TicketMaster: Wanna Collaborate?

Kexin (Colleen) Su (ks3198) & Yilu Fang (YF2505)

A14: Automatic Consumer Sales Leads Finding

02/13/2020

Abstract

In this project, we will design a system that automatically helps ticket sellers promote their concert tickets by recommending potential customers. Our system will utilize methods such as NLP, topic modeling, knowledge graph, etc to analyze twitter user's information, and determine if the user is a potential customer. In the end, the system will return a list of potential customers, their twitter id, and potential contact information.

Background & Literature Review

Sales lead is a person who may eventually become a potential buyer of your product. A high-quality and revenue-producing sales leads can increase the sales of the product. As Hubspot has reported in 2018, 60% of the marketers say it's a great challenge to generate traffic and leads. In reality, a significant amount of salespeople complain that they encounter and waste a lot of time on bad leads.

In our opinion, the first cause of the bad leads is aiming at a customer group that is too big. Companies always assume that their products are for everyone. However, only 22% of the companies are satisfied with their customer conversion rate. They spent a lot on acquiring

customers, but in reality only a small part of customers worth their investment. For example, not all the users on Twitter like Korean music, therefore it would not be efficient to promote Kpop concert tickets to every user. It's essential to distinguish the leads that are the best fit for the products.

Another cause is that those people might not have a need for your product. They may be interested in the products such as Black Pink's concert tickets but have already bought them from another merchandiser, or they may like the performer but don't have the time or the money to go. Under these circumstances, we decide to design our own lead generation process to mine the sales leads for the concert tickets.

Marketing automation is a popular method for creating a personalized customer experience. The Strategic IC(2017) has reported that the revenues of the companies with lead management automation increase 10% or more in just 6 to 9 months. More specifically, Lai et al. (2018) employ collaborative filtering, TFDF, and optimization algorithms to design an AI-empowered traveling support recommendation system. Zhang et al. (2019) have constructed a personalized web-based movie recommendation system, addressing two issues: scalability and feedback, and verification of actual usage. Zeng et al. (2018) design a leads generation robot that automatically mines the leads with steps: employing automatic web crawling, utilizing NLP to extract leads, proceeding intelligence recommendation through word2vec based text analysis.

After doing some investigation, we discovered that the commonly seen personalized product recommendation system recommends products based on the user's preferences. Therefore, we come up with the idea that we should focus on the sale's side. After the seller input the product that he wants to sell, our item-based system will automatically return a list of targeting customers.

Introduction to the Project

As mentioned above, the main difference between our system and the previous personalized recommendation system is that we focus more on the seller's need rather than the buyer's need. The system we want to build in this project is a complete workflow that targets the ticket seller's need for potential customers.

The system is composed of two primary parts. In the first part, we will start by collecting all available concert information from the merchandiser, then our system will automatically determine concerts that need to be promoted. To achieve this, the system will either analyze current entertainment/music news and promote concerts that match the current trend, or it will analyze the ticket information of each concert and promote concerts whose tickets are less likely to be purchased. The merchandiser can decide between the two methods and choose whichever is more beneficial to them.

After we decide the concert that needs to be promoted, the system will proceed to the second part. In this phase, the system will analyze the personal information as well as historical tweets

posted by Twitter users using methods such as NLP, knowledge graph, clustering, etc, and discern which users are more likely to purchase the ticket for that specific concert. In the end, our system will return a list of twitter-id and other contact information of the potential users to the merchandiser, so that they can reach out to them to promote the concert.

For example, suppose we need to choose one out of three concerts to promote (Ed Sheeran, Adele, and Black Pink). After analyzing current entertainment news, the system found that Kpop is very popular recently and decides to promote Black Pink's concert. Then the system will continue to analyze twitter users' information. If a user's location matches the concert's location, and he/she has shown interest in either Korean culture, Kpop songs or Black Pink specifically in his/her past tweets, the user will be considered as a potential customer.

The main difficulty for this project is to choose, out of all the potential analytical methods, the ones that are the most relevant and effective in detecting if a user is a potential customer and match him/her to the corresponding concert. Another challenge is to choose the revenue-producing customers from those who show interest in the concert by semantic analysis.

Introduction to the Dataset

For the first part of our system, we will need data on concerts and ticket information, as well as information on entertainment/music news. For the concert dataset, we found a website called Concert Diary.com that contains all the concert information in the UK starting from the year 2000, but if we want to look at concerts in another country, we will need to find other datasets. A

problem for concert-datasets is that most datasets do not have ticket information that tells us how many tickets are sold. Therefore, after talking to the professor, we decide that we will come up with our own ticket information for the sake of system construction. For news datasets, we found that the GDELT website contains very holistic news information that we can use.

For the second part of our system, we will be using historical Twitter data. We found a dataset on Internet Archive that includes twitter stream information starting from 2011. We also found another dataset, "Million Musical Tweet", that has listening history corresponding to each Twitter user that can be used to further analyze a user's music preference.

A difficulty in using data from social media is that due to the privacy policy of many platforms, it is difficult to find a dataset that matches each tweet to the specific user that posted it.

Plan

Milestone 1: Select the topic of our project, investigate in the field of product recommendation system and marketing, search for the potential innovative aspects, and look for the potential datasets and data API.

Milestone 2: Complete the first part of our system: select concerts to be promoted. Collect and clean the data, look for theories and algorithms that can be applied to concert selection, design and construct the first part of our system.

Milestone 3: Complete the second part of our system: find the target customers. Collect and clean the data, look for theories and algorithms that can be applied to concert selection, design and construct the second part of our system.

Final: Improve the efficiency and accuracy of our system by trying out different mechanisms, implement the front end, finish our final report and video, and make our final presentation.

Reference

Dapergolas, Gerry. "40 Statistics To Inform & Enhance B2B Lead Generation Strategies." *Link to Our Blog*, 2017, blog.strategic-ic.co.uk/lead-generation-statistics.

HubSpot. "2020 Marketing Statistics, Trends & Data - The Ultimate List of Digital Marketing Stats." *2020 Marketing Statistics, Trends & Data - The Ultimate List of Digital Marketing Stats*, www.hubspot.com/marketing-statistics.

J. Zeng, J. Che, C. Xing, and L. Zhang. "LeadsRobot: A Sales Leads Generation Robot Based on Big Data Analytics." In *International Conference on Big Data*, pp. 277-290. Springer, Cham, 2018.

J. Zhang, Y. Wang, Z. Yuan and Q. Jin, "Personalized real-time movie recommendation system: Practical prototype and evaluation" in *Tsinghua Science and Technology*, vol. 25, no. 2, pp. 180-191, April 2020.

K. Lai, N. Y. Yen and M. Chen, "Design of an AI-Empowered Recommender System for Travelling Support: Individual Traveler as an Instance," *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence*

and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), Athens, 2018, pp. 343-346.