

Beyond the Stars: Unmasking Rating Inflation in Atlanta's Restaurants

Team 003: Honglai Peng, Qinyang Song, Xiaoran Zhu, Duanduan Zhou, Xin He

Motivation & Objectives

- The phenomenon of "Rating Inflation," characterized by an increase in average ratings and a decrease in rating variance, which threatens the informativeness of ratings to users and negatively affects their actions.
- While foundational research has comprehensively studied the mechanisms and biases influencing online ratings, the granularity and depth of rating data itself play a pivotal role.
- Our primary objective of this study is to address the causes of Rating Inflation in online restaurant reviews, focusing on aspects like the lack of emphasis on review recency, fake reviews detection, etc.

Data

- Research Dataset.** Our study uses a refined subset of Google Local Data from academic resources, focusing on 1,327 Atlanta restaurants from Georgia's business industries, resulting in a dataset of 760,525 reviews, after filtering and eliminating duplicates and closed businesses.
- Training Dataset.** To address the lack of pre-trained models and datasets in fake reviews detection, our study has compiled a balanced training dataset with over 56,000 entries. This dataset is evenly divided, comprising more than 28,000 authentic reviews and over 28,000 deceptive reviews.

Approaches

- Time Decay Recency Adjustment.** Compared with average rating, our approach involves a novel recency adjustment technique, combining time-decay modeling and Bayesian averaging to emphasize recent reviews.
 - Time-Decay:** Reviews' influence diminishes over time, with a 365-day lifecycle and a 0.90 decay rate. Recent reviews (<90 days) receive a 1.2 boost.
 - Bayesian Average:** Adjusted ratings are computed using Bayesian averaging, favoring median over mean review counts.
- Dual-stage Fake Reviews Removal.** We used CNN-LSTM and BERT models, adept in natural language processing, for our dataset. Initially, BERT detected fake reviews, followed by CNN-LSTM for further verification, effectively isolating genuine reviews.
- Data Visualization & Interactive Map** We recalibrated the rating system, visualized spatial rating distribution, illustrated fake review impact using interactive maps and multiple charts.

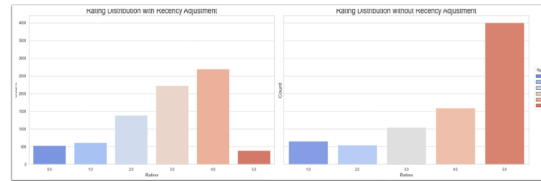


Fig.1 Bar Chart of Rating Distribution for The Varsity w/o Recency Adjustment

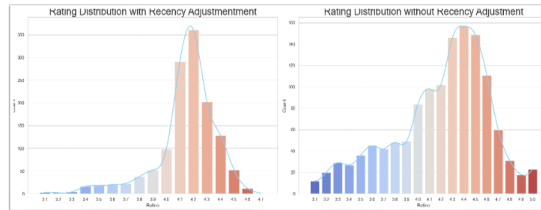


Fig.2 Rating Distribution of The Varsity w/o Dual Strategy

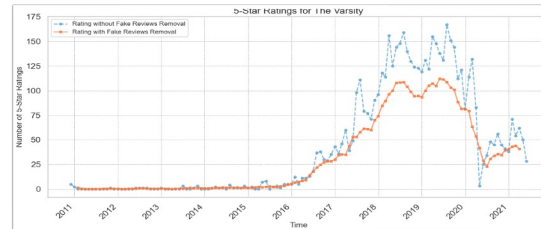


Fig.3 5-star Rating Trend for The Varsity w/o Fake Reviews Removal

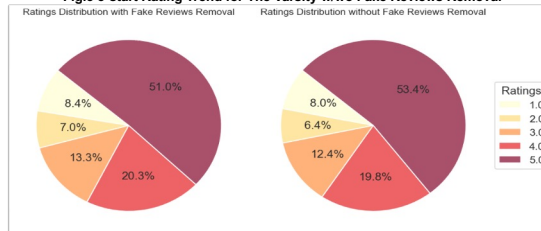


Fig.4 Pie Chart for Rating Distribution for The Varsity

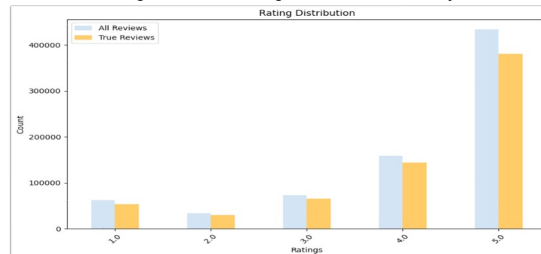


Fig.5 Bar Chart for Rating Distribution w/o Fake Reviews Removal

Experiments & Results Comparison

Implementing our recency adjustment and dual-stage fake reviews removal strategy in the dataset has effectively addressed rating inflation in Atlanta's restaurant scene. These methodologies were focused on enhancing the accuracy and reliability of customer feedback. The figures below evaluate the effectiveness of our approaches in mitigating rating inflation:

- Fig. 1** shows the impact of the Time Decay approach: 5-star ratings in the dataset decreased from approximately 4000 to 400 after recency adjustments. This reduction implies older reviews heavily influenced the original count. Additionally, a new category of 0 ratings was introduced for aged reviews, reflecting their diminishing relevance.
- Fig. 2** demonstrates the Bayesian Average method's effect on our rating analysis, resulting in a more Gaussian-like, balanced distribution compared to the previously skewed pattern, enhancing the rating system's realism and credibility.
- Fig. 3** shows the monthly trends of 5-star ratings for The Varsity. Before and after our review analysis, the data shifted from pronounced peaks (indicative of rating inflation) to a smoother pattern, reflecting a more genuine rating trend and improved reliability in the rating system.
- Fig. 4**, a pie chart case study of The Varsity, demonstrates our strategy's impact: 5-star ratings dropped from 53.4% to 51.0% after removing fake reviews, indicating reduced rating inflation.
- Fig. 5** reveals an 80,000 drop in 5-star ratings for Atlanta restaurants post fake review removal, highlighting the effectiveness of our approach against rating inflation.

Interactive Map

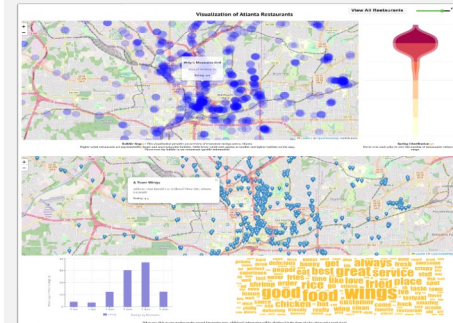


Fig.6 Interactive Map Web Page View

Our web platform, built on React D3, enables an interactive exploration of restaurant ratings. The top part of Fig.6 features a bubble map to filter restaurants by rating categories (4.0 to 5.0 shown here) and a funnel chart showcasing rating distribution. On the bottom part, users can also click on map markers to view a restaurant's rating details and a word cloud of review terms.

Conclusion & Discussion

- Our study enhanced the understanding of rating inflation in Atlanta's restaurant industry, using recency adjustments and fake review removal to recalibrate rating distributions for more accurate customer feedback representation.
- Our study quantified rating inflation's extent and our methods' effectiveness, evident in the significant decrease in older 5-star ratings and the introduction of a zero-rating category, better reflecting current consumer experiences.
- Our study Highlight the need for improved fake review detection models and the potential of integrating sentiment analysis for more nuanced feedback analysis.