Mobility in the age of massive data: opportunities and challenges

Douglas Ferreira
Localiza Rent a Car, Brazil
douglas.sferreira@localiza.com

Fabricio Lucas
Localiza Rent a Car, Brazil
fabricio.lucas@localiza.com

José Gonçalves
Localiza Rent a Car, Brazil
jose.goncalves@localiza.com

Pedro Kassio*
Localiza Rent a Car, Brazil
pedro.kassio@localiza.com

Samuel Oliveira
Localiza Rent a Car, Brazil
samuel.oliveira@localiza.com

Yuri Aurelio
Localiza Rent a Car, Brazil
yuri.aurelio@localiza.com

Gustavo Bergamo
Mobi7, Brazil
gustavo.bergamo@localiza.com

Paulo Born
Mobi7, Brazil
paulo.born@localiza.com

Angelo Ciarlini
Localiza Rent a Car, Brazil
angelo.ciarlini@localiza.com

Abstract

The development of IoT and machine learning in recent years has brought new opportunities to the mobility industry. The amount of data collected from the fleet with telemetry is not only more accurate but also more available with lower IoT costs. Customers' experience and profitability can be dramatically improved by means of using IoT datasets and machine learning models. These opportunities however come together with Big Data challenges. In this paper, we describe some of the opportunities and challenges that Localiza, the largest car rental company in Brazil, has been addressing as part of its digital transformation.

Keywords: IoT, machine learning, artificial intelligence, telemetry, optimization

1 Introduction

The digital transformation using Machine Learning and AI has impacted different industries and created numerous opportunities to improve services, reduce costs and leverage revenue. In the mobility sector this trend is very clear. The improvement of sensors, internet and compute power provides the means to dramatically change the sector. It is now possible to capture, transfer and process data in high volume and speed, amplifying the opportunities to develop new features and services. Real-time remote-controlled devices, online traffic light optimization, self-parking cars and fully autonomous cars are starting to become more common, making the Internet of Things (IoT) a reality of our daily lives, and not a hype [1].

In parallel, there is also a decrease in the barrier of access to technology when compared to previous decades. According to the 2021 Annual Survey of the Use of IT by Fundação Getulio Vargas [3], since 2017 the number of smartphones in Brazil has already exceeded the number of inhabitants. Those smartphones become part of the mobility ecosystem since they work as devices that send real time data information including GPS position, user preferences, incidents on the road and traffic. In line with this phenomenon, there is a change in the customer's behavior, which starts to demand more personalized products and services, in the shortest possible time.

These technological advances have brought countless opportunities for companies and industries in the mobility sector. In the next sections of this article, it will be presented some opportunities and the challenges faced by Localiza to leverage its business through the use of IoT and artificial intelligence. Localiza is the largest car rental company in Brazil with 500 thousand cars and more than 600 branches.

2 Opportunities

Following the technological evolution and the change of customer's behavior, Localiza has identified different opportunities for improvement and innovation in the sphere of products, services and processes. In the next subsections we will talk about some of these opportunities that Localiza has been closely monitoring related to IoT and Data Science.

2.1 Preventive maintenance optimization

When renting a car, one of the fundamental points sought by customers is vehicle reliability. To guarantee a good reliability, it is crucial to ensure that maintenance is up to date, enhancing safety and customer satisfaction. At the same time, an excessive number of maintenance interventions increases operation costs and reduces the number of vehicles available for renting. Identifying the right time to perform preventive maintenance is critical to provide a safe and pleasant experience, as well as to optimize costs and improve operation planning. In a company with several different car models

In: WebMedia in Practice, Curitiba, Brasil. Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia). Porto Alegre: Sociedade Brasileira de Computação, 2022.

ISSN 2596-1683

from different automakers the challenge is even greater as data quality is crucial to achieve good results.

2.2 Car damage detection

With computational advances in recent years, there has been a great evolution in unstructured data processing such as image, voice and text. Deep neural networks had a remarkable development and became able to identify images and objects with high accuracy. To get an idea of this evolution, in 2013 Alexnet [5] achieved an accuracy of 63.3% in the ImageNet benchmark base [2]. At that time this was considered a great achievement; however, today, nine years later, the CoCa neural network [6] reached 93% accuracy. This technological advance has allowed companies to make use of this technology in automating processes, replacing human activities and helping to prioritize activities.

In the car rental process, a prior evaluation of the vehicle is carried out before the customer picks it up and a comparison is made when it is returned. In periods of peak demand, it is necessary to carry out an even faster inspection of the vehicle in order to guarantee a continuous flow of withdrawals and returns, and this can lead to situations where a small damage can go unnoticed. There is then an opportunity to use computer vision algorithms and telemetry data to assist in damage detection and prioritization of vehicles to be inspected by a worker.

2.3 Fleet profiling and optimization

With the evolution of smartphones and automotive telemetry, it is possible to capture traffic information, vehicle position, routes, speed, mileage, fuel consumption and customers' driving patterns and preferences. These kinds of data coupled with business information such as number of vehicles per branch, forecast demand, Net Promoter Score (NPS), customer feedback, fleet age, etc., allows car rental companies to make better decisions regarding the assignment of cars to customers and fleet management

Localiza currently has more than 600 branches and a fleet of hundreds of thousands of cars which is constantly renewed and distributed throughout Latin America, mostly in Brazil. This high number of branches spread across different regions brings unique challenges regarding car allocation, not only by the number of cars but also by the demographic dimension, price, seasonality, culture, digitalization and regional preference of users for specific models. In order to ensure optimal profitability linked to a high level of customer satisfaction, there is an opportunity to use optimization algorithms to make key decisions, such as which car should be rented to each customer, how cars should me distributed over the branches, which car models should be bought and which and when can cars should be deactivated and sold.

2.4 Driver Behavior

Driving pattern behavior is an important factor in car accidents. It can dramatically increase the risk of collision, severity or even fatality of accidents. Approximately 95% of traffic collisions are due to human error [**?**].

Insurance pricing for any asset often considers generalized and non-personalized information regarding the use pattern of the product or service. However, a more customized pricing strategy allows companies to offer a discount and fairer price to customers who drive safer, this is also an incentive for safe driving.

Telemetry data from current and previous rentals can be used by machine learning models to understand drivers' behavior and estimate the risk of accident when a specific customer rents a car. This can be used, even at real time, to notify the drivers and improve their safety. In addition, personalized price offers and rental conditions become viable, reducing risks and improving profitability

2.5 Fraud and theft prevention

As in any type of business, unfortunately in the car rental segment, fraud or theft also happen. The occurrence of fraud makes the business more expensive and processes of checking customer information more time-consuming and exhausting from the user's point of view.

Like many other frauds that occur in other business models, fraudsters tend to follow some type of behavior pattern and in this scenario, telemetry comes in as an alternative for tracking vehicles, providing information to identify patterns. By means of these patterns, we might avoid renting cars to fraudsters and increase the chances of recovering stolen cars.

3 Challenges

Extracting value from massive mobility datasets and making decisions based on them can dramatically improve the way companies offer services and scale their activities. It is a typical Big Data context, in which we must deal with challenges related to volume, variety and velocity of the data.

Telemetry data obtained by means of IoT devices is the main type of dataset we need to process. Challenges related to data gathering, preprocessing and making data available in due time are far from trivial.

The complexity starts in the IoT edge sensors, which often require customization depending on vehicle and model. At the edge, embedded algorithms need to run on high-end tracker devices, to evaluate and respond to every vehicle stimulus in real time ensuring that trustworthy data is being collected at the source constantly.

Once the data arrives to the IoT ingestion platform, it needs to be unpacked, translated, enriched, and normalized. After that, it then can be used to trigger pre-programmed rules and processes. Thus, from this stage on, dealing with volume, variety, and velocity of the data comes into play.

Localiza, for instance, has to process about 100 million daily data points in real time. In addition, all inherent possible IoT issues, such as lack of connectivity and out-of-order samples, need to be properly addressed .

There are also multiple challenges when we need to combine telemetry data, both from mobile and static sensors, with video (provided by cameras), maps and weather information. These datasets are not structured, most of them are continuously generated and need to be synchronized in real time, considering space and time. For instance, we might want to evaluate the probability of accidents due to the weather and specific conditions of the roads (captured by means of cameras) when drivers behave according to a specific pattern.

Many of the applications compute conditions and features in real time, either to score a machine learning model or just to continuously monitor events. In addition, some mobility applications demand a very frequent or even a continuous update of machine learning models. Choices about where and how datasets are stored and processed determine how efficient and cost-effective our applications will be. Data might need to be partitioned according to space, time or any other meaningful feature used to join different datasets. Data locality is crucial as transferring huge datasets on-the-fly might not be viable due to impact on time or on costs. Assume that we need to find out which cars were in the same location for time intervals of more than 10min. In this case, partitioning datasets accordingly increases the level of parallelism and improves performance. This process is not however trivial as we need to consider different granularities of space and time and borders between partitions.

Finally, data privacy is also a key issue in mobility. Information about where customers are at specific times can be very useful to provide services that benefit them, but this is very sensitive data that need to be protected according to the law. Patterns of mobility are a rich source of information for many business applications, but the interest of the customer should prevail over any unintended use of this information. In this way, the implementation of frameworks that reinforce data privacy policies for huge mobility datasets is necessary.

4 Concluding Remarks

In the last 3 years, Localiza has been working intensively on IoT and machine learning algorithms. During this time, Localiza created a Data Science and Analytics team and acquired Mobi7 [4], which became Localiza's IoT and Telemetry branch. Multiple use cases have been implemented and others are under development. Among those that resort to IoT and machine learning we can point out the creation of driving behavior scores and the detection of the probability of theft. Driving behavior scores have been used to encourage drivers to improve their driving pattern by means of bonuses and special deals. The detection of the probability of theft was instrumental in the reduction of the of assets by 59% in 2021 when compared to 2020. In the near future, Localiza intends to continue its effort to fully capture the benefits of combining IoT and machine learning in the mobility sector. Such a combination tends to dramatically improve customers' experience, reduce costs, increase profitability and create new opportunities for the business.

References

[1] Gustavo Bergamo. 2021. IoT: hype ou fruto da evolução natural? (December 2021). https://medium.com/localizalabs/iot-hype-ou-fruto-da-evolu%C3%A7%C3%A3o-natural-d2f0f00fb53d

[2] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 248–255.

[3] FERNANDO S MEIRELLES. 2022. Pesquisa Anual do FGVcia. *Uso de TI nas Empresas* 33 (2022).

[4] Ricardo Novo. 2021. Por que a Localiza adquiriu a Mobi7? A estratégia por trás da aquisição de uma empresa de telemetria. (November 2021). https://medium.com/localizalabs/por-que-a-localiza-adquiriu-a-mobi7-e224dcf11f91

[5] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. 2015. Imagenet large scale visual recognition challenge. *International journal of computer vision* 115, 3 (2015), 211–252.

[6] Jiahui Yu, Zirui Wang, Vijay Vasudevan, Legg Yeung, Mojtaba Seyedhosseini, and Yonghui Wu. 2022. Coca: Contrastive captioners are image-text foundation models. *arXiv preprint arXiv:2205.01917* (2022).