DERIVING COLLECTIVE RECOMMENDATION WITH ASPECT-BASED SENTIMENT AND SOCIAL INFLUENCE

Jyh-Hwa Liou
Department of Information Management
National Taipei University of Business
Taipei, 100, Taiwan
alioujh@ntub.edu.tw

Ssu-Yu Chen
Institute of Information Management
National Yang Ming Chiao Tung University
Hsinchu, 300, Taiwan
sally5272003@yahoo.com.tw

Yung-Ming Li[1]
Institute of Information Management
National Yang Ming Chiao Tung University
Hsinchu, 300, Taiwan
yml@mail.nctu.edu.tw

Guangming Cao
Digital Transformation Research Center
Ajman University,
UAE
g.cao@ajman.ac.ae

ABSTRACT

Despite the vast amount of restaurant information shared on social platforms, users often face difficulties identifying suitable options efficiently. Ratings are constrained by information narrowness, while textual reviews pose challenges due to information overload. Furthermore, restaurant recommendations based solely on ratings lack objectivity, as individual preferences differ. Users cannot judge whether a restaurant is worth visiting without credible information. However, few existing studies integrate semantic analysis with multidimensional orientation and social influence to make recommendations, leaving a gap in objective and comprehensive analysis. This research proposes a synthesis collective recommendation approach utilizing machine learning with aspect-based sentiment and social influence analyses. The proposed approach can appropriately adjust ratings as a basis for deciding the list of recommendations, considering location and preference factors. Experimental results show that the proposed mechanism significantly enhances users' ability to find restaurants that meet their needs, thereby improving business opportunities.

Keywords: Sentiment analysis; Social influence; Online review; Restaurant discovery; Social recommendation

1. Introduction

With changes in family structure and social lifestyles, consumer behavior has also begun to change. More and more consumers choose to eat out, and many dine in restaurants. This trend has significantly boosted the restaurant industry's annual sales to $799 billion in the U.S. in 2021, representing a nearly 20% increase from 2020 (National Restaurant Association, 2022). While this presents numerous opportunities for the catering industry, restaurant owners face the challenge of increasing visibility. On the other hand, the abundance of restaurants poses a persistent challenge for customers, who are often overwhelmed by the multitude of options available to them.

Due to the development of the Internet in recent years and the abundance of online information, people increasingly rely on online guides and social media sites to search for restaurants, particularly mobile phone users. The proportion of mobile phone users utilizing GPS to find restaurants is relatively high, as it allows them to determine the distance between themselves and nearby restaurants, thereby encouraging consumers to visit them (Yang & Wang, 2009). Social media features, including convenience, interactivity, personalization, anonymity, timelessness, and boundlessness, have created a new sales environment in the digital economy (Kotler, 2003). Electronic word of mouth (EWOM) has emerged as one of the most influential sources of information dissemination and significantly impacts customer decisions, especially regarding experiential items such as restaurants (Godes & Mayzlin, 2004). Numerous websites provide information on restaurants, including Yelp, Citysearch, Yahoo Local, Google Maps, TripAdvisor, and Zomato, where general users contribute. These platforms encourage users to enhance the quality of their reviews to assist others in making informed decisions when selecting restaurants (Blanding, 2011).

---

[1] Corresponding author.
Cite: Liou, J.H., Chen, S.U., Li, Y.M., & Cao, G., (2024, Nov.) Deriving Collective Recommendation with Aspect-Based Sentiment and Social Influence, *Journal of Electronic Commerce Research, 25(4).*

Previous studies have highlighted the inherent drawback of personal general ratings, which often fail to assist consumers in making accurate choices (Han, 2022; Jong, 2011; Liu et al., 2023; Sundar, 2008). Public rating is a quick reference when choosing a restaurant (Sundar, 2008). However, the rating has limited information and may not be aligned with the reviews. For instance, platforms like Yelp may feature numerous reviews, but many star ratings may not accurately reflect the restaurant's quality (Jong, 2011). This is because a single-dimensional rating system cannot comprehensively convey information about various aspects of the dining experience (Han, 2022; Liu et al., 2023). Some reviewers may offer an overall rating for the restaurant, while others may focus on specific details. For example, one person might give a five-star rating based on the quality of the waiter, food, and decor, while another might solely consider the food. A multidimensional rating system can enable consumers to make more informed decisions by evaluating restaurants across different aspects, such as food quality, price, service, and ambiance.

Reviewers may exhibit subjectivity when assigning extreme ratings based on personal emotions (Mudambi & Schuff, 2010). The negative impact of this subjectivity is that each person has a different definition of a star rating, which likely results in reduced rating credibility. On the other hand, incorporating more attributes can enhance objectivity in evaluating product quality (Mudambi & Schuff, 2010; Mudambi et al., 2014). Personal general ratings, based solely on the average of numerical scores provided by reviewers, tend to be more subjective. In contrast, aspect-based ratings convert original textual reviews into aspect-based numerical scores through multi-criteria evaluation and objective analysis.

With the characteristics of information narrowness in consumer ratings and information overload in text reviews, many consumers face a choice problem. Existing studies address information overload by investigating how ratings and text reviews influence consumers' purchasing decisions (Büschken & Allenby, 2016; Chakraborty et al., 2022; Jabr & Rahman, 2022; Lei et al., 2022). Most researchers employ text-based analytics and sentiment analysis to analyze consumer textual reviews (Büschken & Allenby, 2016; Chakraborty et al., 2022; Jabr & Rahman, 2022). Some authors have also used empirical analysis to dissect the effects of ratings and reviews (Lei et al., 2022). However, most of these studies focus on using ratings and textual reviews to analyze a company's service rather than conducting a very objective review of the reviews. This research proposes a classification method that utilizes aspect-based sentiment analysis to provide more accurate ratings.

The quality of a restaurant should be assessed based on diverse, detailed, and credible appraisals that are meaningful to the diner. Evaluating restaurant credibility solely based on the traditional star rating system requires further in-depth analysis. Additionally, individuals often place trust in people who have influence. When assessing restaurant quality, social suggestions from individuals with higher social influence are deemed more trustworthy (O'Donovan & Smyth, 2005). Scholars also emphasize the importance of evaluating reviews based on their content and the expertise of the reviewers (Li et al., 2021). Factors affecting the social influence of reviewers should be considered to enhance the credibility of calculating restaurant ratings. However, existing studies primarily address the information filtering issue through semantic analysis (Büschken & Allenby, 2016; Chakraborty et al., 2022; Jabr & Rahman, 2022; Lei et al., 2022) without comprehensively examining multicriteria appraisal, review credibility, and preference alignment to generate a collective recommendation.

To address the literature gap and advance practice, this research proposes a restaurant recommendation system utilizing a machine learning approach incorporating aspect-based sentiment and social influence analyses. The proposed mechanism aims to provide more objective and credible information by considering both objective and subjective ratings and reviews and consumers' locations and preferences. Specifically, we propose a recommendation framework following the principles for designing recommendation artifacts (Herlocker et al., 2004; Xiao & Benbasat, 2007). The framework is divided into two stages: target appraisal and target recommendation, each playing a critical role in the restaurant recommendation mechanism. In the target appraisal stage, we evaluate the quality of the target by incorporating diverse aspects of rating generation from reviews (multicriteria appraisal) and review credibility (including source credibility and content credibility). Some researchers have explored aspect-based sentiment analysis for diverse aspect reviews (María et al., 2019; Zhu et al., 2022). However, few studies utilize diverse aspect ratings generated from reviewer text reviews. Regarding review credibility, some authors consider review objectivity but overlook the reviewer's influence (Asani et al., 2021). Conversely, others consider the reviewer's influence but neglect review objectivity (Divyaa & Pervin, 2019). In the target recommendation stage, we evaluate the fitness of the target (restaurants) for the user by combining multidimensional aspect ratings with preference alignment, which involves both spatial proximity and preference similarity (Asani et al., 2021; Divyaa & Pervin, 2019). Many studies incorporate individual preferences into recommendation systems (Asani et al., 2021; Divyaa & Pervin, 2019). However, they often fail to consider comprehensive factors to ensure that recommendations are trustworthy and objective for users.

To enhance the effectiveness and usefulness of restaurant recommendations, we will address the following issues:

(1) How can the credibility of restaurant ratings be improved by considering diverse aspects of online reviews?

The restaurant rating system on the platform may not always provide informative insights to help users make decisions. For instance, a restaurant may have a five-star rating, but users may not know which aspects of the experience contributed to this rating. Additionally, individual interpretations of star ratings can vary, making them less reliable overall quality indicators. We believe that evaluating restaurants across a diverse range of aspects, based on the analysis of user reviews, would offer a more comprehensive measure of their quality. This approach would enable diners to find restaurants that align with their preferences.

(2) How can the influence of public reviews shared through a community platform be best evaluated?

We know that expert information is generally more credible than that from strangers. This suggests that when searching for a restaurant, we assess the reviewer's credibility based on their expertise and influence. An influential review carries more weight in users' analysis of the restaurant's overall aspects score than a standard review (Zhou & Guo, 2017). Consequently, an influential review should not be given equal weight to a common one in determining the overall rating score. Therefore, we should consider analyzing the reviewer's personal information, such as their platform usage duration and the number of previous posts they have written.

(3) How can diners find nearby restaurants matching their preferences through review analysis?

We collect diners' data, including their previous reviews, and employ aspect-based sentiment analysis to understand their preferences, such as the features they prioritize and the locations they frequent. Additionally, individuals may be influenced by others when making decisions. Reviews and ratings from highly influential experts have a more significant impact on the decision-making process. Once we have obtained the user's preferences and aspect-based ratings, we can rank restaurant candidates that align with individual needs and provide personalized recommendations of restaurants the diners may enjoy.

Nowadays, information is often either overloaded or insufficient, making it difficult to judge the quality of a restaurant based solely on review content. As people prefer simple and useful tips, we propose recommendations based on detailed rating levels that consider various aspects of a restaurant. This approach utilizes sentiment analysis to extract keywords and analyze positive and negative comments within the community (Kim et al., 2016), along with the commenter's social influence analysis. To date, few works have combined the analysis of semantic meaning based on multidimensional orientation and social influence to make recommendations. By compensating for the ratings of reviews that lack objectivity and improving the accuracy of the recommended list, our recommendation system not only helps users save time and effort in searching for restaurants but also assists businesses in reaching valuable consumers.

The remainder of the paper is organized as follows: Section 2 reviews the literature related to the research. In Section 3, we present the system framework of the proposed recommendation mechanism. Section 4 describes the experiments' processes for verifying the proposed mechanism. In Section 5, we present and discuss the evaluation and results of the experiments. Finally, Section 6 summarizes our research contributions and discusses the limitations and future works.

2. Related Works

2.1. Restaurant Rating Problem

With the recent trend of user-generated content, people are increasingly accustomed to sharing their knowledge and experiences. Various factors, such as ratings, reviews, and location, are commonly considered when choosing a restaurant. According to statistics, reviews with star ratings between 4 and 5 stars contain 71% positive and only 6% negative sentences. Conversely, reviews with ratings between 1 and 2 stars have only 5% affirmative sentences and 78% negative ones (Ganu et al., 2009). Additionally, researchers have been addressing the narrowness problem of rating information. For instance, Dai et al. (2018) investigated rating aggregation and proposed an adjusted average rating from Yelp for ground truth quality assessment. Their findings indicated that users often prioritize the average rating. However, users may have differing sentiments when assigning ratings. For example, consider a restaurant with an average rating of 3.5 stars. While patrons may praise the food or atmosphere, they might express dissatisfaction with the pricing, as illustrated by the following excerpt from a review:

- It was our first time there, and overall, it was good. Prices are a bit high, but the food was good. We would consider returning when we are back in that area.
- The atmosphere is nice. We might consider giving more stars if they lowered prices by 10–15%.

If someone is concerned about the price, then he/she may choose not to dine at this restaurant, even if the food is delicious. Beyond the rating, reviews should be thoroughly analyzed and interpreted to assist users in identifying key features of a product (Ganu et al., 2009). For instance, Zagat, a renowned review website in the U.S., employs four categories of rating: food, service, price, and ambiance, which are widely recognized (Tan et al., 2013). These four aspects have recently been utilized to analyze review sentiment (Peng et al., 2022; Nakayama & Wan, 2019).

In this research, we design an artifact collective recommendation mechanism to resolve the rating problem and select suitable restaurants. We derive multidimensional aspect ratings through an aspect-based sentiment analysis approach and utilize the social influence theory to evaluate the reviewers' trustworthiness and the reviews' credibility. The proposed recommendation artifact aims to align the diner's preferences with the aspect quality of the recommended restaurant.

2.2. Aspect-Based Sentiment Analysis

Based on the theory of information transfer (or the information system) (Belkin, 1984), there are three key components: the user, the knowledge resource, and the intermediary mechanism. The rating system on platforms like Yelp can be regarded as an information system, with customers' ratings and reviews serving as knowledge resources. The users believe they can utilize the platform to research reviews or ratings and make informed decisions, while the intermediary mechanism facilitates the interaction between the user's requirements and the knowledge resource (Chen et al., 2018; Liu et al., 2023). The more effective the information, the better consumers' purchasing decisions become. According to the information transfer theory, it is posited that rating diversity enhances the usefulness of reviews within the multidimensional rating system (Liu et al., 2023). Scholars indicate that ratings in a single-dimensional rating system show a downward trend and greater dispersion, while ratings in a multidimensional rating system are significantly higher and more consistent (Chen et al., 2018; Yang et al., 2023). These findings imply that the multidimensional rating system aids consumers in finding products that better match their preferences and boosts their confidence in their choices (Chen et al., 2018). Scholars have also advocated for large-scale ratings to mitigate the potential bias stemming from subjectivity (Yang et al., 2023).

The term "opinion mining" first appeared in Dave et al. (2003) and utilizes feature extraction to classify whether an electronic product's evaluation is positive or negative. With the abundance of text reviews, many consumers struggle to make choices. Most researchers employ sentiment analysis to process large volumes of textual reviews (Chakraborty et al., 2022; Jabr & Rahman, 2022; Lopes et al., 2022; Wan et al., 2023). Sentiment analysis, also known as aspect-based sentiment analysis, was first proposed by Thet et al. (2010). Scholars conduct a comparative review of aspect-based sentiment analysis to provide context for different approaches (Do et al., 2019). To further improve the effectiveness of analysis, SemEval (International Workshop on Semantic Evaluation) held a competition to find better ways to improve accuracy (Pontiki et al., 2016). Through collation, there are currently four main methods of analysis: frequency of words, syntactic relations, supervised learning, and - the best one - grouping extracted aspect words together and lexicographical similarities, synonym relationships, and distances based on taxonomy (Carenini et al., 2005).

As most star ratings do not necessarily represent the most objective or accurate assessment, Jong (2011) utilized traditional sentiment analysis to rectify this issue. Star ratings quickly indicate overall quality, allowing users to discern between positive and negative experiences swiftly. However, completely replacing star ratings may not be necessary by incorporating specific aspects and objective expressions. Wen et al. (2021) measured food services using objective measures from user-generated ratings at the restaurant level and subjective measures from individual customer evaluations of restaurant dining. In this research, we propose a classification method utilizing sentiment analysis to provide more accurate multidimensional ratings across different aspects, thereby balancing objective and subjective measurements of reviews.

2.3. Social Influence

Deutsch and Gerard (1955) distinguished social influence into normative and informational categories. They defined normative social influence as "an influence to conform with the positive expectations of another" and informational social influence as "an influence to accept information obtained from another as evidence about reality." Li et al. (2021) analyzed the reviewer's popularity and expertise based on social influence theories to evaluate the usefulness of a review. They applied social influence theory, considering that reviewer expertise represents informational influence, and reviewer popularity represents normative influence. Reviewer expertise was defined as the number of previous reviews written by a reviewer with rich experience and professional knowledge. In contrast, reviewer popularity was defined as the number of fans of a reviewer. These variables of reviewer expertise and reviewer popularity were utilized in their hypotheses. We applied social influence theory and the concept of reviewer trustworthiness (Banerjee et al., 2017; Deutsch & Gerard, 1955; Li et al., 2021), which integrates the reviewer's expertise and influence across both informational and normative aspects.

In the past, without the availability of online reviews and ratings, recommendations for restaurants primarily relied on word-of-mouth from friends and relatives. However, in modern times, people frequently turn to online review sites for guidance in decision-making. Researchers leverage user-generated content from online forums and employ social network and sentiment analysis methods to make relevant predictions (Colladon et al., 2019; Lu et al., 2020). The quality of customer interactions also significantly influences the overall experience (Wen et al., 2021). Scholars further assess the usefulness of reviews by examining both the review content and the expertise of the reviewers (Li et al., 2021). Moreover, increased search costs may arise without trustworthy information as consumers seek information about products and services. Advice from a credible source holds greater value, as trustworthy data has been shown to enhance the accuracy of recommendations (O'Donovan & Smyth, 2005), consequently reducing additional search costs.

Opinions from individuals with social influence are generally deemed more reliable. Comparatively, recommendations leveraging online social media to understand target users and incorporate suggestions from influential individuals tend to yield more suitable outcomes than traditional recommendations. A prior study suggested that when the number of reviews surpasses 50, there is an increase in the overall perception and rating of the restaurant, with ratings being particularly susceptible to the opinions expressed by "elite" reviewers (Hajas et al., 2014). Evaluating the usefulness of a review, Li et al. (2021) analyze the reviewer's popularity and expertise in line with dual-process and social influence theories. They define expertise as a reviewer's extensive experience, professional knowledge, and popularity as the reviewer's social network reach. This underscores the importance of businesses garnering numerous reviews from individuals with elite status, as these individuals wield nearly double the influence of other commentators. These elite reviewers, as designated by platforms like Yelp, are recognized for consistently posting valuable reviews (Luca, 2011). In this research, the social influence factor of reviewers will be integrated into the recommendation system to enhance the credibility of restaurant rating calculations.

2.4. Recommendation Systems

The recommendation system has been commonly applied to various applications, particularly playing an increasingly important role in e-commerce. Users find the restaurant more relevant to their preferences and tend to show a positive attitude toward those restaurants with a high page ranking. For example, Yelp, the biggest restaurant guild website, did not provide enough information for users to independently judge restaurants' environment, service, food, and other aspects (Yu et al., 2017). Users may have different feelings and tastes, so we need to consider side information. With abundant side information scenes, it becomes hard to work with traditional methods based on matrix decomposition, and feature-based engineering based on human design is very laborious but is required (Zagheli et al., 2017).

Recently, some recommendation systems that use aspect-based sentiment analysis have been proposed. However, most only achieve the aspect extraction of restaurant recommendations (Pronoza et al., 2016). The recommendation system based on aspect-based sentiment analysis has proven more effective than the traditional recommendation system (Chang et al., 2022). Recent studies consider different features to filter the recommended list, such as the calculated polarity and subjectivity of the review (Suresh et al., 2014). Besides, using the term frequency is not optimal, and the time factor should be considered because people's preferences change over time, and the business may also change (Chen et al., 2014).

In this paper, we propose a collective recommendation mechanism for restaurant selection decision-making support, which comprehensively considers multicriteria appraisal (aspect-based sentiment and aspect-based rating), source credibility (reviewer's expertise and reviewer's influence), content credibility (review's objectiveness), and preference alignment (spatial proximity and preference similarity). The comparisons of our study with existing related works are outlined in Table 1.

Table 1: Research Comparisons

Study	Multicriteria Appraisal		Review Credibility		Preference Alignment	
	Aspect-Based Rating	Aspect-Based Sentiment	Content Credibility: Review's Objectiveness	Source Credibility: Reviewer's Expertise & Influence	Spatial Proximity	Preference Similarity
Our study	✓	✓	✓	✓	✓	✓
Zhu et al. (2022)		✓				✓
Asani et al. (2021)		✓	✓		✓	✓
Luo et al. (2020)	✓	✓			✓	
Divyaa and Pervin (2019)				✓	✓	✓
María et al. (2019)		✓				✓
Zhang et al. (2018)	✓					✓

3. The System Framework

For conceptual, the study aims to propose a collective recommendation mechanism for restaurant selection decision-making support, utilizing collective intelligence derived from effective extraction and evaluation aggregation of opinions from the crowd. The proposed mechanism aims to enhance the quality evaluation of target restaurants by incorporating multicriteria appraisal, which includes aspect-based sentiment analysis and aspect-based rating (Luo et al., 2020). Additionally, source credibility, encompassing the reviewer's expertise and influence (Divyaa & Pervin, 2019; Li et al., 2021), has been integrated into the framework. Furthermore, content credibility has been considered, focusing on the objectiveness of reviews (Asani et al., 2021). After quality evaluation, the effective recommendation is enhanced by considering preference alignment, including spatial proximity and preference similarity (Asani et al., 2021; Divyaa & Pervin, 2019). The conceptual framework of our proposed mechanism (see Figure 1) encompasses these elements to provide a holistic understanding of the collective recommendation processes.

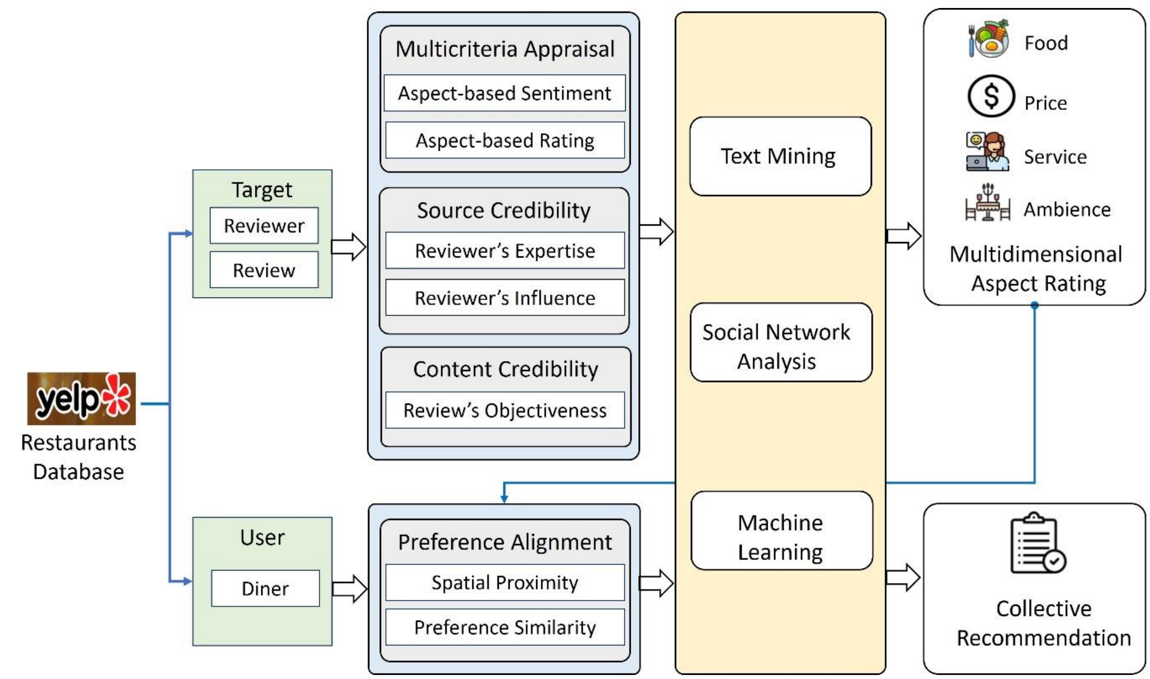

Figure 1: The Conceptual Framework of Collective Recommendation Mechanism

Specifically, the proposed collective recommendation approach includes two stages of the process:

(1) Target Appraisal: The appraisal artifact consolidates crowd opinions by incorporating multicriteria appraisal, source credibility, and content credibility to evaluate the quality of the target. This process begins with collecting users' posts from Yelp's review website. Subsequently, multicriteria appraisal is an analytical method implemented through aspect-based sentiment analysis, which is further transformed into aspect-based ratings. Aspect-based sentiment analysis examines user text reviews, focusing on various aspects of the dining experience. An aspect-based rating is a computed numerical value assigned to different aspects based on pre-processed aspect-based sentiment text analysis. Source credibility is determined by considering the reviewer's expertise and influence through social network analysis (including the reviewer's expertise calculation and the review's influence calculation). Concurrently, content credibility is evaluated by assessing the review's objectiveness using aspect-based sentiment analysis. Finally, the multidimensional aspect rating refers to the numerical scores of various aspects of a restaurant obtained through the analysis of multicriteria appraisal, source credibility, and content credibility.

(2) Target Recommendation: The recommendation artifact aligns the diner's preferences by analyzing spatial proximity (feasible location) and preference similarity (personal preference) using aspect-based sentiment analysis and recommendation methods. The collective recommendation involves generating a ranked list of the most suitable restaurants for the user by combining multidimensional aspect ratings with preference alignment.

The system framework of the proposed mechanism is illustrated in Figure 2.

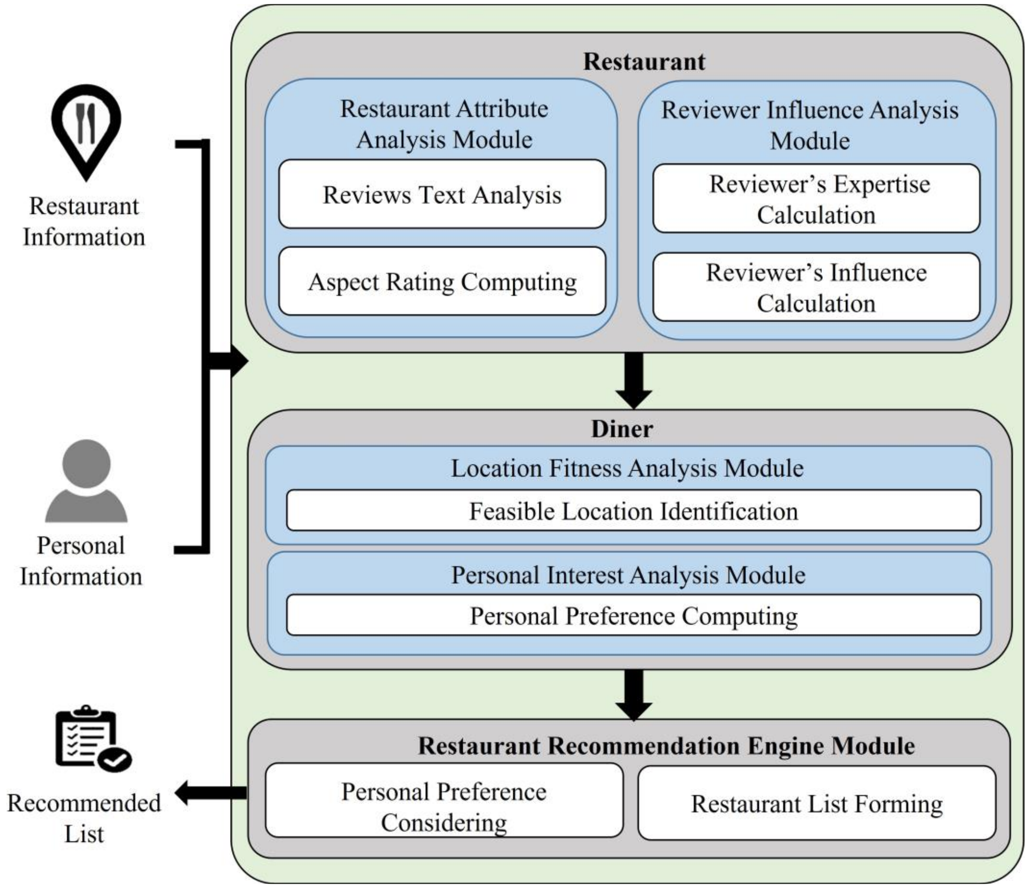

Figure 2: The System Framework

There are five main modules in our system, described as follows:

(1) Restaurant Attribute Analysis Module: This module utilizes aspect-based sentiment analysis to infer the attributes of a restaurant. Restaurants are classified under different types of food to conduct a multicriteria appraisal of aspect-based ratings.

(2) Reviewer Influence Analysis Module: This module assesses the expertise and influence of reviewers, identifying those deemed trustworthy to ensure source credibility. Reviewers with higher trustworthiness exert greater influence, and their reviews are more likely to impact other users significantly.

(3) Location Fitness Analysis Module: Here, we obtain the user's location and locate nearby restaurants based on the user's input or detected location to establish spatial proximity.

(4) Personal Interest Analysis Module: This module calculates a diner's food preferences based on past post contents and check-in places. It classifies this data into different types of food that users may enjoy and identifies similarities in preferences. Spatial proximity and preference similarity ensure preference alignment in the collective recommendation.

(5) Restaurant Recommendation Engine Module: This module integrates the results from the above analysis modules, assigning appropriate criteria weights to calculate suitability scores for the restaurants and generate a recommended list of restaurants for the user.

3.1. Restaurant Attribute Analysis Module

3.1.1 Reviews Text Analysis

The extraction process has been extensively studied in many works. In this research, our approach to feature extraction involves obtaining tokens and performing part-of-speech (POS) tagging on a sentence level. After analyzing syntactic dependency, direct/short dependency, and preprocessing all text files using three different tools—namely, the IXA-pipeline (Agerri et al., 2014) and the Stanford CoreNLP tool (Manning et al., 2014)—, we further pass the named entity data to the Stanford CoreNLP tool to extract additional information, such as lemmatization, POS, and dependency parsing. The output of the dependency parsing step provides insights into the relationship between words, allowing us to determine the intended reference of adjectives. In this review, numerous aspect terms are identified.

Subsequently, we use KNN to cluster similar aspect terms (e.g., "price" and "cost"). We have defined four main aspects: food, price, service, and ambiance. Any aspect term that falls under a specific aspect is categorized accordingly. Finally, we estimate the average sentiment per aspect term using the TextBlob tool.

3.1.2 Aspect Rating Computing

In text analysis, aspects such as food, price, service, and ambiance are dimensions for analyzing a restaurant. Utilizing the TextBlob tool, sentiment and subjectivity scores ranging from 0.0 to 1.0 can be derived. A sentiment score greater than zero indicates a positive aspect, while a score less than zero indicates a negative aspect. A score of zero signifies a neutral position and is not factored into the calculation. Each aspect rating of a restaurant r_j is determined by analyzing reviews associated with the restaurant $\forall p_i \in \Phi_j$. *RntAspectRating* (r_j, A_k) represents the cumulative score of the eigenvalue for each aspect A_k of restaurant r_j, $k \in \{1, \ldots 4\}$, and is formulated as follows:

$$RntAspectRating(r_j, A_k) = \frac{\sum_{p_i \in \Phi_j} fp_i\,(r_j, A_k)}{\sum_{p_i \in \Phi_j} fp_i(r_j, A_k) + \sum_{i=1}^{j} fn_i(r_j, A_k)} - \frac{\sum_{p_i \in \Phi_j} fn_i\,(r_j, A_k)}{\sum_{p_i \in \Phi_j} fp_i(r_j, A_k) + \sum_{i=1}^{j} fn_i(r_j, A_k)} \quad (1)$$

where $f_{p_i}(r_j, A_k)$ represents the weighted number of characteristic words related to aspect A_k of restaurant r_j that are positive for review post *i*, and $f_{n_i}(r_j, A_k)$ represents the weighted number of characteristic words related to aspect A_k of restaurant r_j that are negative. The weighted number of characteristic words is the original number of characteristic words multiplied by the influence of the reviewer u_i, denoted as $Influence(u_i)$ is the reviewer who posted the review p_i and $Influence(u_i)$ is the reviewer $u_i's$ influence measured in subsection 3.2. Since $RntAspectRating(r_j, A_k)$ will have a score range of [-1, +1], and Yelp's rating is presented on a five-point scale (0 points to 5 points), we convert the score of $RantAspectRating(A_k)$ to [0, 5]. If a feature aspect match is not found in an aspect, the original score of 0 points is adjusted to 3 points after conversion. The subjectivity score of 0.0 indicates a very objective opinion, while 1.0 indicates a very subjective opinion.

3.2. Reviewer Influence Analysis Module

3.2.1 Reviewer Expertise Calculation

In determining who has influence, relevant research has established a model with an accuracy of 83% in determining whether a reviewer is trustworthy (Banerjee et al., 2017). Reviewer expertise for a restaurant is often indicated by review platforms highlighting esteemed reviewers, such as Yelp Elite (Banerjee et al., 2017; Nguyen et al., 2021). Scholars have proposed a theoretical model where a reviewer's positivity, involvement, experience, reputation, competence, and sociability influence the reviewer's trustworthiness (McCroskey & Jenson, 1975; Banerjee et al., 2017). The reviewer's expertise for each reviewer u_i is measured by Eq. (2), with the elements defined as shown in Table 2 (Banerjee et al., 2017).

$$Reviewer\ expertise\ (u_i) = \beta_0 Positivity(u_i) + \beta_1 Involvement(u_i) + \beta_2 Experience(u_i) + \beta_3 Reputation(u_i) + \beta_4 Competence(u_i) \quad (2)$$

Table 2: Interpretation of Variables (Banerjee et al., 2017)

Constructs	Measuring variables
Positivity	Average review rating
Involvement	Number of reviews written
Experience	Number of years on Yelp
Reputation	Number of years as an 'Elite' reviewer
Competence	The average number of review helpfulness votes received per review
Sociability	Number of friends

The number of elite years in the above formula will be used as one of the variables in the calculation. "Elite" refers to an exclusive expert on the Yelp system: when a user has enough friends, high votes, and high compliments, the Yelp system classifies them as "elite." The reviewer information is displayed in Figure 3 on Yelp.

The trustworthiness degree for each reviewer u_i is measured by Eq. (3).

$$Trustworthiness\ (u_i) = Reviewer\ expertise\ (u_i) + \beta_5 Sociability(u_i) \tag{3}$$

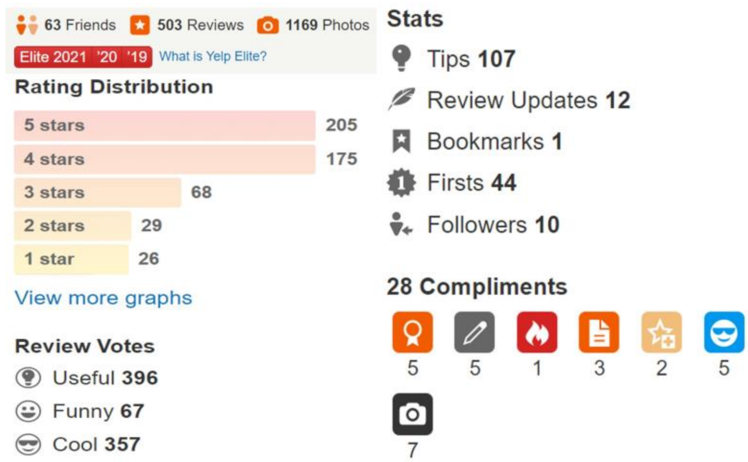

Figure 3: The Information of a Reviewer

3.2.2 Reviewer's Influence Calculation

In this research, we collect all the posts on the restaurant platform and utilize sentiment analysis to evaluate the sentiment polarity of each post. To enhance the accuracy of our mechanism, we employ reviewer influence analysis to assign weight to each post. Eq. (4) is utilized to calculate the weight of each post. If the trustworthiness score exceeds the average value, the post is labeled as having high expertise of the reviewer, whereas if the trustworthiness score is below average, the post is labeled as having low trustworthiness. This implies that a post from a trustworthy reviewer carries a higher weight, approximately double the impact of a general message (Luca, 2011). Individuals are categorized as highly influential when their trustworthiness score surpasses a threshold δ. The formula for determining the influence of each review is as follows:

$$Influence(u_i) = \begin{cases} H, & high\ trustworthiness, \quad trustworthiness\ (u_i) \geq average\ \ trustworthiness \\ L, & low\ trustworthiness, \quad trustworthiness\ (u_i) < average\ \ trustworthiness \end{cases} \tag{4}$$

In this research, we set H=2 and L=1. H is an adjustable parameter representing the importance of influence. We specifically chose H=2 to represent the weight of posts from highly trustworthy reviewers. This value was selected based on the findings of Luca (2011). This parameter can be modified in different contexts to reflect varying degrees of influence.

3.3. Location Fitness Analysis Module

The distance from a user to a restaurant significantly influences the user's decision-making process. With the prevalence of GPS technology, many restaurant-related websites employ location-based services (LBS) to bridge the gap between online information and offline experiences. The widespread adoption of mobile devices, corresponding wireless Internet connectivity, and location-based social network services (LBSNs) has made such services integral to users' dining experiences (Luarn et al., 2015). In this research, we recognize that restaurant location is a crucial factor in users' decision-making processes. Therefore, we analyze users' location data from past posts to enhance our recommendation system.

Feasible Location Identification: Location is pivotal in guiding users to restaurants. Here, we delve into two aspects of location factors influencing a user's willingness to visit a restaurant: location filtering and current location.

The convenience of accessing the city influences a user's willingness to dine out. Therefore, we analyze geographic data extracted from a user's past posts. These posts contain information about restaurant locations, providing details such as city names and coordinates. From this data, we can determine the user's customary area, representing places they frequent. Subsequently, for user u_i and restaurant r_j, we filter restaurants based on their customary area using the following formula:

$$F(u_i, r_j) = \begin{cases} 1, & u_i\text{'s customary area} = r_j\text{'s area} \\ 0, & u_i\text{'s customary area} \neq r_j\text{'s area} \end{cases} \tag{5}$$

Current Location: After filtering users based on their customary area information, we identify restaurants in the same area. We then calculate the distance between the user's location and each restaurant. When a restaurant is closer to the user's current location, their intention to visit that restaurant is likely higher. The user's current location, such as longitude and latitude, is acquired through GPS or mobile location technology. However, if a user does not use GPS, we calculate the latitude and longitude averages of restaurants visited in the past. Afterward, we compute the great-circle distance with the current location value. Let $(\emptyset_u, \lambda_u)$ and($\emptyset_r, \lambda_r)$ denote the geographical latitude and longitude in radians of two points, with u representing the starting point and r representing the destination point. Denoting $\cos \emptyset_r \cos(\Delta\lambda)$ as their absolute differences; then $\Delta\hat{\sigma}$, the central angle between two points, can be given by spherical cosine law:

$$\Delta\hat{\sigma} = \arccos\left(\sin \emptyset_u \sin \emptyset_r + \cos \emptyset_u \cos \emptyset_r \cos(\Delta\lambda)\right) \tag{6}$$

The great-circle distance *d* between these two points can be obtained using the formula for arc length.

$$D(u_i, r_j) = r\Delta\hat{\sigma} \tag{7}$$

The $D(u_i, r_j)$ represents the Euclidean distance between u and r. If d approaches 0, we recognize that the distance between the user and the restaurant is closer. As a result, the location score can be calculated by the following formula:

$$L(u_i, r_j) = F(u_i, r_j) - D(u_i, r_j) \tag{8}$$

According to the result of the $L(u_i, r_j)$, we know the higher value indicates the restaurant is closer to the user. We then proceed to calculate the aspects of the restaurant and the aspects needed by the user.

3.4. Personal Interest Analysis Module

Personal Preference Computing: Each individual has different preferences for various aspects. Analyzing the content of a user's reviews enables us to predict their level of interest in each aspect. Sentiments conveyed in reviews often include both positive and negative expressions. However, neutral reviews are excluded from consideration as they provide limited insight into the restaurant's quality.

Sentiment Analysis: The interpretation of aspect-based sentiment analysis by users and businesses varies somewhat. For instance, in business contexts, a poor rating for food would typically lead to a reduction in the overall rating of the food. However, when a user expresses dissatisfaction with the food, it indicates that they prioritize this aspect. To measure the significance of each evaluated aspect, we count $P(A_k, u_i)$ and $N(A_k, u_i)$, representing the number of positive and negative opinions expressed by user u_i regarding each aspect A_k of previous restaurants they have shared. Here, $k \in \{1, ... 4\}$, indicates the four aspects of the restaurant*: food, price, service,* and *ambiance*.

Afterward, we obtain $PScore(A_k, u_i)$ and $NScore(A_k, u_i)$ though min-max normalized values of $P(A_k, u_i)$ and $N(A_k, u_i)$ (Marrese-Taylor et al., 2013).

$$\begin{gathered} PScore(A_k, u_i) = \frac{P(A_k, u_i) - Min\{P(A_k, u_i)\}}{Max\_P(A_k, u_i) - Min\{P(A_k, u_i)\}}, \\ NScore(A_k, u_j) = \frac{N(A_k, u_i) - Min\{N(A_k, u_i)\}}{Max\_N(A_i) - Min\{N(A_k, u_i)\}} \end{gathered} \tag{9}$$

To calculate the standard deviation of these scores, we use the formula:

$$PNScore(A_k, u_i) = \sqrt{\frac{1}{2}\left(\begin{gathered}\left(PScore(A_k, u_i) - \frac{PScore(A_k, u_i) + NScore(A_k, u_i)}{2}\right)^2 \\ + \left(NScore(A_k, u_i) - \frac{PScore(A_k, u_i) + NScore(A_k, u_i)}{2}\right)^2\end{gathered}\right)} \tag{10}$$

Aspect Preference: After calculating their standard deviations, we obtain a new measure for each aspect, referred to as the relative importance of each aspect, represented by the min-max normalized value of its $PNScore(A_k)$:

$$UserAspectRing(A_k, u_i) = \frac{PNScore(A_k, u_i) - Min\{PNScore(A_k, u_i)\}}{Max\{PNScore(A_k, u_i)\} - Min\{PNScore(A_k, u_i)\}} \tag{11}$$

$UserAspectRating(A_k, u_i)$ represents user u_i 's preference score of the restaurant aspect A_k . Since *UserAspectRating* (A_k, u_i) has a score value in the range of [0, 1] and Yelp's rating is presented on a five-point scale (0 points to 5 points), we will convert the score of *UserAspectRating* (A_k, u_i) to [0, 5].

3.5. Restaurant Recommendation Engine Module

3.5.1 Personal Preference Considering

In this section, we need to consider the personal preference weight of the users who have completed the calculation of $L(u_i, r_j)$ in Section 3.3. The aspect rating score of a restaurant candidate should exceed a user's needs. This implies that restaurant candidates are within their customary range and meet the user's minimum requirements. The $RntAspectRing(r_j, A_k)$ represents the aspect rating of a restaurant r_j based on public reviews, such as food, price, service, and ambiance. The $UserAspectRatinging$ (A_k, u_i) utilized as the proxy for the user's preference criteria weights. The restaurant score with respect to a user is evaluated as follows:

$$RntScore(u_i, r_j) = L(u_i, r_j) \times \sum_{k=1}^{4} \left(UserAspectRing\ (A_k, u_i) \times RntAspectRing(r_j, A_k) \right) \tag{12}$$

3.5.2 Restaurant List Forming

After calculating the data represented above, we obtain the $RntScore(u_i, r_j)$ by considering the user u_i's needs in different aspects. We compute the objective degree of all the reviews associated with a certain restaurant r_j since people trust more objective opinions and objective posts are more plausible. We can determine the subjective of the reviews on a restaurant r_j from TextBlob, and the objectiveness of the reviews on a restaurant r_j can be obtained as follows (Tafesse, 2021):

$$ObjectiveScore(r_j) = 1 - SubjectiveScore(r_j) \tag{13}$$

To balance the factors of personal preference for restaurants and the objectiveness of reviews, we rank the score of a restaurant r_j with respect to user u_i, calculated as follows:

$$RankScore(u_i, r_j) = RntScore(u_i, r_j) \times ObjectiveScore(r_j) \tag{14}$$

Finally, we generate a list of restaurants ranked according to their respective scores based on the potential restaurants.

4. Experiments

This research uses the Yelp platform as our primary reference platform. Yelp is a national restaurant guide website that serves 32 countries. We gather information about restaurants and diners, including reviews, locations, ratings, and other relevant data from the platform. Subsequently, we analyze and categorize the restaurants' aspect ratings by employing sentiment analysis of the reviews and social influence analysis of the reviewers. Finally, the results of the analysis are leveraged for restaurant recommendations.

4.1. Data Collection

Yelp is dedicated to simplifying the finding of reputable local merchants, with restaurant information being one of its primary offerings. Yelp released an academic dataset comprising information about 174,000 local businesses across 11 metropolitan areas spanning 4 countries, sourced from Yelp's open dataset for 2018. This dataset encompasses 5.2 million reviews, which include star ratings, review content, the count of "useful," "funny," and "cool" votes by users, as well as the profiles of the users who submitted the reviews. The dataset also includes information about various types of businesses, such as hair salons, furniture stores, and health and medical care establishments. Following data screening procedures, the dataset contains 54,618 catering-related businesses, with 3.2 million reviews remaining.

4.2. Aspect-Based Sentiment Analysis

Preprocessing: To extract the related restaurant aspects, such as food, price, service, and ambiance, we preprocess the data. Initially, we filter out irrelevant words in each review using IXA-pipe tools built on the natural language toolkit (NLTK). The IXA pipeline involves three modules: (1) IXA-pipe-tok: This module handles token offset and sentence splitting. (2) IXA-pipe-pos: A multilingual part-of-speech (POS) tagger and lemmatizer. (3) IXA-pipe-nerc:

A multilingual named entity tagger, which processes the results from IXA-pipe-tok and IXA-pipe-pos. Notably, the named entity recognition (NER) performance in IXA-pipe-nerc is considered superior (Agerri et al., 2014). Additionally, the opinion target extraction (OTE) of restaurant models available in the IXA-pipe-nerc for SemEval 2015 is employed to enhance the accuracy of named entity extraction. The IXA-pipe-tok is used for offset and sentence splitting. The IXA-pipe-tok handles offset and sentence splitting, IXA-pipe-pos serves as a part-of-speech (POS) tagger on a sentence level, and IXA-pipe-nerc performs named entity recognition based on the results of the POS tagging process. For tokenization, the morphofeat format includes "NN" for nouns, "JJ" for adjectives, and "RB" for adverbs.

Feature extraction: CoreNLP Parser API can help us know which noun this adjective refers to after the preprocessing phase and judge the feature. The result of calculated relevance is shown in Figure 4.

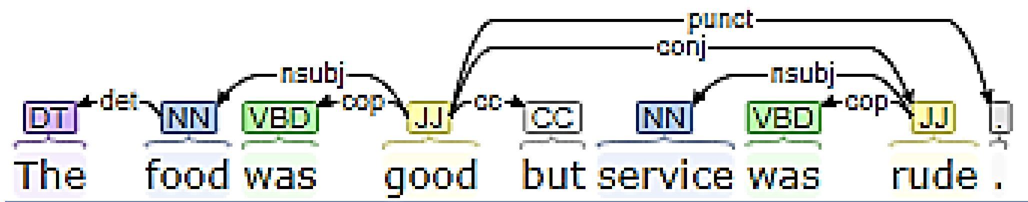

Figure 4: The Result of Relevance Calculation

We calculate the SemEval classification data to determine the aspect to which each feature belongs. Words with similar meanings are assessed for their similarity with existing features in the classification data. The word is identified as an existing feature if the similarity is high. Once all features are obtained, we classify aspects using KNN (K-Nearest Neighbors). Next, we conduct aspect-based sentiment analysis using one aspect with one short sentence containing adjectives. By leveraging TextBlob, a Python library for sentiment analysis, we can obtain the short sentences' semantic polarity value and subjective value. A polarity value greater than 0 indicates a positive sentiment, while a polarity value less than 0 indicates a negative sentiment. A polarity value equal to 0 denotes a neutral sentiment. After aspect-based sentiment analysis, we obtain a polarity label, which is recorded in the database as shown in Table 3.

Table 3: Database Field Presentation

The sentence	**Pos tag**	**Neg tag**
The food was good, but the service was rude.	Food	Service

4.3. Restaurant Aspect Score Calculation

One restaurant may have multiple reviews, each potentially originating from users with varying levels of influence. For instance, consider the following example with two sentences, one posted by a user with high influence and the other by a user with low influence (Table 4).

Table 4: Restaurant Review Sentence

The sentence	**Pos tag**	**Neg tag**	**Reviewer influence**
The food was good, but the service was rude.	Food	Service	High influence
The pizza here made my night... Good people and great pizza.	Food		Low influence

In order to calculate the aspect rating, we tally the occurrences of each aspect on the positive side and count the number of negative occurrences related to each aspect. To enhance the credibility of a restaurant's reviews, if a user shares a highly influential post, the number of tags for that aspect would be doubled. The outcome of the restaurant rating calculation is presented in Table 5.

Table 5: The Result of the Restaurant Rating Calculation

Aspect	Pos	Neg	Rating [-1,1]
Food	3	0	1 [5]
Price	0	0	0 [2.5]
Service	0	2	-1 [0]
Ambiance	0	0	0 [2.5]

4.4. Personal Interest Analysis

4.4.1. User Location Calculation

When determining user preferences, we examine the posts a user has shared in the past, specifically focusing on those related to restaurants. The initial step involves identifying the user's customary areas where they frequently visit restaurants. To mitigate the impact of outliers, we analyze the cities where most of the user's past posts were located.

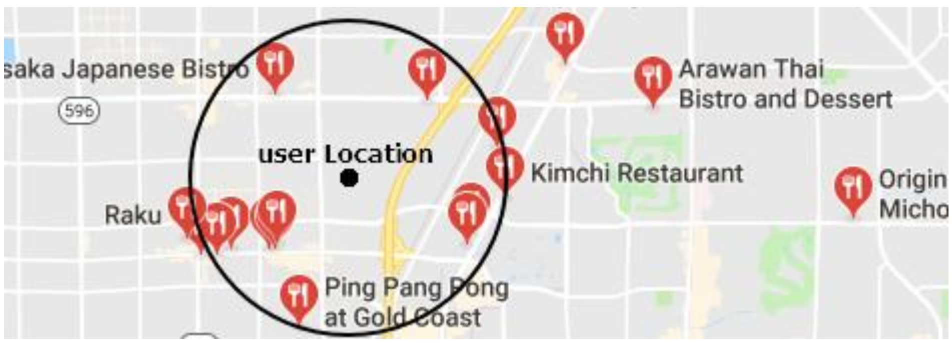

Figure 5: The Location of the Customary Area

As depicted in Figure 5, if the marks represent the locations a user has visited, we can compute the distance between a restaurant and the user's location using the latitude and longitude information of the restaurant. The outcome of the customary location and restaurant distance calculation is presented in Table 6.

Table 6: The Result of Customary Location and Restaurant Distance

Restaurant	Latitude	Longitude	$L(u_i, r_j)$
R1	36.1273	-115.224	0.965
R2	36.1022	-115.17	0.962

The restaurant's proximity to the customary location is considered, and further evaluation is conducted to determine whether it meets the user's needs.

4.4.2. User Aspect-Based Sentiment Calculation

These review posts can also provide insights into what aspects users focus on through aspect-based sentiment analysis. For example, Table 7 illustrates sentences from a user's review.

Table 7: User's Restaurant Review Sentences in the Past

Sentence	Pos tag	Neg tag
The food was good, but the service was rude.	Food	Service
The pizza here made my night... Good people and great pizza.	Food	

After conducting aspect-based sentiment analysis, we tally the number of positive and negative tags associated with each aspect. The difference from calculating restaurants is that even if a negative sentiment is expressed regarding a particular aspect, it does not necessarily mean that the star rating for that aspect should be low; rather, it may require a higher number of stars to compensate for the negativity. We compute the user's rating according to Eq. (5–7). To record user preferences, a new field is created to store the star rating for each user. Pscore is based on the number of positive tags, with food being considered more important than other aspects in the positive sentiment because other aspects were not mentioned. Nscore is based on the number of negative tags, with service being considered more important than other aspects in the negative sentiment because other aspects were not mentioned. The outcome of the aspect score calculation is shown in Table 8.

Table 8: The Result of User Aspect Score Calculation

Aspect	Pos	Neg	PScore	NScore	PNScore	Rating [0,5]
Food	2	0	1	0	0.5	1[5]
Price	0	0	0	0	0	0[0]
Service	0	1	0	1	0.5	1[5]
Ambiance	0	0	0	0	0	0[0]
Min	0	0			0	
Max	2	1			0.5	

4.5. Restaurant Recommendation List

The restaurant recommendation list should meet user requirements, including the lowest star rating of the user aspect and location in the user's customary area. Each user has different weights on various aspects. For example, assume the user aspect requirement is Food: 5, Price: 0, Service: 5, Ambiance: 0, and the score of a restaurant would be calculated as shown in Table 9.

Table 9: The Result of Score Calculation

Restaurant	Food	Price	Service	Ambiance	$L(u_i, r_j)$	RntScore
R1	5	5	4	2	0.965	43.43
R2	2	4	1	4	0.962	14.43
R3	3	3	5	3	0.955	38.2

The restaurant reviews are considered more objective if more people support them. Therefore, the aspect scores and the objective degree are combined as the basis for ranking. The resulting rank scores of the restaurants are shown in Table 10.

Table 10: The Result of Rank Score Calculation

Rank	Restaurant	RntScore	ObjectiveScore	RankScore
1	R1	43.43	0.8	34.74
2	R3	38.2	0.7	26.74
3	R2	14.43	0.6	8.66

5. Results and Evaluation

We compare our proposed recommendation approach with benchmark methods using commonly used performance measures in recommendation evaluation. Precision, recall, accuracy, and F-score are utilized to evaluate the effectiveness of aspect classification, as these metrics are frequently employed in sentiment analysis studies. Subsequently, we assess the accuracy of restaurant aspect ratings and whether they meet or exceed user expectations. We derive restaurant ratings based on their previous posts, and the accuracy indicates that restaurants described as satisfactory in the posts also perform well. Finally, we calculate accuracy, precision, recall, and F-score to evaluate the recommendation performance and determine whether the recommended restaurant aligns with consumer preferences. Higher correct rates indicate better recommendation performance.

5.1 Aspect-Based Sentiment Analysis Evaluation

To evaluate the effectiveness of aspect-based sentiment analysis in the proposed mechanism, we utilized the SemEval 2016 database, which contains 845 reviews comprising 5070 sentences. The dataset includes pre-defined classification aspects and corresponding emotions. Specifically, 342 food-related aspects, 71 price-related aspects, 175 service-related aspects, and 77 ambiance-related aspects are included in the database. The aspect discriminator employed in our mechanism can be compared to manual human classification, indicating a significant degree of practicality. Precision, recall, and F-score are used as metrics for determining the precision of aspect reviews, and they are calculated using the following formula:

Precision is a measure of exactness that determines the fraction of relevant items retrieved out of all items retrieved. It is calculated using the following formula:

$$Precision = \frac{|set\ of\ correctly\ tagged\ clauses|}{|set\ of\ automatically\ tagged\ clauses|} \quad (15)$$

Recall is a measure of completeness that determines the fraction of relevant items retrieved out of all relevant items. It is calculated using the following formula:

$$Recall = \frac{|set\ of\ correctly\ tagged\ clauses|}{|set\ of\ automatically\ tagged\ relevant\ clauses|} \quad (16)$$

F-score, also known as the F1-score, is a measure that attempts to combine Precision and Recall into a single value for comparison purposes. It is calculated using the following formula:

$$F - score = 2 * [(Precision \times Recall)/(Precision + Recall)] \quad (17)$$

The results of accuracy, precision, and recall are shown in Table 11.

Table 11: The Result of Accuracy, Precision, and Recall

Aspect	Precision	Recall	F-score
Food	0.71	0.85	0.78
Price	0.47	0.38	0.42
Service	0.60	0.58	0.59
Ambiance	0.38	0.30	0.33

It can be observed from Figure 6 that the food aspect has the best performance in precision, recall, and F-score, with the service aspect being the second highest. The food aspect demonstrates the highest performance for two potential reasons. First, the number of food-related aspects is larger than other aspects in the SemEval 2016 database. This larger dataset provides more training data, allowing the model to learn and achieve better classification accuracy. Second, reviews often provide more detailed and explicit opinions about food, making classifying it easier. After analyzing these aspects individually, the overall effectiveness of sentiment analysis is evaluated.

Figure 6: The Result of Precision, Recall, and F-score

Precision, recall, and accuracy are used to evaluate the performance of sentiment analysis. They are measured as follows:

$$Precision = \frac{|set\ of\ correctly\ classified\ positive\ or\ negative\ sentences\ or\ clauses|}{|set\ of\ automatically\ classified\ positive\ or\ negative\ sentences\ or\ clauses|} \tag{18}$$

$$Recall = \frac{|set\ of\ correctly\ classified\ positive\ or\ negative\ sentences\ or\ clauses\ |}{|set\ of\ manually\ classified\ relevant\ positive\ or\ negative\ sentences\ or\ clauses|} \tag{19}$$

$$Accuracy = \frac{|set\ of\ correctly\ classified\ positive\ and\ negative\ sentences\ or\ clauses|}{|set\ of\ manually\ classified\ relevant\ positive\ and\ negative\ sentences\ or\ clauses|} \tag{20}$$

We count the number of positive and negative terms in each sentence. Neutral words are excluded as they are not included in the database. The sentiment of a sentence is determined based on the count of positive and negative terms. The overall results are displayed in Table 12.

Table 12: Precision, Recall, F-score, and Accuracy of the Baseline Sentiment Word Count Approach

		Precision	Recall	F-score	Accuracy
Sentiment Level	Positive	0.93	0.77	0.84	0.77
	Negative	0.47	0.79	0.59	
Aspect Level					
Food	Positive	0.92	0.80	0.86	0.78
	Negative	0.73	0.47	0.57	
Price	Positive	0.89	0.65	0.76	0.68
	Negative	0.75	0.40	0.52	
Service	Positive	0.98	0.68	0.80	0.74
	Negative	0.48	0.95	0.64	
Ambiance	Positive	0.95	0.86	0.90	0.84
	Negative	0.50	0.75	0.60	

Figure 7 shows that positive sentiment performance is superior to negative sentiment in sentiment analysis. This difference between the positive and negative classes indicates that the model is more effective at correctly predicting positive than negative class instances. Figure 8 further illustrates that positive performance consistently outperforms negative sentiment across all aspects of sentiment analysis. Overall, the results indicate that the ambiance and food aspects are viewed positively, while the price aspect receives the least positive sentiment. The service aspect has the highest negative sentiment score, suggesting that it may be an area that needs to improve. Figure 9 provides insight into the accuracy of different aspects, with the "ambiance" aspect demonstrating the best performance. This high accuracy of the ambiance aspect could be due to more consistent and clearly expressed opinions in customer reviews.

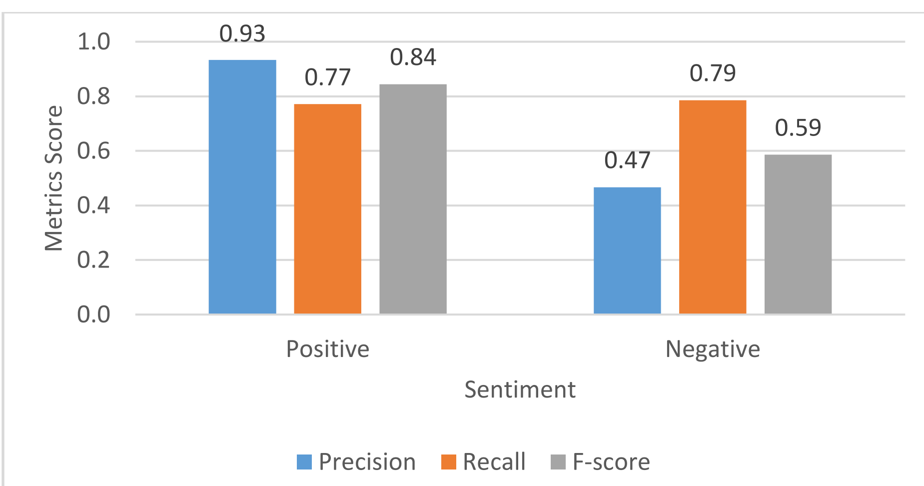

Figure 7: Positive and Negative Performance Metrics (Precision, Recall, and F-score)

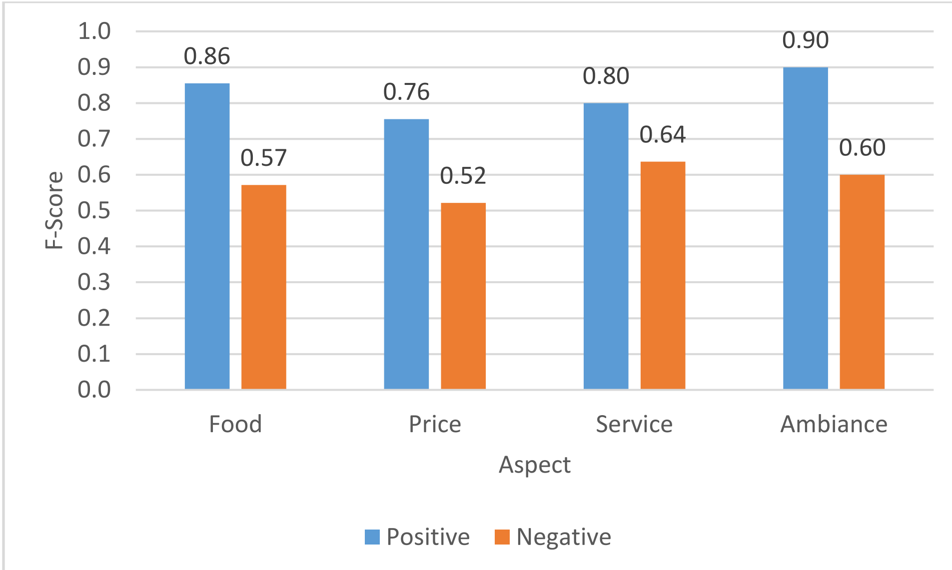

Figure 8: Results of F-score for Each Aspect in Positive and Negative Sentiment Analysis

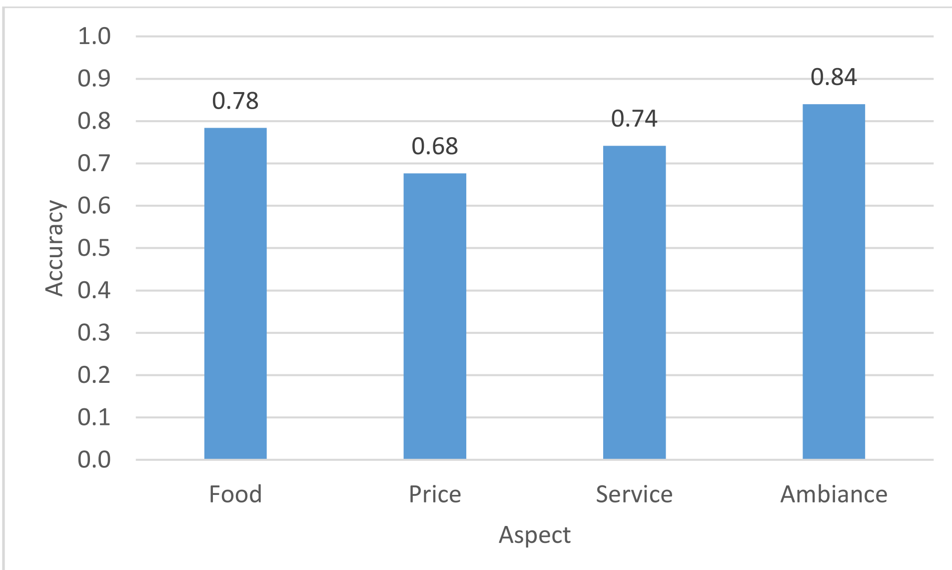

Figure 9: Accuracy Results for Each Aspect in Sentiment Analysis

5.2 User Preference Alignment Evaluation

We conducted further evaluation to assess the accuracy of whether the restaurants recommended by the proposed mechanism align with a user's potential needs. Each restaurant is rated for each aspect and those restaurants with ratings higher than the user aspect rating score are deemed to meet their potential needs. For this evaluation, we collected and analyzed the restaurants reviewed by 100 randomly selected users from the dataset. Accuracy is calculated using the following formula:

$$Accuracy = \frac{|set\ of\ the\ restaurants\ satisfying\ \ user\ demand|}{|set\ of\ all\ restaurants\ \ recommended\ to\ \ the\ users\ |} \tag{21}$$

Through evaluation (Figure 6), the aspects of food and service demonstrate better performance, while the price and ambiance aspects need improvement. The semantic analysis shows good performance in Figure 9. It is essential to use user preference as a benchmark. Individual accuracy in each aspect is shown in Table 13, with an average result of preference alignment accuracy is 0.72. Overall, the results suggest that the recommendation mechanism effectively aligns with the users' preferences across all aspects. This means that most restaurants they went to had met their potential needs. From Figure 10, we can observe the price aspect has the best performance of user demand analysis.

Table 13: Accuracy of Each Aspect

Aspect	Accuracy
Food	0.69
Price	0.74
Service	0.70
Ambiance	0.73

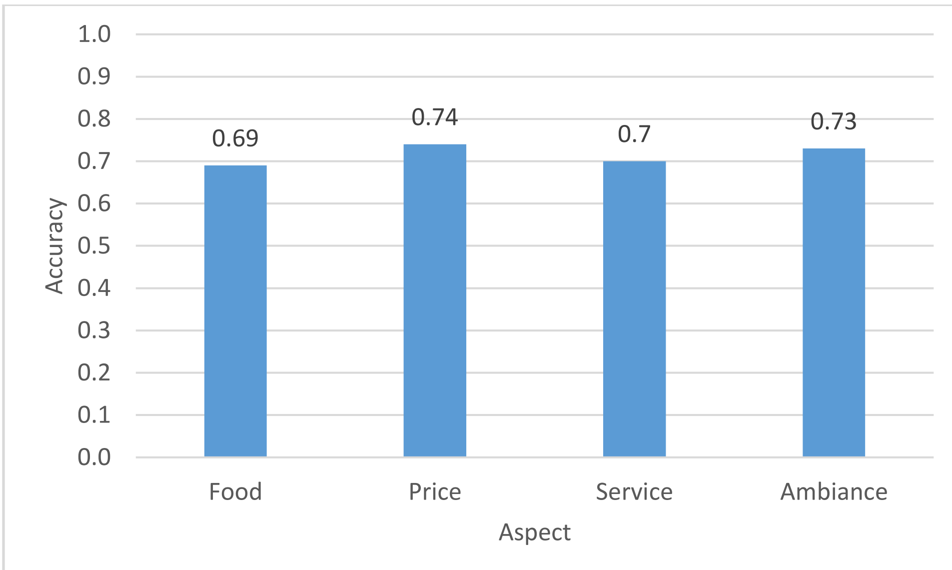

Figure 10: Accuracy of Each Aspect in User Preference Alignment Analysis

5.3. Recommendation Evaluation

To evaluate the recommendation system, users who have shared at least 30 posts before were selected. Their posts were split into training and validation sets. Based on the data from the user's previous posts, if the content of the post fits the user's requirements, the score is deemed "Actually fit." If the content does not fit the user's requirements, the score is deemed "Not actually fit." The evaluation in information retrieval (IR) is shown in Table 14.

Table 14: Information Retrieval

		Reality	
Prediction		Actually fit	Not actually fit
	Fit	TP	FP
	Not fit	FN	TN

Accuracy, commonly used in evaluating recommendation systems, represents the proportion of correct recommendations out of all recommended items. It is calculated by dividing the number of correct recommended items by the total number of recommended items. A higher accuracy rate indicates a better quality of the recommended list.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{|Set\ of\ all\ correct\ predictions|}{|Set\ of\ all\ predictions\ or\ reality|} \quad (22)$$

$$Precision = \frac{TP}{TP+FP} = \frac{|Set\ of\ actually\ fit|}{|Set\ of\ prediction\ fit|} \quad (23)$$

$$Recall = \frac{TP}{TP+FN} = \frac{|Set\ of\ prediction\ fit|}{|Set\ of\ all\ user\ actually\ fit|} \quad (24)$$

We randomly selected 80% of the users' dataset for training sets and the remaining 20% for verification sets. The results of accuracy, precision, and recall are shown in Table 15. We compared the differences between our proposed approach and those in previous studies. Our proposed mechanism demonstrated the significant effectiveness of diverse aspect ratings compared to the general single-dimension public rating, which lacks the detail of quality evaluation and personal preference information. Furthermore, besides the impact of diverse aspect ratings, we demonstrate the influence of review credibility (the social influence of the reviewer) on recommendation effectiveness, comparing existing approaches that did not consider the review source credibility (e.g., a reviewer's expertise and social influence). The results of this paper show that multidimensional ratings generate heterogeneous effects on review helpfulness compared with the features considered by existing related studies. The experimental results show that our proposed model (LPI) outperforms the other modules. The progressive enhancement from module L to module LPI highlights the importance and effectiveness of these extra features in refining the model's predictive capabilities. Module LPI demonstrates the best overall performance, suggesting that adding additional features significantly enhances the model's ability to predict accurately. This improvement highlights these additional features' critical role in enhancing the model's robustness and reliability. The comparison line chart is shown in Figure 11. The comparison benchmarks include:

(1) L approach: Single dimension rating (Munaji & Emanuel, 2021)
(2) LI approach: Single dimension rating +Reviewer's influence (Petrusel & Limboi, 2019; Margaris et al., 2020)
(3) LP approach: Aspect rating (Zhu et al., 2022; Li et al., 2023)
(4) LPI approach: Aspect rating+ Reviewer's influence (our collective recommendation approach)

Table 15: The Result of the Accuracy, Precision, Recall, and F-score

Module	Accuracy	Precision	Recall	F-score
L	0.35	1.00	0.35	0.52
LI	0.61	1.00	0.29	0.44
LP	0.69	1.00	0.43	0.60
LPI	0.73	1.00	0.61	0.76

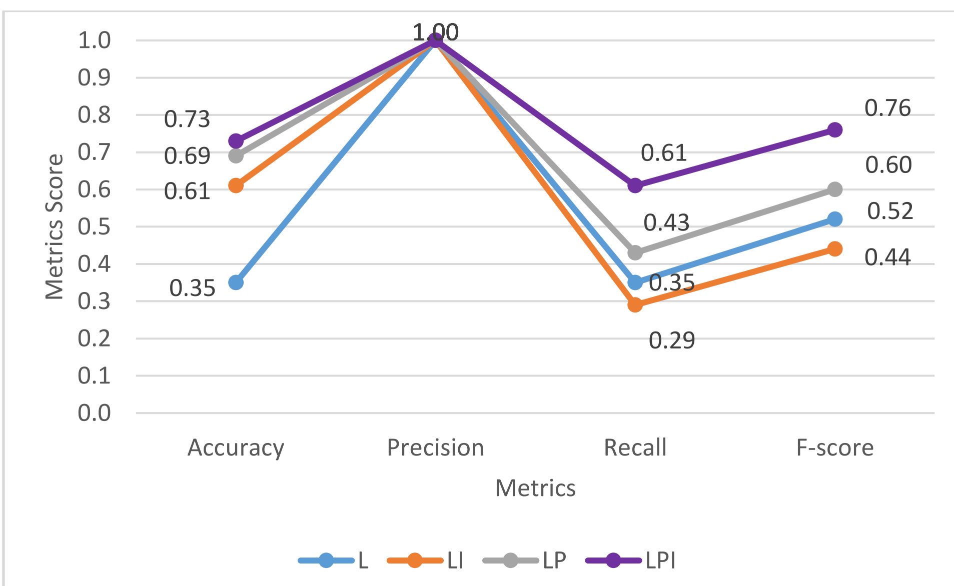

Figure 11: Results of Accuracy, Precision, Recall, and F-score for Each Module

6. Discussion and Conclusion

The proposed recommendation mechanism addresses the following research questions based on our experimental results. Firstly, the single-dimensional rating system cannot comprehensively convey information on different aspects. Our findings indicate that multidimensional ratings produce heterogeneous effects on review helpfulness. By aggregating ratings across these dimensions, a more comprehensive view of the restaurant's performance can emerge, enabling consumers to make more informed dining decisions. Secondly, due to the varying importance for different customers, reviews are weighted based on the reviewers' influence, significantly improving the credibility of the review and recommendation effectiveness. Thirdly, by analyzing the sentiment of user reviews regarding aspects such as food, price, service, and ambiance, the proposed mechanism can identify nearby restaurants and provide recommendations better aligned with a diner's preference.

6.1 Research Contributions

6.1.1 Theoretical Contributions

Existing Yelp review platforms allow consumers to share their opinions about restaurants. However, the abundance of reviews and overall ratings makes it challenging for users to extract meaningful insights from the vast amount of information available. Our mechanism addresses this issue by effectively processing large datasets and generating a list of recommendations to assist users in decision-making. Our research makes several significant theoretical contributions.

First, we propose a theoretical framework to design a novel collective recommendation artifact for restaurant selection decision support, which consolidates the public review opinions by integrating multicriteria appraisal (aspect-based sentiment and aspect-based rating), content credibility (review's objectiveness), source credibility (reviewer's expertise and reviewer's influence), and preference alignment (spatial proximity and preference similarity). While previous studies have explored how ratings and text reviews influence consumers' purchasing decisions (Jabr & Rahman, 2022; Lei et al., 2022), they often rely solely on restaurant attributes with sentiment analysis without considering the social influence and diner preference. Second, to address this literature gap, our research integrates aspect-based sentiment analysis (María et al., 2019; Binder et al., 2019) and social influence (Li et al, 2021) to enhance the quality of recommendations. Unlike previous studies that only utilize a subset of features, our approach considers comprehensive factors across various aspects and mitigates potential biases introduced by influential users. Additionally, we incorporate personal preferences and objectives to adjust the aspect-based rating scores, enhancing fairness and reliability. Third, from the methodological perspective, we integrate text mining (María et al., 2019; Binder et al., 2019), machine learning (María et al., 2019; Binder et al., 2019), and social network analysis (Li et al, 2021) to derive collective recommendations. Taken together by analyzing aspect-based sentiment, extracting user preferences and restaurant attributes, and considering the reviewer's social influence, we provide users with more detailed and objective information, particularly in cases of information insufficiency or overload.

6.1.2 Practical Implications

Our study provides actionable implications for managers, consumers, and platform providers regarding collective recommendation systems. These systems comprise credible ratings and review data, aiding consumers in decision-making. Our research identifies several significant practical implications for deriving collective recommendations. Firstly, enhancing product visibility through diverse rating generation from a collective pool of crowd. This visibility significantly influences purchase decisions, as customers are more likely to trust multidimensional ratings over a single overall rating, thus attracting potential buyers for managers. This enables customers to assess the quality and reliability of products quickly. Secondly, informing service development according to consumer preferences. Our system framework incorporates location fitness and personal preference of aspect computing. This approach not only assists managers in meeting the evolving needs of consumers and market trends but also enables consumers to seek a suitable restaurant. Finally, leveraging the reviewers' expertise improves the credibility of recommendation systems. Encouraging customers to share feedback and contribute can build long-lasting relationships between platform providers and consumers. The proposed collective recommendation approach mitigates the potential bias and limited information stemming from the subjectivity of large-scale ratings and leverages both multidimensional and meaningful aspect ratings derived from credible crowd reviews.

6.2 Research Limitations

This study has several limitations. Firstly, it only extracts data from the Yelp website, whereas numerous other review websites like Google Restaurant and ipean.com exist. Incorporating information from different platforms could potentially enhance the performance of recommendations. Secondly, user preferences are described based on geographic proximity and interest similarity. To improve this, we could further incorporate features related to long-term habitual tendencies and short-term practical needs. Thirdly, each reviewer's influence on users may vary, yet in this study, the influence is treated as a static parameter rather than dynamically adjusting. Evaluating social influence could be enhanced by considering other social activities and relationships. Fourthly, the recommendation list generated in this research primarily focuses on local recommendations due to the reliance on past data. Recommendations for restaurants in other countries need alternative evaluation methods. Lastly, semantic analysis requires sufficient past restaurant reviews and user preferences. If there is insufficient data for the restaurant or the user, methods must be implemented to address the cold start issue.

6.3 Future works

Several related issues can be further studied. Firstly, many factors that may reflect or affect recommendations could be further incorporated. For example, it has been suggested that temperature and season will affect user preferences. Additionally, preference analysis could consider other activities revealed on social media for further analysis of implicit tastes and preferences. Secondly, social influence plays an important role in information diffusion. We can analyze more implicit or sophisticated dimensions. For example, a person's influence is associated with their friends' influence and the physical areas or interest communities, which could be dynamically determined. Thirdly, in our experiment, we used KNN for grouping and the IXA-pipeline and TextBlob for semantic analysis. However, other text-mining and classification approaches could still be exploited to improve the effectiveness of clustering and the accuracy of sentiment analysis. Fourth, the textual review analysis can thoroughly explore language style diversity. Analyzing language style diversity in reviews offers deeper insights into consumer sentiments, preferences, the detection of fake reviews, and cultural contexts. Furthermore, it aids in weighting reviews according to their relevance or credibility. Lastly, this research mainly analyzed rating websites, but other websites may use concepts such as group purchases. In future work, we can analyze the preferences of both parties in terms of which aspects they like and recommend products or coupons that fit the user.

Acknowledgment

This research was supported by the National Science and Technology Council of Taiwan (Republic of China) under the grant NSTC 113-2410-H-141-02.

REFERENCES

Agerri, R., Bermudez, J., & Rigau, G. (2014). IXA pipeline: Efficient and ready to use multilingual NLP tool. *in Proceedings of the 9th Language Resources and Evaluation Conference (LREC2014)*, 26–31.

Asani, E., Vahdat-Nejad, H., & Sadri, J. (2021). Restaurant recommender system based on sentiment analysis. *Machine Learning with Applications*, *6*, 100114.

Banerjee, S., Bhattacharyya, S., & Bose, I. (2017). Whose online reviews to trust? Understanding reviewer trustworthiness and its impact on business. *Decision Support Systems*, *96*, 17–26.

Belkin, N. J. (1984). Cognitive models and information transfer. *Social Science Information Studies*, *4*, 111-129.

Binder, M., Heinrich, B., Klier, M., Obermeier, A., & Schiller, A. (2019). Explaining the stars: Aspect-based sentiment analysis of online customer reviews. *in Proceedings of the 27th European Conference on Information Systems (ECIS2019)*.

Blanding, M. (2011). The Yelp factor: Are consumer reviews good for business? *Harvard Business School*, Weekly Newsletter

Büschken, J., & Allenby, G. M. (2016). Sentence-based text analysis for consumer reviews. *Marketing Science*, *35*(6), 953-975.

Carenini, G., Ng, R. T., & Zwart, E. (2005). Extracting knowledge from evaluative text, *in Proceedings of the 3rd International Conference on Knowledge Capture (K-CAP '05)*, 11-18.

Chakraborty, I., Kim, M., & Sudhir, K. (2022). Attribute sentiment scoring with online text reviews: Accounting for language structure and missing attributes. *Journal of Marketing Research*, *59*(3), 600-622.

Chang, Y. C., Ku, C. H., & Le Nguyen, D. D. (2022). Predicting aspect-based sentiment using deep learning and information visualization: The impact of COVID-19 on the airline industry. *Information & Management*, *59*(2), 103587.

Chen, P. Y., Hong, Y., & Liu, Y. (2018). The value of multidimensional rating systems: Evidence from a natural experiment and randomized experiments. *Management Science*, *64*(10), 4629-4647.

Chen, Y. Y., Ferrer, X., Wiratunga, N., & Plaza, E. (2014). Sentiment and preference guided social recommendation. *in Proceedings of the International Conference on Case-Based Reasoning (ICCBR), 8765*, 79–94.

Colladon, A. F., Guardabascio, B., & Innarella, R. (2019). Using social network and semantic analysis to analyze online travel forums and forecast tourism demand. *Decision Support Systems*, *123*, 113075.

Dave, K., Lawrence, S., & Pennock, D. M. (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. *in Proceedings of the Twelfth International Conference on World Wide Web (WWW '03)*, 519–528.

Dai, W., Jin, G., Lee, J., & Luca, M. (2018). Aggregation of consumer ratings: An application to Yelp.com.

Quantitative Marketing and Economics, *16*(3), 289–339.

Deutsch, M., & Gerard, H. B. (1955). A Study of normative and informational social influences upon individual judgement. *Journal of Abnormal and Social Psychology*, *51*, 629-636.

Divyaa, L. R., & Pervin, N. (2019). Towards generating scalable personalized recommendations: integrating social trust, social bias, and geo-spatial clustering. *Decision Support Systems, 122*, 113066.

Do, H. H., Prasad, P. W., Maag, A., & Alsadoon, A. (2019). Deep learning for aspect-based sentiment analysis: a comparative review. *Expert systems with applications*, *118*, 272-299.

Ganu, G., Elhadad, N., & Marian, A. (2009). Beyond the stars: Improving rating predictions using review text content. *in Proceedings of the Twelfth International Workshop on the Web and Databases (WebDB)*.

Godes, D., & Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing Science*, *23*(4), 545–560.

Hajas, P., Gutierrez, L., & Krishnamoorthy, M. S. (2014). Analysis of Yelp reviews, *arXiv preprint arXiv:1407.1443*.

Han, M. (2022). How does mobile device usage influence review helpfulness through consumer evaluation? Evidence from TripAdvisor. *Decision Support Systems*, *153*, 113682.

Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J.T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, *22*(1), 5-53.

Jabr, W., & Rahman, M. S. (2022). Online reviews and information overload: The role of selective, parsimonious, and concordant top reviews. *MIS Quarterly*, *46*(3), 1517-1550.

Jong, J. (2011). Predicting rating with sentiment analysis. *Stanford University.*, *Stanford*, *CA*, 1-5.

Kim, J., Naylor, G., Sivadas, E., & Sugumaran, V. (2016). The unrealized value of incentivized eWOM recommendations. *Marketing Letters*, *27*, 411–421.

Kotler, P. (2003). Marketing management. *Prentice Hall*.

Lei, Z., Yin, D., Mitra, S., & Zhang, H. (2022). Swayed by the reviews: disentangling the effects of average ratings and individual reviews in online word-of-mouth. *Production and Operations Management*, *31*, 2393-2411.

Li, H., Yu, B. X. B., Li, G., & Gao, H. (2023). Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews. *Tourism Management*, *96,* 104707.

Li, J., Xu, X., & Ngai, E. W. T. (2021). Does certainty tone matter? Effects of review certainty, reviewer characteristics, and organizational niche width on review usefulness. *Information & Management*, *58*(8), 103549.

Liu, X., Li, C., Nicolau, J. L., & Han, M. (2023). The value of rating diversity within multidimensional rating system: Evidence from booking platform. *International Journal of Hospitality Management*, *110*, 103434.

Lopes, A. I., Dens, N., & Pelsmacker, P. D. (2022). Valence and attribute repetition in negative sets of online reviews: (When) Can positive reviews overcome negative ones? *Journal of Electronic Commerce Research*, *23*(1): 1–10.

Lu, X., He, S., Lian, S., Ba, S., & Wu, J. (2020). Is user-generated content always helpful? The effects of online forum browsing on consumers' travel purchase decisions. *Decision Support Systems*, *137*, 113368.

Luarn, P., Yang, J. C., & Chiu, Y. P. (2015). Why people check in to social network sites. *International Journal of Electronic Commerce*, *19*(4), 21–46.

Luca, M. (2011). Reviews, reputation, and revenue: The case of Yelp.com. *Harvard Business School NOM Unit Working Paper (12-016)*.

Luo, Y., Tang L., Kim E. & Wang X. (2020). Finding the reviews on yelp that actually matter to me: Innovative approach of improving recommender systems. *International Journal of Hospitality Management*, *91*, 102697.

Manning, C., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S., & McClosky, D. (2014). The stanford CoreNLP natural language processing toolkit. *in Proceedings of 52nd Annual Meeting of the association for computational linguistics: System Demonstrations*, 55–60.

María, H., Cantador, I., & Bellogín, A. (2019). A comparative analysis of recommender systems based on item aspect opinions extracted from user reviews. *User Modeling and User-Adapted Interaction*. *29,* 381–441.

Margaris, D., Vassilakis, C., & Spiliotopoulos, D. (2020). What makes a review a reliable rating in recommender systems? *Information Processing & Management*, *57*(6), 102304.

Marrese-Taylor, E., Velásquez, J. D., Bravo-Marquez, F., & Matsuo, Y. (2013). Identifying customer preferences about tourism products using an aspect-based opinion mining approach. *Procedia Computer Science*, *22*, 182–191.

McCroskey, J. C., & Jenson T. A. (1975). Image of mass media news sources. *Journal of Broadcasting*, *19*(2), 169-180.

Mudambi, S. M., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, *34*(1), 185-200.

Mudambi, S. M., Schuff, D., & Zhang, Z. (2014). Why aren't the stars aligned? An analysis of online review content and star ratings. *in Proceedings of the 47th Hawaii International Conference on System Science*, 3139-3147.

Munaji, A. A., & Emanuel, A. W. R. (2021). Restaurant recommendation system based on user ratings with

collaborative filtering. *IOP Publishing*, *1077*, 012026.
Nakayama, M., & Wan, Y. (2019). The cultural impact on social commerce: A sentiment analysis on Yelp ethnic restaurant reviews. *Information & Management*, *56*, 271–279.
National Restaurant Association. (2022). Get the facts with the latest National Statistics, National Restaurant Association. Retrieved from: https://restaurant.org/research-and-media/research/industry-statistics/national-statistics.
Nguyen, P., Wang, X., Li, X., & Cotte, J. (2021). Reviewing experts' restraint from extremes and its impact on service providers. *Journal of Consumer Research*, *47*(5), 654-674.
O'Donovan, J., & Smyth, B. (2005). Trust in recommender systems. *in Proceedings of the 10th International Conference On Intelligent User Interfaces(IUI '05)*, 167-174.
Peng, Q., You, L., Feng, H., Du, W., Zheng, K., Zhu, F., & Xu, X. (2022) Jointly Modeling Aspect Information and Ratings for Review Rating Prediction. *electronics*, *11*(21), 3532.
Petrusel, M.-R., & Limboi S.-G. (2019). A restaurants recommendation system: Improving rating predictions using sentiment analysis. *International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*(*IEEE)*, 190–197.
Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., AL-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., Clercq, O.D., Hoste, V., Apidianaki, M., Tannier, Loukachevitch, X., Kotelnikov, N. E., & Bel, N. (2016). SemEval-2016 Task 5: aspect based sentiment analysis. *In Proceedings of the 10th International Workshop on Semantic Evaluation*, 19–30.
Pronoza, E., Yagunova, E., & Volskaya, S. (2016). Aspect-based restaurant information extraction for the recommendation system. *Challenges for Computer Science and Linguistics*, *9561*, 371–385.
Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. *The MIT Press*, 73–100.
Suresh, V., Roohi, S., & Eirinaki, M. (2014). Aspect-based opinion mining and recommendation system for restaurant reviews. *in Proceedings of the 8th ACM Conference on Recommender Systems (RecSys '14)*, 361–362.
Tafesse, W. (2021). Communicating crowdfunding campaigns: How message strategy, vivid media use and product type influence campaign success. *Journal of Business Research*, *127*, 252-263.
Tan, C., Chi, E. H., Huffaker, D., Kossinets, G., & Smola, A. J. (2013). Instant foodie: Predicting expert ratings from grassroots. *in Proceedings of the 22nd ACM International Conference on Conference on Information & Knowledge Management (CIKM '13)*, 1127–1136.
Thet, T., Na, J., & Khoo, C. (2010). Aspect-based sentiment analysis of movie reviews on discussion boards. *Journal of Information Science*, *36*(6), 823–848.
Wan, Y., Nakayama, M., Blodgett, J. G., & Qin, J. (2023). Online service sentiments in transformative society: A cross-cultural analysis. *Journal of Electronic Commerce Research*, *24*(30), 240-252.
Wen, H., Wong, I. A., Kim, S., Badu-Baiden, F., & Ji, K. M. (2021). A multilevel synthesis of subjective and objective measures of foodservices in the experience process. *International Journal of Hospitality Management*, *99*, 103059.
Xiao, B. & Benbasat, I. (2007). E-commerce product recommendation agents: use, characteristics, and impact. *MIS Quarterly*, *31*(1), 137-209.
Yang, F., & Wang, Z. M. (2009). A mobile location-based information recommendation system based on GPS and WEB2.0 services. *WSEAS Transactions on Computers*, *8*(4), 725–734.
Yang, Y., Yang, F., Yi, G., Xia, D., & Li. J. (2023). Product online multidimensional ratings aggregation decision-making model based on group division and attribute interaction. *Engineering Applications of Artificial Intelligence*, *126*, 106835.
Yu, B., Zhou, J., Zhang, Y., & Cao, Y. (2017). Identifying restaurant features via sentiment analysis on Yelp reviews. *arXiv preprint arXiv:1709.08698*, 1-6.
Zagheli, H. R., Zamani, H., & Shakery, A. (2017). A Semantic-aware profile updating model for text recommendation. *in Proceedings of the Eleventh ACM Conference on Recommender Systems (RecSys '17)*, 316–320.
Zhang, C., Zhang, H., & Wang, J. (2018). Personalized restaurant recommendation method combining group correlations and customer preferences. *Information Sciences*, *454*, 128-143.
Zhou, S. & Guo, B. (2017). The order effect on online review helpfulness: A social influence perspective. *Decision Support Systems*, *93*, 77-87.
Zhu, B., Guo, D., & Ren, L. (2022). Consumer preference analysis based on text comments and ratings: A multi-attribute decision-making perspective. *Information & Management*, *59*(3), 103626.