Predicting Airport Pushback Times And Ramp Taxi Times

Audrey Dahlkemper, Lisa Ko, Jihan Song, Nicola Rowe

Table of Contents

Goals:

(1) Review the research about determine the effectiveness of different ML models and different features in predicting airport pushback times

(2) Create and examine the new model created based on the paper

Audience: Federal Aviation Administration, NASA, Aircraft Industry

Purpose: Decrease delays increase efficiency of airport operations and economic profits for airline companies.

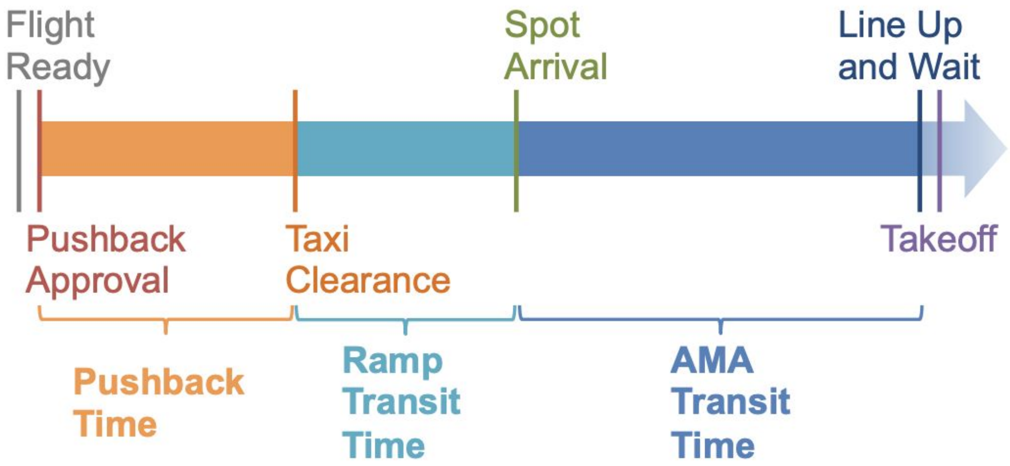
Context: Determining taxi-out time is extremely important for the efficiency for airports; however, there are often lack of accurate data for pushback time and ramp transit time. So, the researchers tried to predict this time using machine learning models.

1. Research Explained

Lee, H., Coupe, J., & Jung, Y. C. (2019). *Prediction of pushback times and ramp taxi times for departures at Charlotte Airport.* AIAA Aviation 2019 Forum. https://doi.org/10.2514/6.2019-2933

- Taxi-out time prediction
  - Require to obtain takeoff time input for runway scheduling
  - Have focused on total taxi time prediction from gate to runway
- Taxi-out time calculation

Taxi-out time = Pushback time + Ramp transit time + AMA transit time



Method: “Prediction of Pushback Times and Ramp Taxi Times for Departures at Charlotte Airport”.

1. Time analysis of taxi time
2. Decision Tree
3. Prediction Models
4. Machine Learning Models: Linear Regression, Support Vector Regression, Lasso Linear Regression, K-Nearest Neighbors, Random Forest, and Neural Networks

Dataset:

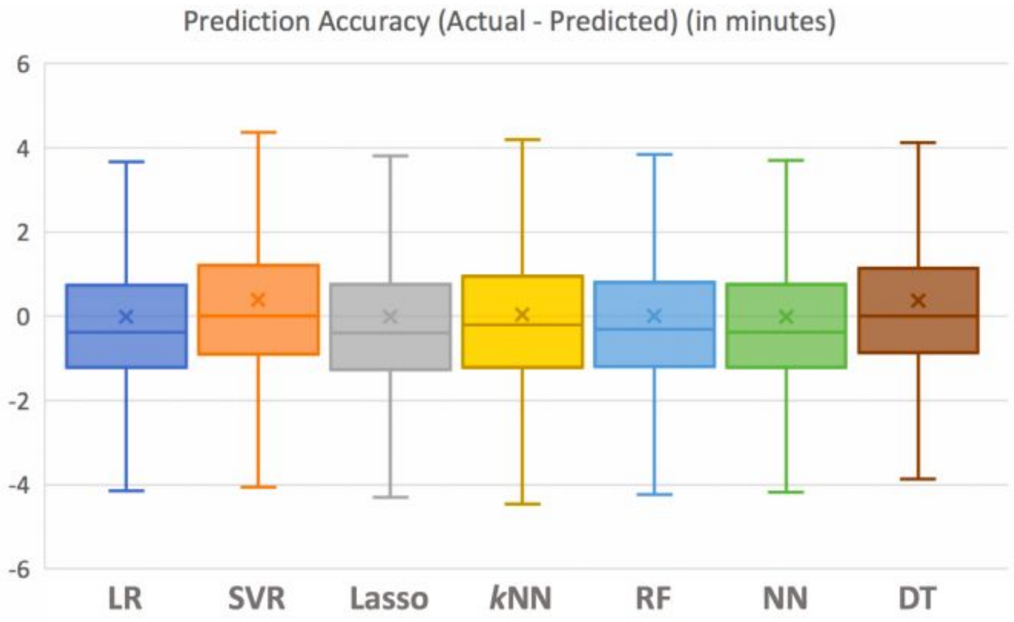
- ATD-2 system data for Charlotte Airport
- August 1st 2018 to August 31st 2018
- 67 features defined for ML algorithms used for prediction models
- 99 features defined for ML algorithm used for ramp taxi time prediction

Features for Pushback Time Prediction

- Ramp area (gate groups): 18 binary variables
- Carrier: 23 binary variables
- Aircraft type: 23 binary variables
- Pushback time of day: 1 numerical variable (hour)
- Gate conflict: 1 binary variable
- Traffic Management Initiative restrictions: 2 binary variables
    - Approval Request (APREQ)
    - Expect Departure Clearance Times (EDCT)

Total of 68 features defined and used for running machine learning algorithms

Prediction Accuracy of Pushback Time Prediction Models



	LR	SVR	Lasso	*k*NN	RF	NN	DT
Mean (min)	0.00	0.39	-0.01	0.05	0.00	-0.01	0.37
RMSE (min)	2.19	2.28	2.22	2.37	2.25	2.20	2.24
Within ± 1min	47.9%	49.4%	45.4%	45.8%	47.9%	47.4%	52.1%
Within ± 3min	90.6%	89.1%	90.4%	87.7%	89.3%	90.5%	89.8%

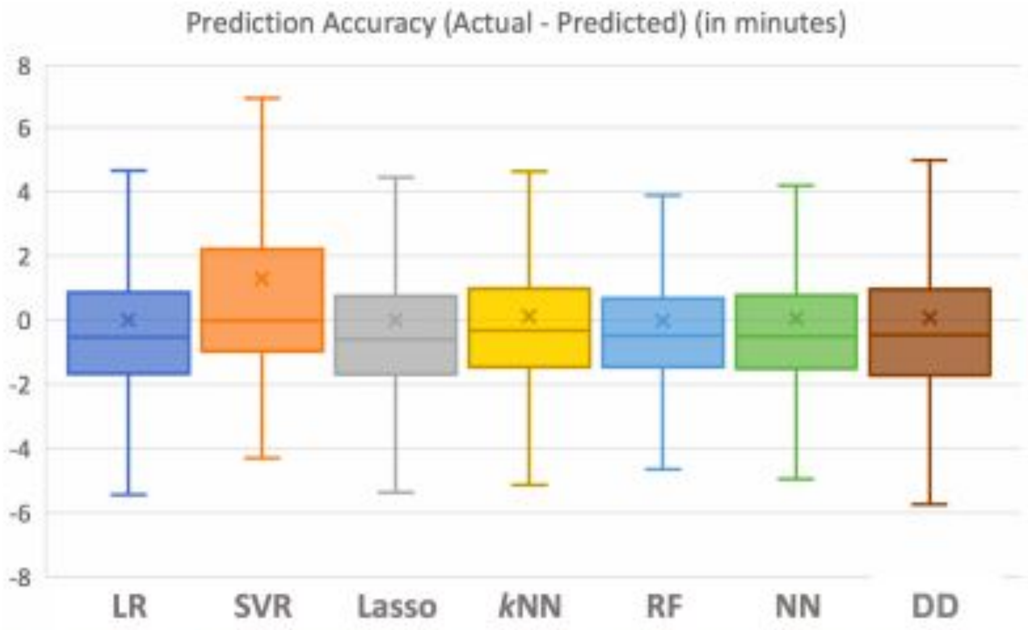
Lee et al., 2019, pg. 8

Features for Ramp Transit Time Prediction

- Ramp area (gate groups): 18 binary variables
- Spot: 25 binary variables
- Carrier: 23 binary variables
- Aircraft type: 23 binary variables
- Runway configuration: 3 binary variables
- Pushback time of day: 1 numerical variable (hour)
- Gate conflict: 1 binary variable
- Traffic Management Initiative restrictions: 2 binary variables
  - Approval Request (APREQ)
  - Expect Departure Clearance Times (EDCT)
- Ramp taxi distance: 1 numerical variable
  - A dominating factor for ramp transit time
- Number of departures and arrivals taxiing in the ramp area: 2 numerical variables
  - to account for ramp congestion level

Total of 99 features defined and used for running machining learning algorithms

Prediction Accuracy of Ramp Transit Time Prediction Models



	LR	SVR	Lasso	*k*NN	RF	NN	DD
Mean (min)	0.02	1.29	0.00	0.13	-0.01	0.04	0.07
RMSE (min)	3.56	4.36	3.60	3.80	3.54	3.52	4.00
Within ± 1min	37.7%	40.2%	37.3%	41.7%	43.0%	41.1%	37.4%
Within ± 3min	80.9%	79.2%	80.8%	78.5%	81.7%	81.1%	76.4%

Lee et al., 2019, pg. 12

Shortcomings

1. Limited time frame: data not representative of the entire year
2. Study only focuses on Charlotte airport
3. Limited number of factors, especially not including weather
4. Limited scope: even 67 features may not cover every factor that can affect airport operations

Solutions to overcome shortcomings in general

1. Incorporate weather data into the analysis: adding different features
2. Different datasets: using data from multiple airports over longer time periods
3. Limited scope: consider adding factors that could impact pushback times: ramp area data, aircraft carrier, aircraft type, etc.

2. Modified Model

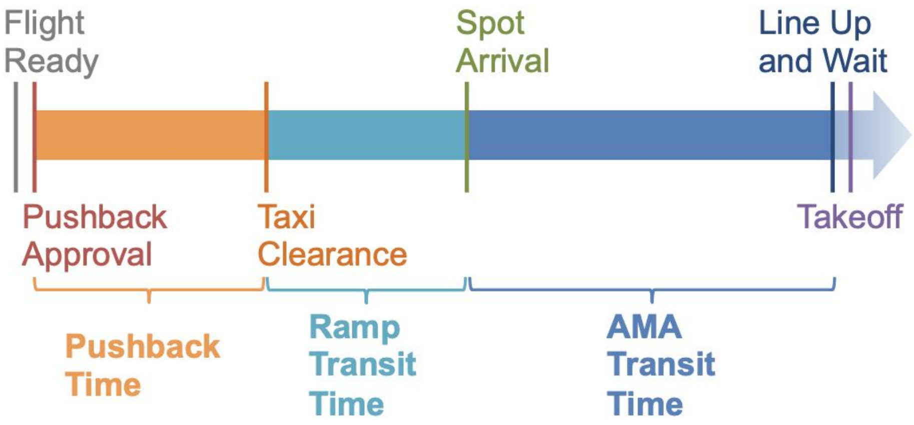
Modified Model For Total Taxi-out Time Prediction

Data:

- Acquired from drivendata.org (only accessible to the competitors)
- Data for the same airport (KCLT) - originally given to predict time until push back
- Still lack of data to create a model
    - The time data for when it arrived to the spot DNE ⇢ Taxi Out Time instead
    - Ramp Area data not available
    - When data was refined to contain essential information, the data did not have any information about aircraft type or major carrier

- Taxi-out time prediction
  – Require to obtain takeoff time input for runway scheduling
  – Have focused on total taxi time prediction from gate to runway
- Taxi-out time calculation

$\boxed{\text{Taxi-out time = Pushback time + Ramp transit time + AMA transit time}}$



➢ Lack of accurate data for pushback time and ramp transit time

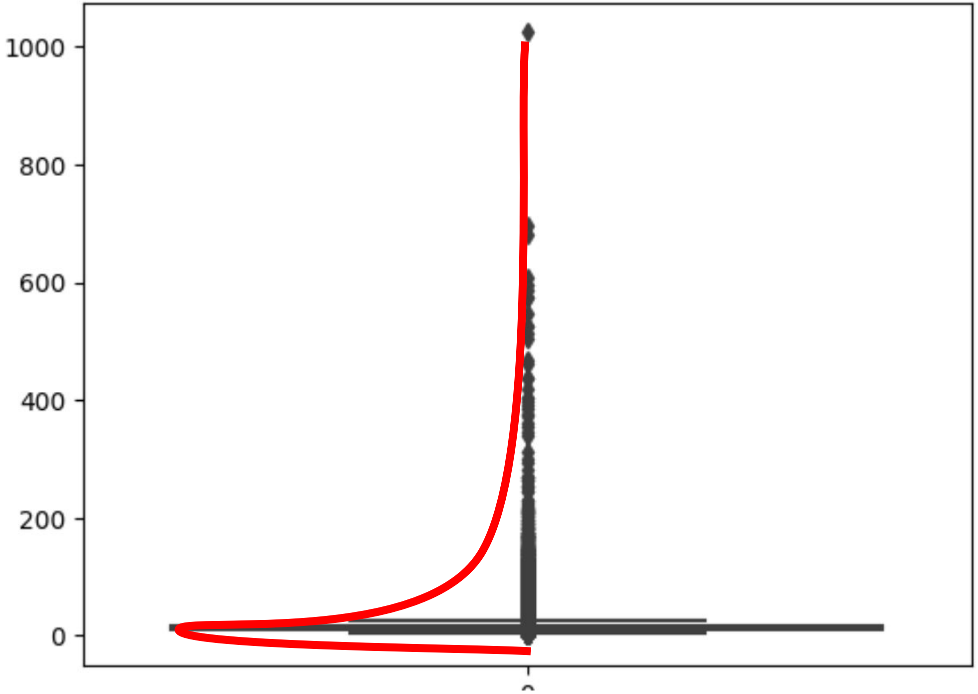
Faced with a problem...

```
aircraft_type : 1 labels NaN
major_carrier : 1 labels NaN
cloud : 6 labels
lightning_prob : 5 labels
precip : 3 labels
```

Two major categorical values used for predicting push-back time data are not available

Modified Model For Total Taxi-out Time Prediction

Label - Total Taxi Out Time




```
count    583735.000000
mean         15.931833
std           8.447546
min           0.020000
25%          11.930000
50%          14.350000
75%          17.730000
max        1024.300000
Name: total_taxi_out_time_min, dtype: float64
```

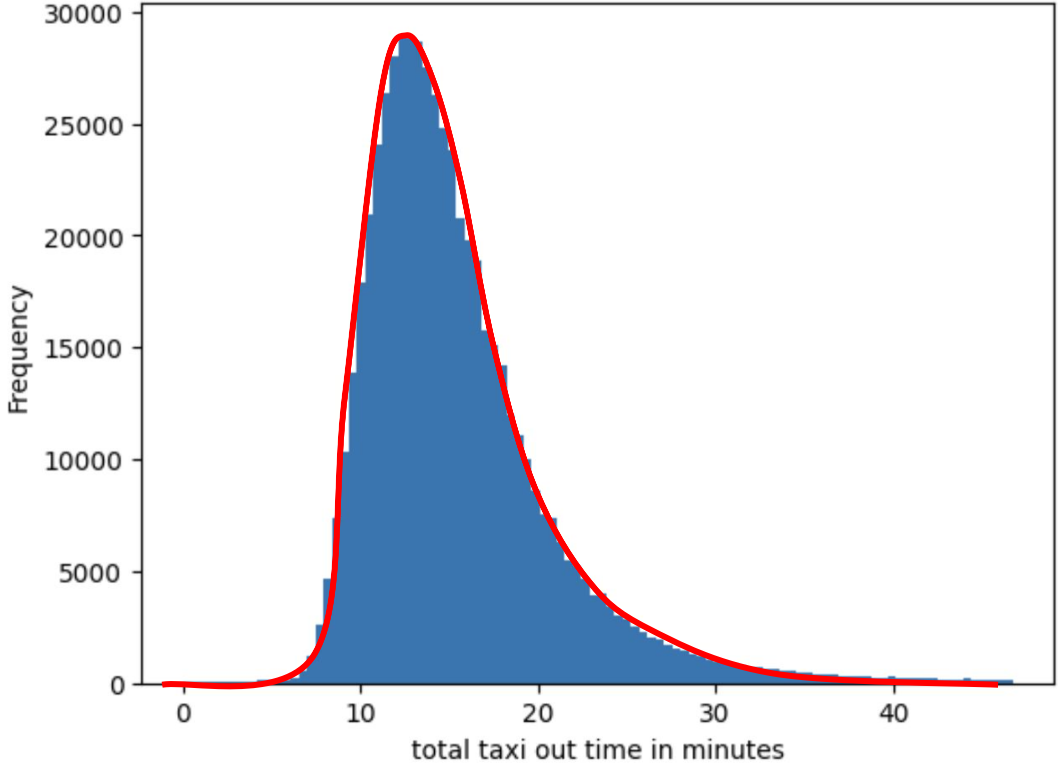
Outliers (0.02 min = 1.2 seconds)

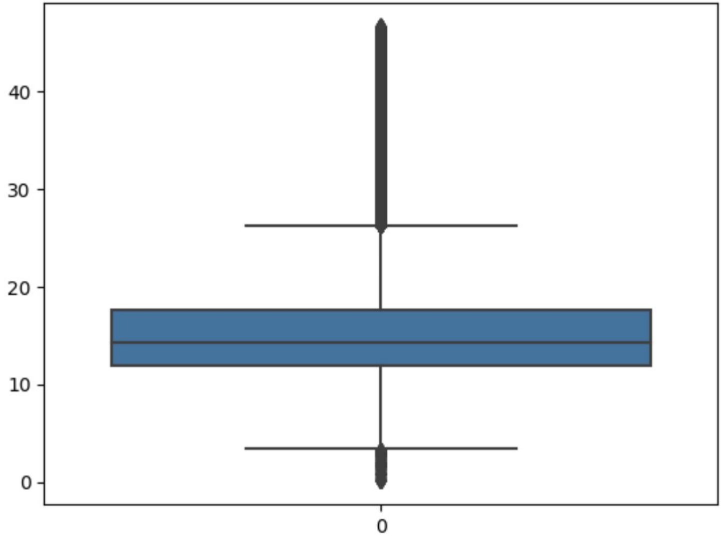
Around 15 minutes...

Outliers

Only remove extreme outlier (over and under 5 * IQR)
⇢ maintain the shape but remove the outliers

Label - Total Taxi Out Time



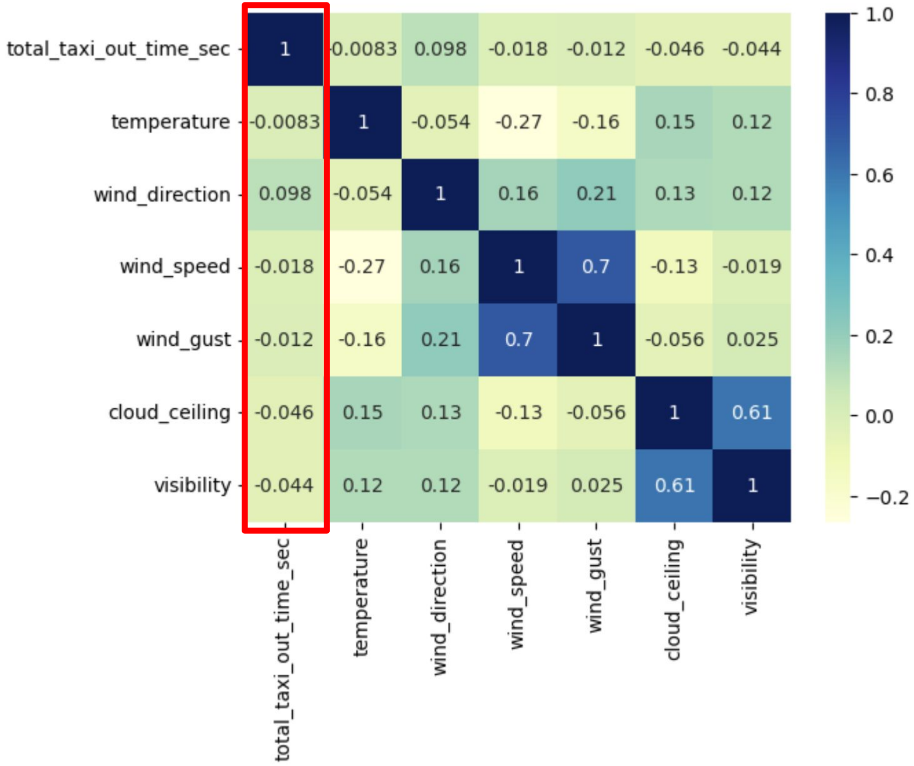


Original - 4720

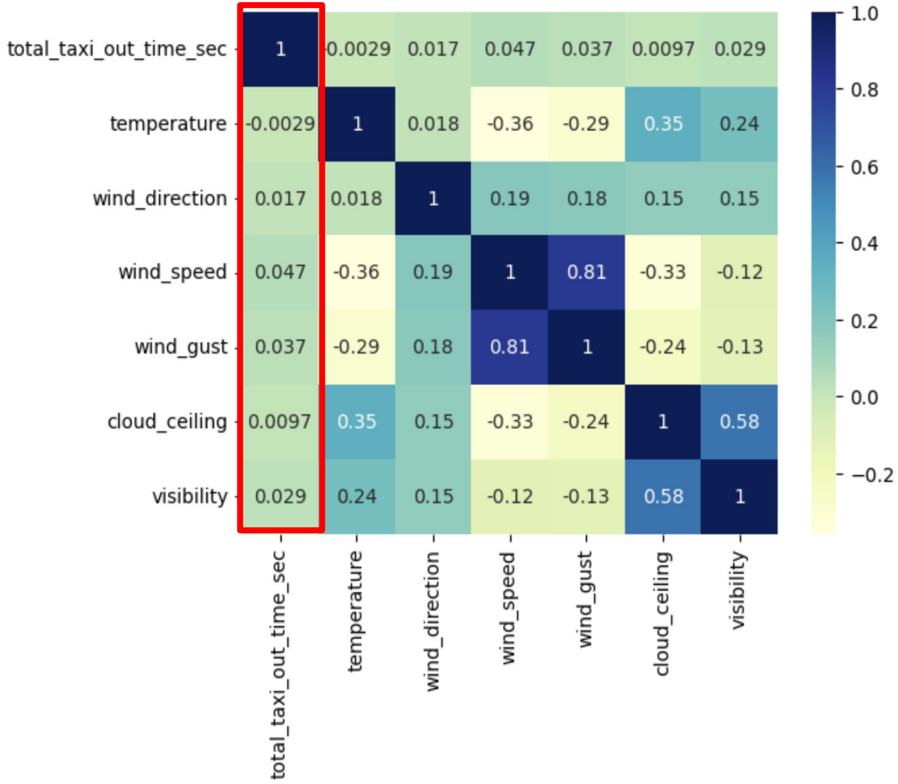
```
count      579015.000000
mean           15.464971
std             5.325209
min             0.020000
25%            11.920000
50%            14.320000
75%            17.620000
max            46.720000
Name: total_taxi_out_time_min,
```

Instead: Using weather data to predict whether it is going to be an outlier (75% + 5 * IQR) or not (Classification)

Data as a whole



Data for taxi out time over 5 * IQR



Label: Outlier (total taxi out time took more than 5 * IQR + 75%)

Features:

- Categorical: cloud, lightning_prob, precip (OHE)
- Continuous: temperature , wind_direction , wind_speed,
- Wind_gust, cloud_ceiling , visibility (Scaler)

```
Model: HistGradientBoostingClassifier
training accuracy: 0.9993494693089598
validation accuracy: 0.9993091775759041
```

Can naturally deal with NaN values

Good, but weird (more in depth study needed)

```
True positive =  57862
False positive =  24
False negative =  16
True negative =  0
```

The fact that FN exists == the model predicts Negative as well

Not Good - might have to reset the threshold for outliers to increase n of negative

Review

- Needs to find (1) better model for classification (2) or use weights
- The lack of data was a quite huge impact
- Increase the number of outliers (can affect the whole research purpose)

Q&A

THANK YOU

After Presentation...

- Lowered the threshold of outlier to 1.5 * IQR to .75 to increase the size of outliers (which is fair since (1) it is a conventional way of calculating outlier and (2) the data itself had a strong tail leading to many outliers, which should be represented in the dataset).
- The result and analysis is included in the report

	Predicted_True	Predicted_False	
Actual_True	55375	2856	58231
Actual_False	27	116	143
	55402	2972	58374

References

Federal Aviation Administration. (n.d.). *Air traffic by the numbers.* Retrieved April 10, 2023 from https://www.faa.gov/air_traffic/by_the_numbers

Stratos Jet Charters, Inc. (n.d.). *Taxi time.* Retrieved April 10, 2023 from https://www.stratosjets.com/glossary/charter-aircraft-taxi-time/

Lee, H., Coupe, J., & Jung, Y. C. (2019). *Prediction of pushback times and ramp taxi times for departures at Charlotte airport*. AIAA 2019-2933. https://doi.org/10.2514/6.2019-2933