Contribution of Knowledge Management to Innovation Capabilities in the Manufacturing Industry Through Machine Learning

Juan Ibujés-Villacís[a] and Michael S. Simba-Herrera[b]
Facultad de Ciencias Administrativas, Escuela Politécnica Nacional, Quito, Ecuador

Keywords: Business Management, Innovation Capabilities, Knowledge Management, Manufacturing Industry, Machine Learning.

Abstract: Knowledge management has been fundamental for organizations to improve their ability to innovate. The objective of this research is to design and develop machine learning models that impact predictive analytics, identifying the determinants of knowledge management (KM) that influence innovation capabilities (IC) in the manufacturing industry. Given the quantitative nature of the research, in a first stage, information on factors related to KM and IC was collected and processed. In a second phase, six models were developed to predict which manufacturing companies innovate in their production processes based on a set of KM factors. Information from 142 manufacturing companies in the province of Pichincha, Ecuador, was used for the study. The results show that all the factors of KM contribute to innovation capabilities, with organizational structure, technology, people and incentives standing out in particular. This study is pioneering in Ecuador and reinforces the strategic value of corporate governance as a driver of industrial innovation and provides a useful empirical framework to guide decision making and business policy formulation. In addition, this study contributes to the field of knowledge management by providing empirical evidence on the key factors on which manufacturing companies should focus their efforts to develop innovation capabilities in processes, products and services.

SCIENCE AND TECHNOLOGY PUBLICATIONS

1 INTRODUCTION

In the business context, innovation is a key mechanism for generating new economic value, by developing novel products, implementing efficient production methods and boosting sales (Nakamori, 2020; Zawislak et al., 2018). This way of creating economic value requires to be properly managed to achieve business objectives.

According to the global innovation index, Ecuador ranks 91st worldwide and 12th in the Latin American and Caribbean (LAC) region (Dutta et al., 2021). These positions reflect the low levels of innovation present in the Ecuadorian business fabric, which is consistent with the results of the latest innovation survey, which determines that only 54.5% of Ecuadorian companies carry out some type of innovation (product, process, organizational or marketing) (SENESCYT-INEC, 2015)

Low innovative capacity negatively affects the competitiveness and sustainability of organizations in developing countries (Qin, 2024) and represents a particular challenge for Ecuadorian industry in the context of the knowledge economy (Aguilar-Barceló & Higuera-Cota, 2019; CEPAL, 2016). This problem is addressed in this research from the perspective of knowledge management (KM) in the industrial setting.

Knowledge management allows the identification and exploitation of best practices in the creation of new products services, in addition to contributing to the prevention of errors and rework (Pagani & Champion, 2024). From this approach, many organizations adopt strategies such as exploitation, acquisition, sharing, exploration and exploitation of knowledge in order to improve their business management (Bolisani & Bratianu, 2018). However, these strategies do not always translate into

[a] https://orcid.org/0000-0001-8439-3048
[b] https://orcid.org/0009-0009-6978-9330

Ibujés-Villacís, J. and Simba-Herrera, M. S.
Contribution of Knowledge Management to Innovation Capabilities in the Manufacturing Industry Through Machine Learning.
DOI: 10.5220/0013673000004000
Paper published under CC license (CC BY-NC-ND 4.0)
In *Proceedings of the 17th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2025) - Volume 2: KEOD and KMIS*, pages 357-369
ISBN: 978-989-758-769-6; ISSN: 2184-3228
Proceedings Copyright © 2025 by SCITEPRESS – Science and Technology Publications, Lda.

innovative results, probably due to the absence of tools that facilitate data-driven decision making.

The purpose of this research is to design and apply machine learning models that allow predictive analysis, identifying which manufacturing companies have innovative capabilities (IC) based on their knowledge management practices. This is a pioneering study in the Ecuadorian context, since no precedents have been found that integrate the proposed models to predict innovative behavior based on KM.

From a methodological perspective, the study has a quantitative approach. A structured survey was applied to a random sample of 142 manufacturing companies in the province of Pichincha, Ecuador, collecting information on factors related to KM and IC, taking as reference previous studies (Ibujés-Villacís & Franco-Crespo, 2024). With these data, several multiple linear regression models were developed, considering the KM variables as independent variables and the IC variables as dependent variables.

This study, by training algorithms with data from medium-sized manufacturing companies, allows us to identify the factors of KM that are determinants for the development of innovation capabilities in this sector. Its results provide relevant inputs for the formulation of strategies aimed at fostering innovation in products, services and processes aligned with business objectives.

The article is structured as follows: first, a conceptual framework on KM, IC and machine learning is presented. Next, six multiple linear regression (MLR) models are described to predict variables associated with IC from factors related to KM. Subsequently, algorithms are developed to identify significant KM variables that impact innovation within the manufacturing industry. Finally, the results are discussed, conclusions and limitations of the study are presented, and future lines of research are proposed.

2 THEORETICAL ELEMENTS

2.1 Knowledge Management

In the field of organizational management, there is practically consensus that at the present time, the most important strategic resource of organizations is (Bolisani & Bratianu, 2018; Davila et al., 2019; Kesavan, 2021; North & Kumta, 2018). By such virtue, KM is one of the most important organizational capabilities of organizations to innovate products and processes (Camisón-Haba et al., 2019; Chang et al., 2017; Ode & Ayavoo, 2020).

KM is a multidimensional category that encompasses important aspects related to the human, technological and political dimensions, which interact in the complex process of value creation in organizations (Espindola & Wright, 2021; Manning & Manning, 2020; Obeidat et al., 2016). According to Ibujés-Villacís & Franco-Crespo (2024), KM is represented by a series of variables grouped into seven factors: Policies and strategies, Organizational structure, Technology, People, Incentive systems, Organizational culture, and Communication.

2.2 Innovation Capabilities

Joseph A. Schumpeter (1883-1950) understood innovation as the introduction of new products or the improvement of existing ones; the introduction of a new or improved method of production; the opening of a new market; the use of a new method of selling or purchasing; the use of new raw materials or semi-finished products; or the introduction of new forms of organization of production (Nakamori, 2020; Szczepańska-Woszczyna, 2021).

To innovate organizations, require new capabilities that are related to the abilities to continuously transform and exploit the potentiality of organizational knowledge, with the objective of value creation through the generation of significant changes in products and processes (Kaur, 2019; Nakamori, 2020; OECD & Eurostat, 2018). By having these capabilities, organizations develop new intra- and inter- organizational learning systems, and focus organizational management towards the market and changing environments (Bogodistov et al., 2017; Bykova & Jardon, 2018; Kodama, 2018; Salmador et al., 2021)

According to Ibujés-Villacís & Franco-Crespo, (2024), IC are represented by a series of variables grouped into six factors: Research and Development, Management capability, Resource availability, Talent management, Staff skills, and Technological capability.

2.3 Machine Learning

Machine learning (ML), predictive modeling and artificial intelligence are closely linked concepts (Shmueli et al., 2023). This field of study allows computers to learn from data without the need to be explicitly programmed. In this context, a computer system improves its performance on specific tasks as it accumulates experience (Akerkar, 2019).

The process usually begins with the construction of a model, a simplified representation of reality that enables the identification of patterns and relationships within the data (Burger, 2018; Kuhn & Silge, 2022). Unlike dashboards, which provide a static view of data, models enable dynamic analysis and prediction of future trends (Burger, 2018). There are several machine learning models, such as regression, clustering, and neural networks, all based on algorithms. To achieve the objective of this research, a supervised learning algorithm was used by applying a multiple linear regression model, a tool widely used in predictive analytics in business (Pagani & Champion, 2024; Weber, 2023).

Machine learning requires training models from a set of data, usually a representative sample of the total available. During this process, model performance is evaluated based on error reduction and fit ability. If errors persist, the model needs to be adjusted and refined (Burger, 2018). A common practice in this process is cross-validation, which consists of dividing the dataset into subsets for training and testing, thus allowing to improve the predictive ability and generalization of the model (Burger, 2018; Hastie et al., 2023).

2.4 Multiple Linear Regression

Multiple linear regression (MLR) is a statistical technique used to model the relationship between a dependent variable and two or more independent variables. MLR seeks to find the best straight line (or hyperplane in higher dimensions) that fits the data optimally. This involves determining the coefficients that minimize the difference between the values predicted by the model and the actual values observed in the data set.

Mathematically, the multiple linear regression model is expressed as equation 1:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \qquad (1)$$

Where

Y is the dependent variable.

$X1$, $X2$, …., Xn are the independent variables.

$\beta 0$, $\beta 1$, $\beta 2$, …., βn are the coefficients representing the slope of each independent variable.

ϵ is the error term, which captures the variation not explained by the model.

Multiple linear regression (MLR) is a statistical technique useful for analyzing how multiple independent factors simultaneously influence a specific outcome. In this study, it is used to examine and predict the relationship between dependent variables associated with IC and independent variables linked to KM, making it possible to identify which KM practices have a significant impact on strengthening innovation within the Ecuadorian manufacturing sector.

The analysis focuses on a group of medium-sized companies in the manufacturing sector located in the province of Pichincha, Ecuador. In this context, the dependent variables - detailed in Table 2 - reflect different aspects of innovation capabilities such as research and development (R&D). The independent variables, presented in Table 1, represent different dimensions of knowledge management implemented by the organizations.

3 METHODOLOGY

This research has a quantitative approach, is correlational, non-experimental and cross-sectional. Figure 1 shows the complete process to achieve the research objective, starting with the determination of the sample and ending with the obtaining of results.

The methodology employed consists of six stages: sample determination, data collection, data exploration and preparation, modeling and analysis, model training and validation, and finally, obtaining results and proposals for action. This approach allows for a rigorous and systematic treatment of the information to ensure the validity of the conclusions.

3.1 Multiple Linear Regression

The study focuses on companies in the manufacturing sector in the province of Pichincha, whose capital is Quito, Ecuador's economic and political center. This sector was selected because of its important contribution to the national economy, representing 14.2% of the country's total production (MIPRO, 2021).

The study population is composed of medium-sized active manufacturing companies with at least five years of operation. These companies are characterized by having between 50 and 199 employees, generating annual revenues between 1 and 5 million dollars and having assets of less than 4 million dollars (SUPERCIAS, 2021). According to records available as of November 2020, there were 338 medium-sized companies in this sector that had submitted their financial reports for 2019 (SUPERCIAS, 2020).

To calculate the sample size, proportional probability sampling for finite population was applied, using simple random selection without replacement. This technique guarantees the representativeness of the sample and ensures that all units have the same probability of being selected (Latpate et al., 2021; Lohr, 2019).

The final determination of the sample size (n) was performed using Equation 2, based on the statistical methods proposed by Lohr (2019) and Ott & Longnecker (2016)

$$n = \frac{Z^2 Npq}{E^2(N-1) + Z^2 pq} \quad (2)$$

The parameters used to calculate the sample were: N = 338 (population size), E = 10 % (margin of sampling error), Z = 1.96 (95 % confidence level), p = 0.5 (probability of success) and q = 0.5 (probability of failure). With these values, a minimum sample size of n = 75 companies was determined. However, the study was conducted with 142 companies, far exceeding the minimum requirement. This made it possible to reduce the estimated sampling error to 6%, maintaining the 95% confidence level.

3.2 Data Collection

Data collection was carried out through a survey directed at senior executives of the companies included in the study sample. A structured questionnaire was applied, based on the innovation management model proposed by Ibujés-Villacís & Franco-Crespo (2022). The instrument assessed a total of 85 items organized into two main sections: on one hand, knowledge management (KM), represented by 35 variables grouped into seven factors; and on the other hand, innovation capabilities (IC), represented by six key variables.

This questionnaire was subjected to content validation by experts, considering four categories: coherence, relevance, clarity and sufficiency of the questions. To ensure these qualities, a pilot test was conducted with the participation of ten experts from academia and industry. Based on the validation and the comments received, the suggested improvements were incorporated and the final version of the questionnaire was prepared.

To respond to the questionnaire, company managers were asked to rate each of the items using the psychometric instrument called Likert scale (Bertram, 2018). A 10-point scale was used, with 1 representing very low agreement and 10 representing very high agreement with the argument presented in each item.

Table 1: Knowledge management factors and variables.

Knowledge management variables	
Factor 1: Policies and strategies (PS)	
Policies for the acquisition and generation of organizational knowledge.	PS1
Policies for the storage, sharing and use of organizational knowledge.	PS2
Implementation of properly documented processes, procedures and routines	PS3
Establishment of alliances with public and private organizations.	PS4
Development of dynamic plans to overcome internal and external barriers.	PS5
Permanent focus on continuous improvement.	PS6
Systematic combination of existing and new knowledge.	PS7
Factor 2: Organizational structure (OS)	
Internal organizational structures dedicated to research and development.	OS1
Regulations established for the access and use of knowledge.	OS2
Agility in the processes to access organizational knowledge.	OS3
Facilities for the horizontal flow of knowledge within the organization.	OS4
Facilities for the vertical flow of knowledge within the organization.	OS5
Factor 3: Technology (TG)	
Use of technology for the methodical storage of knowledge.	TG1
Use of information systems for accessing, sharing and utilizing the organizational knowledge.	TG2
Application of ICT for access, exchange and use of knowledge.	TG3
Utilization of corporate social networks for collaboration and knowledge of the environment.	TG4
Factor 4: Persons (PP)	
Years of employee experience.	PP1
Employees' level of education.	PP2
Age of employees.	PP3
Foreign language proficiency of employees.	PP4
Gender diversity among employees.	PP5
Factor 5: Incentive systems (IS)	
Economic incentives for generating, sharing and using knowledge.	IS1
Training offered as an incentive for generating, sharing and using the knowledge.	IS2
Days off granted as an incentive for generating, sharing, and using the knowledge.	IS3
Public recognition as an incentive for generating, sharing and utilizing the knowledge.	IS4
Factor 6: Organizational culture (OC)	
Importance of personal values.	OC1
Positive attitude towards work.	OC2
Respect for the company's principles and regulations.	OC3
Application of best practices.	OC4
Staff empowerment for decision making.	OC5
Creation of a collaborative and synergistic work environment.	OC6
Factor 7: Communication (CM)	
Formal communication in the work environment.	CM1
Informal communication in the work environment.	CM2
Effective communication with all hierarchical levels.	CM3
Fluent communication in physical and virtual spaces.	CM4

Table 2: Variables of innovation capabilities.

IC variables	
Research and Development (R&D) capability	RD
Management capability	MC
Resource availability	RA
Talent management	HT
Staff skills	SS
Technological capability	TC

The surveys were conducted using a Google form, applied electronically from June to September 2021. A total of 250 questionnaires were sent by e-mail to the companies that were the subject of the study. Each survey complied with ethical research standards: informed consent, voluntary participation, confidentiality and absence of physical or psychological risk to participants.

3.3 Data Exploration and Preparation

Exploratory data analysis is a critical phase in the modeling process as it provides valuable information on the nature and quality of the data (Costa-Climent et al., 2023). This phase is essential because its results can influence decisions made during the modeling process and improve the effectiveness and interpretation of the resulting models.

Since both input and output variables are quantitative, the multiple linear regression model (MLR) was selected to analyze the relationship between knowledge management variables and innovation capabilities. The responses obtained in the survey are on a scale of 1 to 10, so no obvious outliers were identified. Consequently, it was not necessary to perform histograms, boxplots or scatter plots to explore the distribution of the data or to detect possible outliers.

The relationships between each of the variables that make up the different KM categories were explored to detect multicollinearity of the independent variables. Multicollinearity occurs when two or more independent variables in a model are highly correlated with each other (Lantz, 2023). The presence of multicollinearity can cause several problems in regression analysis, including instability in coefficient estimation, increased coefficient variance, and unreliable coefficients.

Correlation analysis between the KM variables identified and eliminated 10 variables with correlation coefficients greater than 0.7, indicating high collinearity. As a result, 25 independent variables were retained for further analysis.

3.4 Data Modeling and Analysis

The methodological approach adopted in this research corresponds to supervised machine learning, which consists of training models with labeled data, i.e., with known values for both independent and dependent variables (Burger, 2018).

To evaluate the impact of KM on IC in the manufacturing industry, we chose to use a MLR model. This choice is based on two main reasons: first, the quantitative nature of all the variables involved; second, the availability and sufficiency of the data set, which allows adequate implementation and validation of the proposed model.

The formal structure of the model is presented in Equation 1. Six multiple linear regression models were developed, the details of which are presented in Table 3.

Table 3: Multiple linear regression models.

Model	IC	Knowledge management factors
	Dependent variable (Y)	Independent variables (X)
1	Y1= RD	PS2, PS3, PS4, PS4, PS5, PS7, OS1, OS2, OS4, OS5, TG3, TG4, PP1, PP2, PP3, PP4, PP5, IS1, IS2, IS3, IS4, OC1, OC5, CM1, CM2, CM3
2	Y2= MC	
3	Y3= RA	
4	Y4= HT	
5	Y5= SS	
6	Y6= TC	

3.5 Training, Validation, Evaluation and Adjustment

The database used contains 142 records and 31 variables, of which 25 are associated with KM and six with IC. To evaluate the performance of the predictive model, the data were divided into two subsets: training data (80 %) and test data (20 %).

To ensure a robust evaluation of the predictive model and reduce the risk of overfitting, it was necessary to apply a technique to estimate its performance more reliably. In this study, cross-validation was used, a common method in statistics and machine learning that allows the predictive capacity of a model to be evaluated. This technique consists of dividing the data set into multiple training and test subsets, training and evaluating the model on different combinations of these subsets (Boehmke & Greenwell, 2020). In particular, the K-fold technique was applied, with ten folds, which allowed obtaining more stable estimates by averaging the results of each iteration, thus strengthening the validity of the analysis.

A recipe was used to define a sequence of data preprocessing steps, systematically applied to the data sets prior to modeling. This recipe acted as a standardized template, ensuring consistency in data pre-processing. A workflow was then built to implement the MLR, integrating both the model and the steps defined in the recipe. This approach allowed the model to be trained and evaluated in a consistent and reproducible manner.

Model validation was carried out by calculating the root mean squared error (RMSE), an indicator that measures the dispersion of the residual errors. The RMSE is obtained by calculating the root mean square error between the model predictions and the actual values observed in the test set. This value provides a direct estimate of the average magnitude of the prediction errors (Kuhn & Silge, 2022).

Where n is the number of observations in the test set, y_i are the actual values of the dependent variable and $\hat{y}_i$ are the predictions of the model for the dependent variable. A model performs well the lower the RMSE value and the closer this value resembles the one obtained between the training and test data (Kuhn & Silge, 2022). Both modeling and data analysis were performed using the RStudio programming language.

4 RESULTS

4.1 Relationship Between KM and Research and Development

The relationship between KM and research and development (R&D) was evaluated using a multiple linear regression model RD=f(X)+Ɛ. Table 4 shows that two KM variables pertaining to factors such as organizational structure and incentive system are significant and have a direct relationship with R&D.

Table 4: KM variables that impact research and development.

KM	Coefficient	Pr(>|t|)
OS1	0.28275	1.2e-05
IS2	0.22511	0.00073

Note: Adjusted R^2=0.791; F=15.2; p-value (model)=< $2e^{-16}$.

The statistical results of the model indicate that it is significant and viable as a whole. The model is represented by the following function: RD= 0.28 OS1 + 0.23 IS2, and presents an RMSE 1.29 which evidences a good fit of the model to the data. Table 5 reviews the statistical assumptions of the model, while Figure 1 shows these results graphically.

Table 5: Statistical assumptions.

Supposed	Value recommended	Value obtained	Evaluation
Normality of waste	P > 0.05	P=0.176	Ok.
Heterocedasticity	P > 0.05	P=0.252	Ok.
Autocorrelated residuals	P > 0.05	P= 0.184	Ok

Note: Statistics obtained from RSudio.

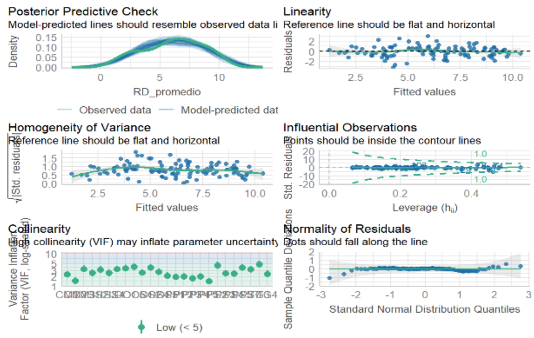

Figure 1: Graphs of statistical assumptions.

Note: Image obtained from RSudio.

4.2 Relationship Between KM and Capacity Management

The relationship between KM and capability management was assessed using a multiple linear regression model MC=f(X)+Ɛ. Table 6 shows that the variable of KM belonging to the factor of organizational structure is significant and have a direct relationship with capability management. In addition, a variable belonging to the Policies and strategies category has a significant and indirect relationship.

Table 6: KM variables that impact capacity management.

KM	Coefficient	Pr(>|t|)
OS1	0.17590	0.0074
PS2	-0.19049	0.0468

Note: Adjusted R^2=0.699; F =11.3; p-value (model)=< $2e^{-16}$.

The statistical results of the model indicate that it is significant and viable as a whole. The relationship is expressed by the following function: MC = 0.18 OS1 - 0.19 PS2, and presents an RMSE of 1.38 which evidences a good fit of the model to the data. Table 7 reviews the statistical assumptions of the model, while Figure 2 shows these results graphically.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

Table 7: Statistical assumptions.

Supposed	Value recommended	Value obtained	Evaluation
Normality of waste	P > 0.05	P = 0.005	Warning
Heterocedasticity	P > 0.05	P= 0.036	Warning
Autocorrelated residuals	P > 0.05	P = 0.008	Warning

Note: Statistics obtained from RSudio.

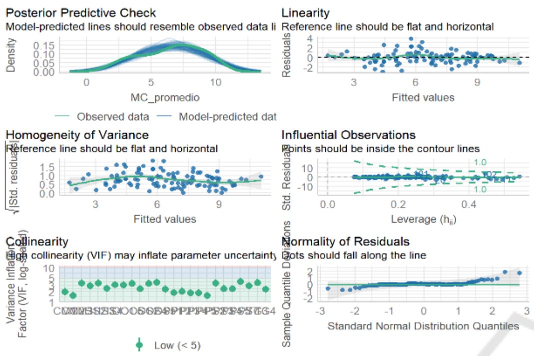

Figure 2: Graphs of statistical assumptions.

Note: Image obtained from RSudio.

4.3 Relationship Between KM and Resource Availability

The relationship between KM and resource availability was evaluated using a multiple linear regression model RA=f(X)+Ɛ. Table 8 shows that a variable of the KM belonging to the organizational structure factor is significant and has a direct relationship with the availability of resources.

Table 8: KM variables that impact resource availability.

KM	Coefficient	Pr(>|t|)
OS1	0.33728	6.9e-05

Note: Adjusted $R^2 = 0.519$; F = 5.84; p-value (model) = < 3.16e^{-16}.

The statistical results of the model indicate that it is significant and viable as a whole. The model is represented by the following function: RA= 0.34 OS1, and presents an RMSE of 1.47 which demonstrates a good fit of the model to the data. Table 9 reviews the statistical assumptions of the model, while Figure 3 shows these results graphically.

Table 9: Statistical assumptions.

Supposed	Value recommended	Value obtained	Evaluation
Normality of waste	P > 0.05	P= 0.444	Ok.
Heterocedasticity	P > 0.05	P= 0.316	Ok.
Autocorrelated residuals	P > 0.05	P = 0.002	Warning

Note: Statistics obtained from RSudio.

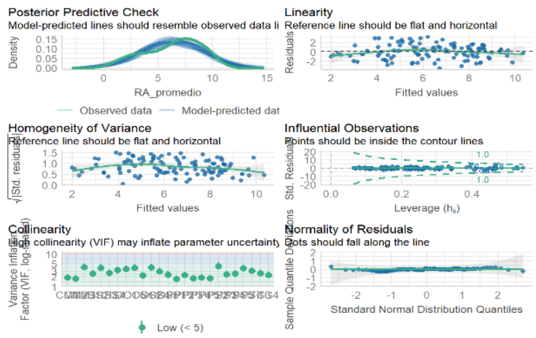

Figure 3: Graphs of statistical assumptions.

Note: Image obtained from RSudio.

4.4 Relationship Between KM and Human Talent Management

The relationship between KM and human talent management was evaluated using a multiple linear regression model HT=f(X)+Ɛ. Table 10 shows that four KM variables belonging to the Policies and Strategies, Organizational Structure, Technology, and Incentive System factors are significant and have a direct relationship with human talent management. In addition, two variables belonging to the Technology and People factors have a significant and indirect relationship.

Table 10: KM variables that impact human talent management.

KM	Coefficient	Pr(>|t|)	KM	Coefficient	Pr(>|t|)
PS7	0.2687	0.0032	TG4	-0.1772	0.0051
OS1	0.1178	0.0481	PP3	-0.1914	0.0027
TG3	0.3611	7.6e-05	IS2	0.1771	0.0059

Note: Adjusted R^2=0.782; F=16.8; p-value (model)=< 2e^{-16}.

The statistical results of the model indicate that it is significant and viable as a whole. The model is represented by the following function: HT= 0.27 PS7 + 0.12 OS1 + 0.36 TG3 - 0.18 TG4 - 0.19 PP3+0.18 IS2, and presents an RMSE 2.13 which evidences a good fit of the model to the data. Table 11 reviews the statistical assumptions of the model, while Figure 4 shows these results graphically.

Table 11: Statistical assumptions.

Supposed	Value recommended	Value obtained	Evaluation
Normality of waste	P > 0.05	P= 0.263	Ok
Heterocedasticity	P > 0.05	P= 0.195	Ok
Autocorrelated residuals	P > 0.05	P= 0.080	Ok

Note: Statistics obtained from RSudio.

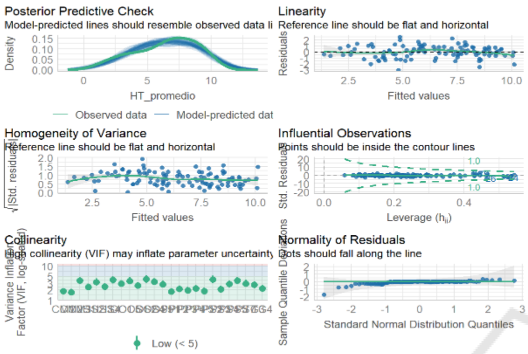

Figure 4: Graphs of statistical assumptions.

Note: Image obtained from RSudio.

4.5 Relationship Between KM and Staff Skills

The relationship between KM and personnel skills was evaluated using a multiple linear regression model SS= f(X)+Ɛ. Table 12 shows that six KM variables belonging to the factors: Organizational Structure, Technology, Incentive System, Organizational Culture and Communication are significant and have a direct relationship with personnel skills.

In addition, two variables belonging to the factors: Policies and strategies, and People have a significant and indirect relationship.

Table 12: KM variables that impact capacity management.

KM	Coefficient	Pr(>|t|)	KM	Coefficient	Pr(>|t|)
PS2	-0.17438	0.0473	IS2	0.16199	0.0117
OS4	0.15421	0.0222	OC5	0.24729	0.0054
TG3	0.32457	0.0018	CM1	0.16080	0.0353
PP3	-0.14495	0.0190	CM2	0.12395	0.0146

Note: Adjusted R^2=073; F=13; p-value (model)=< $2e^{-16}$.

The statistical results of the model indicate that it is significant and viable as a whole. The model is represented by the following function: SS= -0.18 PS2 + 0.15 OS4 + 0.32 TG3 - 0.14 PP3 + 0.16 IS2 +

0.25 OC5 + 0.16 CM1 + 0.12 CM2, and presents an RMSE 1.38 which evidences a good fit of the model to the data. Table 13 reviews the statistical assumptions of the model, while Figure 5 shows these results graphically.

Table 13: Statistical assumptions.

Supposed	Value recommended	Value obtained	Evaluation
Normality of waste	P > 0.05	P= 0.005	Warning
Heterocedasticity	P > 0.05	P < 0.005	Warning
Autocorrelated residuals	P > 0.05	P = 0.030	Warning

Note: Statistics obtained from RSudio.

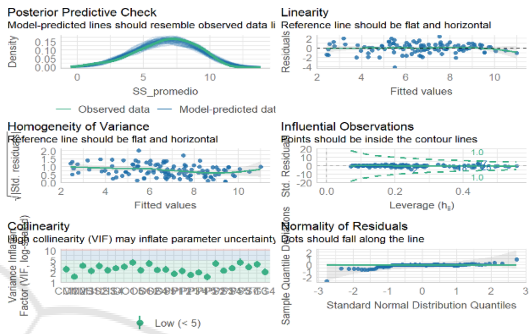

Figure 5: Graphs of statistical assumptions.

Note: Image obtained from RSudio.

4.6 Relationship Between KM and Technological Capabilities

The relationship between KM and technological capabilities was evaluated using a multiple linear regression model TC=f(X)+Ɛ. Table 14 shows that two KM variables belonging to the factors: Technology and People are significant and have a direct relationship with technological capabilities.

Table 14: KM variables that impact capacity management.

KM	Coefficient	Pr(>|t|)	KM	Coefficient	Pr(>|t|)
TG3	0.43720	1.8e-05	PP4	0.16856	0.0024

Note: Adjusted R^2=0757; F=14.8; p-value (model)=< $2e^{-16}$.

The statistical results of the model indicate that it is significant and viable as a whole. The model is represented by the following function: TC= 0.44 TG3 + 0.17 PP4, and presents an RMSE 1.27 which demonstrates a good fit of the model to the data. Table 15 reviews the statistical assumptions of the model, while Figure 6 shows these results graphically.

Table 15: Statistical assumptions.

Supposed	Value recommended	Value obtained	Evaluation
Normality of waste	P > 0.05	P= 0.405	Ok.
Heterocedasticity	P > 0.05	P= 0.440	Ok.
Autocorrelated residuals	P > 0.05	P = 0.376	Ok

Note: Statistics obtained from RSudio.

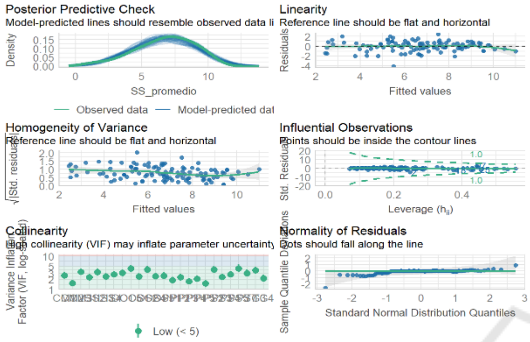

Figure 6: Graphs of statistical assumptions.

Note: Image obtained from RSudio.

In summary, the results allow us to affirm that the six models developed are viable. In all cases, the RMSE values obtained with the training data were similar to those obtained with the test data, which is a favorable indication of the robustness of the models and their adequate generalization capacity.

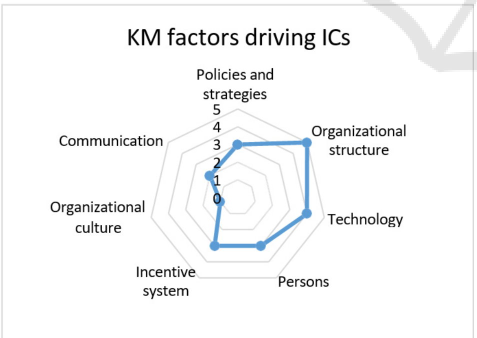

Figure 7: KM factors driving IC.

In addition, of the 35 KM variables, 12 have a significant impact on the innovation capabilities of manufacturing companies. Factors such as Organizational Structure, Technology, People and Incentive System are determinants for developing innovation capabilities in the industrial manufacturing sector, as shown in Figure 7. In addition, the Variance Inflation Factor (VIF) was found to be less than 5 in all six models, indicating the absence of significant multicollinearity. Likewise, a low correlation between the predictor variables was verified and no outliers were identified that would affect the analysis.

5 DISCUSSION

The results show that both organizational units dedicated to R&D and training offered as an incentive have a positive impact on IC in manufacturing firms. This result supports what has been proposed by Nonaka and Takeuchi (1995), who emphasize the importance of structures that facilitate the creation and transformation of knowledge, as well as by Teece (2018) and Jiménez-Jiménez and Sanz-Valle (2011), who point out that continuous training strengthens the organizational dynamic capabilities needed to innovate in competitive environments.

Regarding the relationship between KM and capabilities management, the results show that internal organizational units dedicated to R&D have a positive and statistically significant impact on innovation capabilities, in line with what was proposed by Nonaka & Takeuchi (1995), who emphasize the strategic role of structures that promote knowledge creation. In contrast, policies related to the storage, sharing and use of organizational knowledge show a negative and significant impact, which could reflect deficiencies in their design or an excessively rigid implementation that limits organizational flexibility and learning (Davenport & Prusak, 1998; Teece, 2018).

Regarding the relationship between KM and resource availability, the results are consistent with previous findings in the academic literature: they confirm the need to strengthen internal organizational structures dedicated to R&D to boost IC development within companies. Therefore, these results reinforce the importance of strategically investing in such structures as part of a comprehensive approach to knowledge management for innovation.

The results also show that KM, when articulated with practices such as the combination of existing and new knowledge, the strengthening of internal R&D structures, the use of ICT and incentivized training, has a positive impact on innovation capabilities. This finding is consistent with Nonaka and Takeuchi (1995), who highlight the importance of integrating tacit and explicit knowledge through organized and technological environments, as well as with Teece (2018) and Jiménez-Jiménez and Sanz-Valle (2011), who emphasize the role of continuous training and technological infrastructure as drivers of dynamic capabilities.

On the other hand, negative effects associated with the use of corporate social networks and the age of personnel were observed. In the first case, although some studies highlight their value for the informal exchange of knowledge (Levy, 2009), others warn that inappropriate use can lead to distraction or dispersion of knowledge (Treem & Leonardi, 2012). Regarding age, although experience is a relevant asset (Kanfer & Ackerman, 2004), research has evidenced that aging may be linked to lower technological adaptability or resistance to change (Ng & Feldman, 2012). These findings suggest the need for strategies that integrate generational diversity and optimize the use of digital collaborative tools to enhance organizational innovation.

Regarding the relationship between KM and personnel skills, it was found that variables such as the facilities for the horizontal flow of knowledge within the organization; the application of ICT for access, exchange and use of knowledge; the training offered as an incentive to generate, share and use knowledge; the empowerment of personnel for decision making; formal and informal communication in the work environment have a positive and significant impact on the development of innovation capabilities.

Although knowledge storage, sharing and use policies, along with process documentation, are often considered fundamental in KM (Davenport & Prusak, 1998), their negative impact on innovation may be due to overly formalized and inflexible implementation. This rigidity can limit autonomy, inhibit creativity and move organizational practices away from real collaborative dynamics (Nonaka & Takeuchi, 1995; Teece, 2018). It is therefore recommended to adopt a more adaptive and participative approach, which transforms these policies into facilitators of learning and innovation.

On the other hand, the results show that the application of ICT and staff proficiency in foreign languages have a positive and significant effect on IQ. This finding coincides with Davenport and Prusak (1998), who emphasize the key role of ICT in KM, and with Cohen & Levinthal (1990), who point out that language skills strengthen technological absorptive capacity and organizational learning, key factors for innovation.

5.1 Theoretical Implications

The results support Nonaka and Takeuchi's (1995) proposals on the interaction between tacit and explicit knowledge, and the role of organizational structures in knowledge creation. In addition, the dynamic capabilities model of Teece (2018), which relates organizational infrastructure, learning and strategic change, is empirically validated. The research differentiates the positive impact of variables such as R&D, training and ICT from the negative impact of overly rigid policies and formal structures, providing a more nuanced view of the functioning of KM.

Factors such as staff age and inadequate use of internal social networks are introduced as potential barriers to innovation, extending theories of resistance to change and digital distraction. In addition, the inclusion of language proficiency and ICT management as predictors of innovation aligns this research with the notions of "absorptive capacity" proposed by Cohen & Levinthal (1990).

5.2 Practical Implications

Manufacturing companies should redesign their knowledge storage and use policies, avoiding rigidities that hinder creativity and organizational learning. It is recommended to strengthen organizational structures dedicated to R&D, as their positive impact on innovation has been consistently validated.

The training offered as an incentive should be integrated as a systematic strategy to promote innovation and not as an isolated benefit, as well as technologies should be implemented with strategic criteria, promoting the exchange of knowledge without falling into the noise or digital distraction.

The companies analyzed should design programs that integrate employees of different ages, promoting the intergenerational transfer of knowledge and reducing barriers to technological adoption. On the other hand, encouraging foreign language proficiency among staff can increase access to global knowledge and improve technological absorption capacity.

6 CONCLUSIONS

The main objective of this study was the design and development of predictive models by means of multiple linear regression, applying supervised machine learning techniques, with the purpose of IC from the implementation of practices associated with KM in companies of the manufacturing sector. In the models developed, the independent variables corresponded to KM dimensions, while the dependent variables represented organizational IC.

After the process of filtering and eliminating variables with high collinearity, 25 variables linked to seven key factors of corporate governance were selected: policies and strategies, organizational structure, technology, people, incentive system, organizational culture and communication. Innovation capabilities were operationalized through indicators such as R&D capacity, management capacity, resource availability, talent management, personnel competencies and technological capacity.

Six regression models were developed, in which 12 KM variables with statistically significant effects on IC were identified. Among the factors with the greatest positive impact were organizational structure, technology, people and incentives. These results lead to the conclusion that the systematic, coherent and strategic implementation of KM factors in manufacturing companies can effectively predict the development of innovation capabilities in processes, products and services.

In summary, the findings of this research reinforce the strategic value of KM as a driver of organizational innovation and provide a useful empirical framework to guide decision making and the formulation of business policies that strengthen competitiveness in the industrial sector.

6.1 Limitations and Future Studies

One of the limitations of this study is that corporate governance is a relatively new topic for the management of Ecuadorian business organizations. This resulted in some difficulties in obtaining data from the companies studied. To mitigate this limitation, the surveys included sufficient introductory information to facilitate respondents' understanding and response to the questionnaire.

Based on the findings obtained in this research, it is suggested to extend the methodological approach by applying other machine learning algorithms, such as neural networks, support vector machines (SVM) or random forests, in order to contrast their predictive capacity with the multiple linear regression models used in this study.

On the other hand, it would be pertinent to expand the context of application of the model to other economic sectors, such as services, technology or agribusiness, to evaluate the generalizability of the results. Similarly, comparative studies between countries or regions could provide a broader understanding of how cultural, institutional and technological differences affect the effectiveness of KM practices on innovation capabilities.

REFERENCES

Aguilar-Barceló, J. G., & Higuera-Cota, F. (2019). Los retos en la gestión de la innovación para América Latina y el Caribe: un análisis de eficiencia. *Revista CEPAL*, *127*, 1–20. https://www.cepal.org/es/publicaciones/44570-retos-la-gestion-la-innovacion-america-latina-caribe-un-analisis-eficiencia

Akerkar, R. (2019). *Artificial Intelligence for Business*. Springer. https://doi.org/10.1007/978-3-319-97436-1_3

Bertram, D. (2018). *Likert Scales*. http://poincare.matf.bg.ac.rs/~kristina/topic-dane-likert.pdf

Boehmke, B., & Greenwell, B. (2020). *Hands-On Machine Learning with R*. Taylor & Francis.

Bogodistov, Y., Presse, A., Krupskyi, O. P., & Sardak, S. (2017). Gendering dynamic capabilities in micro firms. *RAE Revista de Administracao de Empresas*, *57*(3), 273–282. https://doi.org/10.1590/S0034-759020170308

Bolisani, E., & Bratianu, C. (2018). Generic Knowledge Strategies. In *Emergent Knowledge Strategies* (4th ed., Vol. 4, Issue July, pp. 147–174). Springer. https://doi.org/10.1007/978-3-319-60657-6

Burger, S. (2018). *Introduction to machine learning with R : rigorous mathematical analysis*.

Bykova, A., & Jardon, C. M. (2018). The mediation role of companies' dynamic capabilities for business performance excellence: Insights from foreign direct investments. the case of transitional partnership. *Knowledge Management Research and Practice*, *16*(1), 144–159. https://doi.org/10.1080/14778238.2018.1428070

Camisón-Haba, S., Clemente-Almendros, J. A., & Gonzalez-Cruz, T. (2019). How technology-based firms become also highly innovative firms? The role of knowledge, technological and managerial capabilities, and entrepreneurs' background. *Journal of Innovation and Knowledge*, *4*(3), 162–170. https://doi.org/10.1016/j.jik.2018.12.001

CEPAL. (2016). Ciencia, tecnología e innovación en la economía digital. La situación de América Latina y el Caribe. In *Segunda Reunión de la Conferencia de Ciencia, Innovación y TIC de la (Comisión Económica para América Latina y el Caribe*. https://doi.org/10.1017/CBO9781107415324.004

Chang, W. J., Liao, S. H., & Wu, T. Te. (2017). Relationships among organizational culture, knowledge sharing, and innovation capability: A case of the automobile industry in Taiwan. *Knowledge Management Research and Practice*, *15*(3), 471–490. https://doi.org/10.1057/s41275-016-0042-6

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, *35*, 128–152.

Costa-Climent, R., Haftor, D. M., & Staniewski, M. W. (2023). Using machine learning to create and capture value in the business models of small and medium-sized enterprises. *International Journal of Information*

Management, *73*(January), 102637. https://doi.org/10.1016/j.ijinfomgt.2023.102637

Davenport, T. H., & Prusak, L. (1998). Knowledge Generation. In *Working Knowledge: How Organizations Manage What They Know* (p. 226). Harvard Business School Press.

Davila, G., Varvakis, G., & North, K. (2019). Influence of strategic knowledge management on firm innovativeness and performance. *Brazilian Business Review*, *16*(3), 239–254. https://doi.org/10.15728/bbr.2019.16.3.3

Dutta, S., Lanvin, B., Lorena, R., & Wunsch-Vincent, S. (2021). Global Innovation Index 2021. In S. Dutta, L. Rivera León, & S. Wunsch-Vincent (Eds.), *Cornell, INSEAD, WIPO* (14th ed.). World Intellectual Property Organization. https://doi.org/10.34667/tind.44315

Espindola, D., & Wright, M. (2021). Leading a Culture of Change. In *The Exponential Era* (pp. 137–151). IEEE.

Hastie, T., Tibshirani, R., James, G., & Witten, D. (2023). An Introduction to Statistical Learning. In *Springer Texts* (2nd ed., Vol. 102).

Ibujés-Villacís, J., & Franco-Crespo, A. (2024). Relationship Between the Factors that Influence the Management of Innovation in the Manufacturing Industry of Pichincha, Ecuador. *Journal of Innovation Management*, *12*(3), 24–55. https://doi.org/10.24840/2183-0606_012.003_0002

Jiménez-Jiménez, D., & Sanz-Valle, R. (2011). Innovation, organizational learning, and performance. *Journal of Business Research*, *64*(4), 408–417. https://doi.org/10.1016/j.jbusres.2010.09.010

Kanfer, R., & Ackerman, P. L. (2004). Aging, Adult Development, and Work Motivation. *The Academy of Management Review*, *29*(3), 440. https://doi.org/10.2307/20159053

Kaur, V. (2019). Review of Literature. In *Knowledge Based Dynamic Capabilities. The Road Ahead in Gaining Organizational Competitiveness* (pp. 21–78). Springer. https://doi.org/10.1007/978-3-030-21649-8

Kesavan, P. (2021). Literature Review. In *Enablers of organisational learning, knowledge management, and innovation. Principles, process, and practice of qualitative data* (pp. 21–66). Springer. https://doi.org/10.1007/978-981-15-9793-0

Kodama, M. (2018). Collaborative Dynamic Capabilities: The Dynamic Capabilities View. In *Collaborative Dynamic Capabilities for Service Innovation - Creating a New Healthcare Ecosystem* (pp. 1–45). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-77240-0

Kuhn, M., & Silge, J. (2022). *Tidy Modelling with R* (1st ed.). O'Reilly Media, Inc. https://www.tmwr.org/

Lantz, B. (2023). *Machine Learning with R* (4th ed., Vol. 4, Issue 1). Packt Publishing Pvt Ltd. www.packt.com

Latpate, R., Kshirsagar, J., Kumar Gupta, V., & Chandra, G. (2021). Simple Random Sampling. In *Advanced Sampling Methods* (pp. 11–36). Springer. https://doi.org/10.1007/978-981-16-0622-9

Levy, M. (2009). WEB 2.0 implications on knowledge management. *Journal of Knowledge Management*, *13*(1), 120–134. https://doi.org/10.1108/13673270910931215

Lohr, S. L. (2019). Simple Probability Samples. In *Sampling. Design and Analysis* (2nd ed., pp. 25–72). CRC Press.

Manning, M. Jr., & Manning, M. Sr. (2020). Knowledge Assets Management. In *Total Innovative Management Excellence (TIME). The Future of Innovation* (pp. 354–398). CRC Press.

MIPRO. (2021). Cifras de industrias. In *Gobierno del Ecuador*. https://www.produccion.gob.ec/wp-content/uploads/2021/06/Presentación-Industria-Junio-2021.pdf

Nakamori, Y. (2020). Innovation Theory. In *Knowledge Construction Methodology. Fusing Systems Thinking and Knowledge Management* (pp. 1–18). Springer. https://doi.org/10.1007/978-981-13-9887-2_5

Ng, T. W. H., & Feldman, D. C. (2012). Evaluating Six Common Stereotypes About Older Workers with Meta-Analytical Data. *Personnel Psychology*, *65*(4), 821–858. https://doi.org/10.1111/peps.12003

Nonaka, I., & Takeuchi, H. (1995). *The knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press.

North, K., & Kumta, G. (2018). Knowledge in Organisations. In *Knowledge Management. Value Creation Through Organizational Learning* (2nd ed., pp. 33–66). Springer. http://www.springer.com/series/10099

Obeidat, B. Y., Al-Suradi, M. M., Masa'deh, R., & Tarhini, A. (2016). The impact of knowledge management on innovation: An empirical study on Jordanian consultancy firms. *Management Research Review*, *39*(10), 1214–1238. https://doi.org/10.1108/MRR-09-2015-0214

Ode, E., & Ayavoo, R. (2020). The mediating role of knowledge application in the relationship between knowledge management practices and firm innovation. *Journal of Innovation and Knowledge*, *5*(3), 210–218. https://doi.org/10.1016/j.jik.2019.08.002

OECD, & Eurostat. (2018). *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation* (4th ed., Issue October). OECD. https://doi.org/10.1787/9789264304604-en

Ott, R. L., & Longnecker, M. (2016). Inferences About Population Central Values. In *An Introduction to Statistical Methods & Data Analysis* (Seventh, pp. 232–299). Cengage Learning.

Pagani, M., & Champion, R. (2024). *Artificial Intelligence for Business Creativity*. Routledge. https://doi.org/10.4324/9781003287582

Qin, W. (2024). How to unleash frugal innovation through internet of things and artificial intelligence: Moderating role of entrepreneurial knowledge and future challenges. *Technological Forecasting and Social Change*, *202*(July 2023), 1–9. https://doi.org/10.1016/j.techfore.2024.123286

Salmador, M. P., Kaminska, R., & McKelvey, B. (2021). Digital dynamic capabilities. In *Management in the Age* *of Digital Business Complexity* (pp. 153–181). Routledge. https://doi.org/10.4324/9780429278211

SENESCYT-INEC. (2015). Principales indicadores de actividades de ciencia, tecnología e innovación. In *Secretaría de Educación Superior, Ciencia, Tecnología e Innovación*. https://www.ecuadorencifras.gob.ec/documentos/web-inec/Estadisticas_Economicas/Ciencia_Tecnologia-ACTI/2012-2014/presentacion_ACTI.pdf

Shmueli, G., Bruce, P., DeoKar, A., & Patel, N. (2023). *Machine Learning for Business Analytics* (Vol. 5, Issue 1). Wiley.

SUPERCIAS. (2020). Ranking de compañías. In *Ranking empresarial del Ecuador*. https://appscvs.supercias.gob.ec/rankingCias/

SUPERCIAS. (2021). Ranking de compañías. In *Superintendencia de Compañías, Valores y Seguros*. https://appscvs.supercias.gob.ec/rankingCias/

Szczepańska-Woszczyna, K. (2021). Innovation and Innovativeness. Determinants of the Innovativeness of an Organisation. In *Management Theory, Innovation and Organisation* (pp. 95–140). Routledge. https://doi.org/10.4324/9781003057123

Teece, D. J. (2018). The Foundations of Dynamic Capabilities. In *Dynamic Capabilities and Strategic Management* (pp. 82–112). Oxford University Press.

Treem, J. W., & Leonardi, P. M. (2012). Social media use in organizations: Exploring the affordances of visibility, editability, persistence, and association. *Communication Yearbook 36*, 143–189. https://doi.org/10.1080/23808985.2013.11679130

Weber, F. (2023). Artificial Intelligence for Business Analytics: Algorithms, Platforms and Application Scenarios. In *Artificial Intelligence for Business Analytics: Algorithms, Platforms and Application Scenarios*. Springer. https://doi.org/10.1007/978-3-658-37599-7

Zawislak, P. A., Fracasso, E. M., & Tello-Gamarra, J. (2018). Technological intensity and innovation capability in industrial firms. *Innovation & Management Review*, *15*(2), 189–207. https://doi.org/10.1108/inmr-04-2018-012.