A PATTERN LANGUAGE APPROACH TO IDENTIFY APPROPRIATE MACHINE LEARNING ALGORITHMS IN THE CONTEXT OF PRODUCT DEVELOPMENT

Sonntag, Sebastian;
Luttmer, Janosch;
Pluhnau, Robin;
Nagarajah, Arun

University of Duisburg-Essen

ABSTRACT

The product development process faces several challenges, such as an increasing and differentiated number of customer requirements, increasing product complexity, and shortened time-to-market. To address these challenges, the implementation of automation approaches in form of machine learning (ML) algorithms appears promising. However, companies lack the implementation of these approaches in their processes, inter alia due to inadequate knowledge and experience in this field. Therefore, the aim of this paper is to develop a structured formulized way of characterising ML algorithms, which can support non-experts in identifying the optimal algorithm to solve a given problem. First, existing approaches covering the determination of appropriate ML algorithms for a given task are examined. Based on this, a pattern language approach is introduced to characterise ML algorithms and problems, allowing matching to be performed to identify the most suitable one for a given task. Due to their broad application, the concept is demonstrated by creating patterns for decision trees and artificial neural networks. A study is conducted to prove that the proposed concept is appropriate to support the ML algorithm selection.

Keywords: Product development, Artificial intelligence, Machine learning, Pattern language, Decision making

Contact:
Sonntag, Sebastian
Universität Duisburg-Essen
Germany
sebastian.sonntag@uni-due.de

Cite this article: Sonntag, S., Luttmer, J., Pluhnau, R., Nagarajah, A. (2023) 'A Pattern language Approach to Identify Appropriate Machine Learning Algorithms in the Context of Product Development', in *Proceedings of the International Conference on Engineering Design (ICED23),* Bordeaux, France, 24-28 July 2023. DOI:10.1017/pds.2023.37

1 INTRODUCTION

Currently, the product development process is faced with an increasing number of novel trends and resulting challenges. For instance, different national and sociocultural frameworks and the shift towards product service systems lead to a rising number of customer requirements that have to be identified and taken into account (Bertoni, 2018; Jetter, 2005). When combined with the high level of innovation necessary to remain competitive in the current market situation, a trade-off arises where more time is needed for innovation development, but the overall development time has to be shortened (Lasi *et al.*, 2014; Baxter *et al.*, 2008).

Implementing automation approaches in the product development process has the potential to overcome the previously presented problems and resulting trade-offs. Thus, there exists the opportunity to automate repetitive tasks, such as the construction of machine parts, which can result in time and cost savings (Entner *et al.*, 2019). Those time savings are available to invest in creative tasks, where automation approaches can support the developer, e.g. in identifying novel solutions, to compensate for one's lack of experience (Chakrabarti *et al.*, 2011).

1.1 Automation in the context of the product development process

The use of automation approaches in the early phase of product development can contribute to speeding up the preparation of quotations, improving cost calculations, adapting products more quickly to customer requirements, improving quality by reducing errors, and building on this, shortening the time to market (Rigger and Vosgien, 2018). Existing approaches range from identifying correlations between technical requirements and product characteristics (Shabestari and Bender, 2017), up to concept generation (Hein and Condat, 2018) and can be assigned to the machine learning (ML) domain.

ML can be described as a subcategory of artificial intelligence and deals with the development of algorithms that can, among other things, analyse a dataset and make predictions based on it (Dunjko and Briegel, 2018).

Artificial neural networks (ANN) and decision trees (DT) are particularly important in this context. Former research identifies them as the most relevant and widely applied ML algorithms during the early phase of product development (Riesener *et al.*, 2020; Shabestari and Bender, 2017).

Both ML algorithms can be assigned to the class of supervised learning, where the algorithm learns the connection between a given input and the corresponding outcome. Therefore, the training data has to be labelled i.e. the result must be known (Zhang, 2010).

The basic idea of DT is to identify the attributes of a dataset that contribute the most information gain to assign the data to a known label. Therefore, different branches are derived that represent different attribute combinations until a label can be assigned, forming a tree-like structure. Since the DT's inner structure, i.e. the decision leading to the classification itself can be further analysed, it is called a "white box". (Reed and Gillies, 2016)

In contrast, ANN are based on three different layers of neurons, representing the input parameters (first layer) and the output parameters (third layer). The neurons are interconnected, with a weight assigned to the connection that strengthens or weakens the signal between them. The weights are adjusted during training until the expected output can be mapped to a new input. Since the decision which leads to the prediction cannot be retraced, ANN can be called a "black box". (Rashid and Langenau, 2017)

Although the benefits of implementing such algorithms in company processes are promising, only a fraction of them actually do so. Recent studies indicate that 75% of the companies surveyed believe that automation approaches, such as artificial intelligence have no significance for them at all (Giering, 2022). Small and medium-sized enterprises, in particular, have not yet introduced automation in product development. Thus only 16% plan to implement automation approaches in their existing processes (Metternich *et al.*, 2021). The reasons for this can be summarized as a lack of experience and knowledge of the companies to recognize opportunities for automation, and based on this, identify suitable implementation options (Lundborg *et al.*, 2019). This applies to determining the appropriate ML algorithm for a given problem under specific conditions, which usually requires an expert. However, the necessary knowledge is contained in existing approaches where certain ML algorithms have been proven effective for various of differentiated tasks.

https://doi.org/10.1017/pds.2023.37 Published online by Cambridge University Press

1.2 Research question

To enhance the use of automation approaches in the field of product development, companies must be enabled to identify the most suitable ML algorithm for potential applications to solve prevailing problems. As mentioned before, ANN and DT are the most relevant ML algorithms applied in the early phase of product development.

Focusing the research on these two ML algorithms supports the algorithm selection for a large portion of the arising automatable problems in this field.

Therefore, this research paper aims to answer the following question:

- How can the knowledge necessary to use ML algorithms in the context of product development be made accessible for non-experts so that, the optimal ML algorithm to solve a given problem can be identified?

To address this problem, the central goal is to introduce a structured form of characterization of the DT and ANN algorithms. The attributes of the formalization cover central common properties, that make it possible to identify the most suitable algorithm to solve a given problem. Therefore, the problem definition has to be formalized in the same way and with the same elements as the characterization of the algorithms, in order to enable a comparison between them.

The work is structured as follows. Chapter 2 presents the existing approaches to support the choice of an appropriate ML algorithm and points out existing gaps. Chapter 3 provides an introduction to the concept of a pattern language. The proposed concept to characterize ML algorithms by the use of patterns is derived from the example of ANN and DT in chapter 4. Chapter 5 provides a study to prove the practicability and functionality of the introduced patterns. Finally, a summary of the outcome and future work will be given in chapter 6.

2 RELATED WORK

To use ML as support for solving a task, one has to identify the most suitable ML algorithm out of a large number of one's available. In general, the procedure is based on choosing potentially appropriate algorithms and testing their applicability to the data set through trial-and-error. This is not only a laborious course of action, the initial guess of suitable algorithms requires a certain amount of experience. (Lickert *et al.*, 2021) To overcome this problem, there are several approaches to help finding the right ML algorithm for a given task (Waring *et al.*, 2020; Riesener *et al.*, 2020; Lickert *et al.*, 2021; Gerschütz *et al.*, 2021; Hashimi *et al.*, 2015).

A variety of the existing approaches can be summarized by the term *AutoML*. Those consider the selection of the appropriate ML algorithm while determining the corresponding hyperparameters as an optimization problem, which is called the CASH problem. Supported ML algorithms are iteratively tested to solve a known task, and their hyperparameters are adjusted for the given input data. In the end, the ML algorithm with the best results is selected. Since it is an iterative process, it is computationally and time intensive. Best known examples are *Auto-Weka* and *Auto-Sklearn*. (Waring *et al.*, 2020)

Riesener et al. (2020) define evaluation criteria to facilitate the selection of ML algorithms in the context of product development. Therefore, the most used ML algorithms in this field have been evaluated by their accuracy, training time, computational effort, tolerance against erroneous and irrelevant parameters, online capability, and transparency.

Lickert et al. (2021) present a similar approach, where supervised learning algorithms for classification problems are analysed and characterized towards their suitability in the context of reverse logistics. The defined criteria are divided into input, implementation, and output. In addition, they include aspects such as the size of the dataset, the stability of the algorithm, and the interpretability of the results.

Gerschütz et al. (2021) established a wiki to assist developers in identifying appropriate ML algorithms in the context of the product development process. The ML algorithms are characterized in terms of potential use cases in product development e.g., requirements identification, required inputs and provided outputs, existing tools for implementation, and their limitations. Thus, this concept corresponds with the creation of a knowledge database for ML algorithms.

Hashimi et al. (2015) define evaluation criteria for the effectiveness of text mining algorithms to facilitate their selection.

The shortcomings of the analysed research above can be summarized as followed:

- The characterization of the problem itself or for what type of problem a ML algorithm can be used is not addressed.

https://doi.org/10.1017/pds.2023.37 Published online by Cambridge University Press

- The characterization of the ML algorithms is incomplete or tailored solely to certain types of ML algorithms, making comprehensible selection difficult. For example in (Lickert *et al.*, 2021) the properties of DT regarding the number of observations and parameters, that have to be available to use the algorithm as intended are missing.
- The possibility to describe a problem to enable the comparison with a ML algorithm is not considered, thus complicating the application. For example, it is not stated how the established evaluation criteria by Riesener et al. (2020) has to be applied to a given problem formulation. Moreover, the comparison between the criteria and the task and the ML algorithm respectively, requires expert knowledge and is only rational to a certain extent.

In summary, no solution allows for an informed selection of ML algorithms based on a well-defined problem within the product development process.

To overcome these shortcomings a structured and formalized characterization of ML algorithms is needed. To ensure completeness, central common properties, as well as general conditions, must be covered, which give the characterization its structured form. A key category has to be the generalized problem that the ML algorithm can ideally be used for. This enables the identification of the ML algorithm as such for a given problem definition.

One approach that offers the ability describing known solutions to common problems in that kind of structured way is pattern language. Pattern language introduces structured documentation of solutions to known problems, enabling problem solving by non-experts. Thus, it is used for the same underlying problem of providing knowledge. The concept of pattern language has been used in the field of software development for several years and has proven to be a robust solution (Borchers, 2001).

3 PATTERN LANGUAGE

The term pattern language was first introduced by Christopher Alexander, who developed patterns for repeating tasks related to urban architecture (Borchers, 2001).

One central goal of defining patterns is to create a common vocabulary, that facilitates communication among experts as well as experts and laymen. In general, patterns enable the formulation and documentation of known and proven solutions for a specific task. Thus, they are intended to make expert knowledge easily accessible and understandable to users. The patterns themselves contain mutual references e.g., to patterns to be applied later or in a different context, creating a hierarchical structure and thus a pattern language.

Common patterns consist of a distinct name to facilitate communication, the context the corresponding problem occurs, a description of the problem including the conditions to which successful deployment of the solution is tied, the solution in a general form such as a blueprint, and references to related patterns. (Borchers, 2001)

Deviating of architecture, patterns are already used in a variety of domains, including software development, human-computer interaction, economics, data mining model evaluation or product lifecycle management (Borchers, 2001; Souza *et al.*, 2002; Feldhusen and Bungert, 2007).

The advantages described above make the construction of a pattern language the ideal solution for providing the necessary knowledge to identify and apply a ML algorithm to a specific task. By analysing existing approaches of applying DT and ANN algorithms in the context of product development, the underlying problems and conditions can be identified and linked to the algorithm found to be applicable. This enables the construction of a pattern language, which in turn enables the identification of proven solution concepts and thus the optimal ML algorithm for new or similar problems. Moreover, based on the definition of references to other patterns i.e., ML algorithms, a combined application can be proposed.

4 DEVELOPMENT OF PATTERNS

To overcome the shortcomings of existing approaches examined in chapter 2 the patterns to be developed are intended to achieve the goals listed in Table 1.

To this end, it is useful to define a structure that departs from Christopher Alexander's and other patterns in use. Thereby, for the description of the individual attributes (e.g., problems for which an ML algorithm is suitable), clearly defined element modules are to be introduced instead of textually varying formulations.

https://doi.org/10.1017/pds.2023.37 Published online by Cambridge University Press

Table 1. Goals to be achieved by the patterns

No.	Goal
1)	Document the kind of problem and circumstances in which the DT and ANN algorithms have been successfully used in the past and proven their worth.
2)	Define the requirements tied to the application of the DT and ANN algorithms.
3)	Hold references to other ML algorithms that could be applied in context.
4)	Provide defined element modules to characterize common problems in the same way as the DT and ANN algorithms, so that matching between problem formulation and pattern can subsequently be performed.

By simultaneously using the element modules to characterize the problem for which an appropriate ML algorithm is being sought, an unambiguous comparison can be made, and thus the most suitable one can be identified.

The construction of the patterns is based on the manual investigation of a total of ten research papers on the use of ML algorithms in the context of product development. Five of the research papers refer to DT and five to ANN. In selecting the publications, attention was paid, among other things, to the fact that good results were achieved through the application of the methods. This implies, that the methods applied are suitable to solve the described problems therein. A list of the examined research papers can be found in Table 2. In addition, one research article covering the theoretical background of ANN (Management Association, 2022) and DT (Ye, 2013) were analysed with respect to possible limitations in the application of the ML algorithms that may not be considered in the approaches proposed in the research papers.

Table 2. Summary of the analysed research papers

Algorithm	Research paper	Underlying problem formulation
ANN	(Chen and Chang, 2016)	Predicting customers perception of the shape of knives.
	(Hsiao and Huang, 2002)	Anticipation of the customer's impression (e.g. practical, elegant, stable) regarding the product design.
	(Seo *et al.*, 2005)	Predicting the environmental impact (over the lifetime) of a product.
	(Tang *et al.*, 2013)	Identifying the optimal design parameters of a cell phone based on customer's expected perception.
	(Kutschenreiter-Praszkiewicz, 2013)	Determining the probability of occurrence of warranty claims for a gearbox.
DT	(Bang and Selva, 2016)	Identifying the design parameters that have a decisive influence on certain user-defined properties of the design.
	(Hoyle *et al.*, 2009)	Identifying the range of parameters for a design that will influence a certain evaluation of the customer.
	(Tucker and Kim, 2009)	Identifying the combination of product attributes to create product concepts that meet customer requirements.
	(Reed and Gillies, 2016)	Identifying the permissible range of parameters for the design of chairs.
	(Wang and Zhang, 2017)	Identifying the product specifications that meet the customer's requirements.

The fundamental question that arises in analysing the research papers is: What aspects of the problems made the application of the ML algorithms successful? In other words: What made them the best choice in the particular cases?

Based on that, following questions can be derived:

- To what kind of problems have the ML algorithms been applied, and what were the expected results?
- What characterizes the problems i.e., what kind of conditions characterize the problems?
- What aspects had to be considered when applying the ML algorithms?

By analysing the research papers and answering the questions above, similar, and common attributes of the ML algorithms have been derived. Those are necessary to identify the ML algorithms as the optimal

https://doi.org/10.1017/pds.2023.37 Published online by Cambridge University Press

solution for a given problem formulation and thus define the basic structure of the pattern. The logical order of the definition of the attributes in the pattern represents a hierarchical structure. Therefore, the most important one is listed first and those whose characteristics can be adapted in relation to the problem (e.g. converting alphabetical values in numerical ones) or which are only relevant in some cases, follow afterwards. The corresponding element modules of the individual attributes were examined from the research and an attempt was made to formulate them in a generalised way.
The defined patterns for DT and ANN can be seen in Figure 1. The following attributes were identified and defined:

Problem
The first element that must be considered to make an informed decision on the use of a ML algorithm is the task or activity that can be solved with it. If the ML algorithm does not match to the problem formulation at hand, it is highly unlikely that the ML algorithm is the optimal one to solve the problem. The literature distinguishes between regression, classification, and prediction (Riesener *et al.*, 2020). However, the analysis of the research indicates, that the task can be further specified and assigned to the ML algorithms. For a concurrent application to characterize the problem, it is crucial to make the formalization as specific as necessary and as generalized as possible.

Input data
The underlying data basis is crucial for the applicability of the analysed ML algorithms. As relevant aspects to be considered, data type, data value, and completeness can be elaborated. Unlike the attribute problem, conversions of the underlying data can be made regarding data type and data value to make the algorithm practicable for use.

Data type
The term data type denotes the characteristic of the data like continuous and discrete variables.

Data value
In contrast to data type, the term data value distinguishes between the kind of data. These can be either numerical or alphabetical values.

Completeness
This attribute refers the completeness of the dataset i.e., those with missing attributes and complete ones, and determines whether or not this is required for the application of the algorithm.

Training data
The attribute training data defines the preconditions that must apply to the data used to train the ML algorithms i.e., the initial data set. Therefore, a distinction is made between labelled and number of datasets and parameters.

Labelled
The term labelled describes whether or not the data in the training set must have a classification or label, i.e., a known desired output beforehand.

Number of datasets and parameters
To be able to train the algorithm properly a sufficient number of datasets have to be available. The number of necessary datasets is highly related to the parameters contained in the data. For example, a small number of datasets can be sufficient if the number of contained parameters is equally small. The range defined in the patterns is based on the combination of the largest number of parameters with the smallest available dataset observed in the research.

Traceability
For the selection of a method, it can be decisive whether, a decision that led to the categorization of a data set is comprehensible and thus the ML algorithm itself is transparent or not. This allows getting deeper insights and interpretations towards the underlying database.

https://doi.org/10.1017/pds.2023.37 Published online by Cambridge University Press

References
The attribute references includes the link to other algorithms that can be used in the context of the application or can be combined with the characterized ML algorithm. In addition, approaches such as Word2Vec can be linked to the pattern, that enables the conversion of alphabetical into numerical data.

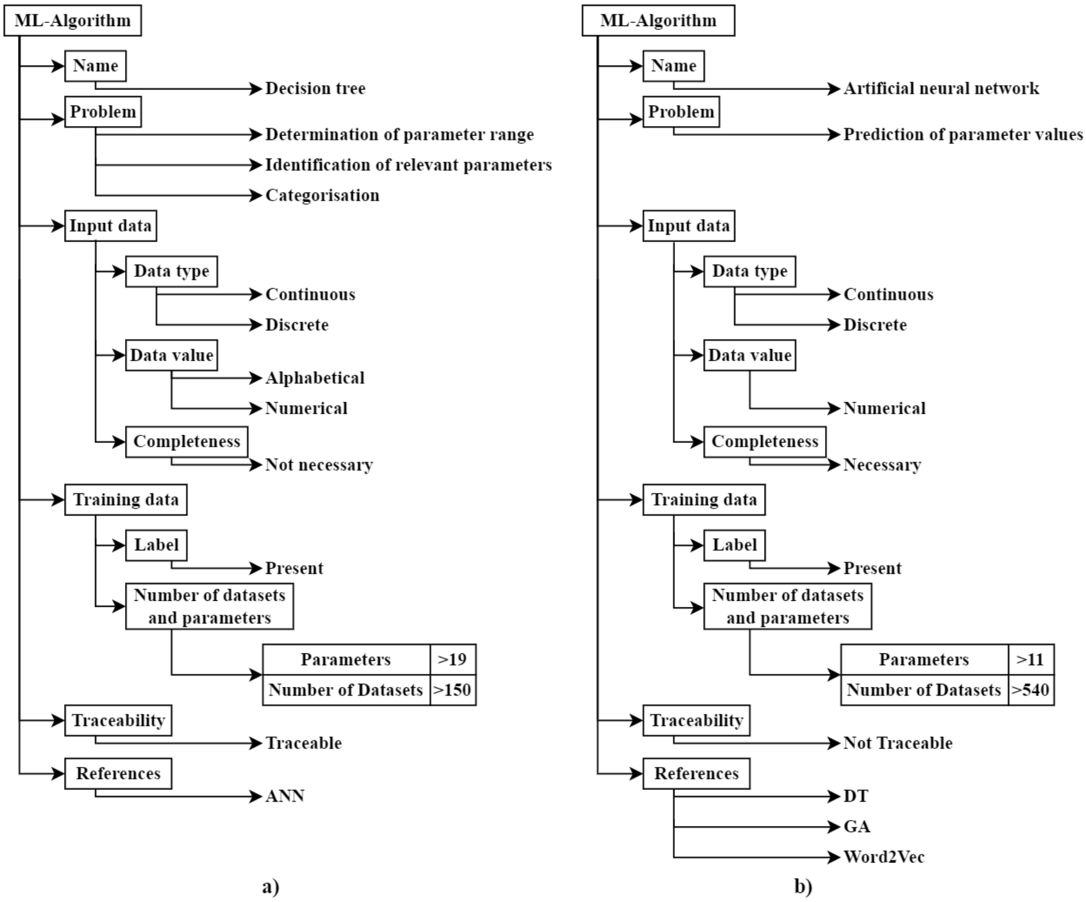

Figure 1. a) Derived pattern for decision trees, b) derived pattern for artificial neural networks

5 VALIDATION

The patterns for ANN and DT presented in Chapter 4 are intended to enable non-experts to identify the most appropriate ML algorithm for a given problem formulation that is solvable with either ML algorithm. To examine the practicability and functionality of the patterns a study has been conducted. The objective was to proof that:

1. It is possible for a user to translate an arbitrary problem formulation into a formulized problem definition using the presented structure and element modules of the patterns.
2. It is possible to determine the optimal ML algorithm to solve the task based on the comparison of the formalized problem definition and predefined patterns for DT and ANN.

Four additional research papers covering the use of ML algorithms during the early phase of product development have been identified. Two of them cover the application of DT (Bae and Kim, 2011; Zhan *et al.*, 2019). The other two cover the application of ANN (Tseng *et al.*, 2012; Sousa *et al.*, 2000). The underlying problems, as well as the associated general conditions were extracted and described in a solution neutral manner.
The derived problem formulations were handed out to seven engineers with a master's degree but without any preliminary knowledge about machine learning. The task was to characterize the problem formulation with the help of the structure and element modules introduced by the patterns. Afterwards the engineers had to identify the most suitable algorithm to solve the formalized problems based on the given patterns for DT and ANN.
A summary of the results of the study can be seen in Table 3. The attributes of the patterns to be identified are juxtaposed against the research papers containing the derived problems. The symbols

https://doi.org/10.1017/pds.2023.37 Published online by Cambridge University Press

indicate whether the subject has been able to derive the correct element module of the corresponding attribute and has been able to identify the optimal ML algorithm (✓) or not (✗).

Table 3. Summary of the study results

Pattern attributes	Research papers containing the derived problem formulations																											
	(Tseng *et al.*, 2012)							(Bae and Kim, 2011)							(Sousa *et al.*, 2000)							(Zhan *et al.*, 2019)						
Problem	✓	✓	✓	✓	✗	✗	✓	✓	✓	✗	✗	✓	✓	✓	✗	✓	✓	✗	✓	✗	✓	✗	✓	✓	✓	✓	✗	✗
Data type	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Data value	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗	✓
Completeness	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Label	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✗	✓	✗	✓	✓	✓	✓
Number of datasets and parameters	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Traceability	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✗	✗	✓	✗	✓	✗	✓
Identified algorithm	✓	✓	✓	✓	✗	✗	✓	✓	✓	✗	✗	✓	✓	✓	✗	✓	✓	✗	✓	✗	✓	✗	✗	✓	✓	✓	✗	✗

Looking at the results in relation to objective number one, the engineers have considerable problems deriving the correct element modules for the attributes *problem*, *data value* and *traceability*. One reason for this can be identified as the given problem formulation itself. The engineers reported difficulties in understanding the general conditions of the problems. This becomes evident regarding the attribute *data value*, which should be trivial to determine. The same applies to the attribute *traceability* where the correct assignment of the element module highly depends on understanding the underlying problem. When it comes to identifying the appropriate generalised *problem* to describe the given task, the applicability of the corresponding element modules does not seem to be sufficiently given. Here, a too general description of the element modules can make a simple assignment difficult. Nevertheless, most of the attributes have been successfully identified, thus the underlying concept seems to be practicable.

With regards towards the second objective, the results imply that the misidentification of the algorithm emerges from the misidentification of the *problem*. This is reflected in the correct identification of the ML algorithm, despite the misidentification of the other attributes. So, as predicted, the *problem* seems to be the most important attribute in determining the optimal ML algorithm. However, the chosen ML algorithms DT and ANN cover the solution of clearly distinguishable tasks. Although, the general conditions gain importance for the decision when covering ML algorithms capable of solving similar or identical generalized problems. However, it should be mentioned that in 64% of the cases, the correct ML algorithm could be identified, confirming a supportive effect of the patterns.

6 CONCLUSION AND OUTLOOK

As the listed research papers in Table 2 indicate, several approaches to automate a task in the context of product development already exist. Nevertheless, these are hardly used by companies, which is partly due to a lack of expertise and knowledge regarding ML algorithms.

In this work patterns have been introduced using DT and ANN as examples, characterized by well-defined attributes, and associated specified element modules, based on the analysis of existing approaches. It has been shown that these schemes can be used to describe a problem and based on that, an optimal fitting ML algorithm can be determined.

However, the contributed study indicates that further research is needed to investigate the importance of the defined attributes that describe the general conditions for a feasible use of the ML algorithms. Therefore, patterns have to be created for ML algorithms suitable for similar generalized problems under different general conditions. Furthermore, the formulation of the element modules of the attribute *problem* must be adjusted based on the difficulties of the subjects in identifying the right one for the given problem.

It should be noted that the construction of the patterns is based solely on ML algorithms, that are used in the context of product development. The reason is to identify the specific problems for which the ML algorithms can be applied successfully. First, this leads to the disadvantage of not capturing other

https://doi.org/10.1017/pds.2023.37 Published online by Cambridge University Press

successful application scenarios for the ML algorithms outside of the product development space. Second, this arises the risk of covering errors regarding the application of the ML algorithms. Due to the multiple amounts of research papers analysed to characterize a ML algorithm and define a pattern, the probability of not noticing and covering such errors is minor.

Regardless of this, the manual analysis approach increases the risk of subjectivity during the process of deriving the characteristics of the problems and hinders the scaling of the process towards the analysis of additional research papers.

Therefore, future work will not only focus on completing the patterns into a pattern language by analysing further ML algorithms used in the context of product development. In addition, an automated approach for analysing the research papers needs to be developed, to guarantee consistency and objectivity during the process. Further, the proposed structure of the patterns can be proven to be sufficient.

REFERENCES

Bae, J.K. and Kim, J. (2011), "Product development with data mining techniques: A case on design of digital camera", *Expert Systems with Applications*, Vol. 38 No. 8, pp. 9274–9280. http://doi.org/10.1016/j.eswa.2011.01.030.

Bang, H. and Selva, D. (2016), "iFEED: Interactive Feature Extraction for Engineering Design", in American Society of Mechanical Engineers Digital Collection. http://doi.org/10.1115/DETC2016-60077.

Baxter, D., Gao, J., Case, K., Harding, J., Young, B., Cochrane, S. and Dani, S. (2008), "A framework to integrate design knowledge reuse and requirements management in engineering design", *Robotics and Computer-Integrated Manufacturing*, Vol. 24 No. 4, pp. 585–593. http://doi.org/10.1016/j.rcim.2007.07.010.

Bertoni, A. (2018), "Role and Challenges of Data-Driven Design in the Product Innovation Process", *IFAC-PapersOnLine*, Vol. 51 No. 11, pp. 1107–1112. http://doi.org/10.1016/j.ifacol.2018.08.455.

Borchers, J.O. (2001), "A pattern approach to interaction design", *AI & Society*, Vol. 15 No. 4, pp. 359–376. http://doi.org/10.1007/BF01206115.

Chakrabarti, A., Shea, K., Stone, R., Cagan, J., Campbell, M., Hernandez, N.V. and Wood, K.L. (2011), "Computer-Based Design Synthesis Research: An Overview", *Journal of Computing and Information Science in Engineering*, Vol. 11 No. 2. http://doi.org/10.1115/1.3593409.

Chen, H.-Y. and Chang, H.-C. (2016), "Consumers' perception-oriented product form design using multiple regression analysis and backpropagation neural network", *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, Vol. 30 No. 1, pp. 64–77. http://doi.org/10.1017/S0890060415000165.

Dunjko, V. and Briegel, H.J. (2018), "Machine learning & artificial intelligence in the quantum domain: a review of recent progress", *Reports on progress in physics. Physical Society (Great Britain)*, Vol. 81 No. 7, p. 74001. http://doi.org/10.1088/1361-6633/aab406.

Entner, D., Prante, T., Vosgien, T., Zăvoianu, A.-C., Saminger-Platz, S., Schwarz, M. and Fink, K. (2019), "Potential identification and industrial evaluation of an integrated design automation workflow", *Journal of Engineering, Design and Technology*, Vol. 17 No. 6, pp. 1085–1109. http://doi.org/10.1108/JEDT-06-2018-0096.

Feldhusen, J. and Bungert, F. (2007), "Pattern Languages: An approach to manage archetypal engineering knowledge", *Guidelines for a Decision Support Method Adapted to NPD Processes*.

Gerschütz, B., Schleich, B. and Wartzack, S. (2021), "A semantic web approach for structuring data-driven methods in the product development process", paper presented at DFX 2021 32nd SYMPOSIUM DESIGN FOR X, available at: https://www.designsociety.org/publication/43575/a+semantic+web+approach+for+structuring+data-driven+methods+in+the+product+development+process. http://doi.org/10.35199/dfx2021.15.

Giering, O. (2022), "Künstliche Intelligenz und Arbeit: Betrachtungen zwischen Prognose und betrieblicher Realität", *Zeitschrift für Arbeitswissenschaft*, Vol. 76 No. 1, pp. 50–64. http://doi.org/10.1007/s41449-021-00289-0.

Hashimi, H., Hafez, A. and Mathkour, H. (2015), "Selection criteria for text mining approaches", *Computers in Human Behavior*, Vol. 51, pp. 729–733. http://doi.org/10.1016/j.chb.2014.10.062.

Hein, A.M. and Condat, H. (2018), "Can Machines Design? An Artificial General Intelligence Approach", in Springer, Cham, pp. 87–99. http://doi.org/10.1007/978-3-319-97676-1_9.

Hoyle, C., Chen, W., Wang, N. and Gomez-levi, G. (2009), "Understanding Heterogeneity of Human Preferences for Engineering Design", *DS 58-5: Proceedings of ICED 09, the 17th International Conference on Engineering Design, Vol. 5, Design Methods and Tools (pt. 1), Palo Alto, CA, USA, 24.-27.08.*2009, pp. 229–240.

Hsiao, S.-W. and Huang, H. (2002), "A neural network based approach for product form design", *Design Studies*, Vol. 23 No. 1, pp. 67–84. http://doi.org/10.1016/s0142-694x(01)00015-1.

Jetter, A. (2005), *Produktplanung im Fuzzy Front End: Handlungsunterstützungssystem auf der Basis von Fuzzy Cognitive Maps, Forschungs-/Entwicklungs-/Innovations-Management*, Gabler Edition Wissenschaft, Deutscher Universitätsverlag, Wiesbaden. http://doi.org/10.1007/978-3-322-82157-7.

Kutschenreiter-Praszkiewicz, I. (2013), "Application of neural network in QFD matrix", *Journal of Intelligent Manufacturing*, Vol. 24 No. 2, pp. 397–404. http://doi.org/10.1007/s10845-011-0604-7.

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T. and Hoffmann, M. (2014), "Industry 4.0", *Business & Information Systems Engineering*, Vol. 6 No. 4, pp. 239–242. http://doi.org/10.1007/s12599-014-0334-4.

Lickert, H., Wewer, A., Dittmann, S., Bilge, P. and Dietrich, F. (2021), "Selection of Suitable Machine Learning Algorithms for Classification Tasks in Reverse Logistics", *Procedia CIRP*, Vol. 96, pp. 272–277. http://doi.org/10.1016/j.procir.2021.01.086.

Lundborg, M., Märkel, C. and Wagner, K. (2019), "Künstliche Intelligenz im Mittelstand. Relevanz, Anwendungen, Transfer", Eine Erhebung der Mittelstand-Digital Begleitforschung, Vol. 2019.

Management Association, I.R. (2022), *Research anthology on artificial neural network applications*, IGI Global, Hershey, Pennsylvania (701 E. Chocolate Avenue, Hershey, Pennsylvania, 17033, USA). http://doi.org/10.4018/978-1-6684-2408-7.

Metternich, J., Biegel, T., Bretones Cassoli, B. and Hoffmann, F. (2021), *Künstliche Intelligenz zur Umsetzung von Industrie 4.0 im Mittelstand: Expertise des Forschungsbeirats der Plattform Industrie 4.0, München*.

Rashid, T. and Langenau, F. (2017), *Neuronale Netze selbst programmieren: Ein verständlicher Einstieg mit Python, Animals*, O'Reilly Verlag, Heidelberg.

Reed, K. and Gillies, D. (2016), "Automatic derivation of design schemata and subsequent generation of designs", *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, Vol. 30 No. 4, pp. 367–378. http://doi.org/10.1017/S0890060416000354.

Riesener, M., Doelle, C., Mendl-Heinisch, M. and Klumpen, N. (2020), "Identification of evaluation criteria for algorithms used within the context of product development", *Procedia CIRP*, Vol. 91, pp. 508–515. http://doi.org/10.1016/j.procir.2020.02.207.

Rigger, E. and Vosgien, T. (2018), "Design automation state of practice - potential and opportunities", paper presented at DESIGN 2018 - 15th International Design Conference, available at: https://www.designsociety.org/publication/40462/design+automation+state+of+practice+potential+and+opportunities. http://doi.org/10.21278/idc.2018.0537.

Seo, K.-K., Min, S.-H. and Yoo, H.-W. (2005), "Artificial Neural Network Based Life Cycle Assessment Model for Product Concepts Using Product Classification Method", in Hutchison, D., Kanade, T., Kittler, J. and et al. (Eds.), *Computational Science and Its Applications – ICCSA* 2005, *Lecture notes in computer science*, Vol. 3483, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 458–466. http://doi.org/10.1007/11424925_49.

Shabestari, S.S. and Bender, B. (2017), "Enhanced integrated sensitivity analysis in model based QFD method", *DS 87-4 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 4: Design Methods and Tools, Vancouver, Canada*, *21-25.08.*2017, pp. 317–326.

Sousa, I., Wallace, D. and Eisenhard, J.L. (2000), "Approximate Life-Cycle Assessment of Product Concepts Using Learning Systems", *Journal of Industrial Ecology*, Vol. 4 No. 4, pp. 61–81. http://doi.org/10.1162/10881980052541954.

Souza, J., Matwin, S. and Japkowicz, N. (2002), *Evaluating data mining models: a pattern language*.

Tang, C.Y., Fung, K.Y., Lee, E.W., Ho, G., Siu, K.W. and Mou, W.L. (2013), "Product form design using customer perception evaluation by a combined superellipse fitting and ANN approach", *Advanced Engineering Informatics*, Vol. 27 No. 3, pp. 386–394. http://doi.org/10.1016/j.aei.2013.03.006.

Tseng, I., Cagan, J. and Kotovsky, K. (2012), "Concurrent Optimization of Computationally Learned Stylistic Form and Functional Goals", *Journal of Mechanical Design*, Vol. 134 No. 11. http://doi.org/10.1115/1.4007304.

Tucker, C.S. and Kim, H.M. (2009), "Data-Driven Decision Tree Classification for Product Portfolio Design Optimization", *Journal of Computing and Information Science in Engineering*, Vol. 9 No. 4. http://doi.org/10.1115/1.3243634.

Wang, Y. and Zhang, J. (2017), "Bridging the semantic gap in customer needs elicitation: a machine learning perspective", *DS 87-4 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 4: Design Methods and Tools, Vancouver, Canada*, *21-25.08.*2017, pp. 643–652.

Waring, J., Lindvall, C. and Umeton, R. (2020), "Automated machine learning: Review of the state-of-the-art and opportunities for healthcare", *Artificial Intelligence in Medicine*, Vol. 104, p. 101822. http://doi.org/10.1016/j.artmed.2020.101822.

Ye, N. (2013), *Data mining: Theories, algorithms, and examples*, *Human Factors and Ergonomics Series*, 1st edition, CRC Press an imprint of Taylor and Francis, Boca Raton, FL. http://doi.org/10.1201/b15288.

Zhan, Y., Tan, K.H. and Huo, B. (2019), "Bridging customer knowledge to innovative product development: a data mining approach", *International Journal of Production Research*, Vol. 57 No. 20, pp. 6335–6350. http://doi.org/10.1080/00207543.2019.1566662.

Zhang, Y. (Ed.) (2010), *New Advances in Machine Learning*, InTech. http://doi.org/10.5772/225.