EXECUTIVE SUMMARY

PROJECT: TABLE FOR FOUR

TEAM: WILDCATTS ANALYTICS

CATHERINE TOLLEY
TAMARA WILLIAMS
TOM ALIG
SHEELA RAO

May 13, 2018

Overview

Recruit Holdings Co., Ltd. is a global holding company, whose core business is divided into 3 primary areas: HR Technology, Media and Solutions (advertising), and Staffing.

Boasting over $16 billion USD in annual revenue, Recruit Holdings collects information and provides services in the housing, bridal, educational, automobile, travel, dining, and human resources industries, among others.

Recruit Holdings' AirREGI point-of-sale (POS) product currently offers cash register and table management services via iOS tablet and mobile devices for restaurants in Japan. A related Recruit Holdings business, Hot Pepper Gourmet (HPG), a restaurant marketing website and mobile app, provides Japanese diners with a searchable interface to locate restaurants and obtain promotional coupons. The company has targeted these two products for continued growth in 2018. Recruit Holdings has hired WildCATTS Analytics to further its vision of connecting customers with businesses, specifically in the Japan restaurant market. The proposed project will provide new business insights and drive shareholder value.

Problem Statement

By tapping into Recruit Holdings breadth of data collection in the Japan restaurant market, WildCATTS Analytics will add predictive modeling to the capabilities of the AirREGI system. This information will help restaurants be much more efficient and allow them to focus on creating an enjoyable dining experience for their customers.

Project Goals

The goals for this project are:

1. Develop a robust predictive model to predict daily restaurant volume up to 30 days in advance.
2. Develop a dashboard for the AirREGI POS solution to deliver predictions in context, enabling restaurant managers to take appropriate actions, including pushing customized promotions to the HPG marketing product.
3. Develop a prototype of a mobile experience for restaurant clients to interact with and potentially respond to volume predictions.
4. Maximize insights from Recruit Holdings' valuable restaurant sector data with meaningful data visualizations.

5. Recommend strategy to monetize prediction-enhanced AirREGI Point of Sale premium-subscription solution.

Achieving success in the project goals will enable Recruit Holdings to adopt the same capabilities across its multiple product channels worldwide.

Data Overview

The data come from a Kaggle competition to forecast restaurant visitors. The data include over 250,000 visits to 829 restaurants from January 1, 2016 - April 22, 2017. The objective is to forecast visitors to each of the 829 restaurants, each day from April 23, 2017 - May 31, 2017. Of note is that these restaurants are located in Japan, and a major Japanese holiday termed "Golden Week" spans from April 29 - May 5 each year. There are 14 distinct restaurant genres and 9 different prefectures in which the restaurants are situated throughout Japan.

The team has done a thorough review of the data. For instance, analysis of number of restaurant visitors based on day of the week and various holidays has been reviewed. Feature engineering data has also been done. As an example, median number of visitors per day, per restaurant, have been calculated.

In addition, WildCATTS Analytics has added outside data to the data set. Number of subway stops within a 500 meter distance of the restaurant, weather, number of attractions such as movie theaters, zoos, and nearby malls, have all been included for evaluation.

Outliers in the data have been identified and are in process of being evaluated for inclusion in model building.

Approach to Analysis

The approach WildCATTS Analytics has taken is to model the data using classic time series and boosted trees. As this is a Kaggle competition, the measurement as defined by Kaggle is to achieve the lowest Root Mean Squared Logarithmic Error (RSME). Additional methods may also be considered. In addition, clustering of the data will be completed to add insights to Recruit Holdings.

Boosted trees are attractive for handling complex patterns and relationships in the data, and time series will address the periodicity of visitor volume. Initial approaches to modeling include two different boosted tree algorithms and two different traditional forecasting algorithms. Initial modeling results are promising, particularly when using the augmentation data.

Forecasting methods were selected to model the time series. Exponential smoothing was chosen for its simplicity and auto ARIMA for its ability to handle seasonality, differencing and auto-regression, which is a linear combination of past values of visits. Time series models require complete historical training data. Missing values in the dataset were imputed.

An Exponential Smoothing (ETS) model was implemented without any transformations in the data. For the best ETS set of models, best AIC was 1191.18. Both ACF and PACF values are within acceptable limits.

The second modelling class used was auto ARIMA, which stands for Auto Regressive Integrated Moving Average. For the best auto ARIMA set of models, the package selected ARIMA(1,0,0) for (p,d,q). The AIC value was better than the ETS set of models, at 292.33.

Proposed RestauRANT Dashboard Scenario

WildCATTS Analytics is developing two Tableau dashboards to address the project goals. The first is a prototype dashboard for incorporating visitor predictions in the AirREGI POS product suite, aimed at enhancing value for restaurant managers. The second is a comprehensive dashboard for RH to assess and monitor their AirREGI business. Both dashboards are interactive and will leverage the enriched data developed during the project engagement.

AirREGI Manager's Scenario

Yuichi Sato is a restaurant manager in the Osaka prefecture. He manages two stores which specialize in serving Teppanyaki. As part of the value add for the Table for Four project, Recruit Holdings can now offer Yuichi a new way to view data on his restaurants. After going thru an RH setup experience to select his restaurants, Yuichi can see a map showing his stores, Yuichi's Original Teppanyaki, and Yuichi's Store Two. Yuichi now can have a feel for his restaurant volume the next few days. He can increase or decrease food orders and staffing and has the option to push out a marketing promotion, if necessary, to increase visitors to the restaurant.

Recruit Holding's Manager Scenario

Himari Mizutani is a Recruit Holdings manager overseeing their Tokyo, Niigata, and Miyagi operations. The new dashboard from WildCATTS Analytics allows her to easily visualize the performance of stores in her area. From the dashboard's landing page, Himari selects one or more of her prefectures. A view opens showing her the visitor predictions for the upcoming week, as well as links to additional charts

which could include: aggregated views of the visitors, reservations, or trends over time. Himari now has an easy way to get visual data on the stores she is responsible for.

Initial Conclusions

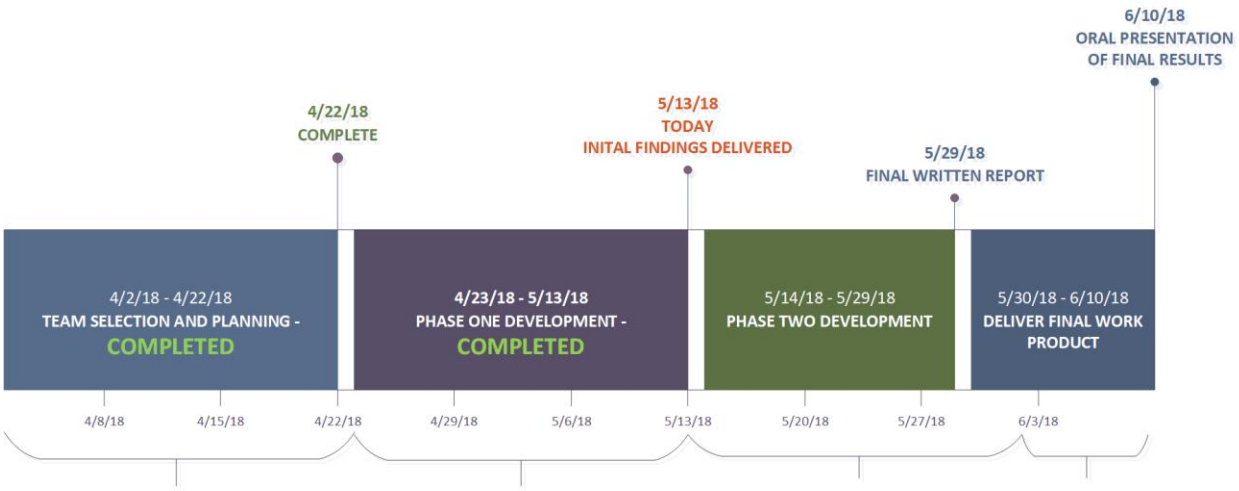
- While providing interesting background on the AirREGI/HPG eco-system, the reservation data is not particularly useful for predictive modeling.
- Imputation of missing values, based on patterns for individual restaurants is a successful approach.
- Time series modelling techniques are outperforming boosted tree models.
    - The addition of cluster membership information to the training feature set is beneficial.
    - Clustering also has uses for helping characterize the population to management.
- Based on our projected uptake and revenue models, the enhancements to the AirREGI product will yield an increase in first-year profits of $2.8 Million.

Project Status

The team is encouraged by the results thus far. There is additional modeling work to be done and it is anticipated the outcomes will further improve. The next phase also includes delivery of predictions and business insights through dashboard development.

The project is on time and on budget. We anticipate delivering all items outlined in our original proposal. These deliverables include a predictive model for guest visits to specified AirREGI restaurant, dashboard visualizations for use by RH management, and a dashboard visualization for use by an AirREGI restaurant manager on a mobile device. At this time, we do not expect any alterations to the original plan. The figure below depicts the project timeline and deliverables.

Project Timeline and Deliverables



Weeks 1 – 3

Initial engagement, formation of project team, proposal of goals

Deliverables to RH

- Initial statement of work and high level project plan

Weeks 4 – 6

Deep Dive into data, business case, and preliminary modelling. Initial results are present to executive at the end of milestone

Deliverables to RH

- Preliminary findings, project status,
- Executive Summary of work to date

Weeks 7 – 9

Final model selection, results collated, and final report along with summary written, and delivered to executive at the end of milestone

Deliverables to RH

- Final written report, predictive model(s), model code, data analysis, dashboard(s), mobile experience, and recommendations
- Executive Summary

Week 10

Results are presented to Executive in live presentation or pre-recorded video

Deliverables to RH

- Presentation for executive staff – slide deck and either live or pre-recorded presentation of the findings