Digital Twins in Product Lifecycle for Sustainability in Manufacturing and Maintenance

Izabela Rojek 1,*, Dariusz Mikolajewski 1 and Ewa Dostatni 2

1 Institute of Computer Science, Kazimierz Wielki University, 85-064 Bydgoszcz, Poland; dmikolaj@ukw.edu.pl
2 Faculty of Mechanical Engineering, Poznań University of Technology, 60-965 Poznań, Poland; ewa.dostatni@put.poznan.pl
* Correspondence: izarojek@ukw.edu.pl

Abstract: A “digital twin” is a dynamic, digital replica of a technical object (e.g., a physical system, device, machine or production process) or a living organism. Using this type of solution has become an integral part of Industry 4.0, offering businesses tangible benefits, in addition to being particularly effective within the context of sustainable production and maintenance. The purpose of this paper is to present the results of research on the development of digital twins of technical objects, which involved data acquisition and their conversion into knowledge, the use of physical models to simulate tasks and processes, and the use of simulation models to improve the physical tasks and processes. In addition, monitoring processes and process parameters allow for the continued improvement of existing processes as regards intelligent eco-designing and planning and monitoring production processes while taking into account sustainable production and maintenance.

Keywords: digital twins; sustainability; manufacturing; maintenance; computational intelligence

1. Introduction

The term digital twin (DT) stems from the concept of a virtual equivalent to a physical phenomenon, developed in 2002 at the University of Michigan. A dynamic, digital replica of a physical system, device, machine, production process or a living organism is more than just a model [1]. Any changes to which the physical object is subjected are detected by sensors and reflected in its digital replica. This offers a more in-depth insight into the processes which occur when using the physical object, rendering it possible to predict events, enabling effective remote management, early malfunction detection, element wear monitoring and predictive responding to such situations. Conclusions drawn from an analysis of a digital twin can subsequently be applied to the original object with reduced risk and increased return on investment. In addition to being implemented in new and already established solutions (e.g., production lines that are planned to operate for many more years), the digital twin technology is also used by businesses to test new products before committing to serial production and use, enabling the introduction of improvements and further development. A key requirement for developing a virtual recreation of the physical world is real-time access to a complete spectrum of appropriate-quality data and that the virtual world can “learn” the behaviors of the physical world. Currently, available dedicated solutions meet this requirement. This is rendered possible as a result of the development of sensors and the Internet of Things, which offers the ability to continuously acquire data and transfer it in large volumes (currently in exabytes). The other end of the process of developing a digital twin involves analytical tools, machine learning and artificial intelligence instruments which make use of this data. Their dynamic development enables the effective use of data to build knowledge about a physical object, its behaviors and reactions to changes in its environment, and continuous verification of the recreation, and in some cases also creating variations for the purpose of analyzing different scenarios. The ability to visualize a technical object (e.g., a machine) and to simulate its operation...
(both functionally and structurally) is expanded to include communication methods that enable teamwork. Digital twins, due to their features, revolutionize industrial and business activities thanks to the results which are possible to achieve with their help, in particular their predictive potential, product monitoring throughout the entire product life cycle and the ability to utilize this knowledge to optimize and develop products, including developing new and improved generations. The availability of technological solutions enabling the implementation of the digital twin concept, combined with the results it generates, makes its implementation a key part of modern digital transformations. In annual reports published by Gartner on major strategic technological trends entitled “Top 10 Strategic Technology Trends for 2019”, “Top 10 Strategic Technology Trends for 2020”, “Top 10 Strategic Technology Trends for 2021”, digital twins and related technologies are frequently in the top ten. According to the report, half of all large manufacturing companies will be using digital twins by 2021, resulting in a 10% increase in effectiveness. In the new generation of the industry—Industry 4.0, digital twins are of key importance. Depending on the industry, DT offers improvements at every stage of the production process, in particular: shorter solution modeling time (up to 20%), shorter construction time (up to 25%), a marked increase in the quality of documentation and reduced amount of design errors (up to 20%) [2–4]. In the middle and long term, it also increases sales and market share, as the enhancement of the above-mentioned processes results in businesses being able to participate in a higher number of projects than in the past, and for specific industries, it also means larger, more complex projects, possibly featuring more advanced technologies and offering a higher return on investment. Monitoring product life cycles renders exchanging information between various departments more effective, which may give rise to completely new possibilities regarding analysis, prediction and creative applications, including faster product line replacement and simultaneous demand creation across multiple markets previously unavailable due to scale of production or logistical limitations, for example. As a result, the reduction in production, service/helpdesk time and/or cost is a natural consequence. The digital twin concept becomes even more attractive in sustainable manufacturing, in both subtractive and additive machining.

Sustainable development is one of the most important issues for current and future generations. The assumption that natural resources are infinite, i.e., that the ability of the environment to regenerate is sufficient to repair all damage caused by the entirety of human activity, is changing. Thus, the idea of sustainable development now influences all organizational aspects of human life, including economic, political, social and environmental. Manufacturing, being one of the foundations of the civilized lifestyle, will be profoundly affected by it and will have an important role to play in creating a sustainable path towards development. Despite this rising awareness, nearly all current production models are based on the old, outdated paradigm. The technology on which manufacturing is largely based must, along with culture and economics, provide tools and opportunities to create new solutions to bring us closer to sustainable production [5,6]. New technologies, business models and lifestyles will become a milestone marking the advent of a new, sustainable world, and nowhere will this be truer than in the manufacturing sector. The inevitable restrictions and increased requirements will impact the entire industrial sector, as well as education and science, as we move closer to sustainable development [7,8]. Research and development will be of key significance, their purpose of enabling society to adequately satisfy the above-mentioned needs in the form of properly trained staff and innovative technologies [9,10]. The main research challenges related to sustainable production have been presented by the authors involved in the international project “IMS2020: Supporting Global Research for IMS2020 Vision”, promoted by the European Commission for the purpose of developing a roadmap for future (2020) production [5,6].

The purpose of this paper is to present the results of the authors’ research on the development of digital twins, which involved data acquisition and their conversion into knowledge, the use of physical models to simulate tasks and processes, and the use of simulation models to improve physical tasks and processes. In addition, monitoring pro-
cesses and process parameters allow for the continued improvement of existing processes as regards intelligent eco-designing and planning and monitoring production processes, including sustainable production and maintenance.

The novelty proposed by the authors, which is missing from the literature, is presented in the following chapters: Materials and methods, Procedures and Results, are the result of the authors’ own research, simulations and calculations. The industry uses DT technology to create benchmarks for predictive analysis of asset performance. As an advanced type of process model simulation, digital twins provide real-time data, and operators can apply them in a variety of ways throughout planning, production and supply. In our study, various support systems are being created, which we have included in the article. As for the digital twins that are being created, they are the research of different researchers, not patterns.

The Introduction section defines the concept of the digital twin in the context of Industry 4.0, sustainable production and maintenance. This part also defines research objectives and methods to achieve them and presents the structure of the article. The rest of this paper is structured as follows: Section two provides a review of the literature related to the use of DTs in enterprises and highlights the importance of machine-learning methods in the context of their use in digitizing data in DTs. The Materials and Methods section presents the own method used to create a digital twin. The Procedures section demonstrates the application of the DT in eco-designing, planning and monitoring of manufacturing processes for sustainability in manufacturing and maintenance. The Results section contains an assessment of the solution, an example application of digital twins and a system architecture incorporating DTs. The final two sections contain Conclusions and References.

2. Literature Review

A comprehensive literature review was carried out using Whittemore and Knafel’s approach. This integrative method of study is the only approach that combines different methodologies (including experimental and non-experimental research) and has the potential to play a greater role in practice. Literature research was conducted in scientific journals and monographs. The most basic criterion for inclusion was the relevance for this group of technologies, both for their theory, simulation and experimental research, and practical implementation in the economy.

The authors of article [7] identify eight DT development prospects, including modular DTs, ensuring modeling coherency and accuracy, incorporating big data analytics into DT models, improving DT simulations, integrating virtual and augmented reality systems with DT, expanding the applications of DT, effective mapping of cyber-physical data and integration with cloud/edge computing. The DT technology is an effective tool that meets the requirements of smart manufacturing by way of recreating physical processes in virtual space [8,9]. This is performed within a broad-spectrum of CPS (cyber-physical systems). The DT paradigm is well suited to a lifecycle-based paradigm [10]. Technology is increasingly considered to be of paramount importance to improving and evolving global production, including its globalization. Because DTs are now being used in new sectors to increase productivity, efficiency and competitiveness, a variety of tools and methods must be used, which include tools for managing data and connectivity, tools for representing and storing data, machine learning tools and analytical methods.

Data acquisition and transmission are of key significance to DT as the technology requires real-time connectivity and data transfer. As an example, Freeman [11] proposes a data stream processing system in which data are analyzed and queried continuously. The above-mentioned data acquisition systems are of key significance to the implementation of DT in production environments containing data collected using temporary (due to the hardware used, type of materials, etc.) and permanent data storage processes (processing systems utilizing real-time sensors) [12]. Heterogeneous data and industry-specific knowledge acquired from production processes require modeling and integration with pro-
duction systems. The most important knowledge representation tools for developing DTs include NoSQL ontologies and databases [13–17], relational databases [18,19], MySQL [20], SQLite 3 [21], real-time databases for storing various UML data structures [22], transactional graph databases [23] and databases handling data transformation and predictive queries [24].

In order to process data into knowledge, it is necessary to utilize machine learning and data exploration methods. The following methods are used in the case of digital twins:

- Neural networks, used in quality prediction and operation control in metallurgy [25]; smart control in the paint and finishing works industry [26,27], error prediction and diagnostics in CNC machining [28];
- Deep neural networks, used in autonomous production in smart production plants [29] and error diagnostics [30];
- Dynamic Bayesian networks, used by DTs to monitor the condition of aircraft wings [31];
- NSGA-II genetic algorithm used to optimize the efficiency of a machine [19,26,28,32].

Businesses utilizing the DT technology also make use of what is referred to as microservices. These are programming tools developed as a set of loosely connected services. Thönes [12] defines them as functions that enable the development of an application as a set of relatively small, coherent, isolated and autonomous services that perform specific tasks. Rojko [33] analyzed virtualization tools in modern production systems, which enable the monitoring and tracking of resource services in a production plant for the purpose of automatic conflict resolution and increasing efficiency by way of facilitating decision-making and control.

Potential applications of DTs in the product lifecycle management process from creation to disposal (PLM) are analyzed within the general competitive process framework proposed in Casadesus-Masanell and Ricart [34], which outlines the structure and process of implementing DTs. It should be noted that introducing a DT renders it possible to limit the number of samples in product assessment drastically and enables businesses to mitigate production risks related to production line imperfections.

From an innovation perspective, process innovation involving a marked increase in the effective utilization of the creative potential of R&D and engineering staff compared to traditional methods translates into product innovations in the long term (increased quality of new product and/or service families). Moreover, the process of training the staff is expedited thanks to increased productivity and the number of projects completed, in addition to better error identification, including among younger, less experienced employees. This is due to the increased effectiveness of DT, including in modeling and analyzing designed objects while accounting for their statics, kinematics and material properties, space optimization thanks to analyzing large-scale designs and multilayered blueprints (electricity, piping, etc.), or creating technical documents or visualizations which take into consideration the technologies used to manufacture the objects (subtractive, additive or hybrid).

Gharaei et al. proposed a systems engineering approach to define DT requirements to formalize the DT concept from a systemic perspective, including the conceptual architecture of DT, which is defined based on ISO 42,010 standard. DT architectures have been identified by capturing formalized requirements using the EARS approach assessed based on a case from the IMPULSE project [35].

Some UE projects are also proposed for this topic, such as FACTLOG—energy-aware factory analytics for process industries, QU4LITY—digital reality in zero-defect manufacturing and the Innosuisse, IMPULSE project on digital twins, etc. We strongly believe that their outcomes will enrich and fuel the future of DT technology and its environment.

3. Materials and Methods

In line with the DT paradigms, research was conducted, which involved data acquisition and the development of the principles according to which these data would be transferred. The data were then processed into knowledge. In our approach, we try to
extract the knowledge and experience of the company’s employees, which are hidden in the developed examples included in databases. Data-mining methods extract knowledge from data. Based on decision trees, decision rules are created and then included in the applications used by employees. You could say that knowledge has been extracted and incorporated into computer applications.

The actual data come from several manufacturing companies in the area of eco-design, process planning and process supervision. These companies were characterized mainly by unit and small batch production, for which the knowledge gained for them is very important. In order to develop models in the form of decision trees with designers, technologists and machine operators, important criteria for solving individual problems were established, and predications (conditional attributes) and decision classes were defined, which were used to file learning examples periodically. To make the data useful for exploration, they were cleaned and transformed. Cleaning the data consisted of standardizing the record, completing the missing data and identifying remote points, then converting the data into normalization and coding.

Physical models were used to simulate tasks and processes. Simulation models were subsequently developed for the purpose of improving physical tasks and processes. Monitoring processes and process parameters render it possible to continually improve physical processes from the perspective of intelligent eco-designing, planning and monitoring manufacturing processes in line with sustainable manufacturing and maintenance principles. Digital twins were created for the above-mentioned areas. The digital twins were developed in stages:

- Analysis of physical processes and input data collected from real processes using sensors and measuring devices;
- Development of artificial intelligence models in the form of neural networks, decision trees and rules, and fuzzy logic;
- AI model assessment;
- Creation of a digital replica of the physical process or machine;
- Simulation of a physical process or machine operation;
- AI model improvement;
- Application of the AI models in existing systems;
- Upgrading and teaching AI models (Figure 1).

We considered individual objects: in eco-design, this is a choice of 3 ways of material. We broke down the planning of the technological process into simpler elements. We divided process supervision into the supervision of selected processes. As far as the size of the teaching files was concerned:

- for eco-design, the files had about 200 examples of learners;
- for design and supervision, there were about 300 examples of learners.

Concerning eco-designing, the focus was primarily on material selection. The goal of the study was to develop material selection methods and models utilizing AI and simulate the implementation of these methods to achieve optimal selection from the perspective of eco-construction. Material selection focused on a multicriteria analysis taking into consideration realistic factors from industry, i.e., determining the compatibility of materials, selecting additional materials while accounting for a set compatibility level and selecting methods of connecting materials. An eco-constructor selects materials and their fastenings in a way that minimizes their negative impact on the environment and renders recycling easier. For example, the compatibility of the materials was determined by the compatibility matrix (Figure 2).
methods of connecting materials. An eco-constructor selects materials and their fastenings in a way that minimizes their negative impact on the environment and renders recycling easier. For example, the compatibility of the materials was determined by the compatibil-

Figure 1. Main phases of the proposed method.

As regards the planning of manufacturing processes, special emphasis was put on tech-
nological process design. The purpose of the study was to develop technological process design methods and models utilizing AI and to simulate the operation of these processes to achieve optimal results from the perspective of a particular business. Decision rules were used to determine the structure of the technological process, i.e., the order of technological operations and procedures. Neural networks rendered it possible to construct models for selecting semi-finished products, instruments, machine tools, tools, tool holders and machining parameters, and decision trees—models for selecting machine tools, tools and tooling. Example selections used to create models were processed, i.e., normalized using fuzzy logic. Subsequently, the models of selecting semi-finished products, tooling, machine tools, tools, tool holders and machining parameters were implemented in the form of a prototype expert system for designing technological processes. The system is dedicated to technologists who do not possess sufficient experience in designing technological processes or have only just begun working at a given production company and thus are not yet familiar with the machines or other means of production.
As regards the monitoring of manufacturing processes, special emphasis was put on machining process monitoring. Simply designing a technological process was not sufficient. The product is manufactured under the planned process; however, various disruptions may occur during this process, which impacts the quality of the final product. Another research goal was to eliminate this type of hazard using our methods and models of monitoring the technological process of machining. The monitoring encompassed:

- Machining process disruptions; a simulation of such disruptions was conducted—when a particular disruption occurs, the operator selects its cause and is provided with a solution to the issue. The knowledge assumes the form of decision rules;
- Machining process stability—control charts were used to develop process stability disruption patterns, and a neural network was used to simulate the monitoring of such a process. After improving the example scenarios, the neural network was learned to react to instabilities properly. This model was then implemented in the system. The network is learned and improved based on actual actions;
- Ra and Rz surface roughness value—using a system consisting of models made up of neural networks and decision trees; the system informs the operator about norms being exceeded, i.e., it signals that the machining parameters require correcting;
- The CNC machine via controlling and compensating for the thermal deformation of ball screws—a system was developed which incorporates neural network-based models for monitoring current speed and load values and predicting the elongation of the CNC machine ball screw based on these values;
- Monitoring and predicting to predict blade wear based on various input data (cutting forces, acoustic emission and mechanical vibration). Selected measurements of physical parameters were made during the processing, and a simulation of blade wear was conducted based on these values. The neural networks were learned to predict the wear level of the blade.

A new element introduced in this study was developing DT as a new approach to selected issues while taking into account maintenance in sustainable manufacturing. Following the paradigm of sustainable production, this change in approach must be implemented on three levels: product, process and system. At the product level, the traditional 3R (reduce, reuse, recycle) concept was transformed into a more sustainable 6R approach (reduce, reuse, recycle, recover, redesign, remanufacture) [37,38], which is being gradually
implemented in eco-designing and manufacturing process planning. At the process level, an attempt was made to optimize technological processes to reduce resource use, the amount of waste generated, and hazards related to the work and working environment, which finds implementation in the designing and monitoring of machining processes.

According to the EN 13306: 2010 standard, maintenance is a “combination of all technical, administrative and managerial actions during the life cycle of an item intended to retain it in, or restore it to, a state in which it can perform the required function.” From the perspective of a business, this means that maintenance operations should focus on, e.g., ensuring the required level of reliability and availability of machines and devices and their efficiency, optimal use of the capital invested, ensuring the required level of safety for users and technical operators, monitoring the environmental aspects of machine operation and operation processes, modernizations ensuring the economic efficiency of the objects used, cooperation with providers of machines, replacement parts and services, regular improvement of technical service employee skills, etc. Therefore, selecting the correct combination of operational strategies for every technical object (corrective, preventive, predictive) must take into account not only economic and technical factors but also environmental and security-related ones and how they relate to the business strategy of the company [39,40]. It is thus evident that maintenance can have a profound impact on the effectiveness of sustainable production. Such major maintenance factors, which are of significance to the development of sustainable production processes, include replacement part and consumable management, cooperation with a machine, device and repair service providers, machine and device modernization, cooperation with design and product development departments, cooperation with production and quality departments, cooperation with OHS and environmental departments, employee competences, use of preventive and predictive operational strategies and systems for collecting and processing operational data [41–45]. Operational data, as analyzed from the perspective of sustainable production, find use in all of the analyzed areas: eco-designing, planning manufacturing and monitoring manufacturing processes. The application of DT in optimizing the physical processes of a business is presented in Figure 3.

Figure 3. Physical process optimization using digital twin (DT).
4. Procedures

The digital twins developed have a profound impact on certain parts of the business in which they were implemented. At the construction level, data acquired from users who test and use products in their everyday life were used. Construction changes were simulated in a digital twin to improve product design via correct material selection. At the production planning and monitoring stage, a DT rendered the process more efficient, reliable and flexible thanks to a wide range of innovative research. This was particularly significant at the technological process design and production process monitoring stage.

DTs can be used to digitize process models to better react to consumer trends. In order to smartly facilitate tasks in a business, digital twins make use of a range of classification models that predict particular parameters used in manufacturing goods. The models were developed in the form of classification trees: C4.5, C&RT, CHAID, reinforced decision trees and random forests.

DT and eco-designing

In the field of eco-designing, the digital twin offers a series of classification models for selecting materials that take into consideration maintenance in sustainable manufacturing.

Recycling-focused eco-designing primarily involves selecting appropriate construction materials and methods of fastening them. Products should be designed to incorporate the maximum possible amount of normalized and recyclable materials. This has a positive impact on the environment in the final stages of the product’s life cycle, including conservation and decommissioning. When selecting materials for use in products, their compatibility should also be taken into consideration: the materials used should be recyclable at the end of the product’s life cycle without the need to separate them. Additional data were also collected about the materials, based on which they can be recovered, redesigned and remanufactured. The authors’ previous papers on the development of neural network classifiers can be found in [46].

Materials can be selected in two ways. The first method involves selecting added materials based on the primary material and the required compatibility level between the two. In the second method, the user defines the primary and added materials of a new element, and the system informs them about the compatibility between these materials. The fastening of materials is another important issue. Material fastening methods must ensure quick and smooth disassembly, especially in cases where the use of incompatible or hazardous materials is necessary due to functional reasons [47]. If the compatibility is high, temporary or permanent fastening can be used. However, if the compatibility is low or if the materials are incompatible, only temporary fasteners should be used.

Files containing examples used to develop the classifiers were created based on an analysis of material properties, including material name (e.g., PVC), a tensile strength in megapascals (e.g., 35.5), density in grams per cubic centimeter (e.g., 7.88), processing temperature in Celsius (e.g., 20.8), elongation at yield (Re) in percentage (e.g., 5.5), Young’s modulus E in gigapascals (e.g., 4.61), dielectric constant (e.g., 2.0), dielectric strength in kilowatts per millimeter (e.g., 22.0), water absorptivity in percentage (e.g., 22.5), recycling cost in PLN per kilogram (1 PLN = 0.23 euros) (e.g., 4.25); a positive value denotes profit from selling the material, while a negative value denotes disposal cost, negative impact on the environment (e.g., true) and the name of the added material (e.g., ABS). The material parameters were established with designers.

For the first problem, i.e., material selection where the user defines a material selected for an element (part) of the product and the required level of compatibility, and the system recommended which material to add, a file was created containing examples with the conditional attributes "main material" and "compatibility", and the decision class "added material". For the second problem, i.e., material selection where the user defines the main and added material of a new element, and the system informs the user about the compatibility between these materials, a file containing examples was created with the conditional attributes "main material" and "added material", and the decision class "compatibility". For the third problem, i.e., selecting material connections, a file containing
Examples were created with the following conditional attributes: "main material", "added material" and "compatibility", and the decision class "connection type". Example classifiers are presented in Figure 4 (in the StatSoft Statistica DataMiner software). The root of the tree shown in Figure 4a,b was the main material attribute, which provided the highest information gain (the division criterion of examples).

Figure 4. Example classifiers: (a) for selecting a second product material, (b) for determining the compatibility of two materials.

DT and planning manufacturing

In the field of planning manufacturing, the digital twin contains many classification models for technological process design while taking into account maintenance in sustainable production. The selection models are based on the following criteria matching the technologies used in a given manufacturing company and related to particular selections:

- Semi-finished products: this selection takes into account production scale, part quality, part shape, semi-finished product availability and type of material;
- Technological process structure: this selection determines the order of technological operations and procedures, taking into account production scale, type of semi-finished product, requirements and shape of the workpiece;
- Workpiece equipment: taking into account production scale, type of semi-finished product, machining method, part shape; also, the technologist assesses the speed of mounting and positioning repeatability;
- Machine tools for technological operations and procedures: taking into account machining accuracy, workpiece dimensions, expected load, production efficiency and hourly cost of machine tool operation;
- Machining tools for technological operations and procedures: taking into account the shape of the machined surface, machining method, scale of production, type and accuracy of the machining, kind of machine tool and type of workpiece material;
- Tooling: taking into account the dimensions of the tooling compatible with the tool and machine;
- Machining parameters are taking into account the type of machine tool, part material, surface quality requirements and type of blade material.

Classification models were developed for all of these selections. The authors’ early work demonstrates certain selection aspects whose classification models were based on neural networks [48]. As an example, an optimal classifier for tool holder selection using decision trees was presented. Tooling selection models were built in a similar fashion to the other classification models. Figure 5 depicts a decision tree, i.e., the optimal classifier for selecting a tool holder. This selection occurred when the tool was not compatible with
the machine. The root of the tree shown in Figure 5 was the milling tool attribute, which provided the highest information gain (the division criterion of examples).

**DT and monitoring manufacturing**

In the field of monitoring manufacturing, the digital twin comprises a range of classification and predictive models for monitoring production processes, which take into consideration maintenance in sustainable manufacturing. This module is constantly being developed. Models were developed which facilitate the elimination of machining process disruptions (simulations were conducted of these disruptions and the methods of their repair), maintaining process stability based on control charts, signaling the need for machining parameter correction based on workpiece surface roughness assessment, CNC machine control including compensating for the thermal deformation of ball screws, predicting blade wear based on various input data, in particular cutting forces, acoustic emission and mechanical vibrations. The authors’ earlier work was related to constructing models in the form of neural networks [49–51] and other authors [52]. This paper demonstrates classification models in the form of decision trees for predicting VBc - tool wear (Figure 6) based on cutting force. The root of the tree shown in Figure 5 was the time of cutting force attribute, which provided the highest information gain (the division criterion of examples).

**Decision tree for Tool holder**

Number of shared nodes: 5, number of end nodes: 6

![Decision tree](image)

Figure 5. Optimal classifier for tool holder selection—the tree with strictly defined nodes.
Figure 6. Tool wear (VBc) level prediction classifier utilizing cutting force analysis.

5. Results

The decision tree models were verified and assessed for the purpose of selecting the most effective models. The models were developed in the form of classification trees: C&RT, CHAID, reinforced decision trees and random forests. Model parameters were changed for every type of tree. The following classification parameters were set for the C&RT model: the cost of erroneous classification, the goodness of fit and a priori probability. The stop criterion incorporated the stop rule: cut according to variance and the parameter of a minimum number of examples per node. For the CHAID model, the cost of erroneous classification was set. The stop criterion required a minimum number of examples per node. For the reinforced tree model, the following classification parameters were set: cost of an erroneous classification and a priori probability. The stop criterion incorporated the parameter of the minimum number of examples per node. For the random forest model, the following classification parameters were set: cost of an erroneous classification and a priori probability. The stop criterion incorporated the parameter of the minimum number of examples per node.

The cost of erroneous classification relates to the distribution of examples across classes. Cost minimization is equivalent to minimizing the overall proportion of erroneously classified cases when a priori probabilities are proportional to class size, and the cost of erroneous classification is equal for every class [53].

The goodness of fit involves finding the split for every predictor, which offers the highest increase in the goodness of fit. How does one define the goodness of fit increase? The goodness of fit can be measured in three ways. The Gini impurity measure reaches zero when a given node contains only a single class (with a priori probabilities estimated based on the size of classes and equal cost of erroneous classification, the Gini measure is calculated as the sum of the products of all class proportion pairs in a given node; and reaches the maximum when the number of classes in a given node is equal). The Gini index
was the preferred measure of goodness of fit for the developers of C&RT [53]. A perfect fit means perfect classification.

A priori probabilities determine the probability that a given case or object fits a given class without any prior knowledge about the predictive variables in the model. They are used in cost minimization and may impact case or object classification. Cost minimization is equivalent to minimizing the general proportion of erroneously classified cases when a priori probabilities are proportional to class size (and the cost of erroneous classification is equal for every class), as predicting should be more effective for larger classes and yield a generally lower rate of erroneous classifications [53].

As for the attributes: attached is a different form of the tree, where you can see what the root is, how these attributes are distributed. Classification matrices demonstrate the effectiveness of the classifiers. These were developed and analyzed for every decision tree to choose the most optimal one. An example classification matrix is presented in Figure 7. The material added to the main material constitutes the observed and predicted class. Table 1 presents classes of added materials, which we can connect to main materials.

**Figure 7.** Decision tree classification matrix for selecting materials based on the added material (the observed and predicted values of the additional material).
Table 1. Classes of added materials.

<table>
<thead>
<tr>
<th>Class</th>
<th>Added Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SAN, PVC, PMMA, PC, PBT, ASA, ABS</td>
</tr>
<tr>
<td>2</td>
<td>PS, PPE, PP, POM, PET, PE, PA</td>
</tr>
<tr>
<td>3</td>
<td>PA</td>
</tr>
<tr>
<td>4</td>
<td>SAN, PS, PP, POM, PMMA, PET, PE, PBT, ASA, ABS</td>
</tr>
<tr>
<td>5</td>
<td>PVC, PPE, PC</td>
</tr>
<tr>
<td>6</td>
<td>SAN, PC, PBT, ASA, ABS</td>
</tr>
<tr>
<td>7</td>
<td>PS, PPE, PP, POM, PMMA, PET, PE, PA</td>
</tr>
<tr>
<td>8</td>
<td>PVC</td>
</tr>
<tr>
<td>9</td>
<td>SAN, PMMA, PET, PC, PBT, ASA, ABS</td>
</tr>
<tr>
<td>10</td>
<td>PS, PPE, PP, PE</td>
</tr>
<tr>
<td>11</td>
<td>PVC, POM, PA</td>
</tr>
<tr>
<td>12</td>
<td>PP, PE</td>
</tr>
<tr>
<td>13</td>
<td>PA, PVC, PC</td>
</tr>
<tr>
<td>14</td>
<td>SAN, PS, PPE, POM, PMMA, PET, PBT, ASA, ABS</td>
</tr>
<tr>
<td>15</td>
<td>PET, PC, PBT, ASA, ABS</td>
</tr>
<tr>
<td>16</td>
<td>SAN, PA, PVC, PE, PS, PMMA, PPE, POM, PP</td>
</tr>
<tr>
<td>17</td>
<td>PC, ABS, PMMA, ASA</td>
</tr>
<tr>
<td>18</td>
<td>PA, SAN, PE, PVC, PET, PS, POM, PP, PPE</td>
</tr>
<tr>
<td>19</td>
<td>PBT</td>
</tr>
<tr>
<td>20</td>
<td>POM</td>
</tr>
<tr>
<td>21</td>
<td>ABS, PVC, ASA, SAN, PA, PS, PBT, PE, PET, PPE, PP</td>
</tr>
<tr>
<td>22</td>
<td>PMMA, PC</td>
</tr>
<tr>
<td>23</td>
<td>PP</td>
</tr>
<tr>
<td>24</td>
<td>PVC, PE, PA</td>
</tr>
<tr>
<td>25</td>
<td>ABS, SAN, ASA, PS, PBT, PPE, PC, POM, PET, PMMA</td>
</tr>
<tr>
<td>26</td>
<td>PS, PPE</td>
</tr>
<tr>
<td>27</td>
<td>ABS, ASA, SAN, PA, PP, PBT, POM, PC, PMMA, PET, PE</td>
</tr>
<tr>
<td>28</td>
<td>PS</td>
</tr>
<tr>
<td>29</td>
<td>ABS, ASA, PA, POM, PBT, PP, PC, PPE, PE, PVC, PET, SAN, PMMA</td>
</tr>
<tr>
<td>30</td>
<td>SAN, ABS, PVC, ASA, POM, PMMA</td>
</tr>
<tr>
<td>31</td>
<td>PS, PP, PE</td>
</tr>
<tr>
<td>32</td>
<td>PET, PPE, PA, PC, PBT</td>
</tr>
</tbody>
</table>

The costs of misclassifications concern the distribution of examples between classes. The minimization of costs corresponds to minimizing the proportion of misclassified cases where probabilities are taken a priori proportional to the size of the classes, and the costs of misclassification equal for each class. In order to assess the models, a cross-match test based on a test sample was used, as well as 10-fold cross-validation. Table 2 compares the cost, risk assessment and standard error of the decision trees.

Table 2. Comparison of classifiers of material selection based on the added material.

<table>
<thead>
<tr>
<th>Classifier Type</th>
<th>Classifier Assessment</th>
<th>Risk Cost/Assessment</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>C&amp;RT</td>
<td>0.035217</td>
<td>0.110509</td>
<td></td>
</tr>
<tr>
<td>CHAID</td>
<td>0.201187</td>
<td>0.016060</td>
<td></td>
</tr>
<tr>
<td>Reinforced decision trees</td>
<td>0.000001</td>
<td>0.000001</td>
<td></td>
</tr>
<tr>
<td>Random forest</td>
<td>0.021846</td>
<td>0.011653</td>
<td></td>
</tr>
</tbody>
</table>

Reinforced decision trees and random forest proved to be the most effective models. Correct classification rates were 100% and 97.64%, respectively. Decision rules were developed based on the trees, which were then implemented in the system. Example rules developed based on the second decision tree are:

1. if added material = “ABS” and main material = “ABS” then compatibility = “good”
2. if added material = “ABS” and main material = “ASA” then compatibility = “good”
3. if added material = “ABS” and main material = “PA” then compatibility = “limited”
4. if added material = “ABS” and main material = “PE” then compatibility = “limited”

The rules were imported into an existing module for determining material compatibility: main and added product (Figure 8).

![Figure 8. Module for determining material compatibility.](image)

The result is a system that facilitates a company’s operations in the following areas: eco-designing, planning manufacturing and monitoring manufacturing while accounting for maintenance in sustainable manufacturing.

Figure 9 depicts a system architecture that incorporates digital twins in areas of importance to the company for which it was developed by us. It provides a dynamic virtual representation of a physical object/process at selected stages of its life cycle. Our DT system includes physical processes chosen area: eco-designing, manufacturing planning and monitoring. In databases, we have descriptions of these physical processes. Systems can both present observations concerning the current state of the system and answer the questions “what-if”. Many decision trees have been defined for each decision problem in these three areas and in their specific tasks. Some trees from various areas are shown in Figures 4–6. Figure 7 shows the evaluation of an exemplary tree. All trees were assessed in this way, and the best ones were selected. The best decision trees are in the knowledge base. Decision rules were created from the trees, which are included in computer applications. The computer programs aid engineers in creating new solutions. Figure 9 shows the architecture of a DT system in the company.

The DT system architecture combines:

- Physical processes that are performed in the departments of the company;
- Digital mapping of physical processes, which is transferred to the database in the form of data;
- Based on standardized and coded data, AI models are created, which are then stored in the knowledge base;
- The system control module supervises all these activities.

When the customer’s order arrives, the control module divides the tasks into individual departments of the company. The order is being processed. New physical processes are the basis for expanding the data and knowledge in the system. This is how the system learns all the time. Data that is relevant to sustainable production and maintenance are collected during the operation of the system.
way, and the best ones were selected. The best decision trees are in the knowledge base. Decision rules were created from the trees, which are included in computer applications. The computer programs aid engineers in creating new solutions. Figure 9 shows the architecture of a DT system.

![Figure 9. The architecture of a DT system.](image)

6. Conclusions

The system supports selected areas: eco-designing, planning manufacturing and monitoring manufacturing, for the purpose of manufacturing goods, and may be of help in sustainable manufacturing. This is of particular utility in situations where green design and manufacturing requirements are difficult to formalize, uncertain, incomplete or unavailable.

Machine-learning proved a precious tool for acquiring knowledge in facilitating operations in companies. It was discovered that designers and technologists frequently make decisions based on intuition and are unable to formulate with precision the rules behind their choices. As they acquire knowledge, machine-learning methods can be used to create classic rules used in smart systems automatically. In the case of sustainable product design, decision tree induction must be used as the classification method due to the large volumes of input data presented in symbolic form. Rules generated based on decision trees are more concise, and the time required to draw conclusions is significantly reduced. Our research
demonstrates that machine learning can be effectively applied in DT. Thanks to DT, the system gains improved connectivity and flexibility, in addition to increased intelligence as a result of introducing knowledge and experience to a computer program. Moreover, manufacturing efficiency is increased, as well as the quality of the goods produced, reducing costs and increasing profits. In addition to improving manufacturing processes, DT renders it possible to transition into individualized production. It also offers the possibility of conducting simulations for the purpose of amending manufacturing process irregularities and accounting for maintenance in sustainable manufacturing. All data are saved in a database, and the knowledge and experience in the system knowledge base. Utilizing data and analyses generated by sensors built into smart products and instruments, it is possible to streamline operations, reconfigurations and maintenance processes. DT facilitates decision-making in multidimensional processes, strategic planning and process prediction with the help of knowledge recycling and experience. DT expands conventional engineering analyses to include information integration to create digital product lifecycles. As monitoring and simulating in real time enables predicting tool and machine damage, all data necessary for this type of prediction are saved in the system database.

Comparing research results with solutions developed by other authors poses difficulties due to the fact that no identical product has been identified on the market, which would include all set elements from the same industry. In addition, the boundary conditions for the application of the solution in question could potentially differ to the degree that would significantly impact the final numerical result, rendering a direct comparison impossible. This indicates the necessity to develop a comparative research standard pertaining to DT solutions, including a division into industries according to their specificity and the level of complexity of the solutions themselves.

Further research will involve integrating the system as a smart module with a system responsible for managing an entire manufacturing company. This will not only enable further verification of current assumptions with the use of large sets of real manufacturing data but also contribute to gaining a more thorough understanding of how to implement DT in an existing business.

From a scientific point of view, we should also monitor the activity of research groups and their progress as published and delivered at conferences and through initiatives integrating the research community into larger umbrella projects. This will allow us to know how our research results are linked to other research, i.e., how they enrich the research carried out by other scientists and also how they affect industrial implementation. This will allow us to better target our own research, including optimizing the effect, speeding up the development of research-based on the experience of other teams, avoiding their mistakes, but also avoiding the risks of even a lack of sustainability.


Funding: The work presented in the paper was co-financed under the 0613/SBAD/4677 grant for Poznan University of Technology and a grant to maintain the research potential of Kazimierz Wielki University.


Conflicts of Interest: The authors declare no conflict of interest.
References


