Carnegie Mellon University
Tepper School of Business

99PLABS RESEARCH: "DYNAMIC VEHICLE SERVICES"

MS Business Analytic Program

Chatchawan Lakkhananukun
Prakhyat Pola
Rebecca Stevens
Jocelyn Wang

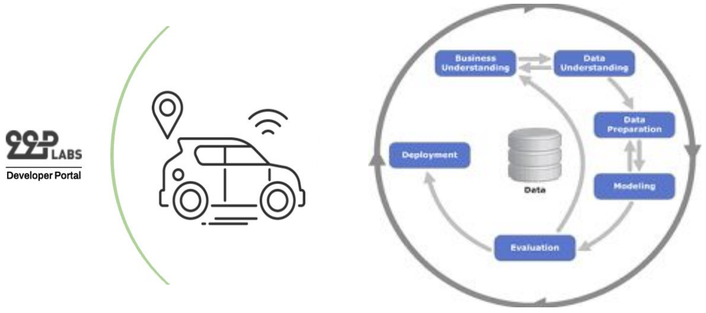
ABSTRACT

In this project, we explored two business models: a car maintenance service and a peer-to-peer car lending platform; we used an analytics-focused approach on the Telematics dataset collected on cars provided by 99PLabs. Dwell time and location predictive models were developed previously by another team, and we provided recommendations on the model improvement. The two business models integrate the predictive models to create potential business values. Because of the limitation of the dataset, two surveys were conducted to understand and validate our business models. Both surveys showed that there are markets for the two proposed business models. We also developed a framework to understand the failure modes of both business models and methods for mitigating the risk. We analyzed the cost-benefit of applying the predictive models to both business models to understand potential values for both 99PLabs and its customers. Finally, we explored the potential of using autonomous vehicles in the peer-to-peer lending service.

INTRODUCTION

Our team had the opportunity to partner with 99PLabs, a research group focused on producing innovative features, concepts, services and designs to remain on the cutting edge of the mobility industry and potentially transform the landscape of transportation. In a recent related project, a team of Berkeley students built a model to predict the dwell time and location of vehicles based on car sensor and infrastructure data from a leading automobile manufacturer. In the problem we've been given, the main question we want to address is: If we can predict the dwell time and location of our vehicles, what business opportunities can we leverage with this information?

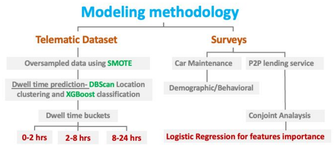
We used the CRISP-DM (Cross Industry Standard Process for Data Mining) framework as a guide as we worked through the project. Within this framework are six consecutive processes that we will discuss in the following sections (see figure below): Business Understanding, Data Understanding, Data Preparation, Modeling, Validation and Deployment.



METHODS AND MATERIALS

We extracted the Telematics dataset from 99PLabs, and cleaned and prepared features using PySpark. We executed a clustering algorithm on a set of important features and applied the XGBoost classification algorithm to classify the dwell time into buckets (0-2hrs, 2-8hrs, 8-24hrs) which outputted a high accuracy model. There were issues with imbalanced classes in the output, so we applied the SMOTE algorithm to oversample the dataset. We focused on the recall metric because the proportion of correctly predicted observations would provide the most value.

We conducted a survey for each business model to gauge the potential of the business use cases. The first survey focused on demographics and car maintenance behavior. Besides the demographic questions, the second survey also comprised of a conjoint analysis that sought to understand the importance of distinct features for both the lending and renting side for the car lending service. Logistic regression was then applied to the conjoint style questions to understand the utility scores. Together with the research on each business case, we constructed two business models that fit the predictive models.

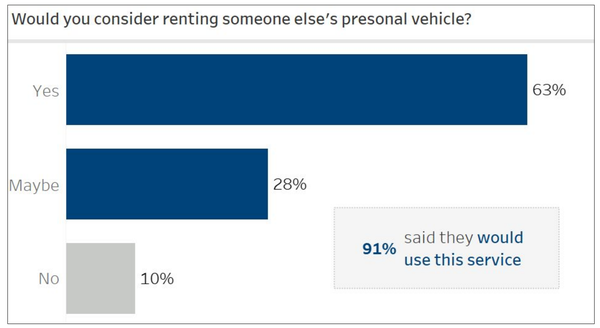


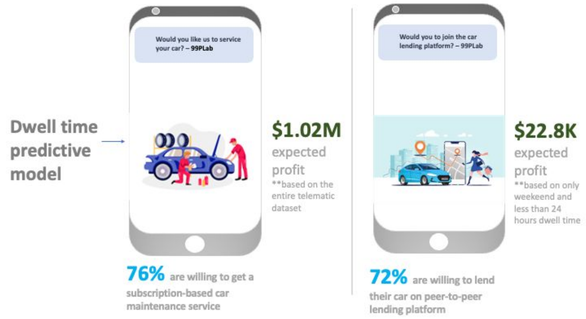
RESULTS

Using the oversampled data, the XGB model performed at 89.9% accuracy and had optimal and well-balanced recall scores (0-2hrs: 94%, 2-8hrs: 78%, 8-24hrs: 98%).

The survey results revealed that there are markets for both business ideas. In the first survey, 76% of survey takers said they would use the subscription-based car maintenance service. In the second survey, 72% of survey takers would consider lending their cars, while 91% would consider renting a car from the platform.

With these results, we proposed two business models that integrate the dwell time predictive model and the survey data – a car maintenance service and peer-to-peer car lending service. Both business models leverage the predicted dwell time as marketing tools. For example, if dwell time is predicted as suitable, customers will be notified to accept the car maintenance service or to participate in the peer-to-peer lending platform.





DISCUSSION

There are risks associated with both business models, especially when both are based on a predictive model. The major concern was that if the predictive model cannot predict, what would be the cost associated? With this in mind, we decided that the lowest risk use case would be to include humans in the loop so that the risk can be mitigated by a confirmation from the potential customers. We conducted research on FMEA style risk and rated different risk scenarios should 99PLabs decide to pursue any of the business models further. We also developed a cost-benefit analysis framework to quantify the cost and benefit when the predictive models are integrated to the business models.

Soon, the number of autonomous vehicles(AV) will significantly increase, and this will create another business opportunity for 99PLab, especially for the peer-to-peer lending service. Instead of relying on the supply from the lender's side, AV can help increase flexibility for lending cars, broaden the reach into remote areas and reduce the risk of incorrect predictions. In this scenario, dwell time predictive models can be used in a supply and demand application.

CONCLUSIONS

The two proposed business models, integrated with the predictive models, depicts great potential for 99PLabs. With the dwell time model, survey results, and the research conducted in this project, both business models present promising opportunities. The model performance and survey response validated the market for 99PLabs. A risk framework consisting of FMEA tables and cost-benefit analysis were developed for quantifying the model risks as well as possible methods to mitigate them.

There is, however, significantly more work that needs to go into each of the business cases. The two weeks of available data were collected in 2018, and the distribution of the data may have changed since then. Another challenge is that the with the constant inflow of data from the car sensors, a strong data engineering infrastructure will be required to handle and stream in this data before 99PLabs can consider deploying the model into production.

CONTACT

Rebecca Stevens
Carnegie Mellon University
rlsteven@andrew.cmu.edu
(978) 290-7174

REFERENCES

1.*https://developer.99plabs.io*
2.*https://medium.com/99p-labs/vehicle-location-and-dwell-time-prediction-conclusion-e5d4b54c25d9*