated platform. This deployment platform is chosen based on: investment cost, speed of computation, ease of deployment and update, among others [1]. Useful information driven by the processing of big data is key of sustainable innovation within an Industry 4.0 and Smart Factory. Such evolution needs the use of advanced prediction tools in order that data can be consistently processed into information that can explain the uncertainties and thereby enable staff members to take decisions better informed. Adopting the "Internet of Things" principle has helped in setting the fundamentals for predictive manufacturing by setting the essential foundations of smart machines and smart networks [11]. The goal of a predictive manufacturing system lies on enable systems and machines with "self-aware" capabilities. Smart computational agent, which contains smart software to direct prediction in modeling functionalities, is the core technology [1].

It is essential for manufactures to transform processes into predictive manufacturing. Most of manufacturing strategies assume continuous equipment availability and constant optimal performance, but those assumptions doesn't have to do in real factories at all. In Fig. 2 is shown the conceptual framework of a predictive manufacturing system, which starts with data acquisition from the monitored assets. This framework proposed by [1] use appropriate sensor installations, various signals such as vibration, pressure, etc. that can be extracted. For this platform is used predictive analytics tool called Watch Dog Agent ®, and it uses algorithms that can be categorized into four sections: signal processing and feature extraction, health assessment, performance prediction, and fault diagnosis. The author also states that when using visualization tools, health information (i.e., current condition, remaining useful life, failure mode, etc.) can be successfully displayed in charts like fault map, radar, health degradation curves, or risk chart.

All that collected data is then accessible to existing company management systems like Enterprise Resource Planning (ERP), Manufacturing Execution System (MES), Supply Chain Management, and Customer Relations Management; to achieve enterprise control and optimization. With predictive capability, equipment, design and customer needs and wants can be managed cost effectively. Then, historical data can be fed back to the equipment designer for closed loop lifecycle redesign.

Covering challenges for what collections of data can mean to i4 also leads to link this from the whole value chain that customers and manufacturers can get from the closed loop of intelligent products.

D. *Business informatics for Industry 4.0*

Customers play a key role in Smart Manufacturing environments. The goal of improving customer-business relationships, leading to customer retention and growth and ultimately driving sales profit is paramount to every business.

In order to achieve this, it is essential for the analysis of business performance and the response to analysis outcomes takes place in real-time throughout the entire customer lifecycle [12]. Nonetheless, this is not a trivial task that can be implemented overnight using existing business informatics models. Two factors can be largely attributed to this: (i) the lack of an automated closed-loop feedback system that can intelligently inform business processes to respond to changes in real-time based on the inputs (for example, data trends, user experience, etc.) received, and (ii) existing analytical tools cannot accurately capture and predict consumer patterns.

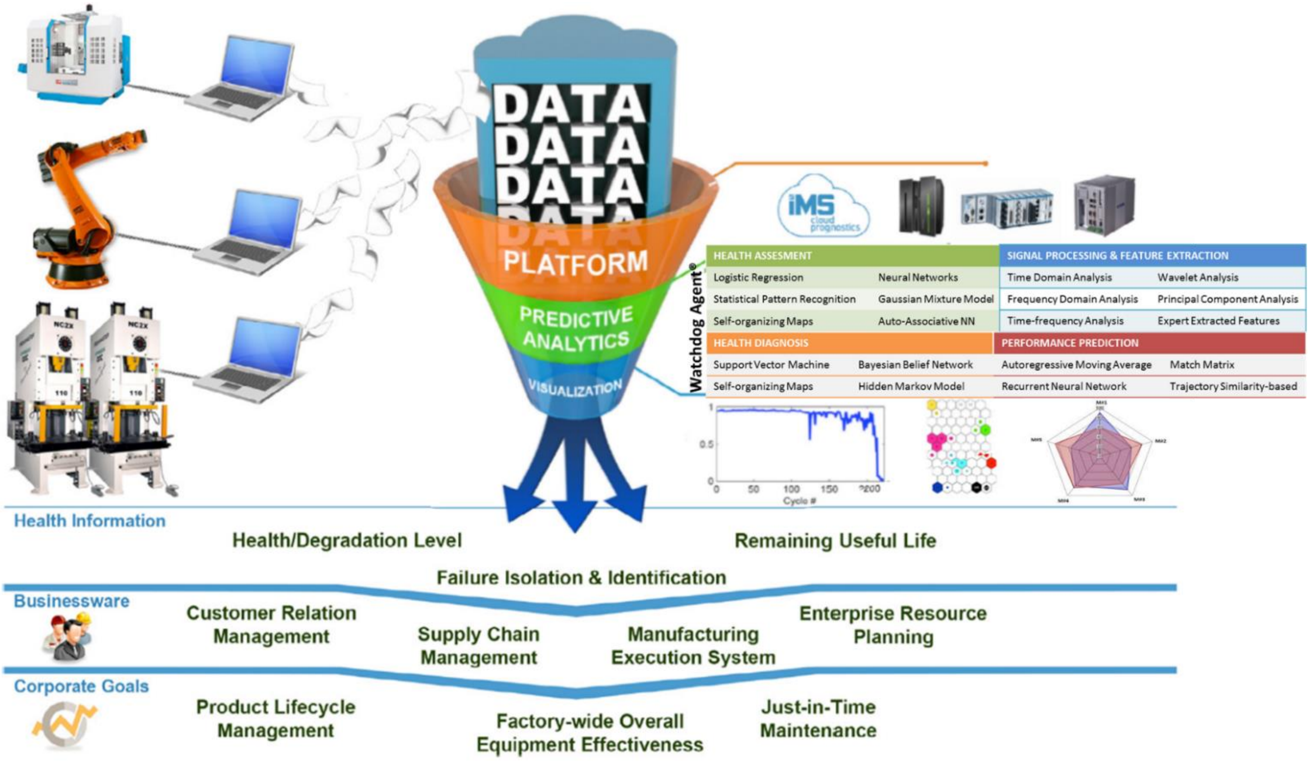

Fig. 2. Predicting manufacturing system framework [1]

A possible way forward for (i) is the use of digital models, a digital model capable of achieve automated close-loop. Leverage on existing web-based technologies, internet market places and internet services where digital products are used as starting points to evolve better designs - this is the vision of the future for i4. To get the most out of i4 technologies companies will have to be prepared for a digital transformation. Data management and cybersecurity will be critical problems to solve.

With the help of these products, they will be able to raise the efficiency both of their machines and of their production to new levels, avoiding the need to invest or buying complete and expensive technology databases as data which is not required does not need to be bought and superfluous investments can be avoided. Then future advantage is that the latest technology data is always used. Comparable business model is the permanent online maintenance of the machines.

A thing that all digital business models share is that there are online information, programs and solutions available for machine tools of manufacturing companies, facilitating their optimal deployment. It is of importance the number of users of the digital model, which determines whether is profitable or not. i4 business models are the same as the business models of digital world in the IoT and, as a result, must also adhere to their approach to controlling, representing a radical change for a conventional company.

Possible ways to resolve (ii) when analyzing business contained in data utilizing intelligence should be viewed as the use of gathered information into data lastly into action. The intelligence in this way, comes from the expert knowledge that can be also integrated in the analysis process, the knowledge-based methods used for analysis, and the new knowledge created and communicated by the analysis process.

Then what is really relevant for prediction in Customer Analytics for making business in i4 principles. Customer Analytics concerned with analyzing data and requires standard techniques like statistics, data mining, machine learning and intelligence Data Analysis.

Exploring the available methods suggested in [12] it is necessary to consider the following statistical tools than can help to achieve prediction:

- Data analysis regularly performed in simple ways, by utilizing linear models, and since from linear statistics are implicit numerous assumptions about mutually independence between variables and normally distributed values, therefore nonrealistic in real-world problems. However, as mathematical function non-linear models are typically unidirectional, i.e., those cannot be inverted then this is why nonlinear approaches are powerful.

- The hidden Markov models (HMM) [13] can be used for creating predictions on time-stamped events. Stochastic methods are represented by Markov models focused on the analysis of temporal sequences of separate (discrete) states.
- Using Bayesian networks [14] for analyze customer satisfaction, which is based on a graphical model representing inputs as nodes with directed associations among them. Nonetheless, because is aimed for educational level and don't provide needed levels of intuition, automation, and integration into corporate environments; accessible Bayesian network software is not suitable, enabling this can create them accessible to business users.

Enabling Real-Time Business Intelligence (RTBI) is key for i4, and since there are lots of generic ideas, Business Intelligence (BI) is not a much characterized term. Some consider BI as data coverage and visualization whereas others include business management performance. Data transformation, extraction and integration is highlighted by database vendors. Data mining and statistical analysis are emphasized by analysis tools vendors. These distinctive perspectives make it clear that BI has numerous facets. Defined in [12] is that BI as the framework for comprehend, access, and analyzing one among the foremost valuable assets of an enterprise – raw data – and turning it into actionable information in order to boost business performance.

Many issues can be found in current BI systems, those can be referred as follows:

- The moving from knowledge into data hindered by the shortage of analysts and specialists who are needed to configure and run analytical package.
- Bottleneck within the transition from information into action. This transition has historically been of a manual nature as a result of the shortage of automatic links back to the business method layer facilitating fast modification of method parameters to enhance execution.
- The capacity to fuse and relate the large quantity of data from various sources into a convenient and important source of information, together with the ability to approve the data and deal with quality problems.

In particular, RTBI ought to give three critical parts, denominated, real-time action on the business processes, real-time business performance analysis, and real-time information delivery. It should be emphasized here that the conception of real-time doesn't essentially equate to zero-latency in the operation of these three segments. The idea of real-time demonstrates the timeliness of the information-decision-action cycle that is important to the particular business environment [15].

Information flow among strategic, operational and key layers is broken by manual intercession, in current BI systems. The challenge is to utilize smart technologies to obtain mathematical representation of the manual intervention actual state in current systems and then automate both the actions necessary to translate strategic objectives back to operational drivers to influence strategic selections in real time, and therefore the flow of information from operational to tactical and strategic layers, representing data to the information stage of RTBI, like is presented in Fig. 3.

III. Self-Organizing and Clustering Framework

As described in the above sections, one of the many challenges before the Smart Factory environment can become a reality is the integration of tools to enable an automated predictive closed-loop system for manufacturing products. In this section, we describe an unsupervised machine learning approach, Self-Organizing Maps (SOM) [17] and how it is used in this work.

A. Self-organizing maps for clustering

To realize the i4 principles, it is recognized that a full integration of CPS and powerful tools for optimization, clustering, modeling, selection and prediction, leading to a complete analysis is crucial [1], [16]. The use of adaptive learning and data mining algorithms creates a knowledge base representing the scenario performance, when it's either considered characteristics of a product or attributes needed to be personalized, and then those mechanisms can be automatically populated. Knowledge base will be able to grow with new data to eventually enhance its fidelity and capability of representing complex working conditions that happen in real-world scenarios.

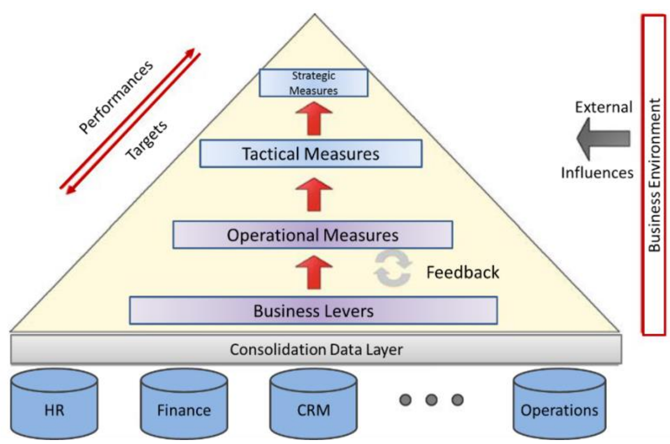

Fig. 3. Real-time business intelligence [12]

SOM is a type of Artificial Neural Network (ANN) which is trained through unsupervised learning, i.e. clustering. A SOM is made up of neurons (nodes), each with an associated weight vector. It is used in dimensionality reduction problems. Through adjusting the neurons and the associated weight vector, it is able to produce low-dimensional cluster representations (2D map) of a set of high-dimensional input data.

The obtained map is a NxN space, where the data are scattered and arranged. The number of neurons is set as the square of the map. The function can be summarized in four steps [17]:

1) Initialization: all connection weights of each cluster are initialized
2) Competition: for each input pattern, the neurons compete to each other in order to win this input. Wins the neuron that adapts its value closest to the input. It can be defined the discriminant function to be squared Euclidean distance between the input vector x and the weight vector w_j for each neuron j as:
$$d_j(X) = \sum_{i=1}^{D}\left(X_i - W_{ji}\right)^2 \qquad (1)$$
3) Cooperation: once it has been selected a wining neuron, it follows the creation of a neighborhood located close to the previous winner. Therefor the winning neuron creates a neighborhood with other neurons, in order to cooperate with each other and win future inputs. If S_{ij} is the lateral distance between neurons i and j on the grid of neurons, we define a topological neighborhood $T_{j,I(X)}$ where $I(x)$ is the index of the winning neuron:
$$T_{j,I(X)} = exp\left(\frac{S_{j,I(X)}^2}{2\sigma^2}\right) \qquad (2)$$
4) Adaption: this last stage is when each neuron creates a neighborhood or becomes a member of a neighborhood and self-organizes, so that the feature map between inputs is formed. The equation that describes the appropriate weight update is:
$$\Delta W_{ji} = n(t).T_{j,I(X)}(t),\left(X_i - W_{ji}\right) \qquad (3)$$

For every step, all neurons adapts their weights to the current input, but not as much as the winner neuron and its neighborhood. Visualization of the map presents, in this way, each neighborhood suitable for approximated values ordered and shaped.

B. Framework for predicting potential customer needs and wants

Fig. 4 presents the framework that can help to solve several of the afore-mentioned challenges in i4. This is based on i4 and Smart Manufacturing key objective, i.e. achieve self-prediction, and self-configurable in order to manufacture products and provide services that highly bespoke at mass production rates.

The users' inputs and requirements are key to any successful design. In the first block of the proposed framework, customer needs and wants are first captured and processed to extract key design characteristics. These information are then fed into a Computer Automated Design (CAutoD) engine [18] where the design requirements, features and performance objectives are mapped into 'genotypes' for further analyses. This process, which is commonly known as rapid virtual prototyping uses intelligent search algorithms such as the Genetic Algorithm (GA) or Particle Swarm Optimization (PSO) to explore the design search space for optimal solutions. In the proposed framework, this process takes place over the Cloud and produces a set of optimized virtual prototype at the end of the search.

The second block of the closed loop in Fig. 4 shows the virtual prototype, which is obtained from the selection and design process in CAutoD Through the integration of CPS or Cyber-Physical Integration (CPI), the virtual prototype in the second block is transformed into a physical product, i.e. the Smart Product as shown in Fig. 4.

The next part of the framework refers to Business Informatics and how the Smart Product is connected to the IoT. Here is where big data takes part, through the performance of the product and the feedback from the customer, more features can be considered, this covers the necessary attributes that makes the product be manufactured in optimal ways.

Following this, the response obtained from the customer is automatically fed back to the system for further analysis and to fine-tune the virtual prototype. It is necessary to perform analysis. This analysis is related to prediction, by using node or dynamic analysis that can perform clustering, selection and detection of patterns and visualize it. After that, the SOM clustering completes update of selected attributes by comparing the latest input to the existing cluster and tries to identify one cluster that is most similar to the input sample using multidimensional distance measurement. Then several features are fed back into the cloud again.

The analysis can result in two outcomes: (i) Similar clusters found. If it is the case, this will be reflected as an existent attribute and the algorithm will update the existing cluster using information from the latest sample. (ii) Non-similar clusters found. The algorithm will hold its operation with the current sample until it sees enough count-of-cluster samples.

When the number of out-of-cluster samples exceeds a certain amount, it means that there exists a new behavior in the data that has not been modeled, then the algorithm will create a new cluster to represent such new behavior.

Solving a clustering problem with a SOM network designed in Matlab® toolboxes was used for training the data that is presented in the following section. Widely used to group data by similarity, the SOM consists of a competitive layer that can classify a dataset of vectors with any number dimensions into as many classes as the layer has neurons [19].

For those cases, SOM clustering can be very adaptive to new conditions. Self-grow clusters will be used as the knowledge base for customization assessment.

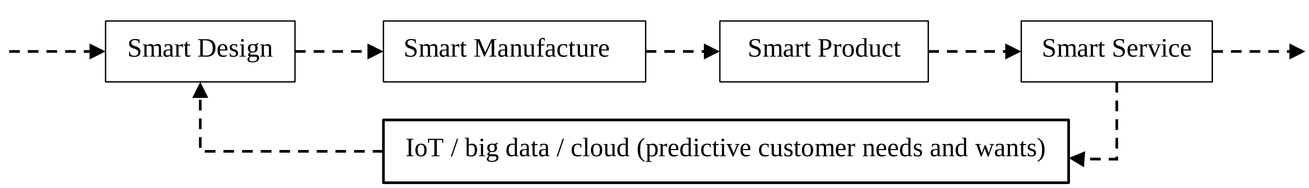

Fig. 4. Industry 4.0 value chain with predictive customer needs and wants fed back for automated customization

IV. CASE STUDY

It has been accessed to trained data found in a machine-learning repository [20] in order to run some tests, the information of the data shows an evaluation model of cars by acceptability, overall price, buying price, price of maintenance, technical characteristics, comfort, number of doors, persons capacity to carry, and safety of the car. The data set comprises of 1728 inputs and each record contains the subsequent attributes: safety, capability describing the persons to hold, buying price, maintenance price, number of doors, the dimensions of baggage boot, and car acceptance. For this data set, the attribute of car acceptance is a category label used to classify the level of the car that customers accept, then different attributes are seen as predictive inputs.

Database contents are shown in Table I.

TABLE I. CAR EVALUATION DATA [20]

Attribute name	Description	Domain
safety	Safety evaluation	Low / med / high
person	The number of passenger	2 / 4 / more
b_price	Buy Price	v-high / high / med / low
m_price	Repair price	v-high / high / med / low
size	Suitcase capacity	Small / med / big
door	The number of the door	2 / 3 / 4 / 5-more
class	level of customer acceptance	Unacc / acc / good / vgood

In order to examine the distributions for knowing this dataset better in Fig. 5 is presented the distributions for the car evaluation set.

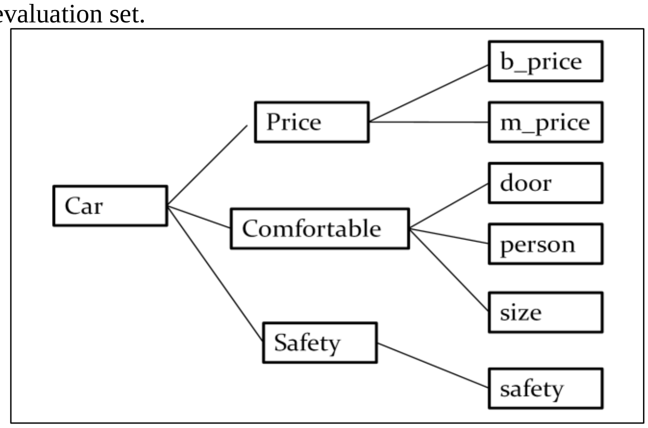

Fig. 5. Distributions of car evaluation dataset for customization

As a methodology to follow it was also considered the business problem as the following question: what reasonably cars can get good assessment? This question then this evaluations are used as target attributes, depending on which attribute. Then as the data mining problem, which is: find out the rules form other attributes.

The results of the SOM clustering are shown in Fig. 6. The neighborhood weight distance figure provides evidence about neighborhoods created: the darker colors represent larger distances, and the lighter colors represent smaller distances [21]. For which the first neuron in the inferior corner on the left results to be the strongest one, meaning that attribute selected is “safety”, if the input “safety” is low it will directly fall under unacceptable (“unacc”). Whatever estimation of safety is, if “person” value is 1, the entry will fall under unacceptable. This is represented in the right part of the figure presented below, where it shows the assigned clusters.

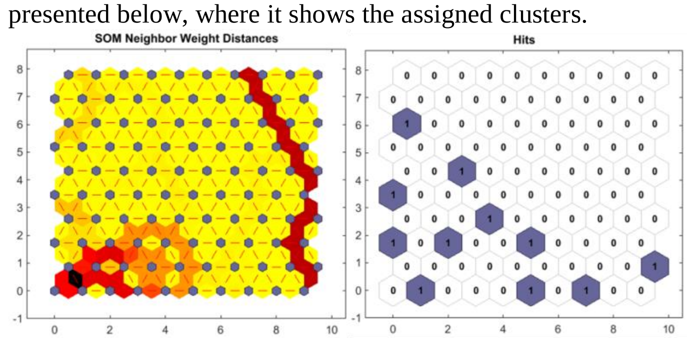

Fig. 6. Results of tested data. SOM weight distances on the left, and SOM clusters found on the right

The SOM worked with 10 hidden neurons, and 200 iterations. The confusion matrix is shown in Table II resulted from the analysis with Matlab®.

TABLE II. CONFUSION MATRIX FOR THE SOM

a	b	c	d	Classified
1171	28	0	3	a= unacc
7	292	4	9	b= acc
0	0	44	0	c= vgood
0	5	5	37	d= good

Then, it was also tested the average clustering coefficient with a value of 0.833. This mean the degree to which nodes in a graph tend to cluster together. Meaning that from the clusters 4/5 can be clustered together.

To test the accuracy of the model obtained, Table III shows in detail each class evaluated, this table as well is created from the Matlab® analysis.

TABLE III. MODEL ACCURACY BY CLASS

TP Rate	FP Rate	Preci -sion	Recall	F-Measure	Classified
0.974	0.017	0.994	0.974	0.984	unacc
0.936	0.026	0.898	0.936	0.917	acc
1	0.006	0.83	1	0.907	vgood
0.787	0.008	0.755	0.787	0.771	good

From the results presented above, it can be inferred that the model performs good, from all assessing values followed with less serious miss-classification, that there was sixty one entries that show wrong classification, it can be told from Table II that even those values are in a wrong category, most of them are lead a category close their actual categories.

Finally, the next section gives the discussion and what future directions can be taken in order to continue improving this loop and fulfilling the integration of technologies and concepts for Industry 4.0 achievement.

V. Discussion and Conclusion

Industry 4.0 value chain aims to enable a new level of mass customization for enhanced manufacturer offerings. The current lack of smart analytics tools for achieving this has been investigated in this paper. It has developed a predictive analytics framework for integration with cloud computing, big data analysis, business informatics, communication technologies, and CPS in a digital industrial system. In the same context, with this approach it was considered to tackle one specific research topic: how computational intelligence and evolutionary computing can contribute to Industry 4.0.

The utilization of advanced prediction tools helps obtain and process self-predictive and self-aware features in subconscious customer needs and wants, which can then be utilized to achieve smart design. In summary, the self-organizing predictive tool offer the following features:

1. The feedback design process is suitable for smart automation with CAutoD.
2. Intelligent search within the design process allows needs and wants to be predictively covered, with virtual prototypes further tunable by the customer, as well as increasing customer knowledge of the company (firm).
3. CPS interconnected to the designed virtual prototypes implements their mass manufacturing efficiently.
4. A smart product may be obtained with business informatics and reliable data constantly, which can be fed back to smart design again with IoT in the loop of the i4 value chain.
5. Since the “Internet of Everything” facilitates connection through the cloud, it makes it faster to satisfy customer needs and wants.
6. Decision for the manufacturer becomes easier to make, with customer-driven informatics, design and automation.
7. SOM clustering reflects the attributes of the car, for example, as shown in the case study, where the customer cares less about the “door” attribute.
8. In the case study, the results reveal that for a car customization restrictions v-good and good cannot be so easily met. Hence, it is predicted that the manufacturer should focus on the attributes on car sealing, offerings of high-security, and not other attributes.
9. Big data analytics (nodes) help visualize the influence of product characteristics, clustering and interpretation of subconscious customer needs and wants.
10. This approach can also contribute from the management perspective of the company to: enable innovation according to customers’ needs and wants, and help companies to avoid unnecessary product differentiation.

Future work includes testing the framework developed in this paper further in a more highly scalable environment. More intelligent tools with big data analysis will also be developed for i4 smart design. For example, intelligent test node analysis from social media may be used for estimation and prediction, for characteristics or features hidden in subconscious customer needs and wants.

References

[1] Lee, J., et al., Recent advances and trends in predictive manufacturing systems in big data environment. Manufacturing Letters, 2013. 1(8): p. 38-41.

[2] Lee, J., H.-A. Kao, and S. Yang, Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. Procedia CIRP, 2014. 16(1): p. 3-8.

[3] MacDougall, W., Industrie 4.0 Smart Manufacturing for the Future. Mechanical & Electronic Technologies, Germany Trade & Invest, 2014(2): p. 40.

[4] Hermann, M., T. Pentek, and B. Otto, Design Principles for Industrie 4.0 Scenarios: A Literature Review. Technische Universitat Dortmund, 2015. 01(3): p. 4-16.

[5] Kagermann, H., Change Through Digitization—Value Creation in the Age of Industry 4.0, in Management of Permanent Change, H. Albach, et al., Editors. 2015, Springer Fachmedien Wiesbaden. p. 23-45.

[6] Kagermann, H., W. Wahlster, and J. Helbig, Recommendations for implementing the strategic initiative INDUSTRIE 4.0. ACATECH NATIONAL ACADEMY OF SCIENCE AND ENGINEERING, 2013(6).

[7] Schönemann, M., et al., Simulation of matrix-structured manufacturing systems. Journal of Manufacturing Systems, 2015(5).

[8] Ciotti, G. (06/03/2013). "Why Steve Jobs Didn't Listen to His Customers." Help Scout Net Retrieved 05/04/2016, 2016, from https://www.helpscout.net/blog/why-steve-jobs-never-listened-to-his-customers/.

[9] Bohlouli, M., et al., Towards an Integrated Platform for Big Data Analysis, in Integration of Practice-Oriented Knowledge Technology: Trends and Prospectives, M. Fathi, Editor. 2013, Springer Berlin Heidelberg. p. 47-56.

[10] Lee, B.E., et al. MTConnect-Based Kaizen for Machine Tool Processes. in ASME 2010 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. 2010. Montreal, Quebec, Canada: American Society of Mechanical Engineers.

[11] Chui, M., M. Loffler, and R. Roberts, The Internet of Things, in McInsey quarterly No.2, McKinsey&Company, Editor. 2010, McKinsey&Company: United States of America. p. 4.

[12] Nauck, D., et al., Predictive Customer Analytics and Real-Time Business Intelligence, in Service Chain Management, C. Voudouris, D. Lesaint, and G. Owusu, Editors. 2008, Springer Berlin Heidelberg. p. 205-214.

[13] Rabiner, L. and B.H. Juang, An introduction to hidden Markov models. ASSP Magazine, IEEE, 1986. 3(11): p. 4-16.

[14] Heckerman, D. and M.P. Wellman, Bayesian networks. Commun. ACM, 1995. 38(12): p. 27-30.

[15] Azvine, B., et al. Real Time Business Intelligence for the Adaptive Enterprise. in E-Commerce Technology, 2006. The 8th IEEE International Conference on and Enterprise Computing, E-Commerce, and E-Services, The 3rd IEEE International Conference on. 2006.

[16] Flores Saldivar, A.A., et al. Industry 4.0 with Cyber-Physical Integration: A Design and Manufacture Perspective. in International Conference on Automation & Computing. 2015. University of Strathclyde, Glasgow: IEEE.

[17] Bougoudis, I., L. Iliadis, and S. Spartalis, Comparison of Self Organizing Maps Clustering with Supervised Classification for Air Pollution Data Sets, in Artificial Intelligence Applications and

Innovations: 10th IFIP WG 12.5 International Conference, AIAI 2014, Rhodes, Greece, September 19-21, 2014. Proceedings, L. Iliadis, I. Maglogiannis, and H. Papadopoulos, Editors. 2014, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 424-435.

[18] Li, Y., K.H.A., G.C.Y. Chong, Wenyuan Feng, Kay Chen Tan, Hiroshi Kashiwagi, CAutoCSD-Evolutionary Search and Optimisation Enabled Computer Automated Control System Design. International Journal of Automation and Computing, 2004. 1(16): p. 76-88.

[19] Arumugadevi, S., and V. Seenivasagam. “Comparison of Clustering Methods for Segmenting Color Images", Indian Journal of Science and Technology, 2015.

[20] Lichman, M., UCI Machine Learning Repository, U.o. California, Editor. 2013, School of Information and Computer Science.: Irvine, CA: University of California, School of Information and Computer Science.

[21] Shuai, S., Z. Laibin, and L. Wei. Condition monitoring and fault diagnosis of rolling element bearings based on wavelet energy entropy and SOM. in Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQR2MSE), 2012 International Conference on.