

## **Categorization of European Countries for Industry4.0 Inequality by Unsupervised Machine Learning Techniques**

**Huseyin Fidan**

*Burdur Mehmet Akif Ersoy University  
Istiklal Campus, Burdur, Turkiye  
E-mail. hfidan@mehmetakif.edu.tr*

<https://doi.org/10.5755/j01.ee.37.3.35976>

*Industry4.0 applications have become a strategic issue for countries because of giving competitive potential. In this context, studies are realized on ranking and grouping countries to determine the usage levels of Industry4.0 technologies. However, containing different variables, indexes and techniques in these studies presents different results. The lack of an agreed technique for grouping countries based on Industry4.0 technologies leads to method confusion on the subject. Clustering is one of the most suitable options for sorting or grouping a dataset. The purpose of the study is to categorize the countries according to usage of Industry4.0 technologies by using machine learning techniques. The clustering algorithms were applied to the ICT Usage in Enterprises data compiled from Eurostat. In the study, the most appropriate clustering algorithm was specified, the distance values within and between clusters were calculated, and the level differences in Industry4.0 technologies of the countries were determined.*

*Keywords: Industry4.0; Technology Management; Digital Diversity; Machine Learning; Clustering.*

### **Introduction**

Technological developments throughout the historical process have led to significant changes affecting social, cultural and economic life. The transition to conscious agriculture, the invention of the loom and internet are considered as turning points in human history. These turning points brought major changes in social and business life, and these developments were described as "revolution" due to their effects (Drucker, 1986). Some technologies in these revolutions have come to the fore as symbols such as plough and stem machine. In Information and Communication Technologies (ICT), which is considered the third revolution, the internet is the prominent technology. Today, a new revolution called Industry4.0 (I4.0) is mentioned, referring to systems that operate completely automatically and communicate with each other in production and service activities. Since this revolution utilizes all the possibilities of ICT, it would not be wrong to emphasize that ICT is its driving force. In this context, the Internet of Things (IoT), Cloud computing (CC), Big data (BD), 3D-printing (3D), Artificial Intelligence (AI) are considered crucial technologies within the scope of I4.0 (Wang *et al.*, 2015).

The I4.0 concept, introduced by the German government in 2011, is defined as production systems supported and shaped by ICT (Kagermann *et al.*, 2013). The concept, situated within the framework of smart production systems, smart factories (Kagerman *et al.*, 2013), and integration tools of product lifecycle (Dalenogare *et al.*, 2018), can be characterized as production systems working autonomously, communicating with each other through ICT integration. This integration process encompasses the entire life cycle of the product, starting from raw material procurement to final consumer evaluations (Wang *et al.*, 2016). Therefore, the concept spans both classical production processes and service processes (Mariani &

Borghi, 2019). In other words, I4.0 is a technology applicable not only in the production sector but also in service sectors such as tourism, logistics, and health. It establishes connections between these sectors. Similar to previous technological revolutions, countries aware of the advantages of I4.0 are making significant investments in this field. Developments in technologies within the scope of I4.0 indicate that countries avoiding these advancements may face serious losses in the long term.

The concept of digital divide, defined as differences in ICT ownership levels, is recognized as a new form of inequality that negatively impacts business processes (Fidan, 2016). Various factors, including social, cultural, technological, and educational aspects, underlie this type of inequality (Kraftova *et al.*, 2011). I4.0-based technologies that have developed in the last decade can be considered as factors that contribute to the widening of these inequalities. So, the digital divide originating from I4.0 technologies (I4.0-DD) is seen as a new inequality problem (Kuruczleki *et al.*, 2016). Differences in the adoption and utilization of I4.0, which significantly contributes to the economic and technological developments, lead to an increase of digital divide between countries. It is emphasized in all studies that countries giving less importance to the technologies developing in I4.0 will have disadvantages. According to Tutak and Brodny (2022), I4.0 technologies have features that provide significant competitive advantages to companies and countries. According to researchers, countries and companies that cannot catch up this revolution will experience significant losses in the long term. According to the 2019 report of the United Nations (UN), the I4.0 revolution has the power to create sustainable industrial value, also has the potential to negatively affect the technological development of countries by further deepening the digital divide (UN, 2019). In 2021 report of United Nations Industrial Development Organization

(UNIDO), attention is drawn to the difference in technological development levels and it is emphasized that the digitalization process of countries should be balanced (UNIDO, 2021). Countries those are aware of the problem aim to increase their digital competence levels in order to avoid the cost of falling behind I4.0 technologies. For this purpose, they provide significant support to technological developments.

The primary issue in digitalization process is to determine the technological level and differences (Senna *et al.*, 2023). Thus, the roadmap, action plan and policies can be specified according to the level of digital differences. For this reason, measuring the level of I4.0-DD is crucial for implementing effective precautions. Many studies have been realized to measure I4.0-DD differences between countries. All of these studies claim that their methods produce reliable results. However, the common view in the literature is that model and method-based some problems are observed in the studies.

The main problem in these studies is which method accurately measures the inequalities. In the studies, country rankings are established by developing measurement tools using statistical methods, scores and certain indexes such as Digital Opportunity Index (DOI), Digital Access Index (DAI), ICT Development Index (IDI), Digital Economy and Society Index (DESI). In addition, new indexes for measuring technological levels, called technological maturity (TM) and digital maturity (DM), have been widely used in recent years (Antony *et al.*, 2021; Tutak & Brodny, 2022; Senna *et al.*, 2023). In TM and DM studies, it is aimed to calculate a score value for the level indicator in indexes created with various variables (Hizam-Hanafiah *et al.*, 2020). However, it is frequently stated in the literature that there are deficiencies in measurement methods determining technological level by calculating score values (Thordsen & Bick, 2023). The use of different approaches, variables and weightings in these indexes leads to different scores and results (Fidan, 2016; Cagnet *et al.*, 2023). A country that ranks first in one index may rank third in another. These differences in country rankings create methodological confusion regarding which index provides superior accuracy.

Another challenge observed in studies is the inconsistent results of Multi-Criteria Decision Making (MCDM) methods frequently used in studies for rankings. Due to the different calculation techniques, MCDM methods create different rankings even in the same datasets (Larichev, 1999). The fact that confirmatory analyses are not included in studies including MCDM creates a contradiction as to which MCDM method provides a healthier ranking (Asadabadi *et al.*, 2019). Different determinations of the importance levels of the variables and weights affect the results (Kumar *et al.*, 2024). However, the fact that the technologies used within the scope of I4.0-DD are interrelated requires that the subject be evaluated holistically. For this reason, it would be a more appropriate approach to conduct analyses by grouping countries in determining the difference levels. Instead of determining a country reaching digital maturity level, determining the difference values within and between groups by grouping countries that have reached the digital maturity level will provide a broader perspective. Thus, it will be possible to develop more comprehensive and appropriate policies related for solutions.

Although all studies on I4.0-DD emphasize that there are differences between countries, another deficiency observed in the studies is that the level of differences is not determined. In I4.0-DD analyses, it should not be neglected which countries need support and to how much they should be encouraged. Since no approach has been found to address this deficiency, an applicable framework for supportive and encouraging policies cannot be presented in the studies. In cluster analyses, both intra-cluster and inter-cluster distances can be calculated (Han *et al.*, 2012). In this context, this study, which proposes cluster analysis for determining I4.0-DD levels, also aims to eliminate the deficiency regarding the calculation of differences.

These problems create methodological confusion in the literature in measuring I4.0-DD, and lead to the questioning of studies in terms of methodology and results. In order to solve these problems, grouping countries and analyzing by unsupervised machine learning techniques that have wider analysis possibilities will reveal more consistent, interpretable and verifiable results. Finding the most successful method by performing the accuracy analysis will put an end to the method confusion in the literature. So, the objective of this study is to find an alternative solution based on machine learning technique for the problems of measuring I4.0-DD.

Clustering is an unsupervised machine learning technique that is one of the most suitable methods for grouping items in a dataset (Han *et al.*, 2012). This method allows the structural analysis of clusters formed by determining related items. In this way, it can be discovered the cluster analysis produces healthy results or not. In addition, in cluster analysis, by finding the distance values within and between clusters, the levels of differences between countries can be determined. While the calculation methods and tools in cluster analysis are sufficient to overcome the shortcomings of the methods used in previous studies on the subject, it is surprising that they are not preferred much in the literature. This study claims that unsupervised machine learning is the most suitable method to overcome the deficiencies in the literature. So, aim of this study is to determine the I4.0-DD level between countries by using clustering analysis. Data on Cloud Computing (CC), Big Data (BD), 3D printing (3D), Internet of Things (IoT), Artificial Intelligence (AI), ICT-Employment, and ICT-Training usage in company processes were compiled from Eurostat's ICT Usage in Enterprises dataset for European countries. The novelty of this study lies in successfully determining I4.0-DD using unsupervised machine learning methods instead of traditional measuring methods such as indexes, scores, and MCDM. Consequently, clustering analysis was conducted to group countries according to their I4.0-DD levels. The performance of clustering was calculated by using internal validity indexes, namely the Silhouette Index (SI) and Dunn Index (DI). The results indicated that the most successful clustering algorithm was K-Means with 4 clusters. Distances between clusters were then calculated to determine the I4.0-DD level between countries. According to the results, Finland and Sweden were identified as the leading countries. In contrast, Cluster 3, which includes Turkiye, Lithuania, Romania, and Bulgaria, was found to lag behind. The I4.0-DD level difference between these two clusters was calculated to be 156 %.

### Industry 4.0 and Technologies

I4.0, introduced by the German National Academy of Science and Engineering in 2011, encompasses the concept of ICT-based technological innovation, laying the groundwork for the development of new technologies and organizational methods (Kloviene & Uosyte, 2019). Technologies such as IoT, AI, CC, BD, and 3D contribute to increased performance, high efficiency, and a competitive advantage in companies' fields of activity (Sauter et al., 2015). In other words, I4.0 holds significant potential to impact entire company processes. Built on the idea of automating company process controls and enabling devices to communicate with each other, I4.0 is a concept encompassing human-to-human (C2C), human-to-machine (C2M), and machine-to-machine (M2M) data communications (Cooper & James, 2009). In this context, it can be noted that I4.0 generally includes technologies covering data production, data transmission, and data analysis.

With I4.0, ICT-based technologies have gained vital importance in company processes (Kloviene & Uosyte, 2019). Countries that recognize this significance aim to enhance their ICT competencies by assessing their level of ownership of I4.0-related technologies used in business life. To support this process, companies that prioritize the employment of ICT experts provide their employees with ICT training to boost the skills of their workforce. Companies, seeking to avoid falling behind in this revolution, place particular emphasis on strategic I4.0 technologies.

### Internet of Things

IoT technology is founded on the acquisition of real-time data within manufacturing processes and its subsequent dissemination to various endpoints, including workers (Bi et al., 2014). Gubbi et al. (2013) outlined that IoT systems contain four units: data production, wireless sensor network (WSN), data storage-analysis, and application layer. An alternative perspective claims that an IoT system can be built by five units: Radio Frequency Identification (Rfid), WSN, middleware, CC, and software (Lee & Lee, 2015). Taking a broader approach, Patel & Patel (2016) claimed that IoT systems have four units as sensor, network, management, and application layer. It can be stated that IoT system consists of three layers: data production, data transmission and storage-analysis (Fidan & Yuksel, 2020). An IoT's architecture is given in Figure 1.

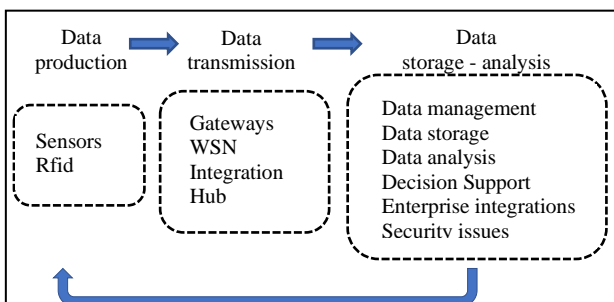


Figure 1. Architecture of IoT

Source: Fidan & Yuksel, 2020

Figure 1 shows the general operation of an IoT system. Data obtained from objects by sensors are brought together in the data transmission layer and made available for use in company processes. The storage-analysis layer contains data storage, analysis, security and data sharing tasks. In this layer, integrations are also made with data security, company decision support systems and applications used by the company such as Supply Chain Management (SCM), Enterprise Resource Planning (ERP), and Customer Relation Management (CRM).

### Artificial Intelligence

AI is defined as systems that perform processes such as learning, decision-making, control and regulation of a human intelligence by using relevant data (Peres et al., 2020). The application of these systems in industry is carried out by processing the data obtained from all activity processes of companies and is described as Industrial AI (Lee et al., 2015). AI, which has subfields such as machine learning, deep learning, natural language processing, and computer vision, enables the autonomous functioning of machines used in production processes (Ahmed et al., 2022). The industrial application areas of AI, which is also used in decision support units of IoT systems, are becoming widespread.

### Big Data

Data types in computer systems are evaluated in two groups as structured and unstructured. Data that can be designed as a table according to a certain number of rows and columns is called structured data. However, in current technological systems, there are also data such as pictures, songs, videos, and text that cannot be directly used in the tabulation method. Analysis of unstructured data, which has become the focus of attention especially with the spread of social media applications, is a challenging process because it requires additional processes and techniques. BD is defined as the analysis of high volume and unstructured data (Shi, 2022). There has been an increase in using BD for company processes in recent years. It is an important tool for companies' market analysis because it provides access to instant data (Xu et al., 2016). Some business processes such as CRM (Hallikainen et al., 2020), brand value (Shirdastian et al., 2019), demand forecasting (Seyedan & Mafakheri, 2020), which are vital analyses for companies, can be realized in a different way by using BD. These analyses performed with real market data provide significant contributions to companies.

### Cloud Computing

CC is an information access model for using internet-based computing resources without time and device restrictions (Benlian et al., 2018). CC, which allows widespread use of ICT resources, is becoming an important business model in ICT resource distribution (Sunyaev, 2020). In this context, it is described as a technology that will enable the reshaping of business models within the framework of collaboration (Xun, 2012). CC technology can transfer the resources of businesses to cloud providers through virtual servers and services. In this way, companies

have advantages such as low cost, effective accessibility and sustainability.

### **3D Printing**

It is a technology used to turn virtual objects created electronically through the layering method into 3D physical objects (Ngo *et al.*, 2018). Production can be made using different materials with the method called additive manufacturing (Jandyal *et al.*, 2022). Today, 3D printers can work with many materials such as plastic, resin, metal, ceramics, candy and chocolate. In this respect, it has found the opportunity to be used in many areas such as health, automotive and education. It provides significant production, cost and time advantages, especially in prototype, personalized and finished products (Ngo *et al.*, 2018).

### **Advantages of I4.0 for Business Processes**

Companies focus to provide high-quality products and cost minimization. Thus, they seek to enhance their market influence through strategic industrial planning. The fundamental objective in gaining market power lies in outperforming other companies in terms of competitive power. To achieve this goal, companies focus on innovating methods and strategies that align with technological advancements in production (Fidan & Yuksel, 2020). In this context, the main concern is that “Can I4.0 systems bring competitive advantages to business?”

It is widely accepted in the literature that technology provides competitive advantages for the companies. It is accepted that technology has great competitive potentials that can change sectors (Porter, 1985). So, I4.0 brings many opportunities for companies (Jernigan *et al.*, 2016). The sections below explain potential ways in which I4.0 systems could contribute to various business processes.

#### ***Automated Systems***

Studies carried out with the aim of ensuring inter-object communication and self-determining machines in company processes reveal different approaches to the idea of automation in all sectors. In this context, changes are observed in the self-functioning system structures designed according to the results of the data collected by people, analysis and decision processes. As the analysis, decision-making, and implementation processes increasingly align with the principles of Industry 4.0, the notion of automation enters an entirely new realm (Ahmed *et al.*, 2022). It is crucial to acknowledge that the dynamic system structure in automation stems from the comprehensive integration of data collection, analysis, decision-making, and guidance within the designed systems (Patel & Patel, 2016). For example, a machine that works in the production process but says that it will malfunction after a while, a production line that detects that the quality level of production outputs will be low and even lists the possible reasons for the lack of quality, a machine that detects the reason for an employee's decrease in productivity by monitoring social media posts and offers solutions. Employee monitoring systems make significant contributions to the production processes of the production line balancing system by taking into account the warehouse capacity according to product demands.

#### ***Real-time Monitoring***

Efficient and optimal utilization of production resources within company operations represents a significant concern in both engineering and economics. Attaining maximum production efficiency relies on the effective utilization of resources throughout the production process. It is fundamental to monitor the capital and labour inputs of companies comprehensively, spanning from the inception to the conclusion of the production cycle, for the overall success of the company. However, challenges such as resource limitations and technological gaps make the establishment and sustenance of monitoring systems costly and often unfeasible, particularly in service sectors (Fidan & Yuksel, 2020). The capabilities of I4.0 systems, empowered by advanced data collection and communication functionalities, offer a viable solution for the seamless implementation of real-time and automated monitoring systems.

#### ***Reducing Data Errors***

The basis of successful business analysis lies in acquiring adequate and high-quality data. Human-generated data is prone to errors, particularly in manual collection and recording during production processes, leading to a decline in data quality, especially when dealing with large datasets. Employing I4.0 technologies for data collection helps mitigate errors arising from human factors. The utilization of technologies like environmental sensors and Rfid in data collection leads to a reduction in error rates, consequently enhancing the overall quality of data analysis. This, in turn, results in more precise analysis outcomes. The efficacy of quality control processes is bolstered as potential errors from individual assessments, particularly in quality control procedures conducted by different individuals, are minimized.

#### ***Increasing Efficiency***

The assessment of productivity, a concept examined across a broad spectrum, lacks a standardized measurement tool due to its evaluation based on diverse parameters (Sauer mann, 2016). Productivity, typically defined as the ratio of outputs to inputs in either quantity or monetary terms, can be analysed separately with respect to labour and capital or comprehensively by considering all production factors collectively (Fisher, 1990). Nonetheless, the prevailing perspective is that enhancing the efficiency of labour, capital, methods, or processes correlates with increased productivity for a company or sector (Reynolds, 1998). Consequently, the measurement of worker or machine efficiency emerges as a crucial consideration for companies. In productivity analyses, I4.0 systems ensure that the entire process is monitored, data is collected completely and analyses are made clearly. In this way, it enables the accurate identification of employees or machines with low productivity.

#### ***Cost Reduction***

Although initial setup costs are high, the operating costs of company processes are reduced with I4.0 (Ghobakhloo, 2020). The uninterrupted operation of automated production systems and with minimal errors provides significant

reductions in companies' costs (Schroeder *et al.*, 2019). It reduces analysis error rates in data entry, access, analysis, control and decision processes and accordingly accelerates system evaluations (Ghobakhloo, 2020). In addition, real-time monitoring of production processes makes positive contributions to preventing work accidents, thus preventing costs that may occur due to malfunctions and work accidents (Dalenogare *et al.*, 2018).

### **Increasing Cooperation**

Due to the system structure, data collection and data storage processes are among the basic processes of I4.0 technologies. The system operation data obtained is the basic component for every analysis to be implemented within I4.0 (Lu, 2017). In this way, it is possible to share data between different units within the company and to create an environment for data exchange with different companies. Shared usage of data produced in different units within the company increases internal process efficiency (Salkin *et al.*, 2018). On the other hand, I4.0 technologies provide increased business connectivity globally. Especially IoT, CC and BD technologies expand companies' data and application sharing opportunities (Han & Trimi, 2022).

### **I4.0-DD**

Digital divide, defined as differences in access to information technologies (Fidan, 2016), can be also used to express differences for usage of I4.0 technologies. In this context, I4.0-DD will be formed between countries that cannot adequately adapt to I4.0 technologies. This problem will lead to inadequate use of the business collaboration opportunities, delay the implementation of I4.0 technologies in business processes, and reduce the effectiveness of global collaboration. The report prepared by the United Nations (UN) in 2019 emphasizes that I4.0 transformations will negatively affect less industrialized countries, development differences between countries will increase and these countries will be more disadvantaged in terms of competition (UN, 2019). In this context, determining the I4.0-DD level between countries becomes an important issue.

### **Related Studies**

The digital divide, defined as differences in using information technologies, is not a new issue. The concept, which has attracted the attention of researchers with the widespread information technologies, is expressed as a serious social and economic problem (Attewell, 2001; OECD, 2001). With entrance of new technological developments into our lives within the framework of I4.0 in the last decade, it is accepted that the concept that creates new inequalities is a global problem, especially for business (Kuruczleki *et al.*, 2016; Pacchini *et al.*, 2019; Branco *et al.*, 2023).

A large number of studies have been conducted to determine the differences in I4.0-DD levels between countries. When the literature is examined, it can be said that the studies generally focus on determining I4.0-DD levels. In studies using different approaches, it is observed that there is no consensus on the method. The first studies on the subject, which started in the 2000s, aimed to create a

metric. In the following years, composite indexes were developed as a measurement tool and used in analyses (Ali *et al.*, 2020). In some of these studies, index values based on various indicators are calculated (Tutak & Brodny, 2022), while in others, country rankings are determined with MCDM methods (Balkan & Akyüz, 2023). It is seen that I4.0 technologies are also included in studies focusing on determining the DM level, especially after 2016 (Elibal & Ozceylan, 2022).

Since the first years of the studies, there is serious method confusion in measuring digital gap level between individuals, companies and countries. Drawing attention to this method confusion, Fidan (2016) states that different measurement tools are applied such as DOI, DAI, IDI, Networked Readiness Index (NRI), Digital Evolution Index (DEI) and DESI. Since different components, variables and weights are used in these indexes, different results are obtained (Fidan, 2016). Stating that IDI has inadequacies in quantitative measurements due to the subjectivity of sub-index weights, Gerpott and Ahmedi (2015) emphasized the deficiency in interpreting socio-economic results. DESI, which is generally used to assess the digital development of European Union countries, covers 4 main components: connectivity, human capital, digital technology integration and digital public services (European Commission, 2021). Despite its widespread use, DESI has some methodological deficiencies too. Stiftung (2020) emphasizes that regional differences are not taken into account in the weighting of variables and argues that DESI has scope problems. According to Yakymova *et al.*, (2022), weighting and expert evaluations create methodological problems in index calculations. Researchers emphasize that this deficiency is the main problem of composite indexes such as DESI. Studies conducted to address the deficiencies in these indexes aim to improve the existing index or design a new one. Kuruczleki *et al.* (2016), who conducted one of the first studies to determine I4.0-DD levels, created a readiness index using eight indicators. Atik & Ünlu (2019) stating that Kuruczleki *et al.* (2016) neglected I4.0 technologies have expanded the study by including I4.0 technologies. According to Gerpott and Ahmedi (2015), who emphasized that the sub-index weightings of the IDI index used in determining the ICT levels of countries are subjective, IDI is insufficient in quantitative measurements. To eliminate this deficiency, they suggested using the Least Squares method and Structural Equation Model. Santos and Martinho (2020) remark that studies on determining technological levels do not have a theoretical infrastructure, and suggest that technology adaptation theory should be taken as the basis in I4.0-DD studies. Hajoary *et al.* (2023), who stated that the evaluation frameworks of the indexes used in the literature are narrow and do not include confirmatory analyzes, developed a model with a wider range of indicators. In a similar view, Senna *et al.* (2023) claimed that environmental factors are ignored and designed a composite index including technology, organization and environment. Stating that All Composite indexes are created with different indicator weights and that the simple averaging method is used in calculations, Mahdiloo *et al.* (2023) argue that this situation prevents sub-indicator interpretations and reduces the usefulness of composite

indexes. According to Cognet et al. (2023), a few models and indexes have a structure that allows quantitative comparisons. So, main source of the method complexity in literature is the inability to associate different models and methods. Researchers have recommended machine learning-based methods to solve the method confusion (Cognet *et al.*, 2023).

Another common method used to rank countries according to their digital technology competencies is MCDM. In some of these studies, MCDM methods are used as a hybrid with indexes. Samaranayake et al. (2017), who argued that the technologies highlighted in I4.0 should be ranked with MCDM and that comparisons should be made according to this ranking, applied the AHP method for the ranking of the highlighted technologies. Honti et al. (2020), who asserted that it would be more appropriate to use MCDM methods together with indexes, developed a composite index using the PROMETHEE method. Contieri et al. (2021) suggested the TOPSIS method, and Balkan and Akyuz (2023) suggested the PROMETHEE method to determine the levels of difference in the use of I4.0 technologies. Marti and Puertas (2023) ranked the countries with the TOPSIS method using the Global Innovation Index (GII) together with DESI. In a recent study, AHP and CoCoSo methods were applied in ranking the performance outputs originating from I4.0 technologies (Kumar *et al.*, 2024). It is seen that various MCDM methods are applied in the literature. However, the fact that accuracy analyses are not included in these studies leads to difficulties in choosing the appropriate method. In addition, different rankings are obtained with the MCDM methods. Reaching different results causes confusions and questions the MCDM methods.

In recent years, studies called Digital Maturity (DM) have been seen aiming to determine digital technology competencies (Thordsen & Bick, 2023). DM studies are on the same level as studies on the digital divide in terms of their research subject and generally aim to measure the digital competencies of companies based on some indicators (Pauli *et al.*, 2021). Especially after the developments of I4.0, the concept of DM has attracted great interest from researchers (Balkan & Akyüz, 2023), and studies have been conducted in different fields such as education (Kobets *et al.*, 2020), healthcare (Shaygan & Daim, 2023), agriculture (Scandurra *et al.*, 2023), manufacturing (Kamble *et al.*, 2020; Konstantinidis *et al.*, 2023), and tourism (Carlisle *et al.*, 2023; Jain *et al.*, 2023). In this context, a large number of indexes and models have been developed (Tutak & Brodny, 2022). Despite some shortcomings, DM studies are accepted to contribute to the digitalization processes of companies (Thordsen & Bick, 2023). In this context, it is emphasized that the technologies developing with I4.0 should also be added to DM analyses (Hsu & Yeh, 2017; Sharma et al., 2020). However, a similar situation on method confusion in I4.0-DD studies is also seen in DM studies. The general notion in the literature is that the methods are not sufficient for a comprehensive evaluation of DM. Hellweg et al. (2021) accepts the reason for this inadequacy as the lack of a theoretical infrastructure for DM. According to researchers, theoretical inadequacies constitute an obstacle to the development of a valid measurement method. This causes a methodological confusion in DM studies (Pavel *et al.*, 2021). In addition to

the method confusion, Cinar et al., (2021) emphasizes that the developed models are not capable of covering all sectors. Although DM models have similar structures, Cognet et al., (2023) claim that each one follows different directions and state that some models focus on technology and some models focus on marketing. For this reason, there is no homogeneity in DM models. It can be said that I4.0-DD and DM studies are related to technological capacity and both research areas face methodological problems. The development of a method with theoretical foundations will solve the problems encountered in these studies. In this context, there is a need to develop a measurement method having scientific basis, questionable, verifiable and generally accepted.

There are few studies using machine learning to identify differences within the scope of I4.0-DD. These studies include methods such as logistic regression (Kovacic & Vukmirovic, 2008), decision trees (Painsky & Rosset, 2017), and classification (Hidalgo *et al.*, 2020) instead of unsupervised machine learning. However, the most appropriate method that can be used to group a dataset according to similar features is clustering (Han et al., 2012). Only a few studies that include clustering analysis have been found in the literature. Applying hierarchical clustering to identify differences between firms, Frank et al. (2019) identified 3 clusters for company levels: low adapters, medium and advanced. Eurostat data and DESI were used in a study examining the relationship between digitalization of EU-28 countries and global competition and labor productivity (Polozova *et al.*, 2021). K-Means clustering was applied to DESI values to group countries. Researchers, who created 4 clusters: leaders, perspective, followers and transition, determined that the leading countries are located in northern Europe. Applying factor analysis to Eurostat data, Branco et al., (2023) divided European countries into 5 clusters using the Ward clustering. Kolupaieva and Tiesheva (2023), who examined the asymmetry between I4.0 technology usages in EU-28 countries, stated that Fuzzy clustering is the most effective method in grouping countries. Researchers, who determined the significance of the clusters created with the fuzzy approach using the F-test, determined that northern European countries have significant advantages.

It is seen that there is no consensus in the literature for measuring I4.0-DD levels and the methods used are questioned in the literature. Different components, variables and calculation methods used in the created indexes cause the results to be different. The second deficiency observed in the studies is that the validation analyses of the methods are not included. For this reason, there are doubts about which method will be valid and successful. Another deficiency in the literature is that the differences are determined according to the averages on a country basis. However, instead of determining the differences according to a certain score value, the approach of grouping countries with similar characteristics will provide a more reliable interpretation of the results in I4.0-DD studies. The levels of differences within and between groups can also be calculated with cluster analysis. Although some studies applying machine learning methods for grouping countries, clustering is rarely encountered in analyses and performance analyses are not included. The possibilities of clustering

analysis will eliminate the deficiencies in I4.0-DD studies. Interpreting the results according to clustering methods with higher performance will provide more successful inferences. Measuring inter-cluster and intra-cluster distances will determine how different countries are from each other within the scope of I4.0-DD and what level of precautions should be taken. In this way, an appropriate roadmap can be determined for the solution of the I4.0-DD problem and effective policies can be developed.

**Materials and Methods**

*Dataset*

For the research, CC, BD, 3D, IoT, AI, ICT specialist employment (ICT-Emp) and ICT-skill training (ICT-Train) rates of companies in EU-27 countries and Turkiye were used. The data for these criteria was obtained from the ICT Usage in Enterprises database on the Eurostat website in November 2023. After collecting the data, it was observed that some countries had missing data in their criteria values. In order not to negatively affect the analysis, the missing data were completed by taking the average of the relevant criteria according to the column data in which they were located. Thus, the dataset to be used for the research was created with 28 items and 7 criteria and presented in Table 1..

Table 1

**Research Dataset**

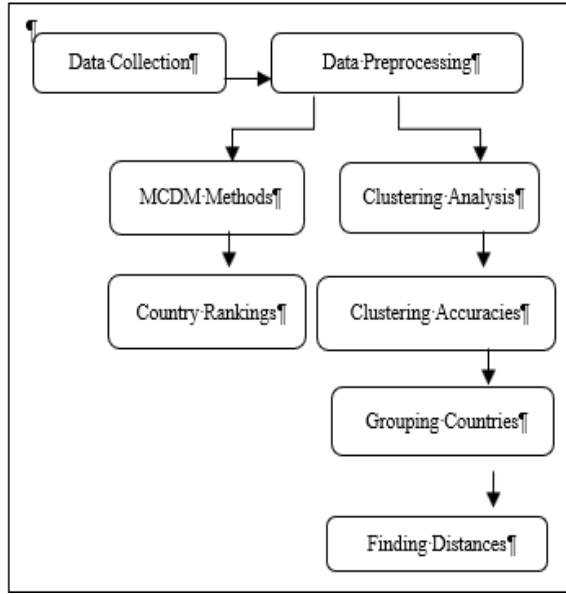
		CC	BD	3D	IoT	AI	ICT-Emp	ICT-Train
C_1	Belgium	<b>40.16*</b>	<b>10.84*</b>	27.10	<b>30.00*</b>	10.60	5.60	<b>22.01*</b>
C_2	Bulgaria	9.70	5.10	9.20	15.20	2.80	3.80	6.90
C_3	Czechia	38.90	8.30	15.50	37.30	4.20	4.50	23.90
C_4	Denmark	62.80	22.60	37.50	21.80	27.30	5.70	32.10
C_5	Germany	38.80	13.10	14.90	36.80	9.40	5.00	28.90
C_6	Estonia	58.20	7.60	12.70	21.10	2.80	6.60	14.40
C_7	Ireland	67.40	22.00	17.80	<b>30.00*</b>	8.50	6.20	28.40
C_8	Greece	18.70	<b>10.84*</b>	9.30	23.50	5.60	2.50	15.00
C_9	Spain	27.90	6.40	18.70	28.30	7.50	4.30	17.40
C_10	France	27.70	16.70	24.80	22.10	5.50	4.30	16.60
C_11	Croatia	34.80	12.20	9.50	24.10	8.40	3.70	19.10
C_12	Italy	61.90	6.30	18.30	36.90	6.70	3.90	19.20
C_13	Cyprus	43.90	4.70	8.30	37.20	1.90	4.60	22.20
C_14	Latvia	26.70	7.70	11.80	33.20	2.60	4.40	12.60
C_15	Lithuania	35.30	7.90	13.50	30.00	5.10	4.40	11.90
C_16	Luxembourg	36.20	13.10	23.20	29.50	12.60	7.70	31.40
C_17	Hungary	24.80	7.40	11.00	23.10	4.00	4.10	18.00
C_18	Malta	53.40	<b>10.84*</b>	<b>16.82*</b>	28.70	7.30	4.80	<b>22.01*</b>
C_19	Netherlands	62.00	23.50	22.00	22.30	12.50	7.20	29.20
C_20	Austria	37.10	9.40	17.00	55.20	9.80	5.00	26.60
C_21	Poland	28.20	6.90	10.40	19.00	2.30	3.60	23.80
C_22	Portugal	29.40	7.00	17.30	23.60	19.10	4.50	19.10
C_23	Romania	11.20	3.50	6.30	9.40	1.20	2.80	6.90
C_24	Slovenia	42.90	7.60	18.70	53.70	12.40	4.50	33.00
C_25	Slovakia	33.80	4.40	16.70	31.40	5.30	4.30	18.50
C_26	Finland	85.20	21.40	26.70	48.50	17.20	7.60	42.30
C_27	Sweden	76.90	17.80	26.70	45.60	9.50	8.60	39.90
C_28	Turkiye	10.60	8.50	9.20	22.40	3.20	<b>4.97*</b>	15.20

Source: EuroStat. \*Values are obtained by the column arithmetic mean

**Research Method**

The schematic representation of the research model is given in Figure 2. After the data was collected from Eurostat, the dataset was preprocessed due to some missing values. In order not to negatively affect the analysis, the missing data were determined by taking the column averages (see in Table 1). First, MCDM methods were applied to the dataset and country rankings were determined. Then, cluster analyses were applied to the same dataset. The clustering performances were calculated by the accuracy analysis of the clustering results. In this way, the most successful clustering algorithm and the number of clusters were determined and the countries were grouped with the highest performing option. In the last step,

according to the clustering results of the countries, distance values were calculated for within cluster and between clusters, and the difference levels between the countries were determined.



**Figure 2.** Research Model

### Clustering Algorithms

Clustering, explained as categorizing the items in a dataset according to similar criteria (Han et al., 2012; Fidan & Yuksel, 2022). The basis of the method is the assumption that there is an item in the dataset that has more similar features than other items (Alpaydın, 2010). In other words, clustering aims to identify items with the closest features. Commonly used clustering approaches are grouped as hierarchical, partitional and fuzzy (Han et al., 2012). Commonly used clustering algorithms for these approaches are Hierarchical Clustering (HC), K-Means (K-M) and Fuzzy C-Means (FCM), respectively (Peters et al., 2013).

#### Hierarchical Clustering:

It is a clustering algorithm that aims to identify similar item pairs by calculating item distances (Han et al., 2012). In the HC algorithm, first of all, each item is accepted as a stand-alone cluster. The nearest item pair is identified by using distance measures such as Manhattan, Euclidian, and Cosine, and this item pair is considered a cluster (See in Equation 1). In other steps, the same pairing processes are applied and repeated until all items in the dataset are in a single group.

$$\min d(i_a, i_b) = \min_{p \in i_a, p^* \in i_b} |p - p^*| \quad (1)$$

In Equation 1, the determination of the minimum value of the distance ( $d$ ) between items is shown.  $i_a$  and  $i_b$  refer to items, and  $p - p^*$  refers to the distance of  $a$  and  $b$ .

#### K-Means:

K-M aims to find the closest items to cluster centres specified by  $k$ . Distance measurements are used to determine the distances of the cluster items, and the mean of these values is accepted as the centre (Peters et al., 2013). Clusters are created with the items closest to the centre. As shown in Equation 2, the decision criterion is the minimization of the sum of squares ( $E$ ) of the distances from the cluster centres of the items.

$$\min E = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - \mu_k\|^2 \quad (2)$$

In Equation 2,  $x_i$  is an item in cluster  $C_k$ .  $\mu_k$  shows the mean of  $C_k$ . The initially determined cluster centres are redetermined according to the mean of the distances. This procedure is iteratively carried out as far as the  $E$  reaches its minimum value (Han et al., 2012; Peters et al., 2013).

#### Fuzzy C-Means:

FCM aims to reduce the weighted sum of the membership values of the items. Euclidean is applied for calculating item distances. Equation 3 shows the clustering criteria that will form  $K$  clusters of  $N$  items.

$$\min J = \sum_{k=1}^K \sum_{i=1}^N \lambda_{i,k}^m |x_i - \mu_k|^2 \quad (3)$$

In Equation 3,  $\lambda_{i,k}$  shows the membership degree of item  $i$  in cluster  $k$ .  $\lambda_{i,k}$  is given in Equation 4 and  $\lambda_{i,k} \in [0,1]$ .  $m$  value is a parameter for fuzzifier and  $m \in (1, \infty)$ . It is suggested in literature that  $m$  value should be set as 2 (Bezdek, 1981).

$$\lambda_{i,k} = \frac{1}{\sum_{j=1}^K \left( \frac{d(x_i, \mu_k)}{d(x_j, \mu_k)} \right)^{\frac{2}{m-1}}} \quad (4)$$

#### Clustering Performance

Validation techniques are used to determine the success level of a clustering analysis. Measurement of cluster performance can be realized by two validations: external and internal. Validation method depends on the including internal data or not (Han et al., 2012). In internal validations that do not have external data, performance is determined according to the structure of the clusters. Despite internal validation methods with different approaches, Silhouette index (SI) and Dunn index (DI) are widely preferred for analysis because they take compactness and separation criteria together (Liu et al., 2010; Gupta & Panda, 2019).

#### Silhouette Index:

SI, which is an internal validation method used to determine the clustering quality, is calculated according to the similarity level of each item to its own cluster (compactness) and its difference values to other cluster items (separation).

$$SI = \frac{1}{N} \sum_{C_k \in C} \sum_{i \in C_k} \frac{b_i - a_i}{\max(a_i, b_i)} \quad (5)$$

where

$$a_i = \frac{1}{|C_k|} \sum_{j \in C_k} d(i, j) \quad (6)$$

$$b_i = \min_{C_t \in C \setminus C_k} \frac{1}{|C_t|} \sum_{j \in C_t} d(i, j) \quad (7)$$

Calculation of SI value is shown in Equation 5 for a data set having  $N$  items for  $C$  clusters. The item  $i$  is located in the  $C_k$  cluster and  $C_k \in C$ .  $a_i$  given in Equation 6 represents the compactness value of SI. It represents the average distance to all items within the cluster containing item  $i$ .  $b_i$ , which represents separation, is the smallest value among the average distances from item  $i$  to the items in other clusters (See in Equation 7). Euclidean distance is used to calculate

distances. The  $d$  value in Equation 6 and Equation 7 is founded by  $d(i, j) = \sqrt{(i - j)^2}$ .

For a multivariate data set having  $m$  features,  $d(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{im} - x_{jm})^2}$  where  $i = (x_{i1}, x_{i2}, \dots, x_{im})$ ,  $j = (x_{j1}, x_{j2}, \dots, x_{jm})$  and SI is  $-1 \leq SI \leq 1$  (Han et al., 2012). When SI approaches to +1, it means better performance in clustering (Rousseeuw, 1987).

*Dunn Index:*

The Dunn Index (DI) assesses the clustering performance by considering the ratio of intra-cluster compactness to inter-cluster separation, as introduced by Dunn in 1973. In this evaluation, the separation between clusters is determined based on the minimum distance between items belonging to different clusters, while the compactness is represented by the maximum diameter of the clusters, as outlined by Arbelaitz et al. in 2013. The Dunn Index is specifically defined as the ratio of the minimum distance among items in all clusters to the maximum diameter of the clusters (Fidan & Yuksel, 2022).

$$DI = \min_{1 \leq i \leq k} \left( \min_{1 \leq j \leq k, j \neq i} \left( \frac{\text{dist}(c_i, c_j)}{\max_{1 \leq l \leq k} \text{diam}(c_l)} \right) \right) \quad (8)$$

where  $x_i$  and  $x_j$  are the items in different clusters,  $\text{dist}(c_i, c_j) = \min_{x_i \in c_i \text{ and } x_j \in c_j} d(x_i, x_j)$  and  $\text{diam}(c_i) =$

$\max_{x_{i1}, x_{i2} \in c_i} d(x_{i1}, x_{i2})$ . A high value of the Dunn Index indicates a strong clustering performance (Arbelaitz et al., 2013).

**Findings**

**Country Rankings**

Multi-criteria decision making (MCDM) methods TOPSIS, VIKOR and PROMETHEE II were used to rank the countries in the data set containing I4.0 technology usage rates. It is assumed that all item criteria in the data set have equal weight. For this reason, the weight calculation of the criteria was not carried out. Country rankings are given in Table 2.

Table 2 are examined, the differences in country rankings are clearly seen. While Finland ranks first in all three methods, Denmark ranks 2nd according to the TOPSIS method and 3rd according to the VIKOR and PROMETHEE II methods. Lithuania, which is ranked 21st in TOPSIS, is ranked 17th according to VIKOR and 25th according to PROMETHEE II. In other words, Lithuania, which is seen in a worse situation than Estonia in TOPSIS in terms of I4.0 usage, is in a better situation than VIKOR, but according to PROMETHEE II, there is a serious difference in favour of Estonia. There are similar implications for the rankings of different countries.

Table 2

**Rankings of the Countries**

	Country	TOPSIS		VIKOR		PROMETHEE II	
		Score	Rankings	Score	Rankings	Scores	Rankings
C_1	Belgium	0.437655	8	0.437078	7	0	21
<b>C_2</b>	<b>Bulgaria</b>	<b>0.086585</b>	<b>27</b>	<b>0.956338</b>	<b>27</b>	0	22
C_3	Czechia	0.316131	15	0.700046	15	0.007881	14
<b>C_4</b>	<b>Denmark</b>	<b>0.742044</b>	<b>2</b>	<b>0.338504</b>	<b>3</b>	<b>0.080091</b>	<b>3</b>
C_5	Germany	0.411147	11	0.514419	9	0.016893	9
<b>C_6</b>	<b>Estonia</b>	<b>0.299161</b>	<b>16</b>	<b>0.761139</b>	<b>19</b>	<b>0.014921</b>	<b>11</b>
C_7	Ireland	0.510215	5	0.431876	6	0.044358	5
C_8	Greece	0.204296	23	0.894635	26	0	23
C_9	Spain	0.284143	18	0.712245	16	0	24
C_10	France	0.350429	14	0.655214	13	0.010852	13
C_11	Croatia	0.294851	17	0.748676	18	0.002519	18
C_12	Italy	0.365994	13	0.666613	14	0.016245	10
C_13	Cyprus	0.272535	19	0.812256	22	0.007856	15
C_14	Latvia	0.219064	22	0.818926	23	0.002588	17
<b>C_15</b>	<b>Lithuania</b>	<b>0.248134</b>	<b>21</b>	<b>0.731518</b>	<b>17</b>	<b>0</b>	<b>25</b>
C_16	Luxembourg	0.49977	6	0.391737	4	0.029598	7
C_17	Hungary	0.198734	25	0.786358	20	0	26
C_18	Malta	0.36756	12	0.578256	11	0.006495	16
C_19	Netherlands	0.566169	4	0.403489	5	0.054092	4
C_20	Austria	0.437128	9	0.480308	8	0.025181	8
C_21	Poland	0.202444	24	0.845707	24	0.001873	19
C_22	Portugal	0.431188	10	0.6508	12	0.012062	12
<b>C_23</b>	<b>Romania</b>	<b>0.014482</b>	<b>28</b>	<b>1</b>	<b>28</b>	<b>0</b>	<b>27</b>
C_24	Slovenia	0.480171	7	0.535816	10	0.034562	6
C_25	Slovakia	0.267499	20	0.799989	21	0.001132	20
<b>C_26</b>	<b>Finland</b>	<b>0.750541</b>	<b>1</b>	<b>0</b>	<b>1</b>	<b>0.099348</b>	<b>1</b>
<b>C_27</b>	<b>Sweden</b>	<b>0.61723</b>	<b>3</b>	<b>0.288366</b>	<b>2</b>	<b>0.080459</b>	<b>2</b>
<b>C_28</b>	<b>Turkiye</b>	<b>0.180346</b>	<b>26</b>	<b>0.87911</b>	<b>25</b>	<b>0</b>	<b>28</b>

**Clustering Results**

HC, K-M and FCM algorithms were used in the cluster analyses based on I4.0 technologies usage rates. With these clustering algorithms determined in accordance with the literature, clustering analyses were carried out so that the countries were formed into 3, 4, 5 and 6 clusters. Clustering performances are given in Table 3.

**Clustering Results**

	HC		K-M		FCM	
	SI	DI	SI	DI	SI	DI
k=3	0.27	0.171	0.307	0.178	0.273	0.143
k=4	0.284	0.198	<b>0.315</b>	0.193	0.07	0.178
k=5	0.302	0.207	0.309	0.193	0.124	0.178
k=6	0.286	0.207	0.305	0.193	0.096	0.178

Table 3

The validation values in Table 3 are examined, it can be said that cluster analyses generally have low performance. According to SI values, it is observed that the most successful clustering is the K-M algorithm at k=4, and the least successful clustering is the FCM algorithm at k=3. According to the DI values, while HC is the most successful in cases k=5 and k=6, FCM is seen as unsuccessful in cases k=3. Since DI values have the same values in k=4, k=5 and k=6, it would be healthier to make evaluations according to SI values. The SI values show that creating 4 clusters by using K-M algorithm is the most successful clustering analysis. K-M clustering results for k=4 are given in Table 4.

**Clustering Results for K-M k=4**

Clusters	Countries	Cluster means	Neighbor clusters	Within clusters Sum of Square
Cluster1	C_1, C_3, C_5, C_13, C_16, C_20, C_24, C_25	20.52	Cluster2, Cluster3	1395.97
Cluster2	C_4, C_6, C_7, C_12, C_18, C_19	23.55	Cluster1, Cluster4	1619.34
Cluster3	C_2, C_8, C_9, C_10, C_11, C_14, C_15, C_17, C_21, C_22, C_23, C_28	13.18	Cluster1	2312.66
Cluster4	C_26, C_27	33.85	Cluster2	78.15

Table 4

The distribution of countries in the clusters is examined, it is observed that the least number of countries is in Cluster4 and the most countries are in Cluster3. Cluster4 has the highest value of Cluster means, which shows the mean of the criteria values of the items and is also centre of the cluster. This determination means that Cluster4 is the highest level in I4.0 technologies. There are two countries in this cluster, Finland and Sweden. Denmark, which ranks 2nd in Table 2 created with TOPSIS, is in Cluster2. Cluster3, which has the lowest Cluster mean, includes the

countries that use I4.0 technologies at the lowest level. In this context, the usage levels of I4.0 technologies were determined as Cluster4, Cluster2, Cluster1 and Cluster3, respectively.

Within clusters Sum of Square values, which show the level of deviation from the average within the cluster, show the difference values of the cluster elements from each other. In Table 4, it is observed that the least deviation occurred in Cluster4 and the highest deviation occurred in Cluster3. This observation means that the countries in Cluster4 are close to each other. The cluster with the highest deviation is Cluster3. The countries in this cluster are the countries with the highest differences.

**Differences between Clusters**

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster1	0			
Cluster2	0.14	0		
Cluster3	0.35	0.44	0	
Cluster4	0.65	0.43	1.56	0

Table 5

Table 5 shows the percentage differences between clusters calculated using Cluster means values. Table 5 shows a serious I4.0-DD of 156% between Cluster4 and Cluster3. The second best, Cluster2, has a difference of 43% with Cluster4, while its difference with Cluster1 is 14%. This value shows that there is less difference between the countries in Cluster2 and the countries in Cluster1, which is at the lower level.

**Discussion and Comparison**

The findings of the study reveal interesting results. According to the clustering results, the Northern European countries of Finland and Sweden were determined as the countries that use I4.0 technologies at the highest level. Germany, which is considered the starting point of the I4.0 idea, was not included in this cluster. Countries such as Romania, Bulgaria and Hungary are among the 12 European countries at the lowest level. The I4.0-DD difference between these countries and the group consisting of Finland and Sweden was calculated as 156%. Similarly, the difference between Cluster4 and the second best group Cluster2 is 43%. These differences reveal that serious precautions need for the issue, especially in countries with low usage values.

In the studies in the literature, it has been observed that some indexes and MCDM methods created according to different parameter and weight values are generally applied for country comparisons. Reaching different results in analyses realized by different indexes and MCDM methods creates confusion in determining the appropriate method. In this study, the clustering method, which was rarely included in previous studies, was used and more interpretable results were obtained. Additionally, it has been observed that MCDM methods such as TOPSIS, VIKOR and PROMETHEE II produce different ranking results. This determination raises a question mark in the grouping operations performed with MCDM methods. MCDM

results also differ from the results of cluster analysis. Denmark, which ranks 2nd in the rankings, was not included in the top-level countries cluster (Cluster4).

Narrow-scope analyses were performed in a few studies that used the clustering approach to determine I4.0-DD levels. Therefore, a comparison of alternative approaches for a healthier result is not presented. In the study, performance analyses of the algorithms were carried out, comparisons were made with the algorithms recommended in the literature, and it was revealed that some algorithms were insufficient. Contrary to the findings of Kolupaieva and Tiesheva (2023), FCM was determined to be the least successful algorithm in grouping countries. It would be an appropriate approach not to use FCM in data sets with a small number of items. Additionally, Frank et al., (2019) suggested 3 clusters and Branco et al., (2023) proposed 5 clusters, it was determined in this study that the highest clustering performance was obtained with the K-M algorithm and 4 clusters. Therefore, using 4 clusters would be the appropriate option for grouping countries.

In determining I4.0-DD values, unlike other studies, both intra-cluster and inter-cluster differences were calculated in this study. So, the research suggests that employing cluster analysis is likely to produce more effective and understandable results for studies related to I4.0-DD. With these calculations, which have not been encountered in previous I4.0-DD studies in the literature, the difference values of the countries in the cluster are determined as well as the difference levels between the clusters are determined quantitatively. According to these values, the highest I4.0-DD differences within the cluster are among the Cluster3 countries that use I4.0 technologies at the least level.

### **Conclusion**

I4.0 technologies have garnered attention from researchers due to the positive effects on business processes. However, disparities in the adoption of these technologies give rise to new inequalities among countries. The use of different variables, indexes, and tools in the literature to categorize countries based on these inequalities has led to methodological confusion on the subject. In addition, the inconsistent results obtained from the rankings with the MCDM methods lead researchers to approach with suspicion, and the lack of accuracy analysis reduces the reliability of the results. Although clustering is deemed the most appropriate approach for grouping countries, it is noteworthy that cluster analyses are not extensively featured in the literature on this topic.

In this study, the unsupervised machine learning method was used with the idea that it would reduce the method confusion, reach inquirable results and enable the levels of difference to be determined more clearly. In this context, cluster analysis and accuracy analysis were conducted in the study to determine the most successful option, than calculated the I4.0-DD levels observed among European countries. The study reveals that I4.0-DD levels can be determined by clustering method. The clustering method was preferred because it has a theoretical basis and it has analysis tools eliminating the deficiencies in the

literature. Clustering analyses were applied to the dataset using HC, K-M, and FCM algorithms, widely used in the literature.

In the study, accuracy analysis was applied to the clustering results in order to quantitatively determine the validation of the analyses. Thus, the most successful clustering analysis was specified and the lack of accuracy analysis in the literature was eliminated. According to the accuracy results, it was determined that the K-M algorithm provided the most successful clustering results for  $k = 4$ . Finland and Sweden, using I4.0 technologies at the highest level, were grouped in Cluster 4, while countries at the lowest level were clustered in Cluster 3.

According to the clustering results, the difference levels between countries were determined quantitatively as within and between clusters. With this method, which is not encountered in the I4.0-DD literature, the differences of countries within the same cluster were calculated according to the cluster mean, and the differences of countries between different clusters were calculated according to the cluster centers. Thus, the I4.0-DD level difference between Cluster4 and Cluster3 was calculated as 156%. Cluster 4 was identified as having the lowest intra-cluster I4.0-DD levels, whereas Cluster 3 exhibited the highest intra-cluster differences. While this determination provides researchers with a new analysis tool in calculating the difference level, it will also provide decision-makers with a broader perspective in the roadmap and policies to be developed.

It was observed in the literature that MCDM methods produced inconsistent results. In this context, the study reveals that the rankings with MCDM methods contain inconsistencies. The study proves that rankings with MCDM methods will not be a suitable option in I4.0-DD analyses. For this reason, it is recommended that confirmatory analyses should definitely be included in studies involving MCDM.

The study, which provides a new perspective on method, accuracy analysis and determination of difference levels, has some limitations. Since the study was carried out with data obtained from Eurostat, the quality of the dataset depends on Eurostat. It was observed that some variables did not have values. Since missing data will negatively affect the cluster analysis, a preprocessing was applied. This intervention to the dataset may cause some countries being not included in the cluster they should actually be in. In this respect, the completeness of Eurostat data is closely related to the quality of the analysis. On the other hand, the algorithms used in the cluster analysis are limited to the algorithms widely used in the literature. In future studies, the inclusion of different clustering algorithms in the analysis will contribute to the literature on the subject by expanding the scope of the study. Another limitation is the low clustering performances. Due to small number of items in the dataset, the clustering performances in the study remained at a low level. In addition, the low performance values reveal that new clustering algorithms that will exhibit higher performance are needed in clustering datasets with small number of items. New clustering algorithms to be developed in future studies would allow for the attainment of more reliable results in I4.0-DD studies.

## References

- Ahmed, I., Jeon, G., & Piccialli, F. (2022). From artificial intelligence to explainable artificial intelligence in industry 4.0: A survey on what, how, and where. *IEEE Transactions on Industrial Informatics*, 18(8), 5031-5042. <https://doi.org/10.1109/TII.2022.3146552>
- Ali, M. A., Alam, K., Taylor, B., & Rafiq, S. (2020). Does ICT maturity catalyse economic development? Evidence from a panel data estimation approach in OECD countries. *Economic Analysis & Policy*, 68, 163–174. <https://doi.org/10.1016/j.eap.2020.09.003>
- Alpaydin, E. (2010). *Introduction to Machine Learning*. London: The MIT Press. <https://doi.org/10.1108/TQM-06-2021-0180>
- Antony, J., Sony, M., & McDermott, O. (2021). Conceptualizing Industry 4.0 readiness model dimensions: An exploratory sequential mixed-method study. *The TQM Journal*, 35(2), 577-596. <https://doi.org/10.1016/j.patcog.2012.07.021>
- Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Perez, J. M., & Perona, I. (2013). An extensive comparative study of cluster validity indices. *Pattern Recognition*, 46(1), 243–256. <https://doi.org/10.1080/23311916.2019.1623153>
- Asadabadi, M. R., Chang, E., Saberi, M., & Zhou, Z. (2019). Are MCDM methods useful? A critical review of Analytic Hierarchy Process (AHP) and Analytic Network Process (ANP). *Cogent Engineering*, 6(1), 1-11. <https://doi.org/10.1016/j.procs.2019.09.123>
- Atik, H., & Unlu, F. (2019). The measurement of industry 4.0 performance through industry 4.0 index: An empirical investigation for Turkey and European Countries. *Procedia Computer Science*, 158, 852-860. <https://doi.org/10.2307/2673277>
- Attewell, P. (2001). Comment: The first and second digital divides. *Sociology of Education*, 74(3), 252–259. <https://doi.org/10.2307/2673277>
- Balkan, D., & Akyuz, G. A. (2023). Technological maturity of the OECD countries: A multi-criteria decision-making approach using PROMETHEE. *Cogent Engineering*, 10(1), pp. 1-19. <https://doi.org/10.1080/23311916.2023.2219097>
- Benlian, A., Kettinger, W. J., Sunyaev, A., & Winkler, T. J. (2018). Special section: The transformative value of cloud computing: A decoupling, platformization, and recombination theoretical framework. *Journal of Management Information System*, 35(3), 719–739. <https://doi.org/10.1080/07421222.2018.1481634>
- Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Algorithms*. New York: Plenum Press. <https://doi.org/10.1007/978-1-4757-0450-1>
- Bi, Z., Xu, L. D., & Wang, C. (2014). Internet of Things for enterprise systems of modern manufacturing. *IEEE Trans Ind Inform*, 10(2), 1537-1546. <https://doi.org/10.1109/TII.2014.2300338>
- Branco, I. C., Henriques, M. A., Jesus, F. C., & Oliveira, T. (2023). Assessing the Industry 4.0 European divide through the country/industry dichotomy. *Computers and Industrial Engineering*, 176, 108925, 1-14. <https://doi.org/10.1016/j.cie.2022.108925>
- Carlisle, S., Ivanov, S., & Dijkmans, C. (2023). The digital skills divide: Evidence from the European tourism industry. *Journal of Tourism Futures*, 9(2), 240-266. <https://doi.org/10.1108/JTF-07-2020-0114>
- Cognet, B., Pernot, J. P., Rivest, L., & Danjou, C. (2023). Systematic comparison of digital maturity assessment models. *Journal of Industrial and Production Engineering*, 40(7), 519–537. <https://doi.org/10.1080/21681015.2023.2242340>
- Contieri, P. G. S., Anholon, R., & De Santa Eulalia, L. E. (2021). Industry 4.0 enabling technologies in manufacturing: Implementation priorities and difficulties in an emerging country. *Technology Analysis & Strategic Management*, 34(5), 489–503. <https://doi.org/10.1080/09537325.2021.1908536>
- Cooper, J., & James, A. (2009). Challenges for database management in the Internet of things. *IETE Technical Review*, 26, 320-329. <https://doi.org/10.4103/0256-4602.55275>
- Cinar, Z. M., Zeeshan, Q., & Korhan, O. (2021). A framework for industry 4.0 readiness and maturity of smart manufacturing enterprises: A case study. *Sustainability*, 13 (12) (2021), 6659. <https://doi.org/10.3390/su13126659>
- Dalenogare, L. S., Benitez, G. B., Ayala, N. F., & Frank, A. G. (2018). The expected contribution of Industry 4.0 technologies for industrial performance. *Int. J. Prod. Econ*, 204(1), 383-394. <https://doi.org/10.1016/j.ijpe.2018.08.019>
- Drucker, P. (1986). *Technology, management and society*. London, England: Routledge. <https://doi.org/10.104324/9780080928876>
- Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3, 32–57. <https://doi.org/10.1080/01969727308546046>
- Elibal, K., & Ozceylan, E. (2022). Comparing industry 4.0 maturity models in the perspective of TQM principles using Fuzzy MCDM methods. *Technological Forecasting and Social Change*, 175, 121379. <https://doi.org/10.1016/j.techfore.2021.121379>

- European Commission. (2021). Digital Economy and Society Index. Available from internet: <https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-index-desi-2021>
- EuroStat. (2023). ICT Usage in Enterprises E-business dataset. Available from internet: [http:// https://ec.europa.eu/eurostat/web/main/data/database](http://https://ec.europa.eu/eurostat/web/main/data/database)
- Fidan, H. (2016). Measurement of the intersectoral digital divide with the Gini coefficients: Case study Turkey and Lithuania. *Inzinerine Ekonomika-Engineering Economics*, 27(4), 439–451. <https://doi.org/10.5755/j01.ee.27.4.13882>
- Fidan, H., & Yuksel, M. E. (2020). Knowledge management and internet of things: Evolutions in business processes. In Emine Yıldız Kuyrukçu (Ed.), *Engineering and Architecture Sciences Theory, Current Research and New Trends/2* (pp. 88-107), Montenegro: Ivpe.
- Fidan, H., & Yuksel, M. E. (2022). A comparative study for determining Covid-19 risk levels by unsupervised machine learning methods. *Expert Systems with Applications*, 190, 116243, 1-9. <https://doi.org/10.1016/j.eswa.2021.116243>
- Fisher, T. (1990). Business productivity measurement using standard cost accounting information. *International Journal of Operations & Production Management*, 10(8), 61-69. <https://doi.org/10.1108/01443579010142243>
- Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. *International Journal of Production Economics*, 210, 15-26. <https://doi.org/10.1016/j.ijpe.2019.01.004>
- Gerpott, T. J., & Ahmadi, N. (2015). Composite indices for the evaluation of a country's information technology development level: Extensions of the IDI of the ITU. *Technological Forecasting and Social Change*, 98, 174-185. <https://doi.org/10.1016/j.techfore.2015.03.012>
- Ghobakhloo, M. (2020). Industry 4.0, digitization, and opportunities for sustainability. *Journal of Cleaner Production*, 252, 1-10. <https://doi.org/10.1016/j.jclepro.2019.119869>
- Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision architectural elements and future directions, *Future Generation Computer Systems*, 29(7), 1645-1660. <https://doi.org/10.1016/j.future.2013.01.010>
- Gupta, T., & Panda, S. P. (2019). Clustering validation of CLARA and K-Means using Silhouette & DUNN measures on Iris dataset. In *International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*. <https://doi.org/10.1109/COMITCon.2019.8862199>
- Hajoary, P. K., Balachandra, P., & Garza-Reyes, J. A. (2023). Industry 4.0 maturity and readiness assessment: An empirical validation using Confirmatory Composite Analysis. *Production Planning & Control*, 35(14), 1779–1796. <https://doi.org/10.1080/09537287.2023.2210545>
- Hallikainen, H., Savimaki, E., & Laukkanen, T. (2020). Fostering B2B sales with customer big data analytics. *Industrial Marketing Management*, 86, 90-98. <https://doi.org/10.1016/j.indmarman.2019.12.005>
- Han, H., & Trimi, S. (2022). Towards a data science platform for improving SME collaboration through Industry 4.0 technologies. *Technological Forecasting and Social Change*, 174, 121242, 1-13. <https://doi.org/10.1016/j.techfore.2021.121242>
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining concepts and techniques* (Third Edition). USA: Morgan Kaufmann Publications.
- Hellweg, F., Lechtenberg, S., Hellingrath, B., Thome, A., & Tavares, M. (2021). Literature review on maturity models for digital supply chains. *Brazilian Journal of Operations and Production Management*, 18(3), 1–12. <https://doi.org/10.14488/BJOPM.2021.022>
- Hidalgo, A., Gabaly, S., Alonso, G. M., & Uruena, A. (2020). The digital divide in light of sustainable development: An approach through advanced machine learning techniques. *Technological Forecasting and Social Change*, 150, 1-7. <https://doi.org/10.1016/j.techfore.2019.119754>
- Hizam-Hanafiah, M., Soomro, M. A., & Abdullah N. L. (2020). Industry 4.0 readiness models: A systematic literature review of model dimensions. *Information*, 11(7), 364. <https://doi.org/10.3390/info11070364>
- Honti, G., Czvetko, T., & Abonyi, J. (2020). Data describing the regional Industry 4.0 readiness index. *Data in Brief*, 33, 106464. <https://doi.org/10.1016/j.dib.2020.106464>
- Hsu, C.-W., & Yeh, C. C. (2017). Understanding the factors affecting the adoption of the Internet of things. *Technology Analysis & Strategic Management*, 29(9), 1089–1102. <https://doi.org/10.1080/09537325.2016.1269160>
- Jain, P., Singh, R. Kr., Mishra, R., & Rana, N. P. (2023). Emerging dimensions of blockchain application in tourism and hospitality sector: a systematic literature review. *Journal of Hospitality Marketing & Management*, 32(4), 454–476. <https://doi.org/10.1080/19368623.2023.2184440>
- Jandyal, A., Chaturvedi, I., Wazir, I., Raina, A., & Ul-Haq, M. I. (2022). 3D printing – A review of processes, materials and applications in industry 4.0. *Sustainable Operations and Computers*, 3, 33-42. <https://doi.org/10.1016/j.susoc.2021.09.004>

- Jernigan, S., Ransbotham, S., & Kiron, D. (2016). Data sharing and analytics drive success with IoT. MIT Sloan Management Review, 8. Available from internet: <http://sloanreview.mit.edu/projects/data-sharing-and-analytics-drive-success-with-internet-of-things>
- Kagermann, H., Wahlster, W., & Helbig, J. (2013). Recommendations for implementing the strategic initiative Industrie 4.0: Securing the future of German manufacturing industry. Final Report of the Industrie 4.0 Working Group Acatech, Forschungsunion. Available from internet: <https://www.din.de/resource/blob/76902/e8cac883f42bf28536e7e8165993f1fd/recommendations-for-implementing-industry-4-0-data.pdf>
- Kamble, S. S., Gunasekaran, A., Ghadge, A., & Raut, R. (2020). A performance measurement system for industry 4.0 enabled smart manufacturing system in SMMEs- A review and empirical investigation. *International Journal of Production Economics*, 229, 107853, 1-15. <https://doi.org/10.1016/j.ijpe.2020.107853>
- Kloviene, L., & Uosyte, I. (2019). Development of performance measurement system in the context of Industry 4.0: A case study. *Inzinerine Ekonomika-Engineering Economics*, 30(4), 472–482. <https://doi.org/10.5755/j01.ee.30.4.21728>
- Kobets, V., Yatsenko, V., & Voynarenko, M. (2020). Cluster analysis of countries inequality due to IT development through macros application. In: Ermolayev, V., Mallet, F., Yakovyna, V., Mayr, H., Spivakovsky, A. (Ed.) *Information and Communication Technologies in Education, Research, and Industrial Applications* (pp. 415-439). New York: Springer. [https://doi.org/10.1007/978-3-030-39459-2\\_19](https://doi.org/10.1007/978-3-030-39459-2_19)
- Kolupaieva, I., & Tiesheva, L. (2023). Asymmetry and convergence in the development of digital technologies in the EU countries. *Quarterly Journal of Economics and Economic Policy*, 18(3), 687–716. <https://doi.org/10.24136/eq.2023.022>
- Konstantinidis, F. K., Myrillas, N., Tsintotas, K. A., Mouroutsos, S. G., & Gasteratos, A. (2023). A technology maturity assessment framework for Industry 5.0 machine vision systems based on systematic literature review in automotive manufacturing. *International Journal of Production Research*, 1–37. <https://doi.org/10.1080/00207543.2023.2270588>
- Kovacic, Z. J., & Vukmirovic, D. (2008). ICT adoption and the digital divide in Serbia: Factors and policy implications. *Proc. Inf. Sci. IT Educ. Conf. (InSITE)*. <https://doi.org/10.28945/3247>
- Kraftova, I., Mateja, Z., & Prasilova, P. (2011). Economic performance: Variability of businesses within each industry and among industries. *Inzinerine Ekonomika-Engineering Economics*, 22(5), 459–467. <https://doi.org/10.5755/j01.ee.22.5.964>
- Kumar, V., Vrat, P., & Shankar, R. (2024). MCDM model to rank the performance outcomes in the implementation of Industry 4.0. *Benchmarking: An International Journal*, 31(5), 1453-1491. <https://doi.org/10.1108/BIJ-04-2022-0273>
- Kuruczleki, E., Pelle, A., Laczi, R., & Fekete, B. (2016). The readiness of the European Union to embrace the fourth industrial revolution. *Management*, 11(4), 327-347.
- Larichev, O. I. (1999). Normative and Descriptive aspects of decision making. In: *Multi-criteria Decision Making, Advances in MCDM Models: Algorithms, Theory and Applications: 5.1–5.24*, edited by T. Gal, T. Stewart and T. Hanne. Kluwer Academic Publishing, Boston, USA.
- Lee, I., & Lee, K. (2015). The IoT: Applications, investments, and challenges for enterprises. *Business Horizons*, 58, 431-440. <https://doi.org/10.1016/j.bushor.2015.03.008>
- Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3, 18-23. <https://doi.org/10.1016/j.mfglet.2014.12.001>
- Liu, Y., Li, Z., Xiong, H., Gao, X., & Wu, J. (2010). Understanding of internal clustering validation measures. *IEEE International Conference on Data Mining*, 911-916. <https://doi.org/10.1109/ICDM.2010.35>
- Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. *Journal of Industrial Information Integration*, 6, 1-10. <https://doi.org/10.1016/j.jii.2017.04.005>
- Mahdiloo, M., Andargoli, A. E., Toloo, M., Harvie, C., & Duong T. T. (2023). Measuring the digital divide: A modified benefit-of-the-doubt approach. *Knowledge-Based Systems*, 261, 110191. <https://doi.org/10.1016/j.knosys.2022.110191>
- Mariani, M., & Borghi, M. (2019). Industry 4.0: A bibliometric review of its managerial intellectual structure and potential evolution in the service industries. *Technological Forecasting and Social Change*, 149, 1-24. <https://doi.org/10.1016/j.techfore.2019.119752>
- Marti, L., & Puertas, R. (2023). Analysis of European competitiveness based on its innovative capacity and digitalization level. *Technology in Society*, 72, 102206, 1-10. <https://doi.org/10.1016/j.techsoc.2023.102206>
- Ngo, T., Kashani, A., Imbalzano, G., Nguyen, K., & Hui, D. (2018). Additive manufacturing (3D printing): A review of materials, methods, applications and challenges. *Composites Part B: Engineering*, 143, 172-196. <https://doi.org/10.1016/j.compositesb.2018.02.012>
- OECD. (2001). Understanding the digital divide. OECD Digital Economy Papers, 49, Paris: OECD Publishing. <https://doi.org/10.1787/236405667766>.
- Pacchini, A. P. T., Lucato, W. C., Facchini, F., & Mummolo, G. (2019). The degree of readiness for the implementation of Industry 4.0. *Computers in Industry*, 113, 103125, 1-8. <https://doi.org/10.1016/j.compind.2019.103125>

- Painsky, A., & Rosset, S. (2017). Cross-validated variable selection in tree-based methods improves predictive performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(39), 2142-2153. <https://doi.org/10.1109/TPAMI.2016.2636831>
- Patel, K., & Patel, S. (2016). Internet of Things-IoT: Definition, characteristics, architecture, enabling technologies. *Application & Future Challenges International Journal of Engineering Science and Computing*, 69(5), 6122-6131
- Pauli, T., Fielt, E., & Matzner, M. (2021). Digital Industrial Platforms. *Business and Information Systems Engineering*, 63(2), 181–190. <https://doi.org/10.1007/s12599-020-00681-w>
- Pavel, E. V., Kudryashova, T. V., & Bykova, P. A. (2021). Research on the Digital Maturity of Mechanical Engineering Companies in Russia. In : International Scientific and Practical Conference “Russia 2020-a new reality: economy and society”(ISPCR 2020), Atlantis Press, pp. 408–411. <https://doi.org/10.2991/aebmr.k.210222.081>
- Peres, R. S., Jia, X., Lee, J., Sun, K., Colombo, A. W., & Barata, J. (2020). Industrial artificial intelligence in Industry 4.0 - Systematic review, challenges and outlook. *IEEE Access*, 8, 220121-220139. <https://doi.org/10.1109/ACCESS.2020.3042874>
- Peters, G., Crespo, F., Lingras, P., & Weber, R. (2013). Soft clustering – Fuzzy and rough approaches and their extensions and derivatives. *International Journal of Approximate Reasoning*, 54(2), 307-322. <https://doi.org/10.1016/j.ijar.2012.10.003>
- Polozova, T., Kolupaieva, I., & Sheiko, I. (2021). Digital gap in EU Countries and its impact on labour productivity and global competitiveness. *Hradec Economic Days*, 11(1), 659-670. <https://doi.org/10.36689/uhk/hed/2021-01-065>
- Porter, M. E. (1985). Technology and competitive advantage. *Journal of Business Strategy*, 5(3), 60–78. <https://doi.org/10.1108/eb039075>
- Reynolds, D. (1998). Productivity analysis: In the onsite food-service segment. *Cornell Hotel and Restaurant Administration Quarterly*, 39(3), 22-31. <https://doi.org/10.1177/001088049803900307>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Salkin, C., Oner, M., Ustundag, A., & Cevikcan, E. (2018). A conceptual framework for Industry 4.0. In A. Ustundag and E. Cevikcan (Ed.), *Industry 4.0: Managing the Digital Transformation* (pp. 3-23), New York: Springer. [https://doi.org/10.1007/978-3-319-57870-5\\_1](https://doi.org/10.1007/978-3-319-57870-5_1)
- Samaranayake, P., Ramanathan, K., & Laosirihongthong, T. (2017). Implementing industry 4.0—A technological readiness perspective. *Proceedings of the 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Singapore (pp. 529–533). IEEE, December. <https://doi.org/10.1109/IEEM.2017.8289947>
- Santos, R. C. & Martinho, J. L. (2020). An industry 4.0 maturity model proposal. *Journal of Manufacturing Technology Management*, 31(5), 1023-1043. <https://doi.org/10.1108/JMTM-09-2018-0284>
- Saueremann, J. (2016). Performance measures and worker productivity. *IZA World of Labor*, 260. Doi: 10.15185/izawol.260
- Sauter, R., Bode, M., & Kittelberger, D. (2015). How industry 4.0 is changing. How we manage value creation. Available from internet: [https://www.horvath-partners.com/fileadmin/horvath-partners.com/assets/05\\_Media\\_Center/PDFs/englisch/Industry\\_4.0\\_EN\\_web-g.pdf](https://www.horvath-partners.com/fileadmin/horvath-partners.com/assets/05_Media_Center/PDFs/englisch/Industry_4.0_EN_web-g.pdf)
- Scandurra, F., Salomone, R., Caeiro, S., & Gulotta, T. M. (2023). The maturity level of the agri-food sector in the circular economy domain: A systematic literature review. *Environmental Impact Assessment Review*, 100, 107079. <https://doi.org/10.1016/j.eiar.2023.107079>
- Schroeder, A., Ziaee A., B., Galera C., Z., & Baines, T. (2019). Capturing the benefits of Industry 4.0: A business network perspective. *Production Planning & Control*, 30(16), 1305-1321. <https://doi.org/10.1080/09537287.2019.1612111>
- Senna, P. P., Barros, A. C., Roca, J. B., & Azevedo, A. (2023). Development of a digital maturity model for Industry 4.0 based on the technology-organization-environment framework. *Computers & Industrial Engineering*, Volume 185,109645, 1-19. <https://doi.org/10.1016/j.cie.2023.109645>
- Seyedan, M., & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data*, 7(53), 1-22. <https://doi.org/10.1186/s40537-020-00329-2>
- Sharma, M., Gupta, R., & Acharya, P. (2020). Analysing the adoption of cloud computing service: A systematic literature review. *Global Knowledge, Memory & Communication*, 70(1/2), 114–153. <https://doi.org/10.1108/GKMC-10-2019-0126>
- Shaygan, A., & Daim, T. (2023). Technology management maturity assessment model in healthcare research centers. *Technovation*, 120, 102444. <https://doi.org/10.1016/j.technovation.2021.102444>
- Shi, Y. (2022). Big data and Big data analytics. In Y. Shi (Ed.), *Advances in Big Data Analytics* (pp. 3-21), Singapore: Springer. [https://doi.org/10.1007/978-981-16-3607-3\\_1](https://doi.org/10.1007/978-981-16-3607-3_1)

- Shirdastian, H., Laroche, M., & Richard, M. O. (2019). Using big data analytics to study brand authenticity sentiments: The case of Starbucks on Twitter. *International Journal of Information Management*, 48, 291-307. <https://doi.org/10.1016/j.ijinfomgt.2017.09.007>
- Sony, M., & Naik, S. (2020). Key ingredients for evaluating Industry 4.0 readiness for organizations: A literature review. *Benchmarking: An International Journal*, 27(7), 2213-2232. <https://doi.org/10.1108/BIJ-09-2018-0284>
- Stiftung, B. (2020). Transformation Index BTI 2020: Governance in International Comparison, Verlag Bertelsmann Stiftung. Available from internet: <https://www.bti-project.org/en/index/political-transformation.html>
- Sunyaev, A. (2020). Cloud Computing. In: Internet Computing (pp. 195-236), Switzerland: Springer. [https://doi.org/10.1007/978-3-030-34957-8\\_7](https://doi.org/10.1007/978-3-030-34957-8_7)
- Thordsen, T., & Bick, M. (2023). A decade of digital maturity models: much ado about nothing?. *Information Systems and E-Business Management*, 21, 947-976. <https://doi.org/10.1007/s10257-023-00656-w>
- Tutak, M., & Brodny, J. (2022). Business Digital Maturity in Europe and Its Implication for Open Innovation. *Journal of Open Innovation: Technology, Market, and Complexity*, 8(1), 27. <https://doi.org/10.3390/joitmc8010027>
- UN. (2019). Structural transformation, Industry 4.0 and inequality: Science, technology and innovation policy challenges, United Nations Conference on Trade and Development TD/B/C.II/43. Available from internet: [https://unctad.org/system/files/official-document/ciid43\\_en.pdf](https://unctad.org/system/files/official-document/ciid43_en.pdf)
- UNIDO. (2021). Entering a New Digital Era. Available online: <https://www.unido.org/unido-industry-40> (accessed on 15 October 2024).
- Wang, L., Torngren, M., & Onori, M. (2015). Current status and advancement of cyber-physical systems in manufacturing. *Journal of Manufacturing Systems*, 37, 517-527. <https://doi.org/10.1016/j.jmsy.2015.04.008>
- Wang, S., Wan, J., Li, D., & Zhang, C. (2016). Implementing smart factory of Industrie 4.0: An Outlook. *International Journal of Distributed Sensor Networks*, 12 (1), 3159805, 1-10. <https://doi.org/10.1155/2016/3159805>
- Xu, Z., Frankwick, G. L., & Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research*, 69(5), 1562-1566. <https://doi.org/10.1016/j.jbusres.2015.10.017>
- Xun, X. (2012). From cloud computing to cloud manufacturing. *Robotics and Computer-Integrated Manufacturing*, 28(1), 75-86. <https://doi.org/10.1016/j.rcim.2011.07.002>
- Yakymova, L., Novotna, A., Kuz, V., & Tamandl, L. (2022). Measuring industry digital transformation with a composite indicator: A case study of the utility industry. *Journal of International Studies*, 15(1), 168-180. <https://doi.org/10.14254/2071-8330.2022/15-1/11>

### Author's Biography

**Huseyin Fidan** is an Associate Professor in Department of Industrial Engineering in Burdur Mehmet Akif Ersoy University (Turkiye). His researches are focused on technology management, Industry 4.0, machine learning and data mining. His research interest has mainly focused on clustering, decision support systems and data analytics. He is also interested in software and studied on new machine learning algorithms.

The article has been reviewed.

Received in January 2024; accepted in March 2026.



This article is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 (CC BY 4.0) License <http://creativecommons.org/licenses/by/4.0>