ECON900–PS1

Yujie Deng*
Clemson University

April 24, 2019

1 Motivation

People play board games across the world. Today's generation play board games on tablets or smart phones. Before the digital age, millions of families spent pleasant evenings around a physical board with dice, tokens, and cards. Before physical board games, people drew the games in the sand and played with seeds and stones.

A series of 49 small carved painted stones were found at the 5,000-year-old Başur Höyük burial mound in southeast Turkey. These are the earliest gaming pieces ever found. Similar pieces have been found in Syria and Iraq and seem to point to board games originating in the Fertile Crescent. The Crescent is comprised of regions around the Nile, Tigris, and Euphrates River in the Middle East.

As centuries of evolvement, nowadays board game can be classified into several categories: Roll and Move, Cooperative, Deck-Building, Area Control, Secret Identity, Legacy, Party, Puzzle and Combat. By product type, it can be segmented into tabletop, card & dice, collectible, miniature and RPG game, among which tabletop games dominated the largest market share in 2017, growing at a forecasted CAGR of over 11% according to a global outlook and forecast research.

2 Data Description

The boardgame dataset is collected from www.boardgamegeek.com. This dataset contains 2388 observations because of the missing value caused by the listed price variable.

*Clemson University 406 Sirrine Hall Clemson SC 29634, 330-274-6822 deng2@g.clemson.edu

Six variables are choosen for my analysis. The dependend variable is the rank of each boardgame on boardgamegeek website. Five explanatory variables are the average rating of each boardgame, the geekrating of each boardgame, the year when each boardgame has been released, the listed price of each boardgame, the number of voters who have been voted for each boardgame. The summary statistics of explanatory variables are shown below:

	mean	std	max.	min.
boardgame_price	40.47	24.65	3.99	299.99
boardgame_avgrating	6.35	0.92	9.6	1.05
boardgame_geekrating	5.7	0.38	8.61	3.46
boardgame_voters	769.66	3049.29	84569	1

3 Analysis

I want to predict the rank of a new boardgame on boardgamegeek.com given the year it releases, the number of voters who vote it, the average rating it gains and also the geek rating.

I use two methods: support vector machine and random forests regressor. Support-vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Here, I use it as regression analysis. Random forests regressor is also supervised learning.

4 Result

After running each model, I get R square score for each model. For this specifc dataset, random forest regressor can predict better because the variable, geek rating, influnce the rank of the boardgame on boardgamegeek.com a lot. If you give me a new boardgame with these 5 characteristics, I can give you a pretty accurate rank of this boardgame on boardgamegeek.com.