Research and Innovation Project for Moobifun Company

Louis Fippo Fitime

Lyon, France

Abstract

Big data and artificial intelligence offer enormous opportunities in the automatic and intelligent control of commercial, financial, industrial and strategic operations. As part of its activities, Moobifun company support its customers in the framework of the management of gaming operations by setting up an architecture for the collection, structuring, storage and exploitation of the massive data generated by these operations.

Exploitation of data as part of Moobifun's activities, consists of designing the best models for customer knowledge in the aim of recommended suitable marketing and commercial products for players. Another challenge under data exploitation is to provide suitable stochastic parameters of payout distribution. The goal is to limit risk and maximize profits.

Keywords: Big data, Machine Learning, Bandit Problems, Recommender systems

1. Introduction

The Moobifun company is a computer company specializing in mobile services with high added value. It provides customers with mobile solutions to deploy their applications or services via mobile technologies. The main applications developed by Moobifun allow customers to connect and benefit from services that can be, for example, lotto games, sports betting or even access to information portals on their mobiles. This interaction between the applications developed by Moobifun and the users on the one hand, and the Moobifun clients on the other hand, generates an enormous amount of data that must be manage and exploit efficiently.

There are many ways to exploit this data. We can quote the most urgent and the most relevant to the needs of our customers. The first track is the implementation of recommendation systems for lotto gaming operations. The idea is to recommend the parameters, the distributions and the payouts in order to retain the players. Ultimately, we want to implement and deploy an artificial intelligence engine that will drive our lotto game operations independently.

Another track is to be able to know the gambling habits of the players in order to develop and offer them new games and new personalize commercial and marketing products.

2. Recommendation system for gaming operations

In this part we will dwell on our most immediate project, namely the implementation of a recommendation system for distributions and game parameters. As we mentioned in the introduction, the idea is to recommend the parameters, the distributions and the payouts in order to retain the players. To achieve this goal, we want to rely on large amount of transactional data collected and stored in our big data Hadoop platform on the one hand. On the another hand, we want to take advantage on new methods develops in the area of machine learning and deep learning to build power full engine capable of generate more precise parameters and distribution based on the observed data.

The main axes of this work are :

- Build an abstraction of our problem as a bandit problem or others related models
- Learned parameters, distribution and payout rate from data
- Build a simulator for generate interesting scenarios

2.1. Build an abstraction of our problem as a bandit problem or others related models

For this part, the idea is to doing a review of the literature to identify the approaches and models that would allow us to abstract our problem in order to derive the interesting properties to improve our distribution and payout parameters.

The bandit models [1, 2] seem very interesting to answer our problems of setting of parameters of distribution and payout. This first exploratory work will therefore aim to carry out, if possible, a first modeling and first experimentations of our problem by using bandit models or more appropriate models.

2.2. Learned parameters, distribution and payout rate from data

This last decade, machine learning, deep learning and reinforce learning methods have attracted an important interest for different applications as customer segmentation [3, 4], as customer retention, as real time decision, as speech recognition [5], natural language processing [6], and image recognition [7]. The idea is to take advantage of the power of these methods and offer an estimate of the most appropriate parameters, winning distributions and payout rates for better player retention while ensuring optimal gains.

Some precautions should nevertheless be taken into account. Firstly, multiple game operations share the same target population. Secondly, the population does not always make the difference between the proposed games. Thirdly, distributions and redistribution rates must be estimated in real-time from the large scale data collected.

2.3. Build a simulator for generate interesting scenarios

In this last part, we aim at build a generic simulator to prior simulate in silico our hypothesis before launch a gaming operation. Ideally, the simulator should reproduce the players behaviors based on parameters estimated from observed data as closely as possible. The objectives are to simulate all possible scenarios before launch a game operation to avoid unpleasant surprises.

3. Customer knowledge

In this section, we are interest in customer knowledge and the income prediction. For that, we don't have personal informations on players and we should only consider historical behavior.

The various tasks to be implemented to meet the operational need are the following:

- Design *RFM Model*

This model means:
Recency When last did a player play? (Hour, Day, Week) Frequency How often does this player play (Once a day, Once a Week, In the morning, afternoon) Monetary How much does the player stake in a session (Per game, per games session, lifetime) LTV Lifetime Value of a player derived by Buy minus Win (GGR) over the age of the account.

- Prospect − > Participant − > Player

  Prospect is defined as any marketable player that we are willing to market to (SMS Black, underage non-prospect).
  Participant can consume the games without any financial stake Player stakes and wagers on our platform and player status is determined by his RFM-model scoring.

- Acquisition − > Retention

  The objective is to retain an acquired player for the maximum value period. The cycle of acquisition, conversion, retention and reacquisition applies to our basket of players.

- Acquisition Channel Tracking

  In order for the business to make the most appropriate investment in marketing channels, tracking behaviours and cost per channel allows the business to make decisions based on the current acquisition sentiment in the business.

- Acquisition Cost − > ROI Period

  By tracking and monitoring the acquisition channels, the business will be able to derive a relatively accurate cost per acquisition and can then decide based on the appetite of the business what level of investment each channel (radio, TC, Free Draws, etc) should perform on.

- Profile − > Behavior

  The current profile is a flat view of the player and his interaction with the casino. We were able to devise methods to track the player and multiple data and trigger points to guide the casino in actions that will maximize the value of the player.

Cluster Behavior per Game. There will be identifiable activities by the player that will allow us to map these groups together to do more intelligent and focused marketing.

- GGR

  Positive / Negative Based on the GGR income of the player including attributable market spend to acquire the player. (E.g. SMS Cost, Free Ticket Cost are GGR negative events) Theoretical / Actual Based on the payout parameters there is a determination of the theoretical values vs the actual values generated during game play (Luck factor / Deviation intentional/unintentional)

To meet all these operational needs, we will use the best machine learning techniques currently available and we will also develop more adapted techniques if the need arises.

From all these questions that we are asking ourselves from an operational point of view, we think we can highlight important research questions.

[1] H. Robbins, Some aspects of the sequential design of experiments, in: Herbert Robbins Selected Papers, Springer, 1985, pp. 169–177.

[2] W. R. Thompson, On the likelihood that one unknown probability exceeds another in view of the evidence of two samples, Biometrika 25 (1933) 285–294.

[3] H. Hwang, T. Jung, E. Suh, An ltv model and customer segmentation based on customer value: a case study on the wireless telecommunication industry, Expert systems with applications 26 (2004) 181–188.

[4] R. A. Soeini, E. Fathalizade, Customer segmentation based on modified rfm model in the insurance industry, in: Proceedings of 4th International Conference on Machine Learning and Computing IPCSIT. Singapore: IACSIT Press, volume 25, pp. 104–108.

[5] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, et al., Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups, IEEE Signal processing magazine 29 (2012) 82–97.

[6] J. Weston, S. Chopra, K. Adams, # tagspace: Semantic embeddings from hashtags, in: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp. 1822–1827.

[7] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, pp. 1097–1105.