Minimizing Warehouse Time Waste

TO HELP FEED THE WORLD

CAPSTONE PARTNERS

Harry Channing
Jessica Nunez

ADVISORS

Michael Li
Daniel Wintz
Kaitlyn Garifi
Caitlin Voegele

PROBLEM STATEMENT

Use historical data to predict how long individual tasks will take in a warehouse. Use these predictions to

- Improve the warehouse tasking engine "Athena"
- Identify customers that take disproportionate amounts of time
- Create a baseline metric to evaluate warehouse workers

BUSINESS IMPACT

Athena Algorithm Improvement

Approximately **30¢ per task saved** across 3.4 million yearly tasks

Up to **27,600 hours** saved per year across mega-warehouses

$20.7 Million enterprise value provided per year

Customer Analysis

Identified best and worst customers by absolute time cost and efficiency metrics

Dynamic customer pricing becomes possible

Labor Baseline

Good workers get promoted and bonuses faster

OVERVIEW

INPUT

Pulled **data** from a selection of hundreds of tables

~ 1/2 Billion rows → 3.4 Million rows
92 work types → 6 work types
9 Gigabytes

A large portion of our data joins came from **pallet history**, **locations**, and **intersections** datasets

PREDICTIVE MODELING

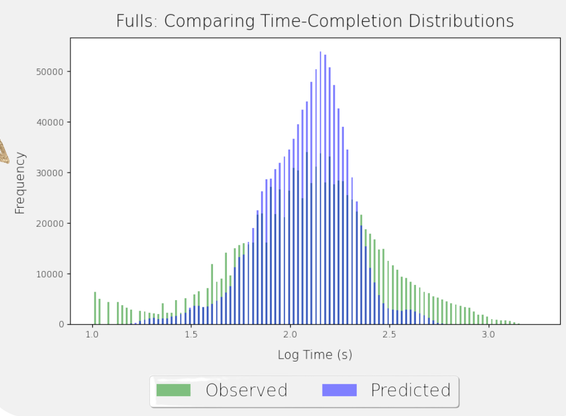
We **ensembled** a wide variety of regressors and used:

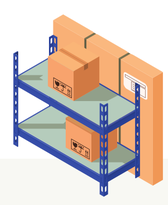
Models
LASSO
Ridge
Elastic Net
Decision Trees
Random Forest
XGBoost
...

Features
Locations
Manhattan Distance
Seasons
Customer
Equipment Used
Worker Experience
...

OUTPUT

Time estimates in seconds

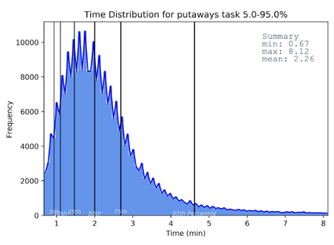




DATA WORK

EXPLORATORY DATA ANALYSIS

Visualized time distribution of different task types

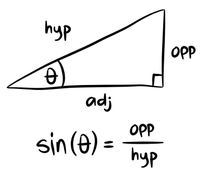


Discovered **significant** amounts of **noise** in data

Discovered distance has little bearing on time taken in task

FEATURE ENGINEERING

Calculated variety of task type and warehouse specific features

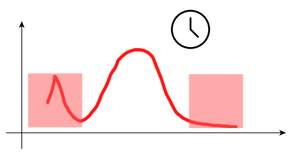


Calculated task **completion-time** as difference between start of task and start of following one

SORT -> CALCULATE: NEXT-CURRENT

DATA CLEANSING

Extreme Outliers
Removed fastest and slowest 5% of tasks



FORKLIFT 1-5 BECAME FORKLIFT

Categorized modeling features

Corrected data recording mistakes

DOOR 5 BECAME DR005

MODELING

Unique Models for Each Task Type

We hyperparameter tuned and 5-fold cross validated 96 different models

Task Type	Drops	Fulls	Loads	Moves	Partials	Putaways
Ensemble	Lasso Ridge XGBoost	Decision Tree XGBoost	Ridge Decision Tree XGBoost	Decision Tree XGBoost	Lasso XGBoost	XGBoost Random Forest Regressor
Explained Variance Increase	16.6%	30.1%	6.0%	25.0%	30.6%	16.5%
Potential Savings	$50,370.14	$260,337.73	$35,682.99	$61,930.59	$323,109.87	$95,999.23

Best scenario estimates: 1-second time differences from observed completion times

RESULTS

$827,430.55
Approximate Savings

~30 ¢
Saved per Task

~3.4 Million
Total Amount of Tasks

Impact

Our **customers** range from small **family-owned** businesses to larger **Fortune 500** companies like Starbucks, Walmart, McDonald's, and Tyson.

Lineage moves over **90 billion pounds** of food per year in their tech-heavy facilities worldwide.

Since about **90% of the food** we eat in America is refrigerated and **40%** of those goods **run through Lineage** warehouses, optimization of any of their warehouses could create an **enormous impact** in the world's leading temperature-controlled logistics company.

Inefficient Customers
6168 34364 2350 17296

Best Customers
126053 114938 2422 1050

*Anonymized for Confidentiality

NEXT STEPS

1. Implement Athena in additional warehouses
2. Improve data gathering methodology in warehouses
3. Remodel on cleaner data
4. Test dynamic pricing with customers
5. Test objective bonus/promotion schemes for worker performance using baseline labor expectations

3.4 MILLION TASKS IMPACTED ANNUALLY

UP TO 27,600 WORKER HOURS SAVED ANNUALLY

$20.7 MILLION ENTERPRISE VALUE ANNUAL INCREASE