Proceedings of the 2018 Pre-ICIS SIGDSA Symposium

Special Interest Group on Decision Support and Analytics (SIGDSA)

12-12-2018

Dynamic Aggregation of Consumer Ratings with Bayesian Non-Parametrics

Christof Naumzik
ETH Zurich, cnaumzik@ethz.ch

Stefan Feuerriegel
ETH Zurich, sfeuerriegel@ethz.ch

Markus Weinmann
University of Liechtenstein, markus.weinmann@uni.li

Follow this and additional works at: https://aisel.aisnet.org/sigdsa2018

Dynamic Aggregation of Consumer Ratings with Bayesian Non-Parametrics

Completed Research Paper

Christof Naumzik
ETH Zurich
Weinbergstr. 56/58
8092 Zurich, Switzerland
cnaumzik@ethz.ch

Stefan Feuerriegel
ETH Zurich
Weinbergstr. 56/58
8092 Zurich, Switzerland
sfeuerriegel@ethz.ch

Markus Weinmann
University of Liechtenstein
Fürst-Franz-Josef-Strasse
9490 Vaduz, Liechtenstein
Markus.weinmann@uni.li

Abstract

Customer reviews reflect the quality of products and services (Gao et al., 2015), thereby reducing the uncertainty associated with them (Hong and Pavlou, 2014) and helping customers make informed decisions (Muchnik et al., 2013). Online reviews have thus become a decisive factor in customer decision-making. According to a recent study, 85% of the customer base trusts online reviews as much as personal recommendations.[1] Of particular importance to customers are aggregated ratings scores (Dai et al., 2018). For instance, a study on behalf of TripAdvisor confirmed that the majority of customers rely upon aggregate scores (83 percent for choosing accommodations and 70 percent for restaurants).[2] Hence, it is of great importance to IS research to find effective strategies for rating aggregation.

The prevailing approach to rating aggregation, such as that used by Yelp, builds upon a simple mean (Dai et al., 2018). In using the unweighted average over all ratings, relevant information about timing, source, and popularity is lost, which are some of the clear disadvantages associated with this approach. These shortcomings become immediately evident when considering the course of ratings over time. For instance, let us imagine a restaurant with an extensive history of several hundred medium ratings that finally replaces its mediocre chef with a celebrity. To adjust to the new quality, it will require a massive number of top ratings. Short of this, the mean may be lifted only slightly and still not reflect the current quality of the celebrity chef. Consequently, a simple average of all ratings entails a misleading indication of quality, since it is likely to deviate from the reality that should be expected by customers at any given moment and does not, therefore, reflect the current truth. Given these drawbacks, we develop a better means of estimating the expected quality in the form of a dynamic mechanism for rating aggregation.

This work develops a dynamic rating aggregation based on Bayesian non-parametrics, which results in the following latent Gaussian process model (LGPM). Our LGPM models the latent dynamics behind a sequence of ratings in order to obtain an indicator of the currently perceived quality as a novel approach to dynamic rating aggregation. The LGPM is dynamic, so it can adapt to drifts in the rating sequence or even structural changes (e. g. firing a chef). Our model takes additional characteristics of ratings into account to further improve the aggregated rating: On the one hand, it models the actual duration between individual ratings and can thus incorporate this information into the aggregation process. For instance, the longer the time interval between ratings, the higher the propensity to downgrade earlier ratings by placing additional weight on the new rating. On the other hand, rating sequences are highly variable and potentially noisy. This is why we must assume a stochastic relationship between ratings and quality, a compelling reason to

---

[1] See https://www.brightlocal.com/learn/local-consumer-review-survey/, last accessed September 6, 2018

[2] TripAdvisor: https://www.tripadvisor.com/TripAdvisorInsights/w810, last accessed September 6, 2018

include latent dynamics. At the same time, the use of latent dynamics represents a means of smoothing the noisy part of a rating sequence in a highly effective manner.

We are aware of earlier works that develop dynamic aggregation mechanisms. The paper by Dai et al. (2018) suggests a weighted average, which leverages the complete rating history but puts more weight on recent ratings. Ivanova and Scholz (2017) utilize a moving average, yet their approach primarily targets a different objective, namely identifying fabricated ratings. Both approaches can theoretically adapt to structural changes or trends in the rating history; however, both approaches function as heuristics and thus fall short of recovering the perceived quality. In practice, this leads to situations where the two rating aggregations are either largely insensitive to changes in the valence or overly sensitive to noise. Both approaches further neglect various properties of consumer reviews that could yield valuable information, such as the length of time between any two reviews.

We evaluate our rating aggregation based on 28,309 restaurant reviews from Yelp. Our empirical evidence confirms the superiority of our latent Gaussian process model over alternative dynamic rating aggregations from the literature. To evaluate our model, we follow the procedure from Dai et al. (2018) and estimate the expected quality from the past rating history by considering different rating aggregations – that is, the overall mean, a weighted average, and a moving average. We then compare the deviation between each aggregation mechanism (i.e. the expected quality) against the quality that was actually perceived by a customer (i.e. her rating). Our latent Gaussian process decreases the mean absolute error of an overall mean (as used, e. g., on Yelp.com) by 6.8%, that of a weighted average by 6.6%, and that of a moving average by 6.5%. This demonstrates the significant accuracy gains of our proposed rating aggregation and points directly towards practical benefits.

Our findings have direct implications in terms of both theory and practice. From a theoretical perspective, we contribute to the growing literature stream concerning online ratings, especially the relatively new stream of rating aggregation. From a practical perspective, our results help to yield more accurate estimates about the expected quality of a product or service, thereby increasing customer satisfaction.

Keywords

user ratings; rating aggregation; Gaussian process; latent dynamics; e-commerce.

References

Dai, W., Jin, G., Lee, J., and Luca, M. 2018. "Aggregation of consumer ratings: An application to Yelp.com," Quantitative Marketing and Economics (16:3), pp. 289–339.

Gao, G., Greenwood, B. N., Ritu, A., and McCullough, J. 2015. "Vocal Minority and Silent Majority: How Do Online Ratings Reflect Population Perceptions of Quality?" MIS Quarterly (39:3), pp. 565–589.

Hong, Y., and Pavlou, P. A. 2014. "Product Fit Uncertainty in Online Markets: Nature, Effects, and Antecedents," Information Systems Research (25:2), pp. 328–344.

Ivanova, O., and Scholz, M. 2017. "How can online marketplaces reduce rating manipulation? A new approach on dynamic aggregation of online ratings," Decision Support Systems (104), pp. 64–78.

Muchnik, L., Aral, S., and Taylor, S. J. 2013. "Social influence bias: A randomized experiment," Science (341:6146), pp. 647–651.