

Improving AI-based modelling for mass customisation strategy: Insights from automotive industry

Justyna Patalas-Maliszewska^{1*}, Pająk Grzegorz¹, Katarzyna Kowalczyńska²

¹ Institute of Mechanical Engineering, University of Zielona Góra, ul. Licealna 9, Zielona Góra, Poland,

² The Doctoral School of Exact and Technical Sciences and Institute of Mechanical Engineering, University of Zielona Góra, ul. Licealna 9, Zielona Góra, Poland

* Corresponding author's e-mail: J.Patalas-Maliszewska@iim.uz.zgora.pl

ABSTRACT

The potential of the mass customisation (MC) strategy is continuously expanding through the implementation of Industry 4.0/5.0 technologies. In this study, a novel approach to enhancing the potential of Artificial Intelligence-based techniques for MC level prediction was presented. Firstly, the empirical research was conducted among more than 150 European manufacturing companies in the automotive sector concerning factors influencing the increase in the level of MC, as well as the Industry 4.0 or 5.0 technologies applied. Next, the collected survey data was pre-processed to eliminate infrequent responses and subsequently analysed using the k-means clustering method for labelling purposes. The resulting dataset was then used to fine-tune parameters and train a classifier based on a multilayer perceptron (MLP), aimed at identifying the MC level in automotive enterprises. The developed classifier achieved average accuracies of 0.94 and 0.93 for training and testing, respectively. The proposed approach was validated through real-life case studies from the automotive industry, demonstrating the adaptability of our concept and its potential for diverse industrial applications.

Keywords: production strategy, mass customisation strategy, automotive industry, clustering, artificial neural networks.

INTRODUCTION

The concept of Industry 5.0 (I5.0) expands the assumptions of the Industry 4.0 (I4.0) concept and enables automation and digitisation of manufacturing by integration of I4.0 technologies: Cyber-physical systems (CPS), internet of things (IoT), artificial intelligence (AI), and big data analytics to support decision-making [1]. In the concept of I5.0 production is moving towards adapting to variations in product [2]. The mass customisation (MC) is a production strategy designed to meet individual customer requirements while benefiting from large-scale manufacturing. It is treated as a strategic approach that seeks to combine the advantages of custom manufacturing – particularly the ability to offer individualised products – with the best characteristics

of mass production, including cost efficiency, timely delivery, and high quality performance [3].

The application of the I4.0/I5.0 technologies as AI for monitoring and forecasting customer needs in alignment with the MC strategy remained a highly relevant and widely discussed topic in the literature. AI-based solutions enable processing data to improve prediction performance [4]. The approach presented in [5] applied k-means clustering, self-organising maps, fuzzy k-means clustering, and decision trees, to evaluate as well as extract both conscious and subconscious customer needs. It proposed a model for predicting customer preferences and utilised k-means clustering with nine variables: functionality, purpose, frequency of use, usage location, style, ease of use, form factor, price range and type of use. The concept of using a cloud-enabled customer

profile to predict preferred product functionalities and appearances by leveraging Microsoft AI as a service was developed in [6]. Moreover, several approaches have been proposed for assessing the maturity level of MC. For instance, three MC maturity levels – low, medium and high – based on factors such as: product platform development, product modularisation, IT-based product configuration, part standardisation, group technology, form postponement, process modularity, and concurrent product-process-supply chain engineering were presented in [7].

The shift to MC poses significant challenges for automotive production companies and here are still new perspectives for works in statistical experimental analysis [8]. For predictive models to be effective and practically useful in industry applications, it is essential to ensure the appropriate quality of data. Therefore, as a first step in the presented research, empirical studies were conducted to create a database regarding the MC level in the automotive production companies. Thus, firstly, the MC level in European manufacturing enterprises was analysed, using companies from the Polish automotive sector as a case study. Data was collected from European manufacturing companies in Poland's automotive sector. Next, cluster techniques were pre-processed and applied to analyse the acquired survey data and to find optimal clustering results. The results enable the design of the data-driven approach for optimising AI modelling to support the enhancement of the MC level.

Therefore, the main objective of this paper was to develop a novel approach that combines clustering and AI-based classification to identify, validate, and predict MC strategy levels. The following research questions (RQ) are thus formulated:

- RQ1: How do manufacturing companies in the automotive industry approach the MC level, and what characteristics determine it?
- RQ2: What types of data-driven techniques can be applied to optimise the potential of an AI-based model for enhancing the MC level?
- RQ3: How can the use of clustering algorithms in real -world applications contribute to enhancing the MC level in the automotive industry?

The main contributions of this work are as follows:

- It utilises survey data collected from over 150 European manufacturing companies.

- It clusters the MC characteristics and identifies the optimal number of groups of MC.
- It identifies the optimal hyperparameter configuration for the proposed AI-based model design.
- It validates the novel approach through real-life case studies.
- It highlights future research directions in AI-based modelling for MC strategy.

The remainder of this paper is structured as follows. Section 2 presents an in-depth review of both AI-based solutions applied to MC and clustering techniques and illustrates the research methodology. Next, section 3 proposes the concept of clustering algorithms for the MC characteristics and develops the optimal number of groups of MC. Section 4 presents the classifier based on a MLP, for determining the MC levels. Section 5 discusses the research results and illustrates the real-life case studies where recently developed clustering techniques are applied. Finally, this study presents possible future research directions within its scope.

BACKGROUND

In the manufacturing sector, MC is achieved through the use of I4.0/I5.0 technologies, mainly flexible processes and production, automation, robotics, and AI, the development of e-commerce, 3D printing, the use of robots, cyber-physical systems, and simulation [9] for flexible production and ensuring business efficiency. For its implementation, particularly in the automotive and electronics sectors in the area of customised products, attention is drawn to the need to use a high level of modular design [10], product configurators [11], and the digital twin's utilisation in the product design phase [12]. Next, in [13] it was stated that MC is linked to technological development.

For AI-based models it is necessary to properly prepare the input and output data. The effectiveness of forecasting models depends on the availability and quality of data. Therefore, the critical focus of the conducted research in the context of improving AI-based modelling for MC strategy is clustering the acquired data of the automotive industry to optimise AI modelling. The clustering techniques are treated as the as appropriate tools for pre-processing acquired

data [14]. A prominent example is the k-means algorithm, which defines clusters based on their centres of gravity and is widely used in pattern recognition. While also examining alternative algorithms that assign data objects to the nearest cluster based on their distance from the cluster centroid, k-medoids/partitioning around medoids (PAM) was considered, as it is more robust to noise and outliers. Unlike k-means, which uses the mean value as the cluster centre, PAM designates the most centrally located data point in each cluster – known as the medoid – as the cluster representative. Next, clustering for large applications (CLARA) is an extension of PAM and is particularly suitable for large datasets. Instead of searching for medoids across the entire dataset, CLARA draws multiple small samples and applies PAM to each of them to identify candidate medoids. After several iterations, the clustering solution with the best performance across all samples is selected as the final result. Clustering large applications based on RANdomized search (CLARANS) replaces the sample-based approach with a randomised search strategy. It allows flexible selection of the number of neighbouring solutions examined at each search step, which can result in faster convergence. Fuzzy c-means (FCM) clustering allows data points to belong to multiple clusters with varying degrees of membership. Widely used density-based clustering methods are the density-based clustering method (DBSCAN), which defines clusters as contiguous regions of high density. Graph-based clustering approach utilises graph theory concepts and properties to find clusters in the data.

However, a research gap remains in the area of pre-processing the data of the automotive industry to optimise AI modelling solutions for forecasting the level of MC in enterprises. Table 1 summarises findings from the literature on AI-based tools that support the enhancement of MC in manufacturing enterprises. Specifically, Table 1 presents recent articles covering (1) the industry branch and/or region analysed, (2) aspects related to enhancing the MC level, (3) data, (4) AI-based tools employed, (4) the effectiveness of the proposed solutions, and (5) challenges identified.

A literature review (Table 1) reveals a research gap in developing the novelty approach to enhancing the level of MC in the automotive industry through the application of clustering techniques and AI-based methods to empirically collected

survey data. Thus, no datasets concerning the factors influencing the MC level and the applied I4.0/I5.0 technologies enabling production under such a strategy were found in the available open databases. Moreover, as shown in Table 1, none of the reviewed studies evaluated the effectiveness of different clustering techniques applied to empirically collected data from the automotive industry for optimising AI modelling. The main challenge lies in acquiring real-world company data for MC enhancement, next processing high-dimensional survey data and identifying the optimal number of groups for MC level classification by developing a systematic and repeatable data labelling model.

A DECISION SUPPORT MODEL FOR THE MC STRATEGY

Input data

Firstly, to bridge the research gap and to address RQ1, a questionnaire was developed. The empirical research on the MC level in the automotive industry was conducted between January 10th 2023 and August 31st 2023. Data were collected from 153 European manufacturing companies in Poland's automotive sector through e-mail surveys and direct interviews. The research was conducted during pre-arranged meetings to minimise disruptions and allow respondents to reflect on their responses. The respondents were board members or senior managers. These employees were decision-makers or have theoretical and practical knowledge of previously made strategic decisions, the implementation of new technologies, and the effects of these implementations affecting the entire enterprise and its individual areas, such as sales, information management, or production planning.

The research sample can be considered representative as the number of respondents exceeded the minimum required sample size of 145 companies – calculated with a 95% confidence level, a proportion of 0.5, and an assumed maximum margin of error of 8%, based on the data from the Central Statistical Office indicating 3.954 companies operating in the automotive industry in 2022.

For MC level classification, data from 100 companies that reported producing products tailored to individual customer requirements were used. These respondents answered a series of

Table 1. Comparative overview of this study on AI-based Tools Supporting MC

Industry realising MC	Problem	Data	Applied AI-based tools supporting MC enhancement	Challenges or limitations	Ref.
Customised packaging- candy wrapping, no country data	Dynamic reconfiguration of the manufacturing system	Data from one company	Machine learning, deep learning, reinforcement learning, multiagent systems, convolutional neural network	Effective information interaction; dynamic reconstruction of manufacturing resources	[15]
Clothing production, no country data	Identification of customer preferences	conclusions from the scientist's experiences	Machine learning	Postponing the decision to address the market uncertainty	[16]
Fixture and jig manufacturing industry, SMEs, no country data	Data-driven production planning, process exploration	images of 2798 fixture parts from 53 categories, covering various bolts, plates, and bars; event log consists of 375 events with an average of five events per case and a mean duration per event of 47 min	Convolutional Neural Networks, k-means clustering	Availability of cost-effective information systems providing relevant data; systems integration, limited scalability of deep learning-based feature extraction and clustering	[17]
Automated manufacturing, reconfigurable manufacturing, no country data	Manufacturing systems need to be readily reconfigured to keep up with the mass customisation trend	data prepared for modelling purposes for two examples	Gated graph neural networks	Testing required on production lines	[18]
Dynamically interconnected assembly systems, no country data	Online scheduling problems	a large data set for training and validation (10944 scenarios) was generated using DOE approach	Artificial neural networks, supervised machine learning	Generalisability of the learned strategies	[19]
Italy, France, Belgium: smart factories	Responding to market changes by reconfiguring	public cloud computing platform	Artificial neural networks	Shifting in modern maintenance systems	[20]
Implant manufacturing, no country data	Product design	data from the Korea Institute of Science and Technology Information, subjects obtained from the Osteoarthritis Initiative	Convolutional Neural Networks	Eliminates the need for manual templating	[21]
Manufacturing, no country data	Fluctuations in demand and high competitive pressure; scheduling complexity	data from 129 reviewed papers (most of the magazines Int. Journal of Production Research, Procedia CIRP and Journal of Intelligent Manufacturing)	Graph Convolutional Neural Networks, Feed-Forward Neural Networks	Scalability and reproducibility	[22]
Traditional manufacturing sectors, SMEs, no country data	Product design, multi-objective machine scheduling	-	Generative Adversarial networks	Necessity to collect large volumes of data	[23]
Car body production (no country data)	Control of a flexible matrix-structured production system	Data modelled and simulated for one case study	Reinforcement Learning, Graph Neural Networks	Flow management of products and optimal utilisation of production cells	[24]

questions, based on which the companies' MC levels were assessed (Appendix 1).

Next, in line with the goal of producing an objective description and dataset suitable for an AI-based model, the authors proposed the following methodology (Figure 1).

Figure 1 illustrates the process of developing of classifier hyperparameters using ANNs for pre-processing the acquired surveyed data, with the goal of producing an objective description and dataset suitable for an AI-based model. The process begins with a preliminary analysis (Step 1.1) aimed at eliminating survey questions

that do not contribute meaningfully to the determination of the MC level. In the next step (1.2), the respondents' answers are uniformly encoded so that each possible response is represented as a distinct variable. Subsequently, variable selection is performed (Step 1.3) to eliminate the variables associated with the responses that occur infrequently to serve as reliable indicators of the MC level. To create a labelled dataset for training an AI-based model, the data described by the selected variables were than group into clusters representing distinct MC levels (Step 1.4). To ensure an objective division based solely on

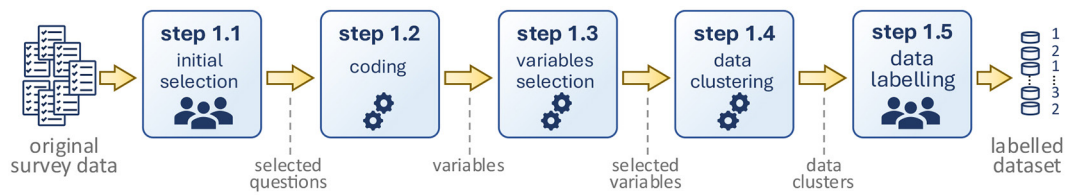


Figure 1. The methodology for elaborating dataset

features describing individual enterprises, the k-means method was proposed. In the final step (1.5), the resulting clusters were subjected to expert evaluation to assign the appropriate MC level. This process produces a labelled dataset in the next stage to train and test a classifier that identifies the MC level in any enterprise.

During the initial selection (Step 1.1, Figure 1), Questions 5 (Appendix 1) and 11 (Appendix 1) were excluded as irrelevant for determining the level of the MC strategy. Consequently, further analysis was based on data obtained from Questions 1–4 and 6–10. The questionnaire included both single-choice (Questions 1, 2, 3, 4, 9) and multiple-choice questions (Questions 6, 7, 8, 10), which required appropriate processing for subsequent analysis. Single-choice responses were encoded using the One-Hot method, while multiple-choice responses were coded as binary variables with values of 0 or 1. Additionally, the cases where respondents did not provide an answer (“undefined”) selected “other” - which introduced new, inconsistent categories potentially

distorting the results - were removed (Step 1.2, Figure 1). As a result, each surveyed enterprise was described using 59 variables representing all possible responses. The distribution of responses to the questions selected for further analyses is shown in Figure 2.

As shown in Figure 2, the distribution of responses is not uniform; for some questions, certain options were selected sporadically. Given the relatively limited size of the dataset, including such infrequent data could negatively impact the accuracy of classifying a company’s MC strategy level. To address this issue (Step 1.3, Figure 1), a response frequency-based criterion was adopted. Specifically, the mean number of responses was calculated for each question, according to the following formula:

$$\bar{N}_i = \frac{\sum_{j=1}^{m_i} N_{ij}}{m_i} \quad (1)$$

where: $\bar{N}_i$ is a mean count of responses of i – th question, m_i denotes the number of

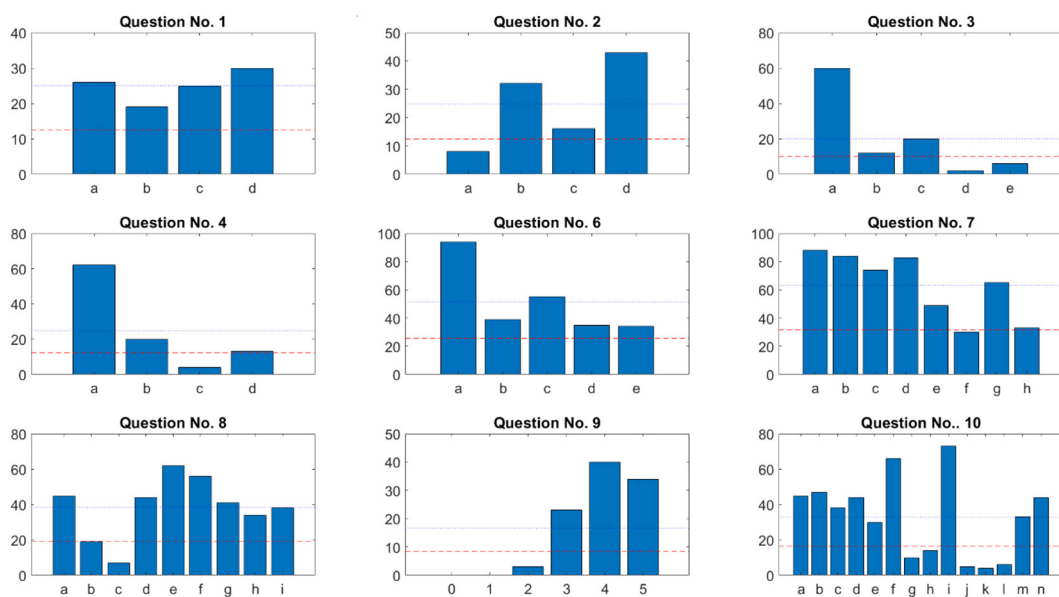


Figure 2. Distribution of responses to questions selected for analysis (numbering and labelling of responses as specified in Appendix 1)

possible responses for i – th question, and N_{ij} means the count of j – th responses for i – th question.

Responses that did not meet the criterion:

$$m_i \geq 0.5\bar{N}_i \quad (2)$$

were considered rare and excluded from further analyses. The results of this verification process are presented graphically in Figure 2. The blue dotted line indicates the average number of responses for each question, while the red dashed line represents the threshold defined by criterion (2). As can be seen, this method led to the rejection of 15 variables corresponding to infrequently chosen responses. Consequently, the classification of the MC strategy level presented in this section was based on 44 variables (Appendix 2). This approach appears effective in eliminating random results caused by insufficient data frequency. In future studies, conducted on a larger dataset, it will be possible to repeat the analysis and re-evaluate the significance of the survey responses.

Data clustering process

A key step in the proposed method is selecting an appropriate technique for identifying companies representing individual MC levels, as the resulting clusters will be used for supervised training of an ANN-based classifier (Step 1.4, Figure 1). For this purpose, the performance of hierarchical clustering, DBSCAN, and k-means is analysed [25].

The hierarchical clustering algorithm treats each element as a separate cluster and then iteratively combines the most similar pairs until a dendrogram encompassing all the data is obtained. The key to this method is the choice of distance metric and merging criterion, which significantly influence the final result. The undoubted advantage of this approach is the lack of the need to pre-define the number of clusters and the repeatability of the obtained results for the adopted distance metric and the merging criterion. Next, the DBSCAN method employs a different approach to identify similar elements using the concept of density. The algorithm groups elements that have many neighbours within a certain distance (creating a high-density region), while elements located in low-density regions are treated as noise and rejected. Initial conceptual analysis

indicated that, given the sample size and lack of clear density-based structures, the DBSCAN method would either combine distinct MC profiles or classify a significant number of companies as noise, which is undesirable for assessing MC levels. The k-means method is one of the most commonly used segmentation methods for identifying clusters in a set of objects, including survey data, and was ultimately selected as the primary method. The approach based on repeated clustering and analysis of multiple solution variants with different numbers of clusters allowed reducing these drawbacks and obtaining stable results. The detailed analysis presented in the next section confirms the validity of the selected clustering strategy.

According to the adopted research methodology, the acquired data were clustered (Step 1.4, Figure 1). To this end, the k-means method was applied, enabling estimation of the optimal number of clusters corresponding to the MC level of the examined companies.

For this purpose, a series of tests was conducted to assess the possibility of dividing the dataset into $k = 2, 3, 4, 5, 6, 7, 8$, or 9 clusters. Based on the results of the experiments presented in the authors' earlier work [26], the Gap Value measure was used as the evaluation criterion:

$$Gap_n(k) = \frac{1}{B} \sum_{b=1}^B \log(W_{k,b}^*) - \log(W_k) \quad (3)$$

where: W_k is the pooled within-cluster dispersion measurement, defined as follows:

$$W_k = \sum_{r=1}^k \frac{1}{2n_r} D_r \quad (4)$$

where: B is the number of generated random datasets, $W_{k,b}^*$ denotes the pooled within-cluster dispersion measurement for b – th random dataset, n is the sample size, k is the number of clusters, n_r is the number of data points in cluster r , and D_r represents the sum of squared Euclidean distances between all points in cluster r .

All calculations were performed using the Matlab 2023b environment. Due to the relatively small size of the sample tested and the random nature of the dispersion measure $W_{k,b}^*$ used by the evaluation criterion (3), the clustering test was repeated 100 times. The results presented are based on the average values of the indicators

obtained. A significant outcome was observed for $k = 3, 4, 5$: the division into 4 clusters occurred in approximately 50% of cases, while the division into 3 and 5 clusters occurred in about 20% of cases each. Other values of k resulted in less than 5% of cases (with 0% for $k = 2$ and 9). In the next step, the dataset was clustered using the k-means method for $k = 2, 3, 4$. Since the results depend on the initial random placement of cluster centroids, the algorithm was repeated 10 times, with each run consisting of 1.000 repetitions (i.e., specifying the number of times clustering was performed with different initial cluster centroid positions). The clustering results are presented in Table 2 and reflect the average values obtained in all experiments. The “Distance” column shows the minimum and maximum squared Euclidean distances between cluster elements and their respective centroids. The “Probability” column shows the lowest and highest probability of an element belonging to a given cluster, calculated on the basis of the outcomes of repeated tests. It should be emphasised that the numbers assigned to the cluster in each case are arbitrary and, at this stage, cannot be interpreted as corresponding to specific MC levels.

As shown in Table 2, only the division with $k = 3$ produced repeatable results, with all objects consistently assigned to the same cluster in each repetition. For $k = 4$ and $k = 5$, the method identified two stable clusters (no. 1 and 4 for $k = 4$ and no. 2 and 5 for $k = 5$), the composition of which remained unchanged across repetitions. However, the probability of assignment to the remaining clusters decreased to 50–70%. It is also worth noting that the degree of cluster compactness – understood as the minimum-to-maximum distance between objects and centroids – is similar for $k = 3$ but clearly varied in the cases of $k = 4$ and 5. The stable clusters with repeatable content were characterised by greater compactness. This

may be related to the sample size, but this results suggest potential difficulties in dividing the dataset into more than 3 clusters.

To further analyse the clustering results for $k = 3, 4$, and 5, a comparison of the centroid locations determined by the k-means method was also performed. Specifically, the distances between the centroids of the individual clusters were calculated. The results of the comparison for the clustering with $k = 3$ vs. $k = 4$ and $k = 3$ vs. $k = 5$ are presented in Tables 3a and 3b, respectively.

As shown, in each case, the k-means method consistently identified 3 clusters grouping objects with similar features – the centroids for these clusters are located very close to each other across the compared settings (highlighted in blue in the tables). In both $k = 4$ and $k = 5$ cases, a new cluster was formed with a centroid clearly separated from the three previously established ones (see the last columns in the tables). This suggests that further subdivision beyond three clusters may not be justified. Specifically, Table 3 shows that for $k = 5$ the centroid of the next cluster (the first column in the table) is located close to centroids no. 1 and 3, identified for the $k = 3$ solution, indicating that this additional cluster groups objects already well represented by the existing classes. These findings support the earlier conclusions and indicate that in the analysed dataset, 3 classes can definitely be distinguished. The presence of a fourth class warrants further analysis and seems to merit investigation. Therefore, in the AI-based model discussed in the next section, only the divisions into 3 and 4 clusters were considered.

For the division into 3 and 4 clusters, the following criteria were used to identify the characteristics of each cluster: the percentage of goods produced using the MC strategy, the areas of MC implementation, the outcomes of MC application, the technologies employed in MC processes, and the sources of information on customer

Table 2. Averaged clustering results

Cluster no.	Number of clusters					
			$k = 4$		$k = 5$	
	Distance (min-max)	Probability (min-max)	Distance (min-max)	Probability (min-max)	Distance (min-max)	Probability (min-max)
1	6.1 – 11.0	1.0 – 1.0	2.7 – 8.7	1.0 – 1.0	4.9 – 10.1	0.7 – 1.0
2	3.2 – 9.1	1.0 – 1.0	5.4 – 10.6	0.5 – 1.0	2.7 – 8.7	1.0 – 1.0
3	4.5 – 10.6	1.0 – 1.0	4.7 – 10.4	0.6 – 1.0	5.0 – 10.7	0.5 – 1.0
4	–	–	3.0 – 6.6	1.0 – 1.0	4.7 – 8.2	0.5 – 1.0
5	–	–	–	–	3.0 – 6.6	1.0 – 1.0

Table 3. Centroid distance matrix for 3-Cluster vs. 4-Cluster solutions and vs. 5-Cluster solutions

4-Cluster Solutions					5-Cluster Solutions					
Centroid	1	2	3	4	Centroid	1	2	3	4	5
1	6.79	0.18	2.26	7.34	1	1.88	6.84	0.54	3.10	7.34
2	0.14	5.32	10.12	4.86	2	9.70	0.14	5.14	11.02	4.86
3	11.07	3.49	0.28	5.19	3	2.18	11.11	3.98	0.35	5.19

preferences. Each cluster was characterised based on the most frequently indicated level of customised goods sold, as well as the shared features of the 3 enterprises closest to the cluster centroid. Finally, in the three-cluster solution, Cluster 1 represented a good, Cluster 2 a low, and Cluster 3 the best level of MC implementation. In the four-cluster solution, the clusters correspond to the following MC level: Cluster 1 – low, Cluster 2 – good, Cluster 3 – best, and Cluster 4 – very good. For the purposes of subsequent studies, the results in both cases were ordered according to decreasing MC levels. The labelling of the classes (data labelling) corresponding to each cluster used in the subsequent stage of the research is shown in Table 4 (Step 1.5, Figure 1).

The adopted MC levels (Table 4) are consistent with the adopted definition of MC maturity levels, in which five levels of MC strategy implementation are aligned with the Technology Readiness Level (TRL) framework [27]. Accordingly, MC level low corresponds to TRL 1–2; MC level good to TRL 3–4; MC level very good to TRL 5–7, and MC level best corresponds to TRL 8–9.

AI-based model design

To address RQ2 specifically, to optimise the potential of the AI-based model for enhancing the MC level, it is essential to clearly determine whether three or four MC level classes are defined. Therefore, the survey data selected, processed and labelled (section 3) were used to develop a universal classifier capable of identifying the MC

level in enterprises. The MLP was applied due to its capabilities of solving any classification problem, irrespective of the feature distribution or data complexity [26]. So, a two-layer classifier with a general structure shown in Figure 3 is proposed in the present study.

The choice of this classifier architecture reflects a trade-off between the model's approximation ability and generalisation capability, which is particularly important when working with limited survey datasets. Preliminary experimental results showed that increasing the number of layers did not improve the generalisation performance; instead, it resulted in rapid overfitting. A systematic architecture selection process was carried out to identify the optimal classifier architecture, as shown in Figure 4.

Firstly, repeated tests are conducted, employing multilayer perceptrons with a varying number of neurons, different activation functions, and diverse regularisation methods during training. Next, the final evaluation of each classifier's performance is based on averaged results. This approach enabled the identification of the optimal ANN structure for each parameter. In subsequent steps, the selected classifiers are re-trained using a validation set to prevent overfitting and then re-evaluated using multiple performance metrics. Based on these results, the optimal classifier structure can be chosen and applied to develop a universal tool for assessing the MC level in any enterprise.

Preliminary tests indicated that the model trained on the available dataset tended to overfit

Table 4. Assignment of MC classes to clusters

Division into 3 clusters			Division into 4 clusters		
Cluster	Class no.	MC level	Cluster	Class no.	MC level
3	1	Best	3	1	Best
1	2	Good	4	2	Very good
2	3	Low	2	3	Good
–	–	–	1	4	Low

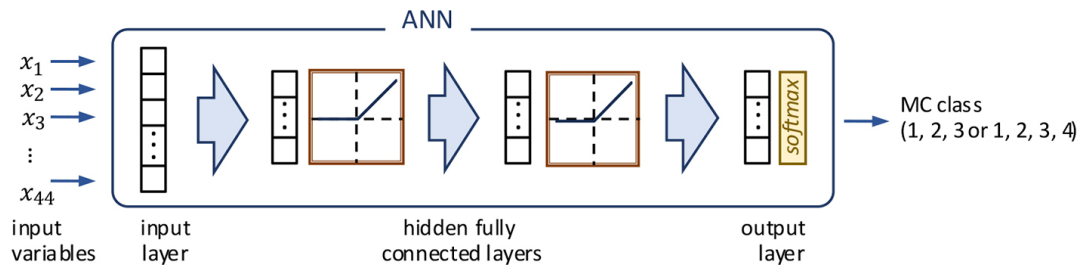


Figure 3. General architecture of the proposed classifier

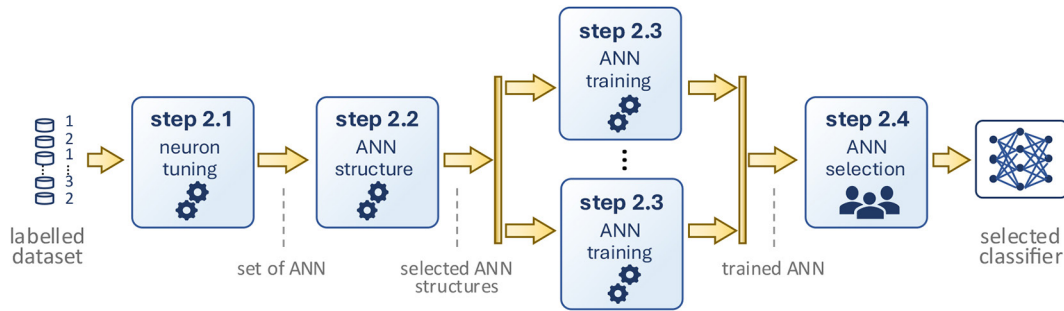


Figure 4. Tuning classifier hyperparameters

rapidly. To mitigate this undesirable effect, dropout regularisation was applied during training [28]. Dropout works by randomly disabling individual neurons, which helps the model learn more general and significant patterns in the training data while reducing the influence of noise and dataset-specific peculiarities. Additionally, L2 regularisation was applied [29], introducing a penalty for large weight values. This encourages the development of simpler models that generalise better to new, previously unseen data.

Due to the imbalance in the dataset – particularly evident when dividing the set into 4 classes – a class weighting technique was applied during training. This approach ensured that each class had an equal influence on the performance of the designed classifier. The weight estimation for each class was calculated based on the following relationship:

$$w_i = \frac{\sum_{j=1}^k n_j}{n_i k}, i = 1 \dots k \quad (5)$$

where: n_i is the number of samples in i -th class and k is the number of classes.

Separate tests were conducted for both $k = 3$ and $k = 4$ (corresponding to networks with 3- and 4-output neurons, respectively). Each series of tests involved the modification of the following network parameters:

- number of neurons in the hidden layers: ranging from 2 to 16 neurons per layer,
- type of activation function (ReLU and Leaky-ReLU with parameter $\alpha = 0.01$),
- dropout probabilities (0.3 and 0.5).

The decision to limit the number of neurons in each layer to 15 was based on preliminary studies conducted for several representative cases, in which extending the size up to 32 neurons did not yield any improvement in classification accuracy and, in fact, resulted in overfitting. The decision to select an activation function stems from the nature of the data being processed. ReLU is widely considered a function with high computational efficiency and good convergence capability. In certain cases (e.g., in a four-class classification scenario, where class imbalance and label uncertainty are more pronounced), the use of ReLU can lead to the phenomenon of “neural death”, which involves a loss of activity when neurons reach a state where the gradient is zero during the training phase. To counteract this, testing the LeakyReLU function, which demonstrates greater learning stability owing to its ability to maintain nonzero gradients for negative inputs, was proposed. This feature reduces the risk of inactive neurons and maintains the learning ability of the network throughout the training phase. To avoid network overfitting and

ensure robust generalisation, the implementation of a strong regularisation technique was proposed. The choice of coefficient values was based on the results of preliminary experiments, which showed that a dropout of 0.3 to 0.5 was optimal for preventing neuronal co-adaptation and reducing the risk of learning patterns specific to a given dataset.

Finally, for each case, 225 network configurations with varying numbers of neurons in hidden layers were tested, each evaluated with two different activation functions and two dropout rates (Step 2.1, Figure 2). To assess the performance of each model, the micro F1-score is proposed as the evaluation metric. This measure accounts for the balance between precision and recall across all classes, without bias towards classes with lower cardinality:

$$F1 = 2 \frac{\mathcal{P} \mathcal{R}}{\mathcal{P} + \mathcal{R}} \quad (6)$$

where: $\mathcal{P} = TP/(TP + FP)$ is the precision and $\mathcal{R} = TP/(TP + FN)$ is the recall of the model; TP is the number of correctly classified samples in all classes (true positives), while FP and FN represent the number of false positives and false negatives, respectively.

To minimise the influence of the training-test split on the performance evaluation of models with given hyperparameters, stratified 5-fold cross-validation was applied. Additionally, to account for the randomness in weight initialisation and to reduce the variance of the results, each test was repeated ten times, and the final model evaluation was based on the average micro F1-score value. During training, the number of epochs was set to 750, and the batch size to 16 in all cases.

Finally, the network with the optimal number of neurons in the hidden layers was selected based on the F1-score achieved during both the training and testing phases. The primary criterion was to obtain an F1-score of at least 0.9 in both phases. If this threshold was not met, networks with progressively lower F1-scores (≥ 0.89 , 0.88, 0.87, etc.) were considered. At this stage, a trade-off between the complexity and classifier performance was analysed. The cross-validation results indicated that increasing the number of neurons per layer above 9 in the three-class model and above 14 in the four-class model did not result in a statistically significant improvement in the test accuracy. Additionally, these higher complexity architectures showed significantly higher variance in the cross-validation results, suggesting that they were starting to model noise rather than patterns in the processed data. A summary of the network hyperparameters for each case is presented in Table 5.

For the final selection of hyperparameters (the last step of the procedure presented in Figure 4), each of the classifiers selected in the previous step was trained on the available data, divided into training, validation, and test sets in the proportion of 70:15:15. To prevent overfitting, early stopping was applied during training, while all other training parameters remained consistent with those used previously. To reduce the influence of randomness on the results, the training and testing process was repeated ten times, and network performance was evaluated based on the averaged outcomes. The final evaluation and model selection relied on multiple criteria, including classic accuracy, F1 score, and confusion matrices illustrating the network's performance during both training and testing phases.

Table 5. Summary of network hyperparameters

No.	Number of outputs (k)	Activation function	Dropout	Number of neurons		F1-score (train, test)
				I layer	II layer	
1	3	ReLU	0.3	9	4	0.995, 0.903
2	3	ReLU	0.5	13	9	0.996, 0.920
3	3	LeakyReLU	0.3	9	5	0.997, 0.904
4	3	LeakyReLU	0.5	12	9	0.997, 0.914
5	4	ReLU	0.3	13	13	0.999, 0.890
6	4	ReLU	0.5	13	14	0.987, 0.876
7	4	LeakyReLU	0.3	13	10	0.998, 0.892
8	4	LeakyReLU	0.5	13	14	0.987, 0.873

For the classifier dividing enterprises into three MC levels, the individual networks yielded similar results. The highest accuracy during both training and testing was observed for class 2 (medium MC level), while the accuracy for classes 1 and 3 was a few percentage points lower. For the classifier dividing into four classes, the highest training accuracy – exceeding 95% – was achieved for classes 1 and 4 (low and best MC level, respectively). Classes 2 and 3 (good and very good MC level) showed varied accuracy in the range of 85–95%, depending on the network parameters. During training, accuracy for classes 1 and 4 was consistently similar, with classes 3 and 4 lagging by a few percentage points. Finally, based on the analysis of accuracy metrics and confusion matrices from both training and testing phases, the following classifier architectures were selected:

- For the division into three classes: number of neurons in hidden layers: 9 and 4; activation function – ReLU, dropout rate – 0.3,
- For the division into four classes: number of neurons in hidden layers: 13 and 14; activation function – LeakyReLU with $\alpha = 0.01$, dropout rate: 0.5

The first classifier achieved average training and testing accuracies of 0.95 and 0.93, respectively, while the second achieved 0.94 and 0.93. The averaged confusion matrices for both the training and testing stages are shown in Figures 5 and 6.

As can be seen, despite favourable average results, the classification accuracies for individual classes, which is particularly noticeable in the case of the second classifier and class 3. Therefore, the adoption of the 3 classes corresponding to 3 levels of MC in enterprises within the automotive industry was confirmed, respectively:

1 MC level: the best, 2 MC level: good, 3 MC level: low (Table 5). The three MC levels allow for more explicit and specific characterisation of the features of enterprises representing each class, taking into account the areas of MC application, I4.0 technologies recommended for implementation at each level, the effects of MC implementation on sales, production planning, and the production workspace, as well as the methods of data processing related to each MC level. This approach enables the practical use of the tool as decision-making support for managers.

Methodological robustness and limitations

The sample size of 153 companies is representative of the Polish automotive sector, but it poses a challenge for traditional deep learning methods. To ensure the reliability of the results and generalisability of the neural classifier, rigorous data preparation and network training strategies were adopted in this study, minimising the risk of overfitting.

During the initial data preparation stage, responses with small sample sizes were eliminated, which could have led to the identification of non-existent patterns in the data. Additionally, to mitigate the problem of an unbalanced dataset (particularly important for a four-class classifier), a class weighting technique was used during network training. The risk of overfitting was minimised by combining different techniques. First, stratified 5-fold cross-validation was used to evaluate the model performance across different data subsets while maintaining class balance. To further eliminate the influence of stochastic initialisation, each experiment was repeated 10 times, and the results were evaluated based

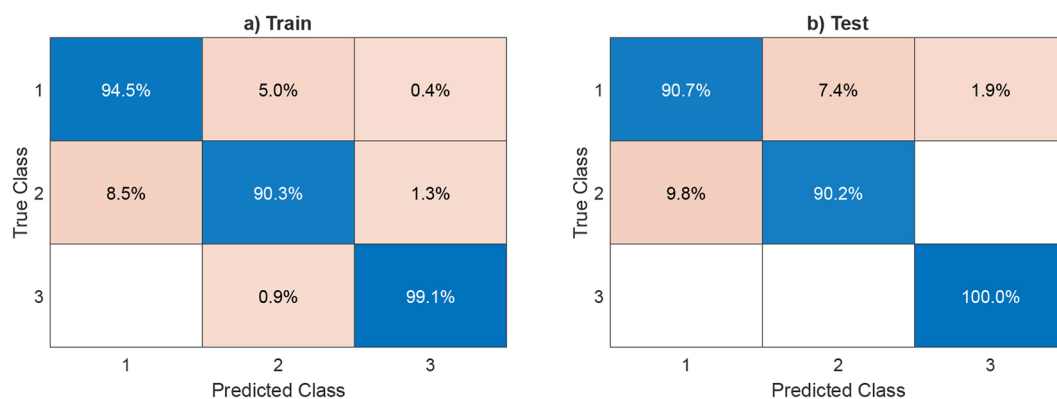


Figure 5. Averaged confusion matrices for the classifier with three classes

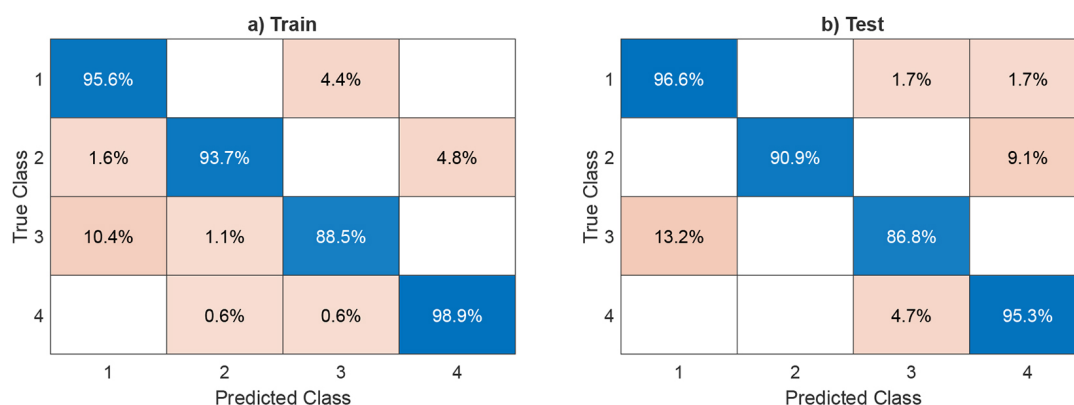


Figure 6. Averaged confusion matrices for the classifier with four-classes

on the average. Second, a combination of two regularisation methods was used: dropout with coefficients of 0.3 and 0.5 and L2 weight regularisation to prevent excessive model complexity. Furthermore, the training process utilised an early stopping method based on monitoring validation losses, preventing the network from overfitting to the subjective nuances in the survey data.

Although the survey data can introduce subjectivity, the high consistency between the training and testing accuracy achieved (especially evident in the three-class model) suggests that the classifier effectively captured the underlying patterns characterising MC levels, rather than noise. This level of robustness indicates that the proposed model can serve as a reliable diagnostic tool, although future research should aim to expand the dataset to validate these findings in different industrial geographies.

However, the main goal of this study was to develop a tool with high robustness and generalisability for assessing the level of MC. However, the presented classifier should not be viewed as a typical “black box.” In this case, partial interpretability was achieved in the unsupervised learning stage, where k-means clustering yielded centroid-based MC profiles that could be examined and evaluated by the experts. Furthermore, the input variables were not chosen randomly but were verified by experts and subjected to a rigorous selection procedure (steps 1.1–1.3). This integration of domain-specific knowledge moves the proposed model from a pure black box to what [30] termed a “gray box” approach, in which the input data have a clear physical meaning. It should also be emphasised that, as noted in [31], the depth or structural complexity of a neural network cannot be equated with deeper semantic understanding

or explainability. Neural network layers represent mathematical transformations, rather than explicit knowledge representations. In this context, the proposed neural network acts as a mapping function between the observed features and MC levels, which is sufficient for the purposes of this study.

FINAL REMARKS

The main finding of this stage of the conducted study is that, by adopting the k-means method and ANN, the proposed approach demonstrates that monitoring and predicting the MC level in the automotive industry is a complex issue and requires a framework for understanding and assessing the MC level.

Firstly, it encompasses a wide range of factors such as goods sold under the MC strategy, areas of MC implementation, effects of MC implementation, technologies employed in MC implementation, and sources of information regarding customer preferences.

Next, the presented study analysed the potential for developing an AI-based approach for predicting and monitoring the level of MC. The resulting model enables the identification of the specific characteristics of MC corresponding to each MC level:

- The best MC level: gaining knowledge of customer preferences is identified as a key factor. MC is implemented in the following areas: product configuration, product design, product cost estimation, forecasting the feasibility of customer-specific orders, production process visualisation, production process control, and automatic planning of production processes. These implementations are supported by tools

such as product simulation, material simulation, production process simulation, systems integration, automated machines, human-machine interaction, and cybersecurity. This enables the observation of several effects, including increased workspace flexibility, increased production flexibility, reduced waste, and decentralised decision-making.

- The good MC level: typically sells 76–100% of customised goods. In their case, knowledge of customer preferences and needs is primarily obtained through direct communication with customers. Robotics is employed, and MC is implemented through product design, product cost estimation, forecasting the feasibility of customer-specific orders, and production process control. Companies at this level report increased production flexibility and reduced waste.
- The low MC level is associated with the sale of only 1–25% of customised products. In such cases, customer preferences are usually identified through interviews with customers. By configuring products and using automated machines, these companies report annual sales increases of up to 10%.

The findings summarises the Table 6. Next, the importance of pre-processing, labelling, clustering and validation of the acquired survey data as a main basis of developing a universal tool for assessing the MC level in any enterprise is highlighted.

However, the conducted research has some limitations. The first concerns the limited dataset, which was obtained through survey research conducted in companies. A key challenge was the relatively small sample size of enterprises obtained through surveys. Hence, special attention was given to minimising random effects caused by the selection of training data and initial weights. Despite this, the adopted classification model, based on a three-level MC scale, enabled clear categorisation of the surveyed enterprises. There is some evidence that introducing a fourth MC level could be viable; however, the associated error margins are currently too high for practical application.

This suggests a need for continued research, either with a larger empirical sample or synthetic input data to improve the robustness and accuracy of the model. While this approach has the advantage of being grounded in expert knowledge, the data are restricted to enterprises operating in the automotive industry and located in the European region, specifically in Poland. future research

could sample for further geographies and industries to acquire additional data for inclusion in the developed AI-based model. Next, while the current model focuses on achieving high classification accuracy, future work will focus on implementing Explainable AI (XAI) frameworks. Implementation of methods such as SHAP (SHapley Additive ExPlanations) or LIME [30] is planned, which will allow for the identification of specific features (e.g., levels of IT integration or product modularity) that act as catalysts or barriers to achieving higher levels. This approach will allow obtaining a model that can be used not only as a classifier but also as a tool to support strategic decision-making.

However, the novelty of the conducted research is also demonstrated by the acquisition of data from real-world manufacturing enterprises and by the application of clustering techniques in AI for pre-processing the collected data. Additionally, the criteria used in the questionnaire were derived from a literature review. In future studies, it would be beneficial to also incorporate feedback from companies directly implementing MC strategies.

Applications in production and managerial insights

While addressing RQ3, guidance to managers on determining the MC level was provided. The defined criteria (Table 7) can be treated as the proposal for providing preliminary audit due to improving MC level. To illustrate the proposed approach to determine the MC level in-depth interviews were analysed and conducted in two real-word companies from automotive industry (Table 7).

Based on the criteria (Table 6) the first company achieved the 1 level of MC, and the second company the MC level: “good”. It means, that the manager receives a card, that presents the level of MC implementation in the five distinguished areas (Table 7). Specifically, it indicates which areas (Table 7) require improvement and which are at a good level. In the analysed company 1, no modifications are required in the production of personalised products. Regarding the analysed company 2, the following recommendations are formulated, divided by areas of characteristics of MC:

- Considering Area 1: Further analysis is necessary, along with the identification and selection of products that could potentially be personalised.
- Considering Area 2: Further analysis is necessary in the context of implementing tools

Table 6. The specific characteristics of MC corresponding to each of the three MC levels

MC Level/ Characteristics of MC	The Best	Good	Low
Products marketed within the MC	76–100%	26–100%	1–25%
Domains of MC implementation	- Product: configuration/ design/cost estimation. - The feasibility of orders tailored to customer needs. - Visualisation/ control/ planning of production processes.	- Product: design/ cost estimation. - The feasibility of orders tailored to customer needs. - Control of production processes.	Product: configuration.
Effects of MC implementation	- Increasing the flexibility of the workspace. - Increasing production flexibility. - Reducing waste. - Decentralising decision-making.	- Increasing production flexibility. - Reducing waste.	Annual sales growth of up to 10%.
I4.0/I5.0 technologies in MC	- Simulation tools. - An integrated information system. - Automated machines. - Human-Machine interaction solutions. - Cybersecurity.	Robotics.	Automated machines.
Sources of information about customer preferences	- Conversation with the customer. - Survey on customer preferences/ satisfaction. - Scientific literature analysis. - Statistical data.	Conversation with the customer.	Conversation with the customer.

Table 7. Results of the preliminary audit on determining the MC level

Characteristics of MC/Real-word companies from automotive industry	Company 1: A small-sized company manufactures PVC vehicle ramps	Company 2: A Medium-sized company manufactures valves, couplings and cylinders
Area 1: Products marketed within the MC	Products in MC account for 76–100% of sales.	Customised goods constitute the majority (51–75%) of the sold goods.
Area 2: Domains of MC implementation	In the last 5 years: product configuration, design, cost estimation, forecasting of feasibility of customer-specific orders, visualisation and control of production processes, automatic production planning	Product configuration, design, cost estimation, forecasting of feasibility of customer-specific orders, monitoring of production processes.
Area 3: Effects of MC implementation	Increasing workspace flexibility and production flexibility, reducing waste, and electricity consumption.	Reduction of waste, more efficient analysis of big data, decentralisation of decision-making, increase in sales of more than 25%.
Area 4: I4.0/I5.0 technologies in MC	Efficient analysis of big data, robotics, material simulation, production process simulation, automated machines, e.g. 3D printers, cyber-security.	Analysis of big data, robotics, product simulation, Integrated system integration, cloud computing, 3D printers, human-machine interaction.
Area 5: Sources of information about customer preferences	Conversation with the customer and survey on customer preferences.	Conversation with the customer.

enabling the visualisation and control of production processes, as well as automatic production planning.

- Considering Area 3: It is recommended to maintain a decentralised decision-making strategy.
- Considering Area 4: It is recommended to develop tools for data and information analysis.
- Considering Area 5: It is recommended to implement solutions that allow real-time monitoring of customer conversation data in order to maintain a high level of MC.

This primary audit serves as a support mechanism during creating a detailed company profile according to MC strategy. It enables development of tailored growth scenarios in MC within the automotive industry and validates the proposed framework through the real-life case studies.

CONCLUSIONS

In this study, AI-based classifier for assessing the level of MC in the automotive industry is successfully developed and tested. By combining unsupervised learning (k-means clustering), used for objective data labelling, with supervised learning (ANNs), the study provides a robust framework for the comparative analysis of manufacturing enterprises. The developed model achieved high classification accuracy (93% for the three-class variant), demonstrating that survey-based data – when subjected to rigorous pre-processing and regularisation – can serve as a reliable foundation for intelligent diagnostic tools.

The theoretical contribution of this study lies in the systematisation of the MC assessment process by shifting from subjective self-assessment toward data-driven classification using appropriate algorithms. From a managerial perspective, the proposed model offers a strategic tool for identifying a company's current MC level, thereby enabling more informed decision-making regarding digital transformation and investments in I4.0/I5.0 technologies.

Although the proposed model demonstrates high predictive accuracy, its internal logic remains partially opaque. Future studies will also address the dynamic evolution of MC levels by developing predictive tools that not only classify current states but also suggest optimal improvement pathways. Additionally, the authors plan to complement self-assessment survey data with objective data from

manufacturing execution systems (MES) or ERP indicators. This integration would further reduce potential subjectivity and enable a more detailed analysis of the relationship between technical infrastructure and mass customisation capabilities.

By incorporating these research directions, the proposed AI-based approach may evolve into a comprehensive decision-support system, assisting enterprises in managing the complexity of modern, highly customised manufacturing environments.

Acknowledgements

This work was supported by a program of the Polish Ministry of Science under the title 'Regional Excellence Initiative', project no. RID/SP/0050/2024/1.

REFERENCES

1. Oztemel E., Gursev S. Literature review of industry 4.0 and related technologies. *Journal of Intelligent Manufacturing*. 2020; 31(1): 127–182. <https://doi.org/10.1007/s10845-018-1433-8>
2. Verna E., Puttero S., Genta G., Galetto, M. A Novel diagnostic tool for human-centric quality monitoring in Human-Robot collaboration manufacturing. *Journal of Manufacturing Science and Engineering*. 2023; 145(12). <https://doi.org/10.1115/1.4063284>
3. Trentin A., Sandrin E., Suzic S., Grosso C., Forza C. Reconciling product flexibility with cost, delivery, and quality: The importance of bundling mass customization practices. *Global Journal of Flexible Systems Management*. 2025: 1–25.
4. Pizoń J., Gola A. Vertical Integration Principles in the Age of the Industry 5.0 and Mass Personalization. In: Burduk, A., Batako, A.D.L., Machado, J., Wyczółkowski, R., Dostatni, E., Rojek, I. (eds) *Intelligent Systems in Production Engineering and Maintenance III. ISPEM 2023. Lecture Notes in Mechanical Engineering*. Springer, Cham. https://doi.org/10.1007/978-3-031-44282-7_26
5. Saldivar A.A.F., Goh C., Li Y., Yu H., Chen Y. Attribute identification and predictive customisation using fuzzy clustering and genetic search for Industry 4.0 environments. In 2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA), 79–86.
6. Aheleroff S., Philip R., Zhong R.Y., Xu X. The degree of mass personalisation under industry 4.0. *Procedia CIRP*. 2019; 81: 1394–1399.
7. Suzic N., Forza C. Development of mass customization implementation guidelines for small and medium

- enterprises (SMEs). *Production Planning & Control*. 2023; 34.6: 543–571.
8. Schindlerová V., Sýkora V., Stančková D., Mrázik J., Using the Design of Experimental Method for Problem-solving of Noisy Parts in the Automotive Industry. *Adv. Sci. Technol. Res. J.* 2024; 18(5): 65–74. <https://doi.org/10.12913/22998624/190060>.
9. Aron C., Sgarbossa F., Ballot E., Ivanov D. Cloud material handling systems: A cyber-physical system to enable dynamic resource allocation and digital interoperability. *Journal of Intelligent Manufacturing*. 2024; 35(8): 3815–3836.
10. Yang H., Jiang L. Modular design of new products and remanufacturing of used products under mass customization. *RAIRO-Operations Research*. 2024; 58(1): 103–128. <https://doi.org/10.1051/ro/2023151>
11. Wang Y., Zhao W., Wan W.X. Needs-based product configurator design for mass customization using hierarchical attention network. *IEEE Transactions on Automation Science and Engineering*. 2020; 18(1): 195–204.
12. Abanda F. H., Jian N., Adukpo S., Tuhaise V. V., Manjia M. B. Digital twin for product versus project lifecycles' development in manufacturing and construction industries. *Journal of Intelligent Manufacturing*. 2025; 36(2): 801–831.
13. Kim S., Lee K. The paradigm shift of mass customization research. *International Journal of Production Research*. 2023; 61(10): 3350–3376.
14. Singh J. Singh D. A comprehensive review of clustering techniques in artificial intelligence for knowledge discovery: Taxonomy, challenges, applications and future prospects. *Advanced Engineering Informatics*. 2024; 62, Part C: 102799 <https://doi.org/10.1016/j.aei.2024.102799>.
15. Wan J., Li X., Dai H.N., Kusiak A., Martinez-Garcia M., Li D. Artificial-intelligence-driven customized manufacturing factory: key technologies, applications, and challenges. *Proceedings of the IEEE*. 2020; 109(4): 377–398.
16. Piller F.T., Euchner J. Mass Customization in the Age of AI. *Research-Technology Management*. 2024; 67(4), 10–1080.
17. Mehdiyev N., Mayer L., Lahann J., Fettke P. Deep learning-based clustering of processes and their visual exploration: An industry 4.0 use case for small, medium-sized enterprises. *Expert Systems*. 2024; 41(2).
18. Dwivedi A., Khan N.R. Reconfiguration of Automated Manufacturing Systems Using Gated Graph Neural Networks. *The Scientific Temper*. 2022; 13(02): 317–326.
19. Göppert A., Mohring L. Schmitt R.H. Predicting performance indicators with ANNs for AI-based online scheduling in dynamically interconnected assembly systems. *Prod. Eng. Res. Devel.* 2021; 15:619–633. <https://doi.org/10.1007/s11740-021-01057-z>
20. Foresti R., Rossi S., Magnani M., Bianco C.G.L., Delmonte NSmart society and artificial intelligence: big data scheduling and the global standard method applied to smart maintenance. *Engineerin.* 2022; 6(7), 835–846.
21. Burge T. A., Jeffers J. R., Myant C.W. Development of an automated mass-customization pipeline for knee replacement surgery using biplanar X-rays. *Journal of Mechanical Design*. 2022; 144(2).
22. Panzer M., Bender B., Gronau N. Neural agent-based production planning and control: An architectural review. *Journal of Manufacturing Systems*. 2022; 65: 743–766.
23. Jimeno-Morenilla A., Azariadis P., Molina-Carmona R., Kyratzi S., Moulianitis V. Technology enablers for the implementation of Industry 4.0 to traditional manufacturing sectors: A review. *Computers in Industry*. 2021; 125: 103390.
24. Steinbacher L.M., Wegmann T., Freitag M. Production control with Reinforcement Learning for a matrix-structured production system. *International Journal of Production Research*. 2024; 1–23. <https://doi.org/10.1080/00207543.2024.2436126>
25. Jain A.K. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*. 2010; 31: 651–666.
26. Patalas-Maliszewska J., Pajak G., Kowalczywska K. Driven data analytics for enhanced mass customization in production. *IFAC-PapersOnline Proceedings*. 2025; 59(10): 2106–2111.
27. Patalas-Maliszewska, J., Kowalczywska, K., Pajak, G. Mass customization in practice: Strategic implementation and insights from Polish small and medium sized enterprises, *Advances in Production Engineering & Management*. 2025; 20, 1, 18–28, <https://doi.org/10.14743/apem2025.1.524>
28. Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R. (Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*. 2014; 15: 1929–1958.
29. Kalina J., Tumpach J., Holeňa M. On combining robustness and regularization in training multilayer perceptrons over small data. In 2022 International Joint Conference on Neural Networks (IJCNN): 1–8, IEEE.
30. Wu Z. F., Li J., Cai M. Y., Lin Y., Zhang W. J. On membership of black-box or white-box of artificial neural network models. 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), Hefei, China, pp. 1400–1404, <https://doi.org/10.1109/ICIEA.2016.7603804>
31. Zhang W. J., Yang G., Lin Y., Ji C., Gupta M. M. On Definition of Deep Learning. 2018 World Automation Congress (WAC), Stevenson, WA, USA, 1–5. <https://doi.org/10.23919/WAC.2018.8430387>