Kaggle Days Porto 2019

talkdesk kaggle DSPT DATA SCIENCE PORTUGAL

1 month ahead hourly Contact Center Agent Forecast

Team DevScope
Dewan Fayzur, Rui Quintino, Luís Costa, Luís Maia

devscope

Kaggle Days Meetup Porto 2019

Forecasting the optimal number of agents for a Contact Center

Data Science Portugal · 16 teams · 13 hours ago

Overview Data Notebooks Discussion Leaderboard Rules Team My Submissions **Late Submission**

Overview

- Description
- Evaluation
- Timeline
- Prizes
- Frequently Asked Questions

Context

Workforce Management (WFM) solutions have been in the contact center space for decades solving needs on volume forecasting, agent scheduling, and intraday resources management.

Forecasting the number of agents to staff a contact center next day, week, or month is key to ensure there is a balance between the capacity to ensure operational requirements but prevent having agents idle increasing overall cost and operational excellence.

More about time series forecast in Forecasting - Principles and Practice: https://otexts.com/fpp2/

Problem

The number of agents available is crucial to the Contact Center success operation, where a low number of agents will lead to long waiting queues for the contacts and a high number will increase the operation cost by having agents doing nothing.

https://www.kaggle.com/c/kdm-porto-2019/overview

Overview Data Notebooks Discussion Leaderboard Rules Team My Submissions Late Submission

#	△pub	Team Name	Notebook	Team Members	Score	Entries	Last
1	—	DevScope			8.69781	33	2d
2	—	EZ			9.32422	37	1d
3	▲2	mayonnaise			9.58480	41	2d
4	—	Inglorious Kagglers			9.58862	24	1d
5	▼2	Nilg.Ai			10.15169	22	2d
6	—	Sample guys			12.13121	18	10d
7	—	ladies & gambles			17.57493	10	2d
8	—	Mantis Shrimp Punch			17.71968	5	4d
9	—	pedromag			21.71968	5	2d
10	▲1	(╯°□°) ╯︵ ┻━┻			24.63855	1	18d
11	▼1	Manuel e os seus muchachos			25.19880	4	12d
12	—	Inês Gomes			31.81709	3	15d
13	—	'); drop table Team;--			59.52087	6	4d
		sample_submission.csv			82.87673		

https://www.kaggle.com/c/kdm-porto-2019/leaderboard

Challenges

- Non-Stationary Time Series
  - (obvious trends/change points)
- Small Test Set (744 rows, 31 days, 24 hours)
- 23 months, but only Dec 2017 available
  - (hard for models to model proper seasonality for Dec 2018 prediction)
- Hard to have an internal validation strategy
  - We used Nov 2018 as validation set, closest to the target month Dec 2018
  - Besides that, we also trusted on submission visualization and PLB score
  - Time splitted validation for bayesian optimization
- Time! ☺ (lack of!)

What "seemed" to work ☺

- Blend of very different approaches
  - (regression, time series, human knowledge)
- Post/processing, human in the loop forecast
  - Christmas holidays, model obvious errors
  - Obsessive & Constant EDA of Train & Submissions
- LightGBM, Prophet, Bayesian Search
- Ceiling raw predictions (to be confirmed)
  - "The time series with the real values will be constructed using a simple agent headcount average **rounded to the next integer**"
- Probe the leaderboard (as feedback is important)
  - Yes, It's a competition!

Submissions

Model	Public Score	Private Score	
Blend of Top 3 below +ceiling	8.53526	**8.69781**	1st
Prophet Based	10.39806	**9.86448**	-
31 Models (LGBM regression)	10.44760	**10.51587**	-
LGBM regression	10.23931	**10.34396**	-

+Adjustments values for Christmas holidays/special days, based on 2017 Dec

+Other Post Processing fixing specific model limitations

Winning Submission

- Simple Avg Blend of 3 model outputs (top 3 from LB)

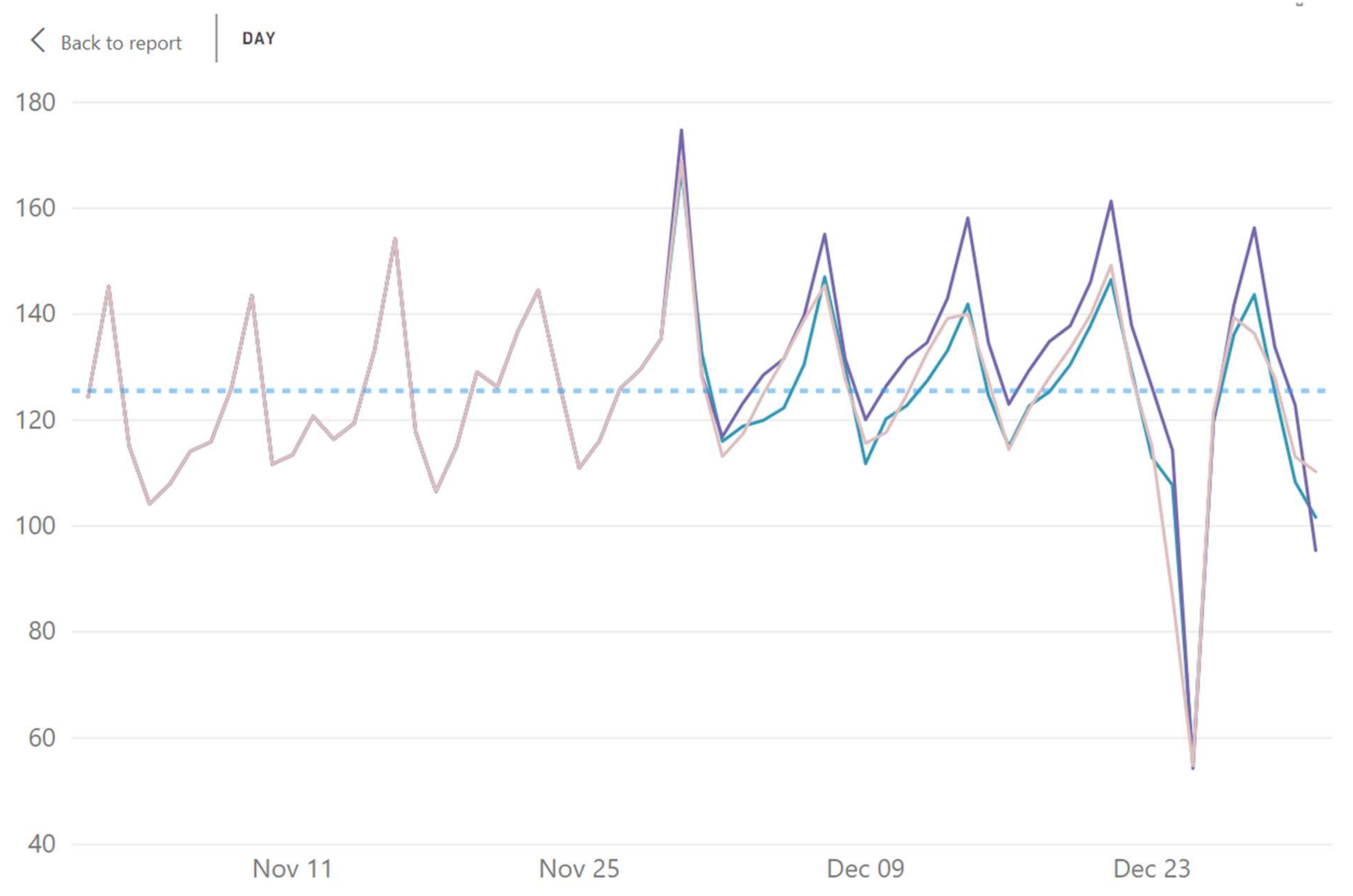
  i) **Supervised LightGBM** model

  ii) Similar to prev but with <u>**31 models**</u>

  (day 1 model -> lag 1 … day 31 model -> lag 31)

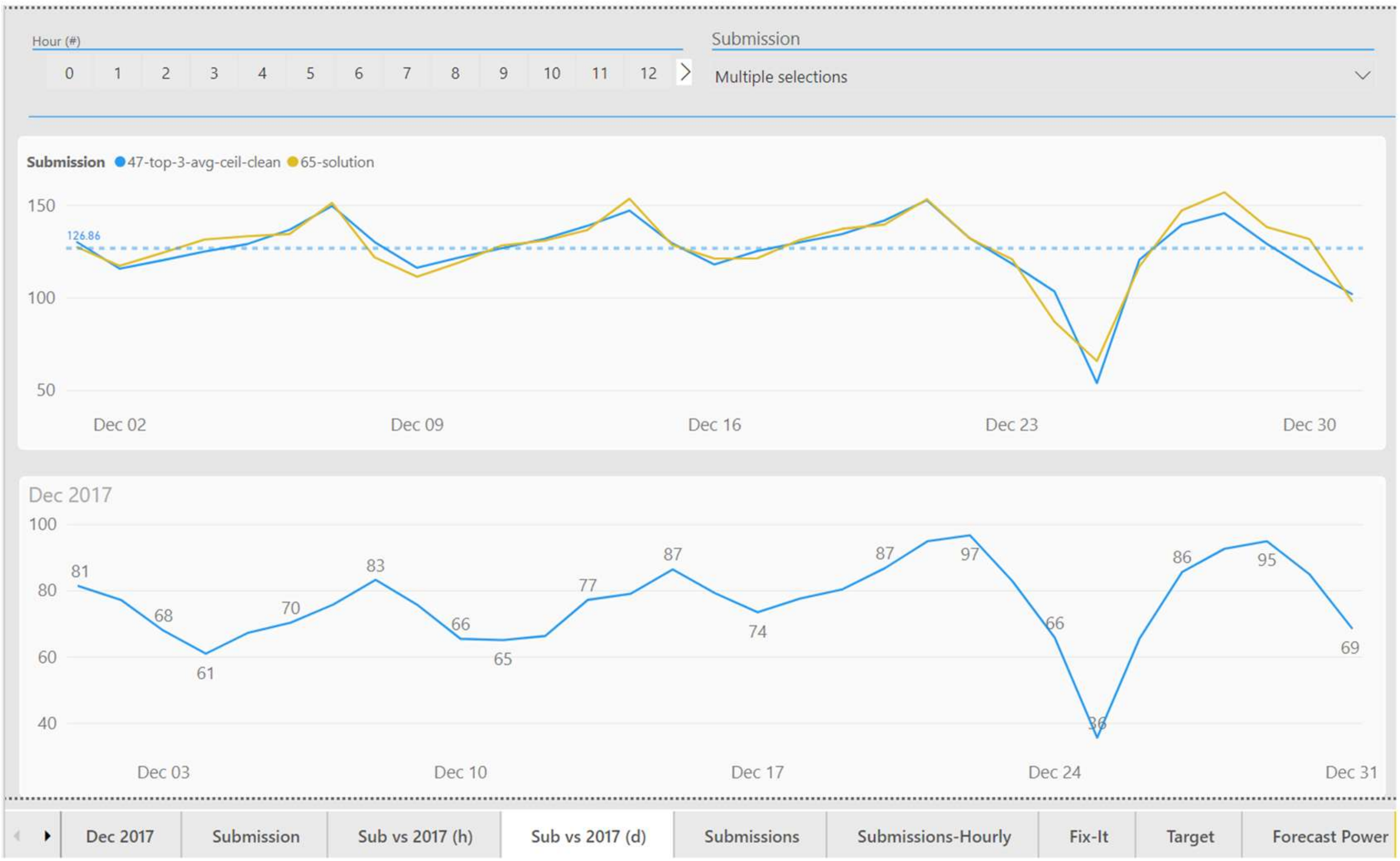
  iii) Facebook **Prophet based** model

- Plus ceiling next integer

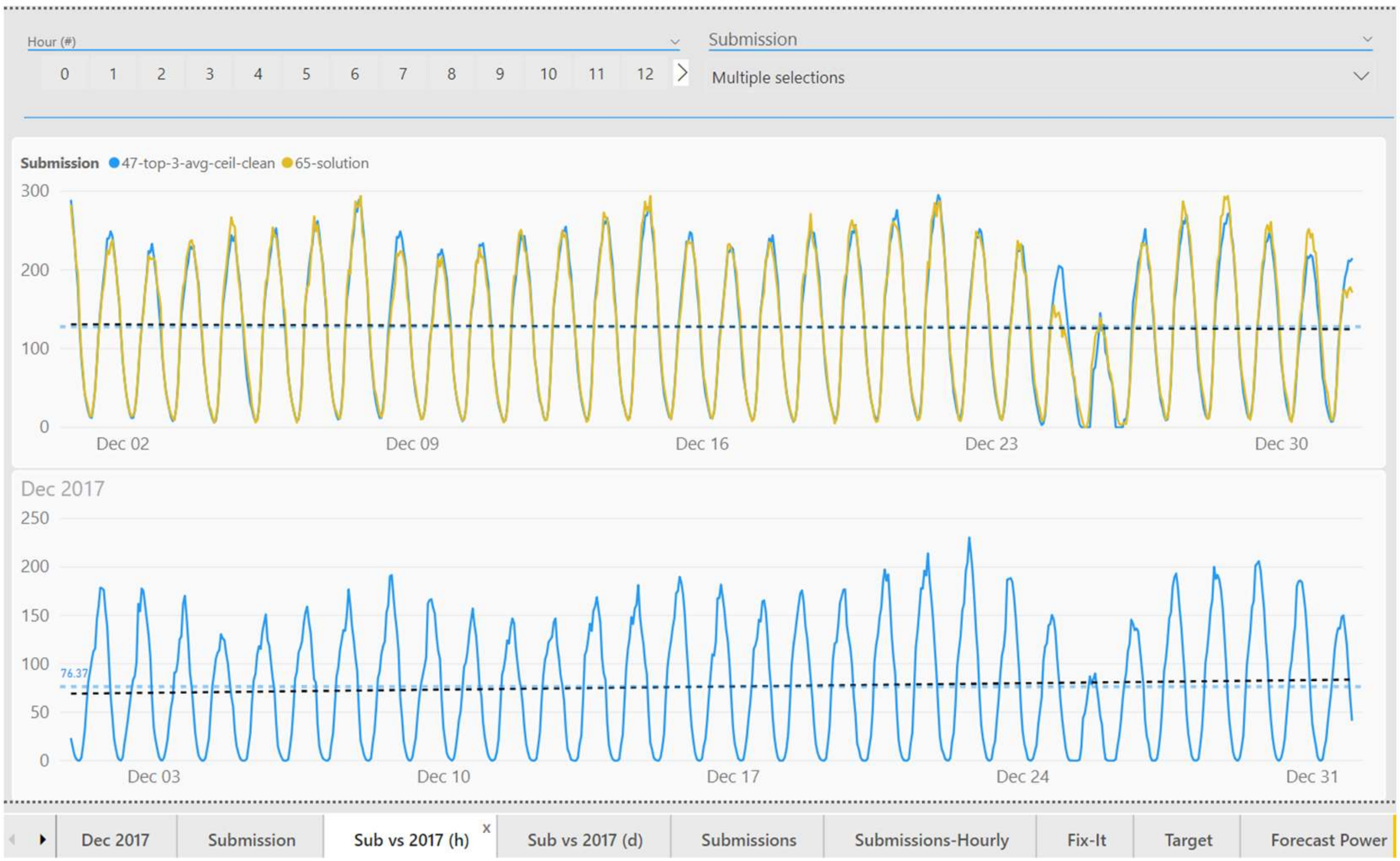


Back to report
DAY
180
160
140
120
100
80
60
40
Nov 11
Nov 25
Dec 09
Dec 23

Final Submission vs Solution/Test Set (daily)



Final Submission vs Solution/Test Set (hourly)



Time Series problem to Supervised problem

- ***Interval*** is converted to US central.
- Converted to (24 x 31 = 744) hours ahead point forecast problem.
- In the dataset, we considered next 744 hours (31 days) of forecast for every given day.
- Reformulated the problem from 744 multi output forecast to single output forecast (per hour) by converting to rows.
- Thus we end up having approximately 500K rows from 16.776 rows.

Features

- Considered only past 744 hours of ***agent_headcount*** data for every 744 hours of forecast.
- Time related variables: hours, days, day of week, weekend, months, years.
- Sine and cosine of time related variables, if the variables are circular.
- Central US holidays.
- Leadtime (1 to 744).

Training & Validation

- Training data split: data till October 2018.
- Validation data split: November 2018.
- **LightGBM, Light Gradient Boosting Machine, is used.**
- Parameters of LightGBM were found using **Bayesian Optimization**.
- Forecasted agents of 25th December 2018 were proportionally adapted from 25th December 2017.

Prophet Model

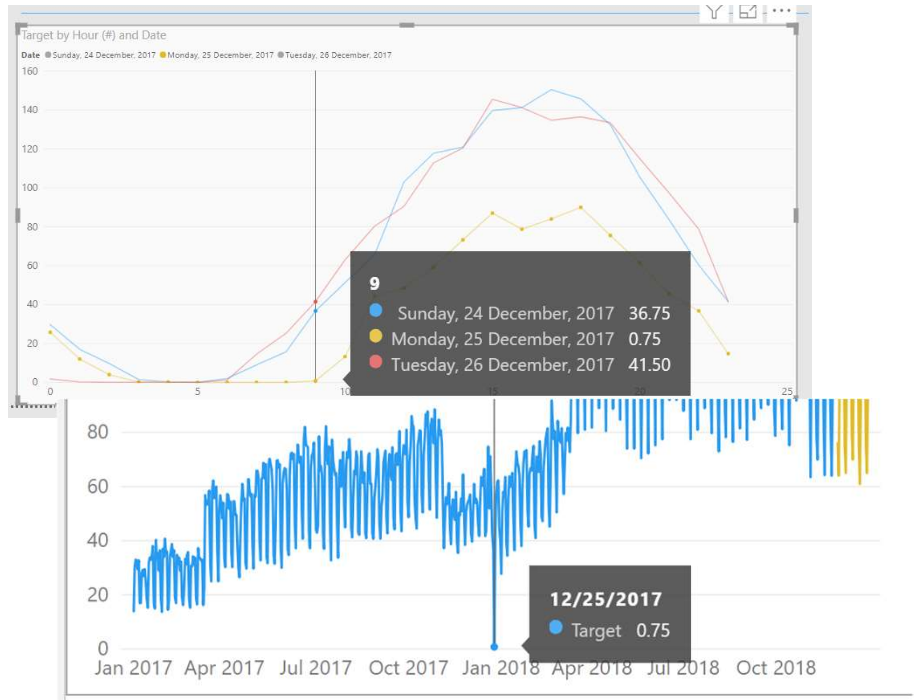
- One fit per hour (24 time series)
- Default parameters
- Training >= Oct 2018
- Used up to 24 dec, blend remaining ("detrend")

Post-Processing

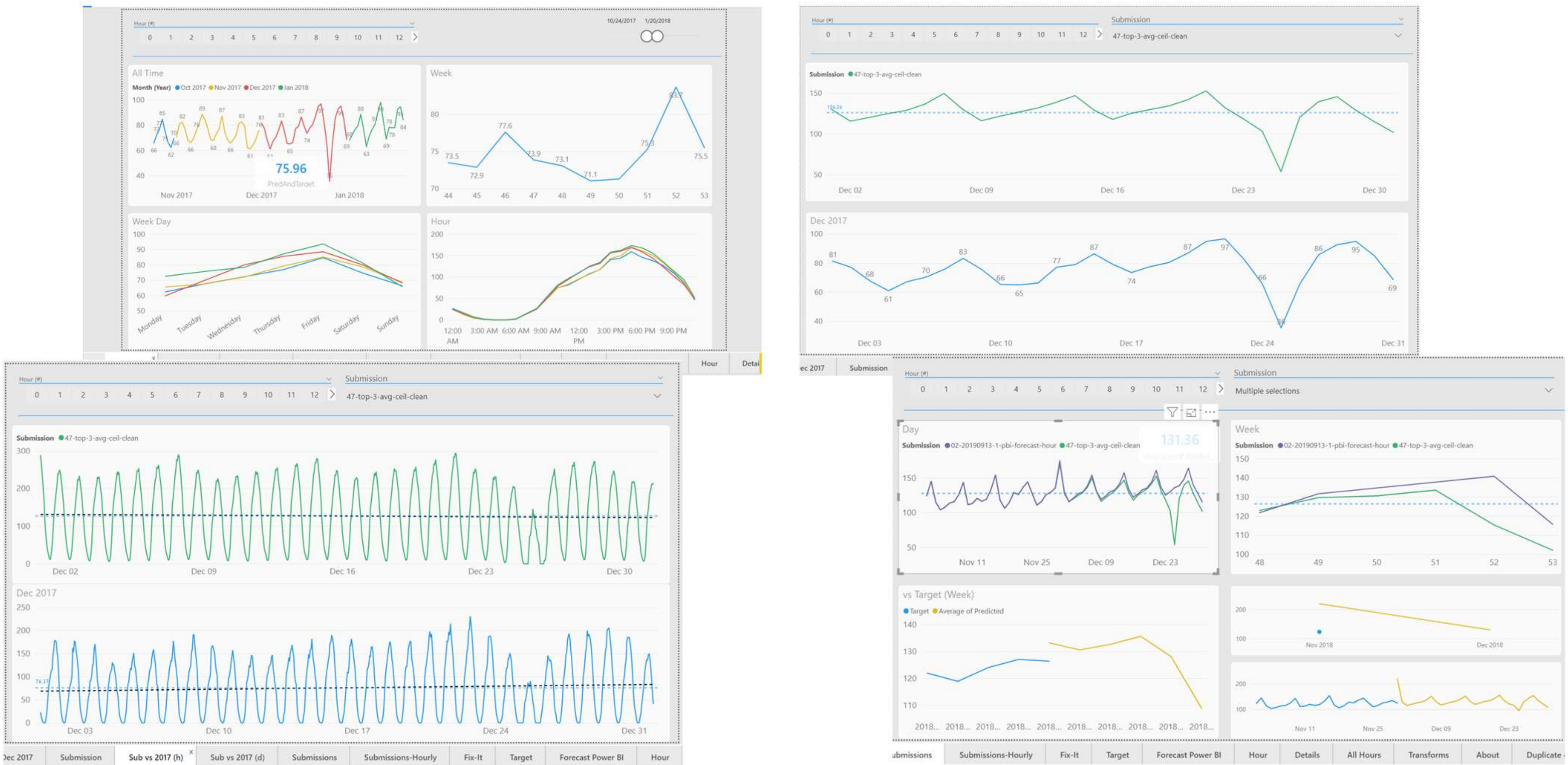
- Post/processing & fixing obvious errors
  - (model limitations, not always time to inspect root causes)
  - Negative predictions -> 0 (as agent head count cannot be negative)

Other Findings

- 25 Dec off until ~9-10am?
- (not lower, but really zero until 9-10 am)



Train, Submissions EDA Dashboards (Power BI)



What we would like to explore further

- Use other features (we mostly used target only)
- Explore feature engineering with additional features
- Understand the predictive power of other features/metrics
- Explore Azure AutoML Forecast
  - promising results but couldn't make it on time to submit ☹
  - Built-in time split cross validation, holidays & model search

What didn't work (so far...)

- Used **Total Calls** as a feature to predict agents.
  - But it did not work. Score private: 15.41 and public: 16.37
- Did not work on local validation
  - Model with 4 variables (On a Call Time, Total Handle Time, Total Calls, and Agent Headcount)
  - One reason for not improving the score considering additional variables, can be that, additional variables are calculated based on **Agent Headcount** of that hour.

Favorite Tools used in the Competition

Power BI

PROPHET

jupyter

python

Microsoft
LightGBM

Automated
Machine Learning

mlflow

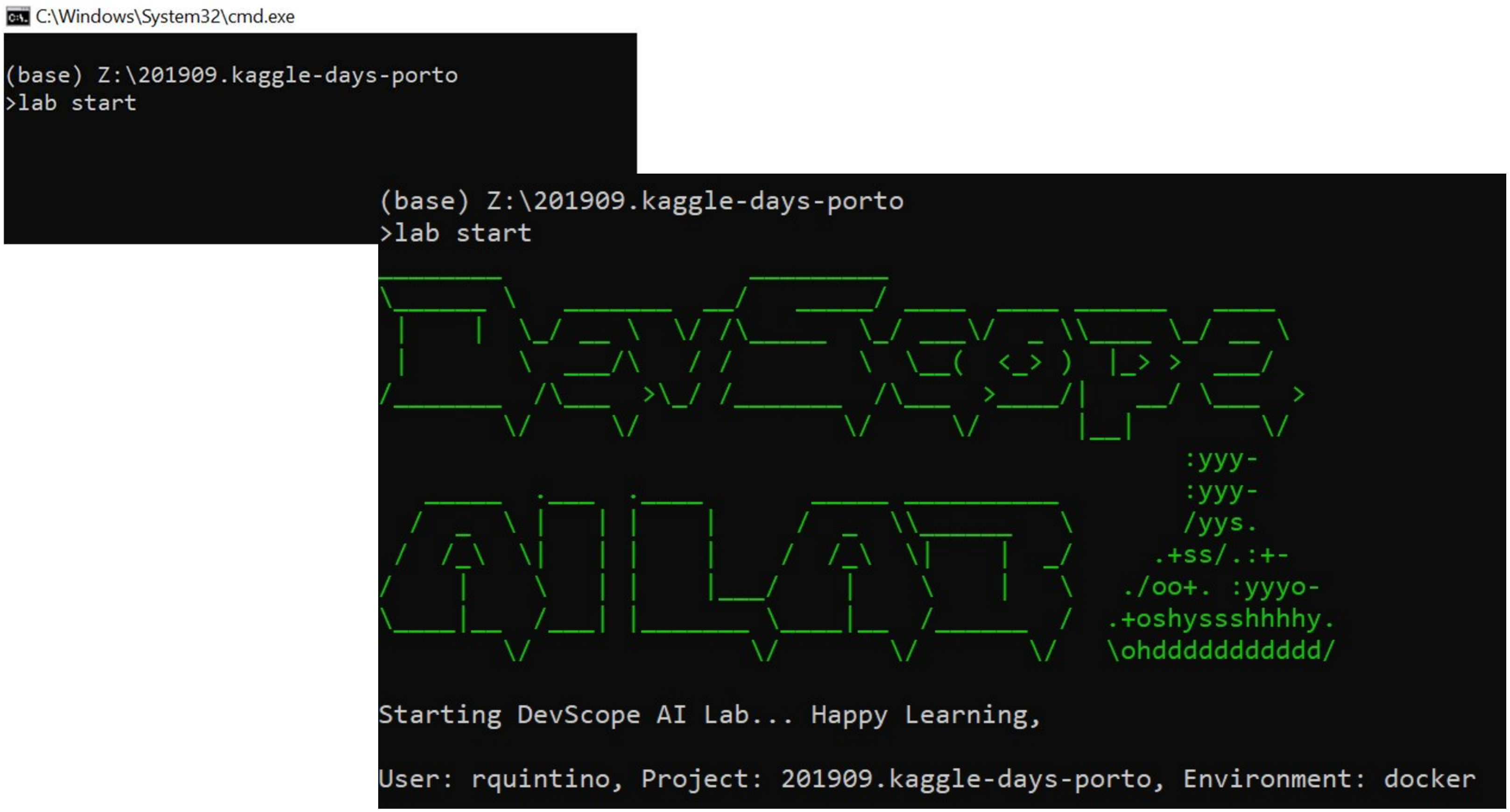
docker

Microsoft Teams
Desktop app

Azure DevOps

git

Some Secret Sauce ☺– Our AI Lab Framework



Final Thoughts

- Human + ML = improved forecasts
- When starting a forecast better to ask
  - do we need a complete automatic forecast?
  - do we need the most accurate forecast?
  - A balance between these two?