Capstone Project Literature Review: Yelp Sentiment Analysis

Joshua Aflleje, Nicole Feil, Andrew Liu, and Jack Oebker

Arizona State University, Tempe, AZ 85281, USA

February 7, 2025

1 Introduction

Online reviews play a crucial role in shaping consumer behavior and business success. Platforms like Yelp, Google Reviews, and TripAdvisor significantly influence purchasing decisions. ReviewTrackers (2025) states that approximately 92.4% of customers rely on online reviews to guide most of their purchasing decisions, and around 63.6% of consumers read reviews before visiting a business. Among these platforms, Yelp is one of the most influential review platforms.

Yelp, a prominent online review platform, hosts millions of user-generated opinions covering restaurants, healthcare providers, and a wide range of local businesses. These reviews provide valuable insights for both consumers and business owners, shaping purchasing behavior and informing service improvements. However, while numerical star ratings offer a quick reference point, they often fail to capture the depth and nuances of consumer sentiment. Yelp's star rating system reduces potentially complex reviews into single numerical values. A five-star rating may indicate general satisfaction, but it does not differentiate which factors in particular contributed to it.

Recent advances in sentiment analysis have attempted to bridge this gap by moving beyond simple sentiment classification to more granular, aspect-based approaches. Researchers now aim to extract actionable insights by identifying specific features—such as food taste, service quality, and ambiance—that influence positive or negative feedback (Hu and Liu, 2004). This transition toward aspect-based sentiment analysis (ABSA) has enabled businesses and researchers to obtain a more detailed understanding of customer experiences.

As sentiment analysis techniques have grown more advanced, researchers have identified key challenges such as data biases and review authenticity. Studies such as Liang (2018) have demonstrated how NLP models can predict Yelp ratings based on review text, while Mukherjee et al. (2021) have explored the impact of fake reviews and how they distort sentiment analysis outcomes. Studies have shown that Yelp's filtering algorithm disproportionately flags extreme sentiment reviews—whether genuine or not—introducing bias into overall ratings Mukherjee et al. (2021).

1.1 Background

Sentiment analysis initially focused on classifying entire reviews as positive or negative, oversimplifying consumer feedback, particularly for products and services where different aspects could receive distinct evaluations.

However, Hu and Liu (2004) introduced aspect based sentiment analysis (ABSA), which identifies specific product or service attributes and evaluates their associated sentiment. Their research involved identifying opinion words (e.g., "delicious," "expensive," "friendly") and linking them to corresponding features (e.g., "food," "price," "service"). Aspect-based sentiment analysis has since been widely used to enable businesses to understand which aspects drive customer satisfaction or dissatisfaction and to allow them to target improvements.

Although Yelp is most commonly associated with restaurant reviews, its influence extends to non food-related businesses, including healthcare. Chen and Lee (2024) examined how Yelp ratings impact physician selection, demonstrating that sentiment-laden text in reviews significantly influences patient decision-making. Chen and Lee (2024) analyzed Yelp physician reviews and found that sentiment strongly influences patient choice. Their research showed that a one-star increase in rating resulted in a 1.9% revenue increase and a 1.2% rise in patient volume (Chen and Lee, 2024). This highlights the real-world impact of sentiment analysis beyond restaurants, making it applicable to professional services where interpersonal interactions shape consumer decisions. It also suggests that businesses in diverse industries can benefit from opinion mining to improve customer engagement and service quality.

Keller and Kostromitina (2020) explored how non-chain restaurants are rated on Yelp, differentiating them from larger franchise operations. Their study analyzed thousands of Yelp reviews to uncover key factors that influence star ratings such as independent restaurants being more likely to be judged on their uniqueness and atmosphere. Their findings emphasize the importance of contextualizing sentiment analysis models based on business type because factors influencing ratings can vary significantly between independent and chain restaurants.

Liang (2018) examined how machine learning models can predict Yelp ratings based on sentiment and topic models. The study employed natural language processing (NLP) techniques to analyze review text and infer numerical ratings, demonstrating the feasibility of automated review classification. By using topic modeling approaches such as Latent Dirichlet Allocation (LDA), Liang's study identified hidden themes in Yelp reviews and mapped them to sentiment scores, highlighting how computational methods can enhance traditional sentiment analysis.

2 Methodologies

2.1 Aspect-Based Sentiment Analysis (ABSA)

Aspect-based sentiment analysis (ABSA) provides a structured approach to extracting important features from Yelp reviews and assigning sentiment labels to those aspects (Hu and Liu, 2004). Deep learning techniques, such as BERT-based transformers, have significantly enhanced ABSA performance. The integration of multi-layered graph convolutional networks (GCNs) with BERT can allow for more effective sentiment dissection by capturing complex relationships between aspects within a review (Aziz et al., 2024). The approach combining aspect-specific embeddings with sentiment propagation across review structures leads to a nuanced understanding of user opinions (Aziz et al., 2024).

2.2 Sentiment-Driven Business Recommendations

Currently, Yelp suggests businesses based on categories and location, without factoring in sentiment-driven preferences. Clustering businesses based on user sentiment trends can offer a more personalized recommendation system (Keller and Kostromitina, 2020). This can be achieved using vector-based similarity models, which compare businesses based on review sentiment patterns rather than just business type. By leveraging sentiment-driven clustering techniques, businesses can be grouped based on their perceived strengths and weaknesses, providing users with more tailored recommendations.

2.3 Synthetic Review Generation

Generative models, such as GPT-based architectures, can create synthetic reviews at different star levels. These synthetic reviews, conditioned on real sentiment trends, allow users to better understand a business's strengths and weaknesses without manually reading thousands of reviews.

3 Conclusion

References

Aziz, K., Ji, D., Chakrabarti, P., Chakrabarti, T., Iqbal, M. S., and Abbasi, R. (2024). Unifying aspect-based sentiment analysis bert and multi-layered graph convolutional networks for comprehensive sentiment dissection. *Scientific Reports*, 14(1):14646–22.

Chen, Y. and Lee, S. (2024). User-generated physician ratings and their effects on patients' physician choices: Evidence from yelp. *Journal of Marketing*, 88(1):77–96.

Hu, M. and Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '04)*, pages 168–177, New York, NY, USA. ACM.

Keller, D. and Kostromitina, A. (2020). Characterizing non-chain restaurants' yelp star ratings: Generalizable findings from a representative sample of yelp reviews. *International Journal of Hospitality Management*, 86:102440.

Liang, Y. (2018). Yelp rating prediction with sentiment and topic models. *North Carolina Libraries*, 76(1):33–40.

Liu, S. (2020). Sentiment analysis of yelp reviews: A comparison of techniques and models. *arXiv preprint*, arXiv:2004.13851.

Mukherjee, A., Venkataraman, V., Liu, B., and Glance, N. (2021). What yelp fake review filter might be doing? In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 7, pages 409–418. AAAI Press.

Olson, A. W., Calderón-Figueroa, F., Bidian, O., Silver, D., and Sanner, S. (2021). Reading the city through its neighbourhoods: Deep text embeddings of yelp reviews as a basis for determining similarity and change. *Cities*, 110:103045.

Rahimi, S., Mottahedi, S., and Liu, X. (2018). The geography of taste: using yelp to study urban culture. *ISPRS International Journal of Geo-Information*, 7(9):376.

ReviewTrackers (2025). Customer reviews statistics and trends.