Ciência e Agrotecnologia
SCIENCE AND AGROTECHNOLOGY

Ciência e Agrotecnologia, 47:e018922, 2023
http://dx.doi.org/10.1590/1413-7054202347018922

Agricultural Sciences
eISSN 1981-1829

Computer vision by unsupervised machine learning in seed drying process

Visão computacional por aprendizado de máquina não supervisionado no processo de secagem de sementes

Romário de Mesquita Pinheiro[1*], Gizele Ingrid Gadotti[2], Ruan Bernardy[2], Rafael Rico Tim[3], Karine Von Ahn Pinto[3], Graciela Buck[3]

[1]Instituto Nacional de Pesquisas da Amazônia/INPA, Núcleo de Pesquisas do Acre/NPAC, Rio Branco, AC, Brasil
[2]Universidade Federal de Pelotas/UFPEL, Centro de Engenharias/CENG, Pelotas, RS, Brasil
[3]Universidade Federal de Pelotas/UFEPL, Faculdade de Agronomia Eliseu Maciel/FAEM, Pelotas, RS, Brasil
*Corresponding author: romario.ufacpz@hotmail.com
Received in December 18, 2022 and approved March 17, 2023

ABSTRACT

Analyzing the impact of harvest-time drying data is crucial for successful storage and maintaining regulatory seed quality. This study aimed to assess the performance of fixed and mobile dryers using machine learning techniques. Data were collected from convective dryers, including the total number of dryers used, drying time (in hours), moisture percentages at the product's entrance and exit, and the humidity difference between them. The study employed the Filtered Clusterer model, which utilizes the Simple *K-Means* technique and the Resample filter to group data based on similarities. The findings indicated distinct differences between fixed and mobile drying systems, with well-defined variations within each system. The algorithm, combined with the applied filters, proved effective in unsupervised classification by identifying and reducing inter-cluster similarity within the fixed system, thereby creating distinct classes within the dataset. In conclusion, the algorithm successfully clustered the scattered dataset and accurately classified and minimized inter-cluster similarity within the fixed system. Conversely, the mobile system exhibited low drying efficiency.

Index terms: Artificial intelligence; digital agriculture; post-harvest technology.

RESUMO

Analisar os efeitos de dados de secagem em época de safra é primordial para o sucesso do armazenamento, além de manter a qualidade fisiológica das sementes. O objetivo com este trabalho foi avaliar o desempenho de secagem em secadores fixo e móveis através de aprendizado de máquinas. Os dados foram coletados de secadores convectivos, com base em: i) total de secadores trabalhados; ii) tempo de duração de secagem (horas); iii) percentual de umidade na entrada e saída do produto (%); e iv) diferenças de umidade entres ambas (%). Foi utilizado o modelo Filtered Clusterer, baseado na técnica Simple *K-Means* e filtro Resample, para agrupar dados com base em suas similaridades. Os resultados apresentados demonstram que o sistema de secagem é distinto entre os equipamentos fixo e móvel, com diferenças bem definidas dentro dos sistemas de secagem estabelecidos. O algoritmo juntamente com os filtros demonstrou ser eficiente na classificação não supervisionada, identificando e minimizando similaridade inter-cluster do sistema fixo definindo em classes distintas dentro do conjunto de dados. Conclui-se, que o conjunto de dados dispersos são agrupados de modo que o algoritmo classifica e minimiza similaridade inter-cluster do sistema fixo com alta precisão. Por outro lado, o desempenho da secagem em sistema móvel apresentou baixa eficiência.

Termos para indexação: Inteligência artificial; agricultura digital; tecnologia de pós-colheita.

INTRODUCTION

Artificial intelligence, particularly through deep machine learning and data mining, is a powerful and interconnected set of tools for optimizing design and performance, analyzing data, and monitoring and controlling systems (Wang et al., 2020). These tools are crucial for addressing various challenges in the seed sector, including decision-making related to germination, vigor, pathogenicity detection, storage quality, and drying conditions.

The increasing demand for smart agriculture has resulted in significant advancements in crop estimation and forecasting, thereby enhancing productivity (Sharma et al., 2022). However, the post-harvest and pre-sowing stages require careful attention due to the extensive information generated to obtain timely insights into seed quality and maximize crop yield.

Obtaining information about cultivars, treatments, harvest, sieving, purity, sanity, percentage of other seeds, germination, and vigor during post-harvest stages is time-consuming and inefficient for making quick decisions (Pinheiro et al., 2021). These factors are crucial in the quality assessment process. However, the processing, drying, and storage stages must not be overlooked, as they are essential for successfully completing the agricultural cycle.

Artificial intelligence has emerged as an indispensable tool for addressing the challenges posed by the dynamic nature of high data generation in the seed sector. It is imperative to address the efficiency issues of agricultural equipment, particularly artificial dryers, in order to ensure product safety and quality. Additionally, the choice of drying method is limited by the quantity of seeds, making artificial drying essential for large volumes. The operating costs of artificial drying are affected by factors such as volume, drying speed, and air temperature (Garcia et al., 2004).

Seed drying plays a crucial role in preserving the physiological quality of seeds during storage, enabling early harvesting and minimizing losses in the production process (Dhurve et al., 2022). Furthermore, data on the water content in seeds and grains aid in decision-making regarding equipment efficiency. Unsupervised machine learning enabled the discovery of unguided responses, allowing the predictive model itself to make decisions. Data mining, as described by Onukwuli et al. (2021), is a technique for extracting essential patterns and knowledge from large datasets. Machine learning encompasses various algorithms that learn predictive rules from historical data to build models for predicting future data (Arumugam et al., 2022).

The seed and grain dryer play a vital role in preserving the physiological quality of seeds. However, it is crucial to comprehend the information generated through agricultural equipment during the drying process, as mishandling can be risky and lead to irreversible damage if conducted without the necessary knowledge and care.

Based on the aforementioned considerations, the objective of this study was to assess the drying performance of fixed and mobile dryers using machine learning techniques.

MATERIAL AND METHODS

The study focused on two types of agricultural dryers based on their installation conditions: fixed dryers, which are installed under a shed, and mobile dryers, which are installed on truck trailers and transported between crops. These dryers operate on a large scale and utilize a gas heating system, along with a control panel.

The configuration of the dryers includes a fan unit powered by an electric motor, flow control valves, a gas heater, a drying chamber, and sensors for measuring air temperature and seed mass. The mobile dryer was transported to the farm and mounted on a cargo transport truck, while the fixed dryer was installed in properties with enclosed sheds (Figure 1).

The studied dryer type has dimensions of 6.65 m in length, 2.50 m in width, and 3.25 m in height, with an octagonal shape. It consists of 12 cells that form a layer of seeds approximately 45 cm thick. The dryer has a total capacity of 17 tons or approximately 300 bags per load. It operates using an LPG gas system, where the gas passes through burners inside the dryer for heating. It is equipped with a 12.5 hp premium W22 model engine (1765 rpm rotation, 220–380 V voltage) and 795 mm propellers with 37.5° blades. The dryer also includes a digital control panel for managing the air temperature and gas pressure.

Data on the drying process of various crops such as soybeans, corn, barley, sesame, and beans were collected for the years 2017 to 2021. The collected drying data included information on the total number of dryers in operation at each client's facility, drying time (in hours), seed moisture percentage at the entry and exit points of the dryer, and humidity differences between the two points. Table 1 presents the drying data collected for both types of dryers across different crops.

The dataset comprised a total of 2028 data points related to drying attributes, emphasizing the importance of understanding the drying duration. Information was collected on entry and exit times, entry and exit humidity in both dryer systems, and temperature. These data were subsequently analyzed and compared between the fixed and mobile drying systems to establish decision-making processes in individual equipment and across the entire system.

To facilitate the initial step, artificial intelligence and data mining techniques were employed. The input data were transformed as per the requirement of the software. The raw data were organized and pre-processed to ensure accurate input and analysis. In this stage, the data received in Excel format (.xls) were processed to make them compatible with the software. This involved placing all attributes in a single row, with each value placed in columns below its respective attribute.

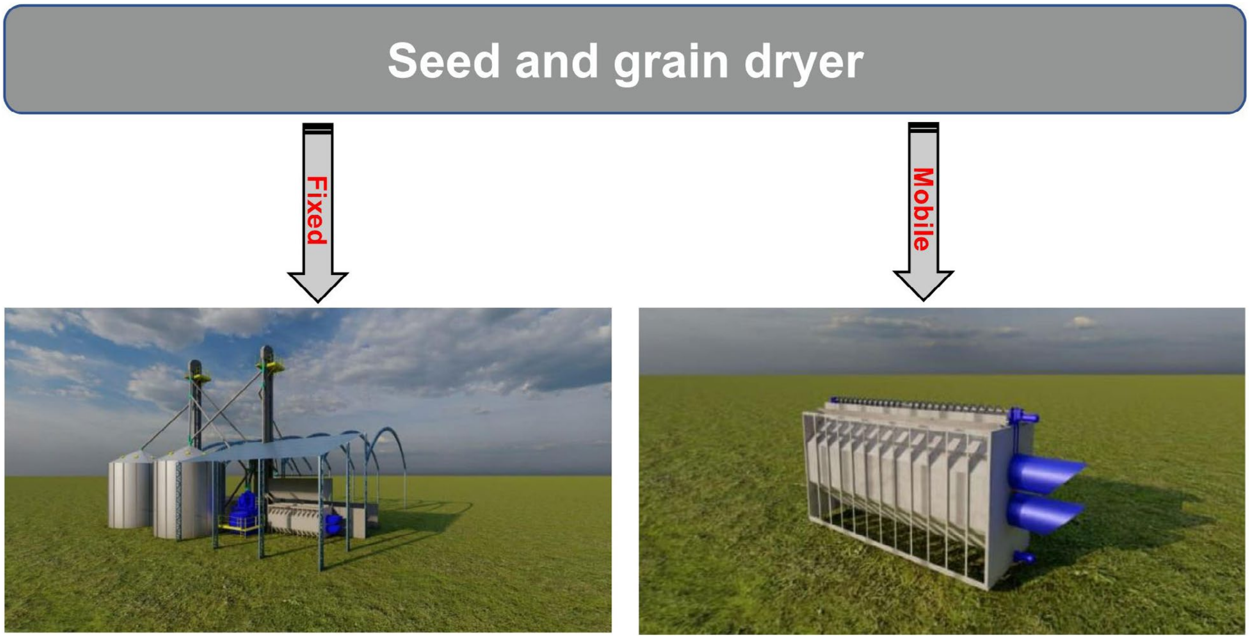

Figure 1: Schematic representation of drying systems, illustrating the installation of agricultural equipment. The fixed dryer is positioned beneath a shed, while the mobile dryer can be located either in the field or on a mobile vehicle. (Source: Adapted from Daurana Oliveira).

Table 1: Data on the analyzed drying quantity in the systems employed for machine learning integration.

Culture	Fixed dryer	Mobile dryer
Beans	71	473
Soybeans	115	84
Corn	161	1075
Barley	5	-
Sesame	2	-

Subsequently, the file was converted to .csv format, and the dataset was processed using Microsoft Windows Notepad. This step required replacing decimal "commas" with "periods" and replacing "semicolons," which divided the attribute columns, with "commas." Any lines containing missing or erroneous data, as identified during the spreadsheet analysis, were excluded during this preliminary processing. The data organization is illustrated in Figure 2.

To in order to use the system type (fixed or mobile) as an attribute, the term "fixed" was replaced with 1, and "mobile" was replaced with 2. This adjustment was necessary because the analysis of the algorithms was performed using the Weka software, version 3.8.5, which only accepts numeric data for attributes when utilizing the K-Means method. A clustering process was conducted to identify patterns in the dataset. The employed algorithm was the "Filtered Clusterer" for data mining.

The training data consisted of 66% of the dataset (1338 randomly selected drying instances). Additionally, filters such as "Resample" were applied, and various evaluation parameters were used in the mentioned algorithms. These parameters included the number of repetitions, the number of clusters, and the mathematical method used to identify data similarities, which were aimed to optimize the results.

The K-means algorithm, a commonly used partition-based clustering method, was employed in this study. This algorithm determines the closest centroid and designates it as the representative for a given training set. The K-Means algorithm identifies k clusters ($C = \{c1, ...ck\}$) from a dataset of n d-dimensional data points ($D = \{x1, ...xn\}$) by minimizing the squared distance between each data point and its nearest cluster center (Kumar; Reddy, 2017). Kumar and Reddy (2017) proposed an equation known as Squared Error Distortion (SED) to quantify this distance, as represented by Equation 1.

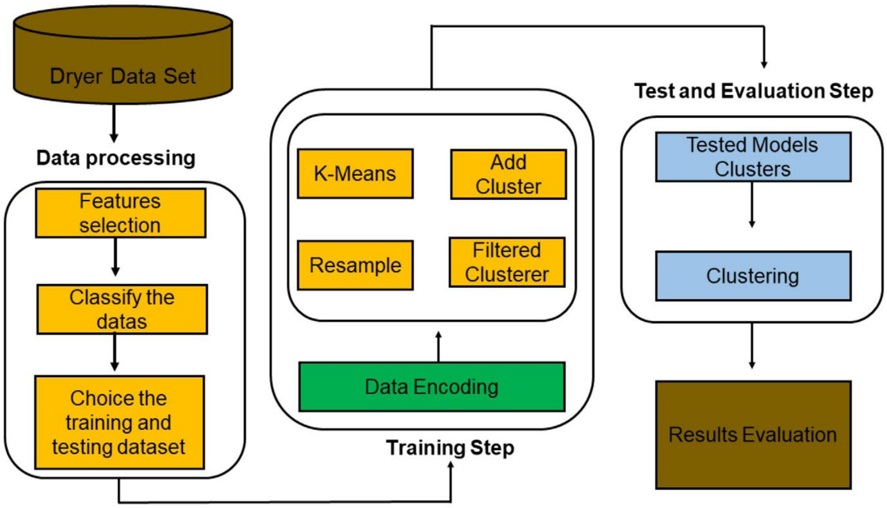

Figure 2: Flow diagram illustrating the sequence of data processing steps in the fixed and mobile drying systems.

$$SED = \sum_{i=1}^{n} \sum_{j=1...k}^{K} \left\| x_i - \overline{c}_j \right\|^2 \quad (1)$$

$\overline{c}_j$ is the mean of the cluster cJ

According to Yao et al. (2013), the K-Means clustering algorithm divides the data into pre-defined classes by minimizing the error function. This algorithm belongs to the category of partitioning feature space clustering algorithms, which are widely utilized.

To validate outcomes derived from the clustering algorithm, the Add Cluster filter was employed. This filter introduces a new attribute known as “Cluster” to the dataset. This filter utilizes the same algorithm to assign each data row to the corresponding cluster, enabling the determination of data accuracy in relation to the confusion matrix.

The calculation steps performed by the clustering algorithm are as follows: First, randomly selects from the dataset points as initial centers of the cluster. Secondly, the algorithm calculates the distance between each sample and the cluster centers, assigning the sample to the closest class. Thirdly, based on the clustering results, the algorithm recalculates the cluster centers by taking the arithmetic mean of all elements. Fourthly, the algorithm repeats this calculation method, updating the cluster centers. Fifthly, using the updated centers, the algorithm regroups all elements in the dataset. These steps are repeated until the grouping remains unchanged, at which point the final result is considered to be obtained.

RESULTS AND DISCUSSION

The findings of this research can assist in the decision-making process through the application of artificial intelligence, specifically unsupervised machine learning. This involves the identification and classification of data pertaining to seed drying into distinct groups. The Filtered Clusterer algorithm is employed on the dataset of both the fixed and mobile drying systems to perform a comprehensive analysis. The filter’s structure is solely based on the training data and ensures that the test instances are processed without altering their structure (Table 2).

Table 2 and Figure 3 illustrate the distinct separation achieved by the filter, resulting in two clusters based on the type of drying system used. This demonstrates notable differences in drying outcomes between the fixed and mobile equipment, with clear distinctions within each drying system.

Table 1 reveals a significant difference in the inlet and outlet humidity in relation to drying time. The fixed dryer achieved faster water removal (4.41 pp/h), even for seeds with a high moisture percentage. In contrast, the mobile dryer required more time to reduce seed moisture and did not meet the desired drying standards of 12% or 13% for processed crops.

After analyzing the results generated by the Filtered Clusterer, it was observed (Figure 3) that scattered data points still persisted. Subsequently, the dataset was reevaluated with the addition of the “X-Means” algorithm, where the software automatically determines the number of clusters and iterations.

Table 2: Results of data grouping for drying systems based on the variables specified in a standardized evaluation model. T. dryers represents the algorithm's predicted total number of dryers required to complete the task.

Attribute	Cluster 1	Cluster 2
Systems	1 (fixed)	2 (mobile)
T. dryers	5.15	6.06
Duration (h)	01:43	01:55
Humidity entry (%)	19.60	19.68
Output humidity (%)	13.28	14.72
Humidity difference (%)	6.31	4.95

Furthermore, the "Resample" filter was utilized to guide the software on the ideal number of repetitions for training. Ultimately, the classifier demonstrated optimal performance in classifying the dataset. However, the evaluation of the data using the K-Means model indicated that grouping into two classes did not account for all the information, resulting in scattered data points (Figure 3). The average total number of dryers operating in each system indicated similar drying times.

The K-Means model selects random points known as centroids from each group, calculates the arithmetic mean, and determines the result representing that specific group. This approach facilitates the separation of clusters from one another (Figure 3). The variables involved in the seed drying process provide crucial information for successful storage. The number of dryers during the drying stages also influences decision-making, leading to faster processing and necessitating efficiency in these drying stages. Figure 4 and Table 3 highlight the differences in drying efficiency between distinct locations and pre-established systems.

Table 3 presents the division into two clusters in the fixed drying system, as the dataset contained more scattered data. The total number of dryers required for water content reduction has minimal influence on this stage, as the drying times are comparable. The advantage of this dryer lies in its high efficiency, with two key features for seed drying: rapid air-drying speed (reaching 26 m/s) and suitable temperature (gas combustion at 35 –37 °C).

In this subsequent analysis, the algorithm classified the fixed system into three clusters based on the number of dryers employed. The first cluster consisted of a higher number of dryers, resulting in a shorter drying time for reducing a greater water content at the end of the process (5.41 pp/h). The second cluster had fewer dryers and a lower initial moisture content, resulting in a drying rate of 2.03 pp/h. Finally, the third cluster, with fewer dryers than the first cluster but exhibiting lower efficiency (3.27 pp/h), confirms the subpar performance of mobile dryers. Overall, the algorithm indicated that to decrease the drying time by an average of 5.06 pp/h, an increase in the number of dryers is necessary. However, high moisture content not only affects the drying process but also increases the labor required for all operational steps, highlighting the influence of humidity and temperature.

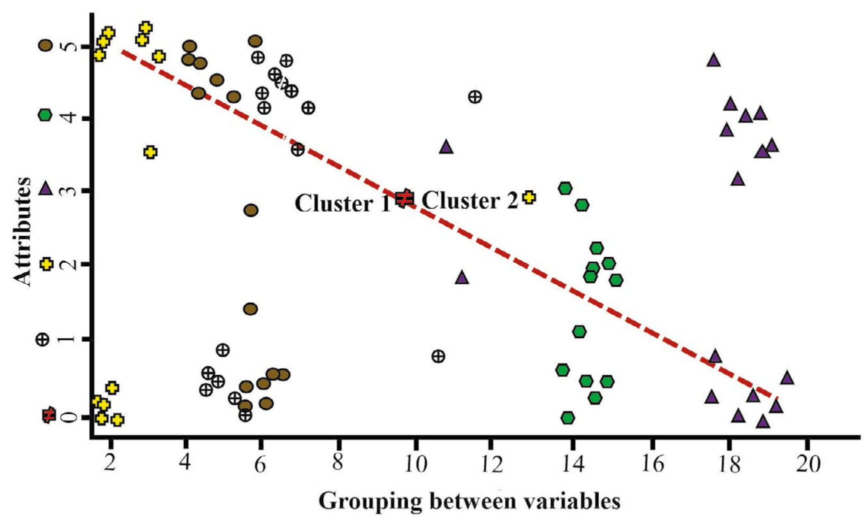

Figure 3: Illustrates the proximity between pairs of calculated data utilizing diverse distance metrics across variables. [0] Represents the difference between the fixed and mobile dryer or between cluster 1 and cluster 2. The following variables are included: [1] Total dryers (unit); [2] Drying time (h); [3] Inlet humidity (%); [4] Outlet humidity (%); and [5] Drying difference between the fixed and mobile systems (%).

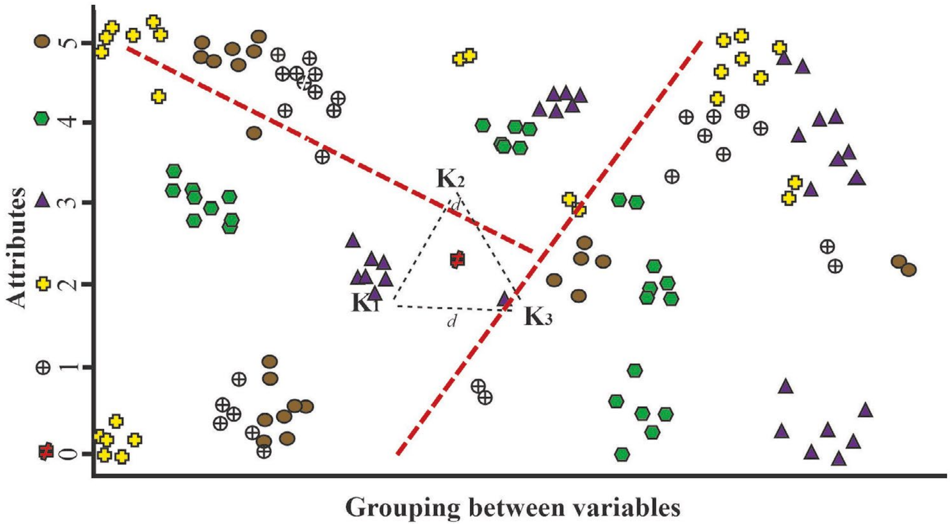

Figure 4: Behavior of data with inter-cluster division in relation to grouping in the fixed (K1 and K2) and mobile (K3) systems. [0] Represents the difference between the fixed and mobile dryers or between cluster 1 and cluster 2. The following variables are included: [1] Total dryers (unit); [2] Drying time (h); [3] Inlet humidity (%); [4] Outlet humidity (%); and [5] Drying difference between the fixed and mobile systems (%). "d" denotes the dimensional distance between the points, while "K" signifies the grouping.

Table 3: Data mining was conducted using the Resample filter and the "Filtered Clusterer" + "K-Means" algorithms to analyze data from both the fixed and mobile systems together. T. dryers represents the total number of dryers required to complete the task, as predicted by the algorithm.

Attribute	Cluster 1	Cluster 2	Cluster 3
System	1 (fixed)	1 (fixed)	2 (mobile)
T. dryers	6.43	3.16	6.11
Duration (h)	01:31	01:47	1:56
Humidity entry (%)	20.32	16.56	19.75
Output humidity (%)	13.23	13.57	14.65
Humidity difference (%)	7.09	2.99	5.10

Despite using similar training parameters, data mining produced different values for the mobile drying system. Consequently, increasing the number of dryers in the mobile system is essential to enhance its efficiency compared to the fixed system. The systems were analyzed separately to discern internal patterns and differences, as the previous disparities between the systems are already known from Tables 2 and 3.

When evaluating the machine learning models solely for the datasets from the fixed dryers (Table 4), the algorithm generated two clusters based on the "total dryers" attribute. Cluster 1 had twice as many dryers as Cluster 2 and exhibited a reduction of 4.27 pp/h in drying rate. In contrast, Cluster 2, with half the number of dryers, experienced a difference of 4.53 pp/h. These findings suggest that doubling the number of dryers does not proportionally increase process efficiency. This analysis is crucial for understanding the efficiency of the production process and the behavior of dryers, as well as for decision-making regarding modular sets or parallel dryers.

The impact of individual stages in the fixed drying system was also examined. The results indicate that the segregation reveals a direct relationship between seed moisture content and the time required to achieve the ideal condition for safe storage (Table 4).

Table 5 presents data from mobile dryers, demonstrating that they are less efficient in achieving the desired drying condition within a short timeframe compared to the fixed system. Cluster 1 and Cluster 2 exhibit drying rates of 3.55 pp/h and 3.02 pp/h, respectively. In this analysis, the most influential attribute is the total number of dryers, indicating that the mobile drying system identifies similarities based on the relationship between the number of dryers. Consequently, the drying efficiency of the mobile system improves with an increased number of dryers, although it requires a longer drying time compared to the fixed system. However, the achieved moisture content does not correspond to the desired temperature (12%), necessitating a longer operating time to achieve the target.

Table 4: Data mining conducted with the fixed system (System 1) using the "Resample" filter. T. dryers represents the algorithm's predicted total number of dryers required to complete the task.

Attribute	Cluster 1	Cluster 2
System	1 (fixed)	1 (fixed)
T. dryers	6.20	3.00
Duration (h)	01:03	02:13
Humidity entry (%)	17.80	22.55
Output humidity (%)	13.40	12.88
Humidity difference (%)	4.40	9.66

Table 5: Data mining performed with system 2 (mobile) using the "Resample" filter. T. dryers is the total number of dryers needed to perform the job, as predicted by the algorithm.

Attribute	Cluster 1	Cluster 2
System	2 (mobile)	2 (mobile)
T. dryers	9.01	4.03
Duration (h)	01:25	02:16
Humidity entry (%)	19.74	19.64
Output humidity (%)	15.30	14.32
Humidity difference (%)	4.44	5.32

Significant differences in initial and final moisture content between clusters indicate that harvesting seeds with higher moisture levels prolongs the drying time. Thus, the dryer's efficiency is more pronounced for crops with moisture levels around 18%. Moreover, the fixed dryer allows for better operational adjustments, ensuring improved drying performance. However, it is essential to analyze other aspects such as energy consumption, gas usage, seed physiological quality, post-drying effects, and other factors to determine a comprehensive cost-benefit ratio.

Upon processing the drying data from various systems, it was crucial to assess the effectiveness of machine learning in unsupervised decision-making. The evaluation of data results was based on the confusion matrix and evaluation metrics (Figure 5 and Table 6) to ascertain the accuracy of the applied model. Groupings were created using the Add Cluster method, and the K-Means algorithm in conjunction with Filtered Clusterer performed further clustering. The software used for cluster classification evaluated the effectiveness of the applied model by constructing a confusion matrix and calculating accuracy metrics. Figure 5 displays the matrices generated through the "Classes to Clusters Validation" option, including the percentage of incorrectly grouped instances.

Subsequently, the confusion matrix was utilized to apply equations that determine the evaluation metrics for assessing the model's learning and efficiency, including recall, precision, and accuracy (Table 6). This enables a more precise evaluation of the model's performance in grouping data for specific stages of the drying process.

The evaluation metrics revealed high values for the clusters, where all parameters surpassed 80% accuracy. This outcome confirms that the unsupervised learning was well-implemented and contributed to the segregation of different types of drying systems, irrespective of the influence of environmental conditions, particularly the water content present during harvest. Recognizing that both drying processes can handle seeds with a high moisture percentage, it becomes necessary to raise the temperature and power of the dryer. Therefore, this type of dryer is recommended for drying grains when there is a need to increase the temperature and flame intensity.

Drying is necessary to handle and store grains and seeds in order to maintain their quality over extended periods and preserve reproductive material from season to season (Moreno et al., 2022). Analyzing and evaluating the efficiency of agricultural equipment for seed drying is crucial for long-term safe storage. It is important to note that the drying equipment should be set at optimal temperatures for the specific cultivars. Viability is the most critical factor to maintain during the drying and subsequent storage of seeds (Maqueda et al., 2018).

The data analysis reveals that the evaluated dryer efficiently reduces water content within a few working hours. Additionally, using multiple dryers in the process helps expedite the drying of products in high demand. Therefore, monitoring water content during harvest is essential. The fixed system dryers demonstrate high efficiency in reducing humidity for the collected cultivars. This finding emphasizes the time required to reduce humidity between the presented systems (fixed and mobile). However, Peske et al. (2019) emphasizes that a decrease in moisture content of 1.2 pp/h can damage seeds. Thus, decisions regarding moisture levels should be made in the field during harvest, considering that harvesting seeds with moisture above 16% increases time, requires more equipment, and compromises the initial seed quality.

Confusion matrix fixed system		Prediction	
		Cluster 1	Cluster 2
Real class	Cluster 1	99	46
	Cluster 2	2	164
Classified incorrectly instances: 15.43% (48)			

Confusion matrix mobile system			
		Cluster 1	Cluster 2
Real class	Cluster 1	586	52
	Cluster 2	52	852
Classified incorrectly instances: 6.743% (104)			

Confusion matrix fixed + mobile system			
		Cluster 1	Cluster 2
Real class	Cluster 1	311	0
	Cluster 2	0	1542
Classified incorrectly instances: 0% (0)			

Confusion matrix fixed + mobile system 3 clusters				
		Cluster 1	Cluster 2	Cluster 3
Real class	Cluster 1	311	0	0
	Cluster 2	0	850	57
	Cluster 3	0	72	563
Classified incorrectly instances: 6.69% (129)				

Figure 5: Confusion matrix showing the clustering prediction results between clusters in the drying system.

Table 6: Evaluation metrics for different approaches to grouping the drying systems.

Drying system	Evaluation metrics (%)			
	Recall	Precision	Accuracy	Clusters
Fixed	0.987	0.974	0.924	1
	1.000	1.000	1.000	2
Mobile	0.987	0.974	0.924	1
	1.000	1.000	1.000	2
Fixed + mobile	1.000	1.000	1.000	1
	1.000	1.000	1.000	2
Fixed + mobile + 3 group	1.000	1.000	0.930	1
	0.937	0.922	0.930	2
	0.887	0.908	0.930	3

The mobile system exhibits lower efficiency in the seed drying process, indicating that the field dryer is susceptible to environmental variations, directly affecting the drying process and resulting in a longer time to reduce water content in seeds. Conversely, the fixed drying system consistently demonstrates superior results, irrespective of the analysis procedures employed. The K-Means algorithm proves to be the most effective in segregating the data, enabling the observation of more dispersed data within the same seed drying system.

The market for drying equipment focuses on achieving the shortest drying time for seeds and grains. However, it is important to consider the cost-effectiveness for producers. Striking a balance between cost reduction and effectiveness is crucial in designing a simple yet highly productive system that minimizes the exposure of the product to high temperatures. Various drying methods, including natural air, infrared, heated air, vacuum, freezing, fluidized bed, and microwave drying, have been employed for seed preservation, particularly for vegetable species. Among these methods, heated air drying is preferred due to its simplicity and cost-effectiveness (Dhurve et al., 2022).

Drying serves as an effective preservation technique by inhibiting the growth of microorganisms, reducing moisture-related degradation reactions (Rashmi; Negi, 2020), and ensuring the maintenance of seed physiological quality. The analysis of data allows us to comprehend the dynamics of unsupervised learning in post-harvest stages, where algorithmic approaches can interact with both computational and human agents (Holzinger, 2016), making it a promising tool to assist seed post-harvest processes.

The algorithm, in conjunction with filters, demonstrates efficiency in unsupervised classification by identifying and minimizing inter-cluster similarities within the fixed system's different classes in the dataset. Cluster analysis, as emphasized by Gersho and Gray (1992), plays a significant role in various applications, such as vector quantization and data compression. The performance of the K-means algorithm can be enhanced by reducing the number of distance calculations involved in finding the cluster center closest to a point (Kumar; Reddy, 2017); thus, grouping the data closer to the centroid, considering that drying data exhibit random points in dimensional space.

The machine learning methods employed in this study exhibited promise in classifying the quality of seed-drying methods. The technique of drying with hot air flow is cost-effective and offers several benefits, such as hygiene, uniform distribution, agility, and the production of high-quality dry products (Onwude; Hashim; Chen, 2016). However, it is important to note that seeds with extremely low water content are more vulnerable to mechanical damage (breakage and cracks) during harvesting and processing operations (Medeiros et al., 2020a).

Other studies, such as Gadotti et al. (2022a, 2022b), Medeiros et al. (2020b), Pinheiro et al. (2022, 2021), have also highlighted the use of artificial intelligence in the seed sector. Producers require decision-making tools for allocating higher-quality seeds, as mandated by Brazilian seed legislation. Therefore, comprehending factors that facilitate proper allocation, speed, and cost-effectiveness poses a challenge in the seed sector. In this regard, the use of machine learning as an aid in decision-making becomes fundamental for the industry.

According to André et al. (2022) and Gadotti et al. (2022b), the utilization of artificial intelligence techniques for data modeling offers a promising alternative to address gaps where conventional statistics fail to produce satisfactory forecasting outcomes for seed quality assessment. Furthermore, analyzing different drying methods can aid in selecting appropriate equipment that yields satisfactory results and ensures secure storage of the final product, thereby extending its quality and lifespan.

The findings indicate that the fixed system is more suitable for expediting the drying process, as it can eliminate a larger amount of water in a shorter duration compared to the mobile system. However, the mobility of the dryer facilitates equipment installation on the premises, as it can be conveniently loaded onto a trailer.

The employment of machine learning techniques, through an unsupervised process, has enabled informed decision-making by identifying avenues to reduce drying time in both types of dryers. Nevertheless, certain adjustments need to be addressed and resolved in seed drying, particularly with regard to the exposure time of seeds to high temperatures. Technical issues can be rectified during equipment manufacturing, especially when dealing with mobile dryers. These dryers are not recommended for seed drying, as seeds are living organisms that require careful handling during the drying process, as an increase in process temperature may adversely impact their viability.

Improving the performance of the drying equipment, such as the fan curve, pressure and flow rate, and gas consumption, can enhance its efficiency as an effective seed drying apparatus. However, the benefits observed in both drying systems are only applicable to grain drying in practical applications. The outcomes are solely based on the provided data and the specific dryer model used, and the conditions may not be applicable to other manufacturers.

CONCLUSIONS

The unsupervised processing model exhibited a remarkable level of data accuracy, enabling the algorithm to acquire knowledge about the rules governing drying processes in different dryers within fixed and mobile systems. Moreover, this approach effectively demonstrates the algorithm's ability to classify the dataset and identify that drying time increases with water content and the installation environment of the equipment.

AUTHOR CONTRIBUTION

Conceptual idea: Pinheiro, R.M.; Gadotti, G. I.; Methodology design: Pinheiro, R.M.; Gadotti, G. I.; Bernardy, R.; Data collection: Pinheiro, R.M.; Bernardy, B.; Tim, R. R.; Pinto, K. V. A.; Buck, G. Data analysis and interpretation: Pinheiro, R. M.; Bernardy, B., and Writing and editing: Pinheiro, R. M.; Tim, R. R.; Pinto, K. V. A.; Buck, G.

ACKNOWLEDGMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001". And we would like to thank the company Boom do Brasil Secadores de Grãos for making the data available and following up on some drying stages.

REFERENCES

ANDRÉ, G. S. et al. Predicting the quality of soybean seeds stored in different environments and packaging using machine learning. Scientific Reports, 12:8793, 2022.

ARUMUGAM, K. et al. Towards applicability of machine learning techniques in agriculture and energy sector. Materials Today: Proceedings, 51(8):2260-2263, 2022.

DHURVE, P. et al. Drying kinetics, mass transfer parameters, and specific energy consumption analysis of watermelon seeds dried using the convective dryer. Materials Today: Proceedings, 59:926-932, 2022.

GADOTTI, G. I. et al. Machine learning for soybean seeds lots classification. Engenharia Agrícola, 42:e20210101, 2022a.

GADOTTI, G. I. et al. Prediction of ranking of lots of corn seeds by artificial intelligence. Engenharia Agrícola, 42:e20210005, 2022b.

GARCIA, D. C. et al. A secagem de sementes. Ciência Rural, 34(2):603-608, 2004.

GERSHO, A.; GRAY, R. M. Vector quantization and signal compression. Kluwer Academic, Boston. The Springer International Series in Engineering and Computer Science. 1992. 737p.

HOLZINGER, A. Interactive machine learning for health informatics: When do we need the human-in-the-loop? Brain Informatics, 3:119-131, 2016.

KUMAR, K. M.; REDDY, A. R. M. An efficient k-means clustering filtering algorithm using density based initial cluster centers. Information Sciences, 418-419:286-301, 2017.

MAQUEDA, R. H. et al. Microwave drying of amaranth and quinoa seeds: Effects of the power density on the drying time, germination rate and seedling vigour. Journal of Microwave Power and Electromagnetic Energy, 52(4):299-311, 2018.

MEDEIROS, A. D. et al. Interactive machine learning for soybean seed and seedling quality classification. Scientific Reports, 10:11267, 2020a.

MEDEIROS, A. D. et al. IJCropSeed: An open-access tool for high- throughput analysis of crop seed radiographs. Computers and Electronics in Agriculture, 175:105555, 2020b.

MORENO, Á. H. et al. Effect of temperature on the microwave drying process and the viability of amaranth seeds. Biosystems Engineering, 215:49-66, 2022.

ONUKWULI, D. O. et al. Comparative analysis of the application of artificial neural network-genetic algorithm and response surface methods-desirability function for predicting the optimal conditions for biodiesel synthesis from chrysophyllum albidum seed oil. Journal of the Taiwan Institute of Chemical Engineers, 125:153-167, 2021.

ONWUDE, D. I.; HASHIM, N.; CHEN, G. Recent advances of novel thermal combined hot air drying of agricultural crops. Trends in Food Science & Technology, 57:132-145, 2016.

PESKE, S. T.; VILLELA, F. A.; GADOTTI, G. I. Secagem de sementes. In: PESKE, S. T.; VILLELA, F. A.; MENEGHELLO, G. E. (Org.). Sementes: Fundamentos científicos e tecnológicos. 4 ed. Pelotas: Becker e Peske, p.355-405, 2019.

PINHEIRO, R. M. et al. Inteligência artificial na agricultura com aplicabilidade no setor sementeiro. Diversitas Journal, 6(3):2996-3012, 2021.

PINHEIRO, R. M. et al. Processamento de imagens como ferramenta importante para inteligência artificial no setor de sementes. Revista Agrária Acadêmica, 5(1):89-101, 2022.

RASHMI, H. B.; NEGI, P. S. Phenolic acids from vegetables: A review on processing stability and health benefits. Food Research International, 136:109298, 2020.

SHARMA, A. et al. Enabling smart agriculture by implementing artificial intelligence and embedded sensing. Computers & Industrial Engineering, 165:107936, 2022.

WANG, Y. et al. Fundamentals, materials and machine learning of polymer electrolyte membrane fuel cell technology. Energy and AI, 1:100014, 2020.

YAO, H. et al. An improved K-means clustering algorithm for fish image segmentation. Mathematical and Computer Modelling, 58(3-4):790-798, 2013.