Acta Infologica,
ACIN 2025, 9 (1): 293–313

https://doi.org/10.26650/acin.1626593
Submitted: 24.01.2025
Revision Requested: 01.05.2025
Last Revision Received: 25.05.2025
Accepted: 05.06.2025
Published Online 13.06.2025

Acta Infologica

Research Article Open Access

Predicting Industry Maturity Index Using Machine Learning Methods

Ayhan Doğan [1] & Cihan Ünal [1]

[1] Hacettepe University, Başkent Organized Industrial Zone, Technical Sciences Vocational School, Ankara, Türkiye

Abstract

Industry 4.0 has become a widely adopted concept in recent years. Maturity and readiness models are commonly used to assess the current state of industrial organizations in relation to Industry 4.0. Companies' maturity levels and index scores are typically determined through structured surveys. However, due to their complexity, time consumption, and high cost, many enterprises lack formally assessed maturity index (MI) scores. To address this limitation, this study initially employed survey data to evaluate the accuracy of the proposed machine learning (ML) framework. A 58-question survey was conducted to calculate the MI scores of the companies. These scores were then used as reference values to be predicted based on five easily accessible enterprise-level variables: company age, industry type, ownership structure, number of employees, and annual turnover. This approach tested whether MI could be accurately predicted without relying on lengthy survey processes, using only a minimal set of key enterprise attributes. The results of this study demonstrate that MI can be estimated successfully using ML techniques without the need for answering long and complex surveys. To reduce the burdens associated with conventional survey-based methods, this study employed multiple ML algorithms, including Support Vector Machines (SVM), Gaussian Process Regression (GPR), Linear Regression (LR), Regression Trees (RT), and Ensemble Tree-based models, and advanced boosting-based methods, such as extreme gradient boosting (XGB) and Light Gradient Boosting Machine (LGBM). The findings demonstrate that the proposed model predicts MI with high accuracy and offers a practical and scalable alternative for enterprises seeking to assess their Industry 4.0 readiness.

Keywords Machine learning · Industry 4.0 · Maturity model · Maturity index

Citation: Doğan, A. & Ünal, C. (2025). Predicting industry maturity index using machine learning methods. *Acta Infologica*, 9(1), 293-313. https://doi.org/10.26650/acin.1626593

Corresponding author: Ayhan Doğan aydogan@hacettepe.edu.tr

Introduction

Industry 4.0 was introduced as a concept at a fair in Germany in 2011. It was used as the name of the state policy used by the German government for digital transformation. According to this fiction, every revolutionary change in the industry is divided into phases: Industry 1.0, Industry 2.0, Industry 3.0, and Industry 4.0. Industry 1.0 represents the invention of steam engines, Industry 2.0 represents electricity and subsequent production, Industry 3.0 represents the beginning of the digital age, and Industry 4.0 represents and symbolizes the fully autonomous production approach.

Industry 4.0 has gained autonomous production capability thanks to many technologies. These include cloud computing, the Internet of Things, artificial intelligence, autonomous robots, sensors, big data, and 3D printers. This way, interaction based on the network, equipment, and data exchange that connect people and machines are established.

Industry 4.0 significantly contributes to businesses thanks to its technology-based production (Saad et al., 2021). However, companies must prepare for Industry 4.0 to adapt to the fourth industrial revolution (Hizam-Hanafiah et al., 2020). Competitiveness depends on efficiency, flexibility, and rapid production. Digitalization and technological transformation are essential (Çınar et al., 2021; Rahamaddulla et al., 2021). However, small businesses may not have the resources and equipment necessary for digital transformation, making competition more difficult (Jesus & Lima, 2020).

Industry 4.0 is a concept that proposes the use of new technologies to increase performance and efficiency (Chonsawat & Sopadang, 2020). Businesses should conduct internal evaluations, examine their competitors, and perform planning to determine Industry 4.0 levels and needs. They require maturity models (MMS) to realize all these processes (Carvalho et al., 2018). While applying for MMS, interviews are held with companies, and various surveys are conducted. Because of these surveys, companies' preparation level for Industry 4.0 can be determined. Appropriate studies should be carried out considering the maturity level of the companies under Industry 4.0. However, companies must be prepared and willing to allocate sufficient resources for this process.

The Industry 4.0 level of many companies is unknown due to reasons such as the time required to conduct surveys and high costs. In this study, the success of machine learning (ML) methods was investigated to make the industrial maturity index (MI) assessments of many enterprises located in industrial zones faster, more economical, and more accessible.

Background and Literature

Candanedo et al. (2018) conducted an extreme case study using the heating, ventilation, and air conditioning systems (HVAC) dataset. Kuo et al. (2017) explored using inexpensive triaxial sensors to monitor machines. To extract meaningful insights from the collected data, they developed a dimension reduction method with a low computational load and a neural network that allows the obtained data to be analyzed automatically (Kuo et al., 2017). Angelopoulos et al. (2019) presented ML solutions related to Industry 4.0 and classified them according to their learning processes. On the other hand, Rai et al. (2021) sought the work of a wide variety of researchers to understand the change that began with the use of ML techniques and to report current research on the fundamental theoretical and experimental aspects of ML and their applications in production and production systems. Brik et al. (2019) recommended a tool for Industry 4.0 to monitor system outages. They used an ML algorithm to develop a predictive model of resource localization, considering the actual task timing in terms of resource localization. Lee and Lim (2021) analyzed numerous journal articles by text mining using unsupervised ML algorithms.

MMS is an evolutionary way to achieve a goal that takes shape in stages. This path guides the reorganization and restructuring of existing abilities to achieve perfection (Finance, 2015). Within the scope of this study, existing maturity and readiness models were examined. Accordingly, the number of dimensions of the existing models is between 3 and 13. Some measures are not preferred for a good analysis. However, in some sectors and exceptional cases, the small number of measurements is not a problem. Although some dimension names are used differently, the evaluation serves the same purpose. When the existing maturity and preparation models are examined, it can be seen that the maturity level (MTL) of the enterprise is between 4 and 6.

The Industry 4.0 revolution can contribute to small and medium-sized enterprises (SMEs), one of the cornerstones of growth and development (Chonsawat & Sopadang, 2020). The readiness of businesses is a measure of the extent to which they benefit from the new technologies that come with Industry 4.0 (Stentoft et al., 2021). In other words, Industry 4.0 preparation concerns companies preparing to use Industry 4.0 technologies (Rais, 2021; Vazire, 2018). Insufficient understanding of concepts and advantages by businesses constitutes a significant problem (Da Silva et al., 2020; Horváth & Szabó, 2019; Rauch et al., 2019).

During the digital transformation phase, organizations' software and hardware requirements also emerge (Haber et al., 2015; Wank et al., 2016). Achieving Industry 4.0 preparation is essential for businesses today (Schaupp et al., 2017). Assume digital transformation is not well managed. Perhaps, it may be faced with the fact that companies that do not attach sufficient importance to digitalization will disappear from the market (Canetta et al., 2018; Ivanov et al., 2019). MMS is used to measure and evaluate the capabilities required by businesses to reach the desired level (Schumacher et al., 2016). It is also used to see the company's progress over time and compare it with that of its competitors (Pessl et al., 2017).

Industry 4.0 Readiness and maturity models

As the concept of 'Industry 4.0' is new, readiness and maturity assessment tools continue to evolve. In this study, 30 different Industry 4.0 readiness and MMS approaches were examined. Experts in the industry developed nine (30%) of the models inspected, while academics developed the remaining 21 (70%). The list of readiness and MMS is presented in Table 1.

Table 1

Widely used readiness and MMS

No.	Model Name	Reference
1	"Industry 4.0 Readiness Evaluation for Manufacturing Enterprises"	(Basl and Doucek, 2019)
2	"Industry 4.0 Maturity Model"	(Bibby and Dehe, 2018)
3	"Future Readiness Level (FRL)/Industry 4.0 Future Readiness"	(Botha, 2018)
4	"E-Business Industry 4.0 Readiness Model"	(Városiné Demeter et al., 2018)
5	"Benchmarking Readiness I4.0"	"Fraunhofer Institute for Systems and Innovation"
6	"SMEs Maturity Model Assessment of IR4.0 Digital Transformation"	(Hamidi et al., 2018)
7	"Readiness for Industry 4.0"	(Horvat et al., 2018)
8	"SSCM Assessment for Industry 4.0"	(Manavalan and Jayakrishna, 2019)

No.	Model Name	Reference
9	"Industry 4.0 Business Model Innovations Tool"	(Müller and Voigt, 2018)
10	"Industry 4.0 Maturity Model"	(Geissbauer et al., 2016)
11	"Manufacturing Companies Industry 4.0 Adoption Model"	(Mittal et al., 2018)
12	"BMS Smart Industry Research Roadmap (Behavioral, Management, Social Sciences)- SIRM"	"University of Twente"
13	"ACATECH Industrie 4.0 Maturity Index"	"Acatech Academy"
14	"Enterprise 4.0 Assessment"	(Valentin, 2017)
15	"Industry 4.0 Maturity Model- SPICE (Software Process Improvement and Capability Determination)"	(Gökalp et al., 2017)
16	"Industry 4.0 Readiness Model for Tool Management"	(Schaupp et al., 2017)
17	"Three Stages Maturity Model in SMEs toward Industry 4.0"	(Erol et al., 2016)
18	"Design Business Modeling for Industry 4.0"	(Gerlitz, 2016)
19	"SIMMI 4.0–System Integration Maturity Model Industry 4.0"	(Leyh et al., 2016)
20	"Industry 4.0 Introduction Strategy"	"Merz Consulting"
21	"Roadmap Industry 4.0"	(Pessl et al., 2017)
22	"Assessment Model for Organizational Adoption of Industry 4.0 Based on Multi-criteria Decision Techniques"	"University of Warwick"
23	"Industry 4.0 Maturity Model"	(Akdil et al., 2018)
24	"Reference Architecture Model for the Industry 4.0 (RAMI4.0)"	(Kannan et al., 2017)
25	"Industry 4.0 Hindering Factors Model"	(Geissbauer et al., 2016)
26	"IMPULS—Industrie 4.0 Readiness"	"Verband Deutscher Maschinen- und Anlagenbau (VDMA)"
27	"Industry 4.0 Barometer"	"MHP Porsche Company"
28	"Roland Berger Industry 4.0 Readiness Index"	"Ronald Berger Consulting"
29	"Fraunhofer Industrie 4.0 Layer Model"	(Geissbauer et al., 2016)
30	"Industry 4.0 Readiness Model for Manufacturing"	(Methavitakul and Santiteerakul, 2018)
31	"Lean modified manufacturing maturity model for Industry 4.0 (LM4I4.0)"	(Sajjad et al., 2024)
32	"Maturity Model-ADIME 4.0"	(Skalli et al., 2023)
33	"Maturity SCAN 4.0"	(Muñoz et al., 2023)

Machine Learning

In a simple definition, ML involves learning from data methods developed for decision-making processes (Jalil et al., 2019). The goal is to find patterns in the data and make better decisions for the algorithm.

ML, a sub-branch of artificial intelligence, is an algorithm that produces predictions about the problem's solution using data about the given situation. Training models were created for these processes. Although various ML methods require various amounts of data, they often require large amounts of data to continually optimize models and make the best predictions (Wei et al., 2019). several algorithms have been developed for ML. Commonly used are linear regression (LR), ensemble trees (ET), gauss process regression (GPR), decision trees (DT), random forest (RF), support vector machines (SVM), regression trees (RT), and artificial neural networks (ANN). ML is used for different purposes, such as clustering, regression, and classification.

The process begins with the collection of the necessary data. The required data are converted into numerical values. Normalization and standardization processes were applied to the data. All data were used in two groups: training and testing.

This study used seven different methods: ET, SVM, GPR, LR, RT, XGB, and LGBM. Estimates were made on the test data using the models obtained from the training. The prediction accuracy can be measured using various metrics. In this study, root mean square error (RMSE), mean square error (MSE), mean absolute error (MAE), and coefficient of determination (R2) metrics were used.

The k-fold cross-validation method was used to evaluate the prediction results of ML. The reason for preferring this method is the relatively small amount of data. The proposed method is suitable for assessing model success because all data are used alternately in training and testing. However, this method requires relatively more computation because training and testing are performed multiple times. In the k-fold cross-validation method, all data are divided into k layers. A layer is used for testing, while the other is used for training. The most preferred number of layers is 5. Too few or too many layers can degrade the performance of the proposed method. In this study, the number of layers was taken as 5.

Linear regression

LR is used to determine the relationship between the dependent variable and the independent variable. This method is generally preferred when a linear relationship exists between inputs and outputs (Luu et al., 2021). Equation 1 defines the LR method.

$$y = a_0 + a_1 X + \varepsilon \tag{1}$$

In Equation 1, ε is the error term, X is the independent variable, and y is the dependent variable. In addition, a_0 is the constant term, and a1 is the regression coefficient showing the slope. If the value of a1 is positive, then the change is increasing, and if the value of a1 is negative, then the difference is decreasing (Bayazit & Oğuz, 1994; Doğan et al., 2023).

The LR method uses one or more arguments. If there is more than one independent variable, the method is called multiple linear regression (MLR) (Memnun & Kalaycı, 2006; Rong & Bao-Wen, 2018). The formula for MLR methods is given in Equation 2.

$$y = a_0 + a_1 X_1 + a_2 X_2 + \cdots + a_n X_n + \varepsilon \tag{2}$$

The values a1, a2, ... an in Equation 2 are multiple regression coefficients. X1, X2, ... Xn values are multiple independent variables (Aslan et al., 2011).

Support vector machines

In the early 1990s, Vapnik developed SVM (Cortes & Vapnik, 1995). The SVM algorithm does not require knowledge of the combined distribution function (Soman et al., 2009). It aims to find the optimal hyperplane to separate classes from each other (Ayhan & Erdoğmuş, 2014; Doğan et al., 2023). SVMs use neural networks and statistical methods together (Haykin, 1999; Tolun, 2008). It is used in various fields, such as identification and classification (Cristianini & Shawe-Taylor, 2000; Schölkopf et al., 1999).

SVM can also work with linear, nonlinear, and multi-class data (Kavzoğlu & Çölkesen, 2010). The equations for the optimal hyperplane are given in 3 and 4.

$$w.x_i + b \geq +1, y = +1 \tag{3}$$

$$w.x_i + b \leq +1, y = -1 \tag{4}$$

In these equations, $x \in RN$ represents the N-dimensional space, $y \in (-1,+1)$ class labels, b the trend value, and w the weight vector (Hearst et al., 1998; Huang et al., 2018). When the data are not linearly separated, the editing parameter and artificial variable are used to classify the data correctly (Doğan, 2024; Kavzoğlu & Çölkesen, 2010).

Gaussian process regression

The GPR method is used to identify unknown functions and to learn and optimize them effectively (Becker et al., 2018). A successful ML method is used primarily in probabilistic, non-parametric problems to produce predictions by solving nonlinear problems (Liu et al., 2019; Mehmet & Doğansoy, 2022). It can also produce successful results with little data (Ateş, 2020; Yesiloglu-Gultekin & Dogan, 2024). Different covariance functions can also be used (Heo & Zavala, 2012). Equation 5 gives the Gaussian process function f(x) (Boran, 2021; Rasmussen and Williams, 2006).

$$f(x) = GP(m(x), k(x.')) \tag{5}$$

In equation 5, m(x) is the mean function, and k(x. x') is the covariance function. The mean function and covariance function are given in Equations 6 and 7, respectively.

$$m(x) = E[f(x)] \tag{6}$$

$$k(x, x') = E[(f(x) - m(x))(f(x') - m(x'))] \tag{7}$$

Regression trees

The RT method is not parametric. RT aims to categorize members of the community into homogeneous subclasses. The basis of this approach is the fact that individuals who are similar are gathered in the same node (Yesiloglu-Gultekin & Dogan, 2024; Yücel, 2017). Nodes or branches with high error values are removed. This process is called pruning (Mendeş & Akkartal, 2009). Estimates for leaves correlate with weighted averages calculated at nodes (Maimon & Rokach, 2005).

Three separation rules apply to RT. The least squares (LS), Least Absolute Deviation (LAD), and Clark & Pregibon (CP) equations were used. The primary goal of homogeneous nodes is (Temel, 2004). The RT method divides arguments into separate rather than continuous intervals.

Ensemble of trees

ET is a step-by-step process that leverages the information from many trees. Each tree was grown using data from previous trees. Fitting to a regression tree was performed using the initial data. The model is constantly updated using the last residuals (Schiltz et al., 2018). Multiple classifiers are widely used. Many studies have demonstrated that multiple classifiers perform better than others (Breiman, 1996; Opitz & Maclin, 1999). Some classifiers may behave erratically. Ensemble classifiers may provide better results (Zhou et al., 2004). RF is an excellent example of an ensemble classifier (Breiman, 2001). It performs well in classification and regression tasks (Katuwal et al., 2018).

Ensemble methods improve prediction success in classification problems. Examples of these methods are RF and gradient boosting. DT, another successful way, are also widely used. The proposed method works faster than neural network classifiers and requires fewer parameters (Shoaran et al., 2018).

Extreme gradient boosting

The XGB method is an advanced ensemble learning method based on the gradient boosting algorithm that is optimized for both speed and performance. The model builds decision trees sequentially, where each new tree attempts to correct the errors of the previous tree. XGB integrates L1 and L2 regularization techniques to prevent overfitting and improve generalization (Chen & Guestrin, 2016). In this study, XGB was employed to evaluate whether MI could be accurately predicted using only five basic enterprise-level variables. The results demonstrate that the method is capable of effectively capturing nonlinear relationships even within relatively small datasets.

Light gradient boosting machine

LGBM is a modern gradient boosting algorithm developed by Microsoft that provides high efficiency and scalability, particularly when working with large datasets. Unlike traditional level-wise tree growth strategies, LGBM uses a leaf-wise growth approach, which allows the model to minimize loss more rapidly (Ke et al., 2017). It also supports parallel computation and is optimized for handling categorical variables. In this study, LGBM was applied to test its ability to predict MI scores with high accuracy and low computational cost.

Algorithm Design

The framework proposed in this study enables the prompt determination of enterprise maturity levels without the need for conducting a survey, resulting in significant time and cost savings.

In this study, a 58-question questionnaire was administered to enterprises, and the MI score was calculated using the data obtained from the questionnaire. The MI score was then calculated using five basic data (age, field of activity, capital size, number of employees and turnover) that can be obtained directly from enterprises without requiring a questionnaire. The MI score results obtained by these two methods are compared, and the success of the proposed method is demonstrated using performance metrics. The details of the implementation steps of this new framework are described in the following algorithm.

Step 1: Preliminary preparation

a) In the preparation phase, a questionnaire is given to gather information about the enterprises.
b) Next, a maturity model (MM) is selected, and the AHP method is used to determine the weights of its dimensions.
c) Finally, the results are used to calculate enterprises' maturity index (MI) scores.

Step 2: Machine learning model training

a) Data on enterprise age, field of activity, capital size, number of employees, and turnover are obtained from the survey and translated into numerical values for input data.
b) The data undergo standardization and normalization processes to be used in machine learning.
c) Training sessions are conducted, and models are generated using various machine learning methods.

Step 3) Machine learning test process

a) The data adapted for machine learning are used as input data. MI scores of each enterprise are used as output data. Model testing was performed using the input and output data. For this purpose, the k-fold cross-validation method was used.
b) The test results were evaluated using performance metrics.

Step 4) Predicting MI score for new business

a) Basic input data on the age of the enterprise, field of activity, capital size, number of employees, and turnover are obtained directly from the enterprise. Because enterprises already know this basic information, they do not need a survey.

b) The standardization and normalization processes were applied to the data for use in machine learning.

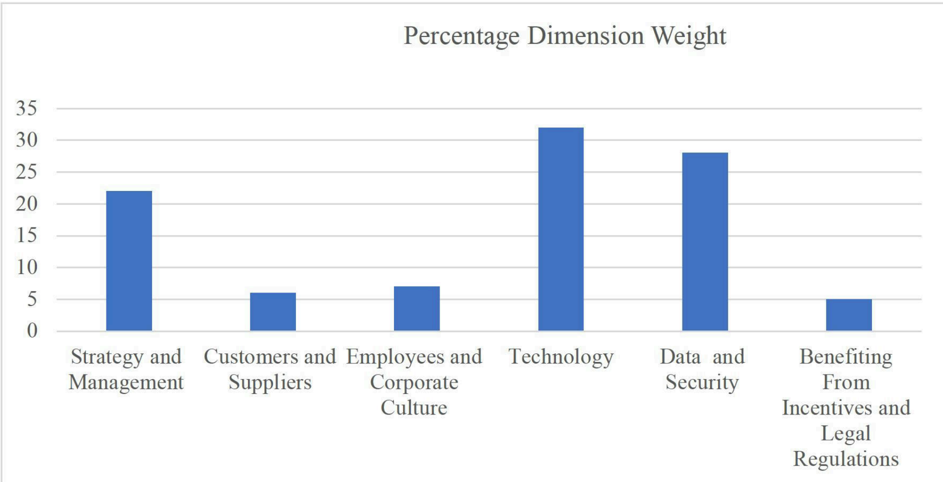
c) Using these data and models previously trained with machine learning methods, the MI score of the enterprise is estimated.

Implementation of the Proposed Method

The SANOL maturity model (SMM) was chosen in this study to test the proposed methodology on a sample SMM. The model measures maturity in six dimensions for Industry 4.0. These six dimensions and percentage weights are shown in Figure 1.

Figure 1

Percentage weight of dimensions



In the SMM, a questionnaire consisting of 58 questions was used as the evaluation method. The weights of the dimensions in the model are determined by the AHP method.

Since the study was initiated during the final stages of the COVID-19 pandemic, and considering the relative ease of both administration and data processing, the survey was conducted online. For this purpose, a dedicated digital form was developed to facilitate the online implementation of the questionnaire. The English version is accessible at https://form.jotform.com/251425426420953. However, no data were obtained from many businesses contacted, and the responses provided were incomplete. As a result, the study took longer than originally planned and was ultimately completed using questionnaires completed by 61 businesses.

The input and output data were prepared for use in machine learning studies. In this study, the machine learning models were trained using five easily accessible enterprise-level variables: company age, industry type, ownership structure, number of employees, and annual turnover. These variables were selected due to their wide availability, simplicity of collection, and conceptual relationship with organizational maturity. Table 2 provides a detailed definition of each variable along with its corresponding categorical groupings, which were used for model training and testing.

Table 2
Description and categorical breakdown of the input variables used in the model

Variable	Definition	Categories
Company Age	Number of years since the company officially began operations	1: 0–9 years 2: 10–19 years 3: 20 years or more
Industry Type	Primary economic sector in which the company operates	1: Electronics 2: Manufacturing 3: Services 4: Logistics 5: Machinery 6: Other
Ownership Structure	Ownership and partnership structure of the company	1: Domestic capital 2: Foreign capital 3: Joint domestic–foreign ownership
Number of Employees	Total number of personnel employed by the company	1: 1–9 (micro) 2: 10–50 (small) 3: 51–250 (medium) 4: 251–1000 5: More than 1000
Annual Turnover	Total annual revenue generated by the company	1: <1M (in the Turkish Lira, TRY) 2: 1–10 M 3: 10–50 M 4: 50–100 M 5: 100–250 M 6: 250–500 M 7: >500M

These variables served as the sole inputs to the machine learning models, replacing conventional survey-based indicators and enabling streamlined estimation of the Maturity Index (MI).

The output data is the MI value, which is calculated using SMM with the survey data. The MI score was obtained from the inputs of the six dimensions shown in Figure 1. Details are available in related publications (Ünal et al., 2022).

Of the 61 enterprises surveyed, 13 were electronics, 13 were manufacturing, 10 were service, nine were logistics, eight were machinery, and eight other enterprises. 26 enterprises were interviewed but could not be surveyed. Among the enterprises surveyed, 17 have more than 250 employees. When the annual turnover of enterprises was analyzed, 33% of enterprises had a turnover of less than 100 million TRY, 46% had a turnover between 100 million and 500 million TRY, and the rest had a turnover of more than 500 million TRY. Of the survey questions administered using the SMM, 12 are related to technology, 11 to strategy and management, ten to supports and incentives, seven to data and security, five to employees and corporate culture, three to customers and suppliers, and the remaining ten are concerned with general catalog information.

MI scores were calculated using enterprises' survey data. The MI score was calculated by adding the weighted dimension scores (Equation 8).

$$E = \Sigma_{i=0}^{n} Bi * gi \tag{8}$$

E: MI score (between 1 and 5).

B: Dimension maturity score (between 1 and 5).

g: Overall dimension weight.

Older enterprises have higher MI scores than newer ones. Considering the establishment dates of the enterprises, approximately 33% are less than 10 years old, 23% are between 10 and 19 years old, and the rest are 20 years old or older.

Seven machine learning methods were used in the studies. Table 3 provides the technical details and parameters of each of these methods.

Table 3

Machine learning methods, subtypes and their parameters

ML Methods	ML Sub Methods	Parameters and technical details
LR	Linear (Selected)	Standard
	Interactions Linear	Standard
	Robust Linear	Standard
	Stepwise Linear	Maximum number of steps:1000
SVM		Manual box constarint:0.769
		Manual epsilon: 0.077
	Linear SVM (Selected)	Manual kernel scale: 1
	Quadratic SVM	Manual kernel scale: 1
	Cubic SVM	Manual kernel scale: 1
	Fine Gaussian SVM	Manual kernel scale: 0.56
	Medium Gaussian SVM	Manual kernel scale: 2.2
	Coarse Gaussian SVM	Manual kernel scale: 8.9
GPR	Rational Quadratic GPR, Squared Exponential GPR (Selected), Matern 5/2 GPR, Exponential GPR	Kernel scale:1.1690659
	Signal standard deviation: 0.4553253	Sigma: 0.4553253
RT		Maximum surrogates per node:10
	Fine Tree	Minimum leaf size:4
	Medium Tree (Selected)	Minimum leaf size: 12
	Coarse Tree	Minimum leaf size:36
ET		Minimum leaf size:4
		Number of learners:30
	Boosted Trees (Selected)	Learning rate:0.1
	Bagged Trees	
XGB	Fine	n_estimators: 700
		learning_rate: 0.02

ML Methods	ML Sub Methods	Parameters and technical details
		max_depth: 4 subsample: 1.0 colsample_bytree: 1.0 gamma: 0.05 min_child_weight: reg_alpha: 0.5 reg_lambda: 1.5
LGBM	Fine	learning_rate=0.01 n_estimators=300 num_leaves=60 max_depth=8 subsample=0.8 feature_fraction=0.9 min_child_samples=15 reg_alpha=0.0 reg_lambda=0.0 lambda_l1=0.0 lambda_l2=0.05

Research Results and Discussion

Summary of Input Data and Model Objective

Within this study's scope, a survey was conducted to evaluate the industrial Maturity Index (MI). There are 58 questions in the survey. Because operating a study is very difficult and time-consuming, all the desired answers were obtained from 61 enterprises, and the industrial MI general score was calculated using SMM. In this study, the success of machine learning (ML) was investigated to make the industrial MI assessments of many enterprises located in industrial zones faster, cheaper, and more accessible. The enterprise age, field of activity, capital size, number of employees, and turnover information were used as inputs to the ML models, and the overall score was estimated as the output.

Evaluation Metrics

The metrics are the feedback of the designed model. The success of the model created according to these metrics was measured, and improvements are made when necessary (Özen et al., 2021).

The coefficient of determination, known as R^2, is a key metric used to evaluate the accuracy of regression models. The value ranges from 0 to 1 and indicates how well the predicted values match the actual data. A value closer to 1 implies a better fit.

The Root Mean Square Error (RMSE) measures the standard deviation of the prediction errors and reflects the magnitude of the model's errors in the same units as the output. The Mean Squared Error (MSE) was calculated as the average of the squared differences between actual and predicted values, emphasizing larger errors due to squaring. The Mean Absolute Error (MAE) computes the average of the absolute differences and provides a more intuitive interpretation of the model's average error per prediction. The coefficient of determination (R^2) shows the performance of the predicted values in the model. R^2 determines how well the data fit a curve. The RMSE returns the standard deviation of the actual and predicted values and measures the magnitude of a model's error. The MSE provides the arithmetic mean of the squared difference between the actual and predicted values. The MAE reflects the average of the absolute distance between the actual and predicted values.

Model Performance and Algorithm Comparison

The predictive performance of all the machine learning models was assessed by comparing their outputs with the results obtained from the field survey. To validate the models on a limited dataset, a 5-fold cross-validation approach was used. As summarized in Table 4, the models demonstrated strong performance with R^2 values generally greater than 0.70. The RMSE and MAE values remained within acceptable ranges, supporting the consistency and stability of the predictions.

Among the seven algorithms, SVM achieved the highest R^2 value of 0.75, followed closely by GPR (0.74), ET (0.73), LGBM (0.72), and XGB (0.71). These results indicate that both traditional and gradient boosting-based models can provide reliable estimations and highlight the robustness and flexibility of the proposed framework across diverse algorithmic strategies.

Table 4 and Figures 2–8 present the detailed performance metrics and visual comparisons of measured versus predicted scores. The performance evaluation revealed that all seven machine learning algorithms demonstrated notable predictive capabilities in estimating the Maturity Index (MI). Among them, SVM achieved the highest R^2 value of 0.75, followed closely by GPR (0.74), ET (0.73), LGBM (0.72), and XGB (0.71). These results indicate that both traditional and gradient boosting-based models can provide reliable estimations and highlight the robustness and flexibility of the proposed framework across diverse algorithmic strategies.

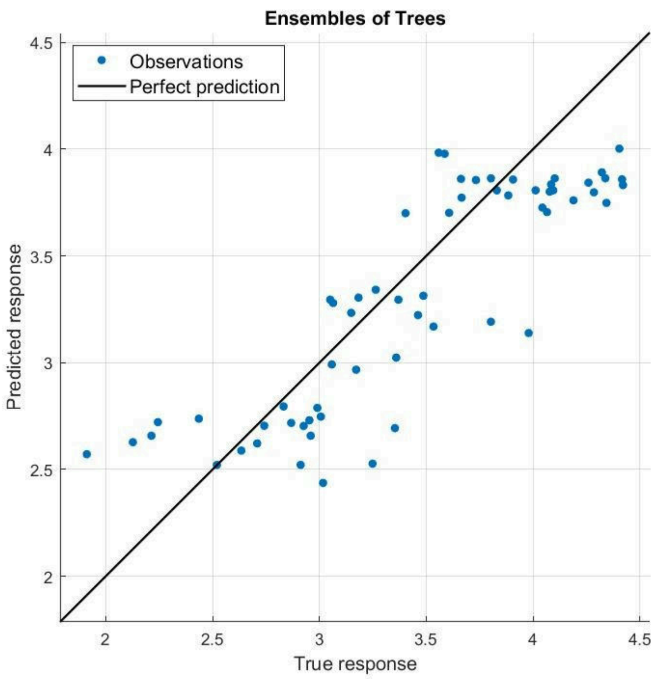
Table 4

Metrics of the test predictions

	LR	SVM	GPR	RT	ET	XGB	LGBM
R^2	0.74	0.75	0.74	0.69	0.70	0.71	0.72
RMSE	0.33	0.33	0.33	0.36	0.36	0.35	0.34
MSE	0.11	0.11	0.11	0.13	0.13	0.12	0.11
MAE	0.28	0.26	0.28	0.28	0.30	0.28	0.28

Figure 2

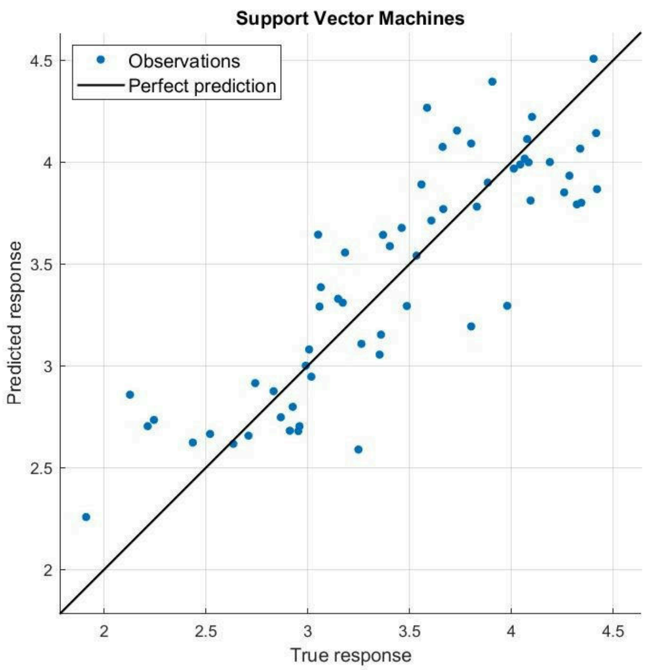
Index scores measured and predicted using the ET method.



The resulting graph of the tests obtained from the SVM studies is shown in Figure 3.

Figure 3

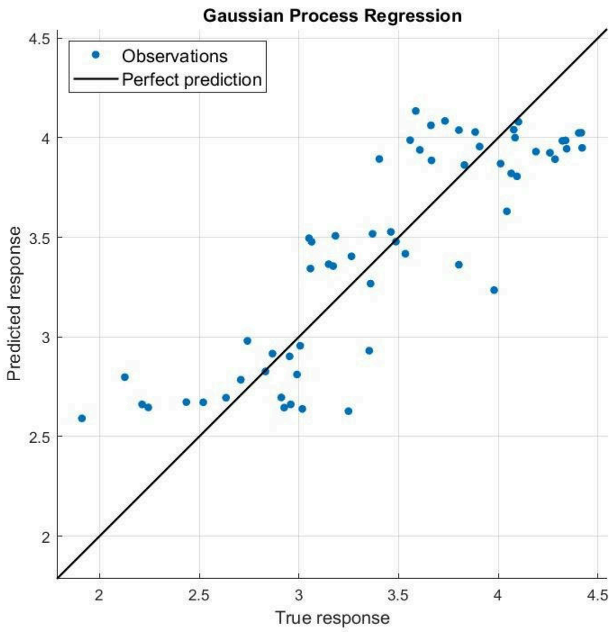
Index scores measured and predicted by the SVM method



The resulting graph of the tests obtained from the studies performed using the GPR is shown in Figure 4.

Figure 4

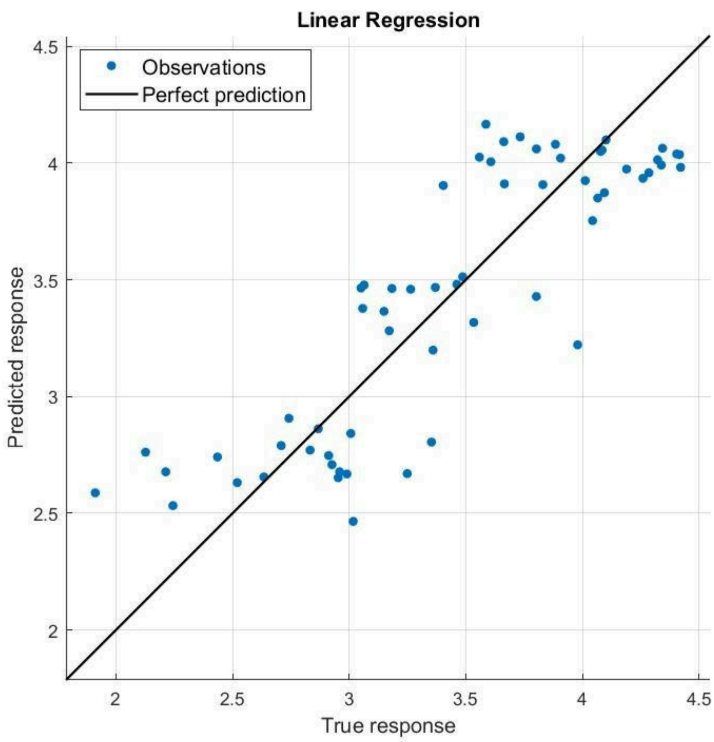
Index scores measured and predicted using the GPR method



The resulting graph of the tests obtained from the studies using LR is shown in Figure 5.

Figure 5

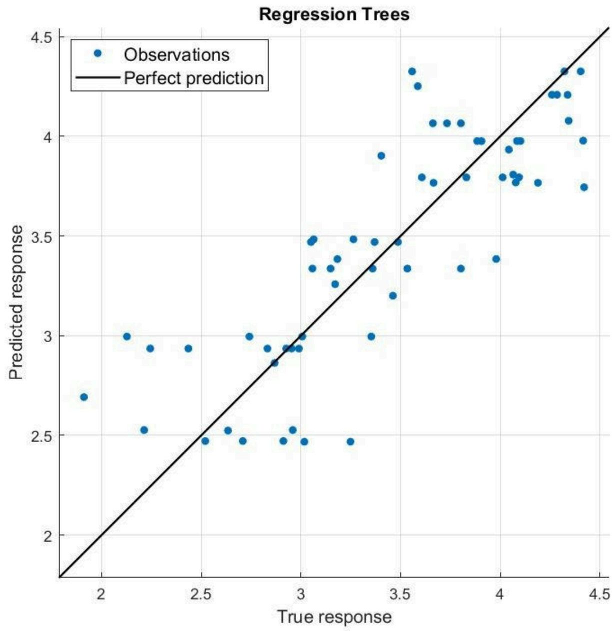
Index scores measured and predicted using the LR method



The resulting graph of the tests obtained from the studies performed with RT is shown in Figure 6.

Figure 6

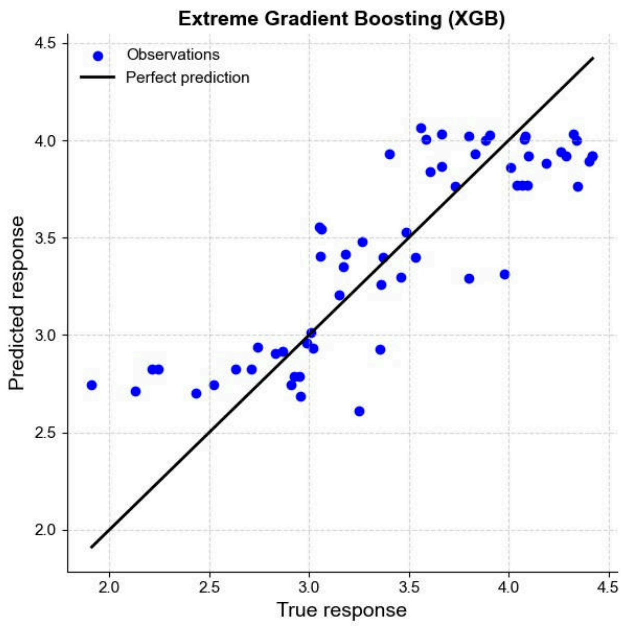
Index scores measured and predicted using the RT method



The resulting graph of the tests obtained from the studies performed with XGB is shown in Figure 7.

Figure 7

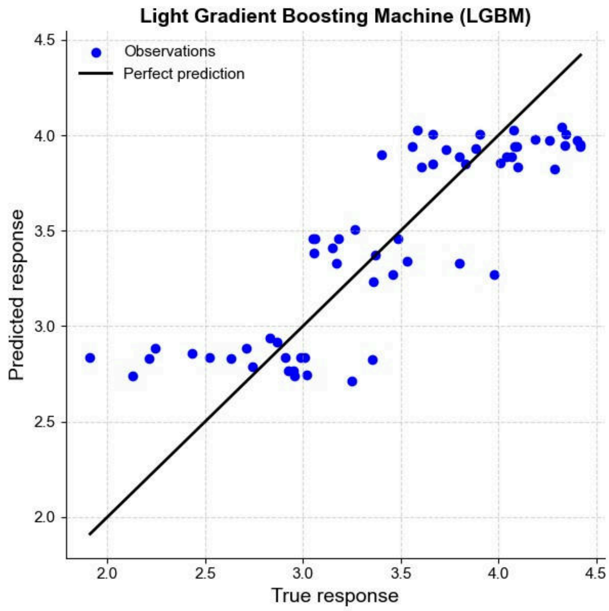
Index scores measured and predicted using the XGB method



The resulting graph of the tests obtained from the studies performed using LGBM is shown in Figure 8.

Figure 8

Index scores measured and predicted using the LGBM method



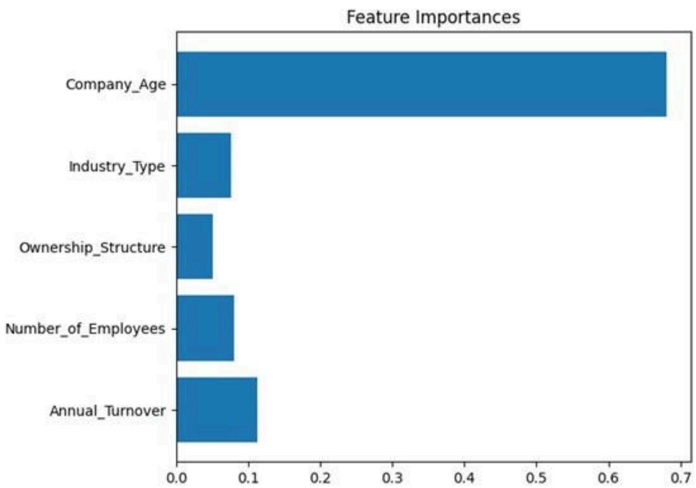
Feature Importance Analysis

To gain further insight into the internal dynamics of the prediction process, a feature importance analysis was conducted to evaluate the relative contributions of each input variable to the prediction of the Maturity Index (MI). As shown in Figure 9, company age is the most influential variable, followed by Annual Turnover and Number of Employees. These variables, associated with enterprise scale and longevity, have a stronger impact on MI predictions. In contrast, Industry Type and Ownership Structure exhibited comparatively lower levels of importance.

This analysis improves the interpretability of the model and supports practical decision-making by identifying the enterprise characteristics that are most influential in determining digital maturity.

Figure 9

Feature importance of input variables in predicting MI



Real-World Implementation and Future Research Directions:

Although the results presented in this study demonstrate the effectiveness of the proposed machine learning framework under controlled conditions, its integration into live industrial environments remains a critical area for future research. Real-time implementation will require not only technical adjustments, such as the automation of data flows and feedback loops, but also organizational readiness and stakeholder engagement. Collaborations with manufacturing enterprises are planned to deploy the model in operational settings and monitor its predictive accuracy and adaptability in dynamic real-world contexts. Such efforts are expected to validate the framework's practical value and guide necessary refinements for broader applicability. These real-world applications are also envisioned as key components of future research activities aimed at scaling up the model and verifying its impact in diverse industrial domains.

Conclusions

It is believed that Industry 4.0 will increase the success and quality of life of production. For this reason, the adaptation of enterprises to Industry 4.0 has gained importance. One measure of compliance is the MI score. Surveys are commonly used to determine MI scores; however, these procedures can be laborious and time-consuming.

This study presents an enhanced framework that integrates seven machine learning methods, including both traditional algorithms (e.g., SVM, GPR, LR, RT) and advanced boosting-based models (XGB, LGBM), to predict MI based on only five easily obtainable enterprise variables. The study demonstrates that the MI scores of enterprises can be predicted with high accuracy using machine learning, without the need for survey-based assessments.

The proposed framework was validated using a five-fold cross-validation technique, and performance metrics (R^2, RMSE, MSE, MAE) revealed that all models achieved strong predictive accuracy, with SVM yielding the highest R^2 value (0.75). Feature importance analysis further clarified the internal mechanics of the model, revealing that company age, turnover, and number of employees were the most influential factors in MI prediction. Thus, companies can reduce costs and accelerate their transition to Industry 4.0 using data-driven estimation tools. Although model performance may vary according to algorithm type, dataset size, and domain specificity, the proposed approach is scalable and practical solution.

In future research, this framework will be integrated into real-time industrial environments to evaluate its dynamic performance. Collaborations with manufacturing enterprises are envisaged to deploy the model in operational settings and monitor its predictive stability and utility across diverse industrial conditions.

Enterprises can rapidly determine Industry 4.0 readiness using the framework proposed in this study.

Peer Review	Externally peer-reviewed.
Author Contributions	Conception/Design of Study- A.D., C.Ü.; Data Acquisition- A.D., C.Ü.; Data Analysis/Interpretation- A.D., C.Ü.; Drafting Manuscript- A.D., C.Ü.; Critical Revision of Manuscript- A.D., C.Ü.; Final Approval and Accountability- A.D., C.Ü.; Technical or Material Support- A.D., C.Ü.; Supervision- A.D., C.Ü.
Conflict of Interest	The authors have no conflict of interest to declare.
Grant Support	The authors declared that this study has received no financial support.

Author Details

Ayhan Doğan
[1] Hacettepe University, Başkent Organized Industrial Zone, Technical Sciences Vocational School, Ankara, Türkiye
0000-0002-9872-8889 aydogan@hacettepe.edu.tr

Cihan Ünal
[1] Hacettepe University, Başkent Organized Industrial Zone, Technical Sciences Vocational School, Ankara, Türkiye
0000-0002-5255-4078

References

Akdil, K. Y., Ustundag, A., & Cevikcan, E. (2018). *Maturity and readiness model for industry 4.0 strategy. Industry 40: Managing the digital transformation*. Springer. pp 61-94

Angelopoulos, A., Michailidis, E. T., Nomikos, N., Trakadas, P., Hatziefremidis, A., Voliotis, S., & Zahariadis, T. (2019). Tackling faults in the industry 4.0 era—a survey of machine-learning solutions and key aspects. *Sensors*, 20, 109.

Aslan, Y., Yavasca, S., & Yasar, C. (2011). Long-term electric peak load forecasting of Kutahya using different approaches. *International Journal on Technical and Physical Problems of Engineering*, 3, 87-91.

Ateş, B. (2020). *Gemi yapılarında gerilme yığılması öngörülerinin kaba ağ yapısı ve makine öğrenmesi ile gerçekleştirilmesi*. Fen Bilimleri Enstitüsü

Ayhan, S., & Erdoğmuş, Ş. (2014). Destek vektör makineleriyle sınıflandırma problemlerinin çözümü için çekirdek fonksiyonu seçimi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 9, 175-201.

Basl, J., & Doucek, P. (2019). A metamodel for evaluating enterprise readiness in the context of Industry 4.0. *Information,* 10, 89.

Bayazit, M., & Oğuz, B. (1994). *Mühendisler için istatistik*. Birsen Yayınevi, İstanbul, 197s.

Becker, B. G., Klein, T., Wachinger, C., & Initiative AsDN (2018). Gaussian process uncertainty in age estimation as a measure of brain abnormality. *NeuroImage*, 175, 246-258.

Bibby, L., & Dehe, B. (2018). Defining and assessing industry 4.0 maturity levels–case of the defense sector. *Production Planning & Control*, 29, 1030-1043.

Boran, P. D. S. (2021). *Makine öğrenmesi yöntemleri kullanılarak kısa dönem rüzgar gücü tahmini*. Sakarya Üniversitesi

Botha, A. P. (2018). Rapidly arriving futures: future readiness for Industry 4.0. *South African Journal of Industrial Engineering*, 29, 148-160.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24, 123-140.

Breiman, L. (2001). Using iterated bagging to debian regressions. *Machine Learning*, 45, 261-277.

Brik, B., Bettayeb, B., Sahnoun, M. H., & Duval, F. (2019). Toward predicting system disruption in industry 4.0: Machine learning-based approach. *Procedia Computer Science*, 151, 667-674.

Candanedo, I. S., Nieves, E. H., González, S. R., Martín, M. T. S., & Briones, A. G. (2018). *Machine learning predictive model for industry 4.0*. Knowledge Management in Organizations: 13th International Conference, KMO 2018, Žilina, Slovakia, August 6–10, 2018, Proceedings 13. Springer. pp 501-510

Canetta, L., Barni, A., & Montini, E. (2018). *Development of a digitalization maturity model for the manufacturing sector*. 2018 ieee international conference on engineering, technology and innovation (ICE/ITMC). IEEE. pp 1-7

Carvalho, N., Chaim, O., Cazarini, E., & Gerolamo, M. (2018). Manufacturing in the fourth industrial revolution: A positive prospect in sustainable manufacturing. *Procedia Manufacturing*, 21, 671-678.

Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Chonsawat, N., & Sopadang, A. (2020). Defining SMEs' 4.0 readiness indicators. *Applied Sciences*, 10, 8998.

Cortes, C., & Vapnik, V. (1995). Support vector networks. *Machine Learning*, 20, 273-297.

Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press, Çınar, 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, 6659.

Da Silva, V. L., Kovaleski, J. L., Pagani, R. N., Silva, J. D. M., & Corsi, A. (2020). Implementation of Industry 4.0 concept in companies: Empirical evidence. *International Journal of Computer Integrated Manufacturing*, 33, 325-342.

Doğan, A. (2024). Prediction of major earthquakes that may affect İzmir using machine learning methods. *Yerbilimleri*, 45, 93-106. doi:10.17824/yerbilimleri.1402618

Doğan, A., Korkmaz, M., & Kirmaci, V. (2023). Estimation of Ranque-Hilsch vortex tube performance using machine learning techniques. *International Journal of Refrigeration*, 150, 77-88. https://doi.org/10.1016/j.ijrefrig.2023.01.021

Erol, S., Schumacher, A., & Sihn, W. (2016). *Strategic guidance toward Industry 4.0–a three-stage process model*. International conference on competitive manufacturing. pp 495-501

Finance, A. (2015). *Industry 4.0 Challenges and solutions for the digital transformation and use of exponential technologies*. Finance, audit tax consulting corporate: Zurich, Swiss:1-12.

Geissbauer, R., Vedso, J., & Schrauf, S. (2016). *Industry 4.0: Building the digital enterprise*.

Gerlitz, L. (2016). Design management as a domain of smart and sustainable enterprise: Business modeling for innovation and smart growth in Industry 4.0. *Entrepreneurship and Sustainability*, 3, 244.

Gökalp, E., Şener, U., & Eren, P. E. (2017). *Development of an assessment model for industry 4.0: industry 4.0-MM*. International Conference on Software Process Improvement and Capability Determination. Springer. pp 128-142

Haber, R. E., Juanes, C., del Toro, R., & Beruvides, G. (2015). Artificial cognitive control with self-x capabilities: A case study of a micro-manufacturing process. *Computers in Industry*, 74, 135-150.

Hamidi, S. R., Aziz, A. A., Shuhidan, S. M., Aziz, A. A., & Mokhsin, M. (2018). *SMEs maturity model assessment of IR4. 0 digital transformation*. International Conference on Kansei Engineering & Emotion Research. Springer. pp 721-732

Haykin, S. (1999). *Neural networks: A guided tour*. Soft computing and intelligent systems: theory and applications 71.

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and Their Applications*, 13, 18-28.

Heo, Y., & Zavala, V. M. (2012). Gaussian process modeling for measurement and verification of building energy savings. *Energy and Buildings*, 53, 7-18.

Hizam-Hanafiah, M., Soomro, M. A., & Abdullah, N. L. (2020). Industry 4.0 readiness models: a systematic literature review of model dimensions. *Information*, 11, 364.

Horvat, D., Stahlecker, T., Zenker, A., Lerch, C., & Mladineo, M. (2018). A conceptual approach to analyzing manufacturing companies' profiles concerning Industry 4.0 in emerging economies. *Procedia Manufacturing*, 17, 419-426.

Horváth, D., & Szabó, R. Z. (2019). Driving forces and barriers of Industry 4.0: Do multinational and small- and medium-sized companies have equal opportunities? *Technological Forecasting and Social Change*, 146, 119-132.

Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics & Proteomics*, 15, 41-51.

Ivanov, D., Dolgui, A., & Sokolov, B. (2019). *The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics*. International Journal of Production Research, 57, 829-846.

Jalil, N. A., Hwang, H. J., & Dawi, N. M. (2019). *Machines learning trends, perspectives and prospects in education sector*. Proceedings of the 2019 3rd International Conference on Education and Multimedia Technology, 201-205.

Jesus, C. D., & Lima, R. M. (2020). Literature search of key factors for developing generic and specific maturity models for Industry 4.0. *Applied Sciences*, 10, 5825.

Kalayci, T. E. (2018) Kimlik hırsızı web sitelerinin sınıflandırılması için makine öğrenmesi yöntemlerinin karşılaştırılması. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 24, 870-878.

Kannan, S. M., Suri, K., Cadavid, J., Barosan, I., Van Den Brand, M., Alferez, M., & Gerard, S. (2017). *Toward industry 4.0: Gap analysis between current automotive MES and industry standards using model-based requirement engineering*. 2017 IEEE International Conference on Software Architecture Workshops (ICSAW). IEEE. pp 29-35

Katuwal, G. J., Baum, S. A., & Michael, A. M. (2018). Early brain imaging can predict autism: Application of machine learning to a clinical imaging archive. *BioRxiv*:471169.

Kavzoğlu, T., & Çölkesen, İ. (2010). Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi. *Harita Dergisi*, 144, 73-82.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30.

Kuo, C. J., Ting, K. C., Chen, Y. C., Yang, D. L., & Chen, H. M. (2017). Automatic machine status prediction in the era of Industry 4.0: Case study of machines in a spring factory. *Journal of Systems Architecture*, 81, 44-53.

Lee, C., & Lim, C. (2021). From technological development to social advance: A review of Industry 4.0 through machine learning. *Technological Forecasting and Social Change*, 167, 120653.

Leyh, C., Bley, K., Schäffer, T., & Forstenhäusler, S. (2016). *SIMMI 4.0-a maturity model for classifying the enterprise-wide it and software landscape focusing on Industry 4.0*. 2016 federated conference on computer science and information systems (fedcsis). IEEE. pp 1297-1302

Liu, K., Hu, X., Wei, Z., Li, Y., & Jiang, Y. (2019). *Modified Gaussian process regression models for cyclic capacity prediction of lithium-ion batterie*s. IEEE Transactions on Transportation Electrification, 5, 1225-1236.

Luu, Q. H., Lau, M. F., Ng, S. P., & Chen, T. Y. (2021). Testing multiple linear regression systems with metamorphic testing. *Journal of Systems and Software*, 182, 111062. https://doi.org/10.1016/j.jss.2021.111062

Maimon, O., & Rokach, L. (2005). *Data mining and knowledge discovery handbook*.

Manavalan, E., & Jayakrishna, K. (2019). A review of Internet of Things (IoT) embedded sustainable supply chain for industry 4.0 requirements. *Computers and Industrial Engineering*, 127, 925-953.

Mehmet, A., & Doğansoy, G. A. (2022). Makine öğrenmesi ve derin öğrenme yöntemleri kullanılarak e-perakende sektörüne yönelik talep tahmini. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 37, 1325-1340.

Memnun, D., & Kalaycı, Ş. (2006). *SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri*. Ankara: Asil Yayın Dağıtım. İlköğretim Online 12.

Mendeş, M., & Akkartal, E. (2009). Regression tree analysis for predicting slaughter weight in broilers. *Italian Journal of Animal Science*, 8, 615-624.

Methavitakul, B., & Santiteerakul, S. (2018). *Analysis of key dimension and sub-dimension for Supply Chian of Industry to fourth Industry Performance Measurement*. 2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI). IEEE. pp 191-195

Mittal, S., Khan, M. A., Romero, D., & Wuest, T. (2018). A critical review of smart manufacturing & Industry 4.0 maturity models: Implications for small- and medium-sized enterprises (SMEs). *Journal of Manufacturing Systems*, 49, 194-214.

Muñoz, L. P., Hagemann, M., Palandella, L., & Cagliari, P. (2023). *Maturity SCAN 4.0: An analysis and discussion of results from scanning of digital capabilities adopted by Argentinian CPG companies*. ESSN: 2701-6277, 854-863.

Müller, J. M., Voigt, K. I. (2018). The impact of industry 4.0 on supply chains in engineer-to-order industries-an exploratory case study. *IFAC-PapersOnLine*, 51, 122-127.

Opitz, D., & Maclin, R. (1999). Popular ensemble methods: An empirical study. *Journal of Artificial Intelligence Research*, 11, 169-198.

Özen, N. S., Saraç, S., & Koyuncu, M. (2021). COVID-19 Vakalarının Makine Öğrenmesi Algoritmaları ile Tahmini: Amerika Birleşik Devletleri Örneği. *Avrupa Bilim ve Teknoloji Dergisi*, 134-139.

Pessl, E., Sorko, S. R., & Mayer, B. (2017). Roadmap Industry 4.0–implementation guideline for enterprises. *International Journal of Science, Technology and Society*, 5, 193-202.

Rahamaddulla, S. R. B., Leman, Z., Baharudin, B. H. T. B., & Ahmad, S. A. (2021). Conceptualizing smart manufacturing readiness-maturity model for small and medium enterprise (sme) in malaysia. *Sustainability*, 13, 9793.

Rai, R., Tiwari, M. K., Ivanov, D., & Dolgui, A. (2021). *Machine learning in manufacturing and industry 4.0 applications*. Taylor and Francis. pp 4773-4778.

Rais, R. B. (2021). *Sharing the Pie The Fourth Industrial Revolution and the Sharing/Platform Economy: Distributive justice implications*. Beyond Free Market. Routledge. pp 125-140

Rasmussen, C. E., & Williams, C. (2006). *Gaussian processes for machine learning, vol. 1*. MIT press Cambridge MA.

Rauch, E., Dallasega, P., & Unterhofer, M. (2019). Requirements and barriers for introducing smart manufacturing in small- and medium-sized enterprises. *IEEE Engineering Management Review*, 47, 87-94.

Rong, S., & Bao-Wen, Z. (2018). The research of regression model in machine learning field. MATEC Web of Conferences. *EDP Sciences*. 01033, 2019.

Saad, S. M., Bahadori, R., & Jafarnejad, H. (2021). The smart SME technology readiness assessment methodology in the context of industry 4.0. *Journal of Manufacturing Technology Management*.

Sajjad, A., Ahmad, W., Hussain, S., Chuddher, B. A., Sajid, M., Jahanjaib, M., Ali, M. K., & Jawad, M. (2024). Assessment by Lean Modified Manufacturing Maturity Model for Industry 4.0: A Case Study of Pakistan's Manufacturing Sector. *IEEE Transactions on Engineering Management*, 71, 6420-6434. https://doi.org/10.1109/TEM.2023.3259005

Schaupp, E., Abele, E., & Metternich, J. (2017). Potentials of digitalization in tool management. *Procedia CIRP*, 63, 144-149.

Schiltz, F., Masci, C., Agasisti, T., & Horn, D. (2018). Using regression tree ensembles to model interaction effects: a graphical approach. *Applied Economics*, 50, 6341-6354.

Schölkopf, B., Williamson, R. C., Smola, A., Shawe-Taylor, J., & Platt, J. (1999). *Support vector method for novelty detection*. Advances in neural information processing systems 12.

Schumacher, A., Erol, S., & Sihn, W. (2016). A maturity model for assessing Industry 4.0 readiness and maturity of manufacturing enterprises. *Procedia Cirp*, 52, 161-166.

Shoaran, M., Haghi, B. A., Taghavi, M., Farivar, M., & Emami-Neyestanak, A. (2018). Energy-efficient classification for resource-constrained biomedical applications. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 8, 693-707.

Skalli, D., Charkaoui, A., & Cherrafi, A. (2023). *Design of Industry 4.0 Maturity Model and Readiness Assessment Under the Digital Transformation Toward the Implementation of LSS4.0*. In T. Masrour, I. El Hassani, and N. Barka, Artificial Intelligence and Industrial Applications Cham.

Soman, K., Loganathan, R., & Ajay, V. (2009). *Machine learning with SVM and other kernel methods*. PHI Learning Pvt. Ltd., New Delhi.

Stentoft, J., Adsbøll Wickstrøm, K., Philipsen, K., & Haug, A. (2021). Drivers and barriers for Industry 4.0 readiness and practice: empirical evidence from small and medium-sized manufacturers. *Production Planning & Control*, 32, 811-828.

Temel, G. O. (2004). *Sınıflama ve Regresyon Ağaçları*. Yayımlanmamış Yüksek Lisans Tezi, Mersin Üniversitesi, Sağlık Bilimleri Enstitüsü, Bioistatistik Anabilim Dalı, Mersin 11:23-24.

Tolun, S. (2008). *Destek vektör makineleri: Banka başarısızlığının tahmini üzerine bir uygulama*. İktisadî Araştırmalar Vakfı

Ünal, C., Sungur, C., & Yildirim, H. (2022). Application of the maturity model in industrial corporations. *Sustainability*, 14, 9478.

Valentin, B. A. (2017). *Methods Of Assessment And Training Of A Company Toward The Enterprise 4.0*. Annals of DAAAM & Proceedings 28.

Városiné Demeter, K., Losonci, D., Szász, L., & Rácz, B. G. (2018). *Assessing Industry 4.0 readiness: a multi-country industry level analysis*.

Vazire, S. (2018). Implications of the credibility revolution for productivity, creativity, and progress. *Perspectives on Psychological Science*, 13, 411-417.

Wank, A., Adolph, S., Anokhin, O., Arndt, A., Anderl, R., & Metternich, J. (2016). Using a learning factory approach to transfer Industrie 4.0 approaches to small-and medium-sized enterprises. *Procedia Cirp*, 54, 89-94.

Wei, J., Chu, X., Sun, X. Y., Xu, K., Deng, H. X., Chen, J., Wei, Z., & Lei, M. (2019). Machine learning in materials science. *InfoMat*, 1, 338-358.

Yesiloglu-Gultekin, N., & Dogan, A. (2024). Estimation of the elastic modulus of basaltic rocks using machine learning methods. *Earth Science Informatics*, doi:10.1007/s12145-024-01472-7

Yücel, Y. B. (2017). *Yaşam memnuniyetini etkileyen faktörlerin sınıflama ve regresyon ağacı ile belirlenmesi*. Fen Bilimleri Enstitüsü.

Zhou, G., Zhang, J., Su, J., Shen, D., & Tan, C. (2004). Recognizing names in biomedical texts: a machine learning approach. *Bioinformatics*, 20, 1178-1190.