Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 53 (2021) 748–759

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

49th SME North American Manufacturing Research Conference, NAMRC 49, Ohio, USA

Quality 4.0 — Green, Black and Master Black Belt Curricula

Carlos A. Escobar[a,*], Debejyo Chakraborty[a], Megan McGovern[a], Daniela Macias[b], Ruben Morales-Menendez[b]

[a]*Global Research & Development, General Motors, Warren, MI, USA*
[b]*Tecnológico de Monterrey, Monterrey NL, México*

Abstract

Industrial Big Data (IBD) and *Artificial Intelligence (AI)* are propelling the new era of manufacturing – smart manufacturing. Manufacturing companies can competitively position themselves amongst the most advanced and influential companies by successfully implementing *Quality 4.0* practices. Despite the global impact of COVID-19 and the low deployment success rate, industrialization of the *AI* mega-trend has dominated the business landscape in 2020. Although these technologies have the potential to advance quality standards, it is not a trivial task. A significant portion of quality leaders do not yet have a clear deployment strategy and universally cite difficulty in harnessing such technologies. The lack of people power is one of the biggest challenges. From a career development standpoint, the higher-educated employees (such as engineers) are the most exposed to, and thus affected by, these new technologies. 79% of young professionals have reported receiving training outside of formal schooling to acquire the necessary skills for *Industry 4.0*. Strategically investing in training is thus important for manufacturing companies to generate value from *IBD* and *AI*. Following the path traced by *Six Sigma*, this article presents a certification curricula for Green, Black, and Master Black Belts. The proposed curriculum combines six areas of knowledge: statistics, quality, manufacturing, programming, learning, and optimization. These areas, along with an ad hoc 7-step problem solving strategy, must be mastered to obtain a certification. Certified professionals will be well positioned to deploy *Quality 4.0* technologies and strategies. They will have the capacity to identify engineering intractable problems that can be formulated as machine learning problems and successfully solve them. These certifications are an efficient and effective way for professionals to advance in their career and thrive in *Industry 4.0*.

Keywords: Quality 4.0; Certification; Smart manufacturing; Artificial intelligence; Big data

1. Introduction

To successfully deploy *Quality 4.0* practices and technologies on manufacturing systems, engineers, managers, and directors must be trained on the basic principles of *Artificial Intelligence (AI)*. Modern technologies, such as *Industrial Big Data (IBD)* and *AI*, are propelling a new era of manufacturing –smart manufacturing– within the context of the fourth industrial revolution (Industry 4.0). According to Forbes, the lack of people power is one of the biggest challenges facing these technologies in business [42]. Brookings estimates that the higher-educated, higher-paid workers such as engineers will be most exposed to, and therefore affected by, these new technologies. Other employees, such as those in either the lower-payed roles or those in senior executive positions will be somewhat unaffected [45]. According to Deloitte [3], 79% of young professionals have reported receiving training outside of formal schooling to acquire the necessary skills for *Industry 4.0*. In this context, this paper presents a *Quality 4.0* initiative and a certification program for quality/manufacturing engineers, managers, and directors. Certified professionals in this initiative will have the tools and skillset required to take the lead in deploying *Quality 4.0* practices and technologies.

> "*Quality 4.0* is the fourth wave in the quality movement (1.*Statistical Quality Control*, 2.*Total Quality Management*, 3.*Six Sigma*, 4.*Quality 4.0*). This quality philosophy is built on the statistical and managerial foundations of the previous philosophies. It leverages industrial big data and artificial intelligence to solve an entirely new range of intractable engineering problems. *Quality 4.0* is founded on a new paradigm based on empirical learning, empirical knowledge discovery, and

* Carlos A. Escobar. Tel.: +1-586-662-7501
E-mail address: carlos.1.escobar@gm.com (Carlos A. Escobar).

10.1016/j.promfg.2021.06.085

real-time data generation, collection, and analysis to enable smart decisions [24]."

Despite the global impact of COVID-19 and the low deployment success rate (13%-20% [52, 58]), industrialization of the *AI* mega-trend has dominated the business landscape in 2020 [32, 46]. According to Quality Digest, a significant portion of quality leaders do not yet have a clear deployment strategy [15], and universally cite difficulty in harnessing these technologies [49]. To prepare tomorrow's workforce for *Industry 4.0*, strategically investing in training approaches is crucial for manufacturing companies to generate value from *AI* and *IBD* technologies [3]. According to Montgomery [44]:

> "quality professionals are going to have to master some of the skills of computer science, such as understanding the structure of large databases, basic data mining techniques, image processing, and data visualization techniques."

Certification is a formal process for recognizing a person that has achieved competency (experience, theory, knowledge, education) in a specific area [35]. Certification organizations are formed by knowledgeable, experienced, and skilled professionals with the capacity to both identify the required competencies and develop the curriculum to achieve them. *Certified Systems Engineering Professional (CSEP)*, *Six Sigma Black Belt (SSBB)*, and *Professional Engineers (PE)* are good examples of certifications. These certifications are issued by the *International Council on Systems Engineering (INCOSE)*, *American Society of Quality (ASQ)* and *National Society of Professional Engineers (NSPE)*, respectively.

Competences acquired during the certification process set certified engineers apart from others. Usually, unique skills generate responsibilities that come with more authority and greater earning potential. According to the *American Society of Mechanical Engineers (ASME)*, mechanical engineers with a *PE* license have a 15% greater median income than unlicensed engineers [39]. On the other hand, in the quality domain, *Six Sigma* certified engineers are expected to lead a change within an organization and play a strong leadership role. Their salary is often directly related to their belt color [31]. Black Belts, on average, earn significantly more money (34%) than Green Belts; whereas Master Black Belts earn significantly more (30%) than Black Belts [5, 34, 53]. Certifications are an efficient and effective way to advance to new professional levels.

Six Sigma is a quality philosophy founded on statistics and *DMAIC (Define, Measure, Analyze, Improve, Control)*, a 5-step problem solving strategy. *Six Sigma* is globally applied across the manufacturing industry, because it delivers measurable, tangible economical benefits with a customer focus. Following the same convention of the Japanese sport, Karate, belt colors (Green, Black, Master Black Belts) are used to recognize proficiency in this philosophy. Today, whereas the *Six Sigma* philosophy is still necessary, *Quality 4.0* is the next natural step in the evolution of quality.

The factory of the future is driven by manufacturing systems that exhibit fast increasing complexity, hyper-dimensional feature spaces, as well as non-Gaussian, pseudo-chaotic behavior, which counteracts orthodox statistical methods, but opens a whole new avenue of opportunities for *AI*. In this context, the manufacturing industry is in the initial stages of adopting *Quality 4.0*. Therefore, certified professionals in this area will be agents of change and may position themselves in leadership positions.

Process Monitoring for Quality (PMQ)–the *Quality 4.0* initiative–is an *IBD*- and *AI*-driven quality philosophy that uses process-data for real-time defect detection, where defect detection is formulated as a binary classification problem. It is a blend of process monitoring and quality control. Empirical knowledge discovery aimed at process redesign and troubleshooting augmentation are at the core of this philosophy. *PMQ* proposes a 7-step problem solving strategy to identify and solve high value engineering intractable problems; *IADLPRR–Identify, Acsensorize, Discover, Learn, Predict, Redesign, Relearn*.

This paper anthologizes several publications [1, 19, 22–24, 27] of the authors and their experience studying complex problems–as part of the Manufacturing Systems Research Lab of General Motors–to propose the certification curricula and requirements for Green, Black, and Master Black Belts in *Quality 4.0*. As described in Fig. 1, the curriculum combines six areas of knowledge–statistics, quality, manufacturing, programming, learning, optimization–and the *PMQ* 7-step problem solving strategy that must be mastered to obtain a certification.

Fig. 1: Areas of knowledge of *Quality 4.0*.

The rest of the paper is organized as follows. A review of *Six Sigma* is presented in Section 2. A brief description of *PMQ* and its applications are presented in Section 3. An overview of the problem solving strategy proposed by *PMQ* is presented in Section 4. The requirements and competences for the Green, Black and Master Black Belt certifications in *Quality 4.0* are described in Section 5. Finally, conclusions and future research are contained in Section 6.

2. Six Sigma

In this section, the *Six Sigma* certification program context and basic background is presented. *Six Sigma* was developed in the 80s by Motorola, one of the world's leading manufacturers of electronics, and it became part of the company's DNA. In 1988, Motorola won the first Malcolm Baldrige National Quality Award. The Motorola University offered *Six Sigma* certifications to engineers around the globe. Soon after, *General Electric (GE)* adopted *Six Sigma* from Motorola in 1995, and under Welch, it became corporate doctrine. The company invested more than one billion USD to train thousands of employees, and the system was adopted by every *GE* business unit. In early 2000, *GE* surpassed Microsoft to become the world's most valuable company [59].

Today, *Six Sigma* is embedded with lean management [16], supply chain [13], design [69], and various other process improvement approaches to achieve synergized benefits.

2.1. DMAIC – Six Sigma Problem Solving Strategy

In *Six Sigma*, problems are solved following the 5-step problem solving strategy *DMAIC (Define, Measure, Analyze, Improve, Control)* proposed by "the father of *Six Sigma*," Bill Smith [17]. It is an effective process improvement cycle for structured solution development. The 5-step strategy is as follows:

- **Define** is the first step, where the problem is selected and its potential benefits assessed. This stage includes various activities, such as team members selection, project charter development, scope and goal identification, problem statement, business case development, bottom line impact estimation, and project schedule preparation.
- **Measure** is the second step, in which the problem is translated into a measurable form, and measurements of the process are taken. Activities in this step include: data collection chart creation, current level of process performance assessment, and quality cost calculation.
- **Analyze** is the third step, where the *Critical To Quality* (*CTQ*) influencing factors and causes are identified. These potential root causes are determined through a cause-and-effect matrix.
- **Improve** consists of designing and implementing adjustments to the process to improve the performance of the *CTQ*s. The main goal of this step is to develop the problem solution. Intense experimentations are performed to statistically validate solutions and the cause-and-effect relationships previously hypothesized in the matrix.
- **Control** consists of the empirical verification of the project's results. The objective is to ensure that the solution created in the Improve phase is well-implemented and maintained. Moreover, the opportunity of replicating the solution to other processes is also evaluated.

2.2. Belt Color Convention – Six Sigma Mastery Levels

The belt color has its roots in the realm of martial arts (Karate). This naming convention is used to describe a level of mastery of *Six Sigma* and it is obtained through a certification process. The most common levels are Green, Black, and Master Black Belts. Certified people conduct projects and implement improvements at different levels, depending on their respective belt color. In descending order of responsibilities:

- *Master Black Belt*: Trains and coaches Black Belts and Green Belts. Functions more at the *Six Sigma* program level by developing key metrics and the strategic direction. Acts as an organization's *Six Sigma* technologist and internal consultant.
- *Black Belt*: Leads problem-solving projects. Trains and coaches project teams.
- *Green Belt*: Assists with data collection and analysis for Black Belt projects.

The *American Society for Quality (ASQ)* offers external certifications in the different belt levels. In each certification, the candidate is required to pass a written examination that consists of multiple-choice questions that measure comprehension of the body of knowledge [4].

2.2.1. Six Sigma in the Age of Big Data

Quality 4.0 is the next natural step in the evolution of quality, as the *Six Sigma* paradigm based on traditional statistics is not designed to efficiently/effectively address the challenges posed to *IBD* [61]. Statistics draws population inferences from a sample. It focuses on analyzing and summarizing experimental data under assumptions and it is more suited for processing lesser amounts of linear, repeatable data derived from systems where relationships are relatively stable [10]. Whereas *Machine Learning Algorithms (MLA)* automatically learn predictive patterns from huge data sets. They learn from observational data, complex non-linear patterns that usually exist in hyper-dimensional spaces, without assumptions or a predefined model form. To cope with the big volumes of data, *MLA* have embedded computational efficiency concepts to enable computational feasibility, e.g., stochastic gradient descent [18], XGBoost [11]. Machine learning programs usually improve with more data, and they are intrinsically designed to automatically learn dynamic relationships e.g., new trends, patterns, or sources of variations. But, this come at the expense of explain-ability [57], as they are considered black boxes with little interpret-ability or understandability capacity. Both paradigms (traditional statistics and machine larning) are complementary, therefore, quality engineers need to combine them in a way that plays to each of their strengths.

According to Montgomery [44], quality professionals need to learn computer science skills and data mining techniques to thrive within *Industry 4.0*. While it is clear that *MLA* can deal more effectively with the volume and complexity posed by big data, there is still little research or documentation aimed at creating guidelines for integrating big data with *Six Sigma* [2]. Recently, leading academic and research scholars have turned their attention towards addressing this research gap.

3. Process Monitoring for Quality – a Quality 4.0 Initiative

In the context of *Industry 4.0*, *PMQ* [1] is a *Quality 4.0* initiative that systematically guides the application of *AI* to *IBD* to generate value, Fig. 2. It is a blend of *Process Monitoring (PM)* and *Quality Control (QC)* (Fig. 3) aimed at real-time defect detection and empirical knowledge discovery, where detection is formulated as a binary classification problem. *PMQ* is founded on *Big Models (BM)* [19], a predictive modeling paradigm that applies machine learning, statistics, and optimization to process data to develop the classifier, Fig. 4. Data mining –empirical– results help to identify the driving features of the system and uncover hidden patterns. This information is further investigated by domain knowledge experts to generate a new set of hypotheses that are tested by experimental means, i.e., design of experiments. Discovered information is used to augment human troubleshooting and guide process redesign and improvement. Fig. 5 shows the conceptual framework of *PMQ*.

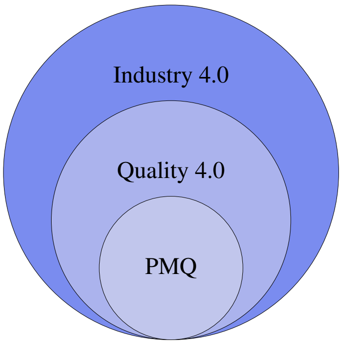


Fig. 2: *PMQ* in the context of *Industry 4.0* [24].

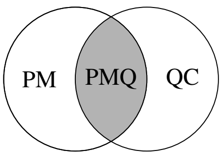


Fig. 3: Process monitoring for quality [1].

In the era of *Industry 4.0*, quality benchmarks are very high. However, although most manufacturing processes generate only a few *Defects Per Million of Opportunities (DPMO)*, customers expect perfect quality. A single warranty event can significantly impact the company's reputation. Therefore, rare quality event detection is one of the most relevant challenges addressed by *PMQ*. The *BM* learning paradigm is founded on

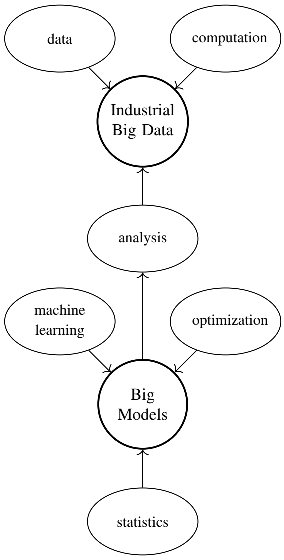


Fig. 4: Industrial big data big models concept [1].

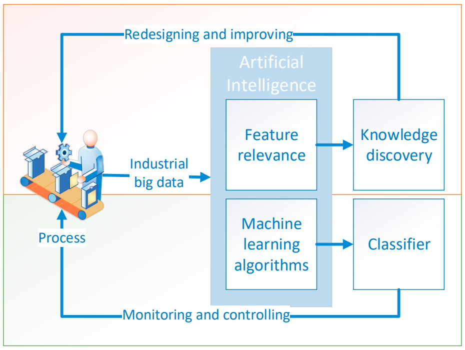


Fig. 5: *PMQ* conceptual framework.

ad hoc learning methods to effectively analyze these data structures.

3.1. Binary Classification of Quality

In a binary classification of quality problem, a positive result refers to a defective item, and a negative result refers to a good quality item, Formulation 1.

$$Label_i = \begin{cases} 1 & \text{if } i^{th} \text{ item is defective } (+) \\ 0 & \text{if } i^{th} \text{ item is good } (-) \end{cases} \tag{1}$$

The confusion matrix [29] is a table used to summarize the predictive performance of a classifier, Table 1. The *TP* is a defective item correctly classified, the *FP* is a good quality item classified as defective. The *TN* is a good quality item correctly classified, whereas the *FN* is a defective item not detected. The type-I (α) error refers to the *FP* rate, and the type-II (β) error describes the *FN* rate (i.e., missing rate). In this context, β is the probability that a defective item will be missed by the classifier.

Table 1: Confusion matrix.

	Predicted good	Predicted defective
Good item	True Negative (TN)	False Positive (FP)
Defective item	False Negative (FN)	True Positive (TP)

3.2. Applications

Though quality inspections are widely practiced before, during, and after production, they still highly rely on human capabilities. According to a recent survey, almost half of the respondents claimed that their inspections were mostly manual (i.e., less than 10% automated) [8]. *PMQ* proposes to use real-time process data to automatically monitor and control the processes, i.e., identify and eliminate defects.

This application has the desirable characteristics of a machine learning project. The basic objective is to learn a repetitive, simple mental concept performed by inspectors, where the task is formulated as a binary classification problem.

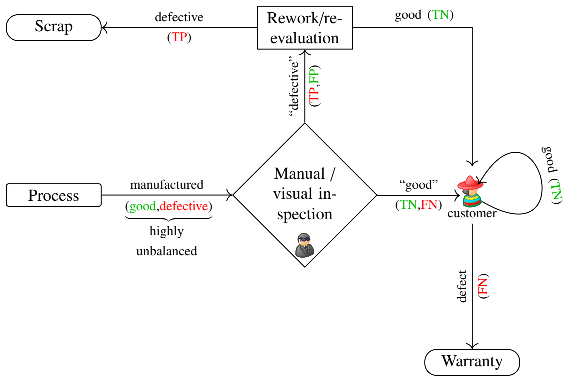
To demonstrate how *PMQ* advances the state of the art of quality, three traditional *QC* scenarios without *AI* are analyzed in Fig. 6. Then, their counterparts are presented in Fig. 7.

A typical manufacturing process generates only a few *DPMO*, Fig. 6. The majority of these defects are detected (*TP*) by either a manual/visual inspection, Fig. 6(a) or by a *SPC/SQC* system, Fig. 6(b). Detected defects are removed from the value-adding process for a second evaluation, where they are finally either reworked or scrapped. Since neither inspection approaches are 100% reliable[55, 68], they can commit *FP* (i.e., call a good item defective) and *FN* (i.e., call a defective item good) errors. Whereas *FP* create the hidden factory effect by reducing the efficiency of the process, *FN* should always be avoided.

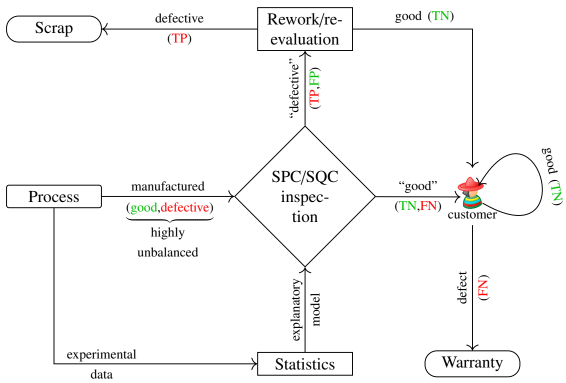
In extreme cases, Fig. 6(c), time-to-market pressures may compel a new process to be developed and launched even before it is totally understood from a physics perspective. Even if a new *SPC/SQC* model/system is developed or a pre-existing model or system is used, it may not be feasible to measure its quality characteristics (variables) within the time constraints of the cycle time. In these intractable or infeasible cases, the product is launched at a high risk for the manufacturing company.

The *BM* learning paradigm is applied to design a classifier with high defect detection capacity to be deployed at the plant, e.g., *final model*, Fig. 7. This data-driven method is applied to eliminate manual or visual inspections, as well as to develop an empirical-based *QC* system for the intractable and unfeasible cases, Fig. 7(a).

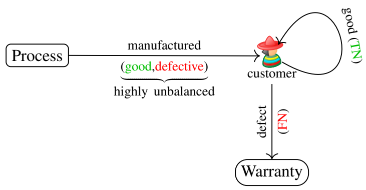
In a process statistically under control, *PMQ* is applied to detect those few *DPMO* (*FN*) not detected by the *SPC/SQC* system to enable the creation of virtually defect-free processes through perfect detection [22], Fig. 7(b). A full analysis of the *PMQ* applications in [24, 27].



(a) Manual/visual control.



(b) Statistical control.

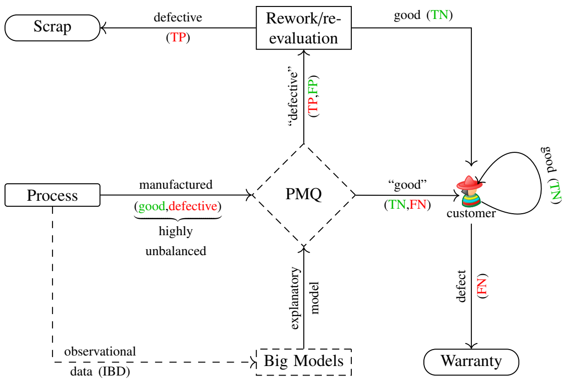


(c) Intractable/unfeasible control.

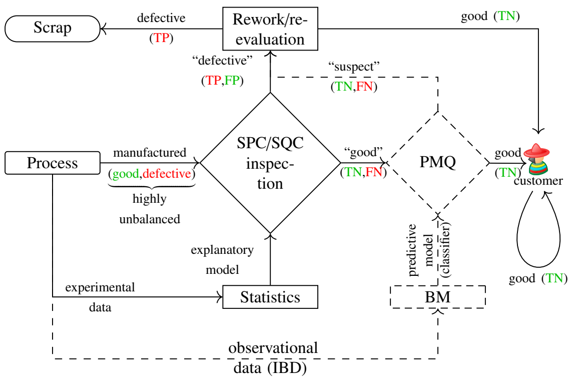
Fig. 6: Traditional quality control scenarios.

4. The Problem Solving Strategy

IADLPRR–Identify, Acsensorize, Discover, Learn, Predict, Redesign, Relearn– is the *PMQ*'s signature framework for value creation Fig. 8. This 7-step problem solving strategy systematically drives innovation, process control, and improvement. It helps to identify and select high value projects that can be formulated as machine learning problems with a high likelihood of success. For illustrative and numerical applications of these steps refer to [1, 19, 22–27].



(a) Empirical-based (data-driven) control.



(b) Boost a process statistically under control.

Fig. 7: *PMQ* applications.

4.1. Identify

Project selection is the main goal of this step. The success of a *Quality 4.0* project relies primarily on the ability of management to identify and select high value/impact projects with a high likelihood of success. Including "quick win" projects is also important to develop momentum and create trust around the new technologies.

The hype and success stories surrounding *AI* has generated interest from quality leaders in deploying an *AI* initiative. But according to recent surveys, 80-87% of projects never make it into production [52, 58]. Improper project selection is the primary cause of this discouraging statistic, as many of these projects are ill-conditioned from the launch.

To address this challenge, an ad hoc approach was developed: a weighted project decision matrix for *Quality 4.0* based on 18 questions. To access these questions, readers are referred to [24]. This approach evaluates many aspects of the potential projects (value, feasibility, availability of data, strategy, expertise, and time) to develop a prioritized portfolio. This activity should be led by the Master Black Belt in collaboration with a cross-functional team with the domain knowledge of the application, to ensure Black and Green Belts are working on the right projects.

4.2. Acsensorize

To generate the capability of observing the process is the main goal of this step. *PMQ* proposes to acsensorize[1] (i.e., observe) the system to generate the raw empirical data. The level of difficulty in the *Identify* phase in Figure 8 is also influenced by the availability and observability of data. Assessment entails enlisting all the available data streams that an existing process ($MRL \geq 5$) may have, or is possible to have, for a process that has not yet been deployed ($MRL < 5$). Sensors could pertain to a specific stage in a manufacturing process, or the process in its entirety. In either scenario, sensors should cover various sensing modalities and aspects with minimum, but non-zero, overlap to ensure some redundancy.

Sensors add complexity, computational, and archival burden on the system. It's thus imperative to make a judicious choice of sensors. Some observable physical parameters like dimension, current, voltage, temperature, pressure, etc. could be sensed directly. Non-observable parameters could be inferred from a combination of observable ones, as in soft sensors [28]. The best sensors and inference methods are the ones that have maximum separation between signal and noise subspaces and contain sufficient discriminative capacity.

Data collected could be streamed wirelessly if the bandwidth requirement is low enough. A true wireless sensor however, is the one that either has onboard energy source or could harvest energy. Once collected, data should be communicated using non-proprietary standard protocols. Every equipment and transducer manufacturer today has their own method of choice, which makes it nearly impossible for the customer to integrate equipments from various vendors. To avoid this, it is strongly recommended to store data in standard architectures in databases and open format binary files. Note, non-proprietary and open formats should not be confused with unencrypted/unsecured.

4.3. Discover

Feature creation and data labeling are the main goals of this step. Feature engineering is one of the most important steps for machine learning. Features are low dimensional representations of data [12]. Their origins are traced back to a data set in which the readings are usually vertically and horizontally fragmented. Then, features are extracted and transformed into a vertical representation to create the matrix with the training set, Fig. 9.

For the time-series domain with one measured parameter, features are based on statistical moments, amplitude, and entropy. For the frequency domain, the most important features

[1] Act of adding a multitude of dissimilar sensors, generally of a variety of sensing modalities, to an existing system that may or may not already have sensors. Acsensorizing plays a significant role in big data research and machine learning.

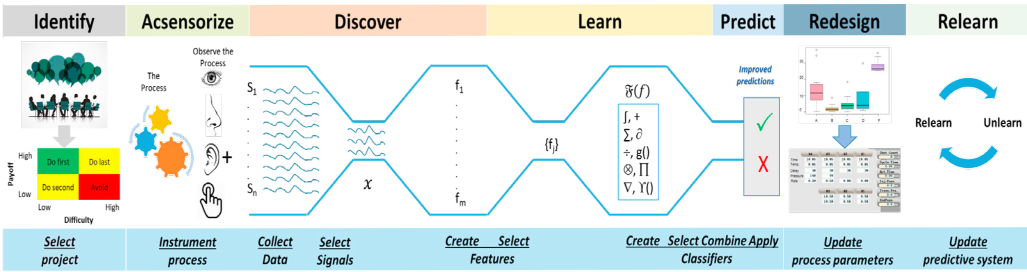


Fig. 8: Comprehensive problem solving strategy for Quality 4.0.

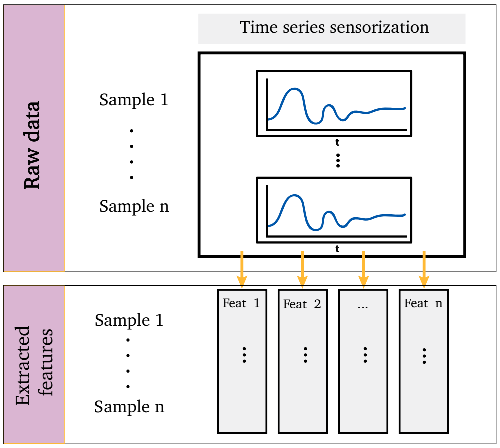


Fig. 9: Feature creation example for time series data [12]

are based on the power spectrum, spectral information, and entropy. For the combination of time-frequency domain, features are based on energy, instantaneous frequency, and entropy, statistical equations in [12]. For image processing, shape, color, and texture are good descriptors [48].

Data labeling is the process of assigning classes to each sample so that *MLA* can learn from it. Usually human intelligence is required to label the data. The combination of the features and labels generates the learning data set, Fig. 10.

Sample	Feature1	Feature2	Feature3	.	.	.	Featuren	Label
1	-0.57	-0.88	-1	-0.86	-0.51	-1.01	-1.03	0
2	0.17	0.17	0.25	0.03	0.67	0.39	0.09	1
3	1.11	1.26	1.1	1.31	1.42	1.17	1.15	0
.	0.24	0.18	0.3	0.36	0.32	0.55	0.22	0
.	-1.56	-1.15	-.33	-1.52	-1.34	-1.27	-1.24	1
.	-1.73	-1.56	-1.47	-1.52	-1.55	-1.44	-1.37	0
m	-0.41	-0.5	-0.52	-0.14	-0.46	-0.33	-0.23	1

Fig. 10: Learning data

4.4. Learn

Classifier development is the main goal of this step. The classifier uses real-time observational data to virtually project each manufactured item into a hyper-dimensional space where those rare quality events can be detected. Manufacturing systems are dynamic and complex entities that pose specific challenges that must be understood and addressed from a technical perspective. General insights from other domains where prediction is the main goal (e.g., Netflix recommendation system) tend not to effectively transfer into manufacturing.

Manufacturing-derived data for binary classification of quality poses the following challenges: (1) hyper-dimensional feature spaces, including relevant, irrelevant, trivial, and redundant, (2) highly/ultra unbalanced (minority/defective class count $< 1\%$), (3) mix of numerical and categorical variables (i.e., nominal, ordinal or dichotomous), (4) different engineering scales, (5) incomplete data sets, and (6) time-dependency. To effectively address these challenges, this section is broken down into three sub-activities majorly performed in a lab-environment: (1) preprocessing, (2) classifier development, and (3) deployment challenges.

4.4.1. Preprocessing

After feature creation, preprocessing is the next step for data cleaning and improving. After the preprocessing, the training set is obtained and presented to the *MLA*. This section provides an overview of the general steps and tools used for this purpose.

- **Exploratory data analysis**, class distribution analysis [40], feature distribution and pairwise analysis [50], and outlier identification [41] help to develop an effective data set and learning strategy.

- **Transformation of numerical data**, since most *MLA* work internally with numerical data, it is important to develop a strategy to deal with categorical variables [66] (i.e., encode them in a numeric form). For binary features, it is recommended to use *effect coding* (-1 and 1) instead of *dummy coding* (0 and 1) [54]. Moreover, since features tend to have different engineering scales, it is important to normalize or standardize the data before presenting it to the algorithm. Feature scaling generally speeds up learning and leads to faster convergence [36]. More insights about when to normalize or standardize can be found in [38, 43].

- **Missing records analysis**, it is important to understand the missing data mechanism to effectively deal with the missing data. Deleting rows or columns with missing

Table 2: Characteristics of the *MLA*

Index	*MLA*	Linear	Nonlinear	Parametric	Nonparametric	Stable	Unstable	Gen	Dis
1	SVM	✓		✓		✓			✓
2	LR	✓		✓		✓			✓
3	NB		✓*	✓		✓		✓	
4	KNN		✓		✓	✓			✓
5	ANN		✓	✓**			✓		✓
6	SVM(RBF)		✓	✓	✓	✓			✓
7	RF		✓		✓	✓			✓
8	RUSBoost		✓		✓	✓			✓

* with numeric features
** with a set of parameters of fixed size
Gen: Generative
Dis: Discriminative.

data is a widely used approach. This method is not advised unless the proportion of eliminated records is very small (<5%) [37]. Imputing the missing records is a better approach, this statistical technique refers to the process of replacing missing data with guessed/estimated values. Although imputation can be applied to preserve all samples, it relies on specific assumptions often unrealistic which can potentially bias results. A review of imputation methods can be found in [6, 37, 51]. Finally, permuting the rows and columns to maximize the information is a different approach [20]. This method does not induce any bias to the data set.

- **Feature selection**, eliminating irrelevant and redundant features improves generalization, eases data collection and information extraction, reduces computing times and the effect of dimensionality [33, 47, 56, 62, 67, 70]. At this stage, filter methods are applied for this task. For highly/ultra unbalanced data containing only numerical features, this separability index-based feature selection method shows superior performance [21].

4.4.2. Classifier Development

To induce information extraction and engineering model trust, the *BM* learning paradigm is founded on the principle of parsimony [9]. Since there is no a priori distinction between *MLA* [65], eight common and diverse *MLA* are proposed[2], Table 2. Due to the high conformance rate in manufacturing, binary classification of quality data sets tend to be highly/ultra unbalanced. To address this challenge the *MPCD* is used to evaluate classification performance and *MLA* hyperparameter tuning. Due to the time effect, models are usually validated following a time-ordered hold-out scheme. The training set is partitioned in training, test, and holdout sets, where the latest set is used to emulate deployment performance and compare it to the learning targets to demonstrate feasibility.

[2] Authors acknowledge that some algorithms can change their taxonomy (e.g., from parametric to non-parametric) depending upon their definition.

4.4.3. Deployment Challenges

In-lab solutions tend to be an overoptimistic representation of predictive system capabilities, since lab data is generated under highly controlled conditions. These conditions are likely not entirely representative of the plant environment. Moreover, the data snooping effect should be considered, as the *MLA* are very powerful and can over-fit (i.e., learn spurious patterns [9]) the training data, which would result in a low generalization performance. For this reason, lab-generated data is useful for developing proof-of-concept models. Pilot runs are required to obtain an unbiased generalization performance.

4.5. Predict

Prediction optimization is the main objective of this step. Once a diverse group of classifiers has been developed, the next natural step is to explore different combination schemes to improve prediction. A *Multiple Classifier System (MCS)* is a powerful solution to difficult pattern recognition problems, as they usually outperform the best individual classifier [14]. To design a *MCS*, an appropriate fusion method is required to combine the individual classifier outputs optimally to determine the final decision (classification). An ad hoc *MPCD* prediction optimization algorithm is presented in [23]. It addresses four specific questions: (1) which classifiers should be included? (2) how should their predictions (labels) be combined? (3) which fitness function should be optimized? and (4) which optimization solver should be used?.

4.6. Redesign

Empirical knowledge discovery is the main goal of this step. Data mining results are presented to the engineering team for analysis and interpretation. These empirical studies can provide rich, deep contextual data valuable at understanding a phenomenon, but cannot be generalized to establish prevalence of the underlying physics of the system. To do so, experimental data must be generated. Therefore, extracted information is used to generate useful hypotheses about possible connections between the features and the quality of the product. Then, statistical analyses (e.g., randomized experiments) can be devised to establish causality to augment root-cause analyses and to find optimal parameters to redesign the process. More insights about this concept in [24] and a real case study in [22].

4.7. Relearn

Relearning strategy development is the main goal of this step. Developing an in-lab model with the capacity to predict well the items in the test set and satisfy the learning targets for the project, is only the beginning since the model will usually degrade after deployment.

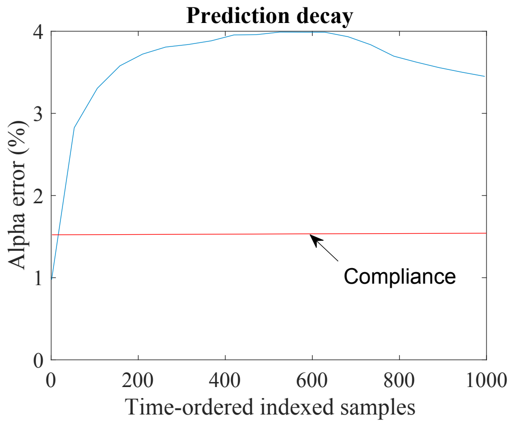
In machine learning, the concept of drift [63, 64] embodies the fact that the statistical distributions of the classes of which the model is trying to predict, change over time in unforeseen ways. This poses difficulties, as the predictive models assume a static relationship between input and output variables. This static assumption is rarely satisfied in manufacturing. The transient and novel sources of variations cause manufacturing systems to exhibit non-stationary data distributions. Consequently, the prediction capability of a trained model tends to significantly degrade overtime.

In Fig. 11(a), a real situation is presented, in which the model exhibited less than 1.5% of α error (target set by the plant) in the test set (lab environment) to satisfy the defect detection goals ($\beta < 5\%$) of the project. Immediately after deployment, the α error increased to 1.78%. A few days later, the α increased to 4%, an unacceptable *FP* rate for the plant. Although this model exhibited good prediction ability and made it into production, it was not a sustainable solution, and therefore, never created value.

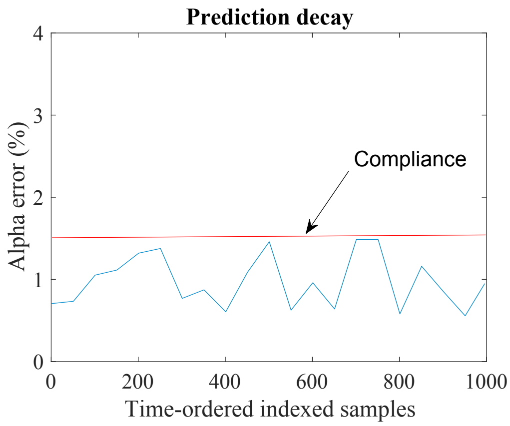
The main goal of relearning, is to keep the predictive system in compliance with the restrictions set by the plant (α error), Fig. 11(b), and the detection goals (β error). This is accomplished by ensuring that the algorithm is learning the new statistical properties of both classes (good, defective). Continual learning or auto-adaptive learning is a fundamental concept in *AI* that describes how the algorithms should autonomously learn and adapt in production as new data with new patterns comes in. A relearning scheme should include the following four components:

(1) Learning strategy, (2) Relearning data set, (3) Relearning schedule, and (4) Monitoring system. The full relearning scheme and insights presented in [24].

Figure 12 describes the online deployment and offline relearning concept. As described in this image, on deployment, the predictive system uses the process data to monitor quality. If a defective item is detected, it is sent to the rework/re-evaluation station for a more detailed inspection. On the other hand, the offline learning concept refers to the relearning schedule, which can occur, for example, between shifts or every night. The relearning procedure follows the learning strategy previously defined–in the lab–by the data science team. The data generated from the rework/re-evaluation station along with the new process-data generated during deployment is used to generate the relearning data set. This strategy includes the ad hoc analytical tools for feature creation, model creation, and classifier fusion.



(a) Without relearning scheme.



(b) With relearning scheme.

Fig. 11: Error analysis after deployment [24].

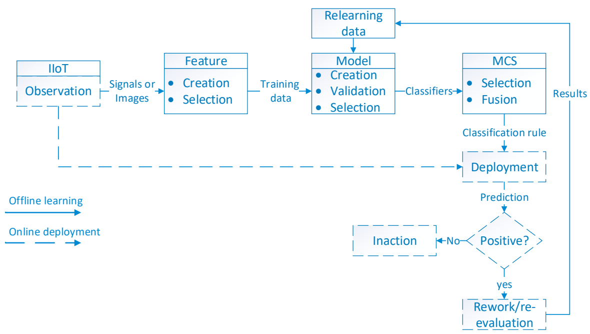


Fig. 12: Online deployment and offline learning framework.

5. Green, Black and Master Black Belt Certifications in Quality 4.0

Quality 4.0 is built on the statistical and managerial foundations of the previous philosophies. It also follows a problem solving paradigm that systematically drives innovation and im-

Table 3: Certification requirements by belt.

Belt	Education	Training data	Machine Learning Theory	Machine learning algorithms	Programming	Experience (years)	Problem solving strategy	Project
Green	Bachelor's	Understands the 10 V's, creates features, and preprocessing techniques	Understands model validation and feature selection methods, bias-variance tradeoff, and generalization	Understands and trains	Applies available libraries for learning feasibility analyses	1	Understands the seven steps	1 (3-steps)
Black	Bachelor's or Master's	Understands the 10 V's, creates features, understands the learning curves and preprocessing techniques	Understands model validation and feature selection methods, bias-variance tradeoff, and generalization	Understands, trains, combines, and develops	Applies available libraries for learning feasibility analyses and develops a full sustainable solution	2	Understands the seven steps and coaches Green Belts	1 (7-steps)
Master Black Belt	Master's or PhD	Understands the 10 V's, creates features, understands the learning curves, and preprocessing techniques	Understands model validation and feature selection methods, bias-variance tradeoff, and generalization	Understands, trains, combines, and develops	Applies available libraries for learning feasibility analyses and develops a full sustainable solution	3	Understands the seven steps and coaches Green and Black Belts	1 peer-reviewed published

provement. However, the new curriculum is not only based on traditional statistics, it has been updated with machine learning, optimization, and computer programming.

Following the *Six Sigma* original color-convention, three different levels of competencies for *Quality 4.0* certifications are identified: Green, Black and Master Black Belts. Each color requires different levels of knowledge, education, and experience as described in Table 3. The curricula focuses on training engineers, managers, and directors in *PMQ*, or broadly speaking, in how to generate value out of *IBD* and *AI*.

5.1. Green Belt Requirements

The Green Belt certificate holder must have a bachelor's degree in *Science, Technology, Engineering or Math (STEM)* with at least one year of experience as a data scientist. This scientist should understand the meaning and implications of the 10 V's of *IBD* [1, 30, 60]. They must know how to create features out of signals or images. They also must understand how to use and apply preprocessing techniques to effectively deal with incomplete data sets and different engineering scales. Green Belts must also demonstrate an understanding of basic machine learning theory, including: model validation and feature selection methods, bias-variance tradeoff, and generalization evaluation metrics. They must be able to train the nine *MLA* proposed by- and following the- *BM* learning paradigm. They also must have the ability to write the basic code to run these analyses and understand each of the seven steps of the problem solving strategy to partially contribute in the full solution cycle. They must document a project to demonstrate their ability to successfully apply at least three steps of the problem solving strategy. Data generation and initial feasibility analyses are performed by a Green Belt.

5.2. Black Belt Requirements

Black Belts are tech-savvy with the capacity to develop –in collaboration with a cross functional team– a sustainable solution from scratch. In addition to the Green Belt requirements, Black Belts should understand the learning curves to guide the generation of data either to generate more data, or more features, or both. They must be able to optimize prediction through a decision combination scheme. They should be proficient in writing the deployment code–learning, relearning. Black Belts must be able to identify the driving features of the system to guide process redesign/optimization and augment human intelligence (i.e., trouble-shooting processes). They must have at least two years of experience as data scientists and document the application of the seven steps of the problem solving theory.

5.3. Master Black Belt Requirements

Master Black Belts serve as mentors of cross-functional teams guided by the Black Belt and act as a bridge between Black Belts and organization management. They are the leaders of the *Identify* stage, so they ensure the data science teams are working on high impact/value projects with high likelihood of success. Master Black Belts must have the ability to write new algorithms and libraries to customize the solutions. To demonstrate this skill, one peer-reviewed paper must be published in *AI*, *IBD* or *Quality 4.0* domain[3]. A Master Black Belt must have at least three years of experience and hold Green and Black Belts.

6. Conclusions

Two decades ago, GE's introduction of the *Six Sigma* program, pioneered the philosophy of quality in mass-manufactured products. Today, *Six Sigma* Black Belts are solving complex problems across all business functional areas. While the *Six Sigma* philosophy is still necessary, it cannot efficiently/effectively address some of the challenges posed by industrial big data. *Quality 4.0* is the next natural step in the evolution of quality, which uses *Six Sigma* principles in conjunction with big data methods. Strategically investing in *Quality 4.0* training approaches is highly recommended for manufacturing companies to maintain competitiveness.

This paper presents a *Quality 4.0* initiative and a certification program for quality/manufacturing engineers, managers, and directors. The proposed curriculum combines six areas of knowledge (statistics, quality, manufacturing, programming, learning, optimization) and a 7-step problem solving strategy that must be mastered to obtain a certification.

Following the *Six Sigma* original color-convention, three different levels of competencies for *Quality 4.0* certifications are proposed: Green, Black and Master Black Belts. Each color requires different levels of knowledge, education and experience.

Certified professionals in this initiative will be able to take the lead in deploying a *Quality 4.0* initiative. They will have the capacity to identify engineering intractable problems that can be formulated as machine learning–binary classification–problems and successfully solve them.

From a career growth perspective, obtaining a Black or a Master Black Belt certification, is an efficient and effective way to advance to new professional levels.

Similar to the adoption of *Six Sigma* at GE two decades ago, early adopters of *Quality 4.0* will join the circle of the most influential manufacturing companies in the world and position themselves for success. This paper offers a vision of how to launch a *Quality 4.0* initiative.

A final consideration with respect to deep learning is worth mentioning. Since most of today's problems can be more effectively solved by simple machine learning algorithms rather than deep learning [7], the problem solving strategy is based on machine learning. Therefore, if a problem is more suitable for deep learning, not all steps must be applied. For example, the discover step is embedded in the deep learning architecture; however, the fundamental concepts of the *Quality 4.0* paradigm remain the same.

[3] This publication does not necessarily need to include the seven steps of the problem solving strategy, since that is a Black Belt requirement

References

[1] Abell, J.A., Chakraborty, D., Escobar, C.A., Im, K.H., Wegner, D.M., Wincek, M.A., 2017. Big Data Driven Manufacturing — Process-Monitoring-for-Quality Philosophy. ASME J of Manufacturing Science and Eng on Data Science-Enhanced Manufacturing 139.

[2] Antony, J., Dempsey, M., Brennan, A., Farrington, T., Cudney, E., 2019. An Evaluation into the Limitations and Emerging Trends of Six Sigma: An Empirical Study . The TQM Journal 31, 205–221.

[3] Armstrong, K., Parmelee, L., Santifort, S., Burley, J., Van Fleet, J., 2018. Preparing tomorrow's workforce for the fourth industrial revolution for business: A framework for action.

[4] ASQ-b , 2020. Six Sigma Belts, Executives and Champions – What Does it All Mean? URL: `https://asq.org/quality-resources/six-sigma/belts-executives-champions`.

[5] ASQ-e , 2020. The Difference Certification Makes. URL: `https://asq.org/quality-resources/six-sigma/belts-executives-champions#CertDifference`.

[6] Barnard, J., Meng, X.L., 1999. Applications of multiple imputation in medical studies: from aids to nhanes. Statistical methods in medical research 8, 17–36.

[7] Bathaee, Y., 2017. The artificial intelligence black box and the failure of intent and causation. Harv. JL & Tech. 31, 889.

[8] Belfiore, M., 2016. Automation opportunities abound for quality inspections. Automation World. URL: `https://www.automationworld.com/products/software/article/13315584/automation-opportunities-abound-for-quality-inspections`.

[9] Burnham, K.P., Anderson, D.R., 2003. Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach. Springer Science & Business Media.

[10] Bzdok, D., Altman, N., Krzywinski, M., 2018. Points of significance: statistics versus machine learning.

[11] Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp. 785–794.

[12] Christ, M., Kempa-Liehr, A.W., Feindt, M., 2016. Distributed and parallel time series feature extraction for industrial big data applications. arXiv preprint arXiv:1610.07717 .

[13] Christopher, M., Rutherford, C., 2004. Creating supply chain resilience through agile six sigma. Critical eye 7, 24–28.

[14] Clemen, R.T., 1989. Combining forecasts: A review and annotated bibliography. International journal of forecasting 5, 559–583.

[15] Dan, J., 2017. Quality 4.0 fresh thinking for quality in the digital era. Quality Digest.

[16] De Carlo N, 2007. The complete idiot's guide to lean six sigma. Breakthrough Management Group, New York.

[17] De Mast, J., Lokkerbol, J., 2012. An Analysis of the Six Sigma DMAIC Method from the Perspective of Problem Solving. Int J of Production Economics 139, 604–614.

[18] Demuth, H.B., Beale, M.H., De Jess, O., Hagan, M.T., 2014. Neural Network Design. Martin Hagan.

[19] Escobar, C.A., Abell, J.A., Hernández-de Menéndez, M., Morales-Menendez, R., 2018a. Process-Monitoring-for-Quality — Big Models. Procedia Manufacturing 26, 1167–1179.

[20] Escobar, C.A., Arinez, J., Macias, D., Morales-Menendez, R., 2020a. Learning with incomplete data. To appear in the Conference Proceedings of IEEE BigData .

[21] Escobar, C.A., Arinez, J., Macias, D., Morales-Menendez, R., 2020b. A separability-based feature selection method for highly unbalanced binary data. To apper in International Journal on Interactive Design and Manufacturing .

[22] Escobar, C.A., Arinez, J., Morales-Menendez, R., 2020c. Process-Monitoring-for-Quality—A Step Forward in the Zero Defects Vision, in: SAE Technical Paper. URL: `https://doi.org/10.4271/2020-01-1302`, doi:`10.4271/2020-01-1302`.

[23] Escobar, C.A., Macias, D., Hernández-de Menéndez, M., Morales-Menendez, R., 2021a. Process monitoring for quality — multiple classifier

system for highly unbalanced data. Studies in Big Data, Springer .
[24] Escobar, C.A., McGovern, M., Morales-Menendez, R., 2021b. Quality 4.0 — a review of big data challenges in manufacturing. To Appear in the Journal of Intelligent Manufacturing .
[25] Escobar, C.A., Morales-Menendez, R., 2019a. Process-monitoring-for-quality—a model selection criterion for l1-regularized logistic regression. Procedia Manufacturing 34, 832–839.
[26] Escobar, C.A., Morales-Menendez, R., 2019b. Process-monitoring-for-quality—a model selection criterion for support vector machine. Procedia Manufacturing 34, 1010–1017.
[27] Escobar, C.A., Wincek, M.A., Chakraborty, D., Morales-Menendez, R., 2018b. Process-Monitoring-for-Quality — Applications. Manufacturing Letters 16, 14–17.
[28] EY; Oxford, A., 2019. Sensors as drivers of industry 4.0 - a study on germany, switzerland and austria. URL: https://www.oxan.com/insights/client-thought-leadership/ey-sensors-as-drivers-of-industry-40/.
[29] Fawcett, T., 2006. An Introduction to ROC Analysis. Pattern Recognition Letters 27, 861–874.
[30] Fernando, L., 2017. 7 v's of big data. URL: http://blogsofdatawarehousing.blogspot.com/2017/01/7-vs-of-big-data.html.
[31] Ghias, S., 2020. Should i get a six sigma black belt? Investopedia. URL: https://www.investopedia.com/articles/investing/102014/guide-six-sigma-black-belt.asp.
[32] Goasduff, L., 2020. 2 megatrends dominate the gartner hype cycle for artificial intelligence, 2020. Gartner. URL: https://www.gartner.com/smarterwithgartner/.
[33] Hall, M., 2000. Correlation-based Feature Selection of Discrete and Numeric Class Machine Learning, in: Proc of the 17^{th} Int Conf on Machine Learning, University of Waikato. pp. 359–366.
[34] Hansen, M.C., 2019. Salary survey 2019: The complete report. Quality Progress 52.
[35] INCOSE, 2020. Certification program history. URL: https://www.incose.org/systems-engineering-certification/certification-program-history. accessed on:.
[36] Ioffe, S., Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, in: International Conference on Machine Learning, pp. 448–456.
[37] Jakobsen, J.C., Gluud, C., Wetterslev, J., Winkel, P., 2017. When and how should multiple imputation be esed for handling missing data in randomised clinical trials–a practical guide with flowcharts. BMC medical research methodology 17, 162.
[38] Juszczak, P., Tax, D., Duin, R.P., 2002. Feature scaling in support vector data description, in: Proc. asci, Citeseer. pp. 95–102.
[39] Kosowatz, J., 2018. Salaries are up for mechanical engineers. The American Society of Mechanical Engineers.
[40] Laurikkala, J., 2001. Improving identification of difficult small classes by balancing class distribution, in: Conference on Artificial Intelligence in Medicine in Europe, Springer. pp. 63–66.
[41] Maimon, O., Rokach, L., 2005. Data mining and knowledge discovery handbook .
[42] Marr, B., 2017. The biggest challenges facing artificial intelligence (ai) in business and society. 13th July, available at: https://www. forbes. com/sites/bernardmarr/2017/07/13/the-biggest-challengesfacing-artificial-intelligence-ai-in-business-andsociety/3 .
[43] Mohamad, I.B., Usman, D., 2013. Standardization and its Effects on K-means Clustering Algorithm. Research J of Applied Sciences, Eng and Technology 6, 3299–3303.
[44] Montgomery, D.C., 2014. Big data and the quality profession. Quality and Reliability Engineering International 30, 447–447.
[45] Muro, M., Whiton, J., Maxim, R., 2019. What jobs are affected by ai?: better-paid, better-educated workers face the most exposure.
[46] NewVantage Partners, L., 2019. Big data and ai executive survey 2019: Data and innovation how big data and ai are accelerating business transformation.
[47] Nicodemus, K., Malley, J., 2009. Predictor Correlation Impacts Machine Learning Algorithms: Implications for Genomic Studies. Bioinformatics 25, 1884–1890.
[48] de Oliveira, M.S., Betting, L.E., Mory, S.B., Cendes, F., Castellano, G., 2013. Texture analysis of magnetic resonance images of patients with juvenile myoclonic epilepsy. Epilepsy & Behavior 27, 22–28.
[49] Ortega, M., 2018. Cio survey: Top 3 challenges adopting ai and how to overcome them. Databricks. URL: https://databricks.com/.
[50] Park, S.H., Fürnkranz, J., 2007. Efficient pairwise classification, in: European Conference on Machine Learning, Springer. pp. 658–665.
[51] Rahman, M.M., Davis, D.N., 2013. Machine learning-based missing value imputation method for clinical datasets, in: IAENG transactions on engineering technologies. Springer, pp. 245–257.
[52] Research, G., 2018. Predicts 2019: Data and analytics strategy. URL: https://emtemp.gcom.cloud/ngw/globalassets/en/doc/documents/374107-predicts-2019-data-and-analytics-strategy.pdf.
[53] salary.com, 2020. Six sigma black belt salary in the united states. URL: https://www.salary.com/research/salary/alternate/six-sigma-black-belt-salary.
[54] Sarle, W.S., 2002. comp.ai.neural-nets faq, part 2 of 7: Learning. faqs.org. URL: http://www.faqs.org/faqs/ai-faq/neural-nets/part2/.
[55] See, J.E., 2015. Visual inspection reliability for precision manufactured parts. Human factors 57, 1427–1442.
[56] Shao, C., Paynabar, K., Kim, T., Jin, J., Hu, S., Spicer, J., Wang, H., Abell, J., 2013. Feature Selection for Manufacturing Process Monitoring using Cross-Validation. J. of Manufacturing Systems 10.
[57] Shmueli, G., 2010. To Explain or to Predict? Statistical Science , 289–310.
[58] STAFF, V., 2019. Why do 87% of data science projects never make it into production? URL: https://venturebeat.com/.
[59] Staley, O., 2019. Whatever happened to six sigma? Quartz at Work. URL: https://qz.com/work/1635960/whatever-happened-to-six-sigma/.
[60] Uddin, M.F., Gupta, N., et al., 2014. Seven v's of big data understanding big data to extract value, in: Proceedings of the 2014 zone 1 conference of the American Society for Engineering Education, IEEE. pp. 1–5.
[61] University, V., 2020. Six sigma or big data? why not both? Online. URL: https://www.villanovau.com/resources/six-sigma/six-sigma-or-big-data-why-not-both/.
[62] Wang, F., Yang, Y., Lv, X., Xu, J., Li, L., 2014. Feature Selection using Feature Ranking, Correlation Analysis and Chaotic Binary Particle Swarm Optimization, in: 5^{th} Int Conf on Software Eng and Service Science, pp. 305–309.
[63] Wang, H., Abraham, Z., 2015. Concept drift detection for streaming data, in: 2015 International Joint Conference on Neural Networks (IJCNN), IEEE. pp. 1–9.
[64] Webb, G.I., Lee, L.K., Petitjean, F., Goethals, B., 2017. Understanding concept drift. arXiv preprint arXiv:1704.00362 .
[65] Wolpert, D.H., 1996. The Lack of a Priori Distinctions Between Learning Algorithms. Neural Computation 8, 1341–1390.
[66] Wright, M.N., König, I.R., 2019. Splitting on categorical predictors in random forests. PeerJ 7, e6339.
[67] Wu, S., Hu, Y., Wang, W., Feng, X., Shu, W., 2013. Application of Global Optimization Methods for Feature Selection and Machine learning. Mathematical Problems in Eng .
[68] Wuest, T., Irgens, C., Thoben, K., 2013. An approach to quality monitoring in manufacturing using supervised machine learning on product state based data. Journal of Intelligent Manufacturing, XX .
[69] Yang, K., Basem, E., 2003. Design for Six Sigma: Roadmap to Product Development. The McGraw-Hill Companies, Inc, New York.
[70] Yu, L., Liu, H., 2003. Feature Selection for High-Dimensional Data: A Fast Correlation-based Filter Solution, in: ICML, pp. 856–863.