Leveraging Analytics to Improve the Efficiency of Assembly Line

Abhishek Bhambhu, Deeksha Goyal, Sudarshan Ananthakrishnan, Mengying Sun, Yang Wang

Purdue University Krannert School of Management

abhambhu@purdue.edu; goyal64@purdue.edu; ananthas@purdue.edu; sun907@purdue.edu; yangwang@purdue.edu

Abstract

This study investigates multiple machine learning methodologies that an operations manager can use to allocate appropriate tasks to suitable workstations efficiently. The motivation of the study is to help operations managers device strategies to find the optimal working hours to enhance the productivity of their operations. With the use of predictive models, our goal is to help them not only understand but also evaluate the capability and performance of each workstation and optimize workforce planning.

Introduction

Assembly Line has become omnipresent in the modern era. With a growing population and exponential demand for almost everything around us, assembly lines are a quick means to meet this requirement. It speeds up the manufacturing process dramatically. Also, the assembly line has come a long way since Ford's original conception, and the drive to optimize continues.

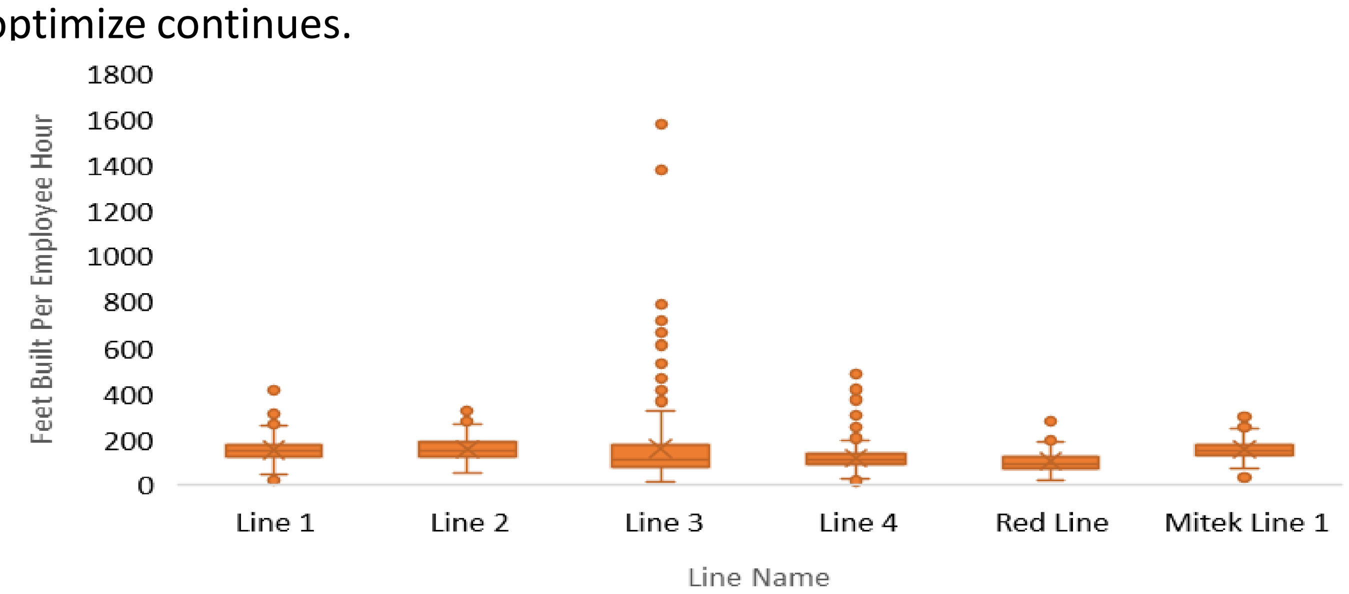


Figure: Feet Built per Employee Hour across various Assembly Lines

This project aims to increase the efficiency of an assembly line by developing analytical models while making use of the operational data provided by our industry partner to predict the completion times of tasks so that the workstations can be scheduled further in advance. Unlike most of the assembly lines today, which are automated and operated by robots, the assembly lines installed at our partner's location involve a higher human element.

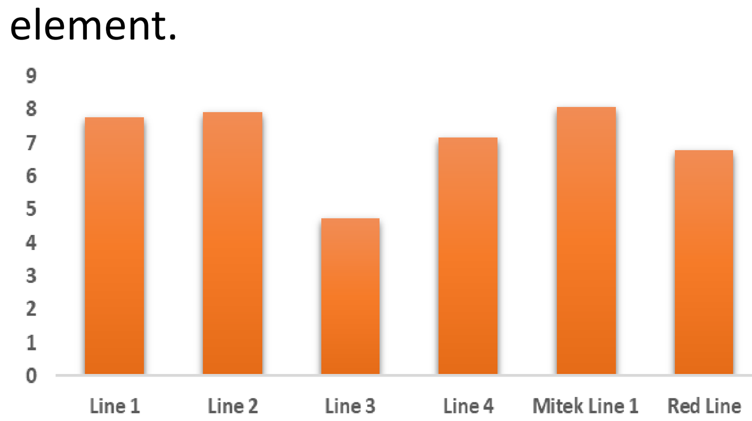


Figure: Average no. of hours each line ran

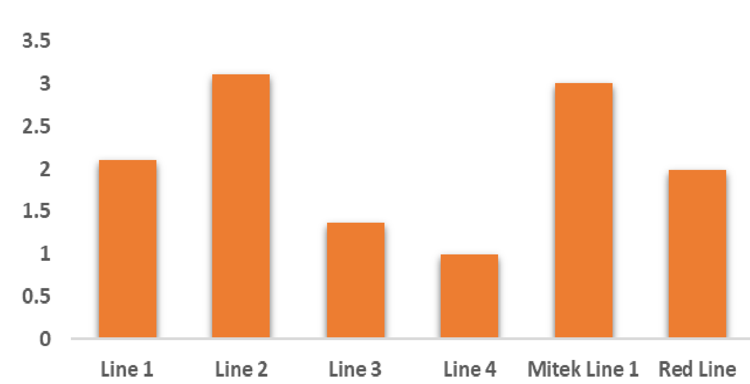


Figure: Average no. of people assigned on each line

Literature Review

Research Paper/ Model	Linear Regression	Lasso	Ridge	Random Forest	Decision Trees	kNN	ANN	SVM
L. et al., 2018	√	√	√	√	√	√	√	√
T., 2011					√		√	
U. et al., 2016							√	
Our Study	√	√	√					

Table: Literature Review summary by models used

Methodology

The data set consists of daily production data points for six different lines for 462 days between 2018-2020. However, the no. of data points are very few for each line. Because of the lack of enough data points, we used the LOOCV (Leave-one-out cross-validation) technique on our training dataset. LOOCV gives an advantage when the training dataset is small by running multiple models on the same dataset with different sets of train datasets.

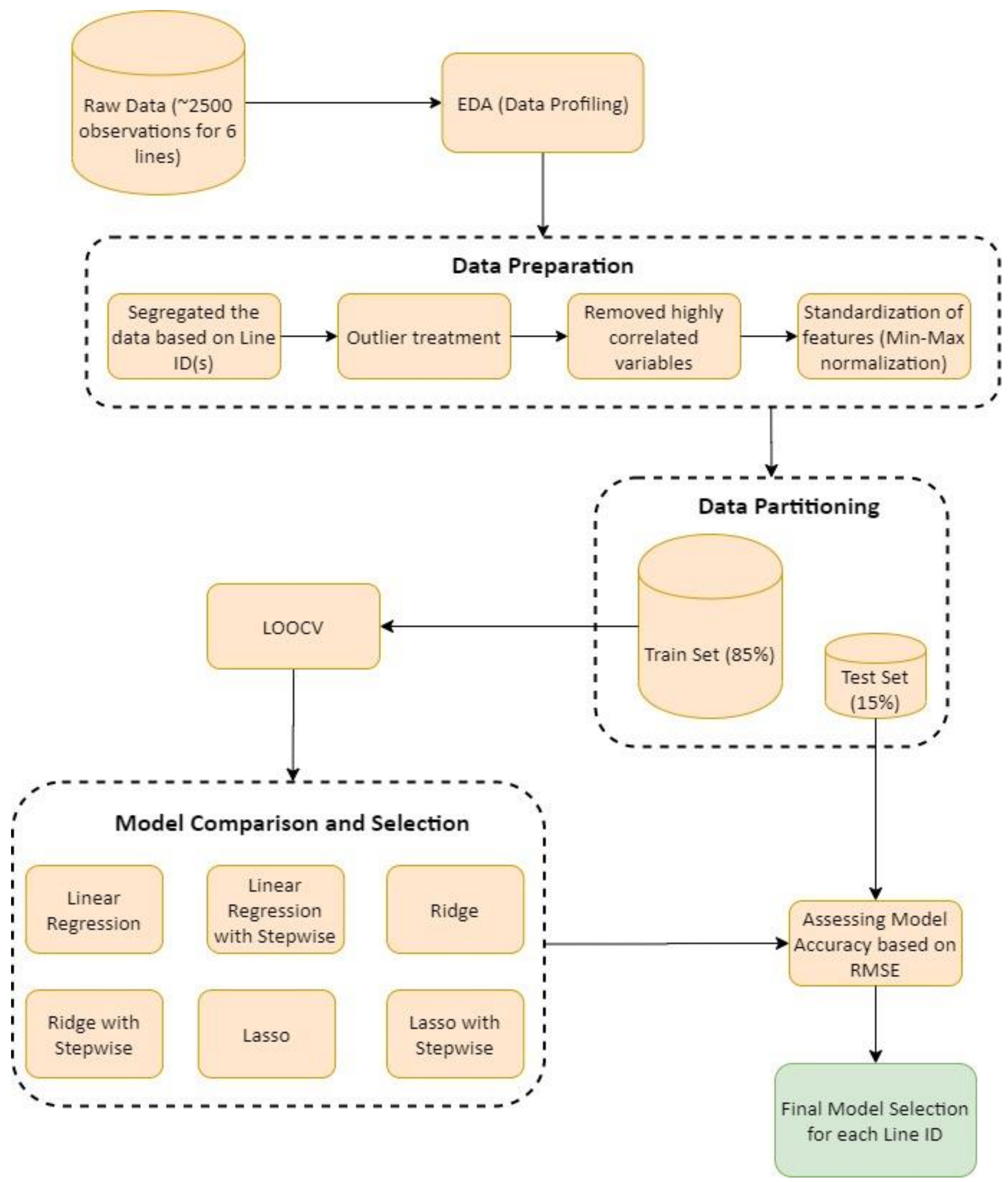


Figure: Study Design

Why RMSE was chosen as a Performance Evaluation Metric:

Root Mean Squared Error (RMSE) is a standard way of measuring error in a quantitative predictive model. It indicates how concentrated the data is around the line of best fit & measures the average magnitude of the error. Since the primary purpose of our model is prediction, RMSE is the most important criterion for fit. Also, RMSE avoids taking absolute values, which is not desirable in many mathematical calculations.

Results

We ran six different models and compared the RMSE for all models and the variation between their test and train RMSE for all the assembly lines. Lower the magnitude of test RMSE; better will be the fit. Also, the variation between Train & Test RMSE was considered so that the model does not overfit.

Line	Best Model Fit	Train RMSE	Test RMSE
Line 1	Ridge	2.54	2.56
Line 2	Ridge step-wise	3.91	3.82
Line 3	Lasso	3.52	3.30
Line 4	Ridge	1.42	1.30
Red Line	Lasso step-wise	3.14	2.84
Mitek Line	LR step-wise	2.89	3.10

Table: Model Evaluation Comparison

2.1 % | **242** HRS

Time saved annually

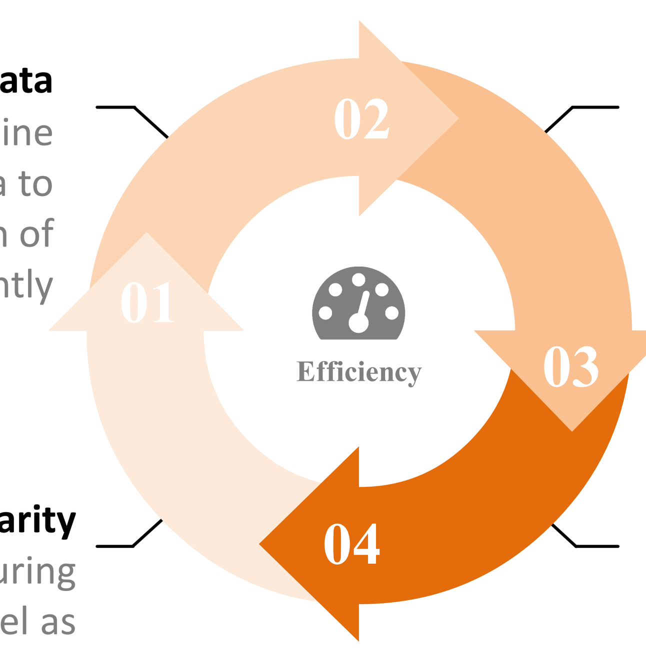
Figure: Working hours saved in manufacturing units

$117K | **1.03%**

Increase in annual revenue on Line 4, Red Line and Mitek Line

Figure: Monetary Benefit Projection

Recommendations



Machine Data
Capture machine utilization data to predict breakdown of machines efficiently

Human Data
Capture performance metrics for the workforce to create a benchmark performance criteria

Higher Granularity
Capture manufacturing data on a unit level as opposed to daily level

Robust Mode
Following the new process, a better model can be created with less errors

Acknowledgement

We thank Prof Yang Wang for constant guidance on this project.