IN SEARCH OF EFFICIENCY

Predicting Garment Employee productivity

By: Tamjid Ahsan

Date: May 27, 2021, 05:00 PM [DST]

Overview

- The garment industry one of the highly labor-intensive industries.
- Often production goals are not met.
- This is a high priority for an organization to achieve deadline and maximize profit by ensuring proper utilization of resources.
- Critical component of Lean manufacturing

Business Problem

- A **garment production pipeline** consists of a handful of **sequential processes**, e.g., designing, sample confirmation, sourcing and merchandising, lay planning, spreading and cutting, sewing, finishing and so on.
- To complete a whole production **within a target time**, these sequential processes need to be to performed **efficiently**. Industrial engineers strategically set a targeted productivity value against each working team in the manufacturing process.
- However, it is a common scenario that the **actual productivity does not align with the target** for several factors, both internal and external.

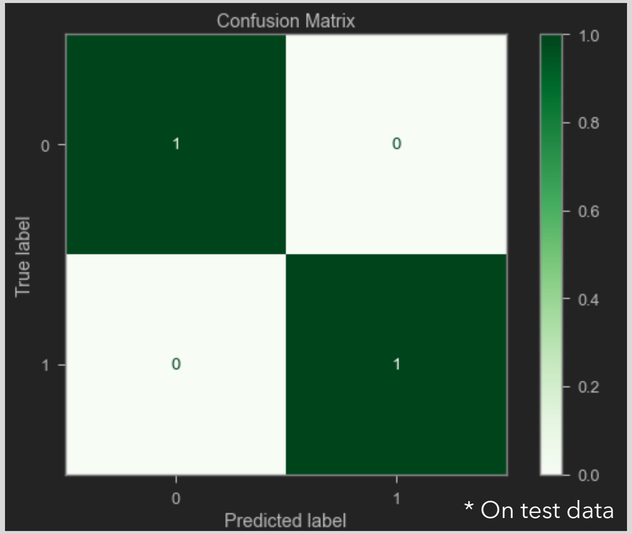
The goal is to predict whether employee met productivity target, with high precision.

Methodology

- The collected dataset contains the production data of the sewing and finishing department for three months from January 2015 to March 2015 of the renowned garment manufacturing company in Bangladesh.
- Cleaned dataset consists of 1197 instances and includes 13 attributes.
- Machine Learning models were used for predicting binary classification.
  - **Target met** OR **not met**.
- Focus was on maximizing prediction of the instances where production target were not met with high confidence.

Methodology - Model

- Using "**random subsampled decision trees**"* modeling technique yielded best performance.
- Accuracy of 100%. (Quite uncommon outcome)

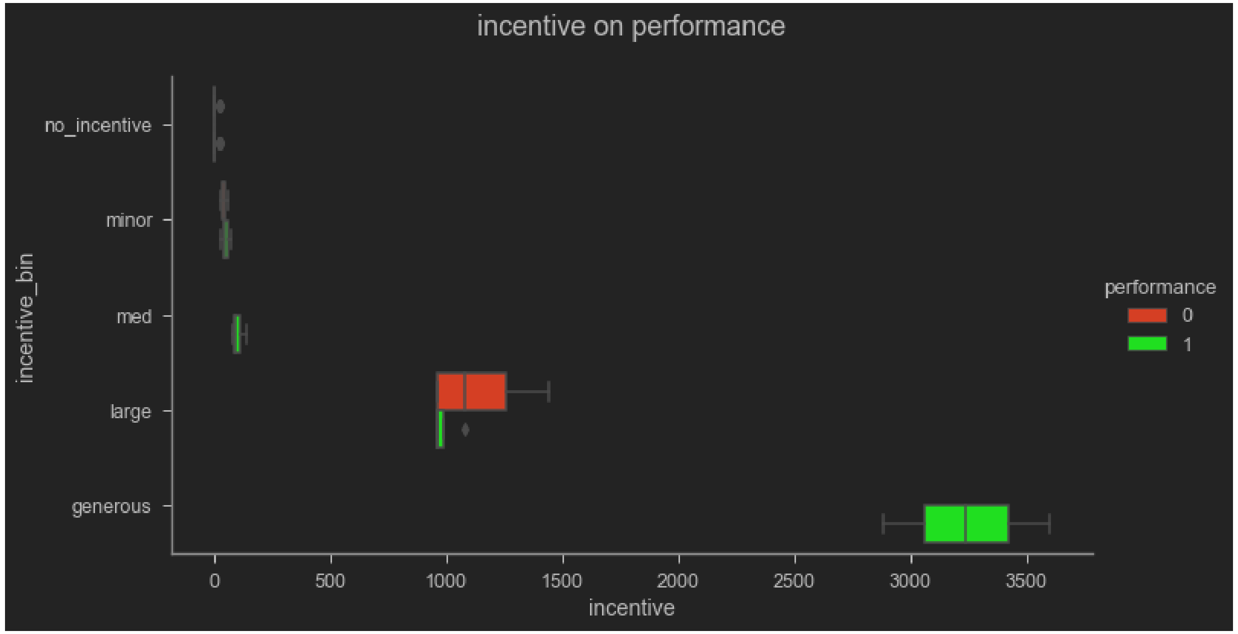


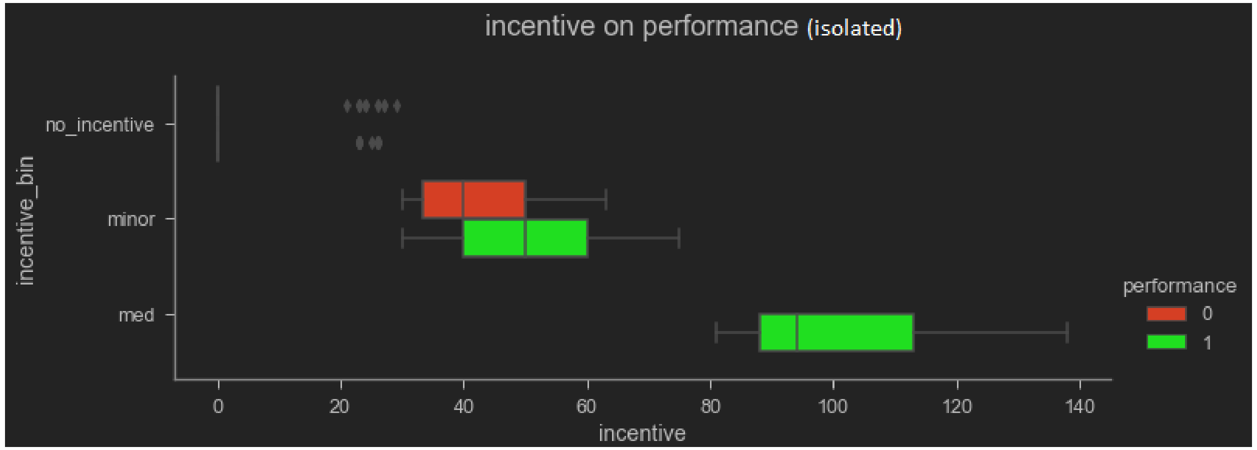
Feature Name
quarter
department
day
team
targeted_productivity
smv
wip
over_time
incentive
idle_time
idle_men
no_of_style_change
no_of_workers
performance

* More details about model is in the appendix

LET'S START WITH SOME EXPLORATION

The Prominent Ones

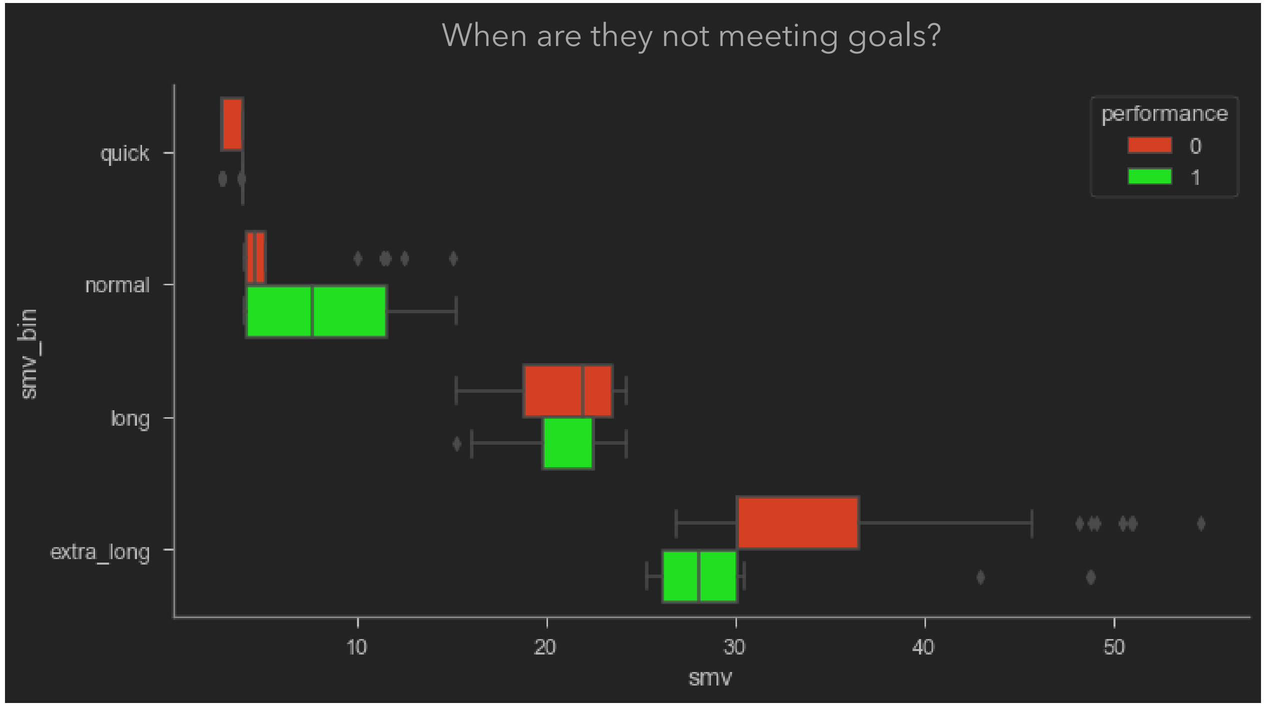




Even a minor incentive have positive impact on performance.

INCENTIVE

TRY TO GIVE INCENTIVE IF POSSIBLE



SMV

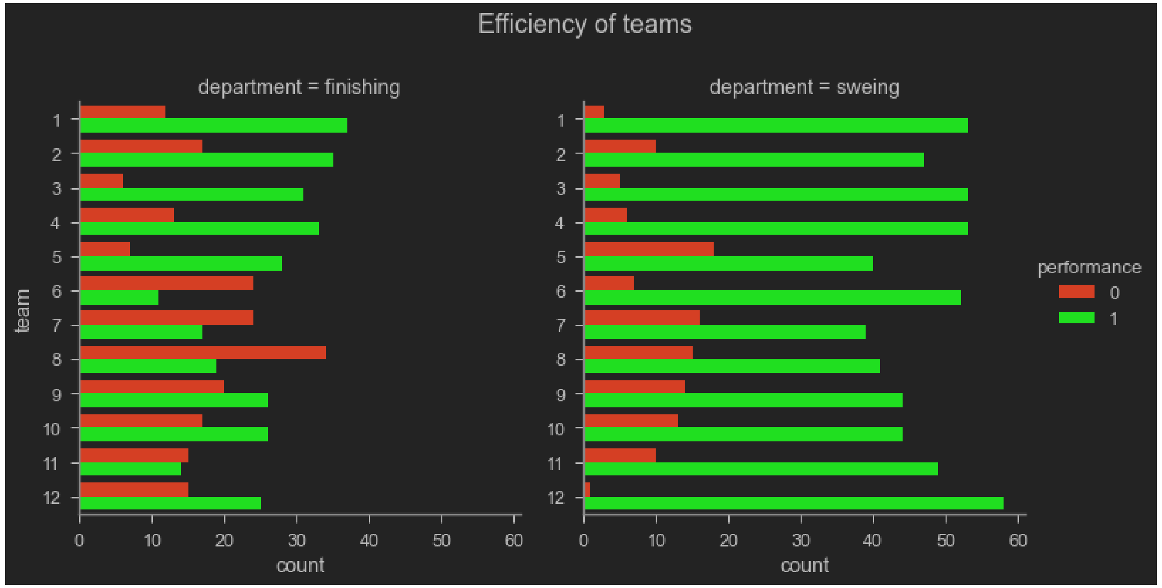
STANDARD MINUTE VALUE

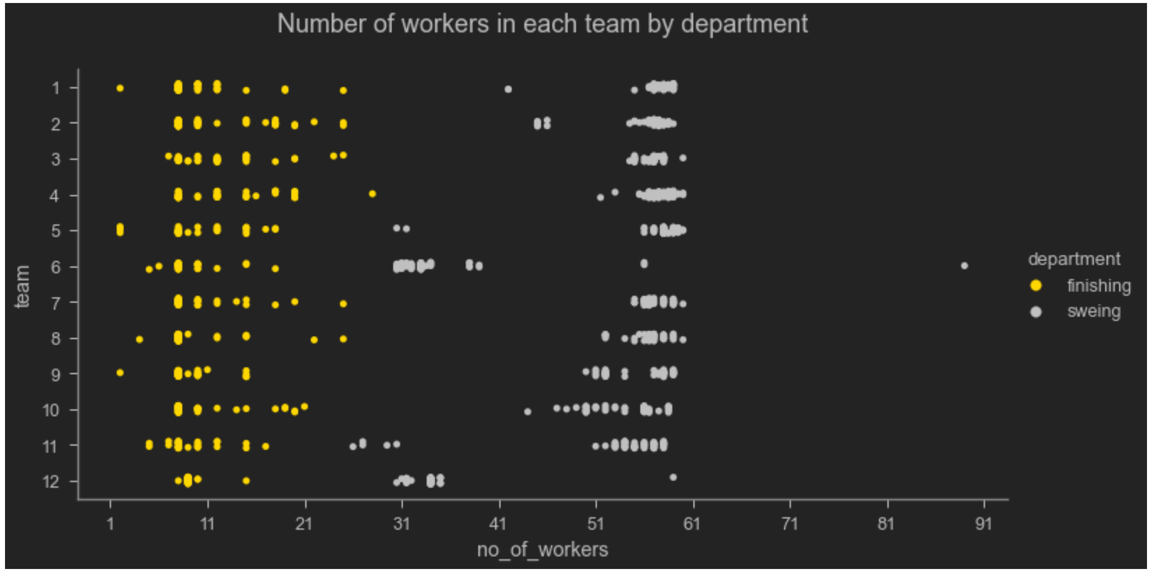
- Generally, employees in can meet their goal.
- Finishing department has hard time meeting goal
- They are not efficient on very short and long tasks.
- Higher or lower SMV tends to lead towards goal not met.

- Re-evaluate assigned minute values because:
  - They might not be well trained quick and long task

OR

  - Assigned SMV is not reflective of the task in hand





Team

Generally finishing department worker size is low.

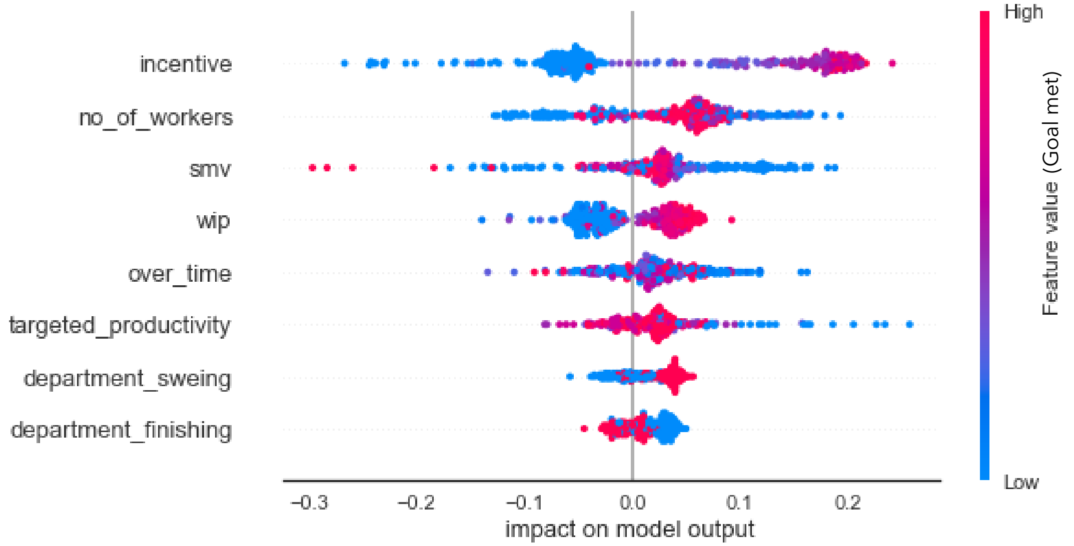
Finishing department fails to achieve goal more often.

Do they have enough manpower?

Random Subsampled Decision Trees

PEEKING INTO A BLACK BOX

Features	Probability of goal met	Probability of goal not met
incentive	High Value	Lower Value
no_of_workers	Above Average	Below Average And High
smv	Above Average	Below Average And High
wip	High Value	Lower Value
over_time	Average	Below Average And High
targeted_productivity	Lower Value	Higher Value
department_finishing	Less often	More often
department_sweing	More often	Less often



* Impact on model output is calculated using SHAPLY values which is based on game theory.

FEATURE IMPORTANCE

Recommendations

- Use insights from features to tune manufacturing process as is.
  - **Train employees** to be more efficient, especially for short and long tasks.
  - **Reassign metrics** if necessary
    - **SMV**, are they reflective of the task?
    - **Targeted productivity**, are they achievable?
  - Evaluate **team size**, especially in finishing department and adjust accordingly.
- Use this model for predicting performance and tune production process accordingly.

Next steps

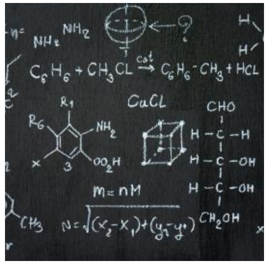
do a multi-class prediction by further binning of target.

fit a model with entire data and prepare for production use.

THANK YOU

GitHub: @tamjid-ahsan

LinkedIn:
linkedin.com/in/tamjidahsan/



Project repo:
https://github.com/tamjid-ahsan/dsc-phase-3-project

APPENDIX

Random Forest™ A.K.A. Random Subsampled Decision Trees

Let's start with a decision tree

- You want to go to grocery to by potato.
- You either go or don’t.
- Factors influence
  - Weather
  - Distance
  - Road condition

A hypothetical scenario

* The actual process is much more complex than that.
This example is more of a flow chart.

Random Forest™ A.K.A. Random Subsampled Decision Trees

Random Subsampling

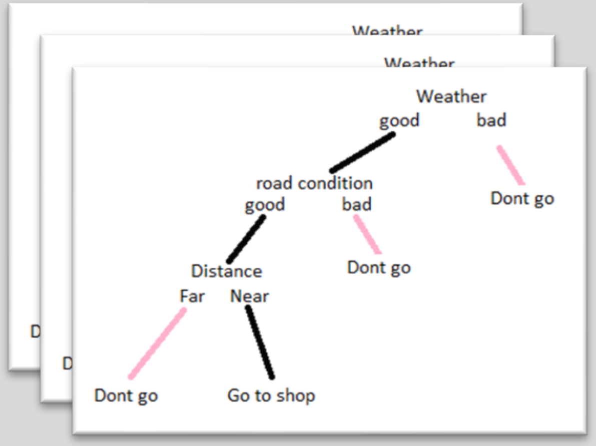
- Different combination of:
  - Weather
  - Distance
  - Road condition

Trees, with an emphasis on the "s"

- Creates a hoard of those trees and measures performance.
- Then selects the *best performing trees.*

Best performing trees?

- You want to buy either an iPhone or android. Then ask three of your friends about suggestion.
  - Friend 1 says to **buy an iPhone** because of the brand value.
  - Friend 2 says to **buy an Android phone** because it is affordable.
  - Friend 3 says to **buy an iPhone** because of the cameras.
- Now you found out that the majority of your friends suggested an iPhone. You decided to **buy an iPhone** after finding out which decision gets the **majority votes**. This is the basic idea behind the **Random forest classifier.**



Performance Metrics of the RF model

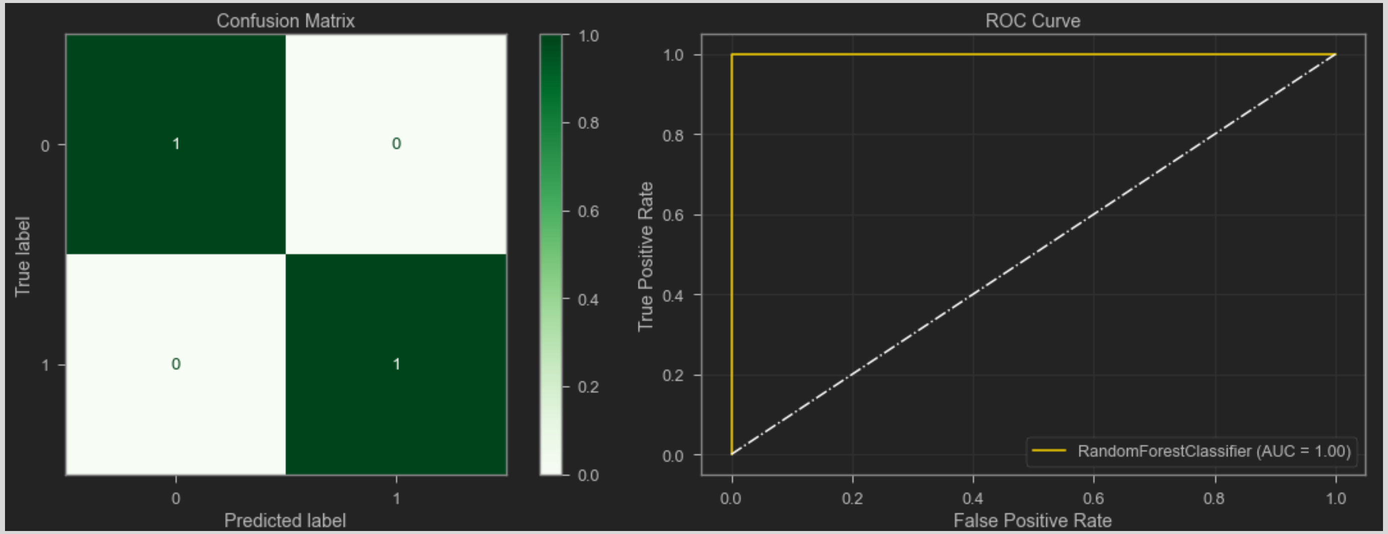
```
Report of RandomForestClassifier type model using train-test split dataset.


  ******************************************************************************************
  Train accuracy score: 1.0
  Test accuracy score: 1.0
      No over or underfitting detected, diffrence of scores did not cross 5% thresh hold.
  ******************************************************************************************

  ************************************************************
  Classification report on test data of:
      RandomForestClassifier()
  ------------------------------------------------------------
                precision    recall  f1-score   support

             0       1.00      1.00      1.00        83
             1       1.00      1.00      1.00       217

      accuracy                           1.00       300
     macro avg       1.00      1.00      1.00       300
  weighted avg       1.00      1.00      1.00       300
```



SMV Binning, minutes

	min_value	max_value
category		
quick	3	4
normal	4	15
long	15	24
extra_long	25	55

Incentive Binning, in BDT

	min_value	max_value
category		
no_incentive	0	29
minor	30	75
med	81	138
large	960	1440
generous	2880	3600