Research and Application of Churn Prediction Model of Meituan Takeaway Rider

Summary

With the acceleration of people's pace of life, takeaways have become one of the main dining options for urban residents. At the same time, the emergence of takeaways is also the general trend of informatization. In the increasingly fierce competition in the take-away market, various types of takeaway platforms have formed their own set of operating modes, and the delivery methods have become more diverse as each business expands. Crowdsourcing rider platform as the main take-away capacity, how to reduce the loss of riders and discover riders who are prone to loss as early as possible has become the core work of more and more food delivery companies' rider operations. The traditional user churn prediction model has shown its deficiencies, including the rough selection of features, the relatively simple churn algorithm, insensitivity to the user's presented group features, and ignoring the information contained in the user's internal connections.

This paper comprehensively analyzes the existing churn prediction algorithms, including logistic regression algorithm, decision tree algorithm, support vector machine algorithm and neural network algorithm. Cluster analysis of the rider is introduced. In the rider clusters, four churn prediction models were established separately. Optimization model for improving the prediction performance of the model was built for each algorithm and the best churn prediction model was selected for each group. Finally, a collaborative filtering algorithm was introduced to iterate the prediction results.

From the actual modeling results, the prediction effect of each algorithm established after clustering has been improved. The prediction accuracy of the whole sample and the prediction accuracy of the loss rider are higher than those of the non-clustering model. For different types of riders show different behaviors, the prediction performance of the four types of prediction algorithms is different in rider subdivisions. The prediction effect of support vector machines was generally poor, and the decision tree model was accurate in the prediction for good balance was achieved between the overall prediction rate and the accuracy of the loss rider's forecast.

Key words: Cluster analysis; Logistic regression; Decision tree; Support vector machine; Neural network; Collaborative filtering