Journal of Information and Data Management, 2026, 17:1, doi: 10.5753/jidm.2026.5929

POI Type Embedding Techniques for Recommendation Systems: A Comparative Analysis

Diogo Alves Silveira [**Universidade Federal de Campina Grande** | ***diogo.silveira@ccc.ufcg.edu.br***]
Salatiel Dantas Silva [**Universidade Federal de Campina Grande** | ***salatiel.dantas@computacao.ufcg.edu.br***]
Nícolas Moreira Nobre Leite [**Universidade Federal de Campina Grande** | ***nicolas.leite@ccc.ufcg.edu.br***]
Claudio E. C. Campelo [**Universidade Federal de Campina Grande** | ***campelo@computacao.ufcg.edu.br***]

Systems and Computing Department Federal University of Campina Grande (UFCG) Campina Grande – PB – Brazil

Received: 15 April 2025 • **Published:** 13 March 2026

Abstract Modern Point of Interest (POI) Recommendation Systems (RSs) employ diverse technologies to enhance user experience and engagement. These systems utilize various techniques to predict user preferences and recommend the next POI, including machine learning algorithms that integrate user behavior with category and time preferences from other users, as well as state-of-the-art embedding techniques that leverage geographic features around POIs to assess user preference likelihood. This study addresses key gaps in POI Recommendation Systems by evaluating algorithms that enhance POI representation through the combined use of check-in data, geographic features, and advanced embedding techniques.

Keywords: Recommendation Systems, Points of Interest, Geographic Embeddings, Location-Based Systems

1 Introduction

Recommender Systems (RSs) play a crucial role in personalizing user experiences, providing highly relevant suggestions across various domains, including e-commerce [Bhuvanya and Kavitha, 2023], entertainment [Kulkarni *et al.*, 2023], and music [Saito and Sato-Shimokawara, 2023]. These systems are essential for helping users navigate through numerous options, recommending products, movies, and music that align with their preferences. Within the vast field of RSs, the recommendation of Points of Interest (POIs) represents a subarea of growing interest. However, effective POI recommendation presents unique challenges due to the inherent complexity of capturing user preferences, which are influenced by cultural, personal and geographical factors.

Modern POI RSs often rely on embeddings (vector representations that capture features of the objects) that are built considering similar information, such as user reviews of locations, check-ins, visiting times, and POI types [Liu *et al.*, 2019; Zhang *et al.*, 2023; Yang *et al.*, 2022]. These data are quite relevant as they represent trajectory patterns and characteristics of the visited POIs. For instance, POI types categorize locations in various ways, such as commercial (e.g., restaurants, hotels), recreational (e.g., parks, museums), and transportation (e.g., airports, train stations), facilitating searches in geographic databases and groupings. However, despite recent researches have shown good results [Yang *et al.*, 2022; Yan *et al.*, 2023; Yin *et al.*, 2023], they do not consider the use of geographic and spatial features in the POIs' vicinity. Geographic features refer to any part of the Earth's surface or any element present on it that can be represented on a map[1]. They are normally regarded as geographic entities with precise boundaries, such as rivers, squares and roads. In geographic databases, they are associated with geometric objects of different types, such as points, lines and polygons. Furthermore, spatial features encompass elements such as the linear separation between POIs or the trajectory necessary to traverse from one location to another.

Geographic and spatial features can be crucial for characterizing user preferences and may even motivate the choice of the next POI to be visited. For example, some people may be more likely to visit restaurants near lakes than restaurants in denser urban areas or having to choose a closer destination due to time restraints. That said, we conducted a study to answer the following questions:

1. Can geographic embeddings of POI types, which encapsulate spatial characteristics, enhance the performance of POI recommendation systems?
2. To what extent does the spatial proximity between POIs influence user preferences in selecting subsequent POIs?
3. How do different neural architectures impact the effectiveness of POI-type embeddings?

To investigate these questions, we propose an approach to integrating such embeddings algorithms into a POI recommendation model and experimentally evaluated our solution by comparing the effectiveness of a state-of-the-art POI RS before and after this adaptation.

To do this, we selected a recently developed approach to incorporating geographic features into POI type embeddings [Silva *et al.*, 2023] generated by Word2Vec and BERT models, which has shown promising results in other application scenarios, and an algorithm based in the Shortest Path approach presented in [Yao *et al.*, 2017], that generates embeddings based on the spatial distribution of the POIs. The model selected for implementing the approach was GETNext [Yang *et al.*, 2022], referenced in the literature as one of the

[1] https://support.esri.com/pt-br/gis-dictionary/search?q=feição

leading options for POI recommendation. The results obtained indicate that the approach appears promising, successfully capturing the importance of geospatial features present in the context of a POI for choosing the next POI a user will visit.

This article is an extended version of [Leite *et al.*, 2024], where the foundation of the work was established, presenting the benefits of adding the geographic features representation to the embeddings of the Graph Convolutional Network proposed by [Yang *et al.*, 2022]. In this manuscript we expand upon the limitations of the previous work, providing broader experimentation scope and different algorithms for POI representation.

The remainder of this paper is organized as follows. Section 2 describes related work in the area of POI representation and POI RSs. In Section 3, we present our methodology, including the model architecture and the method of generating geographic embeddings. Section 4 presents the results of our experiments and Section 4.4 discusses them. Finally, Section 5 concludes the paper and points to future research.

2 Related Work

This section describes related work in the areas of POI representation and POI RSs.

2.1 POI Representation

Several studies have explored POI representation through embedding techniques, highlighting them as a promising approach to enhance the effectiveness of POI RSs. Studies such as [Wang *et al.*, 2020], who introduced Urban2Vec, exemplify significant innovations in this field. Urban2Vec is an unsupervised technique that generates representations of neighborhoods using a combination of street view images and textual information from POIs. This approach is notable because it integrates visual and textual features with geospatial data, outperforming existing baselines and competing with fully supervised methods in downstream prediction tasks such as urban planning, business model development, and social well-being improvement.

In the context of POI RSs, embeddings generated by these techniques are used to capture the essence of POIs and contextualize recommendations according to user preferences and local characteristics. For example, [Feng *et al.*, 2017] introduced the POI2Vec model, which uses embeddings to represent POIs based on user interactions and check-ins, allowing for a deeper understanding of movement patterns and preferences.

Additional advances leverage spatially-aware embeddings to model POI relationships. [Yao *et al.*, 2017] pioneered this by generating POI vectors from shortest-path sequences, capturing functional and geographic proximity. Unlike frequency-based methods (TF-IDF) or topic models (LDA), their approach preserved spatial hierarchies, improving land-use classification by 20%.

On the other hand, [Silva *et al.*, 2023] proposed GeoContext2Vec, a method that leverages the importance of geographic features in the vicinity of a POI for generating embeddings. Unlike methods that rely on check-in data or textual descriptions, GeoContext2Vec evaluates the proportion and uniqueness of the space occupied by these geographic features. This method demonstrated superiority over ITDL ([Yan *et al.*, 2017]), a state-of-the-art method based on POI co-occurrences, in terms of POI type similarity evaluation by human analysis. Additionally, the use of public domain maps such as OpenStreetMap (OSM) for vector creation is highlighted as a practical and reproducible alternative.

Despite advancements in POI representation and recommendation systems, a critical research gap remains in the effective incorporation of fine-grained features to enhance POI characterization within the recommendation process. Existing approaches primarily rely on general geographic information such as POI coordinates [Luo *et al.*, 2021; Wang *et al.*, 2020]. However, existing approaches frequently neglect the influence of hyperlocal geospatial characteristics—including proximity to green spaces, water bodies, or major roadways, as well as the accessibility distance to POIs—on user decision-making. This gap limits the ability of current models to capture the nuanced relationships between users and their surrounding environment, hindering the potential for truly personalized and context-aware recommendations. Therefore, this study employs GeoContext2Vec's Word2Vec and BERT embeddings, along with Shortest Path analysis, to generate enriched POI-type embeddings that incorporate geographic contextual features. These embeddings are subsequently integrated into the POI recommendation framework to enhance recommendation accuracy.

2.2 Conventional POI Recommendation Systems

Initial research in the area of POI RSs utilized Markov chains and matrix factorization techniques [Davtalab and Alesheikh, 2021; Koren *et al.*, 2009; Zhao *et al.*, 2016]. However, these approaches have become limited in representing users' visitation patterns when compared to deep learning and embedding-based approaches [Feng *et al.*, 2020].

Additionally, many approaches model the relationships between visited POIs and potential POIs by incorporating and prioritizing temporal factors. [Yuan *et al.*, 2013] identified a significant gap in existing methods, which often neglected the influence of the specific time of day on user preferences. They proposed a collaborative recommendation model that integrates temporal information, allowing for more accurate recommendations at different times of the day. [Shi *et al.*, 2021] introduced an attentional memory network with correlation-based embeddings (AMN-CE) for time-aware POI recommendation, proposing a temporal attention mechanism to adjust the influence of different times on user preferences. [Wang *et al.*, 2021] developed a model that incorporates users' temporal check-in preferences by designing a cross-graph neural network to control the information flow across different semantic spaces, enhancing recommendation accuracy by considering the relationship between check-in times and POIs.

[Halder *et al.*, 2021] addressed the task of next POI recommendation by considering the queue time users spend entering a POI, a critical factor influencing user mobility behavior.

Using a Transformer model, the TLR-M, the authors not only recommend the next POI but also predict the queue time required for a user to enter a POI.

While temporal factors are important, the geographic influence of POIs is also crucial for improving recommendations. Various geographic embedding generation techniques have been applied considering the context of POIs, neighborhood characteristics, and temporal factors to capture user preferences and enhance recommendations.

2.3 POI Recommendation Systems that consider geographic information

[Luo *et al.*, 2021] proposed the Spatio-Temporal Attention Network (STAN), a model that addresses spatial scarcity and temporal relations by employing a two-layer attention architecture. This model excels in facilitating interactions between non-adjacent locations and non-consecutive check-ins, capturing the explicit spatiotemporal effects influencing user behavior. [Wang *et al.*, 2022] proposed a POI recommendation method leveraging sequential, categorical, and geographic influences. Their process begins by extracting latent vectors of POIs and user preferences from check-in sequences using a word embedding model. Collaborative filtering is then applied to predict user preferences for different POIs based on their behavior.

[Qin *et al.*, 2023] proposed the Disentangled Dual-Graph (DisenPOI) framework for next POI recommendation, enhancing POI recommendation by disentangling sequential and geographic influences, often conflated in traditional approaches. The framework uses two distinct graphs: one to model the user's visit sequence and another to represent the geographic relationships between POIs. Through contrastive learning, the framework extracts disentangled representations of these influences, allowing for a clearer understanding of user preferences and resulting in more accurate and interpretable recommendations. [Yang *et al.*, 2022] presented the Graph Enhanced Transformer (GETNext) model, which also uses graphs for next POI recommendation, leveraging a global trajectory flow map and a Transformer-based architecture. To capture generic user movement patterns between POIs, the authors constructed a graph representing user trajectories, where nodes reflect POIs with attributes such as geographic location, POI category, and check-in frequency. They then used a Graph Convolutional Network to learn POI embeddings encapsulating these global transitions and recommend the next POI to visit.

[Yan *et al.*, 2023] developed the Spatio-Temporal Hyper-Graph Convolutional Network (STHGCN), a model for next POI recommendation that employs a hypergraph to analyze detailed trajectory information. The model learns from historical and collaborative trajectories, addressing the cold start problem [Lika *et al.*, 2014] and improving recommendations for various user trajectory durations. Using a Transformer with hypergraph integration, STHGCN incorporates hypergraph structures with spatiotemporal data, outperforming previous methods in practical tests. [Yin *et al.*, 2023] developed the Sequence-based Neighbour search and Prediction Model (SNPM) for next POI recommendation, using graph embedding techniques and Eigenmap methods to analyze POI relationships from sparse check-in data. The model includes a Dynamic Neighbour Graph and Multi-Step Dependency Prediction, considering both current states and historical visit sequences to POIs.

While the aforementioned approaches have significantly advanced the incorporation of geographic data and graph-based architectures to improve recommendations, there remains a crucial gap in integrating geospatial features into POI recommendation models. Therefore, this study proposes a novel approach that integrates geographic and spatial embeddings of POI types by leveraging both geospatial characteristics and contextual features. The methodology employs GeoContext2Vec and Shortest Path algorithms, building upon established work in geospatial representation learning [Silva *et al.*, 2023; Yao *et al.*, 2017; Li *et al.*, 2022]. We then integrate these embeddings, generated by Word2Vec and BERT models, into the graph and Transformer-based architecture of the GETNext model [Yang *et al.*, 2022] to determine if the addition of these embeddings enhances POI recommendations.

3 Methodology

In our methodological approach, we employ the generation of Geographic POI-type embeddings using various different algorithms and then integrate the different embeddings into the graph-based and transformer architecture of the GETNext model. The details of these processes are presented in the following subsections.

3.1 Generation of POI Type Embeddings

As mentioned earlier, numerous algorithms are used to generate representations of POIs. POI recommendation systems can greatly benefit from the attempt to capture user preferences based on different attributes. To achieve this, we utilize 3 different techniques to represent POI-types, GeoContext2Vec, which utilizes geographic features present in a POI's context to represent it into an Word2Vec embedding format; an Shortest Path algorithm, that utilizes the distances between the POIs to infer the likeliness of the next visit based on the path it would take; and an adaptation of the GeoContext2Vec for the BERT transformer architecture.

3.1.1 GeoContext2vec

GeoContext2Vec was selected for its demonstrated capability to effectively incorporate geospatial contextual features in POI representation. The algorithm utilizes geographic data extracted from OpenStreetMap (OSM) to generate POI-type embeddings by analyzing the co-occurrence patterns of geographic features within a specified proximity radius of each POI. This approach captures the spatial relationships between POIs and their surrounding geographic context through quantitative analysis of feature distributions.

In this context, our implementation employs for primary OSM data tables:

1. `planet_osm_polygons`: Contains areal features including buildings, green spaces (parks), and water bodies (rivers, lakes)
2. `planet_osm_lines`: Comprises linear features such as local streets, minor highways, and railway networks
3. `planet_osm_roads`: Includes major transportation arteries (primary roads and highways)
4. `planet_osm_points`: Consists of point-based features like urban infrastructure (traffic lights) and amenities (fountains, street trees)

Initially, the GeoContext2Vec algorithm considers a radius around the POI (e.g., 100 meters) and identifies the geographic features present in this context (Figure 1). It then calculates the proportion of space occupied by each geographic feature in the context. For example, a river that runs through the entire context will have a higher proportion than a small building.

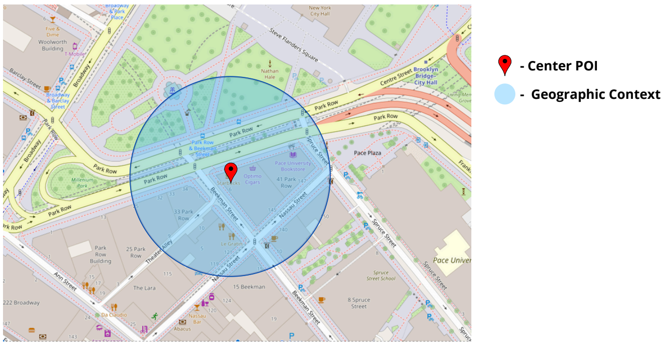

Figure 1. Geographic context of a POI

Additionally, the algorithm considers a parameter ω that determines the emphasis given to the area of geographic features relative to their occurrences. If ω is set to 1.0, only the area of the features will be considered. On the other hand, if ω is 0.0, only the occurrences of the features will be taken into account. Values between 0.0 and 1.0 indicate a combination of the two, where the proportion is determined by the parameter value. The algorithm also evaluates the uniqueness of each geographic feature. For example, a single river will be more relevant for characterizing the context than small trees present in large quantities. By combining space proportion and uniqueness, the algorithm generates a factor that increases the co-occurrence relationship between the POI type and each geographic feature present in the context. Finally, the set of generated co-occurrence relationships is used to train a Word2Vec model [Mikolov *et al.*, 2013], which learns vector embeddings for each POI type. Additionally, the co-occurrence set of relationships was also used to train a BERT model [Devlin *et al.*, 2019], that unlike traditional embedding methods, generates highly robust contextualized embeddings that capture deeper semantic and syntactic information by leveraging bidirectional attention mechanisms, dynamically adjusted based on the surrounding context, resulting in richer representations that encapsulate a broder range of features. Consequently, the enhanced informational capacity of BERT embeddings may lead to superior performance in downstream tasks compared to static or context-independent word representations.

These embeddings capture the contextual relationships between POI types and the geographic features of the context.

Algorithm 1 POI Shortest Path Ordering

Require: Set of N POIs $\mathcal{P} = \{P_1, P_2, \ldots, P_N\}$ with coordinates (x, y)
Ensure: Ordered sequence $path\_list$ representing the shortest path

Step 1: Endpoint Selection
$max\_dist \leftarrow 0.0$
$start\_node \leftarrow$ NULL
$end\_node \leftarrow$ NULL
for all $p_i \in \mathcal{P}$ **do**
  for all $p_j \in \mathcal{P}$ **do**
    $dist \leftarrow$ calc_euclidean_dist(p_i, p_j)
    if $dist > max\_dist$ **then**
      $max\_dist \leftarrow dist$
      $start\_node \leftarrow p_i$
      $end\_node \leftarrow p_j$
    end if
  end for
end for
$path\_list \leftarrow [start\_node, end\_node]$
$remaining\_pois \leftarrow \mathcal{P} \setminus \{start\_node, end\_node\}$
Step 2: Greedy Insertion
while $remaining\_pois \neq \emptyset$ **do**
  $current\_poi \leftarrow remaining\_pois$.pop()
  $min\_path\_len \leftarrow \infty$
  $optimal\_idx \leftarrow -1$
  for $k \leftarrow 0$ **to** length$(path_list)$ **do**
    $temp\_path \leftarrow path\_list$.copy()
    $temp\_path$.insert$(k, current_poi)$
    $current\_total\_len \leftarrow$ calc_path_length$(temp_path)$
    if $current\_total\_len < min\_path\_len$ **then**
      $min\_path\_len \leftarrow current\_total\_len$
      $optimal\_idx \leftarrow k$
    end if
  end for
  $path\_list$.insert$(optimal_idx, current_poi)$
end while
return $path\_list$

3.1.2 Shortest Path

The Shortest Path is employed with the distances between POIs in mind when selecting the next destination, employing geographic coordinates, to compute the distance and possible paths between each POI within the selected area.

Considering the designated area (e.g., a neighborhood, municipal district, or the city itself), the POIs are linked to each other by their spatial relations, revealing spatial distribution attributes and positional relationships of the locations. To associate the points in realistic ways, we start the algorithm by using a greedy approach to find the shortest path in the selected area that passes all the POIs, storing them in sequential order. Afterwards, path-based documents are constructed utilizing words according to the sequence of POIs.

If N POIs exist in the selected area, we can denote the POIs as $P_1(x_{P_1}, y_{P_1})$; $P_2(x_{P_2}, y_{P_2})$; $\ldots$; $P_N(x_{P_n}, y_{P_n})$, with their coordinates (x, y). The specific procedure for determining the order of POIs is detailed in Algorithm 1.

With the path between POIs established, it becomes possible to generate embeddings containing the shortest path between all POIs in the selected area, taking into account both distance and the path length between POIs [Yao *et al.*, 2017]. These embeddings are derived from training a Word2Vec model, thereby capturing the spatial relationships between the paths required to travel from one point to another.

3.1.3 GeoContext2BERT

In light of prior research, a key objective of this study was to examine whether state-of-the-art Natural Language Processing (NLP) models could be effectively utilized to generate embeddings for Points of Interest (POI) types [Silva *et al.*, 2023]. To this end, we selected the BERT model [Devlin *et al.*, 2019] due to its widespread adoption in NLP tasks and its demonstrated superior performance across various applications. Additionally, BERT's Transformer-based architecture enables the capture of intricate linguistic relationships within training documents, making it particularly suitable for this investigation.

To produce POI-type embeddings using BERT, we adopted a SpaBERT [Li *et al.*, 2022] inspired methodology, which incorporates geospatial context by leveraging neighboring entities around a given POI [Silva *et al.*, 2023]. This approach involves converting geospatial entities and their adjacent counterparts into structured documents, which are then used to train the BERT model via Masked Language Modeling (MLM). During training, selected entities are masked with special tokens, requiring the model to predict the missing elements, thereby learning the underlying contextual relationships between entities. For POI-types and geographic features, training documents were constructed using binary relations generated by the GeoContext2Vec algorithm [Silva *et al.*, 2023]. Each relation follows the format <*central POI-type, contextual POI-type*>, forming a two-word sentence. A single document aggregates all such relations associated with a specific geographic feature, allowing the model to learn the contextual dependencies between POI-types within a given spatial setting.

Upon completion of data preparation, the BERT model was fine-tuned to generate geographically contextualized POI-type embeddings. The final step involved transferring the trained BERT weights into a Word2Vec-compatible format, preserving the learned representations for downstream applications.

3.1.4 Parameters and Training Sets

Following the foundational methodologies described in the previous subsections, we adopted specific parameter configurations for embedding generation, considering the POI types present in our dataset. For the GeoContext2Vec algorithm [Silva *et al.*, 2023], we defined a radius of 400 meters for the geographic context of a POI, considering all geographic features within this circumference. This value was established through experimentation, with 400 meters producing the most favorable results 1. We also set the parameter ω to 0.8, meaning that 80% of the weight is allocated to the geographic feature's area and 20% to occurrences, which proved most suitable in our experimental setup.

With ω defined, we extracted all relevant features from the OSM tables and generated training sets capturing co-occurrences between each POI type and geographic feature. Each co-occurrence was replicated according to the ω factor, accounting for both the spatial proportion occupied by the feature and its uniqueness in the POI context. The generated training sets capture relationships between POIs and the following geographic entities:

1. **Polygons** (e.g. buildings, rivers, parks)
2. **Lines** (e.g. access roads and small streets)
3. **Roads** (e.g. roads, highways, avenues)
4. **Points** (e.g. trees, lamps, transit signs)

For the shortest-path analysis, we followed the approach proposed by Yao et al. [Yao *et al.*, 2017], designating the entire dataset as the spatial extent for path generation. A contextual neighborhood of 10 POIs—comprising five predecessors and five successors—was used to provide geographic context for embedding generation. The resulting training sets were employed to train a Word2Vec model utilizing the Skip-Gram architecture [Mikolov *et al.*, 2013], producing embeddings that reflect the spatial distribution characteristics of POIs.

Additionally, we employed a BERT-based adaptation following the methodology established in prior work [Leite *et al.*, 2024]. The binary relationships between POIs and neighboring features generated by the GeoContext2Vec algorithm were used to train a BERT model. The training procedure involved converting these relationships into structured BERT-compatible documents, utilized in a Masked Language Modeling (MLM) framework. Following standard BERT training procedures [Devlin *et al.*, 2019], 15% of the tokens were randomly masked during training. The resulting masked documents were then fed into the BERT model, which was trained to predict the obscured tokens, thereby learning to infer relationships between POIs and their associated geographic features. To leverage prior knowledge of linguistic semantics, we initialized the fine-tuning process with the pre-trained weights of *distilbert-base-uncased*.

3.1.5 Impact of Geographic Context Radius

The GeoContext2Vec algorithm requires defining a radius parameter that determines the geographic context around each POI. To validate our choice of 400 meters, we conducted a sensitivity analysis by varying the radius from 100 to 500 meters in increments of 100 meters, while keeping all other experimental settings fixed (e.g., $\omega = 0.8$, Word2Vec architecture). For each radius, we generated POI-type embeddings using the GeoContext2Vec algorithm and integrated them into the GETNext model following the methodology described in Section 3.2.4. The resulting recommendation performance on the New York City dataset is reported in Table 1.

As observed in Table 1, the variations in recommendation accuracy across different radii are minimal. For instance, Acc@1 ranges from 0.236 to 0.241, Acc@20 from 0.650 to 0.666, and MRR from 0.352 to 0.355. The differences

Table 1. Performance comparison for different geographic context radii. Metrics reported on the NYC dataset.

Radius (m)	Acc@1	Acc@20	MRR
100	0.239	0.658	0.355
200	0.238	0.666	0.353
300	0.240	0.650	0.353
400	0.241	0.654	0.352
500	0.236	0.662	0.352

are well within the typical variance expected from stochastic training processes, indicating that the model's performance is robust to the choice of radius within the tested range.

This stability can be attributed to the high density of POIs in the New York City dataset. With over 5,000 POIs concentrated in a relatively small geographic area, even a radius of 100 meters typically encompasses multiple geographic features and neighboring POIs. Increasing the radius further does not introduce substantially new contextual information, as the immediate vicinity already captures the most relevant spatial relationships. Consequently, the embeddings learned from different radii exhibit similar representational capacity, leading to comparable recommendation performance.

Based on these results and the practical consideration that a 400-meter radius provides a reasonable balance between capturing sufficient geographic context and avoiding excessive computational overhead, we adopted this value for all subsequent experiments involving GeoContext2Vec embeddings.

3.2 Model Architecture

The GETNext model [Yang *et al.*, 2022] aims to predict the next POI to be visited by a user, integrating context and sequential information. We selected this model as the foundation for our approach due to its state-of-the-art performance in POI recommendation and its inherent suitability for incorporating new types of embeddings. Specifically, GETNext's graph-based architecture allows us to seamlessly integrate different learned representations, enriching the model's understanding of each location. Furthermore, GETNext's utilization of a Transformer network enables it to effectively capture sequential patterns in user behavior, making it well-suited for incorporating the contextual information provided by our embeddings. The following subsections first describe the original GETNext architecture components, followed by a detailed explanation of our proposed extensions for integrating geospatially-enriched POI embeddings.

3.2.1 Original GETNext Architecture

The original GETNext architecture comprises three main components: Generic Movement Learning, Contextual Embedding, and Encoder-Decoder, as described below.

Generic Movement Learning: In the Generic Movement Learning layer (Figure 2), a Graph Convolutional Neural Network (GCN) analyzes historical check-in data to learn the embeddings of POIs. These embeddings reflect user movement patterns between POIs, considering aspects such as location, category, and visit frequency. Additionally, the layer includes a Transition Attention Map, which explicitly models the transition probabilities between POIs, reinforcing the influence of collective movement patterns on prediction.

Contextual Embedding: In the Contextual Embedding layer (Figure 2), three additional embeddings are generated, in addition to the POI embeddings generated in the Generic Movement Learning layer. They are: user embeddings, category embeddings, and time embeddings. These embeddings are designed to capture, respectively, user preferences, POI categorization, and temporal preferences. However, to extract deeper insights, it is necessary to combine these embeddings. The POI embeddings are then integrated with user-specific embeddings, derived from individual histories, to personalize recommendations. Additionally, temporal embeddings and category embeddings are also combined, recognizing that user preferences may change depending on the POI category and different times of the day. These integrations adjust the model to align with user preferences and the temporal context of POI visits.

Encoder-Decoder: The encoder-decoder layer, subdivided into a Transformer encoder and multi-layer perceptron decoders, processes sequences of check-ins and extracts relevant features using attention layers and fully connected networks. The decoder employs a multi-layer perceptron to make detailed predictions, including the next POI to be visited, its category, and the visit time.

3.2.2 Integration of Geospatial Features Embeddings

As our main extension to the original GETNext model, we incorporate diverse embeddings that encapsulate geographic and spatial attributes of POIs. These embeddings were generated for New York City using implementations of the GeoContext2Vec Silva *et al.* [2023] and Shortest Path Yao *et al.* [2017] algorithms, as detailed in Section 3.1.

The integration occurs within the Contextual Embedding layer (Figure 2), specifically following the concatenation of composite embeddings (POI + user, category + time). At this stage, we augment the representation by appending our pre-trained geospatial embeddings to the existing feature vector. This produces a final enriched representation that jointly encodes visitation dynamics, user preferences, temporal patterns, and geospatial characteristics.

By incorporating these geospatial features, our approach enhances the model's capacity to interpret the spatial context of visited locations, facilitating a more nuanced understanding of the relationship between geographic attributes and user behavior. Consequently, the model can learn to associate specific geographic features with visitation trends, potentially improving recommendation accuracy and contextual relevance.

4 Experiments and Results

This section describes the configuration and results of the experiments conducted for the task of next POI recommendation.

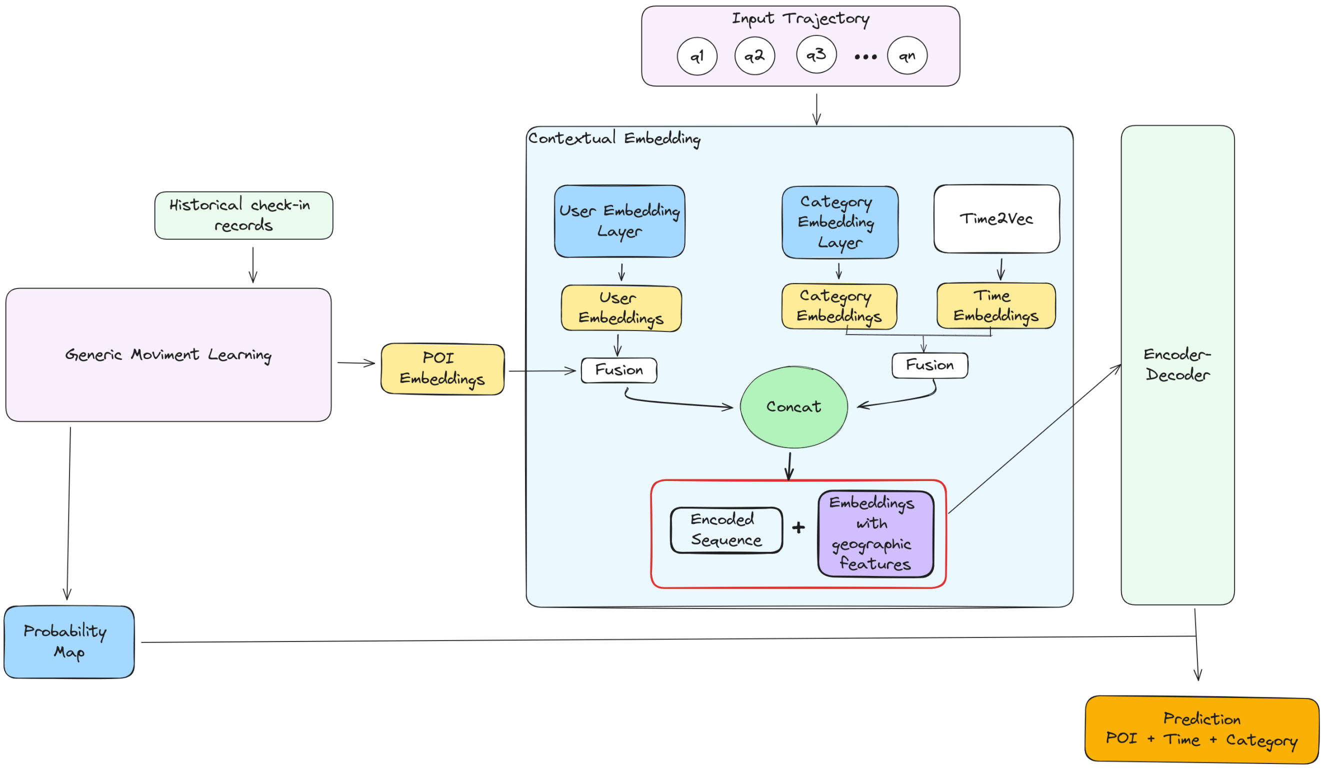

Figure 2. Model architecture based on GETNext with POI type embeddings that integrate geographic features

4.1 Dataset

We conducted our experiments on the public Foursquare New York data set[2], which consists of 1,075 users, 5,099 POIs, 318 categories and 104,074 check-ins collected between April 2012 and February 2013. The total sequence of check-ins for a user was divided into trajectories of 24-hour intervals, generating a total of 14,160 trajectories. Trajectories that contain only one check-in were removed from the dataset. The datasets were split into train, validation, and test sets following a chronological order, with 80% for training, 10% for validation, and 10% for testing. An important point is that if a user or POI did not appear in training but appeared in testing, we ignored it when calculating the metrics.

4.2 Experiment Settings

The model training and evaluation were conducted using the PyTorch and Gensim libraries, implementing both GETNext and Word2Vec architectures. The experiments were performed on a hardware setup consisting of an AMD Ryzen 5 3600X processor, 32GB of RAM, and an NVIDIA GeForce RTX 2080TI (11GB VRAM) GPU. For the GETNext model, the primary hyperparameters included 128-dimensional embeddings for POIs and users, while 32-dimensional embeddings were employed for temporal and categorical features, following the configuration established in the original work [Yang *et al.*, 2022]. In addition to the baseline embeddings, 140-dimensional GeoContext2Vec embeddings and 35-dimensional Shortest Path embeddings were incorporated. The original GeoContext2Vec embeddings, adapted for BERT architecture, exhibited 768-dimensional representations each, totaling 3072-dimensional vectors. However, due to hardware constraints, Principal Component Analysis (PCA) was applied for dimensionality reduction while preserving maximal information integrity.

The training phase of the GETNext model, enhanced with geospatially enriched POI embeddings, was executed over 200 epochs with a batch size of 20. The complete source code and implementation details are publicly available in a GitHub repository[3]. In the experiments, we implemented training phases of the model enhanced by the following combinations of algorithms:

- GeoContext2Vec
- Shortest Path
- GeoContext2Bert
- GeoContext2Vec + Shortest Path
- GeoContext2Bert + Shortest Path

4.3 Results

This study investigates the integration of POI-type embeddings enriched with geospatial features into a POI recommendation model, building upon the framework proposed by Yang *et al.* [2022], with the objective of enhancing recommendation performance. For evaluation, we employ two widely adopted metrics in recommendation systems: Accuracy at k (Acc@k) and Mean Reciprocal Rank (MRR). Acc@k measures whether the ground-truth POI—visited or interacted with by the user—appears within the top-k recommended POIs. However, as Acc@k does not account for the ranking order of relevant POIs within the top-k results, we additionally utilize MRR, which evaluates the positional

[2]https://sites.google.com/site/yangdingqi/home/foursquare-dataset

[3]https://github.com/diogoasilveira/POI-Recommendation

rank of the first relevant POI in the recommendation list. The computation of the metrics is performed as follows: upon the completion of each training epoch, a validation epoch is executed using the validation dataset, and the corresponding epoch-specific metrics are derived. Following the conclusion of the training procedure, the final metric is determined by computing the mean value across all epochs for the respective metric.

Table 2. Performance comparison in Accuracy@K and MRR in New York dataset

	Acc@1	Acc@20	MRR
GETNext	0.2226	0.6600	0.3429
GeoContext2Vec	0,2310	0,6464	0,3436
ShortestPath	0,2263	0,6599	0,3435
GeoContext2Bert	0.2133	0.6417	0.3305
GC+SP	0,2343	0,6505	0,3450
BERT+SP	0,2083	0,5935	0,3138

The primary results of our experiments are presented in Table 2. The baseline GETNext model, which we successfully reproduced under comparable conditions, achieved an Acc@1 of 0.2226, an Acc@20 of 0.6600, and an MRR of 0.3429. Our analysis focuses on how the integration of different geospatially-enriched embeddings alters these metrics.

The combined GC+SP embedding, which merges the contextual geographic features from GeoContext2Vec with the spatial proximity insights from the Shortest Path method, delivered the most robust performance. It achieved the highest Acc@1 (0.2343) and MRR (0.3450), representing relative improvements of 5.26% and 0.6% over the baseline, respectively. While its Acc@20 of 0.6505 was slightly below the baseline, the significant gain in Acc@1, a stricter metric for top-result accuracy, suggests that the fusion of these two geospatial perspectives is particularly effective at helping the model pinpoint the user's most likely next choice.

When examining the individual contributions, GeoContext2Vec alone proved to be a strong performer. It achieved the second-highest Acc@1 of 0.2310, a 3.77% improvement over the baseline, and the highest MRR of 0.3436. This indicates that the rich contextual information about a POI's surroundings (e.g., proximity to parks or water) is a powerful signal for personalizing the very next recommendation. In contrast, the Shortest Path embeddings, which model the transitional costs and paths between POIs, yielded a more moderate but positive gain. Its Acc@1 of 0.2263 represents a 1.66% improvement, and its Acc@20 (0.6599) remains very close to the baseline, suggesting that while path-based logic is relevant, it may be a secondary factor compared to the static contextual features captured by GeoContext2Vec.

A surprising and important finding is the underperformance of the BERT-based embeddings. GeoContext2Bert, despite being generated by a more advanced architecture, fell short of the baseline across all metrics, with an Acc@1 of 0.2133. Its combination with Shortest Path (BERT+SP) resulted in the lowest performance of all tested configurations, with a substantial degradation in Acc@1 (0.2083) and a particularly sharp drop in Acc@20 (0.5935). This suggests that the BERT fine-tuning process, as implemented, failed to effectively capture the geospatial relationships, and that the high-dimensional output may have introduced noise that hampered the downstream recommendation task, even after dimensionality reduction.

In summary, the results clearly show that not all geospatial embeddings are equally beneficial. Methods that capture clear, structured geographic context (GeoContext2Vec) and spatial relationships (Shortest Path) provide tangible improvements, especially when combined. Conversely, more complex, semantically-rich representations (BERT-based) may require more sophisticated integration strategies to be effective in this specific task. All models were trained for 200 epochs, with the final embedding dimensions varying due to the incorporation of geospatial features: GETNext (320), GeoContext2Vec (460), GeoContext2Bert (460), Shortest Path (355), GC+SP (495), and BERT+SP (495).

4.4 Discussion

This study demonstrates that integrating geospatially-enriched POI-type embeddings into the GETNext model improves recommendation performance by up to 5.6%, confirming that geographic context enhances next-POI prediction. The positive outcome suggests the model successfully leverages spatial features to better understand user preferences, answering our first research question in the affirmative.

Examining our second research question—the influence of spatial proximity—reveals nuanced findings. The Shortest Path embeddings, which model distance and connectivity between POIs, yielded a modest 1.66% improvement in Acc@1. However, GeoContext2Vec, which captures broader contextual features (proximity to parks, water bodies, transportation), achieved a more substantial 3.77% gain. This suggests users are influenced more by the character of a POI's surroundings than by pure distance metrics. The synergistic 5.26% improvement from their combination (GC+SP) indicates that both qualitative context and practical transition costs contribute to predictive accuracy.

Regarding our third research question on neural architectures, the results are striking. Word2Vec-based approaches consistently enhanced performance, while BERT-based variants underperformed significantly. We attribute this to a fundamental misalignment: BERT is optimized for semantic relationships from large text corpora, but fine-tuning on sparse geospatial co-occurrence patterns may cause catastrophic forgetting of linguistic priors. Additionally, the high dimensionality of BERT embeddings (768) necessitated PCA reduction to 35 dimensions due to hardware constraints, potentially discarding valuable spatial signals. This finding underscores that architectural complexity does not guarantee superior performance; alignment between the embedding method and data characteristics proves more critical.

Our strategy offers strong replicability, requiring only the addition of precomputed embeddings to existing recommendation frameworks. Models like the hypergraph-based approach proposed by Yan *et al*. [2023]—which currently outperforms GETNext but lacks geospatial features—could potentially benefit from this methodology.

Several limitations warrant acknowledgment. First, New York City's dense, homogeneous urban environment may underrepresent geographic feature diversity, potentially underestimating embedding efficacy in more heterogeneous settings. Second, reliance on a single Foursquare dataset introduces platform-specific biases, limiting generalizability. Third, dimensionality reduction applied to BERT embeddings, while necessary, may have discarded information critical for recommendation performance.

These findings advance our understanding of operationalizing geographic context in recommender systems, moving beyond simple coordinates toward semantically meaningful place representations. The contrasting performance between Word2Vec and BERT methods highlights that successful integration requires careful consideration of both feature characteristics and architectural compatibility.

5 Conclusion and Future Work

This study explored the integration of diverse POI-type embeddings into a POI recommendation model to enhance its ability to capture the influence of geospatial features on user preferences. Our experimental results demonstrated that incorporating these embeddings into the GETNext model led to a performance improvement of up to 5.26%, confirming that the model effectively leverages geospatial context to better predict user movements. These findings suggest that users' POI choices are indeed influenced by surrounding geographic features, reinforcing the importance of contextual information in recommendation systems. Additionally, our approach is highly replicable, as it only requires augmenting existing models with POI-type embeddings that encapsulate geospatial features. For instance, models like the one proposed by [Yan *et al.*, 2023], which currently outperform GETNext but do not incorporate geographic information, could potentially benefit from our methodology. However, extending such models with geospatial embeddings remains an avenue for future research.

Future work should explore the applicability of our approach in more diverse geographic settings, as well as investigate alternative methods for effectively combining geospatial and semantic embeddings. Additionally, testing the model on multiple datasets would help validate its robustness and generalizability. Despite these challenges, our findings contribute valuable insights into the role of geospatial context in POI recommendation systems, paving the way for more sophisticated and context-aware models in the future.

References

Bhuvanya, R. and Kavitha, M. (2023). A real-time e-commerce accessories recommender system by coupling deep learning and histogram features. *J. Intell. Fuzzy Syst.*, 45(1):1179–1193. DOI: 10.3233/JIFS-223754.

Davtalab, M. and Alesheikh, A. A. (2021). A POI recommendation approach integrating social spatio-temporal information into probabilistic matrix factorization. *Knowl. Inf. Syst.*, 63(1):65–85. DOI: 10.1007/S10115-020-01509-5.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding.

Feng, S., Cong, G., An, B., and Chee, Y. M. (2017). Poi2vec: Geographical latent representation for predicting future visitors. In Singh, S. and Markovitch, S., editors, *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA*, pages 102–108. AAAI Press. DOI: 10.1609/AAAI.V31I1.10500.

Feng, S., Tran, L. V., Cong, G., Chen, L., Li, J., and Li, F. (2020). HME: A hyperbolic metric embedding approach for next-poi recommendation. In Huang, J. X., Chang, Y., Cheng, X., Kamps, J., Murdock, V., Wen, J., and Liu, Y., editors, *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2020, Virtual Event, China, July 25-30, 2020*, pages 1429–1438. ACM. DOI: 10.1145/3397271.3401049.

Halder, S., Lim, K. H., Chan, J., and Zhang, X. (2021). Transformer-based multi-task learning for queuing time aware next POI recommendation. In Karlapalem, K., Cheng, H., Ramakrishnan, N., Agrawal, R. K., Reddy, P. K., Srivastava, J., and Chakraborty, T., editors, *Advances in Knowledge Discovery and Data Mining - 25th Pacific-Asia Conference, PAKDD 2021, Virtual Event, May 11-14, 2021, Proceedings, Part II*, volume 12713 of *Lecture Notes in Computer Science*, pages 510–523. Springer. DOI: 10.1007/978-3-030-75765-6_41.

Koren, Y., Bell, R. M., and Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37. DOI: 10.1109/MC.2009.263.

Kulkarni, A., Powell, L., Murphy, S., Rao, N., and Chu, S. L. (2023). Everyday-inspired movies: Towards the design of movie recommender systems based on everyday life through personal social media. In Abdelnour-Nocera, J. L., Lárusdóttir, M. K., Petrie, H., Piccinno, A., and Winckler, M., editors, *Human-Computer Interaction - INTERACT 2023 - 19th IFIP TC13 International Conference, York, UK, August 28 - September 1, 2023, Proceedings, Part III*, volume 14144 of *Lecture Notes in Computer Science*, pages 160–169. Springer. DOI: 10.1007/978-3-031-42286-7_9.

Leite, N., Campelo, C. E. C., and Silva, S. D. (2024). Leveraging geographic feature embeddings for enhanced location-based recommendation systems. In *Anais do XXXIX Simpósio Brasileiro de Bancos de Dados*, pages 354–366, Porto Alegre, RS, Brasil. SBC. DOI: 10.5753/sbbd.2024.240809.

Li, Z., Kim, J., Chiang, Y.-Y., and Chen, M. (2022). Spabert: A pretrained language model from geographic data for geo-entity representation.

Lika, B., Kolomvatsos, K., and Hadjiefthymiades, S. (2014). Facing the cold start problem in recommender systems. *Expert Syst. Appl.*, 41(4):2065–2073. DOI: 10.1016/J.ESWA.2013.09.005.

Liu, B., Su, Y., Zha, D., Gao, N., and Xiang, J. (2019). Carec: Content-aware point-of-interest recommendation via adaptive bayesian personalized ranking. *Aust. J. Intell. Inf. Process. Syst.*, 15(3):61–68.

Luo, Y., Liu, Q., and Liu, Z. (2021). STAN: spatio-temporal attention network for next location recommendation. In Leskovec, J., Grobelnik, M., Najork, M., Tang, J., and Zia, L., editors, *WWW '21: The Web Conference 2021, Virtual Event / Ljubljana, Slovenia, April 19-23, 2021*, pages 2177–2185. ACM / IW3C2. DOI: 10.1145/3442381.3449998.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient estimation of word representations in vector space. In Bengio, Y. and LeCun, Y., editors, *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*.

Qin, Y., Wang, Y., Sun, F., Ju, W., Hou, X., Wang, Z., Cheng, J., Lei, J., and Zhang, M. (2023). Disenpoi: Disentangling sequential and geographical influence for point-of-interest recommendation. In Chua, T., Lauw, H. W., Si, L., Terzi, E., and Tsaparas, P., editors, *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, WSDM 2023, Singapore, 27 February 2023 - 3 March 2023*, pages 508–516. ACM. DOI: 10.1145/3539597.3570408.

Saito, T. and Sato-Shimokawara, E. (2023). Music recommender system considering the variations in music selection criterion using an interactive genetic algorithm. In Saeed, K., Dvorský, J., Nishiuchi, N., and Fukumoto, M., editors, *Computer Information Systems and Industrial Management - 22nd International Conference, CISIM 2023, Tokyo, Japan, September 22-24, 2023, Proceedings*, volume 14164 of *Lecture Notes in Computer Science*, pages 382–393. Springer. DOI: 10.1007/978-3-031-42823-4_28.

Shi, M., Shen, D., Kou, Y., Nie, T., and Yu, G. (2021). Attentional memory network with correlation-based embedding for time-aware POI recommendation. *Knowl. Based Syst.*, 214:106747. DOI: 10.1016/J.KNOSYS.2021.106747.

Silva, S. D., Campelo, C. E. C., and de Oliveira, M. G. (2023). POI types characterization based on geographic feature embeddings. In Hong, J., Lanperne, M., Park, J. W., Cerný, T., and Shahriar, H., editors, *Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing, SAC 2023, Tallinn, Estonia, March 27-31, 2023*, pages 507–514. ACM. DOI: 10.1145/3555776.3577659.

Wang, X., Liu, X., Li, L., Chen, X., Liu, J., and Wu, H. (2021). Time-aware user modeling with check-in time prediction for next POI recommendation. In Chang, C. K., Daminai, E., Fan, J., Ghodous, P., Maximilien, M., Wang, Z., Ward, R., and Zhang, J., editors, *2021 IEEE International Conference on Web Services, ICWS 2021, Chicago, IL, USA, September 5-10, 2021*, pages 125–134. IEEE. DOI: 10.1109/ICWS53863.2021.00028.

Wang, X., Liu, Y., Zhou, X., Wang, X., and Leng, Z. (2022). A point-of-interest recommendation method exploiting sequential, category and geographical influence. *ISPRS Int. J. Geo Inf.*, 11(2):80. DOI: 10.3390/IJGI11020080.

Wang, Z., Li, H., and Rajagopal, R. (2020). Urban2vec: Incorporating street view imagery and pois for multimodal urban neighborhood embedding. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 1013–1020. AAAI Press. DOI: 10.1609/AAAI.V34I01.5450.

Yan, B., Janowicz, K., Mai, G., and Gao, S. (2017). From ITDL to place2vec: Reasoning about place type similarity and relatedness by learning embeddings from augmented spatial contexts. In Hoel, E. G., Newsam, S. D., Ravada, S., Tamassia, R., and Trajcevski, G., editors, *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS 2017, Redondo Beach, CA, USA, November 7-10, 2017*, pages 35:1–35:10. ACM. DOI: 10.1145/3139958.3140054.

Yan, X., Song, T., Jiao, Y., He, J., Wang, J., Li, R., and Chu, W. (2023). Spatio-temporal hypergraph learning for next POI recommendation. In Chen, H., Duh, W. E., Huang, H., Kato, M. P., Mothe, J., and Poblete, B., editors, *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2023, Taipei, Taiwan, July 23-27, 2023*, pages 403–412. ACM. DOI: 10.1145/3539618.3591770.

Yang, S., Liu, J., and Zhao, K. (2022). Getnext: Trajectory flow map enhanced transformer for next POI recommendation. In Amigó, E., Castells, P., Gonzalo, J., Carterette, B., Culpepper, J. S., and Kazai, G., editors, *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, pages 1144–1153. ACM. DOI: 10.1145/3477495.3531983.

Yao, Y., Li, X., Liu, X., Liu, P., Liang, Z., Zhang, J., and and, K. M. (2017). Sensing spatial distribution of urban land use by integrating points-of-interest and google word2vec model. *International Journal of Geographical Information Science*, 31(4):825–848. DOI: 10.1080/13658816.2016.1244608.

Yin, F., Liu, Y., Shen, Z., Chen, L., Shang, S., and Han, P. (2023). Next POI recommendation with dynamic graph and explicit dependency. In Williams, B., Chen, Y., and Neville, J., editors, *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence, IAAI 2023, Thirteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2023, Washington, DC, USA, February 7-14, 2023*, pages 4827–4834. AAAI Press. DOI: 10.1609/AAAI.V37I4.25608.

Yuan, Q., Cong, G., Ma, Z., Sun, A., and Magnenat-Thalmann, N. (2013). Time-aware point-of-interest recommendation. In Jones, G. J. F., Sheridan, P., Kelly, D., de Rijke, M., and Sakai, T., editors, *The 36th International ACM SIGIR conference on research and development in Information Retrieval, SIGIR '13, Dublin, Ireland - July 28 - August 01, 2013*, pages 363–372. ACM. DOI: 10.1145/2484028.2484030.

Zhang, H., Bai, W., Ding, J., and Jin, J. (2023). Time-aware POI recommendation based on multi-grained location grouping. In Shen, W., Barthès, J. A., Luo, J., Vivacqua, A. S., Schneider, D., Xie, C., Zhang, J., Zhu, H., Peng, K., and da Motta, C. L. R., editors, *26th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2023, Rio de Janeiro, Brazil, May 24-26, 2023*, pages 1796–1801. IEEE. DOI: 10.1109/CSCWD57460.2023.10152604.

Zhao, S., Zhao, T., Yang, H., Lyu, M. R., and King, I. (2016). STELLAR: spatial-temporal latent ranking for successive point-of-interest recommendation. In Schuurmans, D. and Wellman, M. P., editors, *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA*, pages 315–322. AAAI Press. DOI: 10.1609/AAAI.V30I1.9986.