

Explainability Analysis of the Evaluation Model of the Level of Digital Transformation in MSMEs based on Fuzzy Cognitive Maps

Jairo Fuentes

Universidad Francisco de Paula Santander, Cúcuta, Colombia

GIDITIC, Universidad EAFIT, Medellín, Colombia

fuentesjairo@gmail.com

and

Jose Aguilar

GIDITIC, Universidad EAFIT, Medellín, Colombia

CEMISID, Universidad de Los Andes, Mérida, Venezuela

IMDEA Networks Institute, Madrid, Spain

aguilar@ula.ve, jose.aguilar@imdea.org

and

Edwin Montoya

GIDITIC, Universidad EAFIT, Medellín, Colombia

emontoya@eafit.edu.co

and

William Hoyos

Grupo de Investigación ISI

Universidad Cooperativa de Colombia

Montería, Colombia

GIDITIC, Universidad EAFIT, Medellín, Colombia

william.hoyos@campussucc.edu.co

and

Diego Benito

IMDEA Networks Institute, Madrid, Spain

diego.benito@imdea.org

Abstract — the concept of digital transformation involves exploiting digital technologies to generate new ways of doing things in organizations, including the creation of new processes, models, and services that produce value based on the digitization of data and processes. The application of digital technologies enables organizations to develop capabilities for innovation, automation, etc., utilizing both established and emerging technologies, including the widespread use of artificial intelligence. This article proposes the implementation of Fuzzy Cognitive Maps (FCMs) based on experts and data for the evaluation of the level of digital transformation in MSMEs (Micro, Small and Medium Enterprises). Additionally, this work carries out an explainability analysis of the evaluation models based on FCMs. The main digital transformation variables used to define our FCMs were classified into five groups, based on the COBIT standard: i) Organization and Culture variables related to strategies, way of working, and ecosystems, ii) Customer variables related to services and digital channels and products, iii) Operations and Internal Processes variables related to supply chain, suppliers, and business model, iv) Information Technologies variables related to innovation, digitization, data and analytic. Finally, the fifth type of variable is the target, which indicates the level of digital transformation of the organization. Our model managed to infer the level of digital transformation of the organizations studied. Furthermore, the explanatory capacity of the FCMs developed in this work was explored using different explainability methods, some general and others specific to the FCMs. In general, the results obtained in the work are very encouraging since the quality metrics obtained with the prediction methods are very good, almost always higher than 98%; and the variability in the explainability techniques (very different between them), allow us to analyze in depth the behavior of the variables in the results obtained, something very important to understand how to improve the levels of digital transformation in organizations.

Keywords: Digital Transformation, Fuzzy Cognitive Maps, Explainability Analysis, Machine Learning.

1 Introduction

Digital transformation is the process of integrating digital technologies into the business model, processes, and culture of an organization to improve its strategies, and performance, among other things [1]. Digital transformation is a long-term endeavour that entails determining how technology is employed within an organization to affect its behaviour. Thus, digital transformation involves envisioning how digital technologies can improve service levels, reduce costs, and optimize inventory management, among other things, through the implementation of organizational changes using these technologies to achieve operational excellence. Advanced technologies, such as artificial intelligence, data analytics, and the Internet of Things, are integrated into the digital transformation process to enable organizations to collect and analyse real-time data, automate processes, and improve the supply chain.

On the other hand, FCMs are graph-based structures that support causal reasoning and enable the propagation of causality in both forward and backward directions [2]. They are particularly useful in domains where the concepts and relationships between them are inherently fuzzy, such as politics, military affairs, and historical events. They have been used in different areas where the modelling of the problem is very complex, such as health [3], analysis of social networks [4], or emerging processes [5], [21], among others. FCMs have had a lot of interest in the area of machine learning due to their explainability capabilities. In the field of explainability analysis methods, a classification can be made according to whether the explanation relates to the model and/or its output. This implies that explainability can be global, providing information on the inner workings

of the whole model for a data set, or local, focusing on a specific input A and its corresponding output Z. In general, many explainability analysis methods are currently being developed, but one of the challenges is knowing which ones are useful for each machine learning technique, such as which ones could be used in FCMs.

1.1 RELATED WORKS

At the level of FCM in classification and prediction tasks, the approach described in [3] combines different indicators used in the standard diagnosis of dengue, such as clinical symptoms, laboratory tests, and physical findings, to determine the severity of dengue. The model achieved an accuracy of 89.4% in diagnosing dengue, indicating performance very close to other well-established machine learning models. Also, to study the vulnerability factors to climate variability in Andean rural micro-watersheds, an FCM was employed in [4], the results of the FCM application did not present an important difference between the two climatic periods to which it was applied. Another work based on an FCM, and a fuzzy linguistic methodology, called VIKOR, determines the vendor selection in Industry 4.0 [5]. VIKOR enables the ranking and selection of suitable suppliers for specific tasks.

The investigation carried out in [6] analyzes the effect of digital transformation on the performance of MSMEs in terms of competitive advantage. The results of this research highlight the positive influence of digital transformation on competitive advantage and the overall performance of MSMEs. In [7], the goal is to develop a holistic maturity model for digital transformation using a theory-based approach, known as DX-CMM. This model aims to help organizations assess their current level of capability and maturity in digital transformation, identify gaps, and create a standardized roadmap for enhancing their digital transformation efforts. According to [6], technological advancements and digital information are expected to be utilized by traditional entrepreneurs. However, this study concluded that all proposed hypotheses were acceptable, and that competitive advantage partially mediated the relationship between digital transformation and MSMEs' performance. In [8], the authors propose a staged digital transformation capability maturity model framework that enables organizations to assess their present digital capability and establish a plan of improvements to move to a higher level. In [7] conducted a systematic literature review and found that none of the 18 existing maturity models in the field of digital transformation met all the criteria of appropriateness, completeness, clarity, and objectivity.

With respect to explainability Analysis for FCMs, the interpretability of FCMs has been confined to the fact that both the concepts and the weights have a well-defined meaning for the problem being modeled. This rather naïve assumption oversimplifies the complexity behind an FCM-based model. The work of Apostolopoulos et al. [35] studies the explainability properties of FCMs and presents critical examples from the literature that demonstrate their superiority in explainability tasks in various domains such as medical precision agriculture, decision support systems, and public policy formulation. Nápoles et al. [36] propose a post-hoc explanation method to estimate the relevance of concepts in FCMs. The proposal considers the dynamic properties of the FCM and uses the values of SHapley additive explanations (SHAP) for scenario analysis. To improve the interpretability of FCMs, in [37] the authors use Liang-Kleeman information flow (L-K IF) analysis, which identifies actual causal relationships from the data using an automatic causal search algorithm. In [38] they propose a symbolic explanation module that allows extracting

information and patterns from the FCM-based model, which can be seen as an inverse symbolic reasoning rule that infers the inputs that must be provided to the model to obtain the desired result. These relationships are then imposed as constraints on the learning process to rule out spurious correlations in the data to improve the aggregate predictive and explanatory power of the model. Mansouri et al. [39] propose using an FCM to develop simplified auxiliary models that can provide greater transparency to an LSTM model built to predict industrial bearing failures based on vibration sensor readings. The FCM mimics the performance of the baseline LSTM model and provides more information about the black-box model, such as (i) what variables contribute to the prediction result and (ii) what values could be controlled to avoid potential failures. Thus, the FCM is used to simplify the deep learning model and offer greater explainability. Finally, the article [40] proposes a temporal FCM (tFCM) model, which combines the predictive power of deep neural networks with the interpretability of FCMs. In the proposed tFCM model, cognitive states are modeled as fuzzy, multidimensional, and interrelated vectors, which are input into a short-term memory network for prediction. This hybrid model combines the ability of deep neural networks to discover latent factors with the ability of FCMs to reveal potential correlations.

According to the literature review, there is no digital transformation approach that allows companies to evaluate their existing digital competencies and create a roadmap to improve and reach a higher level. There are also no applications based on FCMs to analyze digital transformation processes in organizations, or that develop an explainability analysis process to understand the factors that may impact their organizational digital transformation. Therefore, there is a lack of research and products available for maturity models related to digital transformation.

1.2 CONTRIBUTIONS

The main objective of this research is to use FCM to evaluate the levels of digital transformation in MSMEs and explain possible actions to be taken to improve these levels. In this work, the causal relationships between digital transformation concepts are defined in two different ways, using data or according to expert opinions. The FCM is responsible for analyzing the factors associated with digital transformation data and providing a diagnosis. The study is based on five groups of concepts issued by the Governance and Management Objectives in COBIT [7]. The concepts are classified into five types and reflect the behavior of the company at all levels such as: i) Organization and Culture, ii) Customer, iii) Operations and Internal Processes, iv) Information Technology, and v) Objective, which indicates the digital transformation level of the organization. Also, the study considers several explainability analysis algorithms in the context of FCMs. The main contributions of this paper are:

- Develop several FCMs to determine the level of digital transformation in MSMEs.
- Carry out an explainability analysis for the FCMs that allows studying their results in the context of digital transformation.

This paper is organized as follows: Section 2 discusses the theoretical framework employed, Section 3 defines our FCM for the evaluation of the digital transformation levels in MSMEs, and Section 4 details the computational experiments conducted. Section 5 compares our explainability work with other papers in digital transformation, and Section 6 presents the conclusions and future works.

2. THEORETICAL BACKGROUND

This section presents a background of Digital Transformation. In addition, it provides the basic concepts of FCMs and explainability analysis in machine learning.

2.1 DIGITAL TRANSFORMATION

The application of digital technologies to generate significant changes in an organization with the goal of improving its operations is what defines digital transformation. [9]. Digital transformation is based on the adoption of digital technologies in an organization to transform its business processes. The interpretation of "digital technology" in the context of digital transformation differs across the literature. Some authors encompass new Information Technologies solutions [10], and others emphasize in cutting-edge technologies like big data, cloud computing, artificial intelligence, blockchain, and the Internet of Things [11], [20]. Also, some authors include social media analytics, and embedded devices [12].

As described in [13], there are two possible strategies for companies to undertake digital transformation. The first is an offensive approach that involves investments in digital transformation portfolios, while the second is a defensive approach that focuses on cultivating new capabilities within the organization.

COBIT (Control Objectives for Information and Related Technology) is a model for the characterization of business information technologies, promoted by ISACA (<https://www.isaca.org>) since 1996. The COBIT model for digital transformation classifies the concepts into four types:

a) Organization and Culture: variables related to the culture and management of the organization. The concepts of this variable used in this work are:

- *Strategy*, it refers to having an action plan for digital transformation and an understanding of the competitive environment,
- *Culture and management*: it refers to good practices at work, in the management of organizational processes and in the development of leadership, promoting human talent in its different areas and levels, creating an agile and innovative culture,
- *Way of working*: it refers to the type of face-to-face work and/or teleworking, and the practices and procedures implicit in them,
- *Leadership*: it refers to the managerial activities that organizational leaders have to influence the work team, and
- *Ecosystems*: it refers to the environment, not only internal, but also close external to the organization, which impacts the digital transformation process, to analyze its economic/social/cultural effect.

b) Customer: variables related to services (digital or not). The concepts of this variable used in this work are:

- *Digital channels* are the distribution channels of the different organizational strategic messages sent to interested parties.
- *Products*, services, and/or products offered by the organization as a result of its job or operation, and

- *Digital services* refer to the use of digital tools to improve the customer experience, based on a deep understanding of the customer and its needs.

c) Operations and Internal Processes: variables related to the organizational flow. The concepts of this variable used in this work are:

- *Supply chains* are the supply chains of goods and services that an organization requires for its operation,
- *Suppliers* refer to having providers of technological solutions, in addition to having advisors who accompany the digital transformation process of the organization, and
- *Business model* refers to the organizational capacity to explore new businesses based on technology, the development of new products that allow new markets to be captured, among other things.

d) Information Technology: variables related to the organization technology. The concepts of this variable used in this work are:

- *Innovation* refers to the business strategy of organizational transformation and human resource management based on reengineering, optimization, and process improvement,
- *Digitization* is the process of transforming analog or intangible processes, and physical or tangible elements, into digital ones,
- *Data and analytics* refer to the level of access to data that an organization has, the potential to monetize it by learning about its different uses, among other things, and
- Utilization of *technological solutions* involves implementing methods, systems, or tools that aid in the execution of organizational tasks with greater efficiency, as well as automating certain processes.

2.2 FUNDAMENTALS OF FCM.

In 1986, Kosko developed FCMs [14] from the cognitive maps of Axelrod's work in 1976 [27]. FCMs are a popular way of modeling complex systems because of their straightforward construction and interpretation. They consist of a network of concepts and their interrelationships [14], [15], [18].

2.2.1 BASIC DEFINITIONS OF AN FCM

In Figure 1, an FCM is shown, which consists of eight concepts (C1 to C8). Each concept denotes a specific variable of the system to study. The impact of one concept on another is displayed by the weight in the directed edges among them [16]. Different strategies can be employed to establish these relationships within FCMs. In [18], [22], [25] Aguilar et al. describe three ways to define the relationships in FCMs: using fuzzy rules, generic logical rules that describe the causal relationships, or mathematical models to describe the system under evaluation.

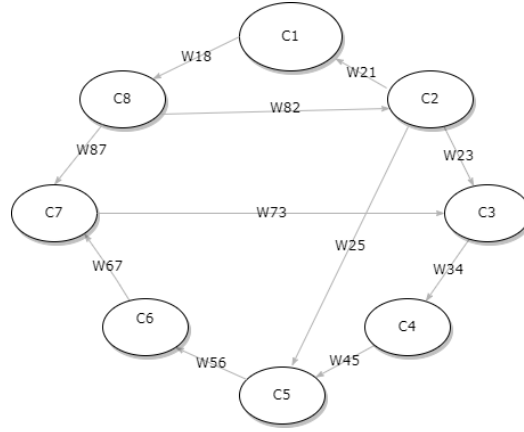


Figure 1. Example of an FCM.

The weight assignment in FCMs can be carried out using different methods, including those involving expert opinion or learning processes [16], [18], [22]. The weights assigned represent the influence of one concept on another, with values ranging from -1 to +1 or 0 to 1, indicating either activating or inhibiting impacts [2]. Thus, different levels of causality are represented. An example of degrees of causality using linguistic terms can be found in Table 1. For example, if the value of the antecedent concept C_i is significantly low, and the value of the consequent concept C_j is substantially high, then the causal link between the two concepts can be categorized as negatively high. Mathematically, an FCM is represented as a tuple of four elements:

$$\Phi = \langle n, f(\cdot), r, W \rangle. \quad (1)$$

where $C \in \mathbb{R}^m$ is the set of m concepts ($C_1, \dots, C_m$), $f(\cdot)$ is the activation function that holds concept values in a determined range r , and $W \in \mathbb{R}^{m \times m}$ is the adjacency matrix employed to capture the connections among the concepts.

Table 1. Example of generic logic rules used to classify the causal relationships based on the values of concepts.

Linguistic term	Numerical value	Generic logic rule	
		Antecedent C_i	Consequent C_j
Positive complete	1	Very high	Very high
Positive high	0.75	High	High
Positive moderate	0.5	Medium	Medium
Positive low	0.25	Low	Low
Null	0	NA	NA

There are various activation functions that can be used to model the activation of each concept in an FCM (see Table 2). The selection of a specific activation function depends on the specific problem being addressed. For example, if the goal is to model disease symptoms, the sigmoid function may be more appropriate than the hyperbolic tangent function since the values of the concepts being modeled will typically range from 0 (absence of the symptom) to 1 (presence of the symptom), making negative values unnecessary [18], [19].

Table 2. Activation functions used in FCMs.

Activation function	Equation	Range
Bivalent	$f(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$	$f(x) \in \{0, 1\}$
Trivalent	$f(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$	$f(x) \in \{-1, 0, 1\}$
Sigmoid	$f(x) = \frac{1}{1+e^{-kx}}$	$f(x) \in [0, 1]$
Hyperbolic tangent	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	$f(x) \in [-1, 1]$

W is used to store the relationships between the concepts in an FCM. For example, see the adjacency matrix for the FCM of Fig. 1 in Fig. 2.

$$W = \begin{matrix} & \begin{matrix} C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 & C_8 \end{matrix} \\ \begin{matrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \\ C_8 \end{matrix} & \left[\begin{array}{cccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{18} \\ w_{21} & 0 & w_{23} & 0 & w_{25} & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{34} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{54} & 0 & w_{56} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & w_{67} & 0 \\ 0 & 0 & w_{73} & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{82} & 0 & 0 & 0 & 0 & w_{87} & 0 \end{array} \right] \end{matrix}$$

Figure 2. Example of Adjacency Matrix.

2.2.2 REASONING IN FCMs

To mathematically define the inference process in an FCM, three key components are required: the weight matrix W, the activation function, and the current state vector of the FCM ($a \in \mathbb{R}^m$) which indicates the current degree of activation of each concept. The activation degree of a concept determines if the concept is activated or stimulated in a particular iteration. Thus, in the case of the range $[0, 1]$, a concept is "activated" when its value is 0 in one iteration (t) but becomes greater than 0 in the next iteration ($t + 1$).

To infer the output in an FCM, it is process is necessary to define a process that involves calculating the state vector through successive iterations of multiplying it with the weights matrix until the system reaches a steady state, as is defined by Equation 2 [14]:

$$a_j(t+1) = f\left(\sum_{i=1, i \neq j}^m w_{ij}a_i(t)\right) \quad (2)$$

where $a_j(t+1)$ is the value of concept C_j at iteration $t+1$, m is the number of concepts, w_{ij} is

the value for the relationship from concept C_i to concept C_j , and $a_j(t)$ is the value of concept C_j at iteration t . The point of equilibrium (steady state) is reached when $a_j(t) = a_j(t-1)$ or $a_j(t) - a_j(t-1) < 0.001$. Thus, the system is considered to have reached a steady state when the difference between the state vectors in consecutive iterations is less than or equal to 0.001.

This is the base of the reasoning mechanism of an FCM. FCMs are a crucial tool in simulation scenarios because they enable experts to analyze the system's performance from various starting points. An initial condition is defined by the set of concepts in the iteration (0) and is denoted as:

$$a(0) = [a_1(0), a_2(0), \dots, a_m(0)] \quad (3)$$

where $a_1(0)$ is the value of concept C_1 at iteration = 0.

2.3 EXPLICABILITY IN FCM

In the current landscape, the ability to provide explanations for results obtained using machine learning techniques is crucial, especially in critical sectors. The literature has developed various explainability methods [28]. Some of these methods are Local Interpretable Model-agnostic Explications (LIME) and feature importance, and the Weight-Filter FCM method proposed in [34]. In this work, we propose an explainability analysis based on these methods.

2.3.1 Explicability Analysis based on Weight-Filter FCM

Much of the existing research on FCM focuses mainly on the constructions of the models and examining the causal relationships between the concepts. There is a notable scarcity of work that delves into a more comprehensive analysis of explainability in FCM. Moreover, the limited studies that do exist often concentrate on employing local sensitivity to quantify the impact of slight perturbations, or changes in input values and weights on the model's predictive capacity.

In the article [34], the authors present a different method. The fundamental concepts of this approach are clarified below. The evaluation of sensitivity between concepts is carried out by eliminating low-impact causal connections between input concepts and target concepts. Thus, this approach aims to examine the causal impact exerted by less impactful concepts on the target concepts to which they are connected. The goal is to determine if this influence significantly affects the performance of the model. To achieve this, the least impactful causal connections (with fewer weights) are eliminated, establishing weight thresholds for this decision. Thus, for a given threshold value u , each causal relationship is characterized by a weight w_{it} where C_i represents an input concept and C_t is a target concept, if the weight w_{it} falls below the specified threshold u , then this causal relationship is eliminated in the model. Once these concepts have been eliminated, the quality of the FCM model with that new configuration is evaluated, and if the performance results do not degrade with that new model, then it is definitively eliminated.

Thus, if notable variations in results emerge between the initial model and the adjusted model, then it can be inferred that the removal of less impactful causal relations influences the performance of the model.

2.3.2 Explicability analysis based on LIME

LIME, introduced by Ribeiro et al. [32], is a method designed to elucidate individual predictions made by a model through the construction of local surrogate models for each specific prediction. The surrogate models are formulated by generating a set of instances in proximity to the target instance and assigning labels to them using the original machine learning model. Subsequently, these instances and labels are employed to train the surrogate model, which tends to be more interpretable. The primary objective of LIME is to provide an explanation for the prediction of a given instance. For each instance requiring an explanation, LIME generates a surrogate model. Given a specific instance, LIME investigates the impact on the prediction when variations occur within the variables of that instance. These variations constitute the neighborhood of the instance. Equation 4 illustrates the bases of LIME to obtain an explanation for a data point x .

$$\varepsilon(X) = \operatorname{argmin}_g (L(f, g, \pi_x) + \theta(g)) \quad (4)$$

Where g is the family of possible explanations, L is the loss function and it measures how close g (surrogate model) is to the prediction of f (original model) in its vicinity π_x , and $\theta(g)$ measures the complexity of the surrogate model

2.3.3 Explicability analysis based on Feature importance

The Feature Importance is a method that calculates the relative importance of each descriptor/feature, given a dataset, in the building of a model. The feature importance method is useful in several contexts such as Feature selection, Model interpretability, Model debugging, and Business decision-making, among others. In this work, we will use it to evaluate the model interpretability. The method provides several ways to classify features according to their contributions to the inference of the target variable. It is possible to use different strategies to analyze the importance of the features, such as the correlation with the target variable, the quality of the model according to each variable, among others. In this work, the sensitivity of the results to the permutation of their values will be used due to its simplicity, easy understanding and low computational cost. The conceptual bases of this method (permutation criterion) are [33]:

Let $X = [x_1, x_2, \dots, x_j, \dots, x_m]$ be an individual with m features, where x_j is the descriptor j of the individual X . Let $C = [c_1, c_2, \dots, c_k, \dots, c_p]$ be an array that contains the labels of the classes that any individual X belongs. In this sense, a possible structure of the dataset is shown in Figure 3.

$$\begin{pmatrix} x_1^1(1) & x_2^1(1) & \dots & x_j^1(1) & \dots & x_m^1(1) & c_1 \\ x_1^2(2) & x_2^2(2) & \dots & x_j^2(2) & \dots & x_m^2(2) & c_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_j^i(k) & \dots & \dots & x_j^i(k) & \dots & x_m^i(k) & c_k \\ x_j^i(p) & \dots & \dots & x_n^j(p) & \dots & x_n^i(p) & c_p \end{pmatrix}$$

Figure 3. Structure of the dataset.

Where $x_j^n(p)$ represents the descriptor j of the individual n in the class p . The permutation criterion algorithm consists of modifying the values of descriptor j , leaving the rest of the values fixed, to determine the performance metric of the model. Thus, each time the model is trained for a new value of the analysed characteristic (e.g., descriptor j) its performance metrics are calculated. This is done repetitively to analyse the sensitivity of the model to the variable.

From the above, the sensitivity of the model to each variable is obtained, such that if the results change a lot, then the model is very sensitive, otherwise, it is not very sensitive. This allows establishing a ranking between the variables, with the most sensitive being the most highly ranked.

3 OUR FCM TO EVALUATE THE LEVEL OF DIGITAL TRANSFORMATION IN COMPANIES.

This section details the process of constructing our evaluation model about the level of digital transformation in organizations utilizing FCMs. Our FCMs are based on experts and data to define the relationships. The construction of the FCMs is based on a methodology that consists of six steps [3], [23], [24]. The rest of this section briefly describes these steps applied to our FCM.

3.1 SELECTION OF EXPERTS AND PREPARATION OF THE DATASET

For the construction of the FCM based on experts, this study included the participation of five experts with more than five years of experience in digital transformation. They have experiences in the fields of management of digital transformation processes in different companies and government entities, in the study of the COBIT Model, in the field of business models and Digital Economy, Government Digitization, among other areas.

On the other hand, the dataset used to build the FCM is from the National Administrative Department of Statistics -DANE-, a Colombian entity responsible for the planning, collection, processing, analysis, and dissemination of statistics of companies and official organizations, at the national and international level [41]. The original dataset has 635 variables and corresponds to data from Colombian MSMEs. We selected from the dataset the 15 variables that the experts defined that are like the concepts of the COBIT model, so they are in turn relevant to characterize the Colombian MSMEs. These variables from the dataset are organizational methods, agile culture, collaborative work, work modality, work team, digital utility, products offered, distribution routes, supply chain, procedure, business pattern, analytics, novelty, digitalization and technologies. Finally, records with more than 80% missing values, for variables, are dropped. Additionally, we used the Synthetic Minority Oversampling Technique (SMOTE) to balance the classes [31].

3.2 CONCEPTS AND RELATIONSHIPS

In the case of the FCM based on experts, the initial phase of this process is carried out by experts in digital transformation. Based on the established concepts of the COBIT model (see Table 3), the experts constructed W with values ranging from 0 to 1 to indicate the relationships between the concepts. Table 3 shows an additional concept to those used in the COBIT model, which is the

concept of the digital transformation level. This concept is the concept that is inferred using the FCM. Using Table 1 as a guide, the experts selected the linguistic term that best represented the relationship between each concept, and then assigned numerical values to these terms to create the weight matrix.

Table 3. FCM Concepts for Digital Transformation.

Node	Name Concept
C1	Strategy
C2	Culture and Management
C3	Way of working
C4	Leadership
C5	Ecosystems
C6	Digital Services
C7	Products
C8	Digital Channels
C9	Suppliers and Supply Chain
C10	Business Model
C11	Processes
C12	Innovation
C13	Digitalization
C14	Technological Solutions
C15	Data Analytics
End Node	Digital Transformation

In the case of the learning approach (FCM based on data), we have used the FCM expert tool [2] to find the best FCM, which is based on PSO.

3.3 FCM DESIGN

In the FCMs developed by the experts in MSMEs, their knowledge is mixed into a single overall map to define the FCM. The process for producing a consolidated global map is described by the equation below [21], [22]:

$$E_{ij}^G = \sum_{e=1}^{NE} \frac{E_{ij}^e}{NE} \quad (4)$$

This equation calculates the overall weight for the FCM, represented by E_{ij}^G . The opinion of each expert e regarding the causal link between concepts C_i and C_j is represented by E_{ij}^e , and NE indicates the total number of experts who participated. For the construction of this FCM, we also have used the FCM Expert package (version 1.0.0) [2] but in the literature, there are others like

[22], [25], [26]. Figure 4 shows the final FCMs.

Also, as we said before, in the case of the FCM based on data, we have used the FCM expert package [2] to train the FCM, which uses a PSO for the training. Specifically, the dataset is divided into a part (80%) for the training of the FCM and the rest for the test phase.

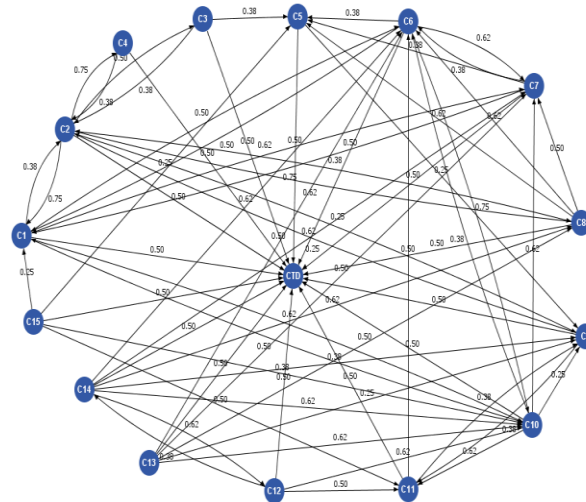


Figure 4. General FCM for evaluating the level of digital transformation in companies.

3.4 INFERENCE

The fourth step involves the utilization of the FCMs defined in this study to analyze different scenarios (for example, in our case, the level of digital transformation of different organizations) [18], [21]. In this step, the inference process is carried out for each case study, following the reasoning process of the FCMs. The experiment design and the results of this stage are shown in section 4.

3.5 INTERPRETATION

The interpretation of FCM inference results for each case study is carried out in this step and is detailed in section 4. In our case, we are going to use four explainability analysis methods, which are LIME, feature importance, Weight-Filter FCM, and an explainability analysis based on Classic FCM. In the latter case, since the way in which FCM establishes influence relationships between the input concepts and the target concepts is through the use of weights assigned to the causal relationships, they are used to identify the most influential characteristics in the model decision-making process.

3.6 DECISION

The last stage is the decision that the user of the FCMs can make to improve the digital transformation process in the organization based on their results. Since FCMs are now extended with explainability analysis methods, this new information can be used in the decision-making

process to identify exactly what to do.

4 EXPERIMENTS

This section aims to evaluate the efficacy of our FCMs in different scenarios to categorize the levels of digital transformation in multiple organizations. Two organizations are selected as case studies for this purpose and two FCMs are defined, one based on experts and another on data. Initially, the COBIT concepts are subject to an evaluation within each organization, to generate the input of the FCM inference process to ascertain the respective digital transformation levels.

4.1 COMPARISON OF THE RESULTS

4.1.1 USING THE TEST DATASET.

The results obtained from the application of the data-driven FCM on the dataset described above correspond to a precision of 1, an accuracy of 1, a recall of 1 and an F1 score of 1, which indicates an excellent performance for the prediction of the concept that represents the digital transformation. The hyperparameter configuration that allowed us to achieve this result consists of an initial population size of 70 individuals, the sigmoid function as the activation function, and the modified Kosko function as the reasoning inference function. In the case of the FCM based on experts, the results are an accuracy of 0.9835, precision of 0.98, recall of 1 and F1-Score of 0.99, which indicates a lower performance for the prediction of the concept that represents the level of digital transformation (see Table 4).

Table 4. Quality of the proposed FCMs.

	accuracy	precision	recall	F1-Score
Data-driven FCM	1	1	1	1
Expert-based FCM	0.9835	0.98	1	0.99

4.1.2 CASE 2. FOR THE CHAMBER OF COMMERCE.

The chamber of commerce acts as a collective body, bringing entrepreneurs and business owners together. In an effort to streamline business operations, Chambers of Commerce are broadening their range of offerings to encompass virtual services, allowing users to conduct their tasks with enhanced speed and convenience.

Table 5 shows the input values for concepts, where 0 is the absence of value in that concept and 1 is the maximum presence of that concept. In this organization, the current values of each concept of the COBIT model were determined as the average of what the company's experts thought. The final values are mentioned below: the concept of Strategy (C1) with a value of 0.7, Culture and Management (C2) of 0.7, Way of Working (C3) of 0.5, Leadership (C4) of 0.3, Ecosystems (C5) of 0.5, Digital Services (C6) of 0.7, Products (C7) of 0.6, Digital Channels (C8) of 0.8, Suppliers and Supply Chain (C9) of 0.4, Model of Business (C10) of 0.6, Processes (C11) of 0.4, Innovation (C12) of 0.8, Digitalization (C13) of 0.8, Technological Solutions (C14) of 0.9 and

Data Analytics (C15) of 0.4. With the weights averaged by the experts, we proceeded to infer the results in each FCM (see Table 6).

Table 5. Input values for the case study 2.

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
0,7	0,7	0,5	0,3	0,3	0,7	0,6	0,8	0,4	0,6	0,4	0,8	0,8	0,9	0,4

Table 6 shows the results of the inference processes. In the final vector, the digital transformation concept (CTD) reached a value of 0.9954 and 0.9979, respectively.

Table 6. Output vector for the case study 2.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	CTD
Expert based FCM	0,936	0,8878	0,7417	0,7754	0,889	0,9532	0,8539	0,8336	0,9502	0,9502	0,8832	0,741	0,638	0,7129	0,659	0,9954
Data driven FCM	0,9854	0,9626	0,9514	0,916	0,9691	0,9863	0,9793	0,9829	0,9761	0,9939	0,9778	0,9834	0,9666	0,971	0,977	0,9979

4.1.3 CASE 3. GALAVIS COFFEE COMPANY.

Galavis Coffe is a north Santander Company, founded in 1918, dedicated to the roasting and distribution of Colombian Coffee. Table 7 shows the input values for the concepts of the COBIT model of digital transformation. These values describe the current state of these concepts in the organization according to the company's experts. With this input, the results of the inference process in each FCM is defined in Table 8.

Table 7. Input table for the case study 3.

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
0,5	0,4	0,3	0,5	0,3	0,4	0,4	0,4	0,3	0,4	0,5	0,3	0,3	0,3	0,3

Thus, Table 8 shows the final vector of concepts. In this case, the CTD decision concept reached a value of 0.9954 and 0.9978, respectively.

Table 8. Output vector for the case study 3.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	CTD
Expert based FCM	0,936	0,8878	0,7417	0,7754	0,889	0,9532	0,8539	0,9636	0,9502	0,9502	0,8832	0,741	0,638	0,7129	0,659	0,9954
Data driven FCM	0,9821	0,9589	0,9501	0,9002	0,9622	0,9791	0,9703	0,9754	0,9712	0,9905	0,9708	0,9788	0,9616	0,9699	0,971	0,9978

The two FCMs reach a very close CTD value. The inference process (causal relationship) of both models in this case leads to very close values even though the behavior of the other concepts is not the same in each model (for the same input). For example, in the expert-based FCM, the relevant variables at the end are C1, C6, C8, C9, and C10, while in the Data-driven FCM are all.

4.2 EXPLICABILITY ANALYSIS

For the explainability analysis, LIME, Feature Importance and classic FCM methods were used for the data-based FCM. For the expert-based FCM, the weight filter FCM was used. In the case of expert-based FCM, the classic FCM method cannot be used since the weights that reached the target variable were all the same. Additionally, the LIME and Feature Importance methods require data in their analysis to update the FCM model after a training phase, which is not possible in expert-based FCM. In the case of the data-based FCM, the weight filter FCM method was not used because in this case, the model depends on all the variables with a certain relevance, in such a way that when applying any type of filter to eliminate some of them, the model no longer behaved adequately.

4.2.1 Explicability Analysis based on Weight-Filter FCM for expert-based FCM

Table 9 shows the adjacency matrix given by the average of the weights given by the experts in each of the relationships between the concepts. In this case, the most important characteristic is "C14" with an influence of 0.44, followed by "C13" with 0.40.

Table 9. Adjacency Matrix Average weights given by Experts.

		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	CTD	Average
C1	Strategy	0	0,312		0,625		0,375	0,25	0,25		0,625	0,125	0,375	0,25		0,562	0,5	0,341
C2	Culture and Imagemanagement	0,125	0	0,875	0,625	0,25			0,438	0,18	0,25	0,375	0,375		0,25	0,125	0,5	0,322
C3	Way of working	0,25	0,75	0	0,375	0,125	0,375	0,25	0,125		0,25	0,375	0,25	0,4375	0,25		0,5	0,293
C4	Leadership	0,5	0,5	0,25	0	0,125					0,25	0,1875	0,25			0,375	0,5	0,271
C5	Ecosystems	0,25				0	0,375		0,625	0,375	0,125	0,25	0,375	0,125	0,125	0,125	0,5	0,250
C6	Digital Services	0,5		0,25		0,625	0	0,5	0,375	0,25	0,75		0,375	0,25	0,25	0,25	0,5	0,365
C7	Products	0,375		0,25		0,25	0,375	0	0,437	0,375	0,125		0,25	0,437	0,25	0,25	0,5	0,281
C8	Digital Channels	0,25	0,625	0,25		0,75	0,375	0,625	0		0,375		0,375	0,437	0,375	0,187	0,5	0,385
C9	Suppliers and Supply Chain		0,375	0,25				0,25		0	0,125	0,75	0,125	0,375	0,437		0,5	0,299
C10	Business model	0,625	0,25	0,25	0,25		0,625	0,375	0,25	0,5	0	0,625	0,625	0,125	0,25	0,375	0,5	0,366
C11	Processes	0,25	0,25	0,25		0,25	0,625	0,25	0,25	0,875	0,25	0	0,25	0,187	0,375	0,375	0,5	0,317
C12	Innovation	0,25	0,25			0,25	0,5		0,25	0,25	0,5	0,625	0		0,375	0,25	0,5	0,318
C13	Digitization					0,25	0,625	0,375	0,375	0,375	0,75		0,25	0	0,375	0,687	0,5	0,406
C14	Technological solutions			0,25		0,25	0,5	1	0,75	0,75	0,5		0,25	0,25	0	0,375	0,5	0,443
C15	Data and Analytics	0,75			0,25	0,375		0,25	0,25		0,5	0,75	0,25	0,5	0,25	0	0,5	0,375

The thresholds have been defined considering the distribution of the weights of the causal relationships given by the experts. This analysis allows us to evaluate the influence of concepts with lower weights by observing whether the behavior of the model changes or remains the same. Table 10 summarizes the filters selected. Each filter is more restrictive than the previous one. For example, Filter 1 is responsible for removing all causal relationships associated with the target concepts (classes) whose weight is less than or equal to 0.125. Filter 2 suppresses causal

relationships with weights less than 0.25, Filter 3 eliminates all causal relationships with weights less than 0.50, while Filter 4 suppresses causal relationships with weights less than 0.75.

Table 10. Filters based on the average of the weights of the causal relationships with respect to the target variable.

Filter 1	≤ 0.125
Filter 2	< 0.25
Filter 3	< 0.50
Filter 4	< 0.75

For filters 1, 2, and 3, the results of the inference process have a small variation between the relationships of the concepts (see the value of CTD in Table 11 for filter 1). This tells us that we cannot clearly extract a feature that is more relevant than the others.

Table 11. Results with filter 1.

Inference results																
Plotter Table																
Step	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	CTD
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
2	0.8741	0.8081	0.7186	0.7549	0.7663	0.8741	0.8267	0.7663	0.8267	0.9284	0.867	0.6926	0.6792	0.6225	0.7548	0.9859
3	0.9635	0.9006	0.8062	0.8589	0.8684	0.956	0.9141	0.8471	0.9253	0.9843	0.9555	0.7812	0.742	0.6508	0.8472	0.999
4	0.9735	0.9216	0.8312	0.8833	0.891	0.966	0.9291	0.8674	0.9394	0.9889	0.9673	0.8016	0.7624	0.6572	0.8697	0.9994
5	0.9748	0.9261	0.8372	0.8877	0.8952	0.9675	0.9316	0.8718	0.9412	0.9896	0.969	0.8053	0.768	0.6586	0.8744	0.9995
6	0.975	0.9271	0.8386	0.8885	0.896	0.9678	0.9321	0.8727	0.9415	0.9897	0.9693	0.8059	0.7695	0.659	0.8754	0.9995
7	0.9751	0.9273	0.8389	0.8887	0.8962	0.9679	0.9322	0.8729	0.9415	0.9897	0.9694	0.806	0.7698	0.659	0.8757	0.9995

For filter 4, when performing the inference, it results in a significant variation in the results of CTD. At the end, the most important feature is “C10” with an influence of 0.91, “C9” with an influence of 0.89, followed by “C11” with an influence of 0.88 (see Table 12). In general, it is observed that the other concepts (“C4”, “C5”, “C6”, “C8”, “C12”, “C13”, “C14” and “C15”) do not have a relevant contribution to the prediction.

Table 12. Result of weights with filter 4.

Inference results																
Plotter Table																
Step	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	CTD
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
2	0.7549	0.7058	0.7186	0.6225	0.7058	0.6225	0.7311	0.7058	0.7879	0.8176	0.7773	0.6225	0.6225	0.6225	0.6225	0.6225
3	0.8319	0.7764	0.7919	0.6508	0.7747	0.6508	0.7947	0.7636	0.8738	0.8966	0.8624	0.6508	0.6508	0.6508	0.6508	0.6508
4	0.8502	0.7974	0.8132	0.6572	0.7937	0.6572	0.8093	0.7776	0.8925	0.9106	0.8814	0.6572	0.6572	0.6572	0.6572	0.6572
5	0.8542	0.8033	0.8192	0.6586	0.7985	0.6586	0.8125	0.7808	0.8963	0.9131	0.8853	0.6586	0.6586	0.6586	0.6586	0.6586
6	0.8551	0.805	0.8209	0.659	0.7997	0.659	0.8132	0.7816	0.8971	0.9135	0.8861	0.659	0.659	0.659	0.659	0.659
7	0.8553	0.8054	0.8213	0.659	0.7999	0.659	0.8134	0.7817	0.8972	0.9136	0.8863	0.659	0.659	0.659	0.659	0.659

4.2.2 Explicability Analysis based on Classic FCM for data-based FCM

According to the FCM model generated with the dataset (see Figure 5), the most important feature is "C1" with an influence of 0.52, followed by "C5" with 0.44. "C12" and "C15" both have an influence of 0.42 and 0.40, respectively. Finally, "C8" has an influence of 0.36. We see that there are no common relevant concepts with respect to those defined in section 4.2.1.

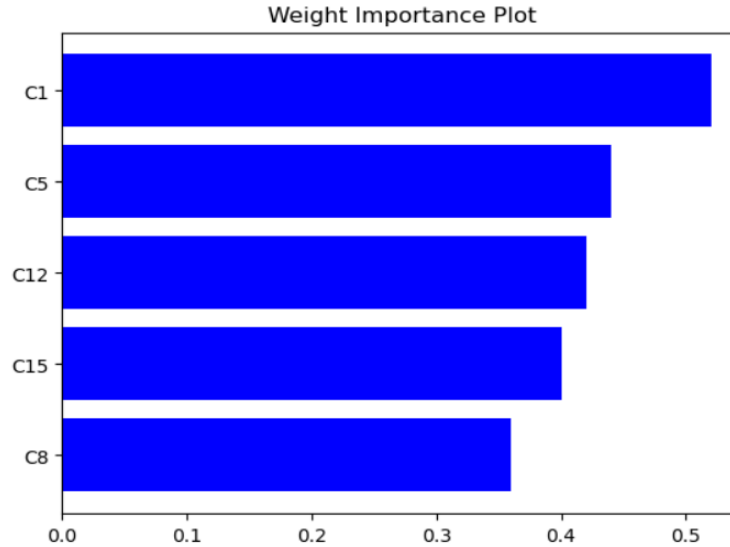


Figure 5. Weight importance Plot.

It is good to highlight that both explicability methods for FCM give different results in the explicability analysis. Now, the reason is that one is based on eliminating the concepts with less weight in their relationship with CTD (section 4.2.1) such that it does not affect the quality of the inferences, while the one in this section is left the concepts with greater weight in the relationship with CTD. Thus, opposite processes follow, but despite that, they agree on some relevant concepts, for example, C1 and C15.

4.2.3 Explicability analysis based on LIME for data-based FCM

Next, we present an example of the analysis that can be done for each class in each dataset, to study the behavior of each variable in a given model. Figure 6 illustrates the calculation of the LIME values for the prediction of the class 1.

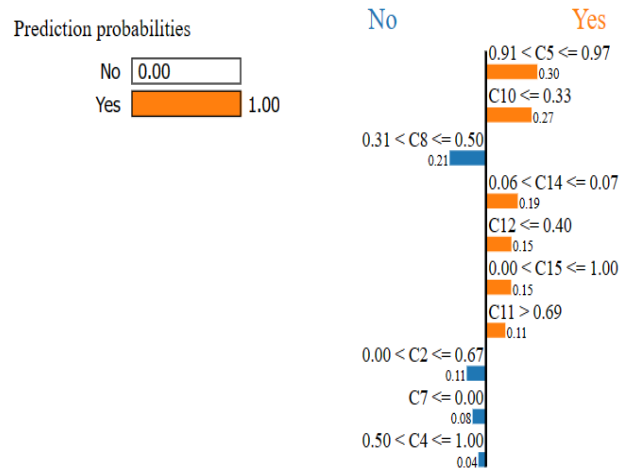


Figure 6. Feature Importance based on LIME for class 1 for data-based FCM.

```
[('0.91 < C5 <= 0.97', 0.3001793956564),
 ('C10 <= 0.33', 0.26649081469982155),
 ('0.31 < C8 <= 0.50', -0.2120230535399748),
 ('0.06 < C14 <= 0.07', 0.1858697854062809),
 ('C12 <= 0.40', 0.1486919740988927),
 ('0.00 < C15 <= 1.00', 0.14711802758805537),
 ('C11 > 0.69', 0.11321779302762842),
 ('0.00 < C2 <= 0.67', -0.11210549710522849),
 ('C7 <= 0.00', -0.07796034786804207),
 ('0.50 < C4 <= 1.00', -0.04413467707423912)]
```

Figure 7. LIME for class 1 of the data-based FCM

To analyze the results presented in Figure 6, the following is taken into consideration:

- The order of the features reflects their weight in the model's prediction. This shows a ranking of which features are more important for obtaining that prediction. In this case, the variable "C5" has 30% importance, "C10" with 27%, "C8" with 21%, "C14" with 18%. and "C12" has 15%. These are the most influential features.
- The colour of each feature and the x-axis indicate how it positively or negatively influences the prediction. The variables "C5", "C10" and "C14" (orange color) have a positive influence on the prediction, while the variables "C8", and "C2" (blue color) have a negative influence.
- The value next to each feature indicates the threshold from which it positively or negatively contributes to the prediction of the class. For example, values between 0.91 and 0, 97 for "C5" have a positive impact on the prediction. Also, values within < 0.31 and ≤ 0.50 of "C8" have a negative contribution to the prediction.
- Figure 7 summarizes all the information shown in Figure 6, showing the list of variables, from which values impact the results, and the degree of importance of each variable.

Based on the information explained above and Figure 7, we can extract the following conclusions in terms of explainability:

- "C5" is the variable with the highest positive weight in the prediction. Values in the interval (0.91, 0.97] contribute to the prediction of class 1 of the dataset.
- "C10" is the second most positively weighted feature in the prediction. Values lower than the 0.33 threshold contribute to the prediction of class 1.
- "C8" is the variable with the highest negative weight in the prediction, Values in the interval (0.31,0.50] contribute to a prediction of class 0.
- "C14" and "C15" are the third and fourth features with the highest positive weight in predicting class 1. Values in the intervals (0.06,0.07] and (0,1], respectively, contribute to the prediction of class 1.
- "C2", "C7" and "C4" have a significant negative impact on the prediction of class 1; however, their influence and weight are lower than the rest of the variables already discussed.
- The rest of the variables have a little significant impact on the prediction.

Based on a similar feature importance analysis for the rest of the classes, we establish the overall importance of the features in each of the datasets. Figure 8 shows the ranking of the most important features for the FCM model.

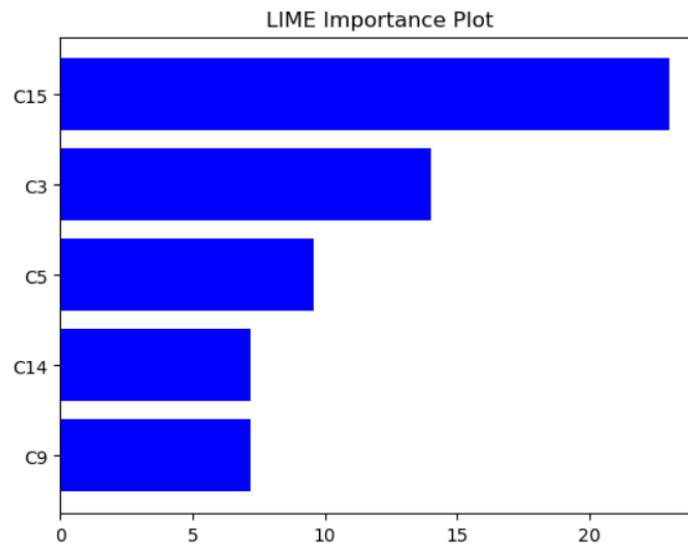


Figure 8. Ranking of the most important features for the FCM model.

It is observed that the "C15" characteristic is the most important for the model, with 23%. It is followed by "C3" with 14%. Next are "C5" with 9.6%, followed by "C9" and "C14" both with 7.2%

4.2.4 Explicability analysis based on Feature Importance for data-based FCM

In this section, we proceed with the explainability analysis using the feature importance method based on permutation. Figure 9 shows the most relevant characteristics according to this analysis.

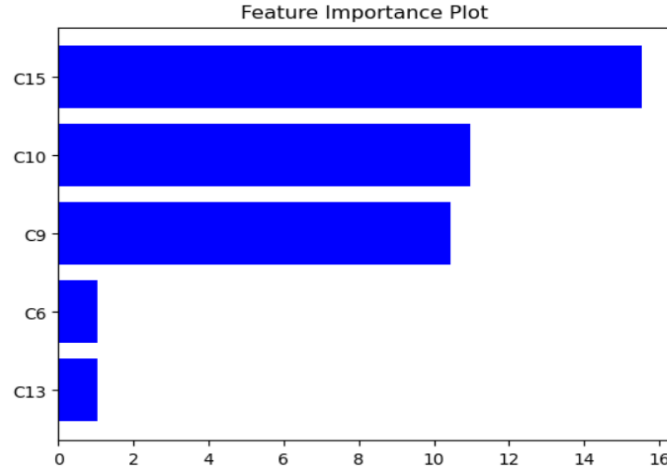


Figure 9. Most relevant characteristics according to the analysis.

In this case, the most important feature is “C15” with a value of 15. It is followed by “C10” with 11, and “C9” with a value of 10. Finally, “C6” with 10, and “C13” with 1.04.

4.2.5 Comparison between Explainability Methods

In Table 13, we see the most important features that each method provides. First, we note that the FCM models give a different ranking of the most relevant characteristics, but even the methods for the same model as well (for example, in the case of FCM based on Data). Now, the only feature that appears in all cases, and even in two with the highest value, is C15. Other features that appear several times are C10 and C5, twice between the first 3, and C3, C9, C6 and C14. In that sense, C15 (Data Analytics) mainly, and C5 (Ecosystems) and C10 (Business Model), are essential according to our analyzes for a high level of digital transformation in an organization.

Table 13. Feature Importance by Method.

Method	First Feature	Second Feature	Third Feature	Others
Expert-based FCM and Weight-Filter FCM	C10	C8	C11	C4, C5, C6, C8, C12, C13, C14 and C15.
Data-based FCM and Classic FCM	C1	C5	C12	C15
Data-based FCM and LIME	C15	C3	C5	C9 and C14
Data-based FCM and Feature Importance	C15	C10	C9	C6 and C13

Explainability methods can produce different results regarding feature importance due to variations in their approaches, assumptions, and limitations. These discrepancies can lead to differences in determining which features are considered most crucial to the model. Numerous authors, including Lundberg and Lee [42], and Ribeiro et al.[43], provide detailed comparisons illustrating how different explainability methods produce variable results. Given these differences in results between explainability methods, an analysis of feature importance performance was performed for each method. The RemOve And Retrain (ROAR) method was used [44]. ROAR is a method used to evaluate the importance of features, in which the subset of the input features considered most critical in a given explainability method is removed, and replaced with a constant average value. By observing how model performance changes when these features are removed and replaced, ROAR provides information about the relative importance of different input features in the model's predictions. A degradation rate is calculated, defined as the difference in performance between the model trained with all features and one trained without a proportion of the features considered most important by each explainability method. In our case, the 3 most important characteristics were considered according to each method, as shown in Table 13. Precision was used as a metric to calculate the degradation rate. The method with the highest degradation rate is considered to suggest the importance of the features that best fit the predictive behavior of the model.

Table 14. ROAR results by method.

Method	Accuracy	ROAR Accuracy	Degradation Rate
Expert-based FCM and Weight-Filter FCM	1	0.6236	0.3764
Data-based FCM and Classic FCM	0.9835	0.4681	0.5154
Data-based FCM and LIME		0.6910	0.2925
Data-based FCM and Feature Importance		0.8357	0.1478

Table 14 shows the degradation indices obtained when evaluating each of the explainability methods. It is noteworthy that *Weight-Filter FCM for Expert-based FCM* and *Classic FCM for Data-driven FCM* have the most important degradations, 0.3764 and 0.5154, respectively. This suggests that these methods provide the strongest explanations in terms of ROAR. This result is consistent since they are methods that arise directly from FCM.

5 COMPARISON WITH OTHER WORKS IN DIGITAL TRANSFORMATION

To compare this study with other works, a set of four criteria were established, which are outlined below:

- Cri1: It uses non-intrusive schemes for the evaluation.
- Cri2: It uses machine learning in the evaluation model.
- Cri3: It explains the reasons for the level deduced of digital transformation.
- Cri4: It uses different types of explainability methods.

Table 15 presents a comparison of this study with previous works based on these criteria that are deemed pertinent in the context of agroindustrial MSMEs and their digital transformation. These criteria are significant as they pertain to the application of automated technologies in agroindustry and the utilization of machine learning to enhance the agroindustrial production process. Using machine learning in the evaluation model allows identifying patterns among the data to make predictions. By explaining the reason for the level of digital transformation inferred will allow organizations to assess their current digital capability and establish an improvement plan to move to a higher level.

Table 15. Comparison with other works.

	Cri1	Cri2	Cri3	Cri4
[3]	√	√	X	X
[4]	X	X	X	X
[5]	X	X	X	X
[6]	X	X	√	X
[7]	X	X	√	X
[8]	X	X	√	X
[9]	X	X	√	X
[29]	X	√	X	√
[30]	X	√	X	√
[31]	√	√	X	√
This work	√	√	√	√

The first criterion is met by the studies [3], [31] and ours since both use computerized methods and models for evaluation. The second criterion is met by [3],[29], [30], [31] and our study, which uses machine learning for different purposes, such as information processing for dengue diagnosis [3]. The third criterion is met by studies [6], [7], [8], [9]. They explain the

relationship between digital transformation and MSME performance, conduct digital transformation maturity assessments, and design a comprehensive understanding of digital transformation and its impact.

With respect to the fourth criterion, in [29] was established that the results obtained through FCMs are strongly linked to relationships with high weights. The work [30] addresses the evaluation of the robustness of the FCM using different uncertainty propagation analysis methods. Both studies use the local sensitivity of the concepts as an indicator of the quality of the robustness, measuring the variations in the predictions by changing the weights of the relationships. Within the scope of fuzzy models, [31] presents a rule-based fuzzy explanatory system designed especially for implementation in deep neural networks.

Our study meets all the requirements of an evaluation model to analyze the digital transformation in an organization. Our model is not intrusive because it uses organizational experience to evaluate the current state of certain organizational variables and, from there, uses a computer tool to make the evaluation. It uses the FCM theory as a machine learning technique, and different methods for an explainability analysis. Using our explainability analysis methods proposed in section 4.2, FCM explains the conclusions reached through the causal relationships that exist between the concepts/variables.

6 CONCLUSIONS

The purpose of this study was to develop FCMs to determine the level of digital transformation in MSMEs, and to carry out an explainability analysis process to understand the results returned by the FCMs. A computational tool is proposed to analyze the variables, which are divided into five specific types: i) Organization and Culture, ii) Customer, iii) Operations and Internal Processes, iv) Information Technologies, and v) Objective, which indicates the level of transformation digital in an organization according to the values in the previous variables in this organization.

The FCM allowed analyzing the behavior of factors associated with digital transformation in MSMEs. Our approach is an explicable and interpretive technique that permits the explanation of results based on the diverse concepts in the digital transformation of the COBIT model. This approach is very useful because it allows the detection of the concepts that need to be strengthened or improved in an organization to strengthen its digital transformation process to be more competitive. The developed model is easily customizable and flexible, making it simple to integrate additional concepts and relationships. For its use, only an initial vector of the current state of the variables related to the digital transformation and the weight matrix is required.

In the case studies of agroindustrial MSMEs and the chamber of commerce, our FCM-based tool allowed determining the variables related to digital transformation to improve. The quality of the FCM-based models for predicting the level of digital transformation in an organization was greater than 98% in all quality metrics. Specifically, it was found that the variables associated with digitization, data and analysis, innovation, and technological solutions, were those that should be improved. In turn, our model allowed determining for the MSMEs of the agro-industrial sector analyzed, the variables that should have a high value at the end because they influence the most to have a high level of digital transformation in companies.

Additionally, an explainability analysis was carried out in the different FCM models developed. In the case of expert-based FCM, the most relevant variables are C10, C8 and C11. In the case of the data-based FCM, where more explainability methods were used, the most relevant variables in general given by them were C5 and C15. We see that the explainability results do not coincide between the methods, and of course, even less so when the different FCMs developed are considered. Due to that, we analyzed the quality of the explainability methods using the ROAR method, and the best were *Classic FCM for Data-based FCM* and *Weight-Filter FCM for Expert-based FCM*.

Among the main limitations of this work is the lack of large databases for the learning process of the Fuzzy Cognitive Map (they were tested with those obtained in the literature). Another limitation has to do with the complexity of testing the model in an organization to assess its level of digital transformation due to the lack of methodologies that allow capturing the required information.

Future work should develop new metrics to evaluate the quality of explainability provided by the methods. With respect to their quality evaluation, Table 14 is one example of that comparison based on a method of evaluating their quality, which should be extended in the future. Also, future work should develop methods that allow hybridizing the analysis provided by them. Other future research should focus on developing methodologies to evaluate the levels of digital transformation in organizations using tools like ours. Furthermore, since the dataset used in this study is from a single city in Colombia, distributed strategies, such as federated learning, must be considered to create a global model, in this case of a country, or world region, that includes data from different sites.

Acknowledgments

All the authors wish to thank the EAFIT University's Vice President for Discovery and Creation for their support in this research. Likewise, the authors would like to thank the HAMADI 4.0 project, code 22-STIC-06, of the STIC-AmSud regional program, for supporting this work.

References

- [1] M. Bach, M. Spremić, D. Vugec, "Integrating Digital Transformation Strategies into Firms: Values, Routes and Best Practice Examples," *Manag. Technol. Challenges Digit. Age*, pp. 107–128, 2018, doi: 10.1201/9781351238922-5.
- [2] G. Nápoles, M. L. Espinosa, I. Grau, K. Vanhoof, "FCM Expert: Software Tool for Scenario Analysis and Pattern Classification Based on Fuzzy Cognitive Maps," *Int. J. Artif. Intell. Tools*, vol. 27, no. 7, 2018, doi: 10.1142/S0218213018600102.
- [3] W. Hoyos, J. Aguilar, M. Toro, "A clinical decision-support system for dengue based on fuzzy cognitive maps," *Health Care Manag. Sci.*, vol. 25, no. 4, pp. 666–681, 2022, doi: 10.1007/s10729-022-09611-6.
- [4] I. Bolaños-Portilla, I. Hurtado-Sánchez, I. Restrepo-Tarquino, "Diffuse cognitive maps for analysis of vulnerability to climate variability in andean rural micro-watersheds," *DYNA*, vol. 87, no. 212, pp. 38–46, 2020, doi: 10.15446/dyna. v87n212.79943.
- [5] A. Hosseini, J. Heidary, J. Antucheviciene, M. Azari, S. Razavi, "Supplier selection in the industry 4.0 era by using a fuzzy cognitive map and hesitant fuzzy linguistic VIKOR methodology," *Environ. Sci. Pollut. Res.*, 2023, doi: 10.1007/s11356-023-26004-6.
- [6] N. Yuliantari, N. Pramuki, "The Role of Competitive Advantage in Mediating the Relationship Between Digital Transformation and MSME Performance in Bali," *J. Ekon. Bisnis Jagaditha*, vol. 9, no. 1, pp. 66–75, 2022, doi: 10.22225/jj.9.1.2022.66-75.
- [7] E. Gökalp, V. Martinez, "Digital transformation maturity assessment: development of the digital transformation capability maturity model," *Int. J. Prod. Res.*, vol. 60, no. 20, pp. 6282–6302, 2022, doi: 10.1080/00207543.2021.1991020.
- [8] T. Aguiar, S. Gomes, P. Da Cunha, M. Da Silva, "Digital transformation capability maturity model framework," in *Proc. IEEE 23rd Int. Enterp. Distrib. Object Comput. Conf.*, pp. 51–57, 2019, doi: 10.1109/EDOC.2019.00016.
- [9] V. Veldhoven, J. Vanthienen, "Designing a Comprehensive Understanding of Digital Transformation and its Impact". In *proc. Bled eConference*. 2019.
- [10] D. Sen, M. Ozturk, O. Vayvay, "An Overview of Big Data for Growth in SMEs," *Procedia - Soc. Behav. Sci.*, vol. 235, pp. 159–167, 2016, doi: 10.1016/j.sbspro.2016.11.011.
- [11] L. Morales Escobar, J. Aguilar, A. Garcés-Jiménez, J. Gutierrez De Mesa, J. Gomez-Pulido, "Advanced Fuzzy-Logic-Based Context-Driven Control for HVAC Management Systems in Buildings," *IEEE Access*, vol. 8, pp. 16111–16126, 2020, doi: 10.1109/ACCESS.2020.2966545.
- [12] A. Singh, T. Hess, "How chief digital officers promote the digital transformation of their companies," *Mis. Exec.*, vol. 16, no. 1, pp. 1–17, 2017, doi: 10.4324/9780429286797-9.
- [13] A. Margiono, "Digital transformation: setting the pace," *J. Bus. Strategy*, vol. 42, no. 5, pp. 315–322, 2020, doi: 10.1108/JBS-11-2019-0215.
- [14] B. Kosko, "Fuzzy cognitive maps," *International Journal of Man-Machine Studies*, vol. 24, no. 1. pp. 65–75, 1986, doi: 10.1016/S0020-7373(86)80040-2.
- [15] J. Aguilar, "A fuzzy cognitive map based on the random neural model," *Lect. Notes in Comput. Sci.*, vol. 2070, pp. 333–338, 2001, doi: 10.1007/3-540-45517-5_37.
- [16] G. Nápoles, I. Grau, L. Concepción, L. Koutsoviti Koumeri, J. Papa, "Modeling implicit bias with fuzzy cognitive maps," *Neurocomputing*, vol. 481, pp. 33–45, 2022, doi: 10.1016/j.neucom.2022.01.070.
- [17] G. Xirogiannis, M. Glykas, C., Staikouras "Fuzzy Cognitive Maps in Banking Business Process Performance Measurement". In *Fuzzy Cognitive Maps. Studies in Fuzziness and Soft Computing* (Glykas M. ed.), Springer, pp. 161–200, 2010.

- [18] H. Sánchez, J. Aguilar, O. Terán, J. Gutiérrez de Mesa, "Modeling the process of shaping the public opinion through Multilevel Fuzzy Cognitive Maps", *Applied Soft Computing*, vol. 85, 2019, doi: 10.1016/j.asoc.2019.105756.
- [19] J. Salmeron, S. Rahimi, A. Navali, A. Sadeghpour "Medical diagnosis of Rheumatoid Arthritis using data driven PSO-FCM with scarce datasets". *Neurocomputing*, pp.:104–112. 2017, doi: 10.1016/j.neucom.2016.09.113.
- [20] L. Morales, C. Ouedraogo, J. Aguilar, C. Chassot, S. Medjah, K. Drira, " Experimental Comparison of the Diagnostic Capabilities of Classification and Clustering Algorithms for the QoS Management in an Autonomic IoT Platform", *Service Oriented Computing and Applications*, vol. 13, no. 3, pp 199–219, 2019.
- [21] N. Perozo, J. Aguilar, O. Terán and H. Molina, "A Verification Method for MASOES," *IEEE Transactions on Cybernetics*, vol. 43, no. 1, pp. 64-76, 2013, doi: 10.1109/TSMCB.2012.2199106.
- [22] W. Hoyos, J. Aguilar, M. Toro, "Dengue models based on machine learning techniques: A systematic literature review", *Artificial Intelligence in Medicine*, vol. 119, 2021, doi: /10.1016/j.artmed.2021.102157.
- [23] J. Aguilar, I. Besseml, M. Cerrada, F. Hidrobo, F. Narciso, "Una Metodología para el Modelado de Sistemas de Ingeniería Orientado a Agentes", *Revista Iberoamericana de Inteligencia Artificial*, vol. 12, no. 38, pp.39-60, 2008.
- [24] M. Sánchez, J. Aguilar, J. Cordero, P. Valdiviezo-Díaz, L. Barba-Guamán, L. Chamba-Eras, "Cloud Computing in Smart Educational Environments: Application in Learning Analytics as Service". In: *New Advances in Information Systems and Technologies* (Rocha, Á., Correia, A., Adeli, H., Reis, L., Mendonça Teixeira, M., eds.), vol. 444. Springer, pp. 993–1002, 2016.
- [25] J. Aguilar, "Multilayer Cognitive Maps in the Resolution of Problems using the FCM Designer Tool", *Applied Artificial Intelligence*, vol. 30, no. 7, pp. 720-743, 2016.
- [26] G. Csardi, T. Nepusz, "The igraph software package for complex network research," *InterJournal Complex Syst.*, no. 1695, 2006.
- [27] R. Axelrod, "Structure of decision: The cognitive maps of political elites". Princeton University Press, 1976.
- [28] W. Stach, L. Kurgan, W. Pedrycz and M. Reformat, "Genetic learning of fuzzy cognitive maps", *Fuzzy Sets Syst.*, vol. 153, no. 3, pp. 371-401, 2005.
- [29] C. Murungweni, M. van Wijk, JA. Andersson, E. Smaling, and K. Giller. "Application of fuzzy cognitive mapping in livelihood vulnerability analysis". *Ecology and Society*, vol. 16, no. 4, 2011.
- [30] A. Al Farsi, D. Petrovic, and F. Doctor. "A non-iterative reasoning algorithm for fuzzy cognitive maps based on type 2 fuzzy sets". *Information Sciences*, vol. 622, pp. 319–336, 2023.
- [31] F. Aghaeipoor, M. Sabokrou, and A. Fernandez. "Fuzzy rule-based explainer systems for deep neural networks: From local explainability to global understanding". *IEEE Transactions on Fuzzy Systems*, vol. 31, no 9, pp. 3069–3080, 2023.
- [32] M. Ribeiro, S. Singh, and C. Guestrin. "Why Should I Trust You?": Explaining the Predictions of Any Classifier". In *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1135–1144. 2016.
- [33] X. Mi, B. B. Zou, F. Zou. "Permutation-based identification of important biomarkers for complex diseases via machine learning models". *NatCommun*, vol. 12, 2022.
- [34] D. Benito, C. Quintero, J. Aguilar, J. M. Ramírez and A. Fernández Anta "Explainability Analysis: An In-Depth Comparison between Fuzzy Cognitive Maps and LAMDA", *Applied Soft Computing Journal*, 2024
- [35] I. Apostolopoulos, P. Groumpos, "Fuzzy Cognitive Maps: Their Role in Explainable Artificial Intelligence. *Appl. Sci.* vol. 12, 2023, <https://doi.org/10.3390/app13063412>.
- [36] G. Nápoles, N. Ranković, Y. Salgueiro, "On the interpretability of Fuzzy Cognitive Maps", *Knowledge-Based Systems*, vol. 281, 2023.

- [37] M. Tyrovolas, X. Liang, C. Stylios, C. “Information flow-based fuzzy cognitive maps with enhanced interpretability”. *Granul. Comput.* vol. 8, pp. 2021–2038, 2023.
- [38] F. Hoitsma, A. Knobien, M. Espinosa, G. Nápoles, “Symbolic Explanation Module for Fuzzy Cognitive Map-Based Reasoning Models”. *Lecture Notes in Computer Science*, vol 12498. pp 21–34, Springer, 2020. https://doi.org/10.1007/978-3-030-63799-6_2
- [39] T. Mansouri, S. Vadera, “Explainable fault prediction using learning fuzzy cognitive maps”. *Expert Systems*, vol. 40, no. 8, pp. e13316, 2023.
- [40] Y. Wei and B. Jiang, "Interpretable Cognitive State Prediction via Temporal Fuzzy Cognitive Map," *IEEE Transactions on Learning Technologies*, vol. 17, pp. 514-526, 2024.
- [41] Departamento Administrativo Nacional de Estadística, “Colombia - Encuesta de Desarrollo e Innovación Tecnológica - EDIT - Industria 2017 - 2018”, [Online]. Available: <http://microdatos.dane.gov.co/index.php>
- [42] Scott M. Lundberg y Su-In Lee, "A Unified Approach to Interpreting Model Predictions," *Advances in Neural Information Processing Systems (NIPS)*, 2017.
- [43] Marco Tulio Ribeiro, Sameer Singh, Carlos Guestrin, "Anchors: High-Precision Model-Agnostic Explanations," *AAAI Conference on Artificial Intelligence*, 2018.
- [44] Hooker, S., Erhan, D., Kindermans, P., & Kim, B. (2018). A Benchmark for Interpretability Methods in Deep Neural Networks. *Neural Information Processing Systems*.