MASTER’S DEGREE IN COMPUTATIONAL SOCIAL SCIENCES

2024/2025

Master Thesis

“Explaining Productivity in the Garment Industry: An Analytical and Machine Learning Approach”

Jorge Ramos Val

Director: Emma Pérez Hernández

Madrid, /06/2025

Abstract

This study investigates the determinants of team productivity in the Bangladeshi garment industry, a sector characterized by intense labor demands and frequent productivity shortfalls. Using real factory data, a supervised classification approach was applied to distinguish teams that met or missed their productivity targets. The analysis combined XGBoost and Random Forest models to identify key organizational and operational factors, with special attention to explainability. Results show that team size, economic incentives, work in progress, and task complexity are the most influential variables for achieving productivity goals. Both models achieved high accuracy and sensitivity, confirming their practical value for early intervention. These insights contribute to a better understanding of productivity dynamics in industrial environments and provide an evidence-based foundation for targeted organizational improvements.

Key words: Productivity, Garment Industry, Performance, Economic Incentives, Machine Learning.

INDEX

TABLE INDEX

LIST OF FIGURES

1. Introduction and Literature review

Productivity is a key element in the economy, as it refers to a company's ability to produce goods or provide services efficiently, optimizing the use of resources and achieving operational objectives. In this study, labor productivity is specifically addressed from an individualized perspective, evaluating the efficiency and effectiveness of workers in relation to factors such as time, resources used, or environmental conditions (OECD, 2001).

Productivity depends on both individual factors (such as motivation, training, or worker health) and organizational factors, which include the work environment, the structure of incentives, or the level of technological innovation. In this regard, Alam, Alias, and Azim (2018) identify up to nine key determinants of productivity: working hours, wages and benefits, holidays, discrimination, harassment and abuse, workplace conditions, forced labor, well-being, and labor relations.

Over the past decades, in the context of economic globalization, the pursuit of higher productivity levels has led to multiple strategies by both states and companies. Since the 1960s, the outsourcing of production to countries in the Global South has been planned by these actors, relocating production where labor legislation is more lax and labor costs are significantly lower.

This phenomenon has also had profound political consequences. Various authors have pointed out how the extraction of surplus value from the periphery to the economic centers reproduces forms of economic domination, conceptualized as new imperialism (Harvey, 2003), and generates a dynamic of global capital that is deterritorialized and depersonalized, making it difficult to assign responsibility and limiting the regulatory capacity of the receiving states.

In these contexts, employee productivity tends to increase, not necessarily due to technical efficiency, but rather because of labor precariousness, productive pressure, and low structural costs (Harvey, 2003). Thus, both active surplus value (through longer working hours and wage reductions) and passive surplus value (through automation and reduced indirect costs) are increased.

The garment industry is a paradigmatic example of this model. In countries such as Bangladesh, a powerful textile industry oriented toward exports has developed, especially to meet the demand for fast fashion in Global North countries. Bangladesh is the world's second-largest exporter of ready-made garments (only behind China), with exports in 2023 reaching $38 billion and representing more than 14% of the country's GDP. The sector employs about 4 million people, the majority of whom are women, often in conditions close to exploitation (20minutos; 2024).

These working conditions began to receive greater visibility after the collapse of the Rana Plaza building in 2013, near Dhaka, which housed several textile factories and where more than 1,100 workers died. Since then, and following strong national and international pressure, safety conditions and labor regulations have partially improved, although dynamics of vulnerability persist (20minutos; 2024).

In this context, the present work proposes the use of machine learning techniques to analyze and classify the productivity of work teams in the textile sector in Bangladesh. Supervised machine learning involves training models with data where the outcomes are already known, enabling them to predict results for new cases. This approach includes regression, which predicts continuous values, and classification, which assigns cases to categories (Kassambara, 2018).

In recent decades, data mining has transformed multiple sectors of industrial engineering and manufacturing, enabling companies to leverage the large volumes of information they generate to optimize decision-making and production planning. In this way, the use of machine learning tools can serve to predict and classify productivity, allowing the discovery of hidden patterns and providing valuable information for the efficient management of the textile industry.

According to Hastie, Tibshirani, and Friedman (2009), these methods aim to identify the underlying function that relates the independent variables to the target variable, minimizing the prediction error. Likewise, Kuhn and Johnson (2013) highlight the importance of correct variable selection, data preprocessing, and cross-validation to ensure model robustness, especially when working with real, noisy databases such as those found in the textile production environment.

In the present study, a classification model will be explored with the aim of categorizing whether teams manage to achieve the expected productivity. The work adopts an explanatory perspective to analyze the factors that enable the attainment of expected productivity. Therefore, this approach allows for the evaluation of the impact of variables such as the number of workers per team, economic incentives, idle time, or the number of overtime hours on observed performance.

Recent research demonstrates that labor productivity analysis in the textile sector has increasingly relied on machine learning methods, especially classification models. Among the most prominent studies, Obiedat and Toubasi (2022) proposed a hybrid approach that combines various classification algorithms and ensemble techniques such as Bagging and Adaboost to predict productivity in the textile sector. Their analysis, conducted on the same dataset used in the present work, showed that Random Forest achieved particularly high and stable performance compared to other algorithms such as SVM, neural networks, or decision trees. These authors emphasize the value of ensemble approaches for improving predictive accuracy and optimizing talent and resource management in highly competitive manufacturing environments.

Similarly, Balla, Rahayu, and Purnama (2021) applied data mining techniques to analyze employee productivity in the textile industry, employing algorithms such as Random Forest, linear regression, and artificial neural networks. They highlighted Random Forest as the best-performing algorithm overall, followed by linear regression and neural networks. Their findings identified idle time, team size, and economic incentives as the most relevant variables for predicting productivity, underscoring the need for precise predictive tools to bridge the gap between expected and actual productivity and thus optimize production management.

The dataset used in this study was originally developed and analyzed by Imran et al. (2021), who collected detailed records from a large garment factory in Bangladesh. Their work applied a variety of machine learning models for both regression (predicting actual productivity) and classification (identifying low, moderate, or high productivity teams). Ensemble methods, particularly gradient boosted trees, were found to offer the best results, with incentives, WIP, and team size emerging as key productivity drivers. This context makes the dataset especially relevant for analyzing the organizational and operational factors that shape productivity outcomes.

Additionally, Imran et al. (2019) developed a deep neural network (DNN) to predict employee productivity, incorporating hidden layers to capture complex nonlinear relationships. Their results suggested that overtime and the proportion of idle time are especially important factors in anticipating workforce performance, highlighting the ability of advanced machine learning models to reveal hidden patterns and deliver accurate predictions, even with moderately sized datasets.

The main objective of this study is to identify and analyze the factors that explain why some work teams do not manage to achieve the expected productivity in an intensive manufacturing environment. To address this question, two secondary research questions are posed: first, what role do organizational factors (such as team size, style changes, or organizational structure) play in productivity? Second, to what extent do economic incentives influence team performance?

Based on these questions, the following hypotheses are formulated:

1. Work teams that do not achieve the expected productivity present significantly different organizational, motivational, and operational characteristics compared to those that do achieve it.

2. Factors such as idle time, the absence of economic incentives, and excessive workloads are negatively associated with the probability that a team will achieve its productivity goal.

2. Data and methodology

2.1 Data

The dataset used in this study originates from the Industrial Engineering Department of a textile factory in Bangladesh, tasked with optimizing the allocation of resources such as labor, machinery, and materials. As documented in the original study by Imran et al. (2021), the data consists of daily efficiency reports at both the line and factory levels, allowing for detailed monitoring of productivity based on a range of operational factors. The unit of analysis is the work team ("team" variable), with each record corresponding to a specific team on a given day.

In total, the dataset comprises 1,197 observations collected between January and March 2015, covering 15 variables. These variables include both expected and actual productivity, incentives, overtime worked, team size, and other operational and organizational features. Notably, the actual_productivity variable is a computed index developed by the original authors, reflecting whether teams meet, exceed, or fall short of their productivity targets.

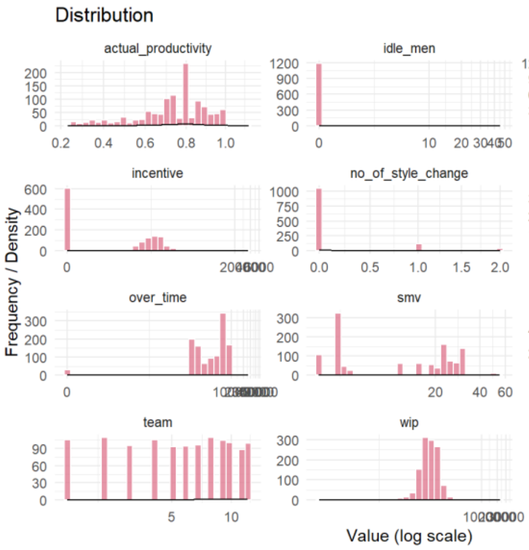
Table 1: Variables descriptions

Variable	Data_Type	Description
date	character	Date variable with MM-DD-YYYY structure.
day	character	Variable containing the name of the days of the week. There are no Fridays.
quarter	character	Variable that divides the month into four quarters.
department	character	Variable that stores the 2 types of departments: Sewing and Finishing.
team_no	numeric	Variable that identifies the team number. A total of 12 teams work in the production departments.
no_of_workers	numeric	Variable that accumulates the number of workers in each team.
no_of_style_change	numeric	Variable that stores the number of changes made to a garment throughout the process.
targeted_productivity	numeric	Variable that records the target production set by the authority for each team for each day.
smv	numeric	Standard Minute Value: Variable representing the time set to complete tasks.
wip	numeric	Work in Progress: Variable that records the number of products to be manufactured, including unfinished items.
over_time	numeric	Variable representing the amount of overtime given to each team in minutes.
incentive	numeric	Variable representing the amount of money allocated to motivate teams to meet production goals.
idle_time	numeric	Variable representing the amount of time the production could have been interrupted.
idle_men	numeric	Variable representing the number of workers who were inactive during production interruptions.
actual_productivity	numeric	Predicted variable recording the actual productivity index of the working team.

Source: own elaboration based on Imran et al. (2021).

Figure 1: Variables distributions

Figure 1:Variables distributions



Source: own elaboration.

A key feature of the dataset is its heterogeneity. The "team" variable, used as a categorical identifier, refers to different groups in each department, with team sizes ranging from 2 to 89 workers (mean: 34.6), illustrating diverse organisational structures. While targeted_productivity is relatively consistent across teams (mean: 0.73), actual_productivity varies more widely (mean: 0.74), reflecting real differences in performance, values above 1 indicate exceeding targets, while those below 1 reflect underperformance.

Other variables such as SMV, over_time, and incentive also show substantial dispersion and skewness: over_time ranges from 0 to 25,920 minutes, and most teams receive minimal incentives, with only a few receiving significant amounts. Variables like idle_time and idle_men are heavily zero-inflated, suggesting that idleness is rare but, when present, can be considerable.

Key variables such as SMV (Standard Minute Value) and WIP (Work in Progress) are calculated using industry-specific formulas, which lend technical rigor to the dataset. WIP is the only variable with missing values (506 cases). As missing WIP does not necessarily indicate inactivity, simple imputation methods like zeros or medians were avoided to preserve the integrity of the data. Instead, regression-based imputation using Random Forest was applied, and an additional indicator variable (wip_missing) was included to adjust for any potential bias (see Figure 7 in the Annex). Following imputation, WIP continued to display substantial variability across teams and time periods.

In the feature engineering section, new variables were created from the available variables in the dataset:

Furthermore, new variables were created from those available in the dataset using feature engineering techniques, in order to enrich the representation of the problem and improve the predictive capacity of the models. These variables are:

1. Labor Exploitation Index: This composite indicator quantifies the extent to which teams exceed planned working time by integrating overtime, economic incentives, and style changes into a single proportion. For comparability, incentive values are divided by 10 (to match the scale of overtime minutes), while the number of style changes is multiplied by 15, based on expert input that each change requires about 15 extra minutes of coordination. The index draws on the frameworks of Chan and Siu (2010) and Patnaik (1972), which relate exploitation to overtime, workload, and compensation. Higher index values indicate greater “extra” work

compared to what was planned, signalling increased labour demands and potential overexertion (see Figure 10 in the Annex). The index is calculated as follows:

$$Normal\ Time = SMV \times nº\ of\ workers$$

$$Labor\ Exploitation\ Index = \frac{Over\ Time + \frac{Incentive}{10} + 15 \times nº\ of\ style\ change}{Over\ Time + Normal\ Time + \frac{Incentive}{10} + 15 \times nº\ of\ style\ change}$$

2. Idle ratio: to assess the efficiency of time use, the idle ratio was constructed as the share of inactive time relative to total available production time (Ahn et al., 2013). This variable helps identify inefficiencies, bottlenecks, or organizational issues. The data reveal that idleness is infrequent, but when it does occur, even at low levels, it negatively impacts overall productivity (see Figure 11 in the Annex). Calculated as follows:

$$Idle\ ratio = \frac{Idle\ Time}{Idle\ Time + SMV + Over\ Time}$$

3. Surplus value rate: Based on Marx's (1867) original formulation and its adaptation in recent studies (Tricontinental Institute, 2019), this metric quantifies the "apparent surplus value" generated by each team. It compares the value produced, including unplanned overtime, against necessary labor, while accounting for economic incentives. The incentive value is divided by 10 to ensure comparability with the other terms in the formula, since incentives are reported in local currency units that are typically an order of magnitude larger than time-based variables. Results indicate that, in most cases, teams generate more value than they receive in compensation, especially in departments where incentive systems are limited (see Figure 12 in the Annex). Calculated as follows:

$$Necessary\ work = SMV \times nº\ of\ workers$$

$$Apparent\ Surplus\ Value = Over\ Time + (Actual\ Productivity \times SMV) - \frac{Incentive}{10}$$

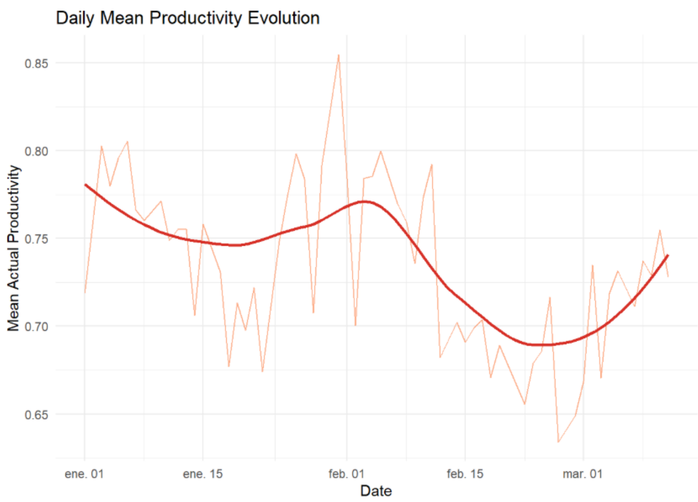
$$Surplus\ Value\ Rate = \frac{Apparent\ Surplus\ Value}{Necessary\ Work}$$

These transformations allow for the introduction of more complex or contextual relationships between the original variables.

The bivariate analysis of the data reveals the profound influence that organizational factors, economic incentives, and operational conditions have on team productivity in the textile manufacturing environment.

Actual productivity shows considerable volatility throughout the analyzed period, with sharp declines and subsequent recoveries that appear to be associated with operational and seasonal factors. In contrast, target productivity remains much more stable, acting as a structural benchmark imposed by the factory for the teams. This dynamic is clearly illustrated in the time series plot of daily mean productivity, where fluctuations in actual productivity are much more pronounced compared to the relatively steady trajectory of the target values (see Figure 8 in the Annex).

Figure 2: Daily Mean Productivity Evolution



Source: own elaboration.

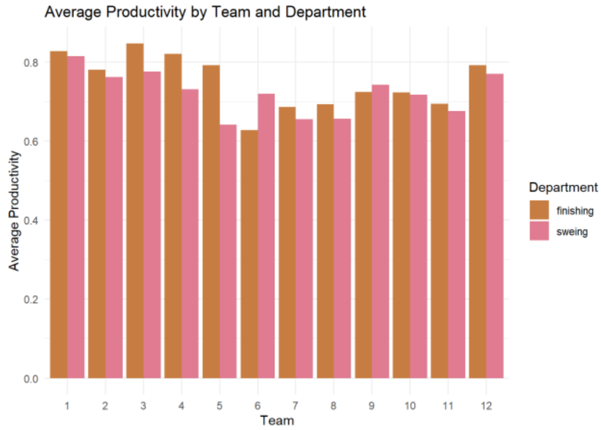
On the other hand, there are marked differences between departments. The sewing area stands out for its larger team sizes and higher incentives, showing more homogeneous productivity, while the finishing area exhibits greater dispersion in results and noticeably smaller teams. At the team level, heterogeneity is significant; the most productive teams tend to combine intermediate or high incentives, low idle time, and a proper balance between size and task complexity

Table 2: Department summary

Department Comparative Table								
Department	**Avg. Productivity**	**SD Productivity**	**Avg. Team Size**	**Avg. Overtime**	**Avg. Idle Time**	**Avg. Incentive**	**Avg. SMV**	**N Obs.**
finishing	0.75	0.20	10.25	1,917.15	0.00	29.64	3.89	506
sweing	0.72	0.15	52.45	6,508.21	1.26	44.48	23.25	691

Source: own elaboration.

Figure 3: Productivity by team and department



Source: own elaboration.

Furthermore, a productivity gap exists: if teams are grouped into those below target, those that meet the target, and those that exceed it, it becomes clear that teams exceeding the target tend to be smaller, record no idle time, and receive the highest incentives. Conversely, teams that do not reach the target are even smaller, with higher inactivity and fewer incentives. Those that merely meet the target do so primarily through volume and time pressure, rather than real efficiency.

Table 3: Performance group summary

Comparative Table by Performance Group				
Performance Group	**Avg. Team Size**	**Avg. Overtime**	**Avg. Idle Time**	**Avg. Incentive**
Below Target	24.86	3,471.24	1.32	10.99
Meets Target	43.76	5,380.97	1.05	42.63
Exceeds Target	30.08	4,307.19	0.00	49.94

Source: own elaboration.

A positive and significant relationship between the level of economic incentives and productivity is confirmed, reinforced by the results of statistical tests. The analysis also shows that a moderate level of overtime can be beneficial, but excessive overtime does not lead to further improvements in performance.

Figure 4: Effect of overtime on productivity

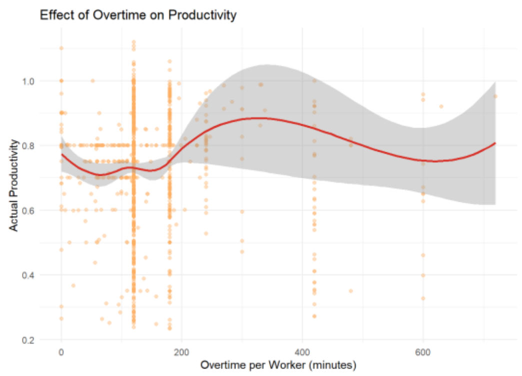
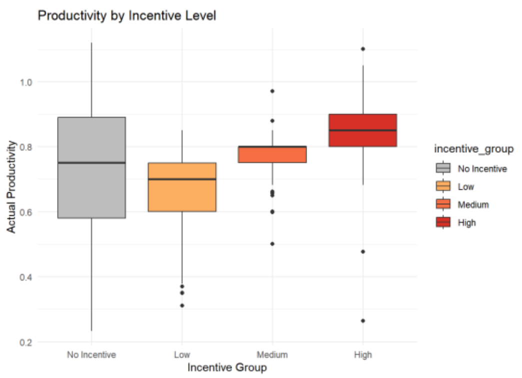


Figure 5: Productivity per incentive level



Source: own elaboration.

2.2 Problem definition

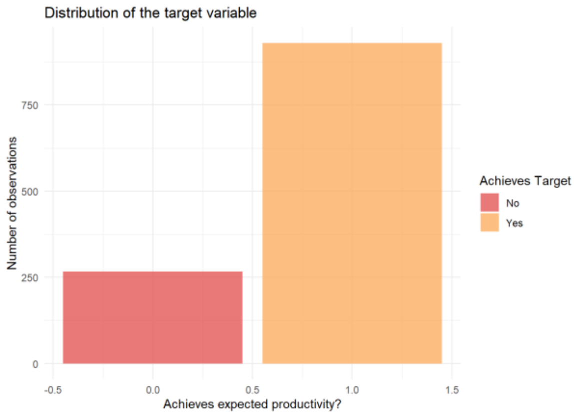
This analysis focuses on identifying which teams achieved their expected productivity by comparing the variable actual_productivity to targeted_productivity. For this purpose, a new binary target variable, achieves_target, was defined: teams that reach or exceed the expected productivity are coded as 1 ("Yes"), while those that do not are coded as 0 ("No").

This approach was chosen for several methodological and practical reasons. Firstly, using a dichotomous variable simplifies the interpretation of results and enables a clearer identification of the factors that distinguish successful teams from those that do not meet the established objectives. This is especially relevant in industrial contexts, where performance monitoring and control typically focus on whether predefined goals are met.

Secondly, the threshold for defining productive success was set at ±0.05 points from the target value. This margin is intended to absorb minor deviations resulting from operational fluctuations, measurement errors, or exceptional circumstances, thereby preventing incorrect classifications due to negligible variations. In this way, only teams whose productivity is clearly below the target are classified as "not achieving" the goal, ensuring a more realistic and robust assessment of productive performance.

It is important to note that, while this binary classification poses challenges such as class imbalance and the potential exclusion of variables with a negative impact on productivity. Nonetheless, this methodology enables a focused analysis of the key factors associated with productive success in a high-intensity manufacturing environment.

Figure 6: Distribution of the target variable (Achieves the target productivity)



Source: own elaboration.

The analysis of this variable shows that 77.7% of the teams achieved the expected productivity level, while the remaining 22.3% did not meet the target. This distribution indicates that most teams are able to adapt to production demands, although there is a significant group that consistently falls short of the goals.

The comparison between both groups reveals significant differences in several key factors. The average team size among those meeting the target is notably higher (37.4 workers versus 24.9; p-value < 0.001), suggesting that larger team sizes may facilitate the achievement of productivity objectives. The analysis of the distributions of economic incentives and overtime (boxplots) also shows that teams achieving the target tend to receive higher incentives and record more overtime, although there is considerable variability within each group.

Overall, these results indicate that team size, the level of economic incentives, and overtime worked are variables associated with the achievement of productivity targets.

However, the internal variability observed within each group highlights the influence of other organizational and contextual factors.

2.3 Method for prediction

To carry out the prediction, machine learning models were selected with the aim of balancing predictive accuracy, interpretability, and computational efficiency. The algorithms selected were Random Forest and XGBoost. Random Forest builds multiple decision trees using random samples of the data and variables, offering robustness against overfitting and allowing for interpretation of variable importance. XGBoost, a boosting algorithm, builds trees sequentially to correct previous errors, excels at capturing complex relationships, and includes regularisation to prevent overfitting. Both methods are well suited for structured data and provide interpretable insights into which factors most influence outcomes.

The literature review included studies that used deep neural networks for productivity prediction (Imran et al., 2019). However, these models require a much larger volume of data to be effective, and in this case, the limited size of the dataset could lead to overfitting and poor generalization. Furthermore, more complex ensemble methods were not employed in this work. Although such approaches may enhance predictive performance in purely competitive or forecasting contexts, the main objective here is to interpret and explain the factors associated with performance. Therefore, the use of individual models, which are more easily interpretable, is more appropriate for extracting useful conclusions to inform decision-making in the industrial context analyzed.

The final selection of variables for the predictive models was informed by a combination of exploratory data analysis (EDA), correlation analysis (see Figure 9 in the Annex), and principal component analysis (PCA), as well as iterative empirical testing of model performance. Initially, the relationship and importance of each candidate variable were assessed using correlation matrices and PCA to capture both direct associations and their overall contribution to variance in the dataset. Subsequently, variables were added and removed in a stepwise manner, empirically evaluating their impact on predictive accuracy and model robustness.

Crucially, no variables that are directly dependent on, or contain information from, the target variable were included as predictors. This decision was made to prevent information leakage and to ensure that the model's performance metrics reflect true predictive capacity, rather than being artificially inflated by circularity or redundancy.

The final set of predictors comprises variables that capture the temporal structure, workload and pressure, organizational characteristics and inefficiencies.

Table 4: Final selection of variables

Description of Variables Used in Final Predictive Modeling		
Variable	**Description**	**Type**
date	Observation date (captures temporal effects and seasonality)	Date
log_over_time	Log-transformed accumulated overtime (workload indicator)	Numeric (log)
department	Department (categorical: sewing or finishing)	Categorical
smv	Standard Minute Value (task complexity and planning)	Numeric
no_of_style_change	Number of style changes (workflow variability)	Numeric
log_incentive	Log-transformed economic incentives (motivation/proxy for bonuses)	Numeric (log)
log_wip	Log-transformed Work In Progress (bottlenecks/flow indicator)	Numeric (log)
no_of_workers	Number of workers in the team (team size/structure)	Numeric
idle_ratio	Proportion of idle time over total available time (inefficiency indicator)	Numeric
achieves_target	Target variable: did the team achieve expected productivity? (1=Yes, 0=No)	Binary

Source: own elaboration.

In this analysis, group-based cross-validation using the team variable was not applied, as the team identifier does not represent a consistent or stable grouping throughout the dataset. As revealed in the exploratory analysis, teams do not appear for the same number of days, and team composition can fluctuate within and between departments over time. This inconsistency means that stratifying or grouping by team during cross-validation would neither ensure meaningful separation of data nor prevent information leakage, since teams are not temporally coherent entities.

Instead, the validation strategy chosen preserves the chronological order of observations by applying a temporal split to the dataset. This approach prevents look-ahead bias and provides a more realistic simulation of operational forecasting, where predictions must be based only on information available up to a certain point in time. By respecting the time sequence, the evaluation remains robust and aligns with the real-world context in which productivity predictions would be deployed in a manufacturing environment.

Specifically, the data was split as follows:

- Training set: 2015-01-01 to 2015-02-15

- Validation set: 2015-02-16 to 2015-02-28
- Test set: 2015-03-01 to 2015-03-11

2.4 Evaluation metrics

To rigorously assess model performance, a set of standard supervised learning metrics was used. In this analysis, class 0 corresponds to “not achieving the productivity target” (positive for underperformance), and class 1 to “achieving the productivity target” (negative class). This labelling is essential for correctly interpreting metrics such as specificity and sensitivity in the results and confusion matrices.

Key evaluation metrics include:

- Accuracy: Proportion of all correctly classified cases. Useful as a global measure, but potentially misleading when classes are imbalanced.
- Sensitivity (Recall) for Class 0: True positive rate for underperforming teams; shows the model’s ability to detect teams that do not reach their targets, supporting early intervention.
- Specificity (Recall) for Class 1: True negative rate for teams meeting productivity targets; important for minimising false positives.
- Balanced Accuracy: The average of sensitivity and specificity, offering a fair measure of model performance in the presence of class imbalance.
- Precision (Positive Predictive Value) and Negative Predictive Value: Indicate how often positive or negative predictions are correct, reflecting the practical reliability of the model.
- Kappa Statistic: Quantifies the agreement between predictions and actual results, beyond what would be expected by chance.
- Area Under the ROC Curve (AUC): Summarises the trade-off between sensitivity and specificity across thresholds, providing an overall measure of the model’s discriminative ability.

To optimize classification performance and balance the trade-off between sensitivity and specificity, the decision threshold was not set arbitrarily. Instead, it was determined using the ROC curve on the validation set, selecting the threshold that maximizes the sum of sensitivity and specificity ("best" threshold). While various manually selected thresholds were tested during model evaluation, the automatically determined threshold consistently provided the best overall performance.

These metrics are widely adopted in the literature on industrial productivity prediction and imbalanced classification tasks (Chawla et al., 2002; Saito & Rehmsmeier, 2015).

3. Results and discussion

Both the XGBoost and Random Forest models exhibited strong and comparable performance in classifying teams according to whether they achieved their expected productivity targets.

- XGBoost produced an overall accuracy of 81.6%, a balanced accuracy of 0.82, and an AUC of 0.80. Sensitivity for the positive class (teams that achieved the target) was 0.81, while specificity was also high at 0.83. The Kappa statistic (0.52) indicates moderate agreement beyond chance. These metrics show that XGBoost provides a well-balanced ability to correctly identify both high-performing and underperforming teams, with a particularly good trade-off between sensitivity and specificity.

- Random Forest achieved a slightly higher overall accuracy of 83.0%, with a balanced accuracy of 0.77 and a higher AUC of 0.87. Its sensitivity for the positive class was 0.87, indicating a very strong capacity to detect teams that reach their productivity goals. However, its specificity was lower (0.68), reflecting a greater tendency to misclassify some teams that did not meet the target as successful. The Kappa statistic (0.49) also indicates moderate agreement.

Table 5: Models evaluation metrics

Model Comparison Table

Model	Sensitivity (class 1)	Specificity (class 1)	Balanced Accuracy	AUC	Accuracy	Kappa
XGBoost	0.81	0.82	0.82	0.82	0.82	0.52
Random Forest	0.87	0.68	0.77	0.87	0.83	0.49

Source: own elaboration.

The comparison reveals that both models deliver robust and consistent results. Random Forest excels in sensitivity (0.87) and overall accuracy (0.83), making it particularly effective for identifying teams that meet or exceed productivity targets, a key consideration in settings where recognising top performers is crucial. XGBoost, on the other hand, achieves higher specificity (0.82) and balanced accuracy (0.82), meaning it is slightly better at correctly identifying both achievers and non-achievers.

In both models, the area under the ROC curve (AUC 0.82–0.87) confirms the strong discriminative power of the predictions, which is in line with previous research in industrial productivity settings. The minor differences in model performance validate the reliability of the results and reinforce the soundness of the modelling strategy.

Importantly, the consistently high sensitivity and balanced accuracy observed across both approaches underscore their practical suitability for real-world applications where the early identification of productivity outcomes is a priority. The slightly reduced specificity in Random Forest is considered acceptable, given the main objective of maximising the detection of successful teams for targeted intervention and process improvement.

The predictive performance of both models varies depending on the department. In the sewing department, an AUC of 0.88 reflects high accuracy and stable relationships between variables, indicating more homogeneous processes. In contrast, the finishing department shows a lower AUC of 0.82, which suggests greater variability and the possible influence of unobserved factors, likely due to the less standardised and more variable nature of finishing tasks.

The analysis of variable importance confirms that both XGBoost and Random Forest consistently identify the same key factors underlying productivity achievement among work teams. Across both models, work in progress and incentives emerge as the most influential predictors. Log_wip reflects the accumulated workload and production pressure, while log_incentive captures the motivating effect of financial rewards on team performance, underscoring the centrality of incentives for meeting productivity goals.

Additionally, Standard Minute Value, which measures task complexity and planned effort, is also highly relevant in both models, highlighting the importance of planning and technical challenge as performance drivers. Variables such as overtime hours and team size provide secondary but still meaningful explanatory power. In contrast, features like department, the number of style changes, or idle ratio have relatively low predictive value.

Taken together, these findings indicate that productivity success is primarily determined by economic incentives, effective workload management, and the appropriate allocation of complex tasks, while other organisational or operational factors play a more limited role.

Potential Improvements and Future Work:

While the current analysis provides robust results and actionable insights, there are several avenues for potential improvement in future research:

- Temporal Cross-Validation for Hyperparameter Tuning: In this study, model selection and hyperparameter tuning were performed using a fixed temporal split to simulate a realistic forecasting scenario. However, implementing temporal cross-validation, such as rolling or sliding windows, would provide a more reliable and unbiased estimate of model performance over time. The approach is especially relevant in time-dependent production environments and could lead to more robust hyperparameter selection.

- Incorporation of Additional Features: The analysis could be enhanced by integrating new variables, such as qualitative indicators (e.g., worker satisfaction, management interventions, or external shocks), which might help explain unexplained variance in productivity, particularly in departments with more unpredictable outcomes.

- Advanced Handling of Class Imbalance: Future work could explore advanced resampling techniques (SMOTE or ADASYN) or cost-sensitive learning to further improve the detection of underperforming teams.

- Given that Work in Progress (WIP, log_wip) emerged as one of the most influential predictors in the models, particular attention should be paid to the treatment of its missing values. Since the imputed values did not fully capture the distribution of the original data and the proportion of missing entries was relatively high, future research should prioritize refining the imputation strategy.

4. Conclusion

This study set out to explain the reasons behind why some teams in high-intensity manufacturing environments fail to achieve their productivity targets, guided by two main hypotheses: (1) that underperforming teams differ in organisational, motivational, and operational aspects compared to successful teams, and (2) that factors such as idle time, the absence of economic incentives, and excessive workloads are negatively associated with productivity.

The results confirm both hypotheses and provide a clear picture of what differentiates high- and low-performing teams. Teams that failed to meet their targets were consistently smaller, received fewer incentives, had higher idle time, and operated with less organisational structure. By contrast, the most successful teams tended to be larger, benefited from stronger incentive schemes, and worked in more structured environments. Notably, economic incentives emerged as the most decisive single factor: teams with higher incentives were significantly more likely to achieve or exceed their targets. Other influential variables included the volume of work in progress (log_wip), which represents ongoing workload and production pressure, and task complexity (SMV), reflecting the importance of planning and technical challenge.

The application of both XGBoost and Random Forest models showed that these machine learning approaches are robust and reliable for this type of analysis, with both models achieving strong and similar predictive performance.

From a broader perspective, the analysis of variable importance deepens our understanding of the drivers behind productivity. Both models highlight that economic incentives and the management of work in progress are fundamental for team performance, whereas organisational aspects such as department or the number of style changes have a secondary impact. Interestingly, the data also show that simply increasing overtime does not lead to better results, and even small amounts of idle time are linked to lower productivity, underscoring the importance of effective workload allocation and operational discipline.

In discussion, these findings suggest that interventions aimed at improving productivity in industrial settings should prioritise well-designed incentive systems and the optimal distribution of workloads. At the same time, attention should be given to the organisational structure and the complexity of tasks assigned, as these can either support or undermine team performance. The results reinforce the need to look beyond surface-level metrics and focus on the underlying mechanisms (likes motivation, team configuration, and workflow management) that ultimately determine whether productivity goals are met.

In summary, this research demonstrates the value of explainable machine learning models for not only classifying productivity outcomes but, crucially, for uncovering the underlying drivers of performance in complex industrial environments.

BIBLIOGRAPHY

20minutos. (2024, May 16). *Bangladesh: Revueltas sociales tensan la fábrica de ropa barata del mundo*.

Ahn, C., Lee, S., & Peña-Mora, F. (2013). A case study: Idle time ratio for construction equipment. *Journal of Construction Engineering and Management*, 139(5), 572–580.

Balla, I., Rahayu, S., & Jaya Purnama, J. (2021). Garment employee productivity prediction using Random Forest. *Techno Nusa Mandiri: Journal of Computing and Information Technology*, 18(1), 49-54.

Chan, J., & Siu, K. (2010). Analyzing exploitation: Excessive overtime and piece-rate work in Chinese export factories. *Critical Asian Studies*, 42(2), 167–190.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357.

Feng, W., Huang, W., & Ren, J. (2018). Class imbalance ensemble learning based on the margin theory. *Applied Sciences*, 8(5), 815.

Imran, A. A., Amin, M. N., Islam Rifat, M. R., & Mehreen, S. (2019). Deep neural network approach for predicting the productivity of garment employees. In *2019 6th*

International Conference on Control, Decision and Information Technologies (CoDIT) (pp. 1402–1407). Paris, France.

Imran, Abdullah Al, Rahim, Md Shamsur, & Ahmed, Tanvir. (2021). Mining the productivity data of the garment industry. *International Journal of Business Intelligence and Data Mining*, 19(3), 319–342.

Kassambara, A. (2018). *Machine learning essentials: Practical guide in R*. STHDA.

Marx, K. (1867). *Capital: A Critique of Political Economy* (Vol. I, Ch. 9).

Obiedat, R., & Toubasi, S. (2022). A combined approach for predicting employees' productivity based on ensemble machine learning methods. *Informatica*, 46(5), 49–58.

OECD. (2001). *Measuring Productivity – OECD Manual: Measurement of Aggregate and Industry-level Productivity Growth*. OECD Publishing, Paris.

Patnaik, U. (1972). Class differentiation within the peasantry: An approach to the analysis of Indian agriculture. *Economic and Political Weekly*, 7(31/33), 1239–1245.

Productivity Prediction of Garment Employees [Dataset]. (2020). UCI Machine Learning Repository.

Ramos J. (2024). Explaining Productivity in the Garment Industry: An Analytical and Machine Learning Approach. Available at Github:

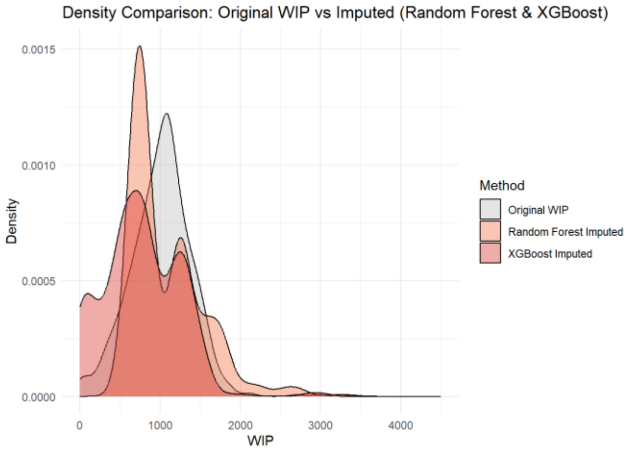
https://github.com/jjorgeramos/Final-Thesis---Jorge-Ramos---Computational-Social-Sciences

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432.

Tricontinental Institute for Social Research. (2019). *The rate of exploitation: The case of the iPhone*.

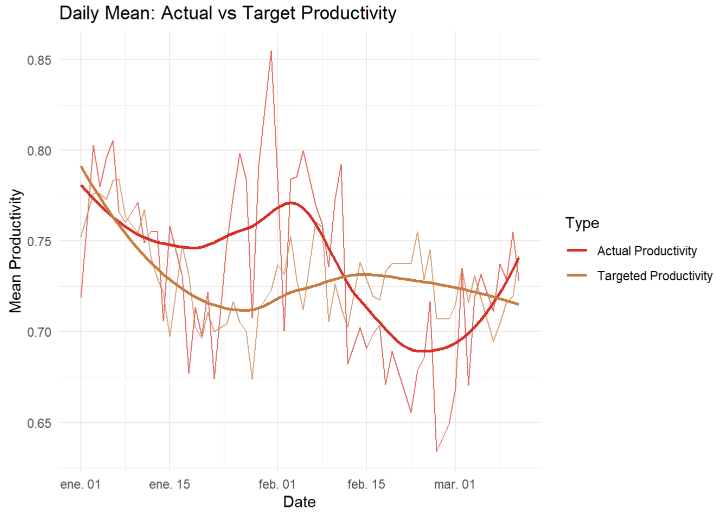
ANNEX

Figure 7: WIP imputation



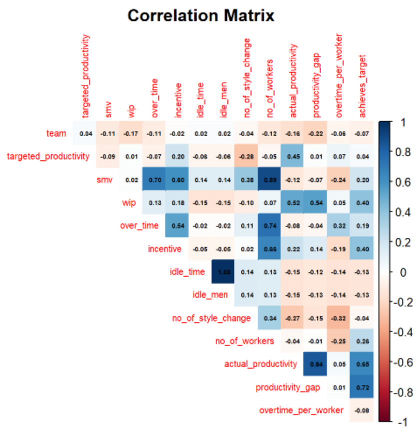
Source: own elaboration.

Figure 8: Daily actual and target productivity



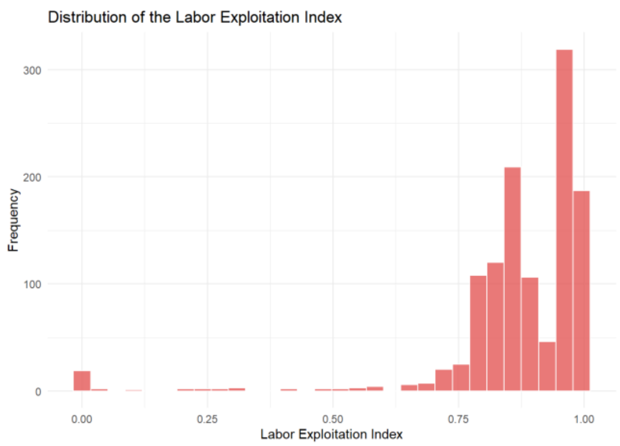
Source: own elaboration.

Figure 9: Correlation Matrix



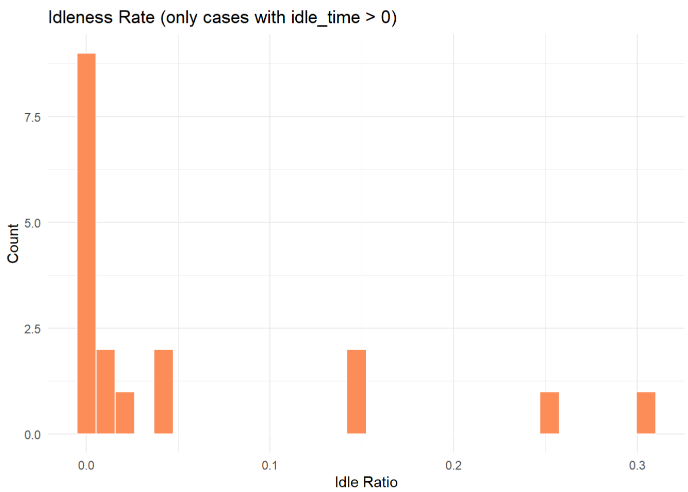
Source: own elaboration.

Figure 10: Labor Explotaiton Index



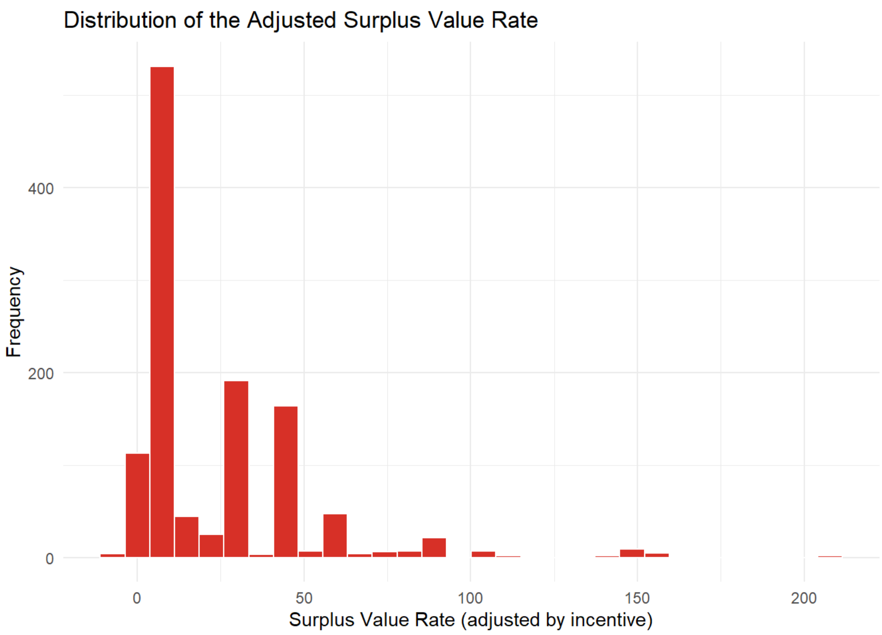
Source: own elaboration.

Figure 11: Idleness Rate



Source: own elaboration.

Figure 12: Surplus Value Rate



Source: own elaboration.