Pre-release New Product Sales Forecast

Using Customers' Dynamic Knowledge Collaborative Activities

Keywords: Pre-release New Product Sales; Dynamic; Knowledge Collaboration; Predictive Modeling

Firms launching new products are keenly interested in the likely sales projections of the products. Most sales forecasts are based on data from the initial periods of product sales. However, for firms having information on likely sales before the product is launched would be valuable as it allows them to modify the product if necessary and make marketing investment decisions more efficiently. Prior research has utilized the wisdom of the crowds to make forecasts regarding stock market prices, election results and quiz shows. These forecasts are based on static information from the crowd that is collected at a single point in time. However, crowd views are dynamic and are often updated. But collecting such data is expensive and potentially infeasible. Recent research in marketing has captured pre-launch consumer sentiment to forecast new product sales. We use Wikipedia, as a platform to capture the crowd (consumers) dynamic collaborative knowledge sharing about product prior to their launch. We address the following research question, can the dynamic process of customers' online knowledge-collaborative activities during the pre-release period be used to develop the opening-week sales forecast of a new product (*hereon* pre-release NPS forecast)? To empirically address this question, we use a sample of 340 new video games launched in the U.S. market from 2011 through 2017.

Formally, knowledge collaboration on an OC platform refers to the "individual acts of offering knowledge to others as well as adding to, recombining, modifying, and integrating knowledge that others have contributed" (Faraj, Jarvenpaa and Majchrzak 2011). In the context of pre-release NPS forecast, an iterative (i.e., dynamic) editing and collaborating process strengthens 1) knowledge contributors' cognitive engagement with the new product as potential customers before the release; and 2) the content created through this process is usually richer and less biased, thus more beneficial to educate regular customers and empower them to evaluate the to-be-released new product when making the purchase decision (Wagner and Majchrzak 2006; Wasko and Faraj 2000, 2005).

However, measuring such a dynamic process is not easy. Customers tend to be more active toward the end of the pre-release period, leading to a very sparse dataset distribution at the beginning, but super dense toward the final moments before the release. Regular time series analysis (TSA) assumes the probability of all time events are uniformly distributed (Cryer and Kellet 1991; Jank and Shmueli 2006). Thus, it is not applicable to this non-stationary and dynamic situation. Furthermore, our dependent variable is the opening week sales for the new products, while the availability of the independent variables covers the entire dynamic process of customers' knowledge collaborative activities. It leads to a small 'n' but large 'p' set-up with our final dataset and eventually makes regression-based econometric approach infeasible (Ramsay and Silverman 2005; Yao et al. 2005).

To conquer these obstacles and improve the predictive accuracy of the pre-release NPS forecast, we configure a unique modeling architecture. 1) We apply the Functional Data Analysis (FDA) and functional principal component analysis (*f*PCA) to convert customers' dynamic knowledge collaborative process for each sampled game into a trajectory, and reduce its dimensionality into a few component scores; 2) We also employ a machine-learning-based Gated Recurrent Unit (GRU) technique to conduct sentiment analysis of customers' emotional involvement (measured by textual posts from an online gaming forum) with the new game. Finally, we use an ensemble random forest (RF) approach to formally conduct the pre-release NPS forecast by incorporating both the dynamic process of customers' knowledge collaboration and emotional involvement, along with other controls.

This study makes the following contributions to the literature: 1) This is the first study incorporating the dynamic evolution of customers' knowledge collaboration for pre-release NPS forecast; 2) Our proposed method is effective in enhancing prediction accuracy by leveraging collaborative knowledge sharing activities, (3) the method out-performs forecasts based on simple cumulative or averaged metrics of the pre-launch knowledge collaborative activities to predict new product sales, highlighting the importance of utilizing the dynamic information and (4) our approach can help accurately predict product sales well in advance of product launch (i.e., early forecasting).