A Critique of: *A Survey on Application of Machine Learning in the Early Phases of Product Development*

Kennard Mah | CID: 01956128

1. Summary

*A Survey on the Application of Machine Learning in the Early Phases of Product Developmen*t reviews the current state of machine learning (ML) in the early stages of product development—namely the requirements, function modelling, and principal concept design—by conducting a systematic literature review. This critique details the use of ML in the denoted stages whilst defining bias and limitations specific to each case; at length, concluding with an analysis of future applications of the mentioned techniques.

2. Definition of Early Phase Product Development

The paper refers to the design process paradox to weigh the benefits of making changes during the beginning of the design process as opposed to subsequent phases since early phase changes are the least expensive [2]. To better define this process, it is divided into three stages: requirements, function modelling, and concept design. To mine the highest quality data regarding the current state of ML applications, certified search engines were selected, and publications were filtered by their. Furthermore, the publication date was set to 1990-2019 and set of keywords used for machine learning and product development were defined to refine the articles. In total, 143 relevant articles were identified, with 40 belonging to the early phases of product development [1].

Requirements: This stage is the starting point where design objectives are specified in the form of 'requirements.' The requirements that define the underlying criteria to shape the final product are processed through four stages: identification, organization, analysis specification, and documentation. In identification, the objective is to mine end-user requirements from online textual content (customer reviews and comments) and classify if the data provides valuable insights through multiple learning processes (supervised and unsupervised.) The objective for the organization stage is to further categorize the identified requirements to extract critical to quality (CTQ) keywords from the customer reviews. The relations between CTQ keywords were modelled, and general rules were clustered based on their respective CTQ. The use of Bayesian networks, where requirements were classified into functional and non-functional classes, proved to be an accurate method on refining the requirements [3]. While the naïve Bayes method surpassed other clustering methods in terms of accuracy, the performance comparison of different clustering methods was not studied in this report. In the next stage, the characteristics of the requirements are analyzed using classification methods to determine if the requirement will be implemented in the final product. Finally, the specification stage is aimed at mapping the requirements to product specifications for function modelling. Another possible application where ML can be used is to approximate the relationship between market demand and design specifications, determining the traits of a profitable product design.

Function Modelling: When developing new products, it is important to build a functional structure that is synthesized from requirements, linking customer requirements to the desired functions. The modelling process involves mapping the basic functionalities and capabilities of the desired products and studying their relations. Neural networks methods can be used to optimize the complexity metrics of a function model concerning its market value [4].

Concept Design: The current state of automatic concept design generation using ML is limited and the potential for ML models to display creativity remains a topic of discussion. The report details a case of using a reinforced learning framework—the Gödel machine—the potential to generate concept attributes based on defined design attributes was observed. Furthermore, by setting predefined values in the design instances, a neural network can be trained to classify design attributes to their respective clusters.

3. Limitations of Proposed Techniques

			Unsupervised Learning		Supervised Learning							Reinforced learning
	Objective of ML Application	Number of case studies in the Literature	Clustering	Bayesian Network	Linear regression	Logistic regression	Näive Bayes classifier	Decision tree	Neural Networks	Support vector machine	Other	
Requirements	Requirements Identification	**3**	**1**							**1**	**1**	
	Requirements categorization	**12**	**9**	**2**			**1**					
	Requirements Analysis	**5**				**1**				**1**	**3**	
	Requirements specification	**8**			**1**			**1**	**5**	**1**		
	Attributes specification from variants	**15**		**1**	**3**			**1**	**7**	**3**		
Functional modelling	Function model characteristics modelling	**1**							**1**			
Concept Design	Concept Generation	**1**										1
	Categorizing and analysing the concepts	**5**	**2**					**1**	**2**			
	Modelling the concepts	**1**			**1**				**1**			
Sum		**51**	**12**	**3**	**4**	**1**	**1**	**3**	**16**	**6**	**4**	**1**

Table 1: Machine-learning applications in the early-phases of the product development

Table 1 displays the number of case studies that use unsupervised learning, supervised learning, and reinforced learning in the early phases of product development. In brief, many of the processes in requirements when gathering user insights and text mining is done using unsupervised learning methods; in terms of function modelling and concept design, however, there is a surplus of supervised learning, meaning that the process itself benefits from human oversight.

Requirements: The initial text mining process—also known as identification—shows a case of support vector method (SVM) to classify customer data and extract relevant insights from textual content from the internet. The article details, 'the more features integrated in the training the more accurate the results can be.' However, the predefined product features in this stage are fragmented and highly subjective. As no predefined list of features characterizing the customer suggestions or opinions exists, there is a lack of accuracy that can be found during the identification process. Likewise, during the organization stage, the clustering method used is guided by external opinions that limit the ability to generalize the definition of the classes. Another limitation of the classification process is the difficulty when repeating a process, especially in the case of unsupervised learning models. For example, the analysis specification stage of requirements involves defining end-user requirement attributes to identify relevant requirements; however, the labelled traits may not apply to different types of products.

Function Modelling: Limitations can originate from the increased complexity of a product and its functionality, raising challenges in the accuracy and interpretability of results generated by ML algorithms. There is also dependence on high-quality data to train the data which can be difficult to prepare for varying product types in the premature field of product development. Therefore, while function modelling can be useful in many domains, it is developed explicitly for the desired system type [4].

Concept Design: Despite attempts at automatic clustering, the report suggests that for selecting the right number of clusters, human clustering surpassed machine clustering when looking at similar designs. Hence, more research must be done on the selection criteria to automate the design generation process. As presented in Table 1, without human oversight in the design process, there can be a lack of innovation or intuition to evaluate the feasibility and practicality of generated concepts. However, these limitations can be reduced by suggesting an objective way to approach 'creativity' [5]. When doing so however, information used to inform the process may be impacted by bias regarding categorizing and analyzing specified concepts for varying cases.

4 Further Applications

While machine learning in the field of early-phase product development is premature and developing, with more practical applications in the field and a larger sample size of performance comparison between different methods, ML can supplement a process that was predominantly thought to be led by creative thinking and human oversight.

Requirements: Many requirement processes have shown practicality in terms of collecting and classifying large amounts of customer data to identify market demands. Table 1 suggests that these processes can be automated with unsupervised learning methods, allowing for better efficiency and accuracy in future applications.

Function Modelling: By quantifying the effort that goes into creating models in varying representations, product designers would benefit from the benchmarking process and be able to choose a functional modelling scheme based on accuracy [4]. Furthermore, automating this process can help to reduce costs while optimizing the performance of the product in future cases.

Concept Design: Though it is previously mentioned that design development and creative thinking are human concepts, ML displays the ability to create new concepts based on its understanding of 'good' and 'bad' design features. While there are limitations regarding intuition, originality or flexibility, ML algorithms can be used to support and amplify creativity by producing design alternatives. Ultimately, the need for human intervention depends on the specific context of product types and desired outcomes during the concept design process. If the goal is to generate novel concepts, unsupervised learning algorithms are capable of proposing outputs that are distinct and original from its training data. For example, in terms of music generation, a ML algorithm that is trained with large dataset of music can generate new tracks that are not based on existing songs [6]. However, the use of human intervention can refine this process if the intent is to steer the direction of concept design towards the desired output.

It is crucial to note that although ML has the potential to efficiently gather insights that drive product development while significantly reducing time and cost in the early phases of product development, it should be used in conjunction with human supervision to overcome the limitations that ML may face. Moreover, unsupervised mistakes in the early phases may birth unintended problems in the later processes of product development that become significantly more expensive to repair. In addition, although supervised learning results in an expenditure of time and costs, it can provide the necessary context for optimizing the classification and clustering processes and ensure accuracy for the results.

References

[1] Shabestari, S. S., Herzog, M. and Bender, B. (2019) "A Survey on the Applications of Machine Learning in the Early Phases of Product Development," Proceedings of the Design Society: International Conference on Engineering Design. Cambridge University Press, 1(1), pp. 2437–2446.
[2] Ullman, D.G. (2010), "the mechanical design process." McGraw-Hill Higher Education, Boston.
[3] Shakeri Hossein Abad, Z., Karras, O., Ghazi, P., Glinz, M., Ruhe, G. and Schneider, K. (2017), "What Works Better? A Study of Classifying Requirements." 25th IEEE International Conference on Requirements Engineering, IEEE, Lisbon, Portugal, 4-8 Sep.
[4] Gill, A.S., Summers, J.D., and Turner, C.J. (2017), "Comparing function structures and pruned function structures for market price prediction." *An approach to benchmarking representation inferencing value*, Artificial intelligence for Engineering Design, Analysis and Manufacturing, Vol. 31 No. 4, pp. 550–566.
[5] Moruzzi, C. (2020) "Learning through creativity: how creativity can help machine learning achieving deeper understanding", *Rivista Italiana di Filosofia del Linguaggio*, 14(2). doi: 10.4396/AISB201904.
[6] Hernandez-Olivan, C., Beltrán, J.R. (2023). "Music Composition with Deep Learning: A Review." In: Biswas, A., Wennekes, E., Wieczorkowska, A., Laskar, R.H. (eds) Advances in Speech and Music Technology. Signals and Communication Technology. Springer, Cham. https://doi.org/10.1007/978-3-031-18444-4_2

Word Count: 1437 (excluding references)