Amaral, T. G; Maciel, C. C. M, Roriz Júnior, M. P & Paula, G. S. (2025). Automated Method for Classifying Wastes by Making-Do Using Machine Learning. In O. Seppänen, L. Koskela, & K. Murata (Eds.), *Proceedings of the 33rd Annual Conference of the International Group for Lean Construction* (IGLC33) (pp. 822-833). https://doi.org/10.24928/2025/0274

METHOD FOR CLASSIFYING WASTES BY MAKING-DO USING MACHINE LEARNING

Tatiana Gondim do Amaral[1], Caio César Medeiros Maciel[2], Marcos Paulino Roriz Junior[3], and Gabriella Soares de Paula[4]

ABSTRACT

The study aims to develop an automated method for classifying making-do wastes using machine learning (ML) techniques. Manual classification of these wastes is prone to inconsistencies, especially in projects with large volumes of data. The automated method makes the process more efficient and accurate. The research is classified as quantitative and empirical, with a descriptive, exploratory and experimental approach. Data was collected using the Melius quality platform and covered six high-end multi-family residential developments from two construction companies in Goiânia. The data was processed in two phases: initially, compliance was checked manually and then the data was adjusted for the ml algorithm. Preliminary results indicate that the main causes of waste are related to lack of labor (67.11%) and problems with materials (15.48%). The highest incidences of waste categories were inadequate, sequencing (26.82%) and lack of equipment (18.21%). In terms of impact, the recurrence of rework (13.37%) and lack of terminality (13.19%) stand out. The neural network model showed unsatisfactory results, with a recall of 55.4% and precision of 53.7%. The study shows the potential of machine learning, but adjustments to the models are necessary to improve their effectiveness.

KEYWORDS

Making-do waste, Machine Learning, Predictive models, Automatic classification, Neural Network.

INTRODUCTION

Based on the initial definition of the Toyota Production System, which includes seven types of wastes (Ohno, 1997), Koskela (2004) introduced a new type of waste: making-do, characterized by the execution of tasks without the complete availability of the necessary resources. Koskela argues that this waste is particularly critical in construction, as its effects are not limited to the task in question but can trigger a series of other wastes throughout the process.

When they are detected, making-do wastes are typically recorded as text fields, which makes them difficult to process. In fact, several authors have reported difficulties in identifying and classifying these wastes Formoso, Sommer, Koskela and Isatto (2011); Fireman and Formoso (2013); Leão, Isatto and Formoso (2016); Amaral, Braga and Barros (2020). The analysis and

[1] Full Professor, Environmental and Civil Engineering Department, Federal University of Goiás (UFG), Goiânia, GO, Brazil. tatianagondim@ufg.br, orcid.org/0000-0002-9746-4025

[2] Master Student, Program in Production Engineering, Federal University of Goiás, Brazil, caiocesar.eng@hotmail.com, orcid.org/0000-0002-6756-4068

[3] Adjunct Professor, Faculty of Science and Technology, Federal University of Goiás (UFG), Aparecida de Goiânia, GO, Brazil, marcosroriz@ufg.br, orcid.org/0000-0003-2795-0009

[4] Student, School of Environmental and Civil Engineering, Federal University of Goiás, Brazil, gabriella.soares@discente.ufg.br, orcid.org/0000-0002-7730-2838

classification of this data largely depends on human intervention, which makes the process laborious and subject to interpretation. The need to interpret textual descriptions and organize different types of information into categories requires objectivity and continuous effort from researchers. As they are typically handled manually, the analysis can lead to inconsistencies or limitations, especially in large-scale studies where the amount of data is significant. These difficulties highlight the need to improve the methods employed and develop more assertive and comprehensive analyses.

At the same time, the popularization of Artificial Intelligence techniques such as machine learning (ML) has helped in the processing of data in various areas. This is because ML enables computer systems to learn to identify data without the need for detailed programming for each specific situation, i.e. without the limits of these classifications being previously defined (Flach, 2012). This makes it possible to speed up and automate the classification process.

However, making-do data is typically textual, which makes it difficult to process. Based on this, this work aims to investigate the feasibility of applying machine learning methods to classify making-do wastes from text-based data in Brazilian Portuguese. Specifically, a dataset of non-compliance tasks from two construction companies was analyzed to identify the class of waste.

LITERATURE REVIEW

THE CONCEPT OF MAKING

The concept of making-do in construction was introduced by Koskela (2004) to describe a waste that occurs when a task is started without all the necessary resources or continues to be carried out even though some of them are missing. Sommer (2010) defines making-do as a form of improvisation, in which an activity is carried out with what is available. According to the author, in practice, workers often deal with a lack of resources, materials or adequate information, trying to complete their activities without causing major damage.

This category of waste was inspired by the work of Ronen (1992), who proposed the creation of the Complete Kit, in which it was pointed out that a task should not be started unless all the resources needed to complete it are available. Within this approach there are two types of complete kit: input and output. The input kit of a task must match the output kit of the previous task.

To understand the occurrence of these wastes, the studies by Koskela (2000), Sommer (2010) and Fireman (2012) identified missing prerequisites, categories of wastes and impacts related to the making-do concept. Koskela (2000) highlighted seven prerequisites, while Sommer (2010) expanded this analysis, resulting in eight prerequisites. Fireman (2012) expanded on these findings and defined eight waste categories and seven impact items. Table 1 summarizes the categorization carried out by the authors.

The waste identification method proposed by Sommer (2010), and then complemented by Fireman and Formoso (2013) and Santos and Santos (2017), presents three groups. The first is used to identify missing prerequisites in work packages, the second is to identify the most affected waste categories, and the last assesses the impacts of waste. It is noteworthy that making-do wastes can occur in different ways, and there are numerous possible combinations of prerequisites, categories, and impacts in the construction environment.

Table 1: Parameters analyzed for making-do waste (Table 1 adapted from Sommer (2010); Fireman and Formoso (2013) and Santos and Santos (2017))

Prerequisite	Categories	Impacts
Information	Access/mobility	Decreased productivity
Materials and component	Component adjustments	Demotivation
Labor	Workspace	Material Wastes
Equipment or Tools	Storage	Rework
Space	Equipment/Tools	Safety reduction
Interconnected services	Temporary installations	Quality reduction
External conditions	Protection (security)	Lack of Terminality
Facilities: workspace infrastructure	Sequencing	

Some authors have highlighted that improvisations are present in all of the construction site stages, thereby making it difficult to identify and avoid them, and requiring strict control of construction processes, investments in cultural change conducive to improvisation and standardization (Amaral et al., 2023; Santos, Fontenele, Machado, Neto and Amaral, 2020; Josephson and Hammarlund, 1999; Horman & Kenley, 2005; Formoso, Sommer, Koskela & Isatto, 2017; Ohno, 1997).

Moreover, some authors reported difficulties in identifying and classifying making-do wastes, pointing to the need to improve the methods used. In addition, the need to develop more quantitative analyses and the acceptable limits of making-do wastes is also highlighted in the literature by several authors (Saurin & Sanches, 2014; Amaral et al., 2019; Santos, Fontenele, Machado, Neto and Amaral, 2020; Josephson & Hmmarlund, 1999; Horman & Kenley, 2005; Formoso et al., 2015; Formoso et al., 2017; Leão, Formoso & Isatto, 2014; Fireman & Formoso, 2013; Kalsaas, 2012).

Making-do waste data is typically textual. In this sense, it requires an interpretative effort on the part of the experts to classify a given waste. In addition, there can be conflicts, i.e. an author determining that a combination of missing prerequisites and waste categories leads to different impacts. To mitigate these shortcomings and speed up the process, the use of machine learning techniques was investigated.

MACHINE LEARNING

Machine learning (ML) is a branch of artificial intelligence that seeks to enable computer systems to recognize and classify data autonomously, without the need for explicit programming that defines the rules or limits. This approach allows systems to learn from the data they receive (Flach, 2012).

The technique works by analyzing a set of data, providing examples of input (data that the system must process) and output (the expected classifications). The system uses these examples to learn the function that relates the inputs to the outputs. Thus, once trained, the system can make predictions or classifications on new data that has not been seen before, applying the knowledge acquired during the learning process.

ML encompasses a variety of methods that can be classified into different categories. Among the classic methods, Linear Regression and Logistic Regression stand out (Amaral, 2024). On the other hand, more advanced methods, such as Deep Neural Networks and Support Vector Machines (SVM), offer greater complexity and the ability to deal with non-linear data. Deep Neural Networks are made up of multiple layers of neurons that allow the model to learn complex representations of the data, making them effective in tasks such as image recognition and natural language processing. Support Vector Machines, on the other hand, work by separating data into different classes using hyperplanes, and are especially useful in high-dimensional data sets.

The main practical difference between these methods lies in the algorithms used to learn the function that relates the inputs to the outputs. This choice of algorithm directly affects the accuracy of the forecasts, as some algorithms are better suited to certain characteristics of the data than others (Zhang et al., 2017).

In addition, the ability of methods to explain the parameters that lead to a forecast varies considerably. Some models, such as Linear Regression, offer clear and straightforward interpretations of the coefficients, allowing us to understand how each variable influences the forecast. In contrast, more complex models, such as Deep Neural Networks, can be considered “black boxes”, where the relationship between inputs and outputs is difficult to interpret (Gonzalez, 2019). This difference is crucial when selecting a method, depending on the need for explainability versus predictive accuracy.

For this work, we chose to explore the Orange Platform (Janez et al., 2013) due to its effectiveness and ease of parameterization. It also includes various methods for creating machine learning models. The tool offers several algorithms, such as K-Nearest Neighbors (KNN), Decision Tree, Random Forest, Support Vector Machine (SVM), Neural Networks and Stacking (which combines several models). Figure 1 shows the screen of the software used to train the models.

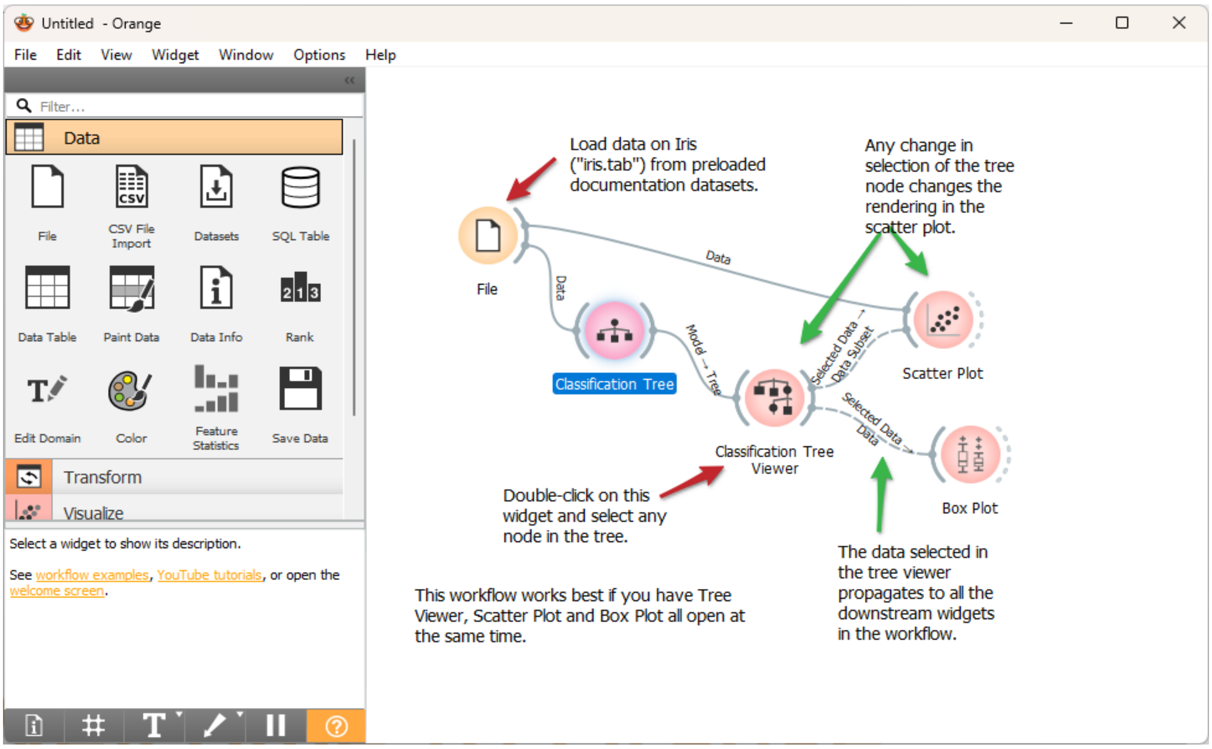

Figure 1: Screenshot of the Orange software (data taken from the Orange Program)

METHOD

RESEARCH CLASSIFICATION

The research adopts a quantitative and empirical methodology, with a descriptive approach, incorporating both exploratory and experimental aspects. The main objective is to investigate phenomena that have not yet been widely studied or understood, using experimental techniques that allow cause and effect relationships to be analyzed. This type of approach makes it possible not only to raise initial hypotheses about the subject, but also to manipulate data to examine their interactions and impacts.

The study was conducted in four distinct stages: 1) Data collection, 2) Data processing, 3) Definition of the machine learning models and 4) Testing and evaluation with the classified database, which will be detailed below.

DATA COLLECTION

A total of 1,137 pieces of data were obtained through the *Melius Qualidade* service and materials management platform, covering information on six projects by two construction companies (marked anonymously as C and S) located in Goiânia/GO/Brazil. Both companies have been in the vertical residential construction business for over 20 years.

To increase the database, it was decided to collect data from work in the final stage. Preliminary data was included at this stage to carry out an exploratory analysis of the results, with the aim of evaluating the initial behavior of the model and identifying possible adjustments before the full collection was completed. Therefore, the results presented should be interpreted as preliminary indications and are subject to refinement as new data is incorporated.

The projects are at similar levels of progress, but with different amounts of data collected (Table 2). The projects that have already been completed have 100% progress have less data collected because of the end of the availability of data from completed projects.

Table 2: Characterization of the enterprises

Project	Construction phase	General progress	Amount of data collected
C-E2	Installations (95%) Finishes (96%)	98%	264
C-E3	Walls and seals (96%) Plaster and subfloor (78%) Installations (88%) and finishes (92%)	87%	413
C-E4	Project completed	100%	342
S-E1	Project completed	100%	6
S-E2	Project completed	100%	66
S-E3	Finishing (97%)	100%	46

DATA PROCESSING

Data processing was carried out in two distinct stages. The first stage involved manual procedures, conducted by the researcher, in which the data is collected, organized and checked to ensure it complies with the study's requirements. During this phase, the researcher filters the information, corrects inconsistencies and eliminates irrelevant data that could interfere with the analysis.

In the second stage, the data was processed and prepared to suit the specific requirements of the ML algorithm. This includes normalizing, categorizing and transforming the data into suitable formats so that the ML model can classify it efficiently. These adaptations are essential to ensure that the algorithm understands the information correctly and can identify patterns accurately, making it easier to classify wastes by making-do.

DATA CLASSIFICATION

Table 3 shows the detailed organization of the data extracted from the Melius Quality platform, followed by its classification using Microsoft Excel software. The first section lists the Service, Stage, Location, Status, Criteria, Problem and Team fields, which provide an overview of the activities and conditions of the services. Next, the information is processed and grouped, including details of Stage, Sub-stage, Activities, Prerequisite, Category and Impact. The stage and sub-stage parameters were organized according to the guidelines established by NBR 12721 (ABNT, 2006).

Table 3: Database model

Metadata extracted for them Melius Quality Platform	Fields used to build the dataset
Service	Team
Stage	Stage
Location	Sub-stage
Status	Activities
Criteria	Prerequisite
Problem	Category
	Impact

MACHINE LEARNING CLASSIFICATION

Table 4 shows the Characteristic/Description column, which details the specific characteristics and descriptions of a task or activity. The Problem column documents errors or failures observed, identifying deviations from what was planned. The Action column contains immediate actions to correct the problems found, mitigating the impact of the failure. Finally, the merge column is added, where the previous three columns are combined, presenting a consolidated version of all the information about the item. This column is useful for ML processes, as it contains a complete and self-explanatory summary of each occurrence analyzed.

Table 4: Input data examples

Description	Problem	Containment Action	Merge
Check the finishing of the faces and alignment of the sheet, portals and planks, ensuring that the joints between the plank and the wall are flush.	The tape on the door frames is coming off due to a defect in the material itself.	The manufacturer will cover the costs and replace the defective items promptly.	Check the finish of the faces and alignment of the sheet, portals and planks, ensuring that the joints between the plank and the wall are flush. The tape on the door frames is coming off due to a defect in the material itself. The manufacturer will cover the costs and replace the defective items promptly.

DEFINITION OF MACHINE LEARNING MODELS

Based on a review of the literature and an exploration of its resources, the phases of the process and the evaluation of the methods were carried out using Orange, a free and open-source platform dedicated to ML and data visualization. This software allows graphical representations to be created in a practical and intuitive way, as well as making it possible to train, validate and compare different ML algorithms. Orange operates with a visual structure based on graphic modules called widgets. Each widget performs a specific function, has its own inputs and outputs and can be linked to other widgets to perform tasks such as data processing and visualization (Janez et al., 2013).

The Orange workflow applied in this research begins with importing the database using the Data Table widget (Figure 2). Next, the relevant columns are selected using the Select Columns widget. To ensure an adequate sample, the Data Sampler widget divides the data set into two parts: the first, with 80% of the data, is used to train the model, while the second, with the remaining 20%, is used to test and evaluate the performance of the models. The Corpus widget is then used to process textual data, quantifying how many times each word is repeated in each text. The data is then processed and cleaned, removing frequent words that don't carry meaning (stop words), such as “and”, “of”, “for” and “or”. The texts are then converted into data vectors using the Document Embedding widget. After processing, the data is presented using the Data Table.

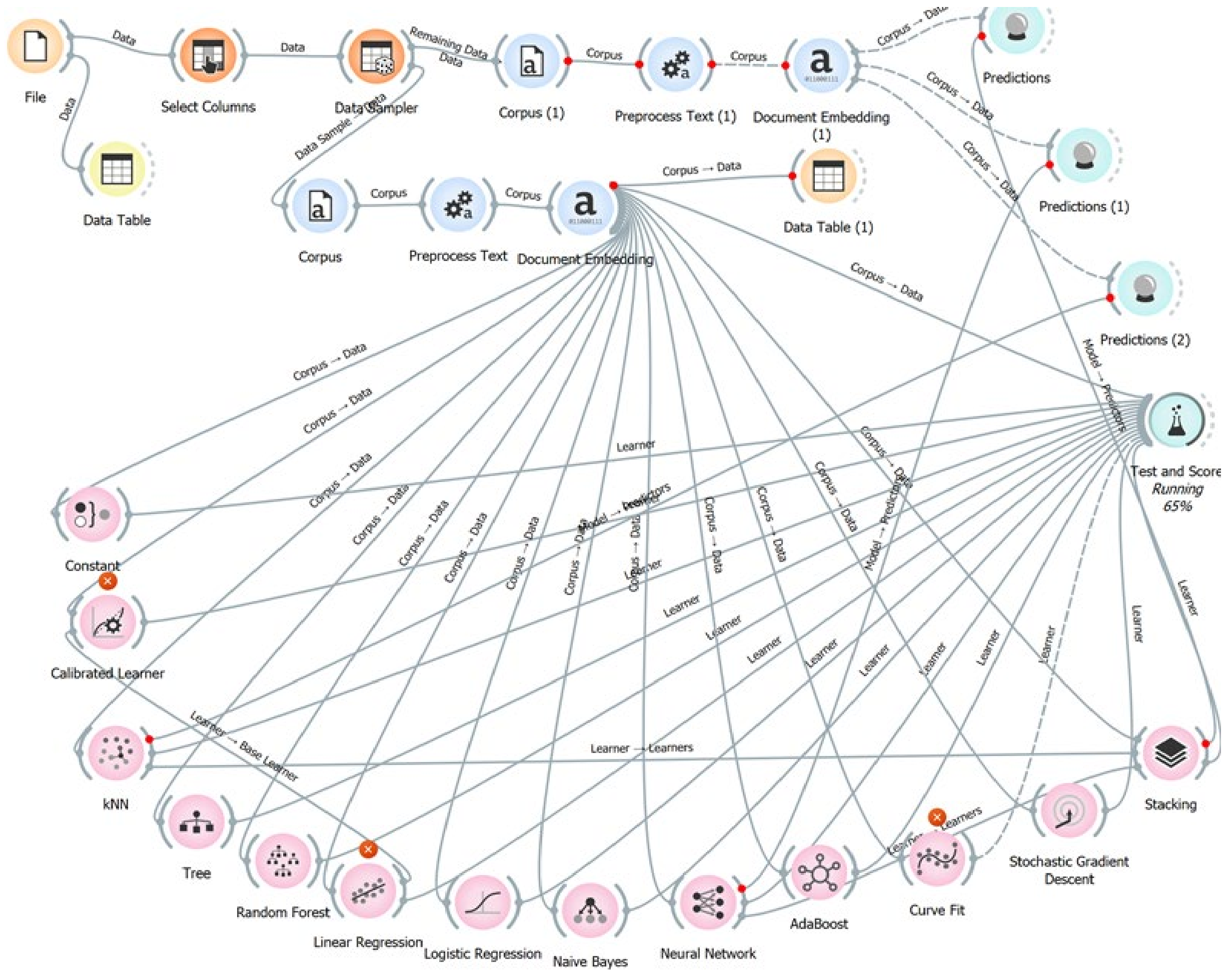

Figure 2: Method’s workflow in Orange

For this study, it was decided to use the 10 models that were compatible with the data format: K-Nearest Neighbors (KNN), Naive Bayes, Random Forest, Decision Tree, Logistic Regression, Neural Network, AdaBoost, Stochastic Gradient Descent (SGD), Stacking and Constant. The ML models are configured and evaluated using the Test and Score widget, which performs cross-validation and compares performances.

Finally, the Predictions widgets are used to generate predictions from the trained models. They take as input both the processed data and the trained models, returning the predicted classifications for each instance. In addition, the Predictions widgets can display these predictions in an organized manner, allowing the researcher to compare the results of different models directly. In this way, it is possible to visualize and analyze the classifications made for the test data set, providing a clear understanding of each model's individual performance.

TESTING AND EVALUATION WITH THE CLASSIFIED DATABASE

The models with the best performance were applied to make predictions on the 20% of the data set aside for testing, using the Prediction tool. The models can then be evaluated using the performance parameters of the ML models. Orange provides seven types of evaluation parameters, such as Area Under the Curve (ROC), Accuracy, Precision, Recall, F1 and Matthews Correlation Coefficient (MCC).

In this study, the precision and recall parameters were chosen to evaluate the models. The choice is due to the relevance of these criteria in scenarios where it is important both to minimize false positives and to maximize true positives.

Precision is calculated by dividing the number of true positive detected (TP) by the total number of positive predictions, including wrong ones, that is true positives and false positives: $Precision = \frac{TP}{TP+FP}$. On the other hand, recall is calculated by the proportion of true positive cases detected (TP) in relation to the total number of true positive cases, even those that were not detected (false negatives), and is expressed as: $Recall = \frac{TP}{TP+FN}$.

Recall measures the model's ability to correctly identify positive instances, i.e. how many of the actual occurrences were predicted as such. Meanwhile, precision measures of how many of the positive predictions made by the model are correct. This is important to ensure that positive classifications have a high degree of reliability, avoiding false positives. Combining these two factors allows for a balanced analysis, focusing both on detecting as many true positives as possible and ensuring that positive classifications are reliable.

RESULTS AND DISCUSSIONS

The previously classified data was imported into Microsoft Power BI to generate interactive graphical analyses. To visualize the data, three chart models were selected: hierarchical tree, funnel and ranges charts, to provide a clear and efficient visual interpretation of the information analyzed.

The left-hand side of Figure 3 represents the hierarchical tree or decomposition tree graph that classifies the making-do wastes into three levels: prerequisite, category and impact.

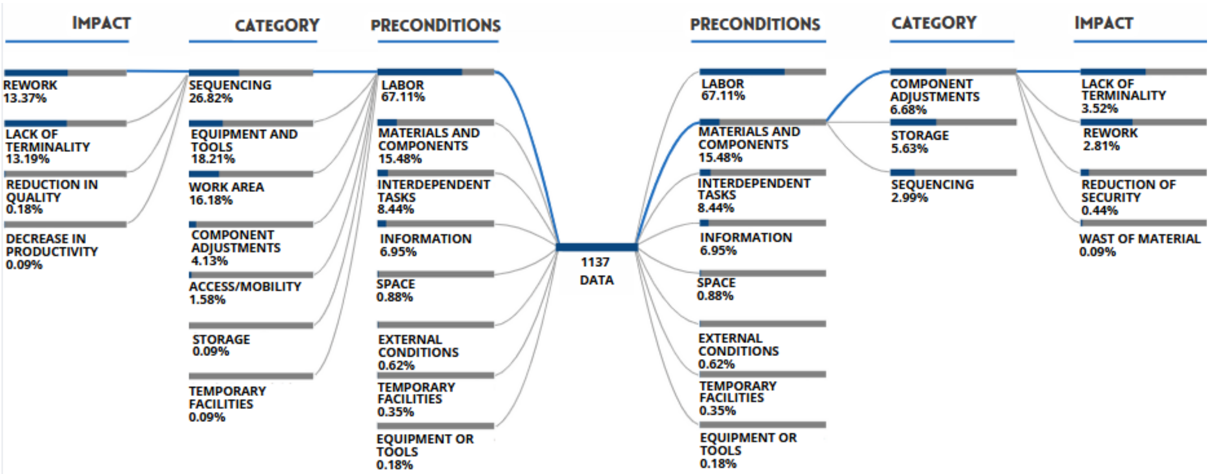

Figure 3: Diagram of the relations between wastes per prerequisite, category, and impacts

The results show that waste is mainly linked to a lack of suitable labor (67.11%) and problems with materials and components (15.48%), as well as failures in interdependent tasks (8.44%) and a lack of information (6.95%). These wastes are grouped into categories, with the most serious being inadequate sequencing of activities (26.82%), lack of equipment and tools (18.21%) and problems with the work area (16.18%), as well as component adjustments and limited accessibility (1.58%). Finally, at the impact level, the main consequences identified were rework (13.37%) and lack of terminality (13.19%), followed by a small proportion of reduced quality (0.18%) and reduced productivity (0.09%).

Analyzing at the right-hand side of Figure 3, which starts with the second largest prerequisite, materials and components (15.48%), this is related to three main categories. The first is component adjustments, which account for 6.68% of waste. This means that component adjustments are often necessary due to a lack of suitable materials or components. The second category, storage (5.63%), indicates that poor management or consumed availability of materials can lead to delays or problems in the progress of services. Finally, sequencing comes in at 2.99%, indicating that a lack of materials has an impact on the order and flow of tasks in the schedule. Thus, these three categories result in different negative impacts.

The biggest impact identified is lack of terminality (3.52%), showing that, in many cases, tasks are not completed mainly due to problems with materials or components. Rework (2.81%) emerges as the second biggest impact, reinforcing the idea that, due to inadequate or unavailable materials, corrections must be made, or stages repeated. There is also a level of reduction in safety (0.44%), which, although less significant, points out that a lack of materials can compromise safety in the process. Finally, the impact of material waste (0.09%) is small but still relevant, reducing waste associated with inefficient material management.

Figure 4 shows that the labor prerequisite has the highest number of impacts, with the impact of rework standing out, accounting for 32.81% of occurrences.

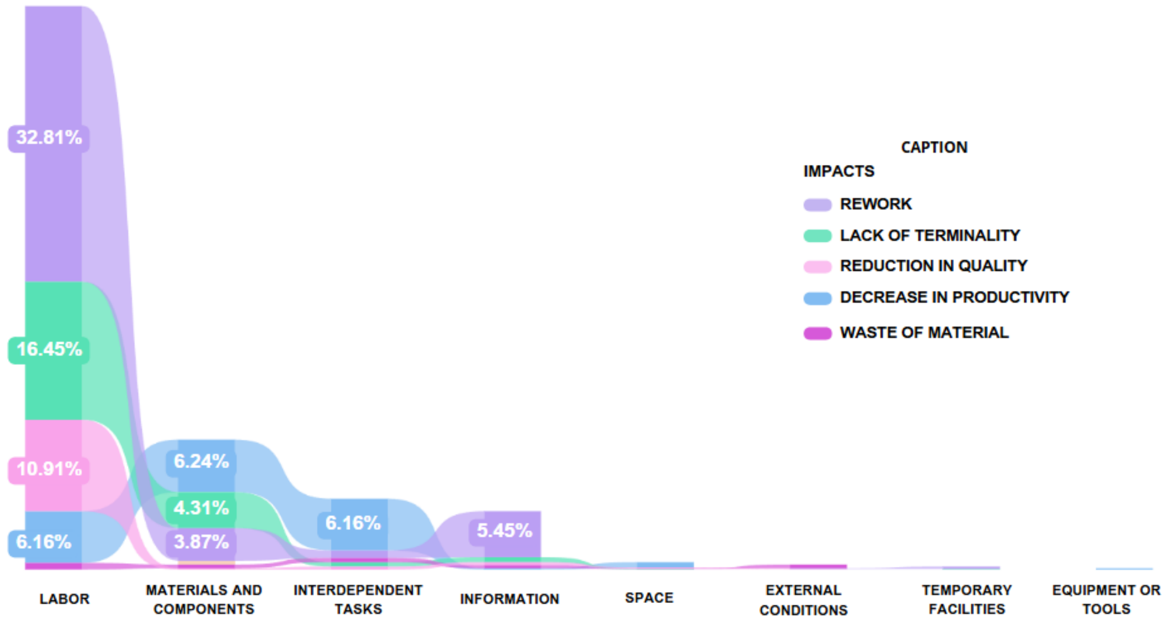

Figure 4: Correlation of wastes between prerequisites and impacts

This high rate suggests that labor-related problems are one of the main causes of rework, which may be associated with the need to correct errors or operational failures. In addition, within this same prerequisite, there are also significant percentages of impact from lack of terminality (16.45%) and reduced quality (10.91%), indicating that labor is a critical factor that directly influences the efficiency and quality of projects. The prerequisite of materials and components also has a significant impact, especially in terms of reduced productivity (6.24%) and lack of terminality (4.31%). These figures indicate that problems with materials or components can cause delays or interrupt the continuity of activities, negatively affecting productivity.

The data was processed using the neural network model to analyze the classification of the database using ML. For the experiment, the prerequisite column was used to be classified, the results of which are shown in Table 5.

Table 5: Preliminary evaluation of data by the Neural Network Model

Model	AUC	CA	F1	Prec	Recall	MCC
Neural Network	60,5%	55,4%	54,1%	53,7%	55,4%	8,8%

The low recall (55.4%) and precision (53.7%) values obtained by the neural network model suggest that it was not able to specifically generalize the learning. This may be related to two possible problems. The first is underfitting, which occurs when the model does not have sufficient capacity to capture the patterns of the training data, resulting in unsatisfactory performance. In this case, the model is unable to represent the complexity of the data, compromising its effectiveness both in correctly identifying positive examples and in preventing false positives.

On the other hand, overfitting could occur if the model presented high accuracy on the training data, but failed when tested on unknown data, or indicated excessive specialization in the training data. This behavior may be even more evident when a new, previously unseen database with different forms of writing is used for classification. In this scenario, the model, instead of generalizing to encompass and classifying new data appropriately, specializes in the data provided during training and finds it difficult to deal with new data that present format variations. This explains the low recall and precision values, since the model is not prepared to correctly identify the new structures that it did not encounter during training.

To correct the identified error, it is assumed that data collection must be fully completed before training the model. Training the model with a complete database ensures that it has access to all the variations and writing forms presented, preventing learning from being limited to an incomplete or outdated subset of data. This allows the model to develop better generalization capacity, which can improve performance in classifying the data obtained and reduce the risk of overfitting. Furthermore, when trained with all the available data, the model will have a more robust representation of the patterns and will be more efficient in dealing with the variations presented in the database.

CONCLUSIONS

By demonstrating the feasibility and accuracy of machine learning in classifying making-do waste, this study contributes to the advancement of data-driven decision-making in construction management. The integration of machine learning into waste classification processes presents a promising avenue for increasing operational efficiency, reducing costs, and improving overall project quality. The results reinforce the need for continued exploration of machine learning applications in construction, paving the way for more innovative and automated methodologies in the construction industry.

This study highlights the feasibility and accuracy of machine learning in classifying making-do waste, contributing to the advancement of data-driven decision-making in construction management. The integration of machine learning into loss classification processes presents a promising opportunity to enhance operational efficiency, reduce costs, and improve the overall quality of waste analysis. The results underscore the importance of continuing to explore machine learning applications in construction, paving the way for more innovative and automated methodologies in the sector.

Beyond improving classification efficiency, this study underscores the broader applicability of machine learning in the construction industry. Future research could explore its use in image-based waste detection on construction sites, further automating the identification and classification of wastes. Additionally, applying these techniques to different types and scales of construction projects could extend the generalizability of the findings, allowing for a more comprehensive approach to waste management.

REFERENCES

ABNT. NBR 12721. (2006). Incorporações imobiliárias - avaliação de custo unitário de construção. *Associação Brasileira de Normas Técnicas*. 1(2), 49-54.

Amaral, T. G. do, Brandão, C. M., Elias, K. V., & Braga, P. B. (2019). Identificação De Perdas Por Improvisação Em Canteiros De Obras. *REEC - Revista Eletrônica De Engenharia Civil*, *15*(1), 245–260. doi.org/10.5216/reec.v15i1.54562

Amaral, T.G., Braga, P. B., Barros Neto, J.P. (2020). Application of Dynamic Spreadsheets in the Analysis of Waste by Making-do. *Proceedings of the 28th Annual Conference of the International Group for Lean Construction (IGLC28*) doi.org/10.24928/2020/0077

Amaral, T. G., Maciel, C. C. M., Filho, R. R. D. G., Pessoni, R. C. S., Paula, G. S. & Silva, S. V. (2023). Results of the Causes and Impacts of Making-Do Wastes in Production in Fortaleza, Ceará, Brazil, *Proceedings of the 31st Annual Conference of the International Group for Lean Construction (IGLC31)*, 1395-1406. doi.org/10.24928/2023/0250

Amaral, T. G., Maciel, C. M., Roriz, M. J., & Paula, G. S. (2024). Modelo de classificação das perdas por making-do com uso de aprendizado de máquina. *Encontro Nacional de Tecnologia do Ambiente Construído. 20(1), 1-13*. 10.46421/entac.v20i1.5732

Braga, A. de P., Ludermir, T. B., & Carvalho, A. C. P. de L. F. de. (2000). *Redes neurais artificiais: teoria e aplicações*. Rio de Janeiro: LTC.

Fireman, M. C. T. (2012). *Proposta de método de controle integrado entre produção e qualidade com mensuração de perdas por making-do e pacotes informais.* Dissertação (Pós-graduação em Engenharia Civil) – Programa de Pós-Graduação em Engenharia Civil, Universidade Federal do Rio Grande do Sul, Porto Alegre.

Fireman, M. C. T. & Formoso, C. T. (2013). Integrating Production and Quality Control: Monitoring Making-Do and Unfinished Work, *Proceedings of the 21th Annual Conference of the International Group for Lean Construction (IGLC21)*.

Flach, P. (2012) Machine learning: The art and science of algorithms that make sense of data. *Cambridge University Press*.

Formoso, C.T., Sommer, L., Koskela, L., & Isatto, E. L. (2011). An Exploratory Study on the Measurement and Analysis of Making-do in Construction. *Proceedings of the 19th Annual Conference of the International Group for Lean Construction (IGLC19).*

Formoso, C., Bølviken, T., Rooke, J. & Koskela, L. (2015). A Conceptual Framework for the Prescriptive Causal Analysis of Construction Waste, *Proceedings of the 23rd Annual Conference of the International Group for Lean Construction* (IGLC 23). https://iglc.net/Papers/Details/1162

Formoso, C. T., Sommer, L., Koskela, L., & Isatto, E. L. (2017). The identification and analysis of making-do waste: insights from two Brazilian construction sites. *Ambiente Construído*, v. 17, n. 3, p. 183-197, jul/set. doi.org/10.1590/s1678-86212017000300170

Gonzalez, L. (2018). Regressão Logística e suas Aplicações. Monografia (Especialização). *Curso de Ciência da Computação, Universidade Federal do Maranhão, São Luis*. 45p.

Horman, M. J., & Kenley, R. (2005). Quantifying Levels of Wasted Time in Construction with Meta-Analysis. *Journal of Construction Engineering and Management*, Vol. 131, Issue 1. 10.1061/(ASCE)0733-9364(2005)131:1(52)

Janez, D., Tomaz, C., Ales, E. Crt., G., Tomaz, H., Mitar M., Martin, M., Matija, P., Marko, T., Anze S., Miha S., Lan U., Lan Z., Jure Z., Marinka Z., & Blaz Z. (2013). Orange: Data mining toolbox in python. *The Journal of Machine Learning Research,* (14), 2349–2353.

https://www.researchgate.net/publication/285018297_Orange_Data_Mining_Toolbox_in_Python

Koskela L. (2000). An exploration towards a production theory and its application to construction. PhD. VTT Technical Research Centre of Finland. https://www.researchgate.net/publication/35018344_An_Exploration_Towards_a_Production_Theory_and_its_Application_to_Construction

Josephson, P., & Hammarlund, Y. (1999). The causes and costs of defects in construction. A study of seven building projects. *Automation in Construction, 8*, 681-687. https://doi.org/10.1016/S0926-5805(98)00114-9

Kalsaas, B. T. (2012). Further Work on Measuring Workflow in Construction Site Production, *Proceedings of the 20th Annual Conference of the International Group for Lean Construction (IGLC20)*.

Koskela, L. (2004). Making-Do — the Eighth Category of Waste, *Proceedings of the 12th Annual Conference of the International Group for Lean Construction (IGLC12)*.

Leão, C. F., Formoso, C. T. & Isatto, E. L. (2014). Integrating Production and Quality Control with the Support of Information Technology, *Proceedings of the 22nd Annual Conference of the International Group for Lean Construction (IGLC22)*.

Leão, C. F., Isatto, E. L., Formoso, C. T. (2016). Proposta de modelo para controle integrado da produção e da qualidade com apoio da computação móvel. *Associação Nacional de Tecnologia do Ambiente Construído,* 16 (4), 109-124. https://doi.org/10.1590/s1678-86212016000400108

Ohno, T. (1997). O Sistema Toyota de Produção: além da produção em larga escala. *Bookman.*

Ronen, B. (1992). The complete kit concept. *The International Journal of Production Research*, v. 30, n. 10, p. 2457-2466. https://doi.org/10.1080/00207549208948166

Santos, É. M. D., Fontenele, A. D., Machado, A. M. L., Neto, J. P. B., & Amaral, T. G. (2020). Analysis of Making-Do Waste at Construction Site in Fortaleza, Ceará, Brazil, *Proceedings of the 28th Annual Conference of the International Group for Lean Construction (IGLC28)*. https://doi.org/10.24928/2020/0082

Saurin, T. A., & Sanches, R. C. (2014). Lean Construction and Resilience Engineering - Complementary Perspectives of Variability, *Proceedings of the 22nd Annual Conference of the International Group for Lean Construction (IGLC22)*.

Sommer, L. (2010). Contribuições Para Um Método de Identificação de Perdas Por improvisação em Canteiros de Obras. Porto Alegre, 2010. 150 f. Dissertação (Mestrado em Engenharia Civil) – Programa de Pós-Graduação em Engenharia Civil, Universidade Federal do Rio Grande do Sul, Porto Alegre. https://lume.ufrgs.br/handle/10183/34763

Santos, P. R., & Santos, D. G. (2017). Investigação de perdas devido ao trabalho inacabado e o seu impacto no tempo de ciclo dos processos construtivos. *Associação Nacional de Tecnologia do Ambiente Construído, 17(2), 39-52.* ISSN 1678-8621 http://dx.doi.org/10.1590/s1678-86212017000200145

Zhang, C., Liu, C., Zhang, X., & Almpanidis, G. (2017). An up-to-date comparison of state-of-the-art classification algorithms. *Expert Systems with Applications, 82, 128–150.*