

# Winning SWISS AI Hackathon 2K19 - Predicting Rotational Delays

December 17, 2019

**Predicting 24 hours in advance flight rotation delays; this was the challenge of 2019`s AI Hackathon organised by SWISS. Read more to learn what we did and how we ended up winning!**

**See the video [here](#)!**

In aviation everything is about moving people between two places as smoothly as possible. Thus, delay is one of the most critical issues that we have to deal with. In 2017 and 2018 Swiss Airlines suffered heavy delays with roughly a quarter of all flights having departed more than 15 minutes late.

This is why Swiss organized the *AI Hackathon 2k19* on the 15<sup>th</sup> and 16<sup>th</sup> of May, 2019. Teams from within and outside of the Lufthansa group were invited to participate on the challenge of building a machine learning model to predict rotational delays. Rotational delays are delays caused by previous flights of an aircraft, e.g., if the aircraft arrives late and, thus, can't depart on time.

We, the team of zeroG, picked up the gauntlet and ultimately won the hackathon.

Below we share about our experience and explain how we approached this challenge.

## **DAY 1**

As the hackathon started at 8 o'clock in the morning, we arrived in Zurich the day before. When we arrived at "The Lab" we had some time to talk briefly with the other competitors and to explore the facility we would spend most of the next two days in. As the competition had not yet started, the atmosphere was relaxed and we met some of the other competitors while sipping our first coffee of the day.

At around 9 AM received a brief on the challenge we had to work on: given multiple sets of data about flight schedules, weather forecasts, passenger bookings, airports and more, we were to predict the rotational delay of each flight one day in advance. All teams were given a set of flights to make predictions on. In best Kaggle-fashion, these predictions were used to

score the teams and find the winner.

Knowing that we were not going to get much sleep, we started with a quick planning session on how we would like to approach the task. We settled on doing data analysis, data transformation, and feature creation first and move to model building on the second day. As we had so many datasets, we went for a divide-and-conquer approach: split the work and update each other on our individual findings later.

As it turned out, the most important features we extracted from the data were historic delays for the flights, the amount of stress at time of departure (indicated by high amounts of passengers and other airplanes at the airport), and the amount of flights an aircraft had to operate before the flight in question.

During the day we worked hard and didn't take any breaks - except to refuel with coffee! As the evening approached, we merely had a set of interesting features but didn't built a proper model yet. Our best prediction at that point was just using the delays from the previous year.

At the end of the first day, we were scoring a lot lower than other teams that already had models working. This reaffirmed our belief that we have strong competition! With a sense of pressure, we decided to continue working a little longer.

## **DAY 2**

After six hours of sleep we met again at "The Lab" in the Swiss headquarters. As planned on the first day, we started building models to predict the rotational delays. Similar to the first phase, we divided the work to try out as many algorithms as possible. At around noon, our scores got closer to those of our competitors and we settled on two models that we wanted to optimize: a RandomForest and a gradient-boosting model.

Our final submission ended up being an ensemble of two RandomForest models, both of which had hyperparameters optimized using grid search. On our way to getting our final result, we spent hours doing grid-searches, tweaking features, and - honestly - panicking because our scores were lower than most others. Only in the last hour or so did we manage to get a decent submission and were excitedly waiting to see our rank in the final standings.

Before the final results were shown, each team did a presentation about their approach. It was very interesting to see what the other teams did. We regret a little bit that there wasn't much exchange between the teams, but a competition is of course a competition.

After the presentations, the final scores were announced. Though we participated for the fun, we really wanted to end up on top. When we were finally named as the winning team we were psyched. After all the hours we put into our model it felt incredibly rewarding to come out first!

The event ended with a get-together and we finally got the chance to chat with some of the other competitors about their approaches and thoughts. We also met a lot of people from Swiss that came and showed interest in the hackathon. All in all a great event!

## **FINAL THOUGHTS**

We thank Swiss, especially Manuel Trunk and René Schmassmann, for hosting this awesome event. We also like to thank our zeroG colleagues for their amazing support! Furthermore, we did not anticipate the level of interest in the hackathon by all the non-IT colleagues at Swiss. It highlights the importance of this topic and their willingness to support innovation.

We had a lot of fun wrapping our heads around this challenge, building a working model in such short time, and growing as a team. We also really enjoyed the friendly atmosphere among the teams.

The use case of delay prediction proved to be more complicated than expected and we think that there is a lot more that can be done on this topic in the future. In the 36h given we only had a chance to scratch the surface of what is possible. We sincerely hope that this topic will be worked on further with models that can unlock the true potential of this critical use case.