Enhancing Dyeing Processes with Machine Learning: Strategies for Reducing Textile Non-Conformities

Mariana Carvalho[1][a], Ana Borges[1][b], Alexandra Gavina[2][c], Lídia Duarte[1], Joana Leite[3,4][d], Maria João Polidoro[5,6][e], Sandra Aleixo[6,7][f] and Sónia Dias[8,9][g]

[1]*CIICESI, ESTG, Polytechnic of Porto, Rua do Curral, Casa do Curral, Margaride, Felgueiras, 4610-156, Portugal*

[2]*Lema-ISEP, Polytechnic of Porto, Rua Dr. António Bernardino de Almeida, 431, Porto, 4249-015, Portugal*

[3]*Polytechnic University of Coimbra, Rua da Misericórdia, Lagar dos Cortiços, S. Martinho do Bispo, 3045-093 Coimbra, Portugal*

[4]*CEOS.PP Coimbra, Polytechnic University of Coimbra, Bencanta, 3045-601 Coimbra, Portugal*

[3]*ESTG, Polytechnic of Porto, Rua do Curral, Casa do Curral, Margaride, Felgueiras, 4610-156, Portugal*

[6]*CEAUL – Centro de Estatística e Aplicações da Universidade de Lisboa, Portugal*

[7]*Department of Mathematics, ISEL – Instituto Superior de Engenharia de Lisboa, Portugal*

[8]*ESTG, Instituto Politécnico de Viana do Castelo, Portugal*

[9]*LIAAD-INESC TEC, Portugal*

Keywords: Textile Dyeing, Non-Conformity, Data Mining, Knowledge Discovery, Prediction, Random Forest, Gradient Boosted Trees.

Abstract: The textile industry, a vital sector in global production, relies heavily on dyeing processes to meet stringent quality and consistency standards. This study addresses the challenge of identifying and mitigating non-conformities in dyeing patterns, such as stains, fading and coloration issues, through advanced data analysis and machine learning techniques. The authors applied Random Forest and Gradient Boosted Trees algorithms to a dataset provided by a Portuguese textile company, identifying key factors influencing dyeing non-conformities. Our models highlight critical features impacting non-conformities, offering predictive capabilities that allow for preemptive adjustments to the dyeing process. The results demonstrate significant potential for reducing non-conformities, improving efficiency and enhancing overall product quality.

1 INTRODUCTION

Nowadays, there has been a notable evolution in the textile sector due to technological progress and a growing focus on quality and sustainability. Among the key areas constantly scrutinized is the dyeing process, which plays a vital role in achieving the desired appearance and meeting strict product requirements. Yet, this procedure is fulled with challenges, including non-conformities such as stains, fading and color mismatches. These challenges not only influence the aesthetic of textile items but also affect customer approval and the ecological impact of manufacturing methods.

Considering this, a Portuguese company in the textile dye sector has proposed a significant challenge. The company's goal is to uncover patters that may lead to non-conformities in the dyeing process. The challenge requires examining numerous variables that may impact the results of dyeing, such as the fabric type, the chemical makeup of dyes and the details of the dyeing equipment. Understanding the complex interplay between these factors is crucial for identifying the root causes of non-conformities, which can vary widely and be influenced by subtle changes in the production process.

[a] https://orcid.org/0000-0003-2190-4319
[b] https://orcid.org/0000-0003-4244-5393
[c] https://orcid.org/0000-0002-4694-933X
[d] https://orcid.org/0000-0001-6828-9486
[e] https://orcid.org/0000-0002-2220-4077
[f] https://orcid.org/0000-0003-1740-8371
[g] https://orcid.org/0000-0002-2100-2844

Carvalho, M., Borges, A., Gavina, A., Duarte, L., Leite, J., Polidoro, M., Aleixo, S. and Dias, S.
Enhancing Dyeing Processes with Machine Learning: Strategies for Reducing Textile Non-Conformities.
DOI: 10.5220/0012992800003838

In *Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2024) - Volume 1: KDIR*, pages 363-370
ISBN: 978-989-758-716-0; ISSN: 2184-3228

Therefore, the authors suggest performing an detailed data analysis and applying machine learning algorithms to prediction of key factors that may lead to non-conformities, such as Random Forest (RF) and Gradient Boosted Trees (GBT) algorithms. Machine learning algorithms allow for the examination of large quantities of data in order to uncover patterns and relationships that may not be readily apparent using conventional analysis techniques. Also, these machine learning models are used to recognize main factors that affect dyeing non-conformities and, since both models have the ability to predict outcomes, it is possible to suggest proactive modifications in the dyeing process, showing considerable potential in decreasing flaws, enhancing productivity and improving the overall quality of the product.

This paper is organized as follows: the background section explores the integration of advanced data analysis techniques and machine learning algorithms in textile dyeing processes, emphasizing the identification of key factors influencing dyeing non-conformities and offering strategies to enhance product quality. Next, the authors present a descriptive analysis of the dataset on non-conformities, highlighting its key features. This is followed by a detailed exploratory data analysis section, organized into several subsections: Analysis of Non-Conformities, Causes of Non-Conformities, Fabrics with Non-Conformities, Colourants in Non-Conformities and Colouring Machines that Lead to Non-Conformities. Subsequently, the paper provides a detailed explanation of the entire process of predicting significant factors that may be resposible for non-conformities using machine learning. Finally, the paper concludes with a discussion and comparison of the results, followed by the Conclusions and Future Work section.

2 BACKGROUND

As stated before, the textile industry has recently advanced due to technological progress and a focus on sustainability and quality control, particularly in dyeing processes. The main challenges include minimizing environmental impacts and addressing non-conformities. Studies like (Zhang et al., 2018) have proposed improved designs for textile production processes based on life cycle assessment, targeting the reduction of environmental impacts by identifying best available technologies and focusing on critical stages like printing and dyeing to improve product quality and reduce resource depletion and ecological influence. (Parisi et al., 2015) emphasize the need for more sustainable production processes, demonstrating the feasibility of alternative dyeing methods that reduce energy, water and raw materials consumption, thereby aligning with consumer demand for eco-friendly products.

In response to these challenges, the integration of advanced data analysis techniques and machine learning into the textile dyeing process represents a significant shift towards more data-driven decision-making. Research by (Park et al., 2020) has developed a cyber-physical energy system that utilizes manufacturing big data and machine learning techniques to improve energy efficiency in dyeing processes without the need for expensive equipment, thereby enhancing process and system efficiency. Furthermore, efforts to incorporate green solvents, as discussed by (Meksi and Moussa, 2017) and to explore the ecological application of ionic liquids in textile processes, offer innovative pathways for reducing the environmental footprint and improving the sustainability of the dyeing process. These developments not only aim to address immediate quality control challenges but also signify a broader movement towards incorporating advanced technologies in traditional textile dyeing industries, setting a new benchmark for sustainability and efficiency.

3 DATASET ON NON-CONFORMITIES

In this section, the authors present the dataset used for the analysis, detailing the preprocessing steps and the comprehensive descriptive statistics of the variables involved.

All preprocessing tasks were conducted using RapidMiner[1] and Python[2]. Missing data were imputed using the K-Nearest Neighbour (Fix, 1985) method to ensure the integrity and completeness of the dataset. This preprocessing step is crucial for accurate and reliable machine learning model training.

In our analysis, the original dataset comprises a total of 5,546 records across 23 distinct variables. But, in order to maintain the confidentiality and anonymity of the textile company, the authors only consider the following set of variables in the subsequent analysis: Fabric, Colourant, Date, Defect (which corresponds to Non-Conformity), Cause and Colouring Machine. The descriptive statistics of all variables are as summarized in Table 1. For categorical variables, the table provides name of variable and unique values. For numerical variables, it includes the name of the variable,

[1] https://altair.com/altair-rapidminer
[2] https://www.python.org/

Table 1: Descriptive Statistics of the Dataset.

Variable name	Unique	Mean	STD	Min	Max
Fabric	13	-	-	-	-
Colourant	-	3.35	2.95	0.00	17.00
Date	754	-	-	-	-
Defect	6	-	-	-	-
Cause	9	-	-	-	-
Colouring Machine	39	-	-	-	-

mean, standard deviation (STD), minimum (Min) and maximum (Max) values.

4 EXPLORATORY ANALYSIS OF THE DATASET ON NON-CONFORMITIES

In this section an exploratory analysis of the content of the database is presented. The authors explore the textile manufacturing non-conformities from January 2020 to July 2023 and show the patterns and trends that emerge from the data, seeking to understand the underlying causes and their temporal dynamics.

4.1 Analysis of Non-Conformities

First, it is important to analyse the evolution of non-conformities occurrences over the years. Figure 1 shows this evolution over the period from 2020 to 2023. The non-conformities considered in this study are 'Stained', 'Oil', 'Other', 'Failed', 'Undyed' and 'Creases'. Overall, the total number of non-conformities (represented by the dark blue line) decreased each year, reflecting an overall improvement in quality control measures. 'Failed' non-conformities (represented by the yellow line) consistently has the highest number of non-conformities. The 'Oil' (represented by the orange line) exhibits variability, with a slight peak in 2021, followed by a consistent decline in 2023. The 'Other' non-conformities occurrences (represented by the grey line), which includes miscellaneous non-conformities, peaked in 2020 and showed a gradual decrease by 2023. 'Stained' non-conformities (represented by the blue line) shows a decreasing trend over the years, starting in 2020 and declining in 2023. 'Undyed' (represented by the light blue line) shows fluctuations, with the highest number in 2022. Despite these fluctuations, the trend appears relatively stable with a slight increase. Lastly, 'Creases' (represented by the green line) shows a slight decrease over the years.

The distribution of non-conformities is detailed as follows: The 'Failed' non-conformity has the highest count, with 2142 occurrences, representing 39% of the total non-conformities. The 'Other' and 'Undyed' categories follow, each constituting 16% of the total non-conformities, with counts of 903 and 915 respectively. 'Stained' non-conformities account for 12% of the total, with 673 occurrences, while 'Creases' represent 9% with 486 occurrences. The 'Oil' non-conformity, although the least frequent, still comprises 8% of the total non-conformities, with 427 occurrences.

4.2 Causes of Non-Conformities

The next step is to analyse the causes that influence non-conformities. The distribution of causes of non-conformities are described as followed: the most significant issue is 'Poorly analysed,' with 1880 occurrences, corresponding a total of 34%. 'Other' reasons have also led to a considerable number of occurrences, totalling 1017 (18%). The 'Poorly executed/monitored process' accounts for 786 occurrences (14%). 'Process phases in different conditions' have contributed to 567 (10%) non-conformities. 'Insufficient disposal by normal process' is the next most frequent concern with 430 occurrences (8%). 'Rope jammed/rebent/running poorly', 'Dyed (folded) accessory together with mesh', 'Lack of machine/cart cleaning' and 'Process interrupted for review' have occurrences over 200 (each one with 4% of total occurrences).

4.3 Fabrics with Non-Conformities

Following this, the analysis of fabrics with non-conformities is also important. The description of fabrics with non-conformities' distribution is as follows. The predominant fabric with non-conformities is Jersey, with 1825 of total occurrences, comprising 33% of the total occurrences. Followed by Rib (with a total of 1337 occurrences) at 24% of the non-conformities. Felpa fabrics represent 14% of the non-conformities and with a total of 754 occurrences, while Golve fabrics contribute 6%. Both Piquet and Screen fabrics account for 7% each. Other fabric types, such as

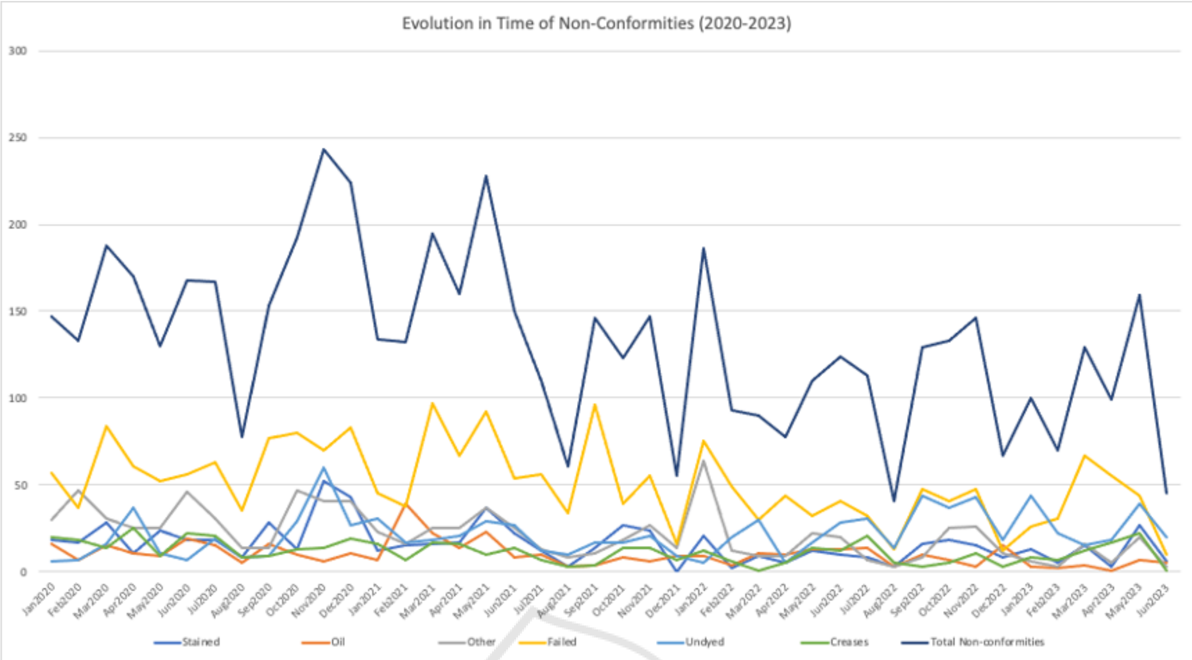

Figure 1: Evolution of Non-Conformities over the years.

Nastro and Interlock, each represent 3% of the non-conformities. Minor categories include Screen, Nets and Cord, each constituting 1% and Turca and Strips have a negligible 0% presence of non-conformities.

4.4 Colourants Presented in Non-Conformities

Next, the authors analyse the distribution of colourants presented in non-conformities. The 'Reactive' colourant has an overwhelmingly high count of non-conformities, totaling 4405, which constitutes 76.46% of the total non-conformities. The colourant 'Reactive/Disperse' also shows a substantial number of non-conformities, with a count of 535, accounting for 9.29% of the total. While significantly lower than 'Reactive', this combination of colourants still represents a considerable source of non-conformities. With 271 non-conformities, 'Colourless' dyes represent 4.70% of the total. 'White' dyes account for 96 non-conformities, with 1.67% of the total occurrences. The colourant 'Acid' has 87 non-conformities, with 1.51% of the total. The combination of 'Reactive/Acid' dyes results in 79 non-conformities, which is 1.37% of the total. Disperse' dyes show a relatively low count of 23 non-conformities, representing 0.40% of the total. The 'Direct' and 'Indefinite' colourants have the lowest counts, with 30 (0.52%) and 10 (0.17%) occurrences respectively. Similarly, 'Cationic/Reactive' colourants also have a low count of 10 non-conformities, which is 0.17% of the total occurrences.

4.5 Colouring Machines

Another important variable that may impact the non-conformities occurrences is the variable colouring machines. This dataset present a total of 39 colouring machines and overall, there's a fluctuation in the percentage of non-conformities for each machine across the four years. Some machines show a reduction in non-conformities over time, while others exhibit an increase or inconsistent patterns. The top 5 colouring machines leading to the most occurrences of non-conformities are: 'TNJT13' with a total of 419 (7.55%), 'TNJT05' with 346 occurrences (6.24%), 'TNJT19' with a total of 304 (5.48%) 'TNJT11' with a sum of 300 occurrences (5.41%) and finally, 'TNJT32' with 287 (5.17% of total occurrences).

This extensive analysis of data helps improve comprehension of the data, leading to better feature engineering and model development in future analysis.

5 PREDICTION OF DYEING NON-CONFORMITIES FACTORS

This section explores the use of machine learning algorithms, specifically RF and GBT, to predict key factors that may contribute to non-conformities and extract feature importance, identifying the most significant factors contributing to these issues. Understanding these key features enables targeted interventions and process optimizations, enhancing product quality and reducing defect rates.

RF combines multiple decision trees to enhance predictive accuracy and control overfitting, making it suitable for datasets with numerous features and non-linear relationships (Robnik-Sikonja, 2004). This algorithm has been effectively utilized in various industrial contexts, such as predictive maintenance, where it anticipates equipment failures by analyzing sensor data and operational logs, thus minimizing downtime and improving productivity (Kusiak and Verma, 2011). Additionally, RF provides insights into feature importance, crucial for understanding key factors influencing non-conformities in dyeing processes (Breiman, 2001).

Conversely, the GBT algorithm builds trees sequentially, with each new tree correcting errors made by the previous ones, thereby significantly enhancing prediction accuracy (Friedman, 2001). GBTs have demonstrated superior performance in industrial applications and in manufacturing. GBTs optimize production processes by identifying critical factors influencing product quality, enabling precise control and reduction of non-conformities (He and Wu, 2018).

5.1 Findings of the Random Forest Model

The Figure 2 shows a representative tree model obtained using the RF algorithm (Ho, 1995). The model configuration chosen was: the number of trees in the forest equals to 100 and the minimum number of samples required in leaf node equals to 50 and the data was divided into 80% for training and 20% for testing. The returned RF model represented in the Figure shows the factors that lead to non-conformities, which was used as classe label. According to the root node, the most important factor is ‘Poorly analysed’ processes mainly resulting in ‘Failed’ classifications. After a thorough analysis, the next significant factor is the ‘Dyed (folded) accessory along with mesh’ process, frequently leading to ‘Undyed’ non-conformities. Next, there are machine-specific factors, especially those involving the colouring machine ‘TNJT25’ and ‘TNJT23’, as well as fabric-related issues like problems with ‘Golves’ fabric, play a significant role in influencing non-conformities. In the Figure, it also possible to see that ‘Colourless’ colourants and ‘Piquet’ fabrics play a major role in ‘Other’ non-conformities. In addition, issues related to the dyeing process such as ‘Process phases in different conditions’ and ‘Insufficient dispossal by normal process’ are important elements.

The obtained RF estimator’s performance measures shows the model’s accuracy in predicting different types of non-conformities. The precision for predicting the non-conformity ‘Crease’ is 0.53, which means that 53% of the predicted creases were correct. The recall is 0.23, suggesting that only 23% of the actual creases were identified. The f1-score is 0.33, reflecting the balance between precision and recall. The ‘Failed’ non-conformity has perfect precision (1.00) and high recall (0.87), resulting in a high f1-score (0.93), which means that 100% of ‘Failed’ predictions are accurate. The precision and recall in the ‘Oil’ non-conformity are both very high (0.97 and 0.99, respectively) and a f1-score of 0.98. The ‘Stained’ non-conformity has a precision of 0.60 and a recall of 0.88. The precision is 0.92 and the recall is 0.81 on the ‘Undyed’ non-conformity, resulting in an f1-score of 0.86. In the ‘Other’ non-conformity the precision is 0.54 and the recall is 0.73, resulting in an f1-score of 0.62. The overall accuracy of the model is 0.79, so it shows that almost 80% of the predictions are accurate. This is a strong performance, indicating that the model is successful in identifying various types of non-conformities.

With the analysis of important features from the RF model one can know which are the features that impact mostly the non-conformities appearances. The cause ‘Poorly analysed’ remains the most influential feature of non-conformities, with an importance value of 0.385060. The second most influential feature is the cause ‘Process phases in different conditions’ which presents an importance of 0.141318. The cause ‘Insufficient disposal by normal process’, rated at 0.135305 in terms of importance, is the third most influential feature. The cause ‘Poorly executed/monitored process’, with a significance rating of 0.117062, is also a major factor in non-conformities.

Additional important factors are the ‘Other’ causes (0.052049), the cause ‘Lack of machine/cart cleaning’ (0.049191) and the cause ‘Process interrupted for review’ (0.042287). While not as influential enough as the other main causes, these factors still greatly affect non-conformities. Other process problems like the cause ‘Dyed (folded) accessory together with mesh' (0.019247) and the cause 'Rope jammed/rebent/running poorly' (0.013083) also play an important role in leading to non-conformities. The colourant 'Colourless' (0.008710), the fabric 'Jersey' (0.007713), the colourant 'Reactive' (0.003967) and the fabric 'Rib' (0.003472) shows that not only the causes and processes influence the non-conformities. While not as significant as cause and process-related factors, these features still contribute to influence non-conformities. Also, Machine-specific features, like the colouring machine 'TNJT05' (0.003181), show that particular machines impact non-conformity rates as well. The fabric 'Piquet' (0.003146) is also considered in the top 15 of the more influential features, suggesting that along with the fabric Jersey and Rib can also lead to non-conformities.

5.2 Findings of the Gradient Boosted Trees Model

Next, the authors apply the GBT algorithm. The chosen model configuration was learning rate ('classifier_learning_rate') are 0.2, maximum depth ('classifier_max_depth') equals to 5, the number of trees ('classifier_n_estimators') equals to 100 and also, the data was split with 80% allocated for training and 20% for testing.

The returned performance metrics in this model are very similar to the ones obtained previously using the RF model. In the 'Creases' non-conformity, the model obtained a precision score of 0.48, which means that 48% of the predicted creases were correct; and a recall score of 0.38, suggesting that only 38% of the actual creases were identified. Within the 'Failed' prediction, the model showed strong results with a precision of 0.96 and a recall of 0.89. The 'Oil' group also showed good outcomes, achieving a precision of 0.97 and a recall of 0.96. On predicting 'Other' non-conformities, the model achieved a precision of 0.59 and a recall of 0.71. Similarly, the 'Stained' non-conformity showed a precision of 0.66 and a recall of 0.79. In the 'Undyed' non-conformity classification, the model reached a precision of 0.90 and a recall of 0.84. Overall, the GBT model achieved an Accuracy of 0.80.

The top 15 factors identified by the GBT

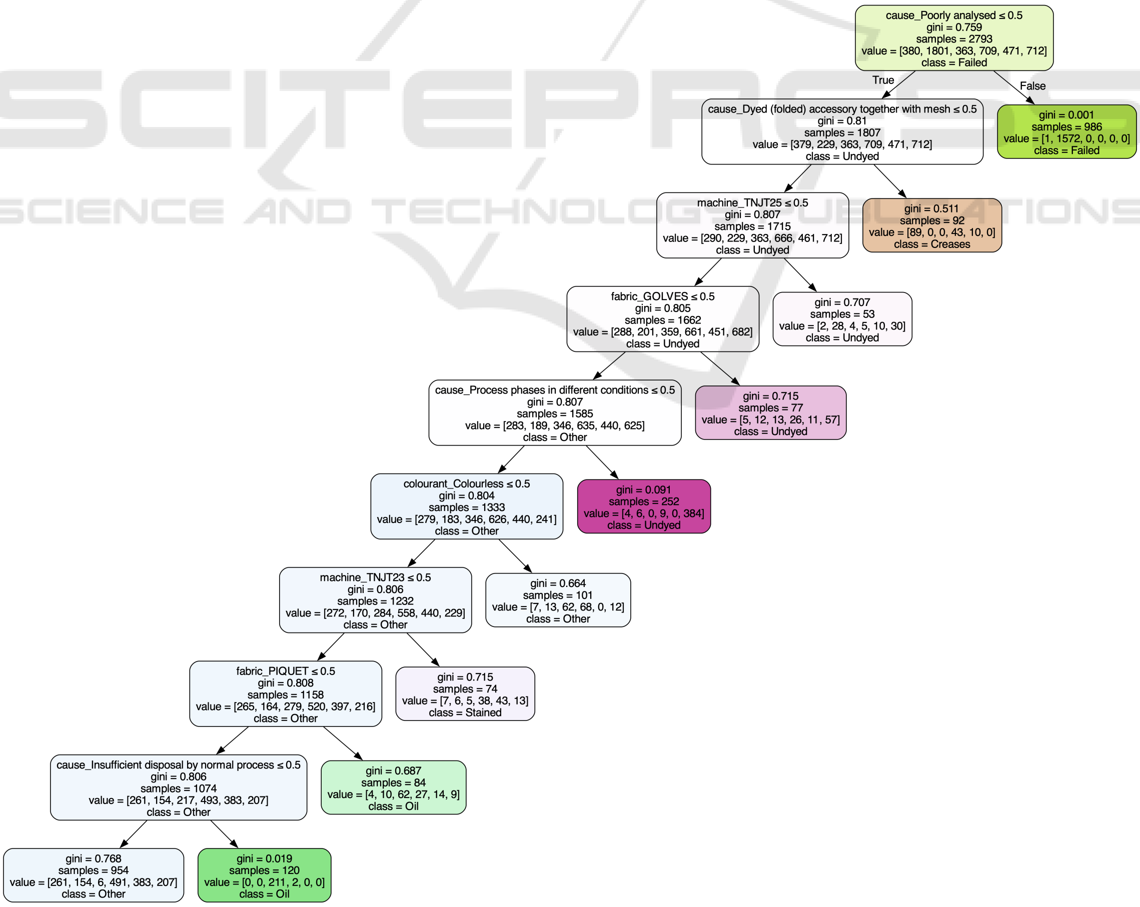

Figure 2: A representative Tree obtained from the Random Forest model.

model that most influence the occurrence of non-conformities are as follows: The ‘Cause_Poorly analysed’ holds the top score of 0.318236, highlighting its major influence on the model’s forecasts. Following this are the phrases ‘Cause_Process phases under various circumstances’ with a significance rating of 0.129308 and ‘Cause_Inadequate disposal through regular procedures’ with a rating of 0.121097. Some other significant characteristics are ‘Reason for lack of cleaning of machine/cart’ (0.083689), ‘Reason for process interruption for review’ (0.052792) and ‘Reason for poorly executed/monitored process’ (0.036939). The factor ‘Cause_Other’ also has a significance level of 0.028393. Further factors like ‘Cause_Dyed accessory folded with mesh’ (0.025745) and ‘Cause_Rope jammed/rebent/running poorly’ (0.018613) also play a role in the model’s predictions. Some characteristics related to particular devices and dyes are also present in the top 15. The listed items are ‘Colouring_Machine_TNJT06’ (0.005783), ‘Colouring_Machine_TNJT32’ (0.005715), ‘Colourant_Colourless’ (0.005656), ‘Fabric_Screen’ (0.005642), ‘Colourant_Acid’ (0.005449) and ‘Colouring_Machine_TNJT25’ (0.005265).

6 DISCUSSION

The application of machine learning algorithms, specifically RF and GBT, to the textile dyeing process has yielded significant insights into the factors influencing non-conformities. Our analysis identified several key variables that impact the occurrence of non-conformities, such as poorly analysed processes, variations in process phases and insufficient disposal methods. These findings are summarized in Table 2.

The RF model’s high importance score for ‘Poorly analysed process’ underscores the necessity for thorough inspections and quality checks at each stage of the dyeing process. This feature’s dominance suggests that many non-conformities could be mitigated by improving the rigor of process analysis. Similarly, the GBT model aligns closely with this finding, reinforcing the critical role of detailed process scrutiny. ‘Process phases in different conditions’ emerged as another significant factor. Variations in these conditions can lead to inconsistencies in dye application, resulting in non-conformities. Both models consistently rated this feature highly, suggesting that addressing these variations could significantly reduce non-conformities.’Insufficient disposal by normal process’ also featured prominently in both models, indicating that the methods used to remove excess materials or byproducts during dyeing can influence the final product’s quality. Optimizing disposal processes to ensure complete removal of unwanted substances could enhance overall dyeing consistency. The ‘Poorly executed/monitored process’ factor, while rated lower in the GBT model, still showed considerable importance in the RF model. This points to the need for continuous monitoring and quality assurance practices during dyeing to prevent errors and ensure uniform quality.

Comparing our findings with existing literature, such as Zhang et al. (2018) and Parisi et al. (2015), reveals a consistent emphasis on the importance of process control and quality management in reducing non-conformities by improving sustainability in textile manufacturing. Our study extends these ideas by providing a data-driven approach to identifying and addressing specific factors leading to non-conformities.

7 CONCLUSIONS

This study demonstrates the potential of machine learning techniques in optimizing the textile dyeing process by identifying and mitigating factors leading to non-conformities. Machine learning models such as RF and GBT provide a detailed analysis of critical features impacting dyeing quality, which is relevant for industry practitioners to enhance process control and quality assurance practices.

The high importance scores for process analysis and conditions suggest that many non-conformities can be mitigated through more rigorous quality checks and maintaining consistent dyeing environments. The alignment of our results with existing literature further validates the significance of robust process control and quality management in the textile industry. Integrating these machine learning results into the dyeing process can lead to substantial improvements in efficiency, waste reduction and overall product quality. This approach not only addresses immediate quality control challenges but also sets a new standard for incorporating advanced technologies in traditional manufacturing processes.

Future work for this study includes expanding the dataset to cover a wider variety of textiles and dyeing methods to improve the accuracy of the predictive models. Incorporating more machine learning algorithms, such as deep learning methods, may yield more precise and reliable predictions. Moreover, utilizing real-time data analysis and anomaly detection systems may facilitate prompt corrective measures during dyeing, thereby enhancing efficacy and minimizing non-conformities occurrences.

Table 2: Top features influencing non-conformities in the dyeing process according to Random Forest and Gradient Boosted Trees models.

Feature	**Description**
Poorly analysed process	Indicates process steps not thoroughly checked, leading to non-conformities.
Process phases in different conditions	Variations in process phases affecting dyeing quality.
Insufficient disposal by normal process	Inadequate removal of materials causing non-conformities.
Poorly executed/monitored process	Indicates issues in the execution and monitoring of dyeing processes.

ACKNOWLEDGEMENTS

This work has been supported by national funds through FCT – Fundação para a Ciência e Tecnologia through project UIDB/04728/2020. The authors thank the textile company for providing the real data used in this study.

The research at CMAT was partially financed by Portuguese Funds through FCT (Fundação para a Ciência e a Tecnologia) within the Projects UIDB/00013/ 2020 and UIDP/00013/2020.

Partially supported by the Centre for Mathematics of the University of Coimbra (funded by the Portuguese Government through FCT/MCTES, DOI 10.54499/UIDB/00324/2020).

Research partially funded by FCT - Fundação para a Ciência e a Tecnologia, Portugal, under the project UIDB/00006/2020, DOI: 10.54499/UIDB/00006/2020 (CEAUL) and ISEL.

REFERENCES

Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–32.

Fix, E. (1985). *Discriminatory analysis: nonparametric discrimination, consistency properties*, volume 1. USAF school of Aviation Medicine.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232.

He, Y., H. Z. and Wu, Q. (2018). Production line control system for the wulanchabu manufacturing plant. In *Proceedings of the 2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pages 1–4.

Ho, T. K. (1995). Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, volume 1, pages 278–282. IEEE.

Kusiak, A., Z. Z. and Verma, A. (2011). Predictive maintenance: Thrust bearing fault classification. *Proceedings of the ASME 2011 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*.

Meksi, N. and Moussa, A. (2017). A review of progress in the ecological application of ionic liquids in textile processes. *Journal of Cleaner Production*, 161:105–126.

Parisi, M., Fatarella, E., Spinelli, D., Pogni, R., and Basosi, R. (2015). Environmental impact assessment of an eco-efficient production for coloured textiles. *Journal of Cleaner Production*, 108:514–524.

Park, K., Kang, Y. T., Yang, S., Zhao, W.-B., Kang, Y.-S., Im, S., Kim, D. H., Choi, S. Y., and Noh, S. (2020). Cyber physical energy system for saving energy of the dyeing process with industrial internet of things and manufacturing big data. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 7:219–238.

Robnik-Sikonja, M. (2004). Improving random forests. In *Proceedings of the 15th European Conference on Machine Learning*, pages 359–370. Springer.

Zhang, Y., Kang, H.-S., Hou, H., Shao, S., Sun, X., Qin, C., and Zhang, S. (2018). Improved design for textile production process based on life cycle assessment. *Clean Technologies and Environmental Policy*, 20:1355–1365.