Garment Worker Productivity Prediction

1.Introduction

The Garment Industry is one of the key examples of the industrial globalization of this modern era. It is a highly labour-intensive industry with lots of manual processes. Satisfying the huge global demand for garment products is mostly dependent on the production and delivery performance of the employees in the garment manufacturing companies. So, it is highly desirable among the decision makers in the garments industry to track, analyse and predict the productivity performance of the working teams in their factories. This dataset can be used for regression purpose by predicting the productivity range.

1.1 Project Overviews:

The garment industry is one of the largest industries in the world, and garment worker productivity is a crucial factor in determining the success and profitability of a company. The development of an accurate garment worker productivity prediction model using machine learning can have significant implications in various domains, including manufacturing, human resources, and supply chain management. This model can help companies identify the factors that affect worker productivity and take corrective actions to improve efficiency, reduce costs, and enhance their competitiveness in the market.

1.2 Objectives:

In this project, we aim to develop a machine learning model that predicts the productivity of garment workers based on a given set of features. Our dataset contains information on various attributes of garment production, including the quarter, department, day, team number, time allocated, unfinished items, over time, incentive, idle time, idle men, style change, number of workers, and actual productivity. We will use this dataset to train and evaluate our predictive model.

2. Project Initialization and Planning Phase

The "Project Initialization and Planning Phase" marks the project's outset, defining goals, scope, and stakeholders. This crucial phase establishes project parameters, identifies key team members, allocates resources, and outlines a realistic timeline. It also involves risk assessment and mitigation planning. Successful initiation sets the foundation for a well-organized and efficiently executed machine learning project, ensuring clarity, alignment, and proactive measures for potential challenges.

2.1 Define Problem Statement:

The manager of a garment industry is trying to improve the profitability of the company but facing challenges to predict the productivity of the workers due to the need of consideration of various attributes, thus realizing the need of deploying a machine learning model for accurate prediction.

Problem Statement Report: https://github.com/Riddhib15/Garment-Worker-Productivity-Prediction/blob/main/1.%20Project%20Initialization%20and%20Planning%20Phase/Define%20Problem%20Statements%20Template.pdf

2.2. Project Proposal (Proposed Solution)

The garment worker productivity prediction model uses machine learning to boost efficiency and accuracy. It has the potential to improve the lives of garment workers by promoting fair and efficient workforce management practices. It can also help manufacturers optimize their workforce management strategies, reduce idle time, and increase productivity. This initiative aligns with our

main objective to enhance decision-making, reduce risks, and streamline operations, ultimately improving worker satisfaction, profitability and operational efficiency.

Project Proposal Report: https://github.com/Riddhib15/Garment-Worker-Productivity-Prediction/blob/main/1.%20Project%20Initialization%20and%20Planning%20Phase/Project%20Proposal%20(Proposed%20Solution)%20template.pdf

2.3. Initial Project Planning

Initial Project Planning involves outlining key objectives and defining scope. It encompasses setting timelines, allocating resources, and determining the overall project strategy. During this phase, the team establishes a clear understanding of the dataset, formulates goals for analysis, and plans the workflow for data processing. Effective initial planning lays the foundation for a systematic and well-executed project, ensuring successful outcomes.

Project Planning Report: https://github.com/Riddhib15/Garment-Worker-Productivity-Prediction/blob/main/1.%20Project%20Initialization%20and%20Planning%20Phase/Project%20Planning%20Template.pd

3.Data Collection and Preprocessing Phase

The Data Collection and Preprocessing Phase involves executing a plan to gather relevant data from Kaggle, ensuring data quality through verification and addressing missing values. Preprocessing tasks include cleaning, encoding, and organizing the dataset for subsequent exploratory analysis and machine learning model development.

3.1. Data Collection Plan and Raw Data Sources Identified

The dataset for "Garment Worker Productivity Prediction" is sourced from Kaggle. It includes attributes of garment production, including the quarter, department, day, team number, time allocated, unfinished items, over time, incentive, idle time, idle men, style change, number of workers, and actual productivity.

Data Collection Report: https://github.com/Riddhib15/Garment-Worker-Productivity-Prediction/blob/main/2.%20Data%20Collection%20and%20Preprocessing%20Phase_/Raw%20Data%20Sources%20And%20Data%20Quality%20Report%20template.pdf

3.2 Data Quality Report

Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive modelling.

Data Quality Report: https://github.com/Riddhib15/Garment-Worker-Productivity-Prediction/blob/main/2.%20Data%20Collection%20and%20Preprocessing%20Phase_/Data%20Quality%20Report%20template.pdf

3.3. Data Exploration and Preprocessing

Data Exploration involves analysing the dataset to understand patterns, distributions, and outliers. Preprocessing includes handling missing values, scaling, and encoding categorical variables. These

crucial steps enhance data quality, ensuring the reliability and effectiveness of subsequent analyses in the loan approval project.

Data Exploration and Preprocessing Report: https://github.com/Riddhib15/Garment-Worker-Productivity-Prediction/blob/main/2.%20Data%20Collection%20and%20Preprocessing%20Phase_/Data%20Exploration%20and%20Preprocessing%20template.pdf

4.Model Development Phase

The Model Development Phase entails crafting a predictive model for productivity prediction. It encompasses strategic feature selection, evaluating and selecting models, initiating training with code, and rigorously validating and assessing model performance for informed decision-making in the lending process.

4.1 Feature Selection Report

The Feature Selection Report outlines the rationale behind choosing specific features for the predictivity model. It evaluates relevance, importance, and impact on predictive accuracy, ensuring the inclusion of key factors influencing the model's ability to discern productivity of the workers.

Feature Selection Report: https://github.com/Riddhib15/Garment-Worker-Productivity-Prediction/blob/main/3.%20Model%20Development%20Phase/Feature%20Selection%20Report%20template.pdf

4.2 Model Selection Report

The Model Selection Report details the rationale behind Linear Regression Model, Lasso Regression Model, Ridge Regression Model, Decision Tree Regressor Model, Random Forest Regressor Model, Gradient Boosting Regressor Model, Extreme Gradient Boosting Model, Bagging Regressor Model and Boosting Regressor Model for productivity prediction. It considers each model's strengths in handling complex relationships, interpretability, adaptability, and overall predictive performance, ensuring an informed choice aligned with project objectives.

Model Selection Report: https://github.com/Riddhib15/Garment-Worker-Productivity-Prediction/blob/main/3.%20Model%20Development%20Phase/Model%20Selection%20Report%20template.pdf

4.3 Initial Model Training Code, Model Validation and Evaluation Report

The Initial Model Training Code employs selected algorithms on the dataset, setting the foundation for regression modelling. The subsequent Model Validation and Evaluation Report rigorously assesses model performance, employing metrics like accuracy and precision to ensure reliability and effectiveness in predicting loan outcomes.

Model Development Phase Template: https://github.com/Riddhib15/Garment-Worker-Productivity-Prediction/blob/main/3.%20Model%20Development%20Phase/Initial%20Model%20Training%20Code%2C%20Model%20Validation%20and%20Evaluation%20Template.pdf

5. Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced accuracy and efficiency.

5.1 Hyperparameter Tuning Documentation

The Bagging Regressor Model was selected for its superior performance, exhibiting high accuracy during hyperparameter tuning. Its ability to handle complex relationships, minimize overfitting, and optimize predictive accuracy aligns with project objectives, justifying its selection as the final model.

5.2 Performance Metrics Comparison Report

The Performance Metrics Comparison Report contrasts the baseline and optimized metrics for various models, specifically highlighting the enhanced performance of the Bagging Regressor model. This assessment provides a clear understanding of the refined predictive capabilities achieved through hyperparameter tuning.

5.3 Final Model Selection Justification

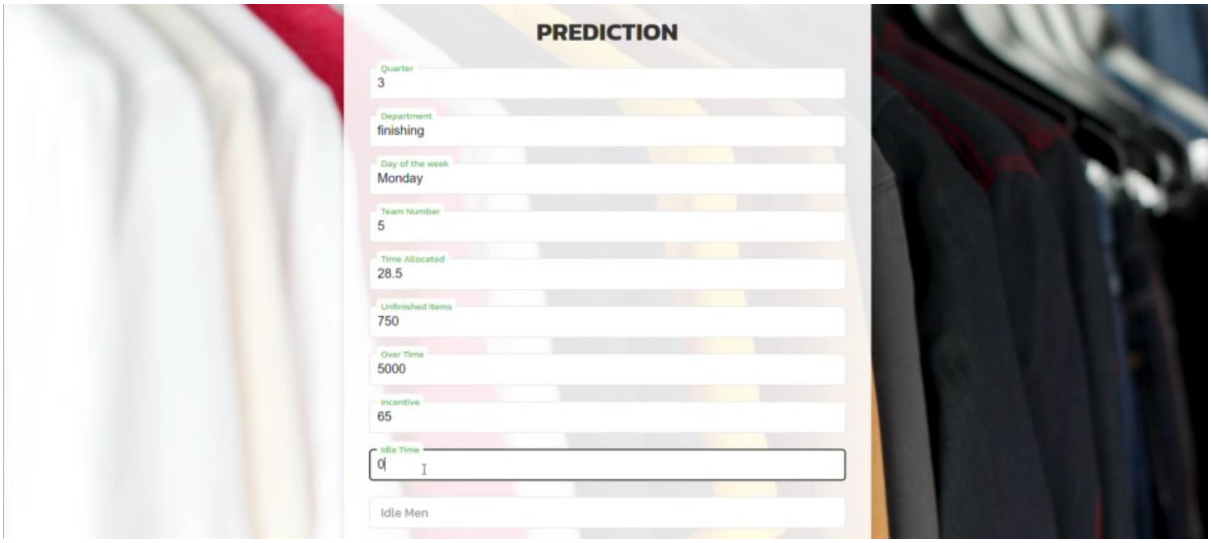
The Final Model Selection Justification articulates the rationale for choosing Bagging Regressor Model as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring optimal loan approval predictions.

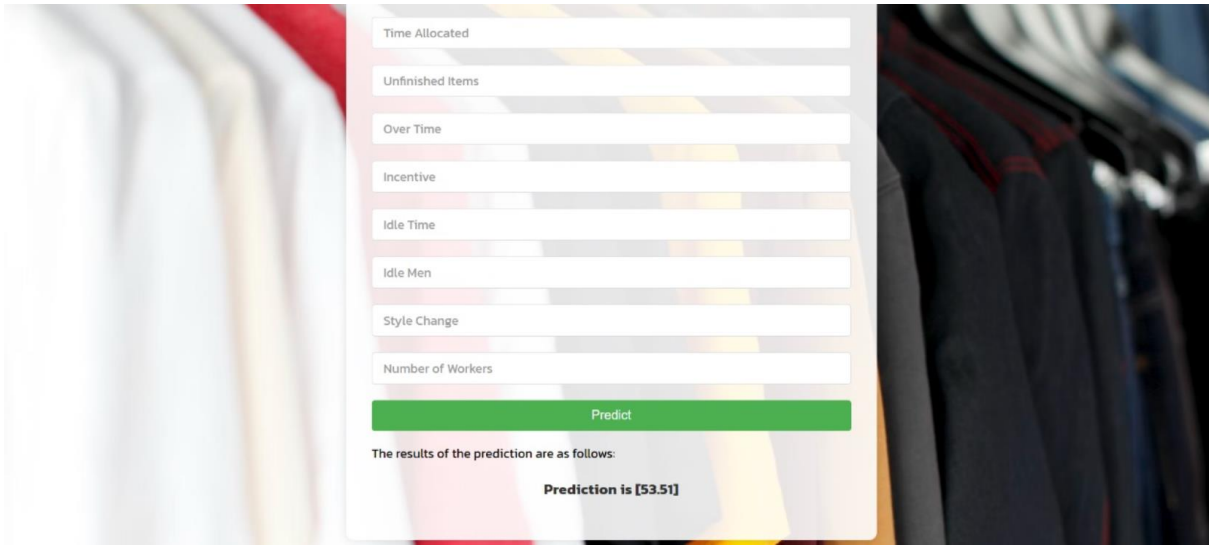
Model Optimization and Tuning Phase Report: https://github.com/Riddhib15/Garment-Worker-Productivity-Prediction/blob/main/4.%20Model%20Optimization%20and%20Tuning%20Phase/Model%20Optimization%20and%20Tuning%20Phase%20Template.pdf

6. Results

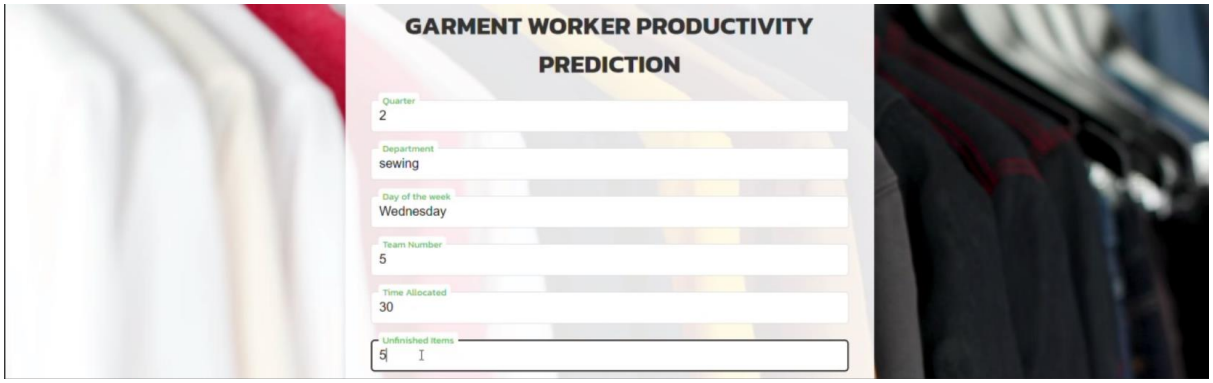
The final output after deployment of flask is given below.

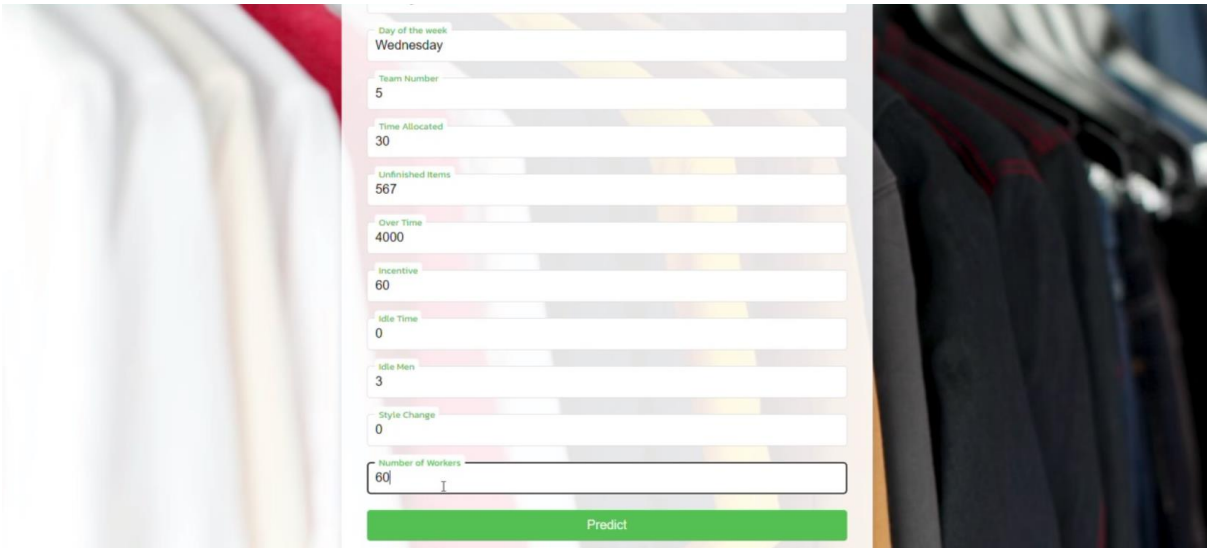
The first prediction example with the prediction given as percentage is:

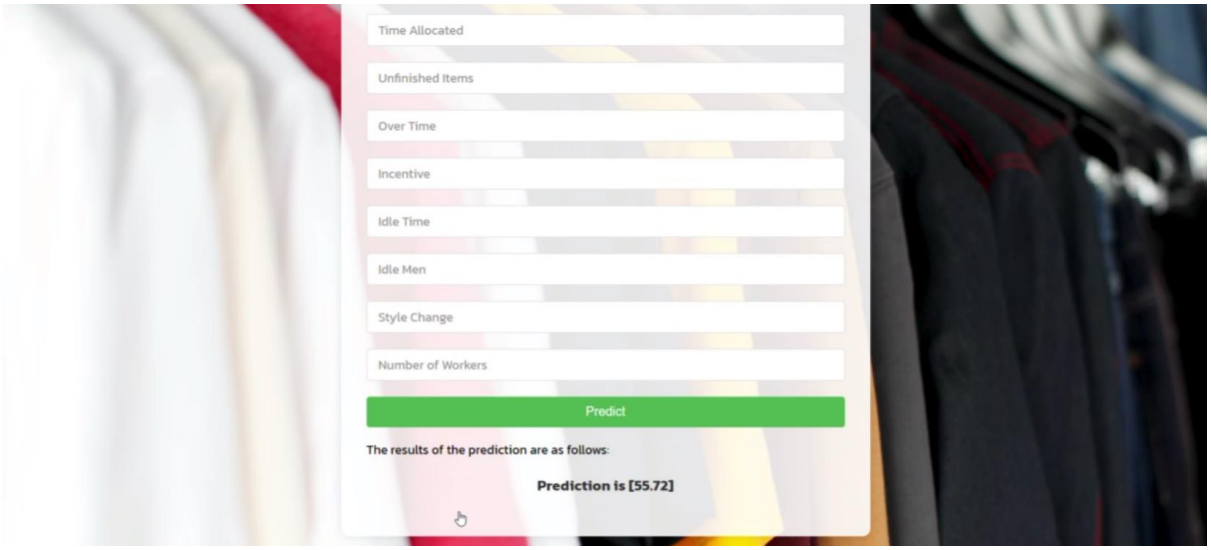




The second example for the given input is as follows:







7. Advantages & Disadvantages

The garment worker productivity prediction model has the potential to improve the lives of garment workers by promoting fair and efficient workforce management practices. It can also have significant implications for the garment manufacturing industry. The model can help manufacturers optimize their workforce management strategies, reduce idle time, and increase productivity. This can lead to higher profits, lower production costs, and improved quality of goods produced.

However, the limitation of the model is that the model has been trained and tested on a very limited dataset, due to the lack of availability of multiple datasets.

8. Conclusion

Thus, correct implementation of the model can lead to higher job satisfaction, better pay, and improved working conditions for garment workers. Also, it will lead to higher profits, lower production costs, and improved quality of goods produced by industry.

9. Future Scope

Additional analysis can be included to identify factors that affect worker productivity and suggest ways to improve working conditions, incentive schemes, and training programs. It can also help manufacturers optimize their workforce management strategies, reduce idle time, and increase productivity.

10. Appendix

10.1. Source Code

Please click on the link given below to view the source code of the project.

Source Code: https://colab.research.google.com/drive/1k1TIutigYG-hMoFIezEL8f_cE2j1QCem?usp=sharing

10.2. GitHub & Project Demo Link

Please click on the respective links to view the GitHub and Project Demonstration link.

GitHub: https://github.com/Riddhib15/Garment-Worker-Productivity-Prediction

Project Demonstration Link: https://drive.google.com/file/d/1lChVe7i_LpgSkJqJDZdBIlnU-05T6zsi/view?usp=sharing