Smart Agricultural Technology 9 (2024) 100658

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Breaking the barriers of technology adoption: Explainable AI for requirement analysis and technology design in smart farming

Kevin Mallinger [a,b,*], Luiza Corpaci [c], Thomas Neubauer [a,b], Ildikó E. Tikász [d], Georg Goldenits [b], Thomas Banhazi [e]

[a] SBA Research, Complexity and Resilience Research Group, Floragasse 7, 1040 Vienna, Austria
[b] TU Wien, Favoritenstraße 7-9, 1040 Vienna, Austria
[c] Christian Doppler Laboratory for Assurance and Transparency in Software Protection, University of Vienna, Kolingasse 14-16, 1090 Vienna, Austria
[d] Institute of Agricultural Economics Nonprofit Kft., Zsil u. 3-5, H-1093 Budapest, Hungary
[e] National Taiwan University, Center for Biotechnology, No. 1, Sec. 4, Roosevelt Rd., Da'an Dist., Taipei City 106319, Taiwan

A R T I C L E   I N F O

Keywords:
User attitudes
Explainable Artificial Intelligence
Technology design
Technology barriers
Precision livestock farming

A B S T R A C T

Understanding the factors that drive and hinder technology adoption is critical for companies that try to access customer segments or governmental agencies that want to foster economic, ecological, or social change. By assessing the technological readiness of customer groups, common and individual barriers or opportunities for technology adoption can be observed and translated into technological requirements, business strategies, or policy interventions. Current approaches to assessing such barriers do not provide information on which factors influence technological readiness more than others, limiting the prioritization of targeted technological or political interventions. This research introduces an Explainable Machine Learning (XAI) approach to overcome this limitation. It exemplifies its usability for the Precision Livestock Farming domain, particularly for smart technologies incorporating novel advances in Artificial Intelligence and Internet of Things. A random forest machine learning model is introduced to identify clusters of different farmers' technological readiness based on the available features (survey questions). XAI techniques are then deployed to understand the influence of individual features on the prediction outcome, highlighting factors that increase or decrease technological readiness of farmers. The results are assessed for their potential for requirement and business analysis while providing targeted suggestions for technology design.

1. Introduction

Precision Livestock Farming (PLF) technologies have the potential to enhance productivity, improve animal welfare, and reduce environmental impacts of farming practices [32,41,33,25,16]. They incorporate novel technologies such as Artificial Intelligence (AI), Digital Twins, or connected sensor technologies (IoT), enabling considerable advances in the monitoring and management of livestock [3]. Hereby, AI and Machine Learning are considered to be the most critical and influential technologies in the next years and decades for PLF use-cases [50–52]. However, the successful implementation of these technologies largely depends on farmers' readiness and willingness to integrate them into their operational procedures. Therefore, finding approaches for formulating technology design and business strategies to overcome current barriers to technology adoption (attitudes, infrastructure environment, etc.) is critical for providing a sustainable impact [47]. Previous research has primarily focused on traditional statistical methods to assess these barriers [34,38,35–37,39,40], without being able to capture the complex dynamics that drive and hinder technology adoption. Traditional statistical analyses typically rely on predefined assumptions about data distribution (normality) or independence among variables [11,12]. These assumptions are often not met in real-world applications as individual attitudes, social and environmental influences, and technological attributes often influence each other in a dynamic way. High-dimensional and non-linear data further limits the ability of such methods to assess the influence of individual variables on the prediction outcome [6], which is essential for understanding barriers at the individual level. They also do not allow for scenario analysis, showing how feature changes might impact cluster assignments. These aspects are particularly critical to assessing individual characteristics' influence on the model outcome and, therefore, for creating targeted and data-driven policies and strategies.

* Corresponding author.
E-mail addresses: kmallinger@sba-research.org, kevin.mallinger@tuwien.ac.at (K. Mallinger).

https://doi.org/10.1016/j.atech.2024.100658
Received 7 October 2024; Received in revised form 12 November 2024; Accepted 15 November 2024
Available online 20 November 2024

This study aims to fill this gap by applying machine learning and Explainable Artificial Intelligence (XAI) techniques. These techniques capture complex, non-linear interactions between features that might be missed by traditional statistical methods without predefined assumptions about linearity, independence, or distribution. To exemplify this, this study analyzes different clusters of technological readiness as a proxy to investigate the associated barriers to technology adoption. These clusters will be used to train a machine learning model to predict cluster association based on survey questions while using Explainable AI techniques to understand the resulting models. Thereby, the introduced techniques capture connections between the survey questions and their influence on technological readiness, highlighting features that are primarily responsible for increasing or limiting technology adoption. The benefits of Explainable AI techniques for identifying barriers will be exemplified through a requirement and market analysis process and translated to potential technology design or business strategies (Fig. 1). In doing so, the following research questions will be assessed in this article:

- To what extent are Explainable AI methods suitable to identify barriers and opportunities influencing technology adoption?
- To what extent can a dynamic requirement analysis approach through Explainable AI support PLF developers in their technology design?
- To what extent can a dynamic market analysis approach through Explainable AI support PLF providers to increase their market access?

2. Related work

User attitudes in precision livestock farming have been studied by large through surveys and statistical analyses, with research presenting a range of factors that influence technology adoption, including economic, socio-demographic, ethical, and institutional aspects [46,4,34–40]. A recent literature review [3] of barriers to smart farming technologies highlighted hereby that the high implementation costs, resistance to new technologies, and lack of necessary infrastructure hinder widespread adoption among small-scale and developing farms. As machine learning and AI are some of the most prominent technologies for precision livestock farming, another literature review [42] summarized the constraints of such technologies for sustainable integration in farms, pointing out the importance of maintainability, reliability, and the integration of specialized knowledge.

However, the analysis of such barriers is mostly done through surveys and interviews, combined with a statistical analysis afterward. These approaches only provide a static picture of attitudes toward smart farming technologies, lacking a description of which factors actually drive technology adoption or serve as fundamental barriers compared to minor issues that hinder technology implementation. In this regard, prior research by Mallinger et al. [1] first analyzed the most important features that distinguish three clusters of technological readiness using a machine learning approach. The authors used a k-means approach and several validation methods (e.g., distance metrics, principal component analysis, focus group, and supervised machine learning) to find meaningful clusters of technological readiness. By using these clusters as labels, they showed that tree-based machine learning algorithms can be used to highlight attributes that separate the clusters well from each other. This information can be used to find attributes that are generally important to include when assessing technological readiness in surveys. However, this information only describes the overall importance of the features but cannot explain what attributes describe and influence individual cluster affiliation. Novel methods must be found to assess how individual attributes positively or negatively influence technological readiness in order to find targeted strategies for technology or policy design.

From the perspectives of technology developers and engineers, limited knowledge is available on improving requirement analysis and defining critical technological functionalities for individual market segments. Kim et al. [43] investigated the use of a KANO matrix for requirement analysis, using technological readiness as a proxy for technology adoption and categorizing the user groups between conservative and early adopters. Considering the use of novel technologies to improve the requirement engineering process, there has been some research about the utilization of natural language processing. Zhao et al. [44] provides an overview of the latest research in this field, particularly for natural language processing techniques that enable the processing of requirement documents. As mentioned above, research predominately captures a broad spectrum of attributes (e.g., economic, socio-demographic, ethical) when assessing barriers to technology adoption. Most of the studies include technological aspects (e.g., data privacy, system compatibility, usability of data) [13–15] as one of many aspects in their assessment. While some research focused particularly on economic or socio-demographic variables [17], there is a lack of studies that assess specific technological and market attributes that technology providers can directly influence, making it difficult for them to identify actionable areas for improvement.

Furthermore, to the author's knowledge, there is no published research on how to utilize Explainable AI techniques to analyze the dynamics and importance of individual features for technology adoption and the definition of technological requirements, let alone for the Precision Livestock Farming technology domain. In order to do so, the authors use the validated clusters by Mallinger et al. [1], which was described above, as a basis for applying Explainable Artificial Intelligence (XAI) techniques to assess the influence of certain attributes on cluster affiliation.

3. Materials and methods

3.1. Survey and data

This study builds on the collected survey data comprising 20 questions of the LivestockSense project[1]. 266 farms across multiple countries in the European Union (such as Sweden, Hungary, Denmark, Poland, etc.) and the Middle East (Israel) have been integrated into this study, in which 121 samples are from the pig and 145 samples from the poultry industry [4]. The questions were designed to capture information about existing infrastructure and attitudes toward smart devices/technologies used in smart farming practices. The survey design incorporates various perspectives to link responses with technological readiness and technology adoption. These perspectives include the availability of infrastructure (as addressed in question blocks 1, 2, and 6), the general presence of expert knowledge and market accessibility of PLF technologies (covered in question block 4), as well as the mindset of PLF technology users towards their capacities, which is reflected in question blocks 3 and 5. Sub-questions were combined into a single feature, with responses evaluated on either a 5-point, 4-point, or 3-point scale to reflect the degree of agreement with the statement. For instance, a rating of 1 corresponded to "Strongly disagree," while a rating of 5 indicated "Strongly agree." The full list of questions included in the analysis is provided in Table A.2. The general description of the project was published recently [45].

3.2. Clustering of technological readiness

The clustering as described in [1] was conducted with a k-means approach. The algorithm takes a set of measurements, where each observation is an n-dimensional vector, and partitions them into k clusters (where $k \leq n$, n representing the total farms contained in the survey) based on their similarity [24]. Two and three different clusters have

[1] https://livestocksense.eu/.

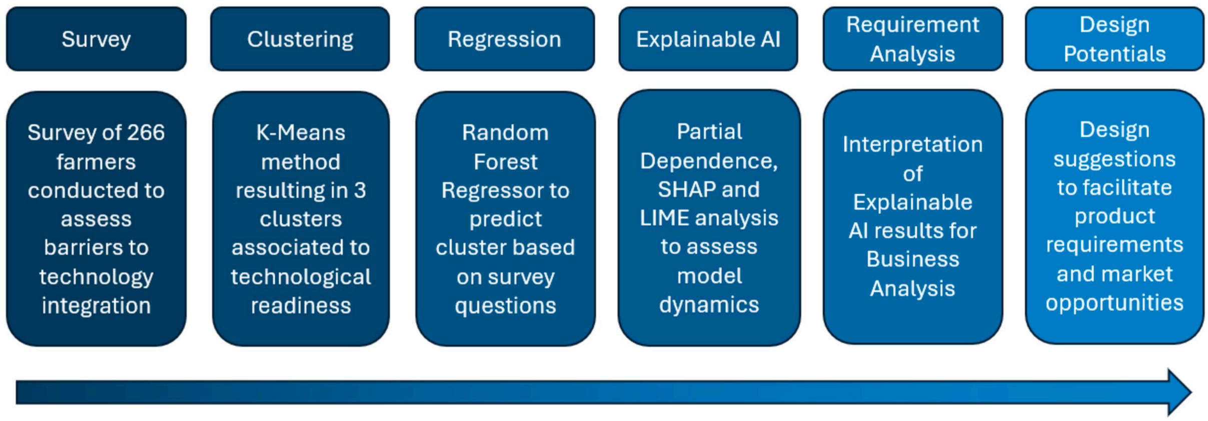


Fig. 1. Research design as described in this article. The chapters of this article are structured following this process.

been evaluated in this study and associated with technological readiness toward precision livestock farming technologies. Several steps have been taken to validate the clusters, such as distance metrics, principal component analysis, focus groups, and the prediction of clusters through a decision tree. The present article will only analyze the three cluster scenario (technologically ready, partially ready, not ready), as we are interested in identifying user requirements for farmers who are neither fully “ready” nor “not ready” to integrate such technologies, as these represent a rather volatile market segment that could be easier expanded on.

As stated in the prior study [1], the clusters highlight perspectives that represent potential barriers to technology adoption, including available infrastructure, access to information, and economic embedding. The identified clusters are:

- **Cluster 1, Not ready**: This subset includes farmers with limited on-site infrastructure availability and limited market accessibility. They tend to question the positive environmental and economic potential, display low levels of trust for smart farming technologies, lack proper education to use such technologies, and critically view their maintainability, operability, and interoperability. This cluster is also the largest group that doesn’t have any smart farming devices incorporated into their farm. It represents an untapped market segment with several barriers to technology integration that need to be overcome in order to gain access to and utilize such technologies.
- **Cluster 2, Partially Ready**: This group displays the highest diversity in infrastructure availability, presence of expert knowledge, market access, and mental attitudes towards PLF technologies. There is a tendency for answers in the mid-range, neither being overly convinced nor particularly critical of PLF technologies. However, this group also shares some aspects with the “not ready” and “ready” group. For example, the infrastructure availability and level of farm automation are similar to the “Not ready” group. However, their positive attitude towards its potential positive impact is more similar to that of the “ready” group. This diversity of answer ranges is also visible in Fig. 3. There, one can see that this group is equally distributed between people who already have such technologies, don’t have them but plan to buy them, or don’t have them. This cluster can be considered the most accessible market segment to increase one’s market share, as they often display positive attitudes towards such technologies but need targeted support for individual barriers to technology adoption.
- **Cluster 3, Ready**: This subset represents farmers with adequate on-site infrastructure availability, easy market accessibility, and people that tend to support the positive environmental and economic potential, display high levels of trust for smart farming technologies, have access to proper education to use such technologies and positively view their maintainability, operability, and interoperability. This cluster is also the biggest subgroup of people who have already bought smart technologies and the second biggest group planning to acquire any. Therefore, this cluster is the primary customer segment for vendors and developers that don’t need much convincing to buy such technologies. However, some barriers remain that can be further improved regarding technology adoption.

To further validate the cluster coherence and validity, we deploy distance metrics and dimension reduction techniques. The distance metrics are as follows:

- **Renyi’s Cross Information Potential (rCIP)**: Renyi’s Cross Information Potential is a metric derived from Renyi’s entropy [7], used to measure the separability and internal coherence of clusters in a dataset [5]. This metric is calculated by estimating the information potential of each cluster, which reflects the distribution density and compactness of points within clusters. Lower values of rCIP indicate better-defined clusters, where data points within each cluster are closer to each other and more distinct from points in other clusters. The Renyi’s Cross Information Potential (rCIP) between two clusters i and j, each represented by their mean vectors c_i and c_j and covariance matrices Σ_i and Σ_j, is defined as:

$$\mathrm{rCIP}_{i,j} = \frac{1}{\sqrt{(2\pi)^d |\Sigma_i + \Sigma_j|}} \exp\left(-\frac{1}{2}(c_i - .c_j)^T (\Sigma_i + \Sigma_j)^{-1} (c_i - c_j)\right)$$

where:
– c_i and c_j are the mean vectors (centers) of clusters i and j, respectively,
– Σ_i and Σ_j are the covariance matrices of clusters i and j,
– d is the dimensionality of the data.
The overall rCIP criterion value for a set of clusters is then calculated as the sum of pairwise rCIP values between all clusters:

$$\mathrm{rCIP} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathrm{rCIP}_{i,j}$$

where n is the total number of clusters. This formulation captures the degree of overlap between clusters, with lower values indicating better separation and compactness. A lower rCIP score signifies more cohesive clusters with minimal overlap, suggesting that the clusters effectively represent distinct groups within the data.
- **WB Index**: The WB Index is a cluster validation metric that combines within-cluster compactness (W) and between-cluster separation (B) to assess the clustering structure [8]. It is computed as the ratio of the sum of within-cluster distances to the sum of between-cluster distances. For k clusters, the WB Index is defined as:

$$\text{WB Index} = \frac{\sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|}{\sum_{i=1}^{k} \sum_{j=i+1}^{k} \|\mu_i - \mu_j\|}$$

where C_i is the i-th cluster, μ_i is the centroid of C_i, x represents a data point in C_i, and $\|\cdot\|$ denotes the Euclidean distance. Lower values of the WB Index indicate a better balance of compact and well-separated clusters.

For visualizing the clustering results, we employ Uniform Manifold Approximation and Projection (UMAP) [9], a nonlinear dimensionality reduction technique that projects high-dimensional data into a lower-dimensional space while preserving the local and global structure of the data. UMAP is particularly useful for understanding the spatial distribution and separability of clusters in a more interpretable two- or three-dimensional space.

UMAP operates based on two main principles: constructing a weighted graph that captures the local relationships between data points in the high-dimensional space, and then optimizing a low-dimensional layout to maintain those relationships. Given a high-dimensional dataset $\{x_i\}_{i=1}^N$ and a lower-dimensional projection $\{y_i\}_{i=1}^N$, UMAP aims to minimize the following cross-entropy objective function:

$$\mathcal{L}_{\mathrm{UMAP}} = \sum_{i\neq j} \left[w_{ij} \log\left(\frac{w_{ij}}{d_{ij}}\right) + (1-w_{ij}) \log\left(\frac{1-w_{ij}}{1-d_{ij}}\right)\right]$$

where:

- w_{ij} represents the probability of a connection between points x_i and x_j in the high-dimensional space, capturing the local similarity based on their distance.
- d_{ij} denotes the corresponding probability of connection in the lower-dimensional space, based on the Euclidean distance between their projections y_i and y_j.

This objective function is optimized to ensure that points that are close in the high-dimensional space remain close in the lower-dimensional space, while distant points are also kept apart. The resulting UMAP projections allow us to visually assess the cluster distribution and separability in a more interpretable form.

3.3. Modeling

This research used a combination of machine learning models and explainability techniques to assess user readiness for technology adoption. First, we utilized a series of unsupervised and supervised machine learning models to predict user readiness levels based on the clusters established by [1]. In the next step, the dynamic of the regression model was further analyzed using Explainable AI techniques to determine the contribution of each feature to the predictions.

Supervised Machine Learning: For the machine learning prediction, a Random Forest Regressor was used. This ensemble learning technique constructs multiple decision trees during training and returns the mean prediction of the individual trees to improve predictive accuracy and reduce overfitting. The scikit-learn library is chosen for the implementation of this method.

The classification performance of the regression model was measured on a 75%/25% train-test split for a rounded version of the predicted values. The predicted values were rounded by splitting the range [0, 2] into three equal intervals corresponding to the classes. The ranges of the classes are set to 0-0.66 for the "not ready" group, 0.67-1.33 for the "partially ready" group, and 1.34-2 for the "ready" group. The model was evaluated using precision, recall, and F1-score, as well as the macro averages of these metrics and accuracy to assess overall performance across all classes.

Precision: Precision per class shows how many of the predicted positive cases for a specific class are actually correct. It is calculated as:

$$\text{Precision}_i = \frac{\text{True Positives}_i}{\text{True Positives}_i + \text{False Positives}_i} \tag{1}$$

Recall (Sensitivity): Recall per class indicates how many of the actual positive cases for a specific class were correctly predicted. It is calculated as:

$$\text{Recall}_i = \frac{\text{True Positives}_i}{\text{True Positives}_i + \text{False Negatives}_i} \tag{2}$$

F1-Score: The F1-score per class is the harmonic mean of precision and recall, balancing the two metrics. This considers class imbalances:

$$\text{F1-Score}_i = 2 \times \frac{\text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i} \tag{3}$$

Overall Accuracy: The overall accuracy is calculated as the number of correct predictions (sum of true positives and true negatives) divided by the total number of predictions. It is given as:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}} \tag{4}$$

Explainable Artificial Intelligence: After this modeling process, several tools in the area of Explainable Artificial Intelligence (XAI) are deployed to analyze the dynamic of the model:

- **Partial Dependence Plots (PDP)**: show the effect of a single feature on the predicted outcome of a model, averaging out the effects of all other features [48]. The average influence of a farmer's characteristics can give valuable insights into how individual features influence the whole study population. Such an analysis is particularly interesting if one wants to find requirements that are important for all subgroups.
- **Individual Conditional Expectation (ICE) Plots**: are an extension of PDPs that display the relationship between the feature and the prediction for individual instances, highlighting the variability of the prediction across the feature's values [49]. Therefore, ICE plots are particularly valuable for analyzing farmers' key characteristics that influence cluster association. By defining different clusters of readiness, one can observe which characteristics might be key enablers that increase the overall readiness level of individuals as well as cluster subgroups.
- **SHapley Additive exPlanations (SHAP)**: explain the prediction of an instance by computing the contribution of each feature to the prediction [21]. These values can be viewed per individual prediction and can give detailed insights into individual subgroups' needs or barriers. However, due to the global settings of SHAP, one can also aggregate the results into subgroups, enabling accumulated analysis of important features influencing readiness for technology integration.
- **Local Interpretable Model-agnostic Explanations (LIME)**: explain individual predictions by approximating the local decision boundary of any classifier with an interpretable model [22]. It helps in understanding why a model made a specific prediction, similar to Shapley values. By sampling individual LIME values within a subgroup, one can detect common patterns that support or hinder technology integration.

3.4. Requirement analysis and market research

Requirement analysis is a systematic process to identify and document the essential requirements of a technology, system, or project. It involves analyzing and validating the needs and constraints of various stakeholders to ensure the final product meets its intended purpose [2]. This process is part of business analysis, serving as the foundation for designing, developing, and implementing effective technologies and solutions [18]. The requirement analysis process comprises several subparts, including stakeholder analysis, requirements elicitation, requirements specification, and requirements validation [10]. Each of these steps plays a crucial role in understanding the business potential, capturing detailed requirements, and ensuring that the final solution aligns with consumer expectations [20]. Closely related to this and also part of the business analysis process is the market research. This process incorporates research to identify potential market segments by understanding the market conditions, competitive landscape, and customer needs.

Understanding barriers to technology adoption for subgroups or even individuals can help producers expand their consumer base by leveraging the readiness of farmers to integrate and use smart technologies. Because of this, the survey and clusters were designed to capture information about existing infrastructure and attitudes toward smart devices/technologies in farming practices that the technology providers directly or indirectly influence. By identifying user qualities associated with different readiness levels and features responsible for increasing or limiting technology adoption, targeted interventions can be taken to improve requirement analysis and, ultimately, product designs and market strategies.

To enhance processes of requirement analysis and market research, we propose the utilization of Explainable Artificial Intelligence (XAI) techniques, including SHAP (SHapley Additive exPlanations) [21], LIME (Local Interpretable Model-agnostic Explanations) [22], and Partial Dependence Plots (PDP) [23]. These techniques offer insights into how individual factors, extracted from survey questions, influence the overall technological readiness of users as well as their barriers to technology adoption. By applying these XAI methods (see further information in Section 3.3), we create a dynamic toolset for requirement analysis and market research that reveals the underlying drivers and barriers of technological readiness for specific subgroups and highlights key factors that need to be addressed to enhance market adoption, product acceptance, and technology integration.

4. Results

4.1. Clustering and predictive modeling

Next to the validation from [1], we extend the analysis of the cluster validity by cluster distance metrics and the visual validity check executed through UMAP (see description in 3.2). As can be seen in Fig. 4, the two and three-cluster solutions are the strongest candidates for the given data set. Fig. 3a shows the Renyi index in which lower values indicate better-defined clusters with higher internal similarity and separation from other clusters. The rCIP is minimized for the two-cluster solution, while the three-cluster configuration is the second best with marginally higher values. On the other hand, the WB Index reaches a minimum at three clusters (lower is better), indicating optimal separation and cohesion for this clustering configuration. This suggests that the three-cluster model best balances the trade-off between intra-cluster similarity and inter-cluster distinction, further validating the conceptual categorization into three readiness levels. Based on the two distance metrics, we can identify the three-cluster solution as a suitable candidate for further analysis. As similarly argued in [1], this is done to assess the attributes of different readiness attributes in more detail, with a particular focus on people that are neither fully ready or not ready. This increases the possibility for more targeted interventions.

Next, we evaluate the three-cluster solution with a UMAP approach, which transforms the multidimensional vector space (every question represents a vector/dimension) into a two- and three-dimensional object. This allows us to assess the cluster distribution visually. Fig. 2 illustrates the three clusters of technological readiness, with the partially ready category positioned between the ready and not ready categories, suggesting a chronological progression. The visualization clearly shows the distinct separation between clusters, with a larger grouping in the middle of the plot and two smaller groups at the top and bottom. Further visualizations of this clustering can be found in the Annex (Figs. C.10, C.11).

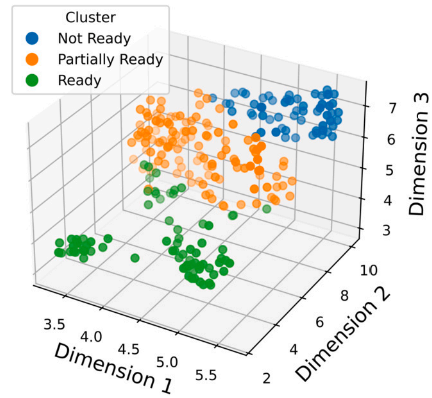


Fig. 2. Three-dimensional UMAP projection illustrating clusters of technological readiness among participants. The clusters are color-coded to indicate different readiness levels: “Not Ready” (blue), “Partially Ready” (orange), and “Ready” (green). Each point represents an individual respondent, and the spatial arrangement reflects the similarity between responses in a high-dimensional feature space, reduced to three dimensions for visualization. The UMAP projection uses Euclidean distance, with 20 neighbors and a minimum distance of 0.1.

To further validate the cluster validity and the respective interpretations resulting from them, we compare the three clusters with the distribution of a subquestion within the survey (“Do you use smart devices at the farm you represent?”). The idea behind this comparison is that clusters of technology-ready people should have a higher tendency to have or plan to integrate smart devices instead of people that have a less favorable attitude/environment to integrate them. On the other hand, people who display challenges of technology adoption should be particularly present in the segment of answers that don’t buy such a technology. However, it is possible that people with limited environmental functionality and critical perspectives towards smart technologies still own or intend to buy smart technologies and vice versa (e.g., technology was bought by someone else, shared, or inherited).

As displayed in Fig. 3, it can be seen that the majority group of people that implemented smart technologies already are in the cluster segment “ready” and “partially ready”. On the contrary, people that have answered that they don’t have any technology are mostly in the category “not ready”. People who answered that they don’t have any smart PLF technologies yet but plan to buy some show a relative equal distribution of attitudes and predispositions towards technologies. This mapping of answers and clusters further strengthens cluster validity, which is crucial for the dynamic analysis of user attitudes and their use for requirement analysis in Section 3.4 and 4. Further statistical information about the distributions of answers can be found in Table B.3.

The random forest model in this experiment predicted the three different classes of clusters based on the survey question results, obtaining an accuracy and recall of 81 percent. The classifier’s precision averaged at 84 percent, indicating that the model predicts outcomes with high consistency or little variability (see Table 1. In comparison, the baseline prediction, if we would choose one class randomly, is 33 percent accuracy. These results indicate that the random forest model provides a stable basis for analyzing the relationship between individual farmer attitudes in the survey and the associated clusters. For this analysis, we deploy three different Explainable AI techniques, as described in Section 3.3 and structure the following subchapters accordingly based on these techniques.

4.2. Explainable AI analysis

After validating the clusters and the machine learning model, we analyze the behavior of the deployed random forest model. For this, several Explainable AI methods (ICE, PDP, SHAP, and LIME) are integrated to highlight the influence of individual features (survey questions) on the prediction outcome (technological readiness) of the model.

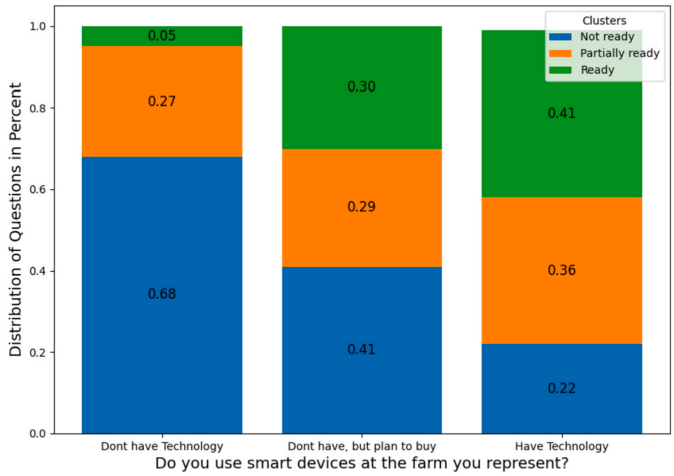


Fig. 3. Clusters within each question category. The values represent the normalized distribution of answers based on cluster size.

Table 1
Performance metrics for the random forest model predicting technological readiness: per-class and macro average.

Class	Precision	Recall	F1-Score	Support
Not Ready	0.93	0.68	0.79	19
Partially Ready	0.77	0.82	0.79	28
Ready	0.83	0.95	0.88	20
Macro Average	**0.84**	**0.82**	**0.82**	**67**

4.2.1. *ICE and PDP - simulating the influence of features on the prediction outcome*

Fig. 5 summarizes the Individual Conditional Expectations (ICE) and Partial Dependence Plots (PDP) for all questions individually in one plot. The ICE simulates the impact on the model's prediction if the individual question is answered differently per farmer while the other attributes stay constant. PDP, on the other side, assesses the average impact of changing one variable and keeping the other features constant over all samples (farmers). Therefore, each blue line represents one farmer whose readiness level increases or decreases if that attitude or environmental factor is changed. Based on this, one can analyze factors that significantly impact the model output or separate the clusters well from each other. Interpreting these changes helps to identify which factors are associated with different levels of technological readiness (clusters). The plots also indicate which factors serve as fundamental attributes in increasing the technological readiness of farmers to the next cluster group (such as that access to information is an important preliminary to understanding the benefits of such technologies). Thereby, the ICE and PDP plots enable us to directly isolate complex dynamics between the feature and their effect on individual readiness as well as per clusters or overall user base.

In this context, Fig. 5 indicates several different questions that influence the association to different levels of technological readiness. Chronologically, the first question that visibly influences the cluster association is within question block 3, which assesses the attitude towards precision livestock farming technologies. Here, for question 3a (Tech helps labor shortage), a slow increase is observable in the "not ready" cluster at the bottom of that subplot between 3 and 5. Still, an even sharper increase of technological readiness can be seen between the "partially ready" and "ready" clusters. One can observe a continuous increase from the middle section until the top of the subplot, starting from 2 until 5. This indicates that the ability to automate farms is an important factor for people who intend to acquire precision livestock farming equipment. As the groups with a prediction of 1 and higher (y-axis) are particularly associated with groups of people that consider buying PLF technologies, this could be a critical focus to expand one's product lines and market shares.

Similar to question 3a, question 3d (Tech helps to meet environmental pollution reduction obligations) displays a sharp increase between 2 and 3 and between 4 and 5. People considered more technologically ready agree that such technologies are important for managing environmental dynamics on their farms (e.g., CO2 or NH4 monitoring). For some farmers that have been considered on the top end (y-axis) of the "partially ready" cluster (between 0.7 and 1.3), this attitude is considered to be an important factor in trusting the technology and potentially integrating them into their farm, but less so for the "not ready" group.

One can also observe an increase in the prediction outcome for question 3e (Technology enables the increase production effectiveness), particularly between 2 and 4. Although most farmers stay within their respective cluster boundaries, we can observe four blue lines that jump between the groups.

Within the attitude block, the last question that displays an effect on the model is question 3h (Technology provides information in a real-time manner). This question has a lesser effect than the first two questions in this block but still has some influence on individual farmers. For three farmers in this survey, an increase in this attitude would have resulted in being classified in the next higher cluster. Further analysis of these samples with statistical and Explainable AI techniques will perhaps show further important attributes as to why these farmers consider this functionality important or if other factors have a higher effect on their association with the individual clusters.

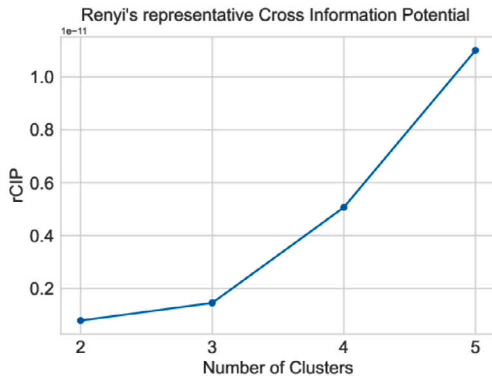
Question block 4 displays the overall presence of expert knowledge and market access to PLF technologies. Within this question category, all questions show some visible influence on the potential to integrate PLF technologies. Accessibility of such technologies (question 4a) begins to show its effects for the "not ready" group between 2 and 3, whereas, for the "partially ready" group, this question is relevant if its answer is higher than 4. This shows that for people with generally lower attitudes toward PLF technologies, accessibility is a limiting factor in developing positive attitudes toward these technologies. This trend is also visible in question 4b (Tech can be purchased at an affordable price). Here, we can see that the "not ready" cluster, in particular, perceives the price to be a limiting factor in adopting the technology (barrier between 3 and 4 on the x-axis).

The most significant factor in this question block is question 4c (It is easy to get information on technology and distributors). Visible changes in the evaluation of readiness for the "not ready" group (below 0.66) are observable between the 3 and 4 on the x-axis. This was also a critical threshold for the "partially ready" group to be classified as "ready". Thereby, this factor is an important consideration for all clusters alike.

Question 4d (It is easy to get technical assistance to smart technologies) is less pronounced than question 4c. Still, it shows a steady increase in importance for the "not ready" and "ready" groups. Many of the farms considered as less inclined to incorporate such technologies can jump to the next highest group, "partially ready", if they reach an answer that is equal to or higher than 4.

The amount of proper education available to use smart technologies (question 4e) is a critical factor that affects particularly the "partially ready" group. Although we only see a few sharp increases in the prediction outcomes if the value is increased, there is a steady increase in most farmers visible in the middle of the subplot and in the lower parts (not ready cluster). This incline is particularly visible between 3 and 5 on the x-axis and indicates that it might be a baseline factor to influence technology adoption.

Question block 5 summarizes the perceived operational functionalities of potential precision livestock farming technologies. Hereby, Fig. 5 shows that the perceived ease of operation is a critical factor that influences technological readiness (question 5b). This is particularly visible between 2 and 3 for all farmers that have been categorized as "not ready". For this group, which displays the highest distrust against smart technologies, functional accessibility is a critical factor that limits their



(a) Renyi's representative Cross Information Potential (lower is better)

(b) WB Index (lower is better)

Fig. 4. Cluster validation metrics for clusters between two and five.

ability to utilize it. However, some sharp increases are visible in the "partially ready" cluster.

Question 5c (Technology can be connected well with other equipment) shows a sharp increase in technological readiness between 2 and 3 for the "not ready" and "partially ready" clusters. This is even more pronounced than ease of operation and highlights that interoperability is a concern for people who are considered skeptics. However, we can also see that interoperability was the limiting factor for some people in the "partially ready" group and that an increase of perceived ability changed their prediction to "ready".

Lastly, we can see in question 5d that reliability (Tech operates in a reliable manner) is a critical limiting factor for the "not ready" cluster. Small increases are visible for all three clusters. However, we can observe several transitions from one group to another in the lower categories.

4.2.2. LIME - localized changes of predictions

Lime analysis provides new predictions based on altered features that are in the vicinity of the original feature values. It then creates an interpretable surrogate model (linear model) on this local feature space. Exploring the local neighborhood of samples and its effect on the surrogate model can be used to understand the behavior of certain clusters better without the influence of values that are typically associated with other clusters. By using a linear model, the interpretation is based on the linearity of features and given by the positive or negative coefficients of the respective surrogate model. Each feature contributes independently, and the overall prediction is the sum of these contributions. The changes in the features are then binned to assess their effect on the output. If a certain bin produces statistically significant changes, this bin is considered a threshold and will be used as an additional feature in the linear model. By calculating thresholds of individual features, one can visualize the non-linear behavior of the attribute on the model prediction. This shows what thresholds for individual questions must be reached in our case to affect users' readiness levels significantly. Figs. 6, 7, and 8 display the aggregated lime explanations for the "not ready", "partially ready" and "ready" groups. Aggregation was done by averaging the impact of the questions for all samples within the respective subgroups. This enables the interpretation of the overall barriers and enablers of technology integration.

For the "not ready" cluster in Fig. 6, one can see that the factor that led to significant decreases in the prediction outcome is question 4c (it is easy to get information on tech and distributors). Here, this question shows the biggest positive impact on the group's prediction. Other critical factors for this group are the accessibility to technology on the market (question 4a), availability of technical assistance (question 4d), the attitude of farmers towards the potential of PLF technology to help with labor shortage (question 3a), and ease of operability (question 5b). Questions 3a and 4a are also primarily answered below 3 and 4 in this cluster and show a significant positive influence if answered above this threshold. This indicates that the predictions of the cluster's samples can be significantly enhanced by increasing this value.

Fig. 7 displays the average changes of predictions of the surrogate model for the cluster "partially ready". Similar to the "not ready" group, we can also see that operability and the potential to subsidize labor shortages are usually answered rather low for this group with equal thresholds around 3 and 4. In general, we have several questions that are overall associated with lower prediction outcomes in this cluster, such as accessibility in question 4a, available technical assistance (question 4d), access to information on tech (question 4c), or available education in question 4e.

Analyzing the local behavior of the "ready" group in Fig. 8 shows that the availability of information (question 4c) influences the prediction outcome significantly if it is answered above 3. The second most positive influential factor is the accessibility of the technology if it's higher than 4, and the third most relevant answer is the assumption about the support of smart technologies for the labor shortage if it's answered above 3. Other factors that significantly enhance the prediction outcome locally are higher answers for the availability of education (question 4e), the ease of operability of smart technologies, and the available technical assistance. Fig. 8 also shows that the interoperability of smart technologies can be a limiting factor for technological readiness (question 5c).

4.2.3. SHAP - calculating prediction shift relative to the mean

The SHAP analysis, as presented in Fig. 9, is calculated by considering all possible combinations of features and measuring the change in the model's prediction when each feature is added to these combinations. If visualized per sample, the SHAP values indicate the influence of a feature compared to an average prediction. Fig. 9 aggregates the individual SHAP values per cluster and displays the average influence of the features per group. The questions are sorted based on the total influence on all clusters. The vertical bars represent how the individual features on the x-axis either increase or decrease the average prediction outcome of the model based on the given answers of the groups. As the "not-ready" group (red) is always predicted below the average and the "ready" group (blue) is always above the average, this analysis enables us to investigate the primary barriers for the "not-ready" group to be considered "partially ready", as well as what attributes distinguish the "ready group" from the "partially ready" cluster. As the "partially ready" group (yellow) is clustered around the center, we can identify the nuanced changes of this group to lean toward the less or more ready group.

As shown in red in Fig. 9, the ease of getting information on the technologies and distributors (question 4c) was, on average, the most significant concern that led to a decrease in prediction for the "not

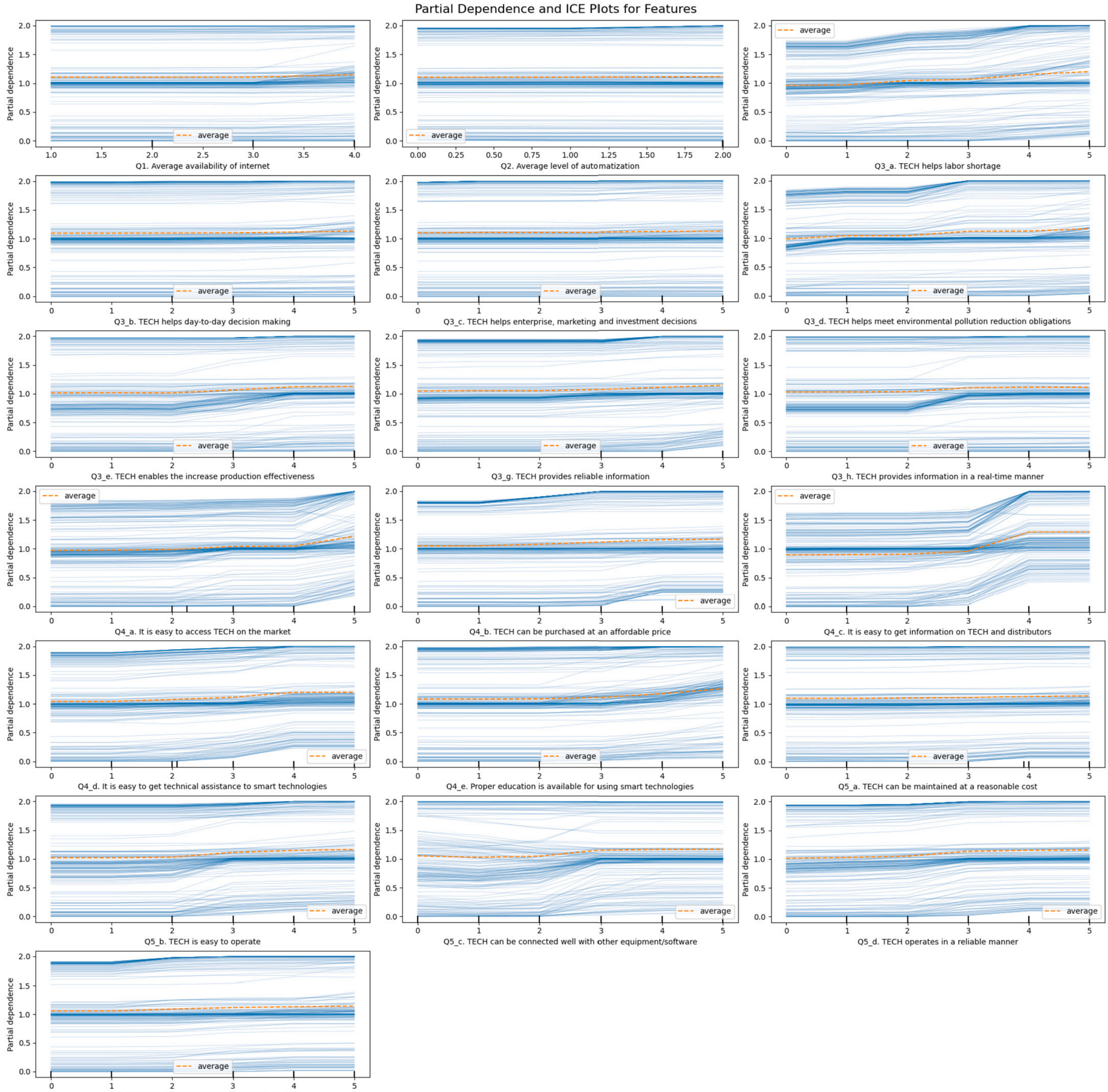


Fig. 5. Individual Conditional Expectation (blue lines) and Partial Dependence Plot (orange dashed line) based on Random Forest Model. Each blue line represents one sample (farmer) and the change of the prediction outcome if the respective question was answered higher or lower.

ready" cluster. This is followed by question 4a (It is easy to access technology on the market), question 5c (Technology can be connected well with other equipment and software), question 3a (Technology helps labor shortage, and question 4a (It is easy to access technology on the market), and question 4d (It is easy to get technical assistance to smart technologies). Three out of the five most influential questions for this group are in the category that describes expert knowledge and market access. However, we can see that many barriers simultaneously limit technology adoption and integration.

Considering the "partially ready" cluster, the assumed potential of smart technologies to support labor shortage (question 3a) was, on average, the most limiting factor. After that, the most limiting factors are, on average, the accessibility to the market (question 4a) and the available education for using smart technologies (question 4e). Contrary to the factors that reduced the readiness for smart technologies compared to the average prediction, this group displayed some characteristics that increased their readiness. Hereby, the most prominent factor was the reliability of smart technologies (question 5d) and the interoperability (question 5c).

To increase the model prediction from "partially ready" to the "ready" group, the most important factors are, on average, the ease of getting information on the technologies and distributors (question 4c), the ease of accessing the technology on the market (question 4a), availability of education (question 4a), and the possibility to help labor

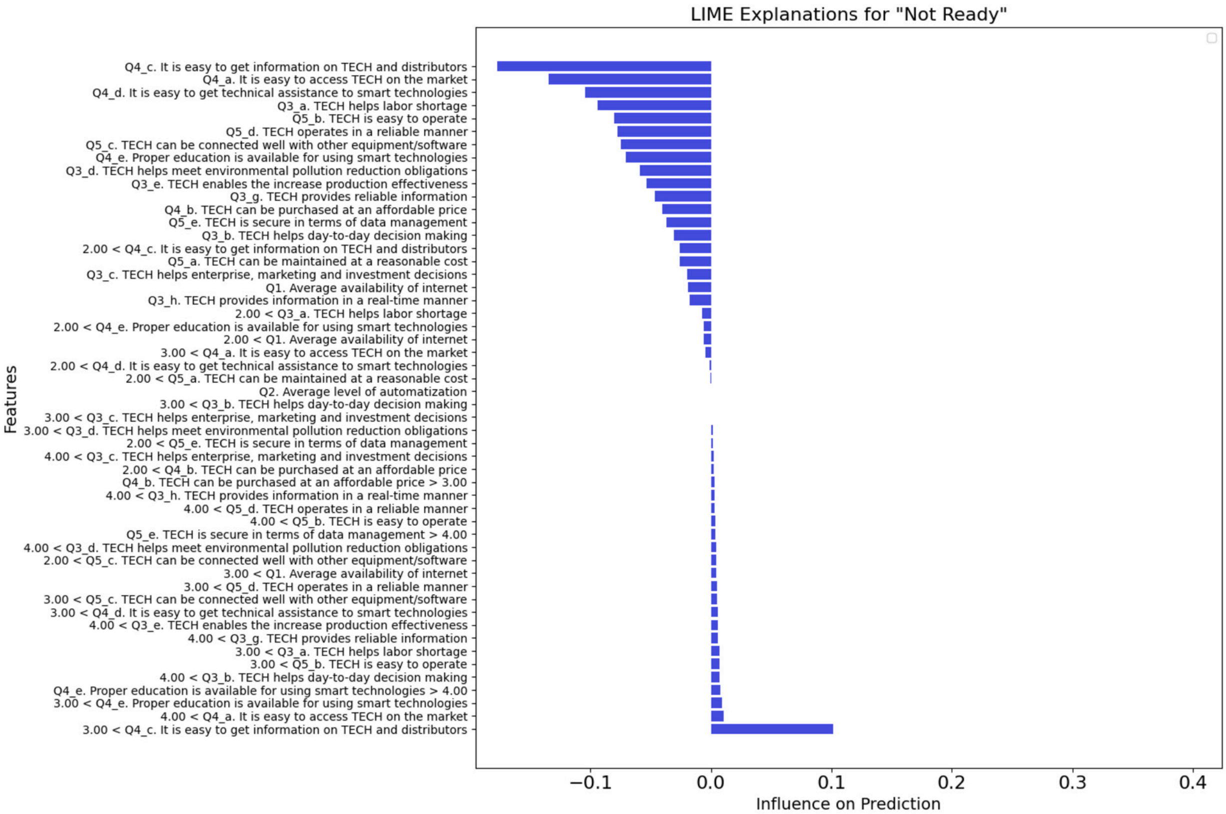


Fig. 6. Results of the LIME analysis for the "not ready" group. The horizontal bars indicate the influence of individual survey questions on the local behavior of the model. Thresholds of questions that provide a significant change on the model output are listed as well.

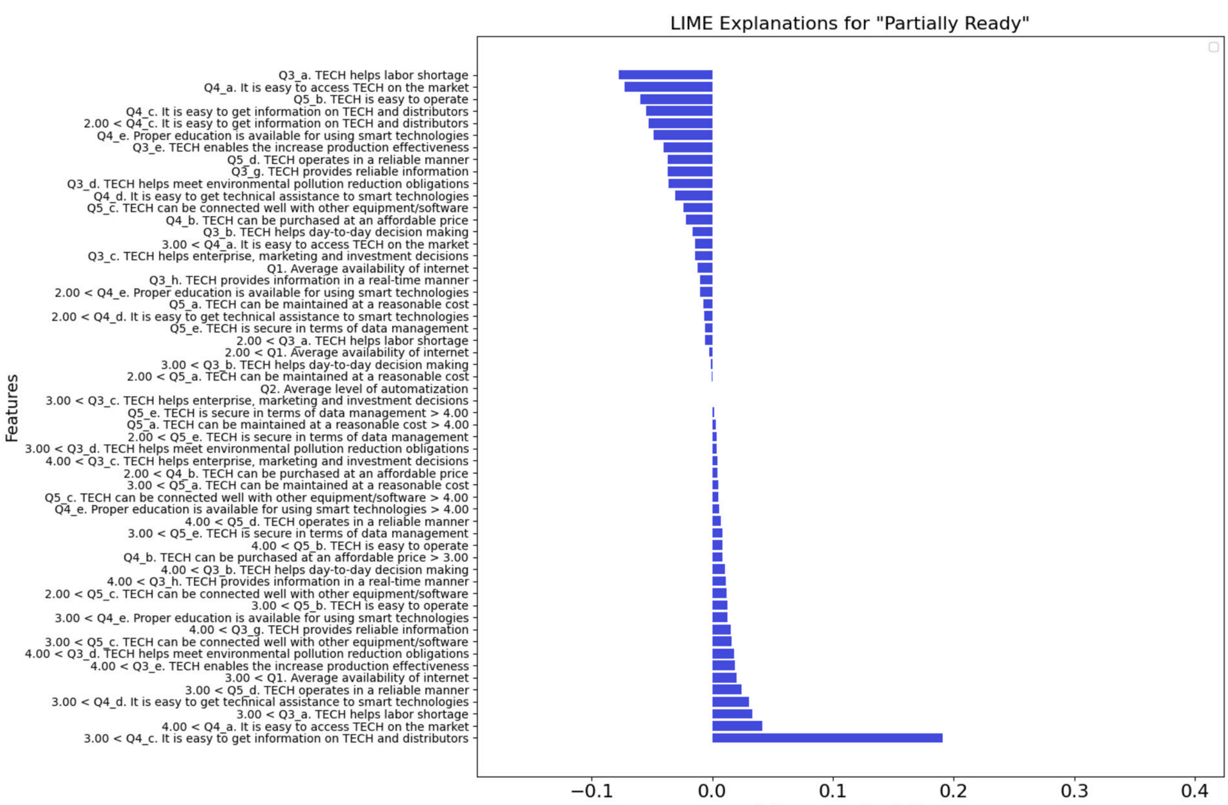


Fig. 7. Results of the LIME analysis for the "partially ready" group.

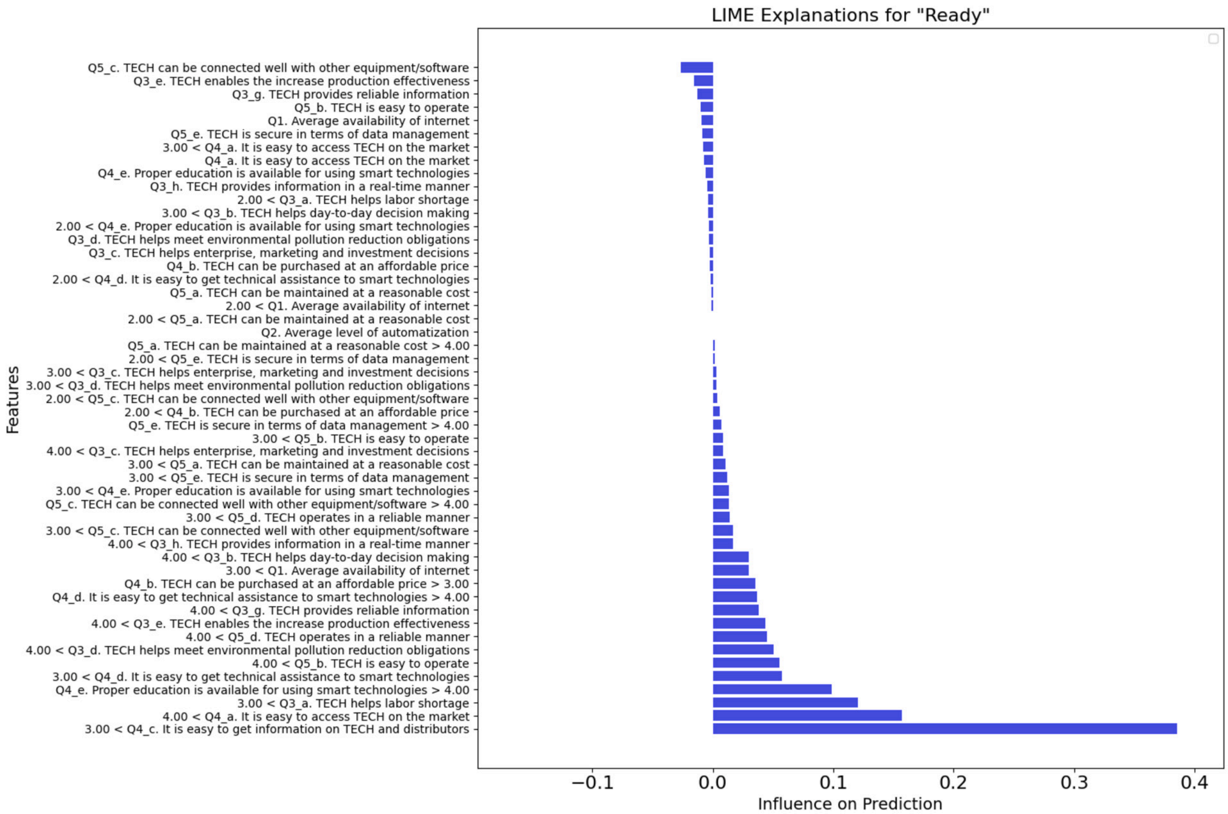


Fig. 8. Results of the LIME analysis for the “ready” group.

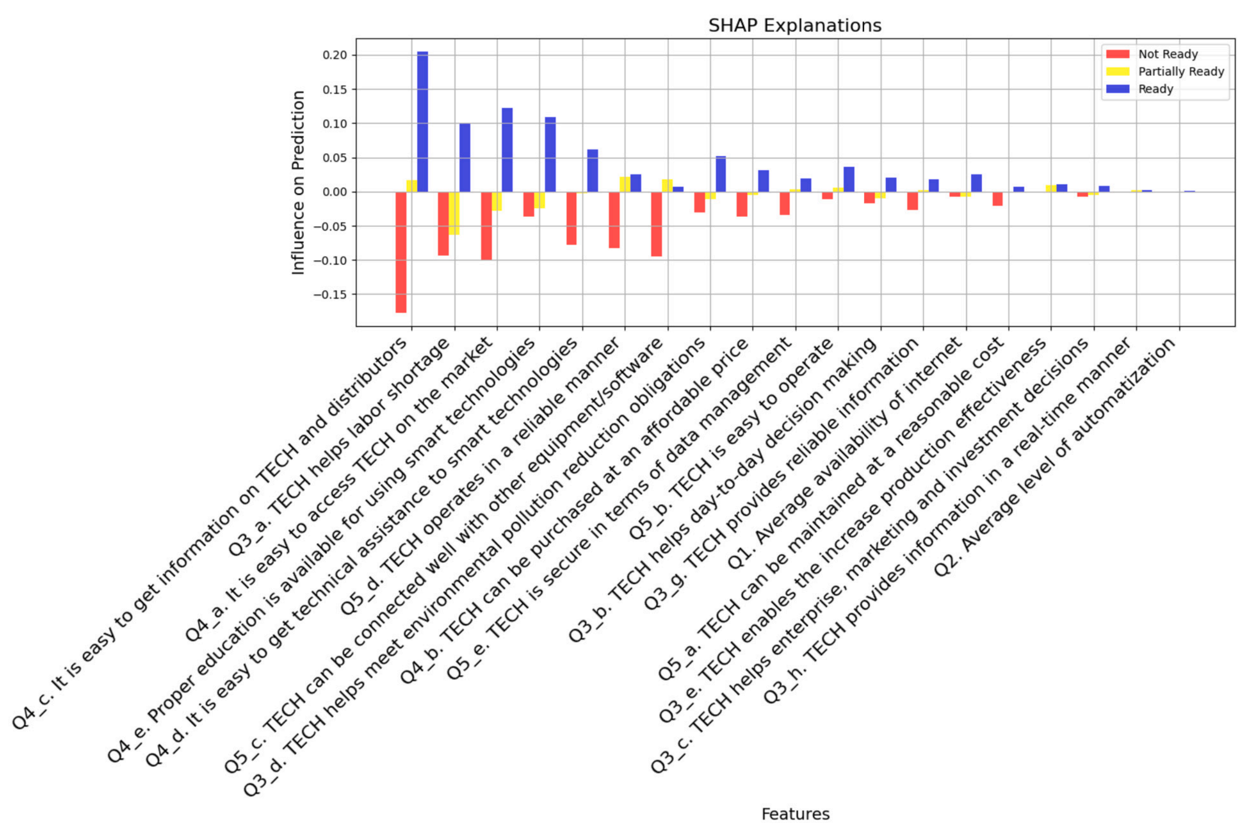


Fig. 9. Aggregated plot of the SHAP analysis by averaging the SHAp values of the individual clusters samples. The red vertical bars mark the “not ready group”, yellow the “partially ready” group, and blue the “ready” group. Values are listed from left to right according to their summed influence on all groups.

shortage (question 3a). Other relevant factors include the availability of technical assistance (question 4d), the ability to meet environmental obligations (question 3d), and the ease of operability (question 5b).

Noteworthy, the assumed potential of smart technologies to support labor shortage is an important factor that influenced the prediction of all three clusters. To a lesser degree, accessibility to the market, the potential to meet environmental obligations, and the available education are also factors that influence all subgroups.

4.3. Implications for technology design and business strategies

4.3.1. Implications considering all clusters

Analyzing all three XAI methods shows that the most influential barrier to technological adoption is the availability of information on smart technologies and distributors (question 4c). This question showed the highest values for the SHAP plot in Fig. 9, the highest values for all three LIME plots, and showed significant fluctuations in the ICE and PDP plots. The second most influential marker was the general accessibility of the market (question 4a), showing the third-highest aggregated values in the SHAP plot and the second-highest positive factor in all three LIME plots while displaying visible changes in the ICE analysis. These two factors can be considered baseline criteria necessary to acquire smart technologies or investigate their potential use. Low values in these categories inhibit other factors that are needed to finally adopt smart technologies, such as a positive attitude or the right farm infrastructure. By analyzing the geographical background of the individual answers on the ICE plot (e.g., by displaying the country per color), one can identify which countries would benefit the most if market and information accessibility were enhanced in these regions. This information provides targeted advice on which countries could lead to a sharp increase in sales and, therefore, technological integration if targeted correctly with marketing strategies and an expansion of the sales area. However, this was not done in this study to anonymize the individual survey participants.

Next to this baseline barrier of market accessibility, another factor that significantly influences technological readiness and adoption is the accessibility to technological assistance for smart technologies (question 4d). It showed high importance in the SHAP plot for the “ready” group and a particularly high influence in the LIME plot for the “not ready” and “partially ready” groups. It also showed high fluctuations, particularly for the “not ready” group and also, to some extent, for the “partially ready” group in the ICE plot. As these groups have a high share of farmers who don’t possess smart technologies yet and are considering buying one in the future, the results indicate that increasing the available support for smart technologies is a crucial factor that should be considered to expand the user base. This can be done by human support or by designing the technology to adapt to the user’s proficiency. Smart interfaces can be a viable way to increase understandability and recognize if a user gets stuck in certain functionalities. This analysis can be further expanded by analyzing other factors in this user group, such as age, location, or other limiting factors that are crucial for identified individuals who show higher readiness if access to support is increased. This could lead to a targeted change in technology design for specific sales regions or market segments. The mentioned suggestions are also relevant for question 4e (Proper education is available), which is the fourth most influential value in the SHAP analysis and the “ready” group in the LIME analysis.

Another factor crucial for enhancing market access and facilitating technology adoption is the ability of smart technologies to subsidize labor shortages. It was the second-highest value in the SHAP analysis, the third-highest positive influence in the “ready” LIME plot, and the highest barrier in the “partially ready” group. Many farmers face challenges finding qualified and persistent personnel to support their farming operations. The ability to automize farming processes is therefore a vital objective to buy smart technologies, particularly in the rapid progress of AI applications. However, in order to do so successfully, smart technologies should focus on a user-centered design. The technology should be intuitive, requiring minimal training for farm operators. This includes user-friendly interfaces, clear instructions, and adaptability to individual levels of expertise. This was also visible in the LIME plots for the question about ease of operation (question 5b) for the “not ready” and “partially ready” clusters. Hereby, the ease of operation was highlighted as one of the strongest barriers to technology adoption, logically supporting the requirements for successful automation. Another supporting factor for automation is the interoperability (question 5c) to existing technologies, such as feeding systems, climate control, and animal health monitoring platforms. This is particularly important for IoT (Internet of Things) environments and AI applications that enable centralized farm management, thereby enhancing automatization capabilities. The ability of smart technologies to be interoperable was particularly important for the “not ready” group, as can be seen in Fig. 6 and Fig. 9.

4.3.2. Targeted intervention considering individual clusters

Requirement analysis often involves understanding the specific needs of different customer segments. This detailed understanding enables more accurate market segmentation during market analysis. By knowing the distinct requirements of various customer groups (clusters), companies can tailor their market strategies and product designs to target better and serve each segment. In this chapter, we distinguish between three different groups of technological readiness and analyze the distinct behavior and attitudes towards smart PLF technologies as well as their intention to acquire said technologies.

In order to acquire information about the importance of individual questions for technological readiness and, consequently, technology design and market analysis, one can a) identify the factors or questions that differentiate a given group from other subgroups, thereby uncovering the critical elements that could help transition less ready groups towards greater adoption of smart technologies, and b) highlight the specific design features or functionalities that are most valued by a given group, thereby generating empirical insights for optimizing technology solutions to align with the requirements and expectations of this segment.

In the case of the **“ready” group**, Fig. 3 shows that this cluster is comprised of farmers who already have the technology or intend to buy it shortly. This subgroup displays positive attitudes towards the benefits of smart technologies and doesn’t need convincing to acquire such technologies. However, by focusing on the ready cluster, companies can gain insights into the specific characteristics and preferences that drive early adoption. This analysis can inform the development of targeted marketing strategies that resonate with potential customers who already have smart technologies or are on the verge of adopting them.

As no comparison with a higher group than ready can be made, the “ready” category must be analyzed based on its own behavior. For example, most farmers in the “ready” group still consider the interoperability of the smart devices as not ideal (see the negative influence in Fig. 8). The negative value in the ready lime plot indicates that the range of answers for this feature has on average a negative effect on the prediction outcome. However, if developers can increase the trust in interoperability above 3, it becomes a positive factor for that group. This shows that for people who either have such technologies yet or are planning to purchase one for the first time 3, higher standards of interoperability could further increase their chances of implementation.

Although the most influential questions for the “ready” group are the accessibility (question 4a) and available information (question 4c) of technologies, this value is not a barrier as the average answers for these questions surpass the given threshold (for accessibility, the average answer of the group is 4.8, and for available information, it is 4.4) [1].

Another noteworthy factor in this group is the attitude to meet environment pollution reduction obligations (question 3d in Fig. 8 and Fig. 9). This segment could, therefore, be targeted with tailored marketing strategies that directly address their specific concerns. For instance, highlighting case studies or built-in visualizations that demonstrate clear benefits for automatization and environmental benefits can reduce their uncertainties. This could be achieved by adding machine learning models that predict the monetary or environmental surpluses of certain strategies (e.g., impact on milk production by changing feeding strategies in combination with local climate models).

In the case of the **"partially ready" group**, Fig. 3 shows that this cluster consists of farmers who are somewhat familiar with smart technologies and may have explored their benefits but have not yet fully committed to adopting them. This subgroup tends to recognize the potential advantages of smart technologies but may still have reservations or face barriers that prevent either full adoption or the will to acquire smart technologies at all. With the right incentives and solutions, they can be the target group that can be converted into active users. Their awareness and interest in the technology make them more accessible than the "Not Ready" group, who may require more effort to educate and convince.

Given that the "partially ready" group is spread across different stages of technology adoption (Fig. 3), modular and scalable solutions that can grow with the farmer's needs are particularly valuable. These products should allow farmers to start small and expand their use of smart technologies as they become more comfortable and see the benefits. This is also reflected in Fig. 7, 5, and 9 in which operability was highlighted as a barrier but also a potential driver for technology implementation for the "partially ready" group. Hereby, the Lime and ICE plots indicate that an increase of that attitude to 3 or higher shows a significant impact on the prediction outcome. Also, the perceived reliability and available education (Fig. 9 and 7) have been identified as specific barriers for that group. Reducing the complexity of smart technologies (e.g., limiting data and model dependencies) can increase the ease of operation but also favor reliability [27]. Interpretable machine learning models that can support users in their assessments can further ease the use of such technologies and foster education [26,29,30].

Another barrier for the "partially ready" group is the attitude toward the ability to subsidize labor shortage (question 3a). This can be seen in all three XAI methods and was already discussed as an important factor for all three groups in subsection 4.3.1.

In the case of the **"not ready" group**, Fig. 3 indicates that this cluster comprises farmers who are either unaware of smart technologies or do not yet see their relevance or value. This subgroup is the hardest to target as they are characterized by their skepticism of the applicability of smart technologies to their specific farming practices and limited access to the market while showing low values on educational support and available physical infrastructure [1]. However, this group may still be an important target as it represents the most significant faction in the question block that does not have a technology yet but intends to buy one in the future. Strategic investments could also increase the readiness of this group in the long run, enabling long-term growth of technology providers.

Fig. 6 indicates that this group has fundamental barriers to technology adoption due to a lack of available information about technologies (question 4c), available technical assistance (question 4d), and accessibility of the market (question 4a). This is also visible in the SHAP analysis (see Fig. 9) as well as the ICE plots (5. Combating these barriers necessitates significant attention to technology design and marketing strategies. Establishing partnerships with local distributors and retailers to ensure that the technology is easily accessible to farmers in remote areas would be a strategy to tackle accessibility barriers. Built-in educational resources, as well as adaptable and intuitive user designs, can support the provision of information and assistance [26]. This could be combined with localized content that addresses specific regional needs and conditions (e.g., as already assessed partly in this study), making the technology more relevant and understandable for farmers. Enhanced understandability of the technology and its impact might be particularly crucial as this group also lacks positive attitudes about the potentials of smart technologies for labor shortage (question 3a), ease of operation (question 3a), and environmental obligations (question 3d), but also production effectiveness (question 3e) to a lesser degree (as seen in Fig. 6). Novel methods in complexity science (e.g., phase space reconstruction, entropy metrics) that allow for increased monitoring capacities of environmental pollutants, greenhouse gases, animal health, or intake efficiency could increase the trust of such technologies [28]. The ICE plot 5 also indicates that the technologies must be perceived as very affordable (equal or higher than 4) in order to influence the decision of adoption.

5. Discussion

This study incorporates several different machine learning modeling approaches (clustering, supervised machine learning, and XAI surrogate models). The choice of algorithms has an influence on the modeling results and should be evaluated carefully with metrics that evaluate the usability of clusters and classification/regression results (e.g., distance metrics, accuracy). A detailed evaluation of the chosen methods can be found in [1]. Compared to the prior study that assessed cluster validity of technological readiness and discussed general barriers to technology adoption [1], this research extended the focus on individual clusters and mechanisms to identify attributes that are particularly important to increase technological readiness for the respective subgroups. This allows for the targeted design of precision livestock farming technologies as well as policies.

The current study underscores prior research findings by [13,37,35,15], which postulated that the trust in the technologies capabilities and the robustness are major barriers for integration. It was shown in section 4 that reliability is an important factor, while the current study narrowed down distinct expectations that are particularly important for technological readiness (e.g., tech helps labor shortage, environmental pollution). Furthermore, interoperability was highlighted for the "not ready" and "partially ready" group, thereby confirming prior studies by [13,19]. However, we could not identify that security is a limitation to technology adoption. This seems to be an ambivalent topic as prior studies either showed its importance [13] or their minor influence [37] for the current precision livestock farming domain. This study also highlighted the ease of getting information as a primary barrier, which was also mentioned particularly by [37] that described the lack of awareness about existing technologies for technology adoption.

This study chose simpler models that are easy to reproduce as well as to limit the randomness and obscurity of more sophisticated machine learning approaches. However, there remains some instability in the predictions and explanations, particularly for methods like LIME. It was shown in a simulated setting that LIME explanations of close points can vary considerably [31]. Therefore, interpretations on single instances should be made cautiously and compared with other explainability approaches. This study tried to combat this as it accumulates the results for each cluster and does not evaluate single instances. Furthermore, recommendations during requirement analysis are mostly based on a combination of Explainable AI methods.

Another potential threat for inconsistency is the rather small sample size in machine learning terms (266 samples) of this study. Bigger surveys would result in more stable models, particularly for the detailed cluster analysis and supervised machine learning approach. Larger sample sizes could also positively affect the accuracy of the results. If companies have access to surveys with several thousand responses, the advantage of explainable machine learning methods becomes even more pronounced.

Each technique presents distinct advantages and limitations in the domain of Explainable AI. ICE plots effectively illustrate the range of influences on the prediction outcome if a single variable is changed. However, such insights may not be as visible in Local LIME plots, which focus on local perturbations of the data. For instance, if the original variable value is low (e.g., 2), the impact of higher values (e.g., 4 or 5) may not be visible in a LIME plot, as it only perturbs the vicinity of the original point. Nonetheless, while ICE plots offer a broader view, their depiction of variable changes could be theoretical, as it may not be feasible to alter one variable independently of others in practical scenarios. SHAP plots, on the other hand, do not directly indicate the effect of individual variable changes but often provide a visualization of how the readiness or non-readiness groups deviate from the average overall prediction of the total sample group.

It should be noted that questions without significant influence on the model prediction might still be general barriers to technology adoption. For example, affordability of smart technologies (question 4b) was not identified as a primary barrier to increasing technological readiness. As this study used a random forest approach, the algorithm uses questions more prominently that separate classes of technological readiness well (e.g., most ready farmers are above 4, while most not ready farmers are equal or lower than 2 for a given question), thereby increasing its importance to the prediction outcome. It could still be the case that affordability is a relevant factor for all three groups but does not provide viable information to separate these classes.

The study's design is focused on assessing technological, market, and psychological factors that influence technological readiness and, ultimately, can be influenced by technology design. This is a general limitation as it doesn't include broader sociodemographic barriers such as age, income, or farm size. Furthermore, this study only investigates farmers' barriers within the European Union (e.g., Sweden, Hungary, Denmark, Poland) or the Middle East (Israel). Different geographical areas will have other barriers to technology integration [47] and should be assessed separately. This is particularly crucial for the cluster analysis prior to model development. Depending on the scope of the analysis, different clusters and associations are possible and required. To find more specific requirements for one's products and services, a tailor-made survey referencing functionalities and aiming at particular countries or target groups would enhance the usability of the results. This is critical as the current study assesses the general barriers of smart technologies in precision livestock farming but doesn't focus on individual technologies (e.g., monitoring or feeding systems). Future work encompasses, therefore, the use of such techniques on different datasets and study goals. In case of bigger sample sizes, more advanced machine learning techniques could be applied (e.g., deep learning methods). However, such methods would be computationally intensive and need adequate computational resources.

Ultimately, a data-driven requirement analysis approach supports the development of precision livestock farming technology based on targeted consumer needs. These methods thereby increase the ease of operation and utilization of economic and environmental opportunities. This includes higher chances that the technologies increase production efficiency or animal well-being.

6. Conclusion

The article presented how Explainable AI approaches can be a valuable tool for companies and researchers to advance their understanding of functional requirements. It identifies user qualities that increase or limit technology adoption, helping companies achieve their business goals/philosophy (e.g., battling climate change), or identifying baseline barriers that might trigger a positive cascade of improved attitudes toward technological innovations in the precision livestock farming domain (e.g., available support, attitude towards automatization capabilities). By calculating clusters of technological readiness as a proxy for technology adoption and using them as labels in a machine learning approach, the authors utilize Explainable AI (XAI) techniques to investigate the influence of individual features on the prediction outcome (technological readiness). In doing so, this study highlights the dynamic interplay between user attitudes, market access, and environmental factors that influence technology adoption and highlight associated barriers. It is shown that individual clusters of readiness display common but also unique attributes that positively or negatively influence their behavior. Fundamental barriers are identified for all groups such as accessibility of the market, availability of information on smart technologies, and the ability to help with labor shortages. Unique barriers include interoperability of smart technologies for the "ready" cluster and operability of smart technologies for the "partially ready" group. The "not ready" group, next to the fundamental barriers, showed particularly low values for technical assistance available to smart technologies. In general, it was shown that a combination of XAI techniques provides a new toolset for targeted requirements and market analysis, building up new opportunities for technology design and business strategies. Associated technological examples to overcome identified barriers have been given. Further work must be done in this regard with a more specified focus on certain technologies, target groups, and novel mechanisms to increase the understandability and operability of said XAI tools.

CRediT authorship contribution statement

Kevin Mallinger: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Luiza Corpaci:** Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation. **Thomas Neubauer:** Supervision, Funding acquisition. **Ildikó E. Tikász:** Data curation. **Georg Goldenits:** Visualization. **Thomas Banhazi:** Writing – review & editing, Validation, Funding acquisition.

Funding

The authors acknowledge TU Wien Bibliothek for financial support through its Open Access Funding Programme. SBA Research (SBA-K1) is a COMET Centre within the COMET – Competence Centers for Excellent Technologies Programme and funded by Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology, Federal Ministry of Labor and Economic Affairs, and the federal state of Vienna. COMET is managed by FFG. Additionally, the financial support by the Austrian Federal Ministry of Labour and Economic Affairs, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association is gratefully acknowledged. Furthermore, the authors acknowledge the funding of the project "Digitale Landwirtschaft", by the Austrian Federal Ministry of Education, Science and Research within the funding scheme "Digitale und soziale Transformation in der Hochschulbildung". This article reports on the result related to the LivestockSense project. This research was made possible by funding from the ICT-AGRI-FOOD 2019 Joint Call. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 862665. The authors acknowledge the contribution of AgHiTech Kft (HU), Institute of Agricultural Economics (HU), Galilei Research Institute Ltd. (IL), SBA Research (AT), Innvite ApS. (DK), Swedish University of Agricultural Sciences (SE), Wroclaw University of Environmental and Life Sciences (PL), Estonia University of Life Sciences (EE), Aarhus University (DK), Denmark and the confounding of the following organizations: NRDI Funds (HU), Hungary, Israel Innovation Authority (IL), Bundesministerium, LRT Fund (AT), GUDP (DK), Ministry of Rural Affairs (EE), The National Centre for Research and Development (PL), Poland and FORMAS (SE).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Kevin Mallinger and Thomas Banhazi reports financial support was provided by ERA-Net Cofund scheme. Kevin Mallinger reports financial support was provided by Austrian Research Promotion Agency. Thomas Neubauer reports financial support was provided by Austrian Federal Ministry of Education, Science and Research. If there are other authors,

Appendix A. Questions

Table A.2
Survey questions used in this study. Question blocks 1-5 have been utilized for the Explainable AI analysis, while question block 6 was used as a proxy to assess cluster validity.

Feature	Question
	1. Please state the average availability of internet access at your farm (0: I don't know, 1: No availability. . . . 4: High availability)
	2. Please state the average level of automatization at your production farm (0: I don't know, 1: Less than 10 y/o, 2: 10-20 y/o, 3: diverse, 4: Over 20 y/o)
	3. Please, indicate how much you agree with the statements on smart devices/technologies (sensors, cameras robots, farm management information system etc.), regardless of whether using them or not in the farm you represent. (0: I don't know, 1: Strongly disagree. . . . 5: Strongly agree) SMART DEVICES/TECHNOLOGY. . .
Q3_a	. . . help/support to cope with labor shortage.
Q3_b	. . . help/support day-to-day decision making in the livestock buildings.
Q3_c	. . . help/support enterprise, marketing and investment decisions.
Q3_d	. . . help/support to meet environmental pollution reduction obligations.
Q3_e	. . . enable to increase the effectiveness of production.
Q3_f	. . . provide reliable information.
Q3_g	. . . provide information in a real-time manner.
	4. Regarding the availability of smart technologies, please, indicate how much you agree with the following statements. (0: I don't know, 1: Strongly disagree. . . . 5: Strongly agree)
Q4_a	It is easy to access smart technologies on the market.
Q4_b	Smart technologies can be purchased at an affordable price.
Q4_c	It is easy to get information on smart technologies and distributors.
Q4_d	It is easy to get technical assistance to smart technologies.
Q4_e	Proper education is available for using smart technologies.
	5. Regarding the operation of smart technologies, please, indicate how much you agree with each of the statements. (0: I don't know, 1: Strongly disagree. . . . 5: Strongly agree) SMART DEVICES/TECHNOLOGY.....
Q5_a	...can be maintained at a reasonable cost.
Q5_b	...are easy to operate.
Q5_c	...can be connected well with other equipment/software.
Q5_d	...operate in a reliable manner.
Q5_e	...are secure in terms of data management.
	6. Do you use smart devices (sensors, cameras, robots etc.) at the farm you represent? (0: I don't know, 1: Yes. 2: No.)

Appendix B. Statistical overview

Table B.3
Statistical overview of cluster results based on the survey answers for $k = 3$ clusters. Higher values for the mean indicate stronger agreement with the question, whereas lower values are associated with disagreement (as seen in [1]). The interquartile range (IQR) is also displayed per cluster to describe the range of 50 percent of the cluster observations.

Feature	Question	Ready		Partially Ready		Not Ready	
		Mean ± Std	IQR	Mean ± Std	IQR	Mean ± Std	IQR
Q1. Average availability of internet access at your farm (Scale: 1–4)		3.225 ± 0.968	1.0	2.752 ± 0.900	2.0	2.618 ± 0.821	2.0
Q2. Average level of automatization at your production farm (Scale: 0–2)		1.960 ± 0.190	0.0	1.899 ± 0.367	0.0	1.824 ± 0.474	0.0
Q3. SMART DEVICES/TECHNOLOGY... (Scale: 1–5)							
Q3_a	... help/support to cope with labor shortage.	4.065 ± 1.036	2.0	3.153 ± 1.469	2.0	2.745 ± 1.405	1.0
Q3_b	... help/support day-to-day decision making in the livestock buildings.	4.513 ± 0.930	2.0	3.148 ± 1.459	2.0	2.335 ± 1.354	1.0
Q3_c	... help/support enterprise, marketing and investment decisions.	4.698 ± 0.489	2.0	4.321 ± 0.842	2.0	3.251 ± 1.271	1.0
Q3_d	... help/support to meet environmental pollution reduction obligations.	4.587 ± 0.584	2.0	4.095 ± 0.779	2.0	3.152 ± 1.229	1.0
Q3_e	... enable to increase the effectiveness of production.	4.738 ± 0.669	3.0	4.479 ± 0.751	1.0	3.415 ± 1.473	1.0
Q3_f	... provide reliable information.	4.852 ± 0.450	2.0	3.505 ± 1.128	2.0	2.118 ± 1.380	1.0
Q3_g	... provide information in a real-time manner.	3.673 ± 0.733	1.0	2.409 ± 1.107	1.0	1.507 ± 0.963	1.0
Q4. Indicate how much you agree with the following statements (Scale: 1–5)							
Q4_a	It is easy to access smart technologies on the market.	4.763 ± 0.484	1.0	3.059 ± 1.120	2.0	1.888 ± 1.245	1.0
Q4_b	Smart technologies can be purchased at an affordable price.	4.304 ± 0.692	1.0	2.746 ± 1.186	1.0	1.393 ± 0.923	1.0
Q4_c	It is easy to get information on smart technologies and distributors.	4.366 ± 1.075	1.0	2.346 ± 1.199	1.0	1.872 ± 1.154	1.0
Q4_d	It is easy to get technical assistance for smart technologies.	3.819 ± 0.933	1.0	2.989 ± 1.238	2.0	1.614 ± 1.129	1.0
Q4_e	Proper education is available for using smart technologies.	4.631 ± 0.662	1.0	3.433 ± 0.946	1.0	2.121 ± 1.151	4.0
Q5. SMART DEVICES/TECHNOLOGY... (Scale: 1–5)							
Q5_a	... can be maintained at a reasonable cost.	3.825 ± 0.921	1.0	2.983 ± 1.225	1.0	1.618 ± 1.131	1.0
Q5_b	... are easy to operate.	4.554 ± 0.828	1.0	3.444 ± 0.992	1.0	1.825 ± 1.272	1.0
Q5_c	... can be connected well with other equipment/software.	3.824 ± 1.089	3.0	2.952 ± 1.455	1.0	1.504 ± 1.459	1.0
Q5_d	... operate in a reliable manner.	4.546 ± 0.830	2.0	3.446 ± 1.004	1.0	1.806 ± 1.246	1.0
Q5_e	... are secure in terms of data management.	3.796 ± 1.075	3.0	2.942 ± 1.433	2.0	1.470 ± 1.246	1.0
Q6. Do you use smart devices at the farm you represent? (Scale: 1–3)		2.756 ± 0.494	2.0	2.568 ± 0.705	1.0	2.110 ± 0.884	0.0

Appendix C. Cluster validation

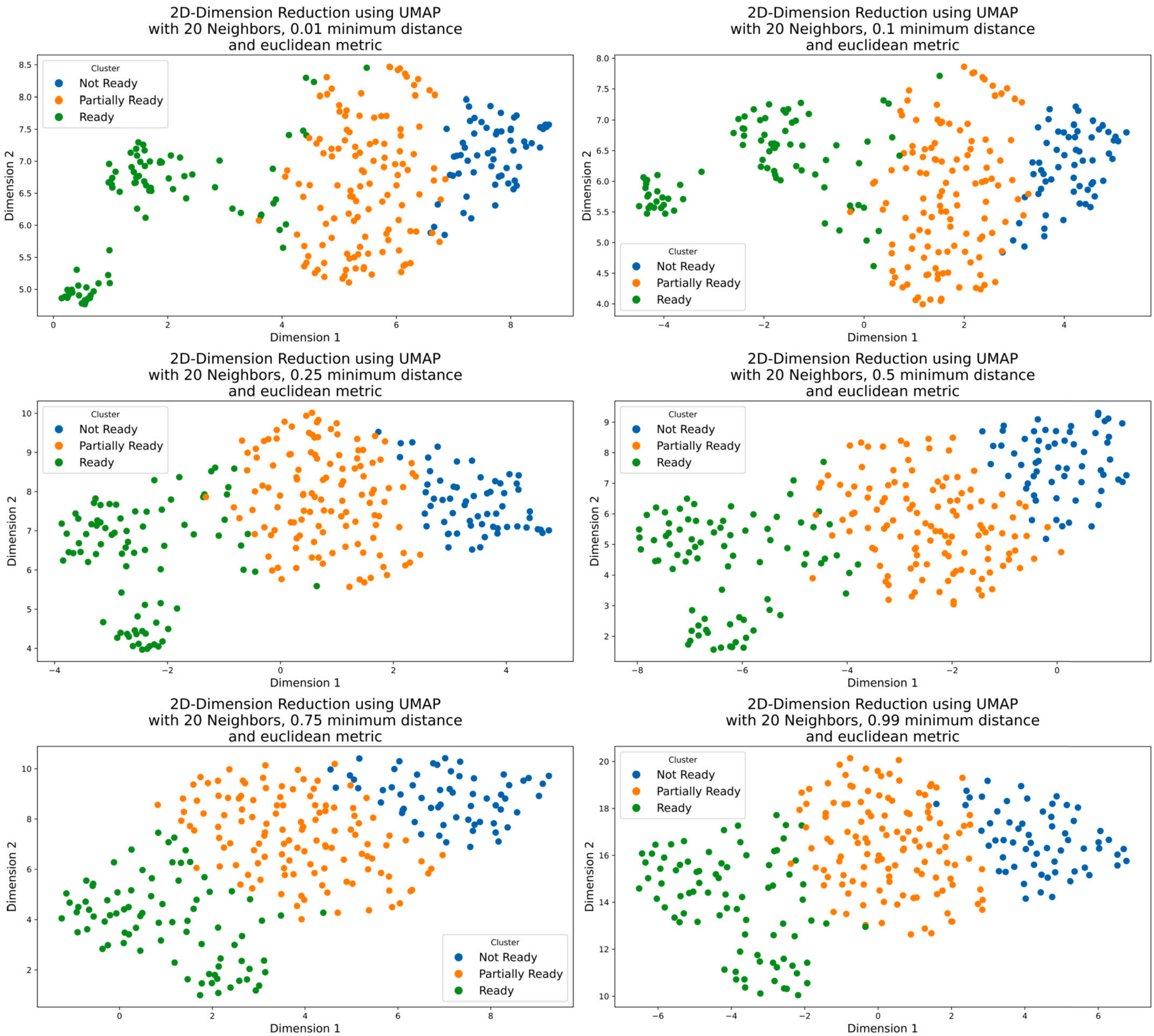


Fig. C.10. Overview of a two-dimensional cluster and sample distribution using UMAP. Neighbors have been chosen with 20. Different variations of distances are displayed to show local and global cluster behavior.

Fig. C.11. Overview of a two-dimensional cluster and sample distribution using UMAP. Neighbors have been chosen with 50. Different variations of distances are displayed to show local and global cluster behavior.

Data availability

The data used is referenced in the article respectively in Section 3.1.

References

[1] K. Mallinger, L. Corpaci, T. Neubauer, I.E. Tikász, T. Banhazi, Unsupervised and supervised machine learning approach to assess user readiness levels for precision livestock farming technology adoption in the pig and poultry industries, Comput. Electron. Agric. 213 (2023) 108239, Elsevier.

[2] K.E. Wiegers, J. Beatty, Software Requirements, 3rd ed., Microsoft Press, 2013.

[3] R.S.R. Kunte, et al., Challenges in implementing AI technology smart farming in agricultural sector–a literature review, Int. J. Manag. Technol. Soc. Sci. 9 (2) (2024) 283–301.

[4] I.E. Tikász, C. Bálint, G. Király, C.A.G. Sorensen, S. Opalinski, S. Gunnarsson, U. Marchaim, E. Kokin, K. Mallinger, T. Banhazi, A survey of pig and poultry farmers' readiness and attitudes towards smart technologies, in: J.C. Barbosa, L.L. Silva, J.C. Rico, D. Coelho, A. Sousa, J.R.M. Silva, F. Baptista, V.F. Cruz (Eds.), Proceedings of the XL CIOSTA & CIGR Section V International Conference in Évora, Portugal, 10-13 of September, Universidade de Évora, 2023, pp. 92–100.

[5] D. Araújo, A.D. Neto, A. Martins, Representative cross information potential clustering, Pattern Recognit. Lett. 34 (16) (2013) 2181–2191, Elsevier.

[6] D.L. Donoho, et al., High-dimensional data analysis: the curses and blessings of dimensionality, in: AMS Math Challenges Lecture, vol. 1, 2000, p. 32.

[7] A. Rényi, On measures of entropy and information, in: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics, University of California Press, 1961, pp. 547–561.

[8] Q. Zhao, M. Xu, P. Fränti, Sum-of-squares based cluster validity index and significance analysis, in: International Conference on Adaptive and Natural Computing Algorithms, Springer, 2009, pp. 313–322.

[9] L. McInnes, J. Healy, J. Melville, UMAP: uniform manifold approximation and projection for dimension reduction, arXiv preprint, arXiv:1802.03426, 2018.

[10] K. Pohl, Requirements Engineering: Fundamentals, Principles, and Techniques, Springer, 2010.

[11] B.G. Tabachnick, L.S. Fidell, J.B. Ullman, Using Multivariate Statistics, vol. 6, Pearson, Boston, MA, 2013.

[12] L. Breiman, Statistical modeling: the two cultures (with comments and a rejoinder by the author), Stat. Sci. 16 (3) (2001) 199–231, Institute of Mathematical Statistics.

[13] J.L. Drewry, J.M. Shutske, D. Trechter, B.D. Luck, L. Pitman, Assessment of digital technology adoption and access barriers among crop, dairy, and livestock producers in Wisconsin, Comput. Electron. Agric. 165 (2019) 104960, Elsevier.

[14] C. Krampe, J. Serratosa, J.K. Niemi, P.T.M. Ingenbleek, Consumer perceptions of precision livestock farming—a qualitative study in three European countries, Animals 11 (5) (2021) 1221, MDPI.
[15] E. Lima, T. Hopkins, E. Gurney, O. Shortall, F. Lovatt, P. Davies, G. Williamson, J. Kaler, Drivers for precision livestock technology adoption: a study of factors associated with adoption of electronic identification technology by commercial sheep farmers in England and Wales, PLoS ONE 13 (1) (2018) e0190489, Public Library of Science, San Francisco, CA.
[16] T. Banhazi, V. Halas, F. Maroto-Molina, Practical Precision Livestock Farming: Hands-on experiences with PLF technologies in commercial and R&D settings, Wageningen Academic Publishers, 2022.
[17] M.Z. Dhraief, S. Bedhiaf-Romdhania, B. Dhehibi, M. Oueslati-Zlaouia, O. Jebali, S. Ben Youssef, Factors affecting the adoption of innovative technologies by livestock farmers in arid areas of Tunisia, FARA Res. Rep. 3 (5) (2018) 22.
[18] International Institute of Business Analysis, A Guide to the Business Analysis Body of Knowledge (BABOK Guide), 3rd ed., 2015.
[19] D. Pivoto, P.D. Waquil, E. Talamini, C.P.S. Finocchio, V.F. Dalla Corte, G. de Vargas Mores, Scientific development of smart farming technologies and their application in Brazil, Inf. Process. Agric. 5 (1) (2018) 21–32, Elsevier.
[20] I. Sommerville, P. Sawyer, Requirements Engineering: A Good Practice Guide, John Wiley & Sons, 1997.
[21] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst. 30 (2017) 4765–4774.
[22] M.T. Ribeiro, S. Singh, C. Guestrin, "Why should I trust you?": explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144.
[23] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer Science & Business Media, 2009.
[24] A.K. Jain, Data clustering: 50 years beyond K-means, Pattern Recognit. Lett. 31 (8) (2010) 651–666, Elsevier.
[25] T. Banhazi, B. Ji, D. Rutley, C.J.C. Phillips, Modelling the effects of environmental stress on weight gain in pigs, in: Practical Precision Livestock Farming, Wageningen Academic Publishers, 2022, pp. 193–210.
[26] K. Mallinger, R. Baeza-Yates, Responsible AI in farming: a multi-criteria framework for sustainable technology design, Appl. Sci. 14 (1) (2024) 437, MDPI.
[27] K. Mallinger, W. Purcell, T. Neubauer, Systemic design requirements for sustainable Digital Twins in precision livestock farming, in: Proceedings of the 10th European Conference on Precision Livestock Farming, Vienna, Austria, vol. 29, 2022.
[28] K. Mallinger, S. Raubitzek, T. Neubauer, S. Lade, Potentials and limitations of complexity research for environmental sciences and modern farming applications, in: Current Opinion in Environmental Sustainability, vol. 67, Elsevier, 2024, p. 101429.
[29] G. Goldenits, K. Mallinger, T. Neubauer, E. Weippl, Tabular reinforcement learning for robust, explainable CropRotation policies matching deep reinforcement learning performance, in: Copernicus Meetings, 2024.
[30] G. Goldenits, K. Mallinger, S. Raubitzek, T. Neubauer, Current applications and potential future directions of reinforcement learning-based Digital Twins in agriculture, arXiv preprint, arXiv:2406.08854, 2024.
[31] D. Alvarez-Melis, T.S. Jaakkola, On the robustness of interpretability methods, arXiv preprint, arXiv:1806.08049, 2018.
[32] C. Pylianidis, S. Osinga, I.N. Athanasiadis, Introducing digital twins to agriculture, Comput. Electron. Agric. 184 (2021) 105942, Elsevier.
[33] D. Berckmans, General introduction to precision livestock farming, Anim. Front. 7 (1) (2017) 6–11, Oxford University Press.
[34] S. Schukat, H. Heise, Towards an understanding of the behavioral intentions and actual use of smart products among German farmers, Sustainability 13 (12) (2021) 6666, MDPI.
[35] A.L. Boothby, D.R. White, Understanding the barriers to uptake of precision livestock farming (PLF) in the UK sheep industry, in: Agricultural Engineering AgEng2021, 2021, p. 572.
[36] T. Groher, K. Heitkämper, C. Umstätter, Digital technology adoption in livestock production with a special focus on ruminant farming, Animal 14 (11) (2020) 2404–2413, Cambridge University Press.
[37] A. Makinde, Investigating perceptions, motivations, and challenges in the adoption of precision livestock farming in the beef industry, Ph.D. thesis, University of Guelph, 2020.
[38] F. Abeni, F. Petrera, A. Galli, A survey of Italian dairy farmers' propensity for precision livestock farming tools, Animals 9 (5) (2019) 202, MDPI.
[39] H.S. Pathak, P. Brown, T. Best, A systematic literature review of the factors affecting the precision agriculture adoption process, Precis. Agric. 20 (2019) 1292–1316, Springer.
[40] A.I. Ugochukwu, P.W.B. Phillips, Technology adoption by agricultural producers: a review of the literature, in: From Agriscience to Agribusiness: Theories, Policies and Practices in Technology Transfer and Commercialization, Springer, 2018, pp. 361–377.
[41] T. Banhazi, V. Halas, F. Maroto-Molina, Practical Precision Livestock Farming (PLF): Hands-on Experiences with PLF Technologies in Commercial and R&D Settings, Wageningen Academic Publishers, 2022, 442 pages.
[42] Y. Nawaz, M. Ullah, I. Ullah, Z. Khan, M.U. Qadir, I.U. Haq, Evaluating the use of artificial intelligence and machine learning in the livestock industry: a review, in: Proceedings of the 1st International Conference on Computing Technologies, Tools and Applications, 2023.
[43] J. Kim, Y. Geum, Y. Park, Integrating customers' disparate technology readiness into technological requirement analysis: an extended Kano approach, Total Qual. Manag. Bus. Excell. 28 (5–6) (2017) 678–694.
[44] L. Zhao, W. Alhoshan, A. Ferrari, K.J. Letsholo, M.A. Ajagbe, E.-V. Chioasca, R.T. Batista-Navarro, Natural language processing for requirements engineering: a systematic mapping study, ACM Comput. Surv. 54 (3) (2021) 1–41.
[45] T. Banhazi, A. Banhazi, I. Tikasz, S. Palotay, K. Mallinger, T. Neubauer, et al., Facilitating PLF technology adoption in the pig and poultry industries, Stud. Agric. Econ. 126 (1) (2024).
[46] I.E. Tikász, G. Király, C. Bálint, S. Opalinski, C.A.G. Sorensen, S. Gunnarsson, U. Marchaim, E. Kokin, K. Mallinger, T. Banhazi, Conditions of applying advanced information technologies in livestock farming, in: J.C. Barbosa, L.L. Silva, J.C. Rico, D. Coelho, A. Sousa, J.R.M. Silva, F. Baptista, V.F. Cruz (Eds.), Proceedings of the XL CIOSTA & CIGR Section V International Conference in Évora, Portugal, 10-13 of September, Universidade de Évora, 2023, pp. 101–110.
[47] I. Kopler, U. Marchaim, I.E. Tikasz, S. Opalinski, E. Kokin, K. Mallinger, et al., Farmers' perspectives of the benefits and risks in precision livestock farming in the EU pig and poultry sectors, Animals 13 (18) (2023) 2868, MDPI.
[48] B.M. Greenwell, et al., pdp: an R package for constructing partial dependence plots, R J. 9 (1) (2017) 421.
[49] A. Goldstein, A. Kapelner, J. Bleich, E. Pitkin, Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation, J. Comput. Graph. Stat. 24 (1) (2015) 44–65, Taylor & Francis.
[50] A. Kumari, M.N. Khan, A.K. Sinha, Harvesting intelligence: AI and ML revolutionizing agriculture, in: Data-Driven Farming, Auerbach Publications, 2024, pp. 126–141.
[51] L. Trapanese, M. Hostens, A. Salzano, N. Pasquino, Short review of current limits and challenges of application of machine learning algorithms in the dairy sector, Acta IMEKO 13 (1) (2024) 1–7.
[52] X. Yang, R.B. Bist, B. Paneru, T. Liu, T. Applegate, C. Ritz, et al., Computer vision-based cybernetics systems for promoting modern poultry farming: a critical review, Comput. Electron. Agric. 225 (2024) 109339, Elsevier.