# Machine Learning and Deep Learning applied to End-of-Line Systems: A review

CARLOS NUNES, E. J. SOLTEIRO PIRES, ARSENIO REIS
Escola de Ciˆencias e Tecnologia, Universidade de Tr´as-os-Montes e Alto Douro, Vila Real, PORTUGAL

Abstract: This paper reviewed machine learning algorithms, particularly deep learning architectures applied to end-of-line testing systems in industrial environment. In industry, data is also produced when any product is being manufactured. All this information registered when manufacturing a specific product can be manipulated and interpreted using Machine Learning algorithms. Therefore, it is possible to draw conclusions from data and infer valuable results that can positively impact the future of the production line. The reviewed papers showed that machine learning algorithms play a crucial role in detecting, isolating, and preventing anomalies, helping operators make decisions, and allowing industries to save resources.

Keywords: Machine Learning, Deep Learning, End of Line, Industry, Predictive Maintenance.

Received: June 17, 2021. Revised: June 16, 2022. Accepted: July 18, 2022. Published: August 5, 2022.

## 1. Introduction

With particular relevance in automotive industrialization, industrial quality assurance tools have been developed [1], [2]. Among these tools end-of-line (EOL) testing systems stand out to be indispensable as before selling a product to a customer, these systems ensure that the product undergoes a series of validation tests.

Given the growing demands connected with quality and reliability standards and the huge amount of data generated by the massive realization of EOL tests, this data can be treated through artificial intelligence/ machine learning algorithms [3], to improve the industrialization significantly. With the latest technological advances that have been made in the areas of Machine Learning (ML), and Computer Science (CS) as well as the increased usage of sensors and monitoring systems, predictive maintenance (PdM) approaches can be applied in any industrial equipment [4].

Analyzing the results obtained from logfiles (which store the information from testing processes) allows for a better understanding on the evolution of the data, which can indicate patterns and trends. Detecting such patterns and trends can help in the preventive detection of system disruptions before a fault occurs in the system, or corrective maintenance after detecting a fault in the system. Having this type of benefits connected to the manufacturing/testing of products allows industries to forecast maintenance issues related to their EOL systems.

This paper is a review of Machine Learning and Deep Learning techniques applied to EOL testing systems. The role of optimizing this systems with use of such techniques, is that it may help address the production line issues by notifying operators early that preventive actions shall be taken prior to a production stop or discarding a product [5]. Regarding the structure of this review, the following sections will describe the functionality of some ML algorithms and DP architectures together with some applications seen mainly in automotive industry, in EOL testing systems. Section III provides an introduction to machine learning in industry, where some of the most common algorithms will be introduced; Section IV is structured the same way as the previous section but regarding the most common deep learning architectures; Section VI draws the main conclusions from the review.

## 2. Review Methodology

This section exposes the methodology adopted by this review. The goal of the review is to answer the following questions:

Q1. How does machine learning impact the future of end of line testing systems?

Q2. What are the most common applications of these technologies in the manufacturing environment?

To answer these questions, this scientific review adopted a systematic methodology. The information collected was obtained from different papers on the following websites: Science Direct, Ieeexplorer, Springer Online, Willey, and Google Scholar. The search query used was: (“Industry 4.0” AND (“Machine learning” OR “Deep Learning”) AND “End of Line Testing Systems” AND (“smart manufacturing” OR “predictive maintenance” OR “preventive maintenance”) AND “Big Data” AND “automotive Industry”). Occasionally, some books and other websites were also considered. Overall, 180 articles were gathered. As for the inclusion criteria, the papers and works obtained from applying the query in the cited websites, written in English, and published from 2010, were selected. As the areas of ML and DL are in constant development, the search considered the latest papers about these technologies to avoid outdated information. After applying the exclusion criteria, a list with the research results was obtained, which returned 180 publications. The exclusion criteria are the following:

1. Posts with duplicate content were removed - a total of 8 papers;
2. After analyzing the title, abstract, and conclusions, 137 publications were discarded;
3. Publications were excluded due to their content (less relevant or not contextualizing with the topic of this review) - a total of 6 papers.

After applying the methodology, 29 scientific papers were left cited in this review. The diagram presented in Figure 3 represents the entire flow performed in this review, as recommended by PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [6]. Figure 1, shows the reviewed papers according to their published year.

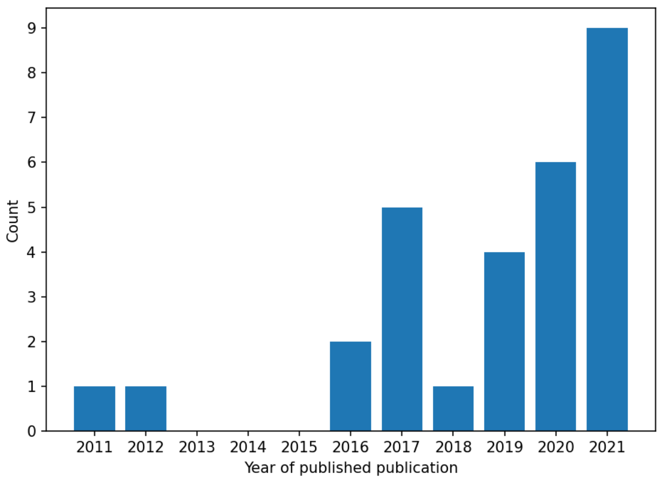


Fig. 1. Year of all the publications reviewed

Figure 2 presents the number of times a specific algorithm or architecture was considered in the reviewed papers. The most reviewed algorithm was the $k$-nn (k-nearest neighbor) algorithm, and the most viewed architectures were CNN (convolution neural networks), RNN (recurrent neural networks), and DT (digital twin).

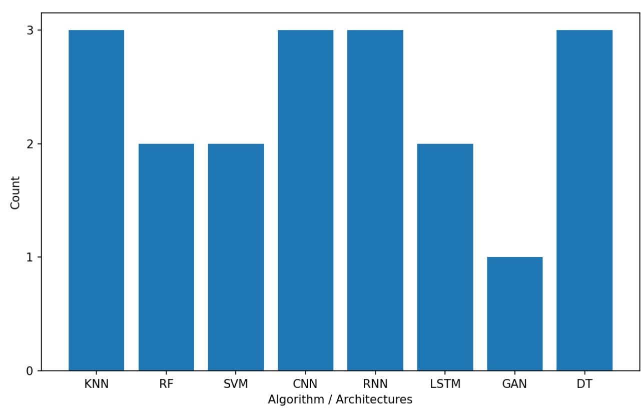


Fig. 2. Algorithms and Architectures reviewed

Table I summarizes the papers considered. Column 1 identifies the article's author, column 2 indicates the methodology used in the work, and column 3 refers to the application used.

# 3. Machine Learning in Industry

Data-driven approaches using machine learning techniques are found to offer promising potential for improved quality control in manufacturing [7]. One of the industry's main

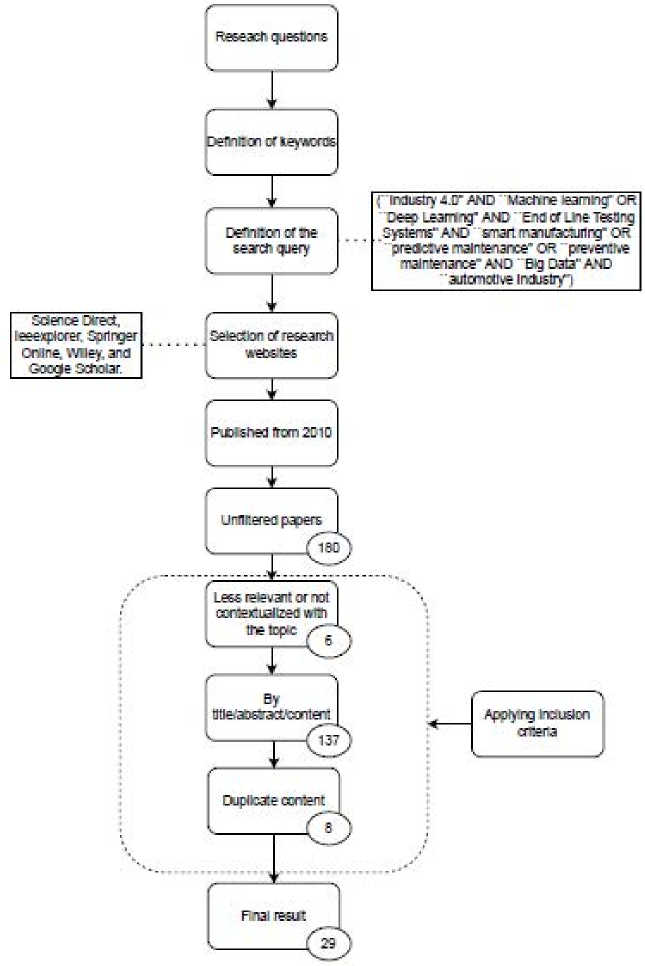


Fig. 3. Review methodology flow diagram adapted from PRISMA [6]

challenges is processing and analyzing large data sets, preferably in real-time. Currently, there are infinite applications of ML algorithms in the industry, from which PdM is the most common form of optimization regarding product manufacturing. Other applications that directly impact a product's reliability are shown to be: either improving industrialization resources, predicting/preventing disruptions and helping with maintenance as shown in [8] and [9]. Nevertheless, eventual approaches will require methods of gathering, treating, and analyzing data. This data usually comes from log files obtained from the production process in EOL systems. The importance of treating this log files has shown to significantly improve the manufacturing processes: making it cheaper, more efficient, less time-consuming, and overall resource-friendly. Before diving into the algorithms, it is also worth mentioning that each algorithm is based on the learning type adopted by the machine. These so-called learning types are supervised learning, unsupervised learning, self-supervised learning, and there are also new emerging ways of learning. Each of the

TABLE I
SUMMARY OF THE REVIEWED PAPERS

| Author | Methodology | Application |
|---|---|---|
| Theirssler et al. [3] | ML | PdM for automotive systems |
| Vicêncio et al. [4] | Data-Driven ML | PdM for EOL systems |
| Hirsch et al. [7] | ML (ensemble learning) | Increase prediction performance |
| Zhang et al. [8] | Data-Driven ML | PdM of industrial equipment |
| Yan et al. [9] | Data-Driven ML | Challenges and applications for PdM |
| Del Rosso et al. [10] | ML | EOL fault detection |
| Verdier and Ferreira [11] | Adaptive Mahalanobis Distance and KNN | Fault detection |
| Zhou et al. [12] | KNN | Fault isolation |
| Wang et al. [13] | RF | Product quality prediction |
| Guo et al. [14] | RF + DT | Fault diagnosis |
| Jalal et al. [15] | SVM | Product Modeling |
| Bodendorf et al. [16] | ML | Estimate product cost |
| Oh et al. [17] | SVM | Real-time quality assessment |
| Elsisi et al. [18] | Deep Learning | Energy management |
| Espinosa et al. [19] | ML + DL | Error detection |
| Vater et al. [20] | CNN | Error detection and correction |
| Park and Yun [21] | CNN | Anomaly detection |
| Huang et al. [22] | RNN | Fault detection |
| Peng et al. [23] | LSTM | PdM under imperfect CSI |
| Lindemann et al. [24] | LSTM | Anomay detection |
| Wang et al. [25] | GAN | Data augmentation |
| Balderas et al. [26] | DT + Simulink model | Printed Circuit Board design and processing |
| Aheleroff et al. [27] | DT | DT as a serice |

types of learning mentioned allows teaching the machine. The algorithm learns through a network (or model) using known examples in the first case. Thus, it can perform classifications or regressions (supervised learning). In the second case, the machine generates clusters that label unlabeled data (unsupervised learning) or even allow the machine to label the data by itself and make predictions (self-supervised learning).

### 3.1 K-Nearest Neighbors

This algorithm is a supervised learning classification type. This algorithm has been enhanced since its reveal [28]. The user defines the number of neighbors considered when classifying a new point, see Figure 4. The uncategorized point is classified as a member of the most crowded class in the neighborhood. The $K$-nn algorithm might not be the best method to classify massive datasets as it decreases run time speed. Still, it is worth looking at how it helps solve problems in the industry due to its simplicity.

*Applications:* The $K$-nn algorithm has played an important role in an EOL quality control procedure based on vibrational analysis [10]. The product under study in this particular paper was induction motors. Some features from the testing process were selected and fed into a classification model based on $k$-nn. This model was able to identify mechanical failures such as damaged bearings or rotor faults. Mainly, $K$-nn has been exploited for fault detection and isolation in different production lines throughout the industry [11], [12]. This type of application can be developed for end-of-line testing systems to detect defective anomalies.

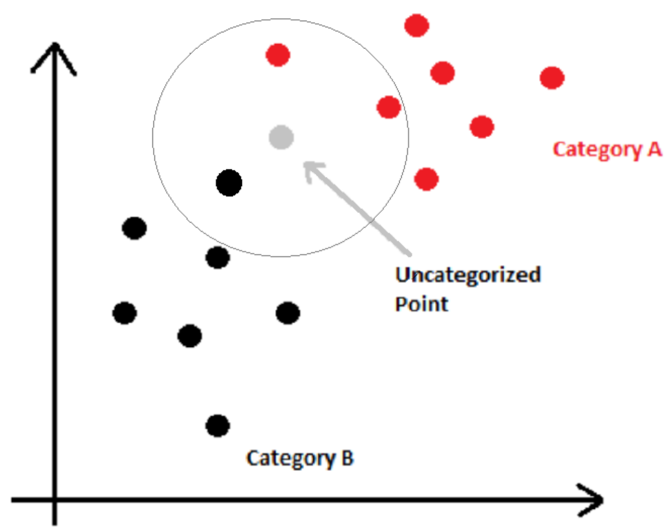


Fig. 4. K-Nearest Neighbours, k = 3

### 3.2 Random Forest Algorithm

The random forest (RF) algorithm used for classification and regression tasks. Even though it can solve both types of problems, it is not as accurate with regression tasks. Proposed in [29], it operates by constructing several decision trees while the training occurs. In classification tasks, the class is determined by the class of the most trees. On the other hand, for regression tasks, the returned prediction is the average prediction of the individual trees, see Figure 5.

*Applications:* The random forest algorithm has been integrated with a Bayesian Optimization for product quality prediction [13]. The authors conclude that fewer but critical features handled by RF-Bayesian optimization can realize satisfactory forecast accuracy as well as cost-effective computing time. This type of application is seen quite often in automated production line systems, it is a powerful application of ML potential to provide managerial insights and operational guidance for product quality prediction and control the real-life process industry. Another application of RF can be seen

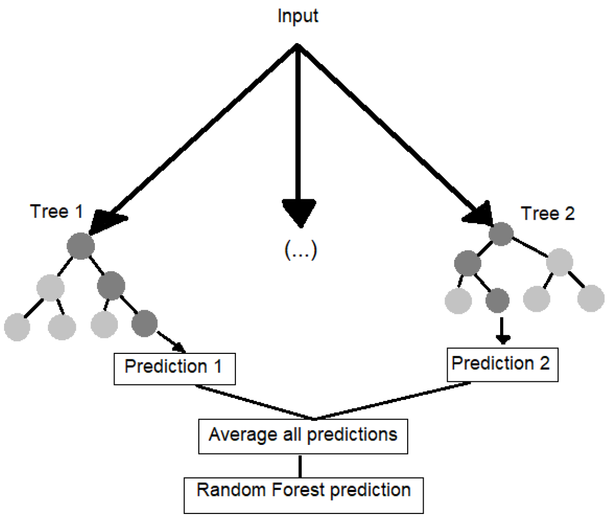


Fig. 5. An example of random forest algorithm

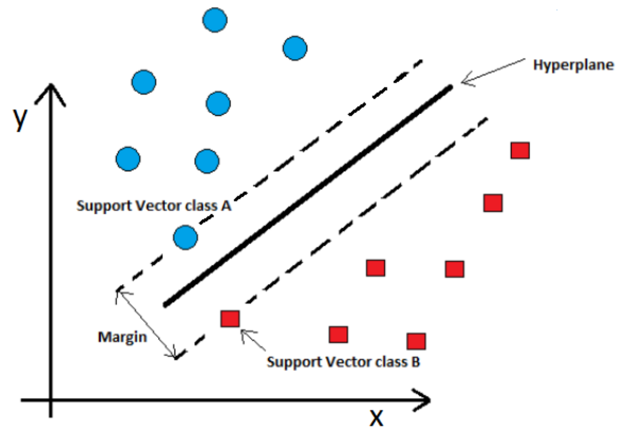


Fig. 6. Support Vector Machine hyperplane example

in [14], where an improved RF algorithm was combined with a digital twin (DT is also reviewed in the DP section, Section IV) for fault diagnosis of intelligent production lines. This architecture was verified through a case study of an automobile rear axle assembly line, with the aim to simulate fault data that is usually difficult to obtain in an actual production line to train a reliable fault diagnosis model.

### 3.3 Support Vector Machines

Support vector machine (SVM) is a supervised machine learning method that analyzes data and recognizes patterns. This method can help to solve classification problems or be used in regression analysis [15]. The built model assigns new samples or data points to the existing categories for a given set of training examples. It can be thought of as a constructed hyperplane in a high-dimensional space that separates different types of data samples [30]. The points closest to the hyperplane are called the support vector points, and the distance of the vectors from the hyper plane is called the margin, see Figure 6. SVM is used as a regression model analyzer, when the data is continuous, usually denoted as SVR.

*Applications:* Bodendorf and Franke [16] use SVR with other machine learning algorithms such as $k$-nn, linear regression, and decision tree to estimate product costs in the early product design phase in the automotive industry. The authors conclude that all the machine learning approaches used to predict costs showed good metric values, proving to be much more efficient than traditional spreadsheet-based cost analysis, supporting the decision-making of costs.

Oh et al. [17] used an adaptive SVM-based real-time quality assessment for the primer-sealer dispensing process of the sunroof assembly line. This work proposed a framework to preprocess the data for an SVM-based decision-making algorithm. This adaptive process is a feedback control that ensures the effectiveness of the SVM, meaning that there was integrated an adaptive loop with human-expert judgment such that when the system detects an anomaly, the human expert would intervene and adequately update the database for effective use of the SVM.

## 4. Deep Learning

Deep learning is a specific area of ML, similar to "a new take on learning representations from data that emphasizes on learning successive layers of increasingly meaningful representations" [31]. This area has been gaining popularity thanks to its promising performance and results. ANNs (Artificial neural networks) have brought a handful of benefits, and there are many types of these Neural Networks, ranging from Convolution Neural Networks, Long Short Term Memory Networks, Generative Adversarial Networks, Recurrent Neural Networks, and others.

Deep learning has been used effectively by enterprises and big companies to save resources and reduce the time needed to complete tasks. An example of such capabilities of saving energy or managing it for smart buildings in the industry is shown in [18]. It is interesting to see the variety of applications of such technologies that can be integrated into the automotive industry production lines, despite the number of robotic systems that help factories to create their products, workers still need to help assemble some tasks. Because operators are human, and physical fatigue is commonly provoked by repetitive tasks, which can interfere with the assembly production line and possible hazards, more specifically, the assembly of electrical harnesses of engines, even though it should be an easy task to perform, there is still a probability of encountering components that have been badly connected. A sound detection system based on deep learning approaches was proposed in [19] to identify click-sounds produced when electrical harnesses are connected, so when electrical harnesses are badly connected, the click-sound should be different from the standard one making it recognizable for the machine.

## 4.1 Artificial Neural Networks

Before diving into the deep learning section, its important to consider artificial neural networks (ANNs). A definition appreciated by many is the one from Howard Rheingold,

> "The neural networks is this kind of technology that is not an algorithm, it is a network that has weights on it, and you can adjust the weights so that it learns. You teach it through trials."

This quote provides a good description of ANN's functioning, in essence, it describes what it is all about. Neural networks are composed of an input layer, hidden layers, and an output layer, see Figure 7.

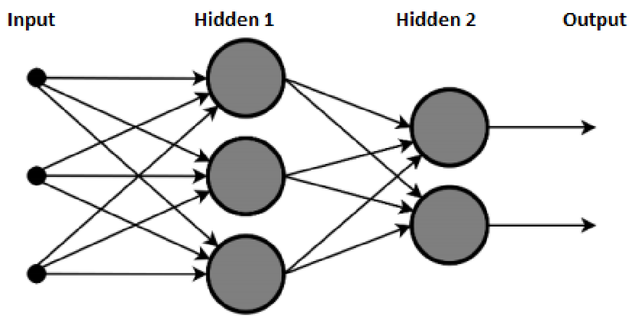


Fig. 7. Neural Network

These layers are composed of perceptrons (mathematical's model of neurons). Figure 8 illustrates a neuron, where the activation function $f(X)$ calculates the output neuron (one input of the next layer). This evaluation depends on the bias and the weighted sum input. Each input variable has a weight assigned to it that is forwarded to the function, adjusted, and passed to the activation function.

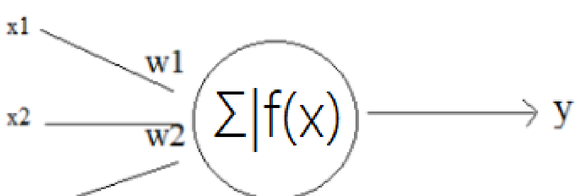


Fig. 8. Perceptron

The classic network learns using the *backpropagation* method of fine-tuning the weights where they are adjusted based on the last iteration error to match what the operator is seeking. This process allows ANN the ability to learn and generalize from data, that is, to mimic the human capability to learn. Now that it is more clear how ANNs work, this technology has been implemented to automate specific tasks and overall to optimize production's pace, resources, and so on. After ANNs, deep learning took place, and what was before a *Multi-Layer Perceptron* can now be put into successive layers reflecting new models.

*Activation Functions:* Activation functions are crucial for determining the progress of a model. Because deep learning benefits with the use of multiple layers of representations [31], it is clear that there must be a type of transformation within layers. Otherwise, the layer could only learn linear transformations [32]. So to benefit from deep representations, a non-linear or activation function is needed. As an example, the *ReLU* or *Rectified Linear Unit* decides whether a neuron should be activated or not. It works is by mapping the input directly into output in case it is positive. Otherwise, the output will be zero. There have been some changes to it, so *LeakyReLU* came up, but both versions are popularly used. LeakyReLU it is just a change to how ReLU previously reacted to negative inputs. Instead of transforming negative inputs into an output of zero, the LeakyReLU allows that input some slight transformation, usually multiplying it by a constant, represented by $\alpha$ in Figure 9.

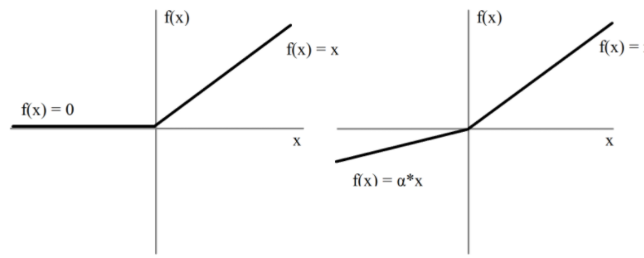


Fig. 9. ReLU and LeakyReLU

The following activation functions are Sigmoid and Tanh. The Sigmoid function is shaped like an S, it exists between values from 0 to 1, and it is mainly used for basic predictions since predictions range between 0 and 1. This function is used in ML for tasks like logistic regression and basic neural network implementation [33].

Tanh is similar to the sigmoid function, also having the S shape. It converts any input's real valueto an output in the range $[-1, 1]$, see Figure 10.

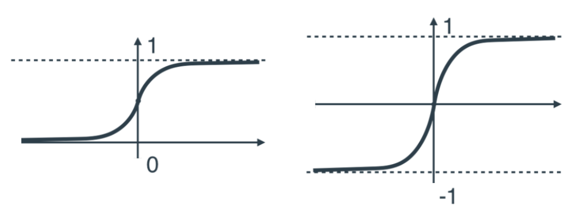


Fig. 10. Sigmoid and Tanh

*1) Convolutional Neural Networks:* The term convolutional neural networks often referred to as "CNN", is one of the most popular architectures of Deep Neural Networks. This network is usually composed of dense, fully connected, convolution, and max-pooling layers (see Figure 11[1]). The fundamental difference between a densely connected and convolution layer is that dense layers learn global patterns in their input feature space [34] (for example, patterns involving all pixels of a certain image). In contrast, convolution layers learn local patterns (in this case, the features represent patterns such as eyes, nose, ears), Figure 12. The max-pooling layers are a way of aggressively down-sampling the extracted feature maps for better overall performance.

*Applications:* CNNs are the core of computer vision as it learns features and recognizes patterns. Valter et al. [20] used an Edge-/Cloud-Architecture and a CNN to detect defects in real-time production. Throughout the production line, the tests consisted in finding anomalies. This architecture can detect anomalies in production line systems and identify the cause of the defection itself. Defects are classified by a convolutional neural network (CNN) and corrected (depending on the defect type) by an automated rework. The implementation of this methodology brought significant advantages, one of them being the reduction time on processing which leads to reduced expenses.

*2) Recurrent Neural Networks:* Recurrent neural network (RNN) architecture allows sequence processing. Basically, it processes sequences by iterating through sequence elements and maintaining a state that contains information relative to what it has seen so far, working as a loop [35] (Figure 13).

*Applications:* This type of Neural Network, RNN, has been applied to production line systems for anomaly detection [21]. The assembly of a surface-mounted device machine manufactures various products on a flexible manufacturing line. The paper proposed an anomaly detection model that can adapt to the various manufacturing environments in a really fast manner.The model is based on a Recurrent Neural Network (RNN) Encoder–Decoder with operating machine sounds. Another fault detection example can be seen in [22], where the manufactured products are motors. The proposed two-stage machine learning analysis architecture can accurately predict the motor fault modes only by using motor vibration time-domain signals without any complicated preprocessing. This architecture is composed of a RNN-based Variational Autoencoder (VAE).

*3) Long Short Term Memory Networks:* Long Short Term Memory (LSTM) network is a change to the standard feedforward neural network. The LSTM unit comprises cells with an input gate, an output gate, and a forget gate [36]. A chain of these units forms the LSTM architecture composes different memory blocks called a cell. The called gates of LSTM perform individual tasks. Figure 14 illustrates an LSTM unit structure [37].

The forget gate gets rid of no longer useful information in the cell.

[1]Convolutional neural network, learn convolutional neural network from basic and its implementation in keras. https://towardsdatascience.com/covolutional-neural-networkcb0883dd6529

*Applications:* Regarding LSTM applications in industry, Peng et al. [23] use the multiple-input multiple-output technique from fifth-gen communication systems (5G) the employ an LSTM to compensate the negative effects of imperfect channel state information (CSI) in the practical radio frequency systems. This CSI imperfection is usually caused by the channel estimation error and the transmission and processing delay, knowing that imperfect CSI severely reduces the system secrecy capacity. An LSTM-based predictor and compensation scheme were designed to alleviate negative effects effectively.

LSTM is also used for anomaly detection seen in [24]. The cited paper presents a novel detection and prediction procedure based on a LSTM architecture to cooperatively predict process outputs and anomalies by using two separate but interacting models. This interesting application of LSTM allows a solution for short-term as well as long-term anomalies.

*4) Generative Adversarial Networks:* The generative adversarial networks (GANs) are composed of a generator and a discriminator networks (often referred to as adversaries). Its essence is to train the generator network to be able to fool the discriminator network, and thus the architecture allows the generator to evolve towards generating increasingly better forms of input [38]. It is worth mentioning that the discriminator network is constantly adapting to the gradually improving capabilities of the mentioned generator. The main issue of GANs is its dynamic system where the optimization process does not seek a minimization but instead an equilibrium between two forces, making them notoriously difficult to train.

*Applications:* Wong et al. [25] introduce a generative adversarial framework based on a game theory for data augmentation. Data augmentation is a strategy used to extend datasets by applying data augmentation techniques such as cropping, padding, and flipping, enabling practitioners to significantly increase the diversity of data available for training models without the need to collect new data samples. This particular technique, increases the diversity of available data, which can be used as a self-learning technique in production lines where the data is scarce. Applying diversity in datasets could find potential unidentified anomaly samples, reducing their impact if such anomalies are detected in a production line.

## 4.2 Digital Twin

Digital Twin is an emerging technology that has brought many benefits since its usage [39]. The digital twin is a virtual representation that serves as the real-time digital counterpart of a physical object or process, working like a computer program that uses real-world data to create simulations [40]. These simulations can simulate how a product or process will perform. Figure 15 from [41] illustrates the manufacturing process of digital twin models.

Figure 15 shows one digital twin configuration that focuses on the manufacturing portion of a specific product's life cycle. When developing a digital twin, the key is to seek this integration and iterative quality of the physical and digital representations, creating a loop from the physical world to the

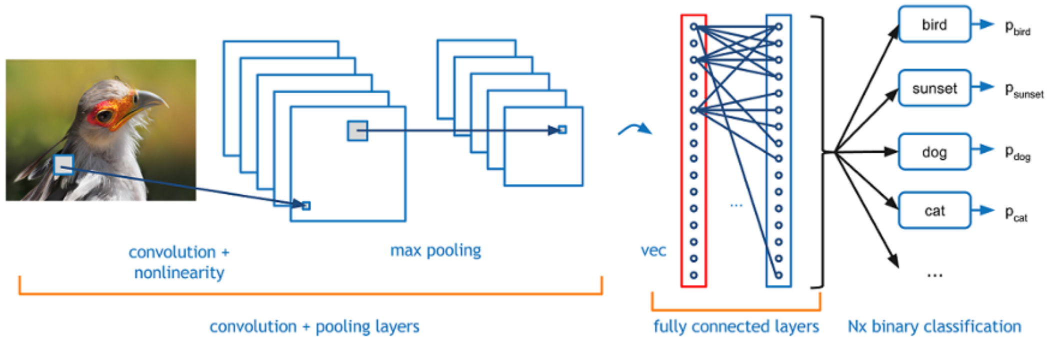


Fig. 11. An example of a CNN

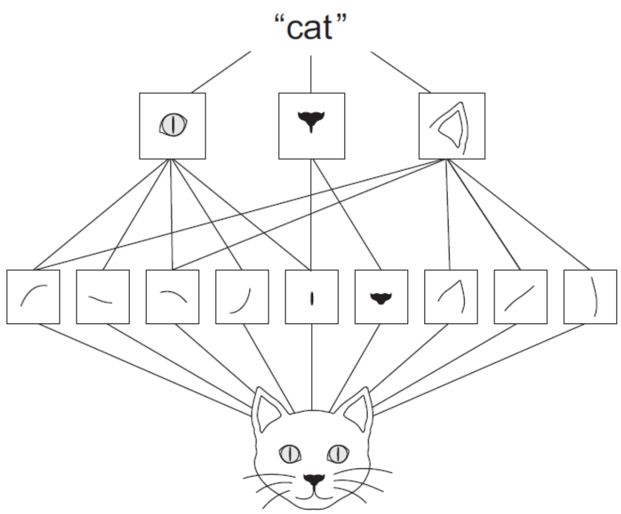


Fig. 12. Hierarchy of local patterns representative of a high-level concept the "cat" from [31]

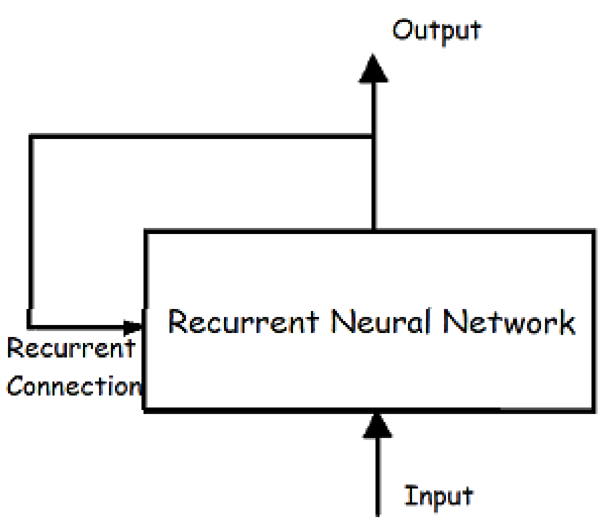


Fig. 13. A Recurrent Network

digital world and back. The product under the manufacturing environment combines advanced techniques with the IoT that helps with tasks such as allowing communication analysis. This feedback and information are used to implement further intelligent actions in the physical production line.

*Applications:* In 2017, Vachálek et al. [43] presented a digital twin of an industrial production line. The concept focuses on optimizing production processes of the physical environment using the virtual one. They perform new tests and norms before submitting the physical environment changes.

Balderas el at. [26] use of a Digital-Twin that integrates a metaheuristic optimization and a direct model for printed circuit boards (PCB) design and processing focused on the drilling process of the manufacturing of such products. The authors conclude that using this type of optimization on DT technology allows for more production at a faster pace.

In [27], Aheleroff et al. developed a DT reference model that can be accessed as an online service. This type of model allowed for considerable advantages, including smart scheduled maintenance, real-time monitoring, remote controlling, and predicting functionalities.

# 5. Discussion of Papers Reviewed

After revisioning the mentioned papers and based on the information retrieved from the research, it is now possible to answer the questions presented in section II, which are:

Q1. How does machine learning impact the future of end-of-line testing systems?

Q2. What are the most common applications of these technologies in the manufacturing environment?

Some applications have shown that by exploiting data from quality control procedures, selecting relevant features, and finding patterns related to faults and anomalies, it is possible to isolate and correct most of them, also simulate fault data which is difficult to obtain in an actual production line. The revisioned papers show that these technologies also estimate product costs and reduce processing time (which reduces expenses). Based on this information, to answer question Q1, ML impacts the future of end-of-line testing systems by assessing predictive maintenance, product quality assurance, demand forecasting, predicting safety issues, and escalating

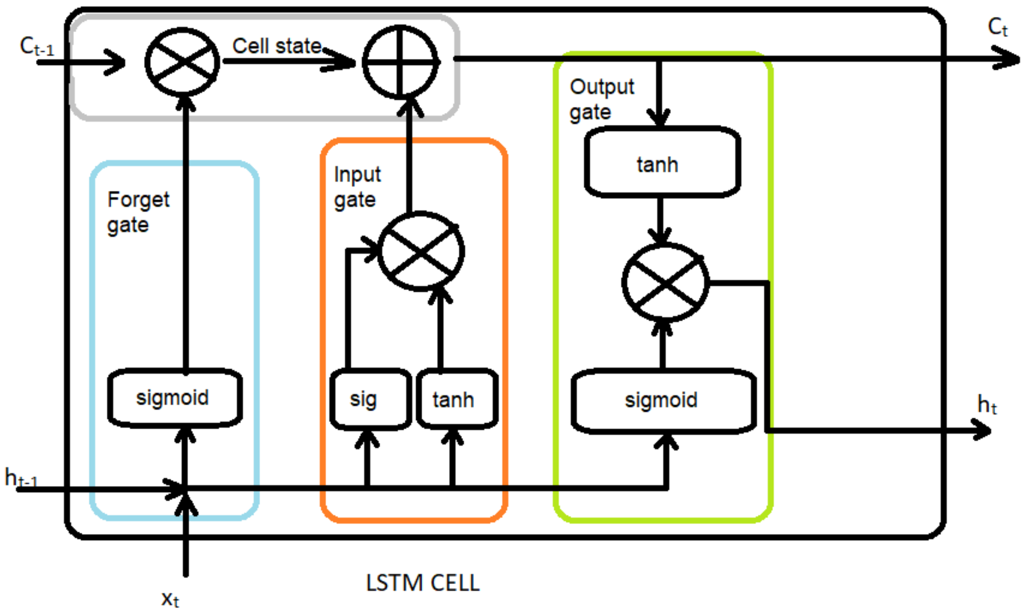


Fig. 14. LSTM gates

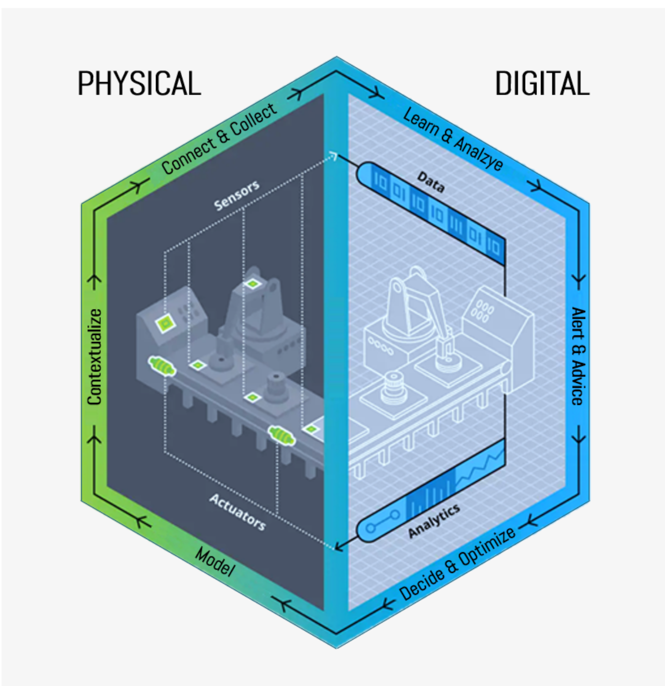


Fig. 15. Manufacturing process digital twin model [42]

resources. Specifically, PdM allows EoL systems to avoid critical equipment downtime with proactive monitoring of a production line's critical devices. Product quality assurance tools grant that potential product quality issues are identified in early stages of the manufacturing process. To forecast demand is also one of the ways that ML impacts the future of EoL systems by accurately predicting product demand to reduce costs and increase profits. Finally, to answer the second question, Q2, the most common applications of ML technologies in the manufacturing environment are related to anomaly detection, anomaly isolation, anomaly correction, proactive maintenance, escalating resources, and reducing production time. These advantages, coupled with the manufacturing environment, allow EoL systems to be more productively accurate, reliable, and less expensive. Another pertinent observation drawn from the revision is the exponential paper's growth using ML in end-of-line systems in recent years.

## 6. Concluding Remarks and Future Trends

This paper reviewed and contextualized the state-of-the-art of ML algorithms, particularly deep learning techniques, throughout the industry, aiming to optimize EOL testing systems mainly in the automotive industry. It is common to see various ML algorithms and deep learning architectures used in the industry. The most common applications reviewed aim to help operators make decisions, predict disruptions, predict maintenance, identify defects and anomalies. Having this type of advantage set to testing environments, it is possible to leverage the product's quality and the functionality of the tests.

Industry has shown that the way forward is to follow PdM approaches based on information collected throughout production. The use o ML and DL allows to innovate every aspect of business such as: design smart products, run smart factories, forecast demand, ensure quality, manage the supply chain, reduce production downtime, reduce the impact of anomalies and understand complex multi-stage processes.

ML's capability will boost quality on factory production lines. In particular, decision-making based on data interpretation is becoming a reality in informed production lines, e.g., decisions based on data obtained from end-of-line testing. These actions will lead to better results, cost savings, and competitive advantages. To conclude, ML is becoming indispensable for the industry, and its future competitiveness depends on it.

# Acknowledgements

This work was supported by the I&D Project “DEoL-TA: Digitalisation of end-of-line distributed testers for antennas operação POCI-01-0247-FEDER-049698”, financed by the Fundos Europeus Estruturais e de Investimento (FEEI), through the Program “Programa Operacional Competitividade e Internacionalização(POCI) / PORTUGAL 2020”.

## *References*

[1] P. Noren, L. J. Foged, and P. Garreau, “State of the art spherical near-field antenna test systems for full vehicle testing,” pp. 2244–2248, 2012.

[2] S. Barreto, J. Ferraz, and E. Martins, “Automated System for Verification and Validation of End-of-Line Tests,” Nov. 2017, pp. 2017–36–0228. [Online]. Available: https://www.sae.org/content/2017-36-0228/

[3] A. Theissler, J. Pérez-Velázquez, M. Kettelgerdes, and G. Elger, “Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry,” *Reliability Engineering System Safety*, vol. 215, p. 107864, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0951832021003835

[4] D. Vicêncio, H. Silva, S. Soares, V. Filipe, and A. Valente, “An intelligent predictive maintenance approach based on end-of-line test logfiles in the automotive industry,” in *Industrial IoT Technologies and Applications*. Cham: Springer International Publishing, 2021, pp. 121–140.

[5] S. Ayvaz and K. Alpay, “Predictive maintenance system for production lines in manufacturing: A machine learning approach using iot data in real-time,” *Expert Systems with Applications*, vol. 173, p. 114598, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417421000397

[6] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, and D. Moher, “The prisma 2020 statement: An updated guideline for reporting systematic reviews,” *International Journal of Surgery*, vol. 88, p. 105906, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1743919121000406

[7] V. Hirsch, P. Reimann, and B. Mitschang, “Data-driven fault diagnosis in end-of-line testing of complex products,” in *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2019, pp. 492–503.

[8] W. Zhang, D. Yang, and H. Wang, “Data-driven methods for predictive maintenance of industrial equipment: A survey,” *IEEE Systems Journal*, vol. 13, no. 3, pp. 2213–2227, 2019.

[9] J. Yan, Y. Meng, L. Lu, and L. Li, “Industrial big data in an industry 4.0 environment: Challenges, schemes, and applications for predictive maintenance,” *IEEE Access*, vol. 5, pp. 23 484–23 491, 2017.

[10] V. Del Rosso, A. Andreucci, S. Boria, M. L. Corradini, and A. Ranalli, “Mechanical fault detection for induction motors based on vibration analysis: a case study,” in *IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society*, 2021, pp. 1–6.

[11] G. Verdier and A. Ferreira, “Adaptive mahalanobis distance and $k$-nearest neighbor rule for fault detection in semiconductor manufacturing,” *IEEE Transactions on Semiconductor Manufacturing*, vol. 24, no. 1, pp. 59–68, 2011.

[12] Z. Zhou, C. Wen, and C. Yang, “Fault isolation based on k-nearest neighbor rule for industrial processes,” *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2578–2586, 2016.

[13] T. Wang, X. Wang, R. Ma, X. Li, X. Hu, F. T. S. Chan, and J. Ruan, “Random forest-bayesian optimization for product quality prediction with large-scale dimensions in process industrial cyber–physical systems,” *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8641–8653, 2020.

[14] K. Guo, X. Wan, L. Liu, Z. Gao, and M. Yang, “Fault diagnosis of intelligent production line based on digital twin and improved random forest,” *Applied Sciences*, vol. 11, no. 16, 2021. [Online]. Available: https://www.mdpi.com/2076-3417/11/16/7733

[15] M. Jalal and H. Jalal, “Behavior assessment, regression analysis and support vector machine (svm) modeling of waste tire rubberized concrete,” *Journal of Cleaner Production*, vol. 273, p. 122960, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652620330055

[16] F. Bodendorf and J. Franke, “A machine learning approach to estimate product costs in the early product design phase: a use case from the automotive industry,” *Procedia CIRP*, vol. 100, pp. 643–648, 2021, 31st CIRP Design Conference 2021 (CIRP Design 2021). [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212827121006120

[17] Y. Oh, K. Ransikarbum, M. Busogi, D. Kwon, and N. Kim, “Adaptive svm-based real-time quality assessment for primer-sealer dispensing process of sunroof assembly line,” *Reliability Engineering & System Safety*, vol. 184, pp. 202–212, 2019, impact of Prognostics and Health Management in Systems Reliability and Maintenance Planning. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0951832017303861

[18] M. Elsisi, M.-Q. Tran, K. Mahmoud, M. Lehtonen, and M. M. F. Darwish, “Deep learning-based industry 4.0 and internet of things towards effective energy management for smart buildings,” *Sensors*, vol. 21, no. 4, p. 1038, Feb 2021. [Online]. Available: http://dx.doi.org/10.3390/s21041038

[19] R. Espinosa, H. Ponce, and S. Gutiérrez, “Click-event sound detection in automotive industry using machine/deep learning,” *Applied Soft Computing*, vol. 108, p. 107465, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494621003884

[20] J. Vater, M. Kirschning, and A. Knoll, “Closing the loop: Real-time error detection and correction in automotive production using edge-/cloud-architecture and a cnn,” in *2020 International Conference on Omni-layer Intelligent Systems (COINS)*, 2020, pp. 1–7.

[21] Y. Park and I. D. Yun, “Fast adaptive rnn encoder–decoder for anomaly detection in smd assembly machine,” *Sensors*, vol. 18, no. 10, 2018. [Online]. Available: https://www.mdpi.com/1424-8220/18/10/3573

[22] Y. Huang, C.-H. Chen, and C.-J. Huang, “Motor fault detection and feature extraction using rnn-based variational autoencoder,” *IEEE Access*, vol. 7, pp. 139 086–139 096, 2019.

[23] T. Peng, R. Zhang, X. Cheng, and L. Yang, “Lstm-based channel prediction for secure massive mimo communications under imperfect csi,” in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[24] B. Lindemann, N. Jazdi, and M. Weyrich, “Anomaly detection and prediction in discrete manufacturing based on cooperative lstm networks,” in *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*, 2020, pp. 1003–1010.

[25] Y. Wang, K. Li, S. Gan, C. Cameron, and M. Zheng, “Data augmentation for intelligent manufacturing with generative adversarial framework,” in *2019 1st International Conference on Industrial Artificial Intelligence (IAI)*, 2019, pp. 1–6.

[26] D. Balderas, A. Ortiz, E. Méndez, P. Ponce, and A. Molina, “Empowering digital twin for industry 4.0 using metaheuristic optimization algorithms: case study pcb drilling optimization,” *The International Journal of Advanced Manufacturing Technology*, vol. 113, 03 2021.

[27] S. Aheleroff, X. Xu, R. Y. Zhong, and Y. Lu, “Digital twin as a service (dtaas) in industry 4.0: An architecture reference model,” *Advanced Engineering Informatics*, vol. 47, p. 101225, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1474034620301944

[28] S. Taneja, C. Gupta, K. Goyal, and D. Gureja, “An enhanced k-nearest neighbor algorithm using information gain and clustering,” in *2014 Fourth International Conference on Advanced Computing Communication Technologies*, 2014, pp. 325–329.

[29] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct 2001. [Online]. Available: https://doi.org/10.1023/A:1010933404324

[30] M. Hearst, S. Dumais, E. Osuna, J. Platt, and B. Scholkopf, “Support vector machines,” *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18–28, 1998.

[31] F. Chollet, *Deep learning with python*. Manning Publications, 2017.

[32] A. Olgac and B. Karlik, “Performance analysis of various activation functions in generalized mlp architectures of neural networks,” *International Journal of Artificial Intelligence And Expert Systems*, vol. 1, pp. 111–122, 02 2011.

[33] S. Sharma, S. Sharma, and A. Athaiya, "Activation functions in neural networks," *towards data science*, vol. 6, no. 12, pp. 310–316, 2017.
[34] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.
[35] W. Yin, K. Kann, M. Yu, and H. Schütze, "Comparative study of cnn and rnn for natural language processing," *arXiv preprint arXiv:1702.01923*, 2017.
[36] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: Lstm cells and network architectures," *Neural computation*, vol. 31, no. 7, pp. 1235–1270, 2019.
[37] G. Singhal. (2020) Introduction to lstm units in rnn. [Online]. Available: https://www.pluralsight.com/guides/introduction-to-lstm-units-in-rnn
[38] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 53–65, 2018.
[39] E. Negri, L. Fumagalli, and M. Macchi, "A review of the roles of digital twin in cps-based production systems," *Procedia Manufacturing*, vol. 11, pp. 939–948, 2017, 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2351978917304067
[40] S. Boschert and R. Rosen, "Digital twin—the simulation aspect," in *Mechatronic futures*. Springer, 2016, pp. 59–74.
[41] D. U. Press. (2017) Industry 4.0 and the digital twin. [Online]. Available: https://www2.deloitte.com/content/dam/Deloitte/kr/Documents/insights/deloitte-newsletter/2017/26_201706/kr_insights_deloitte-newsletter-26_report_02_en.pdf
[42] "Industry 4.0 and the digital twin- manufacturing meets its match," https://www2.deloitte.com/us/en/insights/focus/industry-4-0/digital-twin-technology-smart-factory.html.
[43] J. Vachálek, L. Bartalský, O. Rovný, D. Šišmišová, M. Morháč, and M. Lokšík, "The digital twin of an industrial production line within the industry 4.0 concept," in *2017 21st International Conference on Process Control (PC)*, 2017, pp. 258–262.