BC2407 CASE STUDY PROPOSAL

Group & Team	Group 1, Team 2	Last Updated	Wed 11 Apr 2019

Case Study Title
Powering Airbnb's Growth through Personalisation

Sources and Datasets
The data is obtained from Kaggle: https://www.kaggle.com/c/airbnb-recruiting-new-user-bookings It contains the demographics, behaviours and web records of real-life customers who signed up on Airbnb as a member but may or may not have made their first booking. There are 12 possible outcomes of the destination country: 'US', 'FR', 'CA', 'GB', 'ES', 'IT', 'PT', 'NL','DE', 'AU', 'NDF' (no destination found), and 'other'. All the users in this dataset are from the USA, who visit Airbnb from 2009/03/19 to 2015/06/28.

Analytics Techniques
Power BI will be used to analyse the data for useful business insights and generate visualization for the importance findings. A list of Data Analysis and Machine Learning techniques from BC2406 and BC2407 will be applied in R and explained in the case study. ● Data Visualization ● Association Rule ● Logistic Regression ● Decision Tree ● Random Forest ● MARS

Focus

Airbnb provides convenient homestays and unique tourism experiences for its users. While Airbnb's usage statistics are still growing, its revenue growth is slowing which has caused worries over the sustainability of Airbnb's growth. According to a report from Morgan Stanley Research (Newsdesk, 2018), this is largely due to a saturated market, inactive users and increased competition from online travel agencies (Ting, 2018).

In order to solve the stagnating growth, Airbnb has to improve the **booking rate of users** and **decrease average time to first booking**. This can be achieved by **personalizing user experience** - providing the right content to the tight user at the right time. By accurately predicting **top 5 country destinations** that a new user will book their first travel experience, Airbnb can share more personalized content on the website and app. Though segmentation of users into urgency and non-urgency status, Airbnb can also have a more targeted marketing strategy for each user.

According to a report from RJ Metrics, after a new user makes their first booking, there is a higher probability that they will book again. Hence, it is very important that Airbnb's personalisation effort makes new users book and eventually increase overall booking rate.

In a challenge hosted on Kaggle, Airbnb provides a list of users data along with their demographics and some summary statistics about user behaviour. These data are to be used to predict which country a new user's first booking destination will be and whether the user needs to make a booking urgently.

Student learnings

Students and readers of the case will learn:

- The use of Analytics as a competitive strategy and disruptor to an existing market.
- The use of Analytics to understand customers' preferences and provide better user experience
- The use of Analytics to derive valuable insight for better business decisions.
- The correct development and testing of Analytics models.
- Application of Selected Techniques [as listed above].
- The techniques to improve the performance of models.
- The evaluation of models based on performance, interpretability and efficiency.

Content Headings

1. Case Overview
2. Business Background
3. Business Problem Statement

What business problem is Airbnb facing? Why does analytics playing an important role in the Airbnb operations? What are the business outcome measures and targets.

4. Analytical Problem Statement

What makes it challenging for Airbnb to predict travel destination and urgency status of new users with the data? Why this can help solve the business problem? What are the analytics performance measures and targets.

5. Data Preparation

What is the source of the data? Is it clean? Are there outliers, missing values or duplicated? Are there any extra information that can be extracted from the available features? Should there be factorisation? According to the business problem, should there be data splitting?

6. Data Exploration

What useful business insights can we get from the data through data visualisation to come out with recommendations for Airbnb?

7. Modelling

How to build powerful models for the prediction of country destination and urgency status? How to optimize them?

8. Model Evaluation

How is the performance, interpretability and efficiency of the models compared to each other? Which one is the preferred choice?

9. Recommendation

How to use those insights and models in the business?

10. Feasibility

Are those recommendations financially feasible and effective in achieving business goals? Are there any strength and weaknesses?

11. Limitations and Future Research Direction

Are there any limitations for the analytics and recommendations? How could we further improve the analytics and the model performance?

First Paragraph

As Airbnb celebrates ten years of operation in 2018, its rapid growth is slowing due to increasing competition, saturated market and inefficient marketing. The plateauing of Airbnb's growth highlights the importance of exploring innovative ways to improve its service and user experience such that more users will be engaged with the platform. To achieve this, Airbnb should personalise content on their website, mobile app and email marketing. Apart from providing the general guide to drive the business, analysis from data could also help create experience that is relevant to and targeted towards a user's needs, which could potentially results in higher booking rate, decreased time for new user to make their first booking and eventually, increased revenue growth.

Conclusion

In conclusion, data is valuable in conducting analytics and building models for Airbnb to provide better user experience by personalisation. With an accurate and efficient user travel destination and urgency status prediction models, Airbnb can provide personalized and targeted information with each user, attracting more users to book travel experience and avoiding wasted marketing resources, thus leading to larger revenue growth. The user community also benefit from better user experience.

Meanwhile, the model still needs to be upgraded and modified in the long run to ensure sustainable performance. Feedback, customer action data and customer demographics data needs to be continuously collected in order to measure performance and improve the prediction model if necessary.

Lead Questions

1. Given information about past users' information and booking history, what would a good "solution" look like that will be useful to Airbnb?
2. What could be some obstacles to achieving a good "solution"?
3. What are the analytics performance measures that should be measured?
4. How could one improve analytics performance?
5. What other data, besides the existing customer transactional data, could be useful?
6. What other useful insights could we get from the data provided besides from predicting users' first booking destination.
7. What could be the explanation for those observations?
8. How those observations could be used for better business decision makings?