Predicting Peak Chair Opening at WhistlerBlackcomb

BrainStation Capstone Project - Technical Summary

05 December 2021
William Matthews

Introduction

The Peak chair at WhistlerBlackcomb is a favourite among locals, ski-guides, and tourists alike. Each of these groups are looking to get the most out of each day on the mountain, and knowing if Peak chair is going to open or not can go a long way to helping them plan their day. Knowing the chair will not open can allow locals to frequent their other favourite powder stashes. It can allow ski guides to plan days better for their clients. It can allow tourists to best choose if today is the day to take a break in their 7-day ski vacation and relax at the Scandinave Spa or take a long lunch at the Bearfoot Bistro.

The objective of this project is to provide proof of concept that accurately predicting the Peak chair status on any given day during the ski season is possible. Considering the scarcity of days on the mountain for most people, the threshold of 'accurate' will need to be very stringent. The opportunity costs of either missing an open day or standing in line on a day the chair does not open are high, so we feel an accuracy of 95% or greater is an appropriate threshold to aim for. If that is possible, the next steps are to build a data pipeline to keep daily information flowing in and build a simple website/webapp to serve the predictions to the public. Further natural steps would then be to build models to predict the status of the other alpine lifts at WhistlerBlackcomb, and then to see if the models can be refined to predict the opening times of the lifts on any given day.

Data Sets

Data for this project was collected from three sources, the PowDay online database, the Government of Canada's climate website, and the University of Wyoming's upper air sounding website. The data collected covers the ski seasons at WhistlerBlackcomb from January 2015 through April 2021. With the exception of the Government of Canada's weather website, which appears to be updated monthly, all sources appear to be updated daily and could be used for a model in production.

The PowDay app has a database of push notifications that provided us with each date the Peak chair opened. The website had a handy interface to download all of these records as a CSV file with a few clicks of the mouse. These formed our target variables. We also used the previous day's lift status (opened or closed) as features to predict the lift status of the current day.

The Government of Canada's climate website provided us with temperatures, snow fall, rainfall, and snowpack depth from the Pig Alley weather station at 1,640m elevation on Whistler Mountain. To obtain the data you can either download month by month, or the main website provides a link to a google drive in which there is a text file. That text file contains a bash command which allows you to download the files en masse for any date range and weather station. The biggest disappointment here was that weather data for October through late November of each year was missing. Based on our approach, this became a critical hole in our dataset.

The University of Wyoming's upper air sounding website provides weather balloon flight data from around the world. We were interested in two sites, Port Hardy on Vancouver Island, and Quillayute on the Olympic Peninsula in Washington State. These provided us with temperatures, humidity, mix ratios, wind speed, and wind direction. The university's website would only provide a basic HTML page for one month of flight records for one station at a time. Luckily, the URL of each page was in a standard form, so we were able to build a scraping algorithm to request and parse all of the data we needed.

Data Processing

Each of the three data sets obtained was cleaned and prepared separately before being joined in the final data processing step.

The PowDay data set contained two pieces of information that we wanted to extract, the date/time of the lift opening, and the name of the lift that opened. This involved some simple string parsing. The data provided values only for the days the lift opened. A little research provided the opening and closing dates of WhistlerBlackcomb each season. With these we inferred all the days we did not have a notification of open for had a lift status of closed. There were also a number of days with multiple opening times. After inspection, this was addressed by keeping the first instance of opening only.

The Pig Alley weather station data from the Government of Canada's website provided some key pieces of weather information and a lot of superfluous columns. Columns not pertaining to weather were all dropped. We explored the numerical data for erroneous values that appeared out of range and corrected them. The data set also contained a number of missing values. These were only in one or two day gaps, so we imputed the missing values with a linear imputation between the values on either side of the gap for temperature and snowpack. Snowfall was a little more complicated. We were able to determine that the average contribution of snowfall to snowpack on the subsequent day was five-ninths of the snowfall depth. Using this, and the completed record of snowpack data, we imputed any missing values for snowfall. Finally, we returned to this step during our modelling to attempt to engineer the snowpack depth at an elevation of 2,000m. It was at this stage that we discovered that almost all the data for October through late November was missing.

The data from the two weather balloon stations was the most challenging to acquire and process. Once the scraping process was determined, we needed to extract the date information from each balloon flight and convert it from Zulu (Greenwich Mean) time to local time. The flight data was contained in preformatted text, and required a combination of string parsing and regex to extract. Finally, we captured only the flight data around the 700 milli-bar/3,000m elevation. This data also required one notable piece of processing. Wind direction is a circular variable and was provided in degrees. Machine learning models cannot understand that 1 degree and 359 degrees are near equivalent. We engineered a sine and cosine component of the wind angle from the degrees provided.

Each of the above data sets was saved to CSV and then combined. The data was then structured in a pseudo time-series analysis fashion with each data point copied multiple times as it became a lag for future targets up to 5 days ahead.

Exploratory Data Analysis and Feature Engineering

Our EDA focussed on several features that we believed were going to be critical for model success. Of the features we explored, several interesting results appeared.

First, we explored snowpack depth. Our intuition was there is a minimum snowpack depth that must be achieved before the Peak chair will open. Since an operational goal of the mountain is to have the chair open as much as possible, we assumed the range of snowpack depth on the day the Peak chair first opened each season would be small. Contrary to this, there was a spread of almost 1.3m in snowpack depth on the day the Peak chair first opened each season. With the midpoint of the Peak chair roughly 360m above the Pig Alley weather station, we tried to engineer the snowpack at that elevation from total rainfall and temperature gradients. Unfortunately, the weather data missing from the Pig Alley weather station in October and November stymied our efforts here.

A result that was much more useful was the correlation between the data collected at the Port Hardy balloon station and the Quillayute balloon station. Each of these balloon stations had numerous gaps in the data. Due to the strong correlation between the two stations, we were able to fill missing values from one with data from the other to produce a unified data set with almost no missing values.

Another feature that showed some surprising promise during EDA was a correlation between wind direction and Peak chair open/closed status. The data appeared to show that winds from the south-west had a higher probability of causing the chair to not open. This was not a numerical correlation, but a visual one based on a plot of open/closed intensity for wind angles.

Finally, one feature we added to our data set was day of the week. We consulted with several people very familiar with both Whistler and weather. Each, including ourselves, were convinced that WhistlerBlackcomb makes a greater effort at avalanche control to ensure the Peak chair is open on weekend days compared to weekdays. After adding the feature in, we created a contingency table and ran it through a Chi-Squared test. The results were a p-value of almost 0.34. We retried the test after grouping days into weekend vs. weekdays, and long weekends (Fri/Sat/Sun/Mon) vs weekdays, but none of these iterations produced any statistically significant results.

Modelling

Due to our lack of experience and knowledge in climatology, our approach was to run our data through a suite of 5 classifiers each time we dropped or engineered a feature. We then tracked the scores of the models over each iteration to understand the effects. We tracked five metrics, of which the three most important to us were accuracy, recall of open days, and recall of closed days. The 5 classification models we began with were: Logistic Regression, Decision Tree, K-Nearest Neighbours, Random Forest, and Support Vector Machines.

After the completion of a baseline model and a major clean up of features, 4 further iterations produced minor improvements. Our weakest metric was recall of closed days. This was troubling, since telling people to stand in line for days the Peak chair would not open is a guaranteed way to ensure no one ever trusts your predictions again. We did have a class imbalance in our target of roughly 3:1, so we tried upsampling. The results were quite dramatic, causing a drop in many of our metrics, but importantly, it did increase the recall on our closed days considerably.

With the recall on our closed days improving dramatically on upsampling at the cost of other metrics, we now attempted to improve our scores further by building a stacked model. First we selected the model that had scored best on each metric/data set combination and ran each through a K-folds cross validation to optimize for each model's hyperparameters. This was done using the training data or the upsampled training data, depending on which model we were optimizing. We then trained each of our optimized models on the training data or upsampled data and produced the predicted probabilities. These became the training features for our final Logistic Regression Classifier. Once the final classifier was trained, we ran the test data through each of the first group of models with no upsampling of the data. From this, we extracted the predicted probabilities and formed them into the feature matrix to feed to the final Logistic Regression Classifier. The result of all this work was undeniably disappointing. We saw a several point drop in some of our metrics and no improvement.

With this reversal, we returned to our individual models and selected our strongest candidate, a Logistic Regression scoring 90.1% accuracy. Our next step was to check different decision boundaries to see if we could squeeze any further accuracy out of the model. A quick search showed that a decision boundary of 62% as opposed to 50% improved our model accuracy to 90.8%. Not a huge gain, but with a little rounding, put us within 4 percentage points of our goal.

Evaluation

With our goal almost in reach, we took some time to further evaluate our model to help inform what next steps can be taken. These evaluations produced some interesting insights.

First, we took a look at our mis-classified data points across each season. While not an explicit rule, over all the model was mis-classifying early season days as closed when they were open, and mis-classifying late season days as open when they were closed. We believe the early season mis-classifications can be solved with snowpack data from the Peak chair elevations. Further, we believe that the misclassification of end season days is due to the same effect. March and April can be some of the heaviest precipitation months, but if the freezing level is between the Pig Alley weather station and the Peak chair, what is showing as rain below could very well be snow up top. Snow means avalanche control, and that lowers the chances of the Peak chair opening.

The other interesting insights came as a benefit of the Logistic Regression being the best model. This allowed us to pull out the coefficients and calculate odds ratios. Three of the five strongest features increasing the odds of the Peak chair being open were a status of open for the Peak chair in each of the previous three days. These had odds ratios of 2.21, 1.61, and 1.13 respectively in order of increasing lag days. It is also worth noting that nearly three-quarters of our features had an odds ratio of one, meaning they neither increased or decreased the predictive power of the model and should have been eliminated earlier.

More exciting was that the snowpack feature had an odds ratio of 1.49. If we had been able to engineer the snowpack at 2,000m and get a tight range on Peak chair opening day, we could have binarized this feature. This would most likely have helped with both our early and late season errors. The snowpack at upper elevations tends to continue accumulating throughout the season. This means once we have passed the minimum snowpack required to open, that the snowpack feature will continue to increase the odds of the chair opening with every additional centimeter of snowpack. We actually want it to contribute nothing additional to the odds past the opening threshold.

Also very exciting were the coefficients of the mix ratios at 3,000m both at 4AM day of prediction and 4PM the day before. Mix ratio is the ratio of the mass of water vapour to the mass of dry air. This can be used as an indication of snow. The mix ratios at 4AM day of prediction and 4PM the day before modified the odds of the Peak chair opening by factors of 0.67 and 0.81 respectively. These were followed next in line by snowfall 24 hours before 6AM day of prediction and 48 hours before day of prediction with factors of 0.9 and 0.91 respectively. We believe that the mix ratio at 3,000m being more indicative of the Peak chair being closed than the snowfall at 1,640m is evidence that better snowfall and snowpack data at the Peak chair elevation would provide a boost to this model's accuracy.

Conclusions and Next Steps

This report has shown that while we have not achieved our target of 95% accuracy, we have come very close. We feel the result of 90.8% accuracy along with the issues identified throughout the project warrant a continuation of this project as a proof of concept. Specifically, next steps could be:

- Gather data to engineer snowpack at the Peak chair and binarize the snowpack feature based on this
- Bin the wind direction into NE, SE, SW, and NW cardinal directions
- Remove features of no significance to reduce noise
- Look at the correlation between mix ratio and snowfall to see if there is anything further to explore there
- Consult further with experts to see if more data can be used from weather balloon flights
- Approach the problem with additional models, including a time series ARIMA approach and utilizing neural networks

Acknowledgements

We would be remiss if we did not acknowledge and thank all the people who contributed towards this project in various capacities. First and foremost, Dr. Douw Steyn for all of his advice on weather data specifics and general encouragement, Mr. David Jones for pointing us to the weather balloon data sets and all the details around WhistlerBlackcomb's weather data collection, Mr. Todd Laney for providing us with his PowDay push notification database, Dr. Pascal Haegeli for connecting us to David Jones and steering us away from irrelevant avalanche forecast data, and last but not least, the whole BrainStation Data Science Education Team for all their advice and direction. Our thanks to each and every one of you for your support through this project.