

The LoyalBlue Membership

Jason Chang, Yong Jin Kim, and Bryen Xie

2018.11.30 - 12.02 @YHack 2018

Proposed Question:

When is the optimal time for me to leave my place of residence to catch a scheduled flight such that my time of boredom waiting at the departure gate is minimized?

Inspiration

As frequent travelers, we have all encountered the conundrum of when is early enough to leave for a flight. Wouldn't it be nice if there is an application that performs precise calculation that notifies you to leave your house at the right time, such that you can casually swing by the trip to the airport and following checking-in and security processes without spending one extra minute of boredom sitting at the departure gate? One might argue that there already exists the MyTSA app made by the Transportation Security Administration. However, according to an article published by Barbara Peterson, the experience of travelers interacting with the MyTSA app has been frustrating. The main reason of such frustration is the app's crowd-sourced feature. More specifically, the data collections of the app rely heavily on the crowdsourcing of information from the travelers waiting in the queue at the airport (or some even false-reported away from the airport?!). I argue that the queue times reported by these travelers are oftentimes exaggerated as influenced by the negative sentiments of waiting in the queues. Now, here we introduce an innovative exclusive service to the premium TrueBlue customers of JetBlue airline — The LoyalBlue Membership. Offered as a reward for the commitment of TrueBlue customers, the LoyalBlue serves as a virtual private travel assistant to guarantee that the experience of TrueBlue customers to boarding JetBlue airplane is simple and stress-free.

What it does?

The LoyalBlue Membership allows users to schedule reminders to leave home for future flights, such that it gives them enough time to arrive at the airport on time despite the traffic, check in and collect their boarding pass, check any baggage through, go through the security screening checkpoint and be at the departure gate before their flight starts the boarding process. This web app is capable of delivering such precise calculation of when to send out reminders mainly because of its incorporation of historical time series data and modern statistical machine learning algorithms. Details regarding the computations will be discussed in the subsequent section.

How we built it?

With the exclusive unpublished datasets provided by the sponsor, JetBlue airline, we were able to obtain a complete set of U.S. airports and the general information about each of them, which aids in improving the UI/UX design to include convenient functionalities in the web app, such as drop-down menu and air traffic map visualization. To abandon the problematic crowdsourcing approach of collecting queue time at the security screening line, we consult the official time-series queue data supplied by the U.S. Customs and Border Protection. However, only data for selected airports were available. Hence, we implement modern machine learning methodology to predict the queue time for the remaining airports at specific times. Since limited amount of datasets suitable for our tasks was available publicly at the time, we had decided to manually construct our own datasets with features known to be significant indicators of the popularity of each airport . Although we simultaneously ran into the risk of oversimplifying the dimensionality of the data, same can be said about the outcome of dimensionality reduction in traditional data mining workflow. Furthermore, we tackled the task as a classification problem with known labeling of training samples. Essentially, we treat each airport included in the dataset provided by the CBP as individual centroids of distinct clusters and deployed nearest centroid classifier to accurately predict the labeling of the testing samples in the remaining unlabeled airports. To offset the slight difference between each unlabeled sample point and the centroids, we implemented correction mechanisms that improve accuracy of our prediction based on the statistical measurements of how the popularity of the unlabeled airport compared to that of centroid. In addition, we utilized the cutting-edge technology of Google Maps API to estimate the time it would take for the client to transport from his or her residence to the airport of interest via various different modes of

transportation. Overall, this allows us to provide a comprehensive, yet accurate, estimate of the exact time recommended to the client to enable the smoothest checking-in and boarding experience at the airport. With the exclusive unpublished datasets provided by the sponsor, JetBlue airline, we were able to obtain a complete set of U.S. airports and the general information about each of them, which aids in improving the UI/UX design to include convenient functionalities in the web app, such as drop-down menu and air traffic map visualization. To abandon the problematic crowdsourcing approach of collecting queue time at the security screening line, we consult the official time-series queue data supplied by the U.S. Customs and Border Protection. However, only data for selected airports were available. Hence, we implement modern machine learning methodology to predict the queue time for the remaining airports at specific times. Since limited amount of datasets suitable for our tasks was available publicly at the time, we had decided to manually construct our own datasets with features known to be significant indicators of the popularity of each airport. Although we simultaneously ran into the risk of oversimplifying the dimensionality of the data, same can be said about the outcome of dimensionality reduction in traditional data mining workflow. Furthermore, we tackled the task as a classification problem with known labeling of training samples. Essentially, we treat each airport included in the dataset provided by the CBP as individual centroids of distinct clusters and deployed nearest centroid classifier to accurately predict the labeling of the testing samples in the remaining unlabeled airports. To offset the slight difference between each unlabeled sample point and the centroids, we implemented correction mechanisms that improve accuracy of our prediction based on the statistical measurements of how the popularity of the unlabeled airport compared to that of centroid. In addition, we utilized the cutting-edge technology of Google Maps API to estimate the time it would take for the client to transport from his or her residence to the airport of interest via various different modes of transportation. Overall, this allows us to provide a comprehensive, yet accurate, estimate of the exact time recommended to the client to enable the smoothest checking-in and boarding experience at the airport.

Challenges we ran into

In term of full stack tasks, there were difficulties in communications between the front end and the back end. However, through clear documentations of what JSON fields are transmitted throughout the communications, we were able to overcome this challenge.

Additionally, mobile development was very difficult at many points and required thorough testing in mobile devices connected to a local Python HTTP server. Challenges encountered during the process of data mining have been mentioned in the previous section.

Accomplishments that I'm proud of

We are proud of extensive functionality and the infinite potential of our application. The creation of teamwork behind collaborations between dedicated members was an emerging phenomenon that allowed us to achieve something together that we would've never achieved individually. With various specialized background of expertise in data science and front-end, back-end, and mobile developments, we combined to product an aesthetic product of pure teamwork.

What I learned

From the perspective of front-end development, we practiced parsing complicated JSON objects, working with React.js, and further practicing general Bootstrap skills when writing CSS. In terms of data science, we obtained experience manually constructing comprehensive dataset from scratch and came to realize the precious value behind such establishment of such enormous datasets. Lastly, it was fascinating to become aware of how much computer science can contribute to improving the overall living experience of the society.

What's next for LoyalBlue Membership

With the current state of the LoyalBlue Membership, we provide our premium clients with the most efficient transportation from their place of residence to the selected airports. However, the cost and reliability of the transportation still lie on the shoulder of the clients. Hence, in the near future, we can go one step further by providing clients of proximal destinations with rideshare shuttling service to designated airports via the most optimal routes computed by our algorithm. In addition, we would like to establish long-term partnership with JetBlue to further improve the integrated flying experience of our committed clients by targeting age-old caveat, such as that of the ticket-overbooking

problem.

Built With

GCP • React.js • Firebase • Twilio • JetBlue • JavaScript • CSS3 • HTML5 • Flask • Python • Google Maps API • Scikit Learn • Scipy • Numpy • Pandas • Axios • Apache2 • Wsgi

Try it out

loyalblue.club

Demo

[Link](#)

Snapshots



References

1. Peterson, Barbara. "TSA Knows Its 'My TSA' App Is Useless." Condé Nast Traveler, <https://www.cntraveler.com/stories/2016-07-08/tsa-knows-its-my-tsa-app-is-useless>. Accessed 2 Dec. 2018.