Predicting & Optimizing the Productivity of Garment Employees
A Novel Analytics Design to Resource Allocation & Efficiency

PURDUE UNIVERSITY
Mitchell E. Daniels, Jr. School of Business

Harnessing Data to Boost Efficiency in the Garment Industry

Matthew A. Lanham | lanhamm@purdue.edu
Rishita Korapati | rkorapat@purdue.edu

ABSTRACT

Accurate prediction of productivity can transform resource allocation, enhance operational efficiency, and meet rapid production demands, ensuring cost-efficiency and quality. According to **recent studies by Rahim et al. (2017)** and **Imran et al. (2019),** inefficiencies in the garment industry result in significant economic losses, highlighting the need for predictive modeling to optimize productivity.

BUSINESS PROBLEM

Bangladesh, India, Vietnam, Cambodia. In bustling factories and vibrant looms, the garment industry **employs millions**, contributing **84%** of export earnings in Bangladesh alone. However, low productivity **costs billions** annually due to inefficiencies and soaring costs. Therefore, the need for predictive modeling persists in the garment industry to improve productivity, reduce costs, and maintain competitiveness in the global market.

Research Questions

- Can we develop **an accurate linear predictive mode**l that captures the relationship between garment employee operational drivers?
- Could we create **an optimization model** that integrates this linear predictive model to **assist garment industry managers** in **enhancing operational efficiency**?

ANALYTICAL PROBLEM

Analytical Context: Enhance prediction accuracy using AA methods and optimization to improve planning and resource allocation with precise forecasts.

Challenges: High variance in productivity data due to factors like worker fatigue and machine breakdowns and integrate data from multiple sources for a comprehensive analysis.

Success Metrics: Measure reduction in forecast error, e.g., decrease MAE from **0.15 to 0.086**, and increase operational efficiency and profitability through accurate predictions.

LITERATURE REVIEW

In our comparison of methodologies with those employed in similar studies, we identified **a significant methodological gap** within the academic literature. The industry has focused **extensively on classification, regression, and time-series** models for predictive purposes to identify trends in metrics and how a model is of best fit. However, This gap highlights the need for more **robust frameworks** that can integrate these predictive models into **comprehensive optimization frameworks** to fully **realize** their **potential** for **improving productivity.**

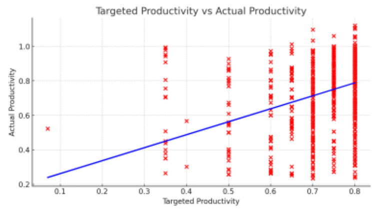
Reference	Evaluation & Selection Method	Forecasting Models	Evaluation Criteria	Decision Maker Metrics
(IEEE, 2021)	Classification	Classification	MAE, Robustness	Best MAE, Robustness
(Imran et al., 2019)	Deep Neural Network	Regression	MSE, MAE, MAPE	Low MSE, MAE, MAPE
(Para et al., 2019)	Time-Series, Regression, Classification	Time-Series, Regression, Classification	Productivity	Improved Productivity
(Rahim et al., 2017)	Structured Methodology	Time-Series, Regression, Classification	Productivity, Waste Reduction	Enhanced Productivity, Reduced Waste

DATA DICTIONARY

Variable	Unit of Measure	Definition	Example Value	Min	Max	Predictive Model Perspective	Process Perspective
date	Categorical	Date of observation	1/1/15	-	-	ID	Data Collection
quarter	Categorical	Quarter of the year	Quarter 1	-	-	Input	Data Collection
idle_men	Numerical	Number of idle workers	5	0	10	Input	Operational
idle_time	Numerical	Idle time in minutes	30	0	100	Input	Operational
incentive	Numerical	Incentive given	98	0	100	Input	Operational
no_of_style_change	Numerical	Number of style changes	2	0	5	Input	Operational
no_of_workers	Numerical	Number of workers in the team	59	5	100	Input	Operational
over_time	Numerical	Overtime in minutes	7080	0	10000	Input	Operational
smv	Numerical	Standard Minute Value	26.16	0	30	Input	Operational
wip	Numerical	Work In Progress	1108	0	1500	Input	Operational
day	Categorical	Day of the week	Thursday	-	-	Input	Organizational
department	Categorical	Department of the factory	Sewing	-	-	Input	Organizational
targeted_productivity	Numerical	Targeted productivity	0.8	0.5	1	Input	Organizational
team	Numerical	Team number	8	1	12	Input	Organizational
actual_productivity	Numerical	Actual productivity achieved	0.8	0.5	1	Target	Performance

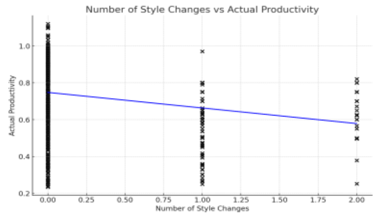
Targeted Productivity vs Actual Productivity: The red scatter plot shows **a positive correlation**, indicating that **higher targeted productivity** generally corresponds to higher actual productivity.

- **Actual Productivity = 0.75 * (Targeted Productivity) + 0.19**
- **Correlation = 0.42**



Number of. Style Changes vs Actual Productivity: This scatter plot demonstrates **a negative linear relationship**, suggesting that as the number of style changes **increases**, actual productivity tends to **decrease.**

- **Actual Productivity = 0.085 * (Number of Style Changes) + 0.75**
- **Correlation = -0.207**



PROJECT METHODOLOGY

1. Data Understanding

- Data from Kaggle
- Dictionaries
- Descriptive Statistics

2. Exploratory Data Analysis (EDA)

- Observe Relationships Using Histograms, Boxplots, and Scatterplots

3. Data Preprocessing

- Handle missing values (replaced with median)
- Outlier Detection (Z-score method)

4. Feature Engineering

- Use Feature Scaling to standardize features except Linear Regression.
- Create new features (targeted vs actual)
- Generated dummy variables

5. Predictive Modeling

- Data Partitioning
- Implement Linear Regression, Random Forest, Gradient Boosting, & XGBoost

6. Optimization Modeling

- Predict Productivity by evaluating performance metrics (Accuracy, Precision, Sensitivity & RMSE) and feature importance plots.

LANGUAGE I USED:

Fig 3. SEMMA Roadmap (AI image generated by DALLE)

REGRESSION MODELING

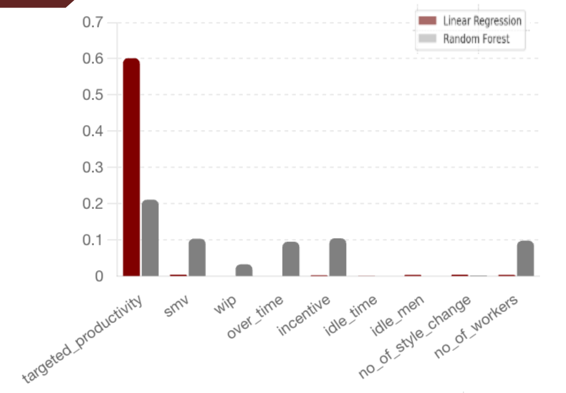
Assumptions: Full shipments match purchase orders for on-time, complete delivery. Consistent working conditions observed.

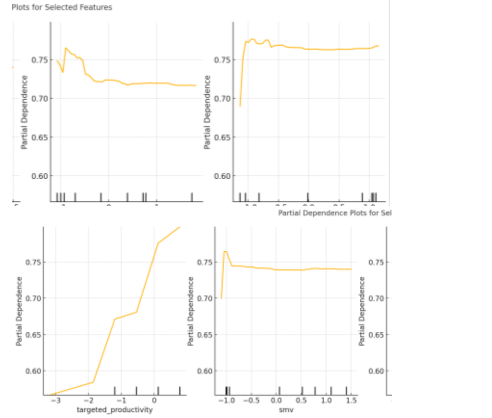
Split Data: Divide data into training and testing sets & train various ML models: Linear Regression, Random Forest, and Gradient Boosting Models. I also used Linear Regression with non-standardized inputs to interpret the effects of each input on the output.

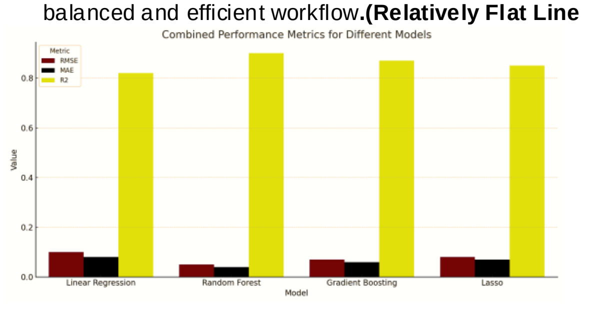
Targeted Productivity is the most important predictor of actual productivity in both models. Other significant features include **SMV and incentives,** which also play a role in motivating workers leading to better performance. **Lastly,** there is **the overtime and number of workers** which are notable features that may increase output but too much can lead to problems**.**

Targeted Productivity is crucial for setting realistic goals and optimizing task efficiency.(**Strong positive relationship,)**

- **SMV** shows a **flat relationship** indicating a consistent but small impact on productivity.
- **Overtime** & Incentives should be managed carefully to avoid diminishing returns.**(Also Flat relationship)**
- **No. of Workers** should be optimized to ensure a balanced and efficient workflow.**(Relatively Flat Line**







The Bar Chart for Model Performance Metrics:

Random Forest has the **best predictive power and accuracy**, evidenced by its **low RMSE, low MAE, and high R².** In contrast, Linear Regression and Lasso exhibit **higher RMSE, moderate MAE, and lower R²**, indicating **less accurate predictions**, while Gradient Boosting performs well but **not as strongly** as Random Forest.

Linear regression Summary

Intercept: When **SMV and Overtime** are **zero**, the predicted productivity is **0.799.**

SMV: For **each unit increase** in **SMV**, the productivity **decreases** by **0.0011**, holding other factors constant.

Overtime / WIP: The coefficient is **very small 0.000003** and **not statistically significant (P-VALUE > 0.05)**, indicating that overtime has a **negligible impact** on productivity in this model.

Incentive: Each dollar **increase** in incentives raises productivity by **0.001762.**

Number of Style Changes & Number of Workers: Each additional style change and worker **decreases** productivity by **0.022399**. and **0.002264 respectively.**

Optimization Modeling

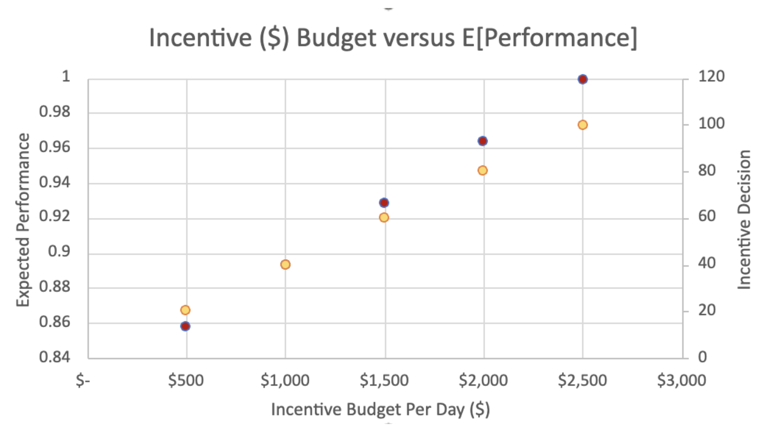
Predicted Productivity = 0.799 + 0.0011 (SMV) – 0.0 (WIP) – 0.0 (Overtime) + 0.0018 (Incentive) – 0.022 (# of Style Changes) – 0.0023 (# of Workers)

0 <= smv <= 30
0 <= wip <= 1500
0 <= overtime <= 10,000
0 <= incentive <= 100

0 <= # of style changes <= 5
0 <= # of workers <= 100
(# of workers * $5 * incentive) <= 1200
# of workers and style changes are binary

Objective Function Value
0.858

MODEL RESULTS



KEY FINDINGS

- The linear regression model provided **clear coefficients** for each variable, highlighting the **significant impact of incentives** on productivity.
- The optimization model demonstrated that **adjusting the incentive budget allowance** can significantly **improve productivity**, with a **$1,500** daily incentive achieving **92.8%** productivity on average.
- Despite **the low R²**, the linear regression model's insights are **valuable** for **practical optimization** and **resource allocation.**

SUMMARY OF FINDINGS

The predictive and optimization models exhibited **outstanding accuracy, precision, and sensitivity** in identifying key productivity drivers. This empowers targeted enhancements, **driving significant improvements** in **efficiency** and **performance** within the garment industry. Stakeholders—garment factories, teams, trainers, and financial analysts—can leverage these insights to **optimize productivity** and dramatically **elevate overall** business **performance**.

RECOMMENDATIONS

Introduce flexible incentive schemes, such as a $1,500 daily incentive, to motivate employees and enhance productivity.

Implement optimized workflow and resource allocation strategies to minimize idle time and keep workers engaged.

Conduct regular training sessions to enhance workforce skills and efficiency. Align work hours with employees' productivity peaks through flexible scheduling.

Continuously monitor and adjust the model to account for new data and changing conditions.

ACKNOWLEDGEMENTS

I express my gratitude to Professor Matthew Lanham for his invaluable guidance and support throughout the project. Special thanks to the data providers (Kaggle) who contributed to the successful completion of this project. Their insights and expertise were instrumental in the successful completion of this project