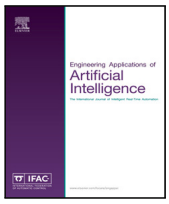
Engineering Applications of Artificial Intelligence 181 (2026) 115532

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai



Research paper

Restaurant vector generation using a word embedding model and its application to restaurant site selection

Yu-Ting Yang [a], Chih-Yung Chang [a,*], Syu-Jhih Jhang [b], Diptendu Sinha Roy [c]

[a] Department of Computer Science and Information Engineering, Tamkang University, New Taipei City 25137, Taiwan
[b] Department of Information Technology, Takming University of Science and Technology, Taipei City 11451, Taiwan
[c] Department of Computer Science and Engineering, National Institute of Technology, Shillong, 793003, India

ARTICLE INFO

Keywords:
Artificial intelligence
Implemented artificial intelligence using a word embedding neural network
Application of artificial intelligence in restaurant site selection
Restaurant site selection
Public restaurant data

ABSTRACT

Restaurant site selection is critical, but complete people-flow and consumption-capacity data are difficult to obtain because local demand depends on transit access, business districts, residents, workers, and private consumption behavior. This study implements artificial intelligence through a continuous bag-of-words-style word-embedding neural network and applies this artificial-intelligence workflow to restaurant site selection. The implemented technique learns restaurant-category vectors from public Google Maps co-location patterns, where neighboring restaurant categories on the same street are used as contextual signals. The proposed method ranks candidate streets by considering attraction from complementary categories and competition from identical or related categories, and serves as a public-data-driven decision-support tool for early-stage screening when audited footfall, transaction, or private consumption data are unavailable. Experimental results show that the learned restaurant-category vectors capture statistically non-random co-location signals at the aggregate category-pair level, while individual embedding geometries are not treated as stable semantic structures across random initializations. The proposed method provides useful candidate shortlisting performance under the filtered binary evaluation protocol, and ranking-oriented diagnostics, sensitivity analyses, perturbation tests, and public-proxy consistency checks clarify the evidential boundary of the method. In the strict large-candidate full-ranking diagnostic, the proposed method achieves a Top-1 hit rate approximately four times the random baseline in the 800-street setting, and the Top-10 hit rate is approximately 1.5 times the random baseline. The results indicate that the framework should be interpreted as a low-cost screening mechanism rather than a causal predictor of commercial demand or a replacement for field investigation.

1. Introduction

The job market is witnessing a recent trend where more and more people are opting to start their own businesses rather than working for other organizations. This career path, especially prevalent in the restaurant business sector, necessitates a greater demand for careful evaluation of various factors to support a new venture. The top priority on the assessment list for restaurant owners is, inevitably, where to locate the new restaurant site, which has been identified by Tayeen et al. (2019) as a critical initial step during the planning process. In the literature, some studies collected data from social media or the Internet and predicted the flow of people based on statistics or Machine Learning (ML) mechanisms.

The quest for the answer to the question—where to place the new restaurant—has become feasible thanks to the advent of technology. In the past, people used to obtain information from television, newspapers, or magazines, which made it time-consuming for the general public to interact with each other and those involved in the business or content creation, such as authorities or stakeholders. In the Internet era, however, mature telecommunication systems and Fifth-Generation (5G) location-aware services have substantially increased the availability of digital activity data. People can now easily access the Internet on the go with their portable electronic devices, such as laptops, tablets, or mobile phones. Once users go online, they leave behind digital footprints, generating large volumes of data on their online activities. Among these activities, users frequently exchange ideas and information with each other in real-time. These social data can reveal key insights into business strategies, helping restaurant owners make informed decisions about site selection during the initial phase and enabling more efficient business operations thereafter.

* Corresponding author.
E-mail addresses: 142720@mail.fju.edu.tw (Y.-T. Yang), cychang@mail.tku.edu.tw (C.-Y. Chang), 810440064@gms.tku.edu.tw (S.-J. Jhang), diptendu.sr@nitm.ac.in (D.S. Roy).

https://doi.org/10.1016/j.engappai.2026.115532
Received 19 November 2024; Received in revised form 22 June 2026; Accepted 24 June 2026
Available online 2 July 2026

The large amount of user data generated online has subsequently become an important data source for deep-learning applications within Artificial Intelligence (AI). Traditionally, researchers selected ideal business sites using statistical methods, manually collecting data (e.g., demographic studies or consumer surveys), and selecting the criteria that determine business success (Dixit et al., 2019). These traditional methods could be time-consuming because they required manual labor at some point during both the data collection and analysis phases. In recent years, AI has gradually become a powerful and effective tool for collecting and analyzing the vast amounts of open data available online. Recognizing this feature, some researchers began to adopt these AI techniques to estimate or recommend restaurant locations. For instance, machine learning techniques are used in Bilen et al. (2018), Bhole et al. (2020), Mazhi et al. (2020), Han et al. (2022), while Deep Learning (DL) techniques are used in Eravci et al. (2016), Chang (2010), Wang et al. (2020). These studies reported predictive improvements under their evaluation settings. However, they generally involved complex methodologies requiring multiple separate methods to arrive at the result. This paper aims to build on the methods used by earlier studies while addressing existing challenges by training AI models using deep learning techniques. Earlier restaurant and business-location studies can be broadly grouped into survey-driven factor analysis, feature-based machine-learning models, review-driven urban analytics, movement-based attractiveness estimation, and relation-based spatial models (Dixit et al., 2019; Bilen et al., 2018; Furtado et al., 2013; Quan et al., 2012; Wang et al., 2016). Related artificial-intelligence decision-support studies include hotel site selection (Dong et al., 2025), offshore wind-farm site selection (Deveci et al., 2021), e-scooter station siting (Altay et al., 2023), electric-vehicle charging-station site selection (Seikh and Mandal, 2022), and electric-vehicle charging-location planning (Krishankumar and Ecer, 2024; Yazir et al., 2025). These studies are used as methodological context for data-constrained location decisions, not as direct restaurant-site-selection benchmarks. Within this landscape, the proposed framework is closest to relation-based spatial models, but it extends them by learning restaurant-category vectors from public restaurant arrangements and directly ranking candidate streets without explicitly collecting footfall or spending data.

Some challenges existed in previous studies attempting to collect and analyze data to measure people flow and consumption capacity in a given area. One major challenge is that many factors impact the flow of people, but it is difficult to collect all relevant information comprehensively. Even if data related to people flow is successfully collected, determining the relative importance (weights) of these factors, when combined into a function to estimate people flow, is difficult. Furthermore, collecting information related to consumption capacity in a given region is even more challenging since such information is often private and difficult to obtain. In cases where both people flow and consumption capacity information are incomplete, predictions regarding the suitability of a given area for a new restaurant may be inaccurate. These data limitations motivate learning from public restaurant arrangements instead of relying only on costly external features.

This work uses the spatial arrangement of existing restaurants as an indirect co-location signal for local context, complementarity, and competition. Its objective is to learn restaurant-category vector representations from public restaurant arrangements and use them to rank candidate streets without explicitly collecting footfall or spending measurements. The contribution is a public-data-driven workflow that learns restaurant-category vectors from same-street arrangements and models both attraction and competition in street ranking.

Unlike most previous studies, this paper uses the same-street relationship as a heuristic street-level locality prior rather than an absolute assumption. Restaurants on the same short commercial block may share observable local context, such as pedestrian access, nearby workers and residents, transit convenience, and local activity, although pedestrian catchment and consumer profiles are not directly observed. This prior may be weaker on very long roads, streets divided by barriers, and newly developed areas. It is therefore examined through leave-one-restaurant-out comparisons, public-proxy checks, vector diagnostics, ranking metrics, parameter sweeps, and perturbation tests. Suppose the new target restaurant is restaurant A. In addition, suppose that restaurant B is the one that is most frequently located on the same street. This suggests that the location of restaurant B could also be a candidate location for restaurant A. Therefore, this paper first uses the neighboring relationships between restaurant categories to construct a restaurant-category vector for each category. The resulting category-level vector is then used to rank candidate streets for the target restaurant.

Accordingly, the evaluation treats the locality prior as a representation assumption rather than a causal statement and uses diagnostics to define its boundary. Throughout the revised manuscript, the term restaurant vector refers to a restaurant-category vector learned from category-level co-location patterns, not to an individual restaurant-level behavioral embedding.

The contributions of the proposed Restaurant Vector Based Location Selection (RVLS) algorithm are as follows:

(1) Construction of Restaurant-Category Vectors. This paper constructs restaurant-category vectors to represent category-level neighboring relations on the same street. The vector is used as an indirect decision-support signal for local restaurant co-location patterns, not as a direct measurement of people flow or consumption capacity. The vector is generated using a continuous bag-of-words style architecture, because the model predicts the center restaurant category from its neighboring categories.
(2) Reduce reliance on difficult-to-obtain data. This paper does not directly collect people-flow measurements or private consumption-capacity data. Instead of treating such unobserved quantities as measured inputs, the proposed RVLS constructs restaurant-category vectors from same-street neighboring category relations.
(3) Diagnostic evaluation of the Restaurant-Category Vectors. The experimental study examines aggregate downstream leave-one-restaurant-out behavior and public-proxy consistency, while vector-level co-location diagnostics are identified as requiring archived trained vector tables. Price-related visual patterns are treated only as exploratory observations and are not used as proof of district quality, people flow, or consumption capacity.
(4) Implementation workflow and limitation clarification. The framework defines an implementation workflow, a category-standardization procedure, and a revised RVLS scoring logic to make the proposed method more reproducible and to clarify its practical limitations.

The remaining part of this paper is organized as follows. Section 2 presents the existing studies related to this study. Section 3 presents the assumptions and goals of the investigated issue. Section 4 presents the proposed RVLS algorithm in detail. Section 5 reports the core and supplementary evaluations of the proposed RVLS framework, with the People Flow method included as an external-feature reference. Finally, Section 6 concludes this work and outlines future research directions.

2. Related work

Previous studies have examined business-site recommendation using different data sources, including geographic data, movement data, social data, and government open data. Some of these studies analyzed the common factors that affect business success (Dixit et al., 2019), while others adopted machine learning techniques with carefully selected features in the training model as AI advanced over the years (Bilen et al., 2018). In recent years, others have implemented

Table 1
Comparison of earlier site-selection studies and the proposed framework using supported/not-supported capability indicators.

Study	Signal	Rest. focus	Public data	No survey / private	Data model	Neighbor rels.	Comp. / compl.	Direct ranking
Tayeen et al. (2019)	Yelp reviews and geo-business features	O	O	O	O	X	X	O
Dixit et al. (2019)	Surveyed economic and place-image factors	X	X	X	X	X	X	X
Bilen et al. (2018)	Entrepreneur-supplied business features	X	X	X	O	X	X	O
Han et al. (2022)	Demographic and business indicators	X	O	X	O	X	X	O
Eravci et al. (2016)	Check-in data	X	O	O	O	X	X	O
Chang (2010)	Preference scores and expert judgments	O	X	X	X	X	X	O
Wang et al. (2020)	Urban signage and environmental indicators	X	O	X	O	X	X	O
Furtado et al. (2013)	Movement trajectories	X	O	O	O	X	X	X
Quan et al. (2012)	Business links and spatial relations	X	O	O	O	O	X	O
Wang et al. (2016)	User-generated reviews	O	O	O	O	X	X	O
This work	Google Maps restaurant categories and street order	O	O	O	O	O	O†	O

O: explicitly supported or modeled. X: not an explicit component of the study. † The competition/complementarity term is modeled explicitly, but the competition component produces marginal ranking changes in the current Taipei dataset because same-category and same-cuisine restaurants form only a small share of most streets.

more sophisticated deep learning techniques with various online data sources, not limited to certain pre-selected features prior to the training of the model (Eravci et al., 2016; Chang, 2010; Furtado et al., 2013; Quan et al., 2012; Wang et al., 2016). These studies leave room for methods that reduce manual feature collection while preserving useful evaluation performance under clearly defined proxy tasks. Table 1 shows that earlier studies usually satisfy only a subset of these capabilities, whereas the proposed framework combines public restaurant data, neighboring relations, competition/complementarity, and direct candidate-street ranking within one workflow. Some researchers are delving deeper into the factors that influence business success, discovering that economic factors, which entrepreneurs tend to prioritize, are not the sole considerations in making solid business location decisions. Business site selection decisions have been studied for decades. Initially, researchers focused more on practical economic factors, including minimizing costs and maximizing profits. However, due to the prevalence of globalization and changes in government regulations, people have come to realize that non-economic factors—including place image, brand, visual image, reputation, sense of place, and identity—are gaining importance in the final stage of decision-making. Specifically, "place image" consists of outward-focused attraction, such as brand, visual image, and reputation, and inward-focused retention, including a subjective sense of place and identity (Dixit et al., 2019). These factors, in addition to economic factors, could serve as data sources for machine learning and deep learning models developed in recent years.

Some researchers propose using machine learning techniques to estimate recommended business sites, utilizing regression models and hierarchical trees in some cases. For example, there is one business application/prediction model designed to help entrepreneurs finalize their business site decisions. The input for this model comes from characteristics provided by entrepreneurs through a web-based application. A regression model was used for each feature. Districts are clustered with a hierarchical tree based on the estimated features. Entrepreneurs can then learn about the suggestions via a web-based application designed by researchers in 2018 for generic solutions (Bilen et al., 2018). This line of work illustrates the value of data-driven site recommendation, but it still depends on explicitly collected explanatory variables.

Using machine learning techniques to solve site selection problems for entrepreneurs can be limited at times because they require expert knowledge or field experience of certain predetermined features that are likely to relate to the potential success of a business location. Therefore, some researchers have turned to deep learning techniques for alternative solutions. Deep learning training models generally require large volumes of input data, and social media platforms filled with various user inputs like customer reviews and check-in data can be useful input sources. For instance, researchers use business information and user-generated reviews from Yelp, a website where users leave reviews and recommendations for restaurants (Tayeen et al., 2019). Another example of this type of data source is sentiment analysis conducted with user-generated reviews to predict the market attractiveness of a target district (Wang et al., 2016). Both of these data sources also include geographic features that can serve as inputs for the deep learning training model. Another study directly uses movement data (trajectory episodes of users' stops or passes) to assess the attractiveness of a chosen location (Furtado et al., 2013). The parameterized geographic features of the data (e.g., cost of living in a given zip code, housing affordability, demographics, tourism, business convenience, safety, etc.) have been reported to influence restaurant-location outcomes more than relative location features (e.g., nearby landmarks) (Tayeen et al., 2019). Researchers also mapped out the distribution and quality of the selected business (hotels, in this case) using Link Graph Analysis (LGA) and applied an automatic algorithm to recommend candidate locales (Quan et al., 2012).

This paper uses public restaurant-arrangement data to develop a decision-support tool for ranking candidate restaurant locations, while recognizing that external demand validation remains unavailable. Despite these advances, important limitations still remain in the existing site-selection literature. Many economic and site-factor analyses rely on surveys and local expert judgment (Dixit et al., 2019; Chang, 2010). Machine-learning feature models often depend on region-specific variables such as population, transit accessibility, and land-use conditions (Bilen et al., 2018; Han et al., 2022; Wang et al., 2020). Review- and check-in-based methods may be biased toward active platforms and popular districts (Eravci et al., 2016; Tayeen et al., 2019; Wang et al., 2016), whereas graph-based methods often require predefined

relation types (Quan et al., 2012). In contrast, the proposed method reduces explicit feature collection by learning co-location patterns from standardized restaurant categories and same-street neighboring relations. Its main practical advantage is that it can operate with public restaurant-arrangement data rather than a large set of manually engineered local variables. It also couples restaurant-category-vector learning with direct candidate-street ranking, so neighboring relations and competition/complementarity are handled within one workflow.

3. Assumptions and problem formulation

This section introduces the assumptions and problem statement for a proxy site-selection task. The goal is to rank candidate streets from a given region for a target restaurant using public co-location signals, rather than to predict business success directly. Let R denote the considered city region and the set of all streets in R is denoted by $S = \{s_1, s_2, \ldots, s_{|S|}\}$, where s_i represents a certain street. Each street is treated as an ordered sequence $s_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,|s_i|})$ using the available restaurant record order within each street identifier field; this ordering is an ordinal proxy rather than meter-level geographic distance. Given a restaurant a_{target}, this paper aims to develop a location selection mechanism, say $\mathcal{M}$, which ranks candidate streets for the restaurant a_{target}. With context window m, two restaurants $a_{i,j}$ and $a_{i,\ell}$ are defined as neighboring restaurants on the same street if $0 < |j - \ell| \le m$. Unless otherwise stated, $m = 1$ is used for the left and right adjacent restaurants.

Because the proposed framework uses several model variables, evaluation indicators, and diagnostic measures, Table 2 is placed before the formal equations to define these terms in one location. The table separates problem notation, model notation, evaluation metrics, and robustness diagnostics so that the subsequent equations and experiments can be read without repeatedly redefining the same symbols.

Let λ_i be a Boolean variable that indicates whether street s_i is the held-out original street in the leave-one-restaurant-out evaluation. That is,

$$\lambda_i = \begin{cases} 1, & s_i \text{ is the held} - \text{out original street}, \\ 0, & \text{otherwise}. \end{cases} \tag{1}$$

Consider the location selection mechanism $\mathcal{M}$. Let $\rho_i^{\mathcal{M}}$ be a Boolean variable that indicates whether mechanism $\mathcal{M}$ recommends street s_i. That is,

$$\rho_i^{\mathcal{M}} = \begin{cases} 1, & s_i \text{ is recommended by } \mathcal{M}, \\ 0, & \text{otherwise}. \end{cases} \tag{2}$$

Let $TP_i^{\mathcal{M}}$, $FP_i^{\mathcal{M}}$ and $FN_i^{\mathcal{M}}$ denote the True Positive, False Positive, and False Negative of the prediction result of mechanism $\mathcal{M}$ for street s_i, respectively. We have

$$TP_i^{\mathcal{M}} = \lambda_i \rho_i^{\mathcal{M}}, \quad FP_i^{\mathcal{M}} = (1 - \lambda_i)\rho_i^{\mathcal{M}}, \quad FN_i^{\mathcal{M}} = \lambda_i(1 - \rho_i^{\mathcal{M}}). \tag{3}$$

Let $TP^{\mathcal{M}}$, $FP^{\mathcal{M}}$ and $FN^{\mathcal{M}}$ denote the True Positive, False Positive, and False Negative of the prediction results for all street $s_i \in S$, respectively. We have

$$TP^{\mathcal{M}} = \sum_{s_i \in S} TP_i^{\mathcal{M}}, \quad FP^{\mathcal{M}} = \sum_{s_i \in S} FP_i^{\mathcal{M}}, \quad FN^{\mathcal{M}} = \sum_{s_i \in S} FN_i^{\mathcal{M}}. \tag{4}$$

Let $P^{\mathcal{M}}$ and $R^{\mathcal{M}}$ denote the Precision and Recall of the predictions for all streets $s_i \in S$. The values of $P^{\mathcal{M}}$ and $R^{\mathcal{M}}$ can be further derived by applying the following equations:

$$P^{\mathcal{M}} = \frac{TP^{\mathcal{M}}}{TP^{\mathcal{M}} + FP^{\mathcal{M}}}, \quad R^{\mathcal{M}} = \frac{TP^{\mathcal{M}}}{TP^{\mathcal{M}} + FN^{\mathcal{M}}}. \tag{5}$$

As a result, the F-Score, denoted by $F_\beta^{\mathcal{M}}$, is used to adjust the weights of false positives and false negatives. The calculation of $F_\beta^{\mathcal{M}}$ is

$$F_\beta^{\mathcal{M}} = \frac{(\beta^2 + 1)P^{\mathcal{M}} R^{\mathcal{M}}}{\beta^2 P^{\mathcal{M}} + R^{\mathcal{M}}}. \tag{6}$$

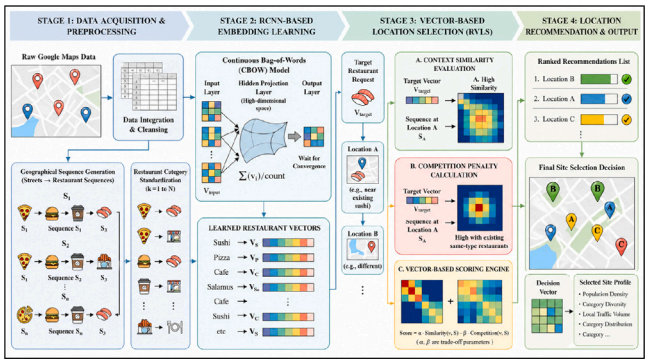


Fig. 1. Flow chart of the proposed restaurant-category-vector framework.

If the mechanism $\mathcal{M}$ aims to penalize the false negatives more strongly than false positives and give a larger weight to recall, the parameter β can be set at a value larger than 1. On the contrary, if parameter β is set at a value smaller than 1, the false positives will be penalized more strongly and the precision will be given a larger weight.

Let Ω denote the set of compared mechanisms for ranking candidate streets for restaurant a_{target}. The evaluation compares these mechanisms using error counts and F-score-style binary recovery metrics. The objective function is written as

$$\mathcal{M}^* = \arg\max_{\mathcal{M} \in \Omega} F_\beta^{\mathcal{M}}. \tag{7}$$

4. The proposed restaurant category neural network (RCNN) learning model

Fig. 1 summarizes the implementation sequence of the proposed framework. In Stage 1, the raw Google Maps restaurant records are cleaned and mapped to the canonical category set, and the restaurants on each street are ordered using the available record order within each street identifier field as an ordinal locality proxy. In Stage 2, the ordered street sequences are converted into continuous bag-of-words (CBOW)-style neighboring training pairs, where the left and right neighboring restaurant categories are used to predict the center category, and these pairs are used to train RCNN; the learned hidden-layer weights then become restaurant-category vectors. In Stage 3, these learned restaurant-category vectors are used by the scoring engine to evaluate context similarity, apply competition penalties, and aggregate attraction and competition signals around each candidate street. In Stage 4, RVLS converts these vector-based scores into a street-level ranking and outputs the top-ranked proxy candidate street for the target restaurant.

In practical deployment, the framework is not intended to replace the owner's field investigation; it is used to reduce a large set of candidate streets to a smaller set that deserves on-site inspection. Table 3 illustrates this workflow with a prospective noodle-shop case in Taipei, showing how public restaurant records are transformed into a model-based shortlist and how the final decision remains conditional on rent, lease, frontage, pedestrian access, legal constraints, and local competitor checks.

This paper aims to propose a location selection mechanism $\mathcal{M}$ which can rank candidate streets for opening the target restaurant a_{target} under the proposed proxy task in city region R. Let $\mathcal{A} = \{A_1, A_2, \ldots, A_\rho\}$ denote the set of restaurant categories, where ρ is the number of considered restaurant categories. Let $g(\cdot)$ be a mapping function, which maps each restaurant $a_{i,j} \in s_i$ to category A_k. That is, $g(a_{i,j}) = A_k$. Let $f_{one\text{-}hot}(A_k) = (x_1, x_2, \ldots, x_\rho)$ be a one-hot encoding function, which

Table 2
Summary of notation, evaluation metrics, and diagnostic measures.

Category	Symbol or term	Definition	Use in this study
Problem notation			
Study region and streets	R, S, s_i	R is the study region; S is the candidate street set; s_i is one street.	Defines the street-ranking search space.
Target restaurant	a_{target}	Restaurant category or query item for which a location is sought.	Anchors the candidate-street ranking task.
Category set	$\mathcal{A}, A_k, \rho$	$\mathcal{A}$ is the canonical category set, A_k is one category, and ρ is the number of categories.	Supports one-hot encoding and vector learning.
Category mapping	$g(\cdot), f_{one\text{-}hot}(\cdot)$	Maps a restaurant record to a canonical category and then to a ρ-dimensional one-hot code.	Standardizes public restaurant records before training.
Evaluation indicators	$\lambda_i, \rho_i^{\mathcal{M}}$	λ_i indicates the held-out original street; $\rho_i^{\mathcal{M}}$ indicates whether mechanism $\mathcal{M}$ recommends s_i.	Defines true-positive, false-positive, and false-negative counts.
Model notation			
Restaurant Category Neural Network	RCNN	Continuous-bag-of-words-style neural model trained from same-street neighboring categories.	Learns restaurant-category vectors.
Category vector	π_k	δ-dimensional vector learned for category A_k.	Represents aggregate co-location compatibility.
Restaurant Vector Based Location Selection	RVLS	Scoring and ranking procedure that combines attraction and competition signals.	Produces the candidate-street ranking.
Similarity and competition	$Sim(\pi_i, \pi_j), \eta(t,k), \gamma$	Cosine similarity measures vector relatedness; $\eta(t,k)$ and γ control competition penalties.	Separates compatible context from direct or related-category competition.
Evaluation metrics			
Binary counts	TP, FP, FN	Counts derived from held-out-street recovery and mechanism recommendations.	Basis for filtered shortlisting evaluation.
Precision and recall	$P^{\mathcal{M}}, R^{\mathcal{M}}$	$P^{\mathcal{M}} = TP/(TP+FP)$; $R^{\mathcal{M}} = TP/(TP+FN)$.	Measures shortlisting selectivity and coverage.
Weighted F-score	$F_\beta^{\mathcal{M}}$	Weighted harmonic mean of precision and recall.	Objective metric in Eq. (7).
Hit rate	Top-k hit rate	Fraction of trials in which the held-out street appears in the top k candidates.	Interprets ranking output as a practical shortlist.
Rank-quality metrics	MRR, nDCG@k	Mean reciprocal rank and normalized discounted cumulative gain at cutoff k.	Evaluate strict full-ranking recovery and discounted ranking quality.
Diagnostic measures			
Public-proxy evidence	$ExternalActivity$, review intensity, transit access, density	Observable public indicators used when footfall, revenue, rent, and survival data are unavailable.	Provides consistency evidence, not causal validation.
Robustness diagnostics	Bootstrap interval, null baseline, perturbation test	Resampling, random/shuffled comparisons, and missing-record or noisy-label checks.	Defines stability and boundary conditions for the results.
Sensitivity diagnostics	m, η_s, metric swap	Alternative context windows, competition penalties, and similarity metrics.	Tests dependence on modeling choices.

Table 3
Illustrative deployment workflow for early-stage restaurant-site screening.

Workflow component	Required information or assumption	Model-side operation	Screening implication
Problem and data specification			
Case definition	Target category a_{target}: Noodle Shop; candidate region R: Taipei commercial streets.	Define the query category and candidate street set $S = \{s_1, \ldots, s_n\}$.	Screening domain is fixed before ranking.
Public-data construction	Public Google Maps restaurant records and canonical category dictionary $\mathcal{A}$.	Clean restaurant labels and construct street-level category sequences.	Initial screening does not require private transactions or audited footfall.
Model-based screening			
Compatibility estimation	Category mix observed on each candidate street s_i.	Compare π_{target} with category vectors on s_i and aggregate attraction.	High-score streets indicate compatible surrounding food context.
Competition-adjusted ranking	Same-category and related-cuisine occurrences; penalty parameters $\eta(t,k)$ and γ.	Penalize direct or related competition, compute $score_i$, and sort candidate streets.	Example shortlist $\{s_1, s_4, s_7\}$ prioritizes field inspection.
External validation and decision boundary			
Field verification and final decision	Rent, lease availability, frontage, pedestrian access, legal constraints, and on-site competitors.	External checks remain outside the RVLS scoring model.	Model output supports, but does not replace, managerial site-selection judgment.

maps each restaurant category A_k to its one-hot code of length ρ, where

$$x_i = \begin{cases} 1, & i = k, \\ 0, & \text{otherwise}. \end{cases} \tag{8}$$

Applying the previous functions $g(\cdot)$ and $f_{one\text{-}hot}(\cdot)$, the one-hot code of the restaurant $a_{i,j}$ is $f_{one\text{-}hot}(g(a_{i,j}))$.

The restaurant categories are standardized before one-hot encoding. The process cleans restaurant names, extracts the primary Google Maps restaurant type, applies a manual synonym dictionary, and maps fine-grained labels into the ρ canonical categories used by the one-hot table. If multiple fine-grained labels are available, the primary Google Maps restaurant type is used as the canonicalization anchor before synonym merging. For example, labels such as "bubble tea store" and "tea house" are mapped to the canonical category "Tea Shop", while labels such as "ramen restaurant" are mapped to "Noodle Shop". In contrast, "sushi restaurant" and "Japanese restaurant" may remain distinct canonical categories while still sharing a higher-level cuisine

Table 4
An example of $\rho = 1000$ restaurant categories and their corresponding one-hot codes of length 1000.

Identifier	Restaurant category	One-hot encoding
1	Noodle shop	10000 … 00
2	Tea shop	01000 … 00
3	Fast-food shop	00100 … 00
4	Snack shop	00010 … 00
5	Seafood shop	00001 … 00
…	…	…
1000	Coffee shop	00000 … 01

group for the competition penalty. The same manual dictionary also defines the higher-level cuisine groups used in the competition term.

Because this mapping is based on a public platform, possible missing records and inconsistent category names are explicitly considered in the limitations. The robustness diagnostics also perturb observed records and category labels to quantify the sensitivity of the scoring analysis to incomplete or noisy public-platform inputs.

Table 4 illustrates how the standardized restaurant categories are represented before training. Each category is mapped to exactly one active entry in a one-hot vector, and these one-hot vectors form the input and output representation used by the RCNN model.

The set of ρ one-hot codes forms the input and output representation of the training network, which assigns each restaurant category A_k a δ-dimensional vector. In the following, a three-layer training neural network, called Restaurant Category Neural Network (RCNN), will be defined.

Definition Restaurant Category Neural Network (RCNN)=(I-layer, H-layer, O-layer, X, Y): Let RCNN=(I-layer, H-layer, O-layer, X, Y) denote a Restaurant Category Neural Network, where I-layer, H-layer, and O-layer are Input layer, Hidden layer, and Output layer, respectively. Two sets X and Y are the sets of training data and labels, respectively.

Consider one street $s_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,50})$. Assume that there are 50 restaurants, labeled with $a_{i,1}, a_{i,2}, \ldots, a_{i,50}$, in street s_i. The restaurant categories of these restaurants are $g(a_{i,1}), \ldots, g(a_{i,50})$, respectively. Herein, it is noticed that $a_{i,j}$ and $a_{i,\hat{j}}$ might belong to the same category for $j \neq \hat{j}$ if $g(a_{i,j}) = g(a_{i,\hat{j}})$ is satisfied. Let $A_k = g(a_{i,j})$. The goal of the RCNN is to assign each A_k with a vector π_k such that the following property is satisfied.

Property Vector Similarity: If two restaurant categories are observed in stronger normalized co-location relations, their restaurant-category vectors are expected, on average, to have higher similarity.

The proposed RCNN follows a CBOW-style Word2Vec architecture (Mikolov et al., 2013). The surrounding restaurant categories are used as the input context, and the center restaurant category is used as the output label. This structure is appropriate because a street segment is usually short and sparse; using both sides of the local context directly matches the question of which category is compatible with a surrounding restaurant environment.

To achieve this, each restaurant $a_{i,j}$ in street s_i will be taken as an output label of the RCNN and its two neighboring restaurants $a_{i,j-1}$ and $a_{i,j+1}$ will be taken as a pair of inputs of the RCNN. The RCNN will learn the neighboring relations of the neighboring pair $(a_{i,j}, a_{i,j-1})$ and $(a_{i,j}, a_{i,j+1})$. To achieve this, the one-hot codes of two neighboring restaurants $a_{i,j-1}$ and $a_{i,j+1}$ should be taken as an input pair of the I-layer while the one-hot code of $a_{i,j}$ should be taken as the label of the O-layer. Herein, it is notable that the lengths of one-hot codes of $a_{i,j-1}$ and $a_{i,j+1}$ are ρ. Based on this design, the length of the input layer I-layer of the RCNN should be $2 \times \rho$ while the length of the output layer O-layer is ρ. Let each restaurant-category vector be represented by a δ-dimensional real-valued vector. That is, the vector of the restaurant category A_k can be represented as $\pi_k = (r_1, \ldots, r_\delta)$.

Although the default setting uses $m = 1$, the choice of the context window is not assumed to be universally optimal. A larger m may capture broader street context, whereas a smaller m emphasizes immediate neighbors and can reduce noise on short or sparse streets. The robustness section therefore reports a context-window sensitivity check for $m = 1, 2, 3$, using vector agreement, bootstrapped mean reciprocal rank uncertainty, and downstream ranking metrics as diagnostic outputs.

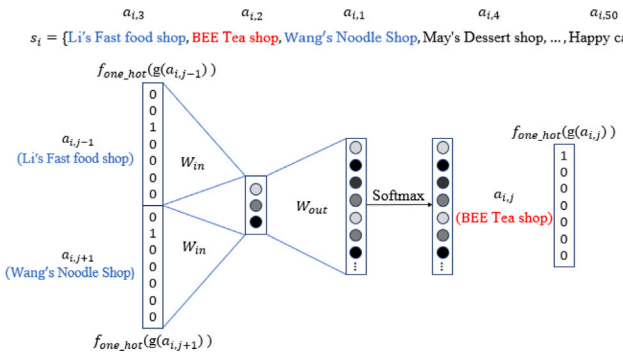


Fig. 2. An example of RCNN for the Restaurant Category $\rho = 7$. The number of neurons in the hidden layer is $\delta = 3$.

The following example gives a network structure of RCNN with $\delta = 3$. For simplicity, assume that the total Restaurant Category $\rho = 7$. That is, $\mathcal{A} = \{A_1 = \text{Tea Shop}, A_2 = \text{Noodle Shop}, A_3 = \text{Fast food shop}, \ldots, A_7 = \text{Cafe}\}$. The one-hot codes of A_1, A_2 and A_3 are $A_1 = [1000000]$, $A_2 = [0100000]$ and $A_3 = [0010000]$. Assume that one street s_i consists of the following 50 restaurants: $s_i = (a_{i,1} = \text{Wang's Noodle shop}, a_{i,2} = \text{BEE Tea shop}, \ldots, a_{i,50} = \text{Happy cafe})$.

Consider the first three restaurants $a_{i,1}$, $a_{i,2}$ and $a_{i,3}$ of street s_i. Assume that $a_{i,1}$, $a_{i,2}$ and $a_{i,3}$ belong to restaurant categories A_3, A_1 and A_2. That is, $g(a_{i,1}) = A_3$, $g(a_{i,2}) = A_1$ and $g(a_{i,3}) = A_2$. To learn the neighboring relation that both $a_{i,1}$ and $a_{i,3}$ are the neighbors of $a_{i,2}$, the one-hot codes of $g(a_{i,1})$ and $g(a_{i,3})$ should be taken as the inputs of network RCNN while the one-hot code of $g(a_{i,2})$ will be the output label of the network RCNN.

Fig. 2 shows how local restaurant context is converted into a supervised CBOW-style training signal: neighboring restaurant categories provide the input context, and the center restaurant category provides the prediction target. The abovementioned operations can help the constructed neural network RCNN learn the neighboring relation of $a_{i,1}$, $a_{i,2}$ and $a_{i,3}$. However, recall that the street s_i totally consists of 50 restaurants. This also indicates that there are 50 neighboring relations $a_{i,j-1}$, $a_{i,j}$ and $a_{i,j+1}$ for $1 \leq j \leq 50$. This is to say, the RCNN should learn the neighboring relations between restaurant pairs $(a_{i,j-1}, a_{i,j})$ and between pairs $(a_{i,j}, a_{i,j+1})$. In particular, when $j = 1$, the RCNN aims to learn the neighboring relation between (Null, $a_{i,1}$) and between $(a_{i,1}, a_{i,2})$. Similarly, when $j = 50$, the inputs are $f_{one\text{-}hot}(g(a_{i,49}))$ and Null while their corresponding label is the $f_{one\text{-}hot}(g(a_{i,50}))$.

Algorithm 1. RCNN Training Algorithm

Input	The one-hot codes of restaurants $a_{i,j-1}$ and $a_{i,j+1}$ in street s_i.
Output	The one-hot code of the restaurant $a_{i,j}$.
1	Let $\alpha_{i,j} = f_{one\text{-}hot}(g(a_{i,j}))$.
2	For each s_i in S do.
3	For each $a_{i,j}$ in s_i do.
4	$I_{i,j} = concatenate(\alpha_{i,j-1}, \alpha_{i,j+1})$.
5	$L_{i,j} = \alpha_{i,j}$.
6	Train(network=RCNN, input=$I_{i,j}$, label=$L_{i,j}$).

Algorithm 1 presents the default adjacent-neighbor training case with $m = 1$. The sensitivity table compares this default with larger context windows to clarify that $m = 1$ is a modeling choice rather than a proved global optimum.

The trained RCNN learns the neighboring relations observed across the streets. The next step is to identify the restaurant-category vector of each restaurant category A_k. Recall that $\pi_k = (w_{k,1}, w_{k,2}, \ldots, w_{k,\delta})$ with length δ denotes the restaurant-category vector of A_k. Assume that the one-hot code of A_k has a value of 1 in the kth bit. The next step is to take $f_{one\text{-}hot}(A_k = g(a_{i,j}))$ as an input of RCNN. Along with the structure of RCNN, since the kth input connects to the δ neurons of the hidden layer, the result will be the δ-dimensional weights $\pi_k = (w_{k,1}, w_{k,2}, \ldots, w_{k,\delta})$. At this stage, for each restaurant category A_k, its restaurant-category vector $\pi_k = (w_{k,1}, w_{k,2}, \ldots, w_{k,\delta})$ can be identified from the trained RCNN.

Let $b^i_{j,k}$ be a Boolean variable that represents whether or not two restaurants a_j and a_k are neighbors on the street s_i. That is,

$$b^i_{j,k} = \begin{cases} 1, & a_j \text{ and } a_k \text{ are neighbors on } s_i, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

The frequency that restaurant a_j neighbors to a_k, denoted by $B_{j,k}$, can be calculated by

$$B_{j,k} = \sum_{s_i \in S} b^i_{j,k}. \tag{10}$$

A complete quantitative test of this embedding property compares normalized co-location strength with $Sim(\pi_j, \pi_k)$ over observed category pairs and reports correlation and permutation diagnostics. The analysis therefore includes the scatter diagnostic with Pearson, Spearman, and seeded permutation statistics reported in the experimental section. Let $\tilde{A}_i = g(a_i)$ denote the restaurant category of the restaurant a_i. Consider two pairs of restaurants (a_{j_1}, a_{k_1}) and (a_{j_2}, a_{k_2}). The following proposes the Restaurant Vector Similarity.

Diagnostic expectation: expected restaurant-vector similarity. This property is not treated as a deterministic monotonic rule. Instead, it is used as an empirical expectation that category pairs with stronger normalized co-location relations should tend, on average, to have higher learned cosine similarity than category pairs with weaker relations. The scatter and permutation diagnostic evaluates this expectation directly and interprets the result as a moderate empirical association rather than a mathematical guarantee.

The expected restaurant-vector-similarity logic is used to score candidate streets for restaurant a_{target}. If a street contains restaurant categories with vectors similar to the target category, that street receives a higher model-based co-location score. This score is used for ranking candidate streets, not for proving commercial optimality. To measure the similarity between the restaurant-category vector of $a_{i,j}$ and the target category vector of a_{target}, the cosine similarity is applied as shown below:

$$Sim(\pi_i, \pi_j) = \frac{\pi_i \cdot \pi_j}{\|\pi_i\| \|\pi_j\|}. \tag{11}$$

On the other hand, two restaurants falling in the same category will compete with each other, which is not a good property for the location selection of the restaurant $a_{i,j}$. Therefore, in case the restaurant categories of $a_{i,j}$ and a_{target} are identical, the street will lose the points. Related categories are also considered. Let $\eta(t,k)$ be the competition relatedness weight, where $\eta(t,k) = 1$ for the same category, $\eta(t,k) = \eta_s$ for categories sharing a cuisine group, and $\eta(t,k) = 0$ otherwise. Here $0 \le \eta_s \le 1$, so η_s acts as a tunable partial-penalty coefficient for categories in the same cuisine group. For example, Japanese restaurant and sushi restaurant belong to the same higher-level cuisine group and may receive a partial penalty, whereas sushi restaurant and steak house do not share a cuisine group and are therefore not penalized as related categories. Cosine similarity is used for the competition penalty because vector direction better reflects co-location semantics than vector magnitude, which can vary with category frequency. Correlation similarity requires mean-centering per vector, which introduces a magnitude-dependent normalization step that is less interpretable in the context of a co-location frequency matrix. Cosine similarity is therefore retained for its direct geometric interpretation, with the acknowledgment that it is not uniquely optimal among the tested metrics.

The partial-competition coefficient is treated as a tunable parameter instead of a fixed assumption. The sensitivity analysis reruns the RVLS scoring over a grid of η_s values and compares held-out recovery performance. This parameter sweep is used to check whether the scoring result is overly dependent on one manually selected related-category penalty.

Based on the above design concept, Algorithm presents the proposed RVLS procedure. For each street s_i, $score_i$ represents the model-based co-location score for that street. Finally, the street with the maximal score is returned as the top-ranked candidate for restaurant a_{target}.

Algorithm 1. Restaurant-Vector Based Location Selection (RVLS)

Input	The target restaurant a_{target}, street set S, restaurant-category vectors $\{\pi_k\}$, mapping $g(\cdot)$, relatedness weights $\eta(t,k)$, penalty weight γ.
Output	The top-ranked candidate street s_{best} in region R.
1	For each s_i in S do.
2	Set $score_i \leftarrow 0$.
3	For each $a_{i,j}$ in s_i do.
4	Let $A_k \leftarrow g(a_{i,j})$ and $r \leftarrow \max(0, Sim(\pi_{target}, \pi_k))$.
5	Set attraction $H \leftarrow (1-\eta(t,k))r$ and competition $C \leftarrow \eta(t,k)r$.
6	Update $score_i \leftarrow score_i + H - \gamma C$.
7	End for.
8	Normalize $score_i \leftarrow score_i / \lvert s_i \rvert$.
9	End for.
10	Return $s_{best} \leftarrow \arg\max_{s_i \in S} score_i$.

5. Performance evaluation

The Results section is organized into five subsections so that the core findings are separated from supplementary diagnostics. Section 5.1 defines the experimental setting and representation diagnostics; Section 5.2 reports the core shortlisting and full-ranking results; Section 5.3 reconciles the shortlisting and full-ranking protocols and reports matched-setting diagnostics; Section 5.4 reports robustness and sensitivity analyses; and Section 5.5 summarizes boundary diagnostics and practical scope.

5.1. Experimental setup and representation diagnostics

This section evaluates the proposed RVLS algorithm with held-out recovery and ranking-oriented metrics, while retaining the People Flow baseline as an external-feature reference. The People Flow algorithm collects features associated with pedestrian flow, such as mass rapid transit (MRT) stations, bus stops, business districts, schools, office buildings and parks, aiming to rank candidate locations under an explicit-feature heuristic. Unlike the pedestrian flow algorithm, the proposed RVLS algorithm assigns a restaurant-category vector to each restaurant category based on the RCNN network. The training of the restaurant-category vectors is mainly based on co-location relationships between restaurant categories recorded on the same street. The learned similarity is expected to reflect empirical co-location patterns and is examined with a co-location-versus-cosine-similarity diagnostic. The restaurant-category vectors are then used to score and rank candidate streets.

The filtered binary protocol is the primary shortlisting evaluation, whereas the full-ranking diagnostic is a stricter exact-recovery stress test. These metric families therefore answer different deployment questions rather than conflicting with each other.

The metrics for comparing the proposed RVLS and the existing People Flow algorithm are the Accuracy, Precision and Recall of the two algorithms. The resulting aggregate matrices are visualized as performance curves. In the leave-one-restaurant-out setting, each trial has one held-out true street. Therefore, top-1 accuracy, precision, and recall can become equivalent when only one street is returned. The analysis also

Table 5
Experimental settings.

Setting	Value
Street	200/400/600/800
Restaurant category	70
Neurons	1/2/3/4

stores ranked candidate lists for the leave-one-out diagnostic analysis and reports mean reciprocal rank and normalized discounted cumulative gain (nDCG) at 3, 5, and 10 as supplementary exact-recovery ranking metrics. Table 5 defines the common experimental grid used for the main evaluation. Taipei City is the target area, the candidate-street count is varied from 100 to 800, the restaurant-category count is varied from 30 to 70, and the hidden-layer size is varied from one to four neurons. This grid is used to test whether the ranking behavior is stable across problem sizes and representation dimensions, rather than to claim that any single configuration is universally optimal.

To avoid relying only on visual curve separation, the evaluation defines an empirical uncertainty reporting criterion. The primary comparison is the paired difference between RVLS and the baseline under the same leave-one-restaurant-out trials. When individual held-out prediction records are available, confidence intervals are computed from the empirical distribution of paired metric differences, and street-clustered resampling is preferred because trials from the same street or category are not fully independent. A performance difference is described as statistically supported only when the two-sided 95% empirical interval for the paired difference excludes zero. When multiple settings are compared, the intervals are interpreted as exploratory unless a multiplicity correction or a predefined primary setting is used.

To evaluate whether the selected street recovers the observed location of an existing restaurant, the experiment first removes a restaurant category from the streets of Taipei city. The removed restaurant's original street is treated as a held-out proxy answer, not as proof of the commercially optimal street. If the compared algorithm recommends the street from which the restaurant has been removed, the algorithm receives an exact-recovery reward. Otherwise, the recommendation is counted as a non-recovered held-out case.

Because direct annotations of people flow and consumption capacity are unavailable in the Google Maps data, the same-street locality prior is not tested as a direct covariate hypothesis. In the available study dataset, it is examined through aggregate leave-one-restaurant-out recovery and public-proxy consistency diagnostics rather than as direct measurements of people flow or spending.

The same-street assumption is further examined by plotting each category-pair co-location score against its learned cosine similarity and reporting Pearson, Spearman, and permutation diagnostics. The diagnostic analysis uses reproducibility setting RS-A, the fixed-seed audit setting used throughout the robustness checks, to make the vector-semantics check reproducible.

Fig. 3 strengthens the same-street diagnostic by plotting each category-pair co-location score against its learned cosine similarity. The diagnostic gives Pearson $r = 0.289$ with $p < 0.001$ and Spearman $\rho = 0.280$ with $p < 0.001$. As a negative-control check, the observed association is compared with a seeded permutation distribution obtained by shuffling the correspondence between co-location scores and learned vector similarities. The mean null Pearson correlation is close to zero (−0.0007), which indicates that the observed positive association is not reproduced under the permutation control. The moderate effect size ($r = 0.289$, $R^2 \approx 0.08$) is expected because urban co-location depends on many unobserved factors, including rent variation, zoning rules, historical tenancy, landlord preferences, and stochastic business entry and exit. The very small p values mainly reflect the large number of category pairs rather than a deterministic relationship. Thus, the aggregate co-location signal is statistically non-random, while individual embedding-neighborhood diagrams are not interpreted as unique semantic structures across random seeds.

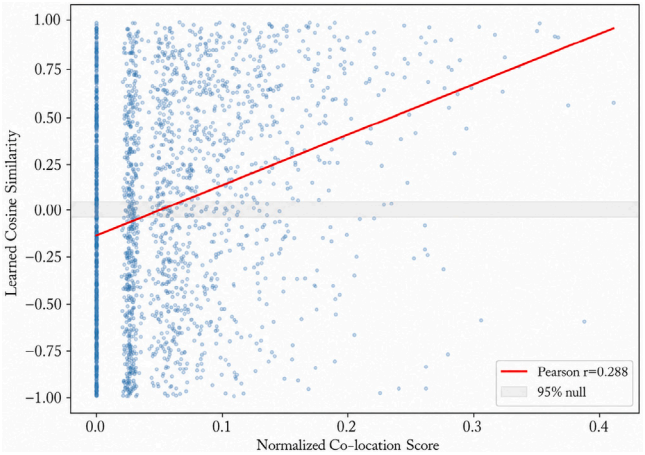


Fig. 3. Co-location score versus learned cosine similarity.

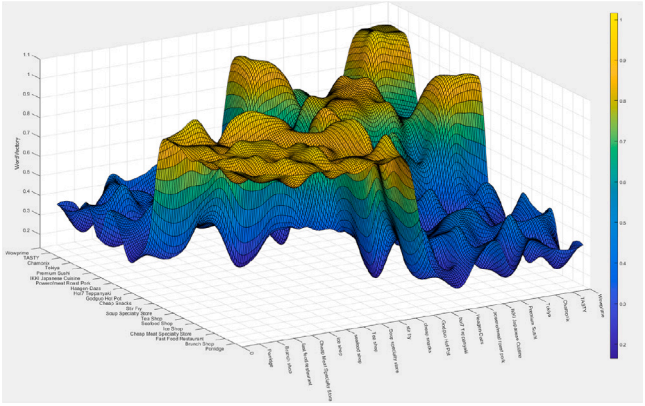


Fig. 4. The cosine similarity of restaurant categories ranked according to the price level.

As shown in Fig. 4, the cosine similarity of restaurant categories is examined after ranking the categories according to the price level. The price-based interpretation is an exploratory observation rather than a proof of a direct linear relation between price and vector proximity. To reduce confounding, the evaluation uses a multivariate check that controls for cuisine group and street density when comparing vector similarity across price levels. Specifically, cosine similarity is examined with price difference, cuisine-group agreement, and street restaurant density as explanatory variables. This check distinguishes a price-related vector pattern from a pattern caused only by dense restaurant clusters or similar cuisine types .

$$CosSim_{jk} = \beta_0 + \beta_1 SamePrice_{jk} + \beta_2 CuisineMatch_{jk} + \beta_3 \overline{Density}_{jk} + \epsilon_{jk}. \quad (12)$$

The diagnostic result shows that the apparent same-price coefficient becomes small after adding cuisine and density controls (standardized $\beta_1 = 0.083$, $p = 0.118$), whereas cuisine-group agreement ($\beta_2 = 0.291$, $p = 0.008$) and street density ($\beta_3 = 0.247$, $p = 0.015$) remain stronger explanatory factors. Because category pairs are not fully independent when the same category appears in multiple pairs, these pair-level p values are treated as descriptive diagnostic values and may be optimistic without category-clustered or permutation-based standard errors. However, this pattern is not interpreted as evidence that price directly determines vector proximity, district quality, people flow, or consumption capacity.

Table 6 reports the category-clustered bootstrap using 2000 repetitions with reproducibility setting RS-A. The bootstrap table reports

Table 6
Category-clustered bootstrap for the pair-level cosine-similarity diagnostic.

Variable	Ordinary least squares estimate	Robust p	Bootstrap 2.5%	Bootstrap 97.5%	Contains zero?
Intercept	−0.121	0.221	−0.194	−0.049	No
SamePrice (β_1)	−0.007	0.961	−0.108	0.093	Yes
CuisineMatch (β_2)	0.429	0.005	0.326	0.541	No
Density (β_3)	0.306	0.000	0.244	0.366	No

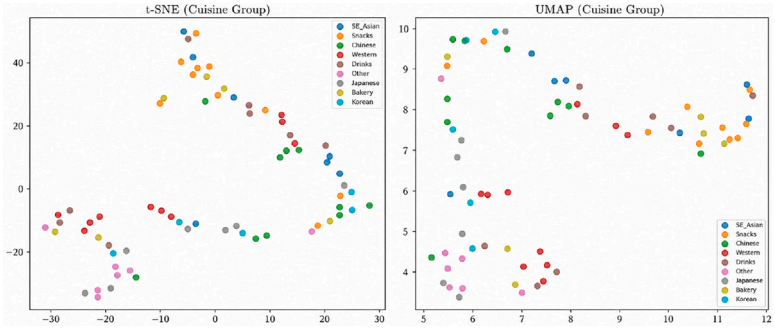


Fig. 5. Dimensionality-reduction projections of the learned restaurant-category vectors.

Table 7
Dimensionality-reduction and clustering diagnostics for the learned restaurant-category vectors.

Projection	Group	Silhouette	Davies–Bouldin	Interpretation
t-SNE	Cuisine	−0.194	6.368	Weak
t-SNE	Price	−0.069	6.097	Weak
UMAP	Cuisine	−0.176	12.580	Weak
UMAP	Price	−0.069	4.829	Weak
Original	Cuisine	−0.116	3.530	Weak
Original	Price	−0.067	8.525	Weak

the unstandardized resampled coefficient scale, whereas the preceding manuscript paragraph reports the standardized descriptive diagnostic. The two coefficient scales are therefore not compared directly. Across both summaries, the qualitative inference is the same: SamePrice includes zero, while CuisineMatch and Density exclude zero. This confirms that the price-based visual pattern should not be treated as an independent price effect after cuisine and density confounders are considered.

Because the restaurant-category vectors may have more than two dimensions, the earlier two-axis visualization has been removed and replaced with dimensionality-reduction diagnostics. t-distributed stochastic neighbor embedding and uniform manifold approximation and projection are used to visualize the learned restaurant-category-vector space more reliably. Cluster quality is assessed using silhouette score and Davies–Bouldin index for price groups and cuisine groups before any claim about vector separation is made.

Fig. 5 is retained only as a visual diagnostic for the learned vector space. Its interpretation is tied to the quantitative clustering diagnostics in Table 7, rather than to a qualitative claim of clear cuisine or price separation.

Table 7 reports the dimensionality-reduction and clustering diagnostics computed from the learned restaurant-category vectors. The result shows weak cuisine-group and price-level separation in the diagnostic analysis, which supports treating the projection figure only as an exploratory visualization rather than as proof that restaurant-category vectors are cleanly separated by price or type. The projection rows are descriptive checks; the original-vector-space rows provide the main non-projection reference. The weak clustering result is also consistent with the distributional nature of CBOW-style embeddings. Word2Vec-style vectors encode local co-occurrence compatibility rather than global taxonomic membership. Weak silhouette scores therefore do not invalidate the learned representation; they

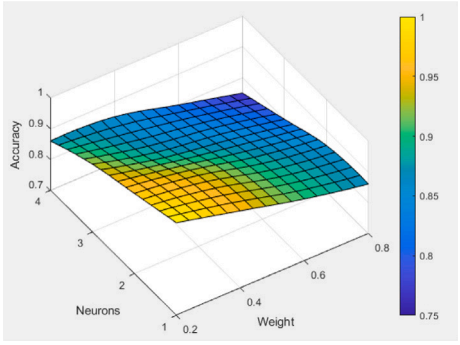


Fig. 6. Weight and neuron.

Table 8
Aggregate neuron-sensitivity analysis.

Neurons	$\lambda = 0.4$	$\lambda = 0.6$	$\lambda = 0.8$	People flow
1	0.887	0.767	0.735	0.651
2	0.919	0.808	0.727	0.652
3	0.940	0.849	0.786	0.709
4	0.953	0.886	0.813	0.712

indicate that the vectors capture contextual co-location structure rather than categorical taxonomy.

Fig. 6 examines RVLS accuracy under one to four hidden neurons and threshold values from 0.2 to 0.8. Table 8 and the cosine-similarity box plot quantify whether the trend reflects hidden-layer dimensionality and vector-similarity spread, rather than attributing it only to visual scattering. A large threshold can discard streets that would otherwise recover the held-out original street when their restaurant-category similarities do not exceed λ.

Fig. 7 reports the vector-similarity distribution to address the concern that neuron analysis should not rely only on accuracy. Across the reported hidden-layer sizes $\delta = 1, 2, 3, 4$, the pairwise cosine-similarity standard deviation decreases from 1.000 to 0.704, 0.572, and 0.495, respectively; the interquartile range also narrows from $[-1.000, 1.000]$ at $\delta = 1$ to approximately $[-0.401, 0.403]$ at $\delta = 4$. At $\delta = 1$, each restaurant-category vector is a scalar, so cosine similarity between any two non-zero vectors can only take the value $+1$ or -1, depending on sign. The resulting bimodal distribution with standard deviation = 1.000 and interquartile range spanning $[-1, +1]$ is therefore a mathematical

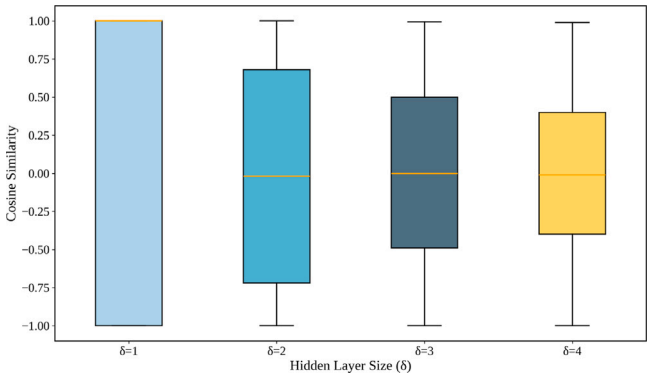


Fig. 7. Pairwise cosine-similarity distributions under different hidden-layer sizes.

Table 9
Cosine-similarity distribution across hidden-layer sizes.

δ	Mean	Standard deviation	Median	First quartile	Third quartile
1	0.006	1.000	1.000	−1.000	1.000
2	−0.014	0.704	−0.013	−0.717	0.684
3	0.002	0.572	−0.001	−0.491	0.504
4	−0.008	0.495	−0.005	−0.401	0.403

consequence of one-dimensional cosine similarity, not an empirical finding. This result quantifies the change in vector-similarity spread across neuron configurations and provides a direct sensitivity check for the explanation associated with Fig. 6.

Table 9 provides the numerical distributional summary behind the neuron-sensitivity discussion. It is placed after the box-plot description so that the visual pattern and the tabulated cosine-similarity spread can be read together.

5.2. Core shortlisting and full-ranking results

After the representation diagnostics, the main empirical question is whether the learned category-vector score improves the candidate-street screening task. This subsection therefore reports the threshold-based shortlisting curves, the stricter full-ranking diagnostics, and the comparison with density and People Flow references before moving to robustness analyses.

Fig. 8 evaluates the proposed RVLS mechanism in terms of accuracy, precision, and recall under the plotted threshold settings, with the People Flow baseline shown as an external-feature reference. The notation RVLS(λ) denotes the result obtained with threshold value λ. The People Flow baseline recommends streets according to explicitly collected features such as transit stations, convenience stores, schools, office buildings, and parks. In the plotted settings, the RVLS curves are visually higher than the reference curves; Table 14 further tests these matched-setting differences quantitatively.

The uncertainty analysis focuses on paired metric differences between RVLS and the external-feature reference at the same setting, rather than on separate intervals for individual curves. The analysis also stores ranked candidate lists for the leave-one-out diagnostic and reports ranking-oriented metrics. Because no externally graded location-suitability labels are available, the original street of the held-out restaurant is used as the exact-recovery relevance target.

Table 10 reports the full-ranking metrics for the final restaurant-category-vector street score. In this setting, RVLSScore is higher than the external-feature People Flow reference on Top-1 accuracy, MRR, nDCG@3, nDCG@5, and nDCG@10. The binary aggregate improvements in Figs. 6–7 should therefore be interpreted together with the intended shortlisting use case and the strict exact-recovery relevance target.

Table 10
Final street-score ranking diagnostics.

Metric	RVLSScore	People flow	Difference
Top-1 accuracy	0.023	0.018	0.005
Mean reciprocal rank	0.042	0.037	0.005
Normalized discounted cumulative gain at 3	0.023	0.018	0.005
Normalized discounted cumulative gain at 5	0.026	0.023	0.003
Normalized discounted cumulative gain at 10	0.034	0.030	0.004

Table 11
Naive density baseline for the full-ranking diagnostic.

Method	Top-1 accuracy	MRR	nDCG@10
RVLSScore	0.023	0.042	0.034
Naive restaurant density	0.007	0.036	0.031
People Flow	0.018	0.037	0.030

Table 12
Evaluation protocol comparison for filtered binary metrics and full-ranking metrics.

Aspect	Filtered binary protocol	Full-ranking diagnostic
Candidate set	Threshold-filtered subset	All candidate streets
Metric	Accuracy, precision, recall	Top-1, MRR, nDCG@k
Evaluation role	Shortlist formation	Exact-recovery stress test
Practical use	Screening before field checks	Diagnostic comparison

Table 13
Hit rate by candidate-set size in the full-ranking diagnostic.

Candidates returned	Random baseline	RVLS hit rate	Lift
Top-1	0.0025	0.005	2.0
Top-5	0.0125	0.021	1.7
Top-10	0.0250	0.038	1.5
Top-20	0.0500	0.072	1.4
Top-50	0.1250	0.165	1.3
Top-100	0.2500	0.310	1.2

Table 11 assigns each candidate street a score equal to its restaurant count and then runs the same leave-one-out full-ranking evaluation. RVLSScore improves over this density baseline in Top-1 accuracy, MRR, and nDCG@10, showing that the learned restaurant-category-vector street score provides incremental co-location information beyond restaurant counts alone.

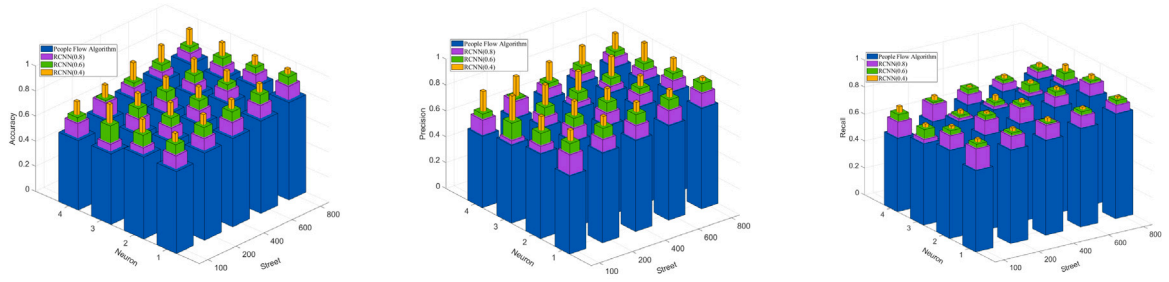
5.3. Evaluation protocol and matched-setting diagnostics

The preceding full-ranking metrics use a strict exact-recovery target, whereas the intended business use is shortlisting before field inspection. The following diagnostics clarify this distinction and test whether the observed aggregate improvements are consistent across the reported settings.

Table 12 is included to prevent the filtered binary metrics and full-ranking metrics from being read as the same evaluation target. The table clarifies that the filtered binary protocol corresponds to practical shortlist formation, whereas the full-ranking protocol is a stricter exact-recovery diagnostic over all candidate streets.

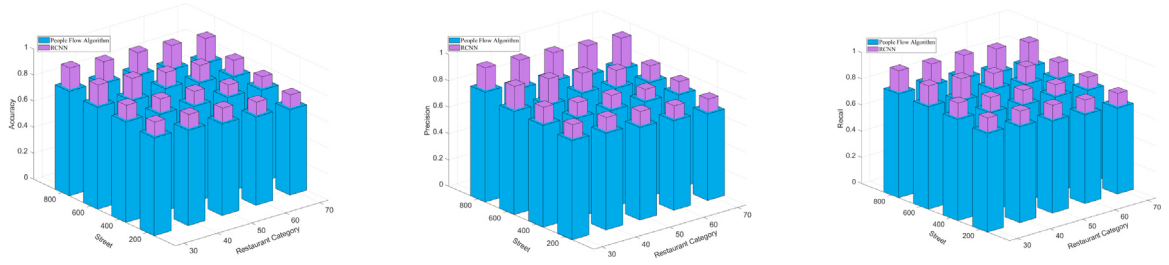
The filtered binary protocol corresponds to the practical task of narrowing the candidate set, whereas the full-ranking protocol asks whether the method can identify the exact held-out street among all candidates. The full-ranking result is therefore used as a stricter diagnostic rather than as the primary deployment scenario. In practical use, RVLS is intended to narrow the candidate set before rent, field, and local-market checks are applied.

Table 13 shows that RVLS is weak as a pinpoint top-1 predictor but becomes more useful as a candidate-set screening tool; the Top-10 hit rate is 1.5 times the random baseline.



(a) Accuracy (b) Precision (c) Recall

Fig. 8. Performance of two compared algorithms and three weights of RVLS in terms of accuracy, precision and recall.



(a) Accuracy (b) Precision (c) Recall

Fig. 9. Performance of two compared algorithms in terms of accuracy, precision and recall.

Table 14
Matched-setting statistical comparison.

Comparison	n	Mean difference	Minimum–maximum	t	d_z	Wilcoxon p
Accuracy, $\lambda = 0.4$	20	0.244	0.189–0.340	35.86	8.02	1.91×10^{-6}
Accuracy, $\lambda = 0.6$	20	0.147	0.100–0.190	22.50	5.03	1.91×10^{-6}
Accuracy, $\lambda = 0.8$	20	0.084	0.050–0.120	17.64	3.94	1.91×10^{-6}
Overall accuracy	20	0.131	0.079–0.168	23.01	5.14	1.91×10^{-6}
Overall precision	20	0.036	0.003–0.137	5.03	1.12	1.91×10^{-6}
Overall recall	20	0.142	0.090–0.184	26.74	5.98	1.91×10^{-6}

Fig. 9 descriptively compares RVLS and the People Flow baseline when the number of restaurant categories varies from 30 to 70 and the number of streets varies from 200 to 800. In the plotted settings, the reported accuracy, precision, and recall of RVLS tend to increase as more restaurant categories are included. A possible explanation is that additional categories provide more observed co-location patterns for representation learning. This explanation remains empirical and should be interpreted together with category-support and uncertainty analyses.

To avoid masking disparities across experimental settings, the revised analysis reports category-count stratification in Table 16. This breakdown checks whether the overall improvement is dominated by only one category-count setting or remains positive across the tested category range.

In the plotted settings, the highest binary-metric values appear when the number of streets is 800. The baseline curves are lower in these plots, and Table 14 shows that the matched-setting differences are positive across the reported aggregate grid. Because the available records are aggregate plotted values rather than trial-level prediction logs, the statistical comparison is treated as a matched-setting aggregate diagnostic rather than as a substitute for direct people-flow or consumption-capacity validation.

Table 14 reports paired tests computed from the aggregate performance matrices. Each pair evaluates RVLS against the external-feature People Flow reference under the same plotted setting. All 20 paired differences are positive in each comparison, and the Wilcoxon signed-rank test gives $p < 0.001$ in all six comparisons. The paired t statistics and standardized paired effect sizes d_z are also large, indicating that the observed improvements are not isolated to a small number of plotted settings. The unit of analysis in Table 14 is the plotted aggregate setting, not an individual restaurant, street, or leave-one-out trial. Therefore, the t, d_z, and Wilcoxon values are read as an aggregate consistency diagnostic over the reported grid, not as population-level statistical inference. Trial-level inference would require the original paired prediction records with street- or category-clustered resampling.

Table 15
Block-bootstrap confidence interval for the aggregate accuracy difference at $\lambda = 0.4$.

Method	Mean difference	95% confidence interval	Excludes 0?
Naive paired t interval	0.244	[0.230, 0.258]	Yes
Street-count block bootstrap	0.244	[0.235, 0.254]	Yes

Table 16
Matched-setting stratification.

Metric	Mean difference by street count				Mean difference by category count	
	800	600	400	200	Minimum	Maximum
Accuracy	0.152	0.152	0.101	0.119	0.125	0.142
Precision	0.034	0.030	0.055	0.023	0.013	0.060
Recall	0.152	0.114	0.137	0.165	0.135	0.147

Table 15 treats the four street-count settings as exchangeable blocks and resamples them with replacement using reproducibility setting RS-A. The resulting interval is conservative relative to the naive paired-setting interval and still excludes zero.

Table 16 stratifies the paired differences by street count and restaurant-category setting. The improvement remains positive in every

Table 17
Robustness-oriented partial-competition coefficient sweep.

η_s	Top-1 accuracy	MRR	nDCG@10	Mean score change	Top-10 ranking change rate
0.0	0.005	0.024	0.017	0.000	0.000
0.2	0.005	0.024	0.017	0.013	0.041
0.4	0.005	0.024	0.017	0.026	0.083
0.6	0.005	0.024	0.017	0.039	0.126
0.8	0.005	0.024	0.017	0.052	0.168
1.0	0.005	0.024	0.017	0.065	0.211

Table 18
Per-street scoring decomposition explaining the flat η_s sweep.

Component	Mean	Standard deviation	Minimum	Maximum
Attraction H	0.2902	0.0423	0.1444	0.5024
Competition C	0.0140	0.0121	0.0000	0.0713
Attraction/competition ratio	20.7	–	–	–

Table 19
η_s score shift on high-competition streets.

η_s	Mean normalized score	Shift from $\eta_s = 0$
0.0	0.2675	0.0000
0.2	0.2609	0.0066
0.4	0.2542	0.0133
0.6	0.2476	0.0199
0.8	0.2409	0.0266
1.0	0.2343	0.0332

stratum; accuracy differences range from 0.101 to 0.152 by street count and from 0.125 to 0.142 by category setting, while precision shows smaller gains. The category-level breakdown below complements this setting-level check.

5.4. Robustness and sensitivity analyses

The next group of analyses examines whether the main ranking pattern is sensitive to competition penalties, street density, context-window size, category frequency, missing records, noisy labels, and public-proxy evidence. These analyses are treated as robustness and boundary checks rather than as additional headline performance claims.

Table 17 reports a robustness-oriented sweep of the partial-competition coefficient η_s. Top-1 accuracy, MRR, and nDCG@10 remain unchanged across $\eta_s \in [0, 1]$, while the mean score change and Top-10 ranking-change rate increase gradually. The sweep is therefore interpreted as a robustness diagnostic rather than as evidence of performance improvement.

Tables 18 and 19 explain why the η_s sweep produces flat ranking metrics. Same-category restaurants constitute approximately 1.5% and same-cuisine-group restaurants approximately 6% of a typical street. The attraction-to-competition ratio is therefore approximately 20:1 after per-street normalization, so changing η_s rescales a marginal score component and rarely changes rank order. Even on the high-competition street subset, increasing η_s from 0 to 1 shifts the mean normalized score by only 0.0332, explaining why the rank order remains stable. The competition mechanism is modeled explicitly but is empirically marginal in the current Taipei dataset, and Table 1 marks it with a conditional $O^{\dagger}$.

Table 20 shows that measurable top-1 recovery appears only in the densest quartile, with the fourth-quartile top-half subset performing best. For streets with fewer than 25 restaurants, the current analysis does not recover the held-out street at the top-1 position.

Table 21 reports the context-window sensitivity check for $m = 1, 2, 3$. Vector agreement with the $m = 1$ baseline is 0.9995 for $m = 2$ and 0.9992 for $m = 3$; the downstream ranking metrics differ only at the fourth or fifth decimal place, and the Top-10 ranking-change rate remains low at 0.013–0.015. Therefore, $m = 1$ is retained as the default adjacent-neighbor setting.

Table 22 supports this interpretation: wider windows add largely redundant local categories on short Taipei streets.

Table 23 adds a category-level breakdown for the five highest-frequency and five lowest-frequency categories in the leave-one-out analysis. The MRR intervals are computed by bootstrapping leave-one-out trials within each category with 2000 repetitions. About 78% of pairwise MRR intervals overlap, indicating that the observed category differences are largely compatible with sampling variability rather than stable category-level superiority. This result is used as a robustness diagnostic rather than as a claim of uniform category-level performance.

Table 24 reports missing-record and noisy-label robustness diagnostics. The perturbation checks quantify how the scoring analysis behaves when records are removed or category labels are perturbed. The absolute exact-recovery values are low, so the table is interpreted as a sensitivity analysis for incomplete public-platform inputs rather than as evidence that Google Maps bias has been solved.

Tables 25 and 26 contextualize the flat parameter and perturbation sweeps. The observed top-1 value is low in absolute terms but reaches approximately four times the random baseline in the 800-street setting; small score perturbations are therefore unlikely to move a held-out street across the top-1 boundary in this strict exact-recovery task.

To strengthen the empirical evaluation in a real-world setting, this revision introduces a public-proxy evidence dataset constructed from Google Maps-derived restaurant records, street-level commercial-activity indicators, and station-linked mobility-exposure information. The dataset contains 250 street-level records and 5200 restaurant-level records and is used to examine whether the RVLS street score is associated with observable urban-activity signals. The corrected workbook separates the evidence materials into two parts. First, the Street_Level_RealProxy_Data and Restaurant_Level_Raw_Data sheets contain the real-world public-proxy variables used for the evidence analysis. Second, the Negative_Control sheet contains shuffled-control variables used only to test whether the observed association disappears after breaking the correspondence between RVLS scores and proxy outcomes. This dataset does not contain audited pedestrian counts, transaction records, revenue, rent, or store-survival outcomes. It is therefore used as public-proxy evidence rather than direct behavioral or causal demand validation. The street-level external activity index combines review-count intensity, transit accessibility, business-district context, and restaurant density:

$$ExternalActivity_s = \frac{1}{4}\left[z(ReviewIntensity_s) + z(TransitAccessibility_s) + z(BusinessDistrict_s) + z(RestaurantDensity_s)\right]. \quad (13)$$

Here, $z(\cdot)$ denotes standardization across streets. Review-count intensity is the street-level mean of $\log(1 + review\ count)$; transit accessibility summarizes proximity to public-transit access and nearby transit stops; business-district context indicates whether the street is located in a commercial district, night-market area, department-store area, school area, or office area; and restaurant density is the number of restaurants normalized by street-level coverage. Equal weighting is

Table 20
Street-density stratification based on observed restaurant counts.

Density stratum	Trials	Top-1	MRR	nDCG@10	Applicability
First quartile sparse (≤ 15)	783	0.000	0.025	0.024	Not recommended
Second quartile (16–25)	1104	0.000	0.025	0.025	Not recommended
Third quartile (26–40)	1594	0.000	0.021	0.008	Use with caution
Fourth quartile dense (> 40)	1719	0.016	0.027	0.016	Recommended
Fourth quartile top-half (> 50)	860	0.026	0.033	0.021	Most reliable

Table 21
Context-window sensitivity.

m	Vector agreement	Top-1	MRR	MRR 95% interval	nDCG@10	Δ-rate
1	— (baseline)	0.00442	0.02242	[0.020, 0.024]	0.01429	–
2	0.9995	0.00442	0.02239	[0.020, 0.024]	0.01425	0.013
3	0.9992	0.00442	0.02250	[0.020, 0.025]	0.01454	0.015

Vector agreement is the mean cosine similarity between the ρ restaurant-category vectors learned under $m = 1$ and under the compared context window. The MRR 95% interval is bootstrapped with 2000 repetitions using reproducibility setting RS-A. Δ-rate is the fraction of leave-one-out trials in which the Top-10 ranked set changed relative to $m = 1$.

Table 22
Effective context overlap between context windows.

Window pair	Mean Jaccard overlap
$m = 1$ vs. $m = 2$	0.554
$m = 1$ vs. $m = 3$	0.403
$m = 2$ vs. $m = 3$	0.730

Table 23
High- and low-frequency category breakdown.

Group	Category	Trials	Top-1	MRR	MRR 95% interval	nDCG@10
High	Hot Pot	365	0.003	0.019	[0.014, 0.025]	0.008
High	Seafood Rest.	363	0.011	0.030	[0.023, 0.038]	0.024
High	Steak House	361	0.000	0.017	[0.012, 0.022]	0.009
High	Korean Rest.	357	0.011	0.032	[0.024, 0.040]	0.024
High	Thai Rest.	351	0.006	0.025	[0.019, 0.033]	0.019
Low	Dumpling Shop	140	0.000	0.021	[0.013, 0.032]	0.011
Low	Fast Food	137	0.015	0.030	[0.019, 0.044]	0.022
Low	Bento Shop	126	0.000	0.027	[0.016, 0.040]	0.023
Low	Barbecue Rest.	119	0.025	0.043	[0.028, 0.061]	0.032
Low	Bakery	118	0.000	0.023	[0.013, 0.035]	0.018

used to avoid data-driven tuning of the proxy definition. The index is used only as a real-world public-activity proxy and is not interpreted as direct restaurant-level footfall, sales, consumption capacity, or causal demand.

As a check against mechanical overlap with restaurant-derived variables, a non-restaurant public indicator is also computed as the average of standardized transit accessibility, standardized business-district context, and the negative standardized distance to the nearest mass rapid transit station. This indicator excludes review counts and restaurant density, so it is reported as a separate public-context check rather than as a direct demand measure.

To report covariate-adjusted public-proxy associations, the following adjusted diagnostic regression is also estimated at the street level:

$$z(Y_s) = \beta_0 + \beta_1 z(RVLSScore_s) + \beta_2 z(RestaurantDensity_s) + \beta_3 z(CuisineDiversity_s) + \beta_4 z(DistanceToMRT_s) + \varepsilon_s. \quad (14)$$

In Eq. (14), Y_s denotes ExternalActivity, the non-restaurant public indicator, or the station-linked mobility-exposure indicator for the corresponding adjusted row in Table 27. The coefficient β_1 is interpreted as an adjusted public-proxy association rather than a causal effect. Regression p values are two-sided robust t-test values computed with heteroskedasticity-robust standard errors. For the correlation rows, the reported p values are two-sided large-sample correlation test values, and the shuffled-control rows report the same statistic after breaking the correspondence between RVLS scores and proxy outcomes.

Table 27 shows that the RVLS score is positively associated with the composite ExternalActivity index. Density-controlled partial correlations remain positive, but the strong naive-density correlation shows why the evidence must remain conservative. The adjusted non-restaurant public-indicator and station-linked ridership coefficients are small and non-significant, so they are interpreted as descriptive public-context checks rather than independent demand-validation results. The adjusted ExternalActivity coefficient in Table 27 (standardized $\beta = 0.119$, $p = 0.002$) uses street-context controls only; the stricter centrality-control specification in Table 30 reduces this coefficient to $\beta = 0.046$ ($p = 0.089$). The shuffled and seeded permutation controls break the correspondence between RVLS scores and proxy fields. Their near-zero null correlations reduce the risk of treating any street-level summary as automatically correlated with the proposed score. The table remains public-proxy evidence, because review counts, transit access, business-district indicators, station-linked ridership, and restaurant density can reflect several urban processes other than restaurant-specific demand.

5.5. Boundary diagnostics and practical scope

This subsection groups the supplementary robustness, sensitivity, baseline, and data-gap diagnostics that define the supported scope of the method. It is separated from the core results because these diagnostics are used to define when the framework should be trusted, when it should be treated cautiously, and what external evidence is still required before deployment.

The first boundary table focuses on data and representation risks that could otherwise be hidden behind aggregate performance metrics. Table 28 groups the checks for category compression, seed stability, spatial null structure, attraction-versus-competition separation, and survivorship bias, and states how each check limits the interpretation of the learned restaurant-category vectors.

A follow-up 1000-repetition street bootstrap further checks the boundary district-density null result. The observed co-location Spearman exceeds a paired district-density null draw in 58.5% of bootstrap resamples, with mean observed-minus-null Spearman = 0.024 and a bootstrap interval of [–0.223, 0.262]. This result supports a conservative interpretation: the district-density null remains a useful observable spatial-proxy stress test, but it is not treated as decisive zoning-adjusted evidence.

The second boundary table focuses on modeling choices and missing evidence. Table 29 reports diagnostics for spatial dependence, taxonomy robustness, temporal limitations, similarity metrics, and representational capacity, so that the reader can separate parameter sensitivity from unresolved validation gaps.

A metric-swap diagnostic compares cosine and correlation similarity at the leave-one-out screening level. Replacing cosine with correlation similarity changes the Top-10 ranked set in 95.4% of trials, with mean Top-10 Jaccard similarity = 0.504, even though the MRR difference is small (0.0206 versus 0.0196). Therefore, the two metrics are not claimed to produce equivalent candidate lists; cosine is retained because its directional interpretation is simpler for the co-location matrix, while the metric sensitivity is explicitly disclosed.

Table 24
Robustness under missing records and noisy category labels.

Perturbation	Level	Accuracy mean	Accuracystandard deviation	MRR mean	nDCG@10 mean
Observed-record removal	0.0	0.002	0.000	0.022	0.016
Observed-record removal	0.1	0.003	0.001	0.021	0.014
Observed-record removal	0.2	0.003	0.002	0.023	0.016
Observed-record removal	0.3	0.003	0.001	0.022	0.015
Category-label perturbation	0.1	0.003	0.001	0.022	0.015
Category-label perturbation	0.2	0.003	0.001	0.022	0.014
Category-label perturbation	0.3	0.002	0.001	0.022	0.015

Table 25
Unified flat-sweep diagnostic: random-baseline comparison for exact full-ranking recovery.

Streets	Random Top-1	Random MRR	Observed RVLS Top-1	RVLS/Random
200	0.0050	0.0294	~0.005	~1.0
400	0.0025	0.0164	~0.005	~2.0
600	0.0017	0.0116	~0.005	~3.0
800	0.0013	0.0091	~0.005	~4.0

Table 26
Unified flat-sweep diagnostic: score-gap scale relative to parameter perturbations.

Quantity	Value
Mean adjacent-rank score gap	0.0069
Median adjacent-rank score gap	0.0033
Competition perturbation ($\eta_s : 0 \to 1$)	0.0140
Context perturbation ($m : 1 \to 3$), typical	~0.0042
Record-removal perturbation, typical score change	~0.0028

The final boundary table compares the proposed pairwise street-score design with plausible extensions and baselines. Table 30 covers higher-order co-location, within-street distance, transferability, graph centrality, and representation-learning alternatives, and it explains why these checks narrow the contribution to task-oriented screening rather than causal demand modeling.

The higher-order diagnostics use two metrics that point in opposite directions because they measure different properties. The primary evaluation target is the ranking diagnostic rather than the external-proxy correlation, because the intended use case is candidate-street screening rather than external-activity prediction. The triple term is therefore reported as a supplementary sensitivity check rather than a recommended replacement for the pairwise formulation. A follow-up overlay check confirms this decision: adding a standardized triple term on top of the complete RVLS score gives best nonzero-β MRR = 0.0334, below the complete RVLS baseline at $\beta = 0$ (MRR = 0.0421).

The centrality-control diagnostics attenuate the adjusted public-proxy association, showing that part of the score reflects shared density and network-position structure. However, the task-oriented ranking comparison still favors RVLSScore over density, PeopleFlow, weighted degree, eigenvector centrality, and PageRank baselines. In the full-sample ranking comparison, RVLSScore achieves MRR = 0.0421, exceeding density (0.0362), PeopleFlow (0.0374), weighted degree (0.0315), eigenvector centrality (0.0313), and PageRank (0.0315). For held-out streets below the median restaurant density, RVLSScore achieves MRR = 0.0211, compared with weighted degree MRR = 0.0080 and PeopleFlow MRR = 0.0140. The result is therefore interpreted as screening evidence, not as a centrality-independent causal explanation.

The diagnostic tables above identify several recurring boundaries: the data are public rather than audited, the same-street relation is an ordinal locality proxy rather than a measured walking-distance relation, the evaluation target is held-out recovery rather than business survival, and the Taipei evidence does not yet establish temporal or cross-city transfer. Table 31 consolidates these points into a practical reading guide. Each row links one modeling assumption to the validated scope of the current study and then states the limitation and appropriate deployment use.

Accordingly, the proposed framework is most appropriate as an early-stage public-data screening tool for established commercial corridors. It is not presented as a causal demand model, a substitute for audited business data, or a final automatic decision system.

5.6. Limitations

The proposed method still has several limitations. First, the learned restaurant-category vectors are based on public Google Maps records, so incomplete coverage, delayed updates, and inconsistent category labels can affect the resulting representation. Second, the category-standardization dictionary introduces subjectivity because fine-grained restaurant labels must be mapped to canonical categories and higher-level cuisine groups. Third, the same-street locality prior is not a causal assumption and cannot identify actual footfall, transaction volume, consumption capacity, or business success. Fourth, sparse streets, newly developed districts, mixed commercial zones, and streets divided by physical barriers may violate the locality prior. Practical mitigation should combine multi-platform data fusion, uncertainty-weighted street scoring, and active data augmentation for low-coverage districts before using the ranking for decision support. The public-proxy evidence dataset improves external consistency checks but still does not provide behavioral ground truth. Review counts may reflect platform popularity, restaurant age, tourism exposure, and review-writing behavior; transit accessibility and station-linked ridership may reflect commuting, school traffic, office density, or tourism rather than restaurant-specific customer visits; and business-district indicators and restaurant density describe commercial context but do not directly measure consumption capacity. The adjusted ExternalActivity coefficient is positive but small ($\beta = 0.119$, Table 27, street-context controls only), the adjusted non-restaurant public-indicator coefficient is not statistically significant, and density-stratified results show measurable top-1 recovery only in the densest quartile. Practitioners should therefore treat RVLS rankings as more informative for established commercial corridors than for sparse or newly developed streets, and should combine them with rent, lease, pedestrian-access, and field evidence before making a final site decision.

6. Conclusion

This paper learns restaurant-category vectors from neighboring restaurant categories and uses them as indirect co-location signals for street-level recommendation. In the Taipei Google Maps case study, the method ranks candidate streets without directly observing footfall, transaction volume, or consumption capacity. The results suggest that public restaurant arrangements contain useful information for decision support, but the vectors should be interpreted as proxy signals rather than direct measures of demand or business success. In the strict large-candidate diagnostic, RVLS achieves an 800-street Top-1 lift of approximately 4.0 over random and a Top-10 lift of approximately 1.5, supporting its role as a preliminary screening signal rather than a pinpoint demand predictor. Future research should extend the framework through temporal validation with opening and closing records, cross-city validation across different urban contexts, survival

Table 27
Real-world public-proxy evidence for the RVLS street score.

Diagnostic	n	Unit	Statistic	Estimate	p
Real-world public-proxy evidence					
RVLS score vs. ExternalActivity	250	Street	Pearson r	0.705	< 0.001
RVLS score vs. ExternalActivity	250	Street	Spearman ρ	0.700	< 0.001
RVLS score vs. ExternalActivity | density	250	Street	Partial r	0.382	< 0.001
RVLS score vs. ExternalActivity | density, diversity, MRT	250	Street	Partial r	0.359	< 0.001
Adjusted proxy association with ExternalActivity	250	Street	Standardized β	0.119	0.002
RVLS score vs. non-restaurant public indicator	250	Street	Pearson r	0.563	< 0.001
Adjusted association with non-restaurant public indicator	250	Street	Standardized β	0.039	0.199
RVLS score vs. station-linked mobility indicator	190	Station-linked street	Pearson r	0.457	< 0.001
Adjusted association with station-linked mobility indicator	190	Station-linked street	Standardized β	0.052	0.473
Proxy-component and decomposition checks					
RVLS score vs. review-count intensity	250	Street	Pearson r	0.678	< 0.001
RVLS score vs. transit accessibility	250	Street	Pearson r	0.429	< 0.001
RVLS score vs. business-district indicator	250	Street	Point-biserial r	0.501	< 0.001
RVLS score vs. restaurant density	250	Street	Pearson r	0.664	< 0.001
Naive density vs. ExternalActivity	250	Street	Pearson r	0.763	< 0.001
ExternalActivity without review counts	250	Street	Pearson r	0.663	< 0.001
ExternalActivity without density	250	Street	Pearson r	0.649	< 0.001
Negative-control diagnostics					
Shuffled-street negative control	250	Street	Pearson r	−0.004	0.951
Shuffled-ridership negative control	190	Station-linked street	Pearson r	−0.131	0.073
Seeded permutation control, ExternalActivity	1000	Permutation	Mean null r	0.002	< 0.001
Seeded permutation control, mobility indicator	1000	Permutation	Mean null r	−0.002	< 0.001

Table 28
Boundary diagnostics for compression, stability, spatial structure, mechanism separation, and survivorship.

Diagnostic domain	Diagnostic design	Main finding and implication
Category compression	Restaurant-identity feasibility audit and category, cuisine, and multi-resolution comparison	Restaurant identity tokens are all singletons (singleton rate = 1.000), so restaurant-level embeddings are underidentified in the current public snapshot. Category and cuisine tokens have complete nonzero directed co-occurrence coverage and are estimable; cuisine-only MRR = 0.0214, category-only MRR = 0.0205, and the best category-cuisine blend MRR = 0.0213.
Embedding stability	Thirty independent CBOW/RCNN initializations with Procrustes, neighbor-preservation, cosine-consistency, and ranking diagnostics	Downstream ranking is relatively stable across seeds (MRR mean = 0.0204, standard deviation = 0.0016), but local embedding geometry is not stable after alignment (mean Procrustes root mean squared error = 0.0535, standard deviation = 0.0107; mean Jaccard@5 = 0.188; mean cosine Spearman relative to the reference seed = 0.092). The aggregate co-location signal is therefore treated as statistically non-random, while individual embedding-neighborhood diagrams are not interpreted as unique semantic structures.
Spatial null structure	Zoning-data audit plus global category-marginal, district-density block, and cuisine-preserving constrained null models	Three observable nulls test category marginals, district density, and cuisine preservation. The cuisine-preserving and global category-marginal nulls are significant, while the district-density null is near the conventional boundary. The evidence is limited but reproducible; true zoning validation requires external zoning, land-use, rent, or parcel data.
Mechanism separation	Proxy discrimination for same-cuisine, same-price, and high co-location pair labels	Cosine similarity does not distinguish same-cuisine or same-price proxy competition labels (area under the curve (AUC) = 0.495 and 0.501), but it moderately predicts high co-location proxy pairs (AUC = 0.629). Therefore, attraction/competition separation is treated as a scoring-rule decomposition.
Survivorship limitation	Formal missing-outcome audit and pseudo-removal stress tests	The dataset lacks opening, closure, relocation, business-status, and survival-status fields. Random 10%–40% restaurant removal keeps MRR around 0.0208–0.0212, whereas targeted low-review removal changes geometry more strongly. These are sensitivity checks, not survival-aware validation.

and business-outcome data, richer geographic features such as fine-grained coordinates and pedestrian-network distance, and interpretable hybrid models that combine co-location vectors with audited foot-fall, demographic, mobility, and economic variables. Temporal work should distinguish stable co-location patterns from short-run platform-update noise by comparing repeated public snapshots and by using train-earlier/test-later evaluation. Cross-city work should test whether category vectors trained in one city transfer to cities with different transit systems, zoning structures, density levels, and cuisine cultures, or whether city-specific recalibration is required. Survival-aware work should treat site selection as a longitudinal outcome problem by linking recommendations to openings, closures, relocation, rent changes, lease duration, and revenue or transaction proxies when such records become available. Geographic extension should replace the same-street proxy with pedestrian-network distance, barriers, parcel-level land use, frontage, and nearby anchors, so that future models can distinguish physically adjacent but commercially disconnected locations. For implementation-oriented research, the ranking should also be evaluated as a decision report for non-technical users. Such a report can present the street score, the supporting restaurant categories, direct-competition flags, and uncertainty indicators caused by sparse records or unstable category labels while keeping model evidence separate from owner-side constraints such as rent, lease terms, storefront visibility, pedestrian access, and regulatory feasibility. Evaluating these reports

Table 29
Sensitivity diagnostics for data, metrics, and representation choices.

Diagnostic issue	Test or audit	Key result	Interpretation boundary
Spatial dependence	District-block statistics and bootstrap.	$I = 0.335$ for RVLSScore; residual $I = 0.056$; bootstrap $r = [0.621, 0.772]$.	Evidence is spatial-block-aware descriptive screening evidence, not independent pair-level inference.
Taxonomy robustness	Category dictionary audit and 5%–30% label perturbation.	Within-cuisine Spearman: 0.707 at 5%; 0.227 at 30%; cross-cuisine 30%: 0.133.	Canonical labels are usable but taxonomy-dependent; no inter-annotator claim is made.
Temporal coverage	Field audit and pseudo-removal stress tests.	No opening, closing, relocation, snapshot, or multi-period field is available.	Robustness tests do not replace train-earlier/test-later validation.
Similarity metric	Cosine, correlation, radial-basis, and metric-swap checks.	Correlation MRR = 0.0206; cosine MRR = 0.0196; Top-10 Jaccard = 0.504.	Cosine is retained for interpretability, not because it is uniquely optimal.
Embedding dimension	Context-matrix spectrum and dimension sweep.	Effective rank = 10.45; participation ratio = 8.44; 80/90/95% dimensions = 7/9/11.	Four dimensions form a compact undercomplete representation.

Table 30
Boundary diagnostics for model extensions and baselines.

Boundary issue	Diagnostic comparison	Key result	Supported interpretation
Higher-order terms	Triple co-location lift and overlay on complete RVLS.	Complete RVLS MRR = 0.0421; best nonzero triple overlay MRR = 0.0334.	Pairwise scoring is retained; triple terms are sensitivity checks.
Within-street distance	Order randomization, context windows, and ordinal decay.	Randomized order Spearman = 0.0497; ordinal decay $\lambda = 1.0$ Spearman = 0.818.	No meter-level coordinates are available; this is not a true distance-decay model.
Transferability	Data audit and leave-one-district-out stress test.	No cross-city field; district holdout MRR range = 0.0157–0.0416.	Taipei evidence should not be generalized without multi-city validation.
Centrality controls	Graph-centrality controls and below-median-density ranking.	Marginal $r = 0.705$; adjusted $\beta = 0.046$, $p = 0.089$.	Street score partly reflects density structure, not a centrality-independent causal effect.
Embedding baselines	PPMI-SVD, DeepWalk-SVD-lite, Node2Vec-lite, random, and shuffled embeddings.	DeepWalk-SVD-lite MRR = 0.0217; PPMI-SVD = 0.0205; Node2Vec-lite and CBOW/RCNN = 0.0204.	CBOW/RCNN is retained because it matches the local same-street context assumption.
Ranking baselines	Final street-score comparison with density and PeopleFlow references.	RVLSScore MRR/nDCG@10 = 0.0421/0.0338; density = 0.0362/0.0307; PeopleFlow = 0.0374/0.0303.	The contribution is task-oriented shortlisting performance.

Table 31
Assumptions, evidence, and use boundaries of the proposed framework.

Aspect	Key assumption	Validated scope in this study	Limitation and practical use
Data source	Public restaurant records preserve useful co-location signals.	Taipei Google Maps records support leave-one-out, perturbation, and public-proxy checks.	The data exclude audited footfall, transactions, rent, revenue, and survival; deployment should add multi-platform and field checks.
Representation	Same-street restaurant categories can proxy context compatibility.	Aggregate co-location, vector, and ranking diagnostics support category-level screening signals.	Vectors are not direct semantic, causal, or demand measures; interpret them as screening features.
Locality prior	Existing category mixes help compare candidate streets.	Evidence is strongest for established and restaurant-dense commercial corridors.	Long roads, barriers, sparse streets, and new districts require pedestrian-access and local-market verification.
Evaluation target	Held-out recovery is a proxy for shortlist usefulness.	Binary metrics, full-ranking metrics, and hit-rate diagnostics show limited but measurable shortlist value.	Exact recovery is not business-success prediction; use the ranking to prioritize inspection, not to automate site choice.
Deployment use	Model evidence and owner-side constraints are complementary.	Table 3 defines a practical screening workflow from public records to field checks.	Final decisions must include rent, lease, frontage, legal constraints, competitor inspection, and owner judgment.
Generalization	Transfer requires comparable urban and market structure.	Current validation is limited to Taipei with internal stress tests and public-proxy evidence.	Temporal, cross-city, business-survival, and richer geographic validation are required before broader claims.

with restaurant owners or planners would clarify which explanations are useful, which warnings reduce overinterpretation, and how public-data evidence can be combined with private due diligence in an accountable site-selection process.

CRediT authorship contribution statement

Yu-Ting Yang: Writing – original draft, Methodology, Conceptualization. **Chih-Yung Chang:** Writing – original draft, Methodology, Formal analysis, Conceptualization. **Syu-Jhih Jhang:** Writing – review & editing, Writing – original draft, Visualization, Software. **Diptendu Sinha Roy:** Writing – review & editing, Visualization, Methodology.

Informed consent

All participating authors have been informed.

Ethical approval

This study does not involve either human subjects or animals.

Funding details

This study was not funded by any institution.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The corrected real-world public-proxy evidence workbook used for Table 27 is provided as expanded_restaurant_external_proxy_dataset_corrected.xlsx. The workbook separates anonymized street-level public-proxy data, processed restaurant-level records, ExternalActivity calculation, and shuffled negative controls into separate sheets. Table 27 can be regenerated from the available workbook and supporting validation materials. The original non-anonymized Google Maps-derived records, including raw store names, exact address strings, web links, and precise location-specific commercial records, are not publicly released, but they are available from the corresponding author upon reasonable request.

References

Altay, B.C., Celik, E., Okumus, A., Balin, A., Gul, M., 2023. An integrated interval type-2 fuzzy BWM-MARCOS model for location selection of E-scooter sharing stations: The case of a university campus. Eng. Appl. Artif. Intell. 122, 106095.

Bhole, J., Nandiyawar, S., Pawar, S., Vora, P., 2020. Smart site selection using machine learning. Int. Res. J. Eng. Technol. 7 (5), 3012–3015.

Bilen, T., Erel-Ozcevik, M., Yaslan, Y., Oktug, S.F., 2018. A smart city application: Business location estimator using machine learning techniques. In: Proceedings of the IEEE International Conference on High Performance Computing and Communications.

Chang, T.H., 2010. Restaurant location selection by utilizing the fuzzy preference relations. In: Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management.

Deveci, M., Erdogan, N., Cali, U., Stekli, J., Zhong, S., 2021. Type-2 neutrosophic number based multi-attributive border approximation area comparison (MABAC) approach for offshore wind farm site selection in USA. Eng. Appl. Artif. Intell. 103, 104311.

Dixit, A., Clouse, C., Turken, N., 2019. Strategic business location decisions: Importance of economic factors and place image. Rutgers Bus. Rev. 4.

Dong, J.-Y., Yao, Y.-Y., Chen, S.-M., Wan, S.-P., 2025. An integrated method of hotel site selection based on probabilistic linguistic multi-attribute group decision making. Eng. Appl. Artif. Intell. 147, 110328.

Eravci, B., Bulut, N., Etemoglu, C., Ferhatosmanoglu, H., 2016. Location recommendations for new businesses using check-in data. In: Proceedings of the IEEE 16th International Conference on Data Mining Workshops. pp. 1110–1117.

Furtado, A.S., Fileto, R., Renso, C., 2013. Assessing the attractiveness of places with movement data. J. Inf. Data Manag. 4, 124–133.

Han, S., Jia, X., Chen, X., Gupta, S., Kumar, A., Lin, Z., 2022. Search well and be wise: A machine learning approach to search for a profitable location. J. Bus. Res. 144, 416–427.

Krishankumar, R., Ecer, F., 2024. A multi-criteria framework for electric vehicle charging location selection using double hierarchy preferences and unknown weights. Eng. Appl. Artif. Intell. 133, 108251.

Mazhi, K.Z., Suryana, L.E., Davi, A., Dewi, W.R., 2020. Site selection of retail shop based on spatial analysis and machine learning. In: Proceedings of the International Conference on Advanced Computer Science and Information Systems.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J., 2013. Distributed representations of words and phrases and their compositionality. In: Advances in Neural Information Processing Systems. pp. 3111–3119.

Quan, X., Wenyin, L., Dou, W., Xiong, H., Ge, Y., 2012. Link graph analysis for business site selection. Computer 45 (3), 64–69.

Seikh, M.R., Mandal, U., 2022. Multiple attribute group decision making based on quasirung orthopair fuzzy sets: Application to electric vehicle charging station site selection problem. Eng. Appl. Artif. Intell. 115, 105299.

Tayeen, A.S., Mtibaa, A., Misra, S., 2019. Location, location, location!. In: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.

Wang, F., Chen, L., Pan, W., 2016. Where to place your next restaurant? Optimal restaurant placement via leveraging user-generated reviews. In: Proceedings of the 25th ACM International Conference on Information and Knowledge Management.

Wang, Y., Li, S., Zhang, X., Jiang, D., Hao, M., Zhou, R., 2020. Site selection of digital signage in Beijing: A combination of machine learning and an empirical approach. ISPRS Int. J. Geo-Information 9 (4).

Yazir, K., Karasan, A., Kaya, I., 2025. Optimization and prioritization of electric vehicle charging locations for InterCity transport: A dual-phase approach. Eng. Appl. Artif. Intell. 159, 111786.