A survey on deep learning based Point-of-Interest (POI) recommendations

Md. Ashraful Islam, Mir Mahathir Mohammad, Sarkar Snigdha Sarathi Das, Mohammed Eunus Ali

Department of Computer Science and Engineering (CSE), Bangladesh University of Engineering and Technology (BUET), Dhaka 1000, Bangladesh

ARTICLE INFO

ABSTRACT

Location-based Social Networks (LBSNs) enable users to socialize with friends and acquaintances by sharing their check-ins, opinions, photos, and reviews. A huge volume of data generated from LBSNs opens up a new avenue of research that gives birth to a new sub-field of recommendation systems, known as Point-of-Interest (POI) recommendation. A POI recommendation technique essentially exploits users' historical check-ins and other multi-modal information such as POI attributes and friendship network, to recommend the next set of POIs suitable for a user. A plethora of earlier works focus on traditional machine learning techniques that use hand-crafted features from the dataset. With the recent surge of deep learning research, we have witnessed a large variety of POI recommendation works utilizing different deep learning paradigms. These techniques largely vary in problem formulations, proposed techniques, used datasets and features, etc. To the best of our knowledge, this work is the first comprehensive survey of all major deep learning-based POI recommendation works. Our work categorizes and critically analyzes the recent POI recommendation works based on different deep learning paradigms and other relevant features. This review can be considered a cookbook for researchers or practitioners working in the area of POI recommendation.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

Location-based Social Networks (LBSNs) offer users a unique opportunity to socialize by sharing their check-ins, opinions, photos, and reviews. All these advantages paired with the wide availability of smartphones have dramatically increased the user base to billions-scale in these LBSNs platforms. Consequently, we have witnessed an explosion of rich multimodal spatio-temporal data collected from these platforms. The availability of this huge amount of data opens up new opportunities in Point-of-Interest (POI) recommendation, a vibrant independent sub-area in a recommendation system that has garnered significant attention from both user and business perspectives in recent years. A POI recommendation technique essentially exploits users' historical check-ins and other multimodal information to recommend the next set of POIs suitable for a user. As the size and modality of the data and the user expectation widely vary, this opportunity of having tons of multimodal data comes up with new challenges enticing the researchers to design novel techniques to better capture mobility patterns and other features (e.g., spatial, social, textual) to improve the POI recommendation performance.

Different types of POI recommendations have already been incorporated in different commercial systems that include Yelp,

Facebook, Twitter, and Google Map. For example, Yelp suggests the places that a user might be interested in based on his/her previous check-ins. Similarly, Google maps use the location history of a user to facilitate improved search results through location suggestions.

Earlier research works in POI recommendation primarily focused on feature engineering and conventional (non-deep learning) machine learning-based methods. Markov Chain based stochastic models have been explored extensively in this regard [1–9]. Due to the success of Matrix Factorization (MF [10]) based methods for recommendation systems in other domains, MF methods [2,11–20] have also been studied for better POI recommendation modeling. To achieve better performance than vanilla MF methods, Bayesian Personalized Ranking (BPR [21]) methods have been employed [22–28,8]. Other traditional approaches like Support Vector Machine (SVM) [29], Collaborative Filtering [30–35], Gaussian Modeling [36], Transitive Dissimilarity [37] have also been exploited in different works for personalized POI recommendation. One major shortcoming of all these approaches lies in feature engineering, and explicit feature engineering requires sufficient domain expertise. The increasing availability of data from different modalities like images, texts, and POI reviews makes these feature engineering tasks even more challenging as manually

crafting the relationship between these unstructured features is not a trivial task. Consequently, deep learning-based methods replaced most of those traditional techniques in recent years.

Deep learning methods like Convolutional Neural Networks (CNN) or Recurrent Neural Networks (RNN) provide a number of advantages in terms of automatic feature extraction eliminating the difficulties in handcrafted features extraction. Furthermore, deep learning-based methods excel in modeling complex relationships between structured and unstructured data, which lets us leverage multimodal data from different domains in POI recommendation. In the last few years, we have seen an unprecedented rise in the number of works leveraging deep learning in POI recommendation in all major venues (e.g., AAAI, IJCAI, SIGIR, CIKM, WWW, etc.). The use of different deep learning paradigms such as CNN [38,39], RNN [40–43], Long Short Term Memory (LSTM) [44–47], Gated Recurrent Unit (GRU) [48–50], and self-attention [51,52] have greatly boosted the performance of POI recommendation models. On top of that, state-of-the-art techniques from Natural Language Processing (NLP) have also been employed for complex modeling of human mobility in POI recommendation. Some recent works have leveraged graph embedding to enrich models with semantic geospatial information [53–55]. The wide varieties of deep learning-based POI recommendation techniques introduced in a short timespan necessitate a comprehensive review of these works (i) to demonstrate how different techniques have been used to handle different features, (ii) to identify the pros and cons of each model, and (iii) to propose a summary guideline for potential gaps and future research opportunities.

1.1. Previous surveys on POI recommendation

Several survey papers exist in the literature on POI recommendation. In an early work, Bao et al. [56] reviewed traditional (i.e., non-deep learning) POI recommendation methods. In a later work, Liu et al. [57] conducted an experimental evaluation of some of the then state-of-the-art traditional POI recommendation models. In another work, Zhao et al. [58] classified POI recommendation models in three taxonomies: influential factor-based, methodology-based, and task-based. All those surveys primarily focused on feature engineering-centric (i.e., non-deep learning) models. Later, Zheng et al. [59] did a comprehensive review of location prediction on the Twitter dataset, where they also acknowledged the uprise of deep learning-centric approaches. In another review of POI recommendation models, Liu et al. [60] also mentioned few neural network models. Since then, a large body of works in POI recommendation has been introduced leveraging different deep learning paradigms. In the last few years, researchers utilized RNN, LSTM, graph neural networks, attention networks in different ways to make use of different features resulting in significant performance uplift. Recently, Wang et al. [61] summarized a handful of deep learning-based models in the spatio-temporal domain. However, since this paper was a summary of the whole spatio-temporal domain, few POI recommendation models were discussed at a high-level. Another recent review of location prediction models [62] mostly discusses non-deep learning models with a very coarse focus on some deep learning models.

Large varieties of recent deep learning-based POI recommendation works largely vary w.r.t. problem formulations, proposed techniques, used datasets, used features, etc. There is no unified study to categorically discuss the pros and cons of different deep learning paradigms on POI recommendations. The wide variety of these techniques can easily puzzle someone willing to explore this field of POI recommendation. This survey work fills up the above gaps of the existing studies.

1.2. Contributions of our survey

In summary, to fill the gaps of existing surveys on POI recommendation techniques and to analyze the growing number of recent papers in this domain, we have made the following major contributions in this review paper.

- We enlist the major challenges that researchers have faced while building POI recommendation models (Section 3).
- We outline the features of all datasets used in this domain and discuss their strengths as well as their limitations (Section 5).
- We categorize the POI recommendation models based on different deep learning paradigms and compare their competitive (dis)advantages (Section 6).
- We identify different factors (i.e., social influence, sequential effect, etc.) that impact the POI recommendations and provide tabular analysis of each factor (Section 8).
- We present the comparison of all state-of-of-art techniques based on their performance metrics (Table 3).
- Finally, we provide comprehensive future recommendations for POI research (Section 9).

2. Problem definition

Point-of-Interest (POI) recommendation is a class of problems that suggest suitable future POIs for a user, given the historical check-in history of users and other associated data of an LBSN. Let $U = \{u_1, u_2, \ldots u_N\}$ be a set of N LBSN users and $P = \{p_1, p_2, \ldots p_M\}$ be a set of M POIs in the LBSN. Users may be linked to each other through a set of connections Ü $= \{\langle u_i, u_j \rangle | u_i, u_j \in U\}$. Each POI p is geo-coded by latitude x_p, longitude y_p, and a set of attributes W_p representing POI semantics. We first define the relevant terms and then formally define the problem as follows:

2.1. Definition 1 (Check-in)

A check-in $\rho^u_{t_i}$ indicates the POI checked-in by user u at time t_i.

2.2. Definition 2 (Check-in List)

Each user u is associated with a list of check-ins $C^u = \{\rho^u_{t_1}, \rho^u_{t_2}, \ldots, \rho^u_{t_T}\}$, where $\rho^u_{t_i}$ denotes a check-in record of user u at time t_i and $1 <= i <= T$.

2.3. Definition 3 (Next POI recommendation)

Given a check-in list C^u of a user u, the next POI recommendation refers to the prediction of the next POI at time t_{T+1}.

2.4. Definition 4 (Sequence of POI Recommendation)

Given a check-in list C^u of a user u, sequence of POI recommendation will recommend the next n POIs which is from t_{T+1} to t_{T+n}.

2.5. Definition 5 (Missing POI Check-in Identification):

Given the check-in list, $C^u = \{\rho^u_{t_1}, \rho^u_{t_2}, \ldots, \rho^u_{t_{m-1}}, \rho^u_{t_m}, \rho^u_{t_{m+1}}, \ldots, \rho^u_{t_T}\}$ of an user u where, the check-in $\rho^u_{t_m}$ at time t_m is missing, then missing POI check-in identification system will predict the POI for time t_m.

3. Challenges

In recent years, extensive research has been carried out in the field of POI recommendation. Researchers had to face a number of challenges in devising effective POI recommendation techniques based on users' past check-ins and other metadata. Major challenges in the area of POI recommendation are listed as follows.

3.1. Data sparsity

Data sparsity is one of the most crucial problems of building an effective POI recommendation system. From a user's point of view, a person travels to a very few locations in his/her lifetime compared to the sheer huge number of POIs available to be visited. Furthermore, exploring different POIs cost significantly higher than the exploration of different options in other fields of recommendation, which further exacerbates the data sparsity issue. Consequently, the relationships between users and POIs described in the datasets are genuinely sparse. Thus, it is always a challenge to design an effective POI recommendation system with sparse datasets.

3.2. Cold start

When a user joins an LBSN network, the lack of proper characterization of that user results in poor initial recommendation performance. Similarly, when a new POI is created to be explored, it heavily lags behind the already existing POIs in terms of recommendation due to the lack of historical trajectories associated with that POI. This problem is common to most of the present recommendation models, which is also known as a cold start problem. Eliminating the cold start problem requires in-depth insight into the field of POI recommendation research.

3.3. Capturing contextual information

Accurate POI recommendation relies on processing POI features as well as user reviews and other text data available from LBSN platforms. These textual data contain information about a person's experience of visiting a place. User data of POI recommendation contains reviews and text data that convey information about a person's experience of visiting a place. The processing of this semantic information requires additional different architectures to be added parallel to the main neural network of POI recommendation. The processing of user-POI embeddings with semantic word embedding layers is thus a quite challenging task for POI recommendation.

3.4. Change of user preferences

A user may show a significant amount of variation in his/her lifestyle. These changes may occur due to major life events such as aging, marriage, migrations, job changes etc. of an individual. These complex events are hard to detect and analyze from a dataset that contains a low amount of information per individual. As a result, an individual user may not get personalized recommendations from social media websites that use non-deep learning models because non-deep learning models do not capture the in-depth insight of user data quite well.

4. Network architecture preliminaries

In this section, we present the preliminary overview of different deep neural network paradigms that include Feed-Forward Network, Convolutional Neural Network, Recurrent Neural Network, Long-Short Term Memory, Gated Recurrent Unit, Attention Mechanisms, and Generative Adversarial Network.

4.1. Feed-Forward Network

Feed-Forward Networks (Fig. 1) are the most basic form of neural networks. Neural nodes are stacked up in layers where every node from a layer is connected to all the nodes in the next layer. The weighted connections combine the features of one layer and pass them to the subsequent layer through a nonlinear function (e.g. ReLU, Sigmoid, tanh, etc.). Stacking up layers of neurons dramatically increases the expressiveness of the network.

Although feed-forward networks can capture highly complex relationships within features, their overly high representational power usually causes overfitting training data resulting in poor generalization. Furthermore, as the number of layers increases, the sizes of the models dramatically increase making them harder to train and deploy. Most importantly, feed-forward networks have no explicit spatial and sequential feature handling capability, which limits their usage in spatio-temporal models.

4.2. Convolutional Neural Network

Convolutional Neural Networks (CNNs) are particularly suited for capturing spatial features from a given input. In CNN, convolutional filters, and pooling layers are systematically used to hierarchically process inputs. Subsets of the inputs are gradually channeled through convolutional filters, and pooling layers are used to scale down the transformed features. This process helps the CNN to gain spatial awareness while keeping the number of parameters significantly lower than feed-forward networks. To extract the spatial patterns in spatiotemporal data, CNNs are thus proven to be highly useful.

4.3. Recurrent Neural Network

Feed-forward networks or Convolutional Neural Networks do not consider any sequential or temporal dependency within the inputs. Recurrent Neural Network (RNN) solves this problem by taking a sequence of inputs and then learning the sequential pattern of the input sequence by using hidden states. We show the basic diagram of RNN in Fig. 2. In this figure x_i, y_i and h_i represent the input, output, and hidden state, respectively. We can see that besides input each RNN block uses a hidden state to produce output. Actually, the hidden states capture the context information of the input sequence which means capture the sequential pattern.

4.4. Long-Short Term Memory

RNN suffers from exploding Gradients and vanishing Gradients problems, as a result, can not capture long-term preferences. To solve the problems Long-Short Term Memory (LSTM) [63] is proposed. LSTM uses a gate mechanism and can capable of capturing long-term preferences. We can see an LSTM cell in Fig. 3. In addition to hidden state h_t which is used in RNN, every LSTM block has a cell state c_t. Also, the flow of the information among consecutive LSTM cells is controlled via three gates: (1) input gate, (2) forget gate, and (3) output gate.

4.5. Gated Recurrent Unit

LSTM resolves the problems of RNN but it has three gates so the training of an LSTM based model is slower and requires a large amount of training data. For solving these issues Gated Recurrent Unit (GRU) [64] is proposed. It uses only two gates, i.e., reset and
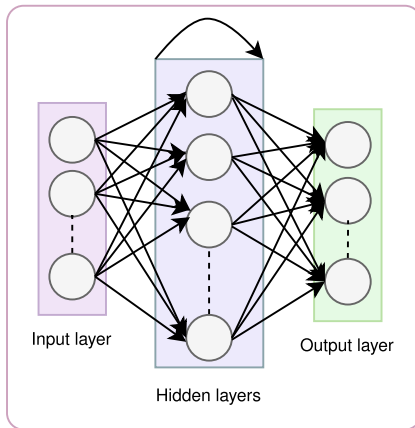
Fig. 1. Feed-forward network.

forget gates. Thus, the GRU-based model can be trained faster and performs better than LSTM when there is less training data.

4.6. Attention mechanism

Sequence models like RNNs or LSTMs process inputs by logical order of sequence. However, this scheme tends to lose features in longer sequences resulting in poor model performance. Attention mechanism [65] largely solves this shortcoming by mimicking a humanlike focus in salient input regions. Humans are prone to giving higher attention to key parts of the input, which in turn helps to break down a complex input into simpler parts that can easily be processed. While Seq2Seq models [66] have taken the advantage of this attention mechanism to improve performance, recent advancements in attention mechanisms have introduced self-attention mechanisms [67] that improves performance as well as allows parallel processing of inputs making them lucrative for various applications. The key idea here is that inputs are mapped to query, key, and value vectors. The outputs are calculated by taking the weighted sum of the value vectors where weights are determined by a function of query and key values. This technique has been highly effective in many areas of NLP research, which shows its potential in other domains involving sequential data.

4.7. Generative adversarial network

Generative adversarial networks (GANs) [68] are a special form of generative machine learning framework where two different
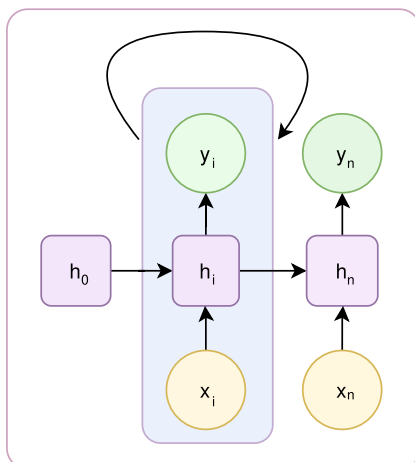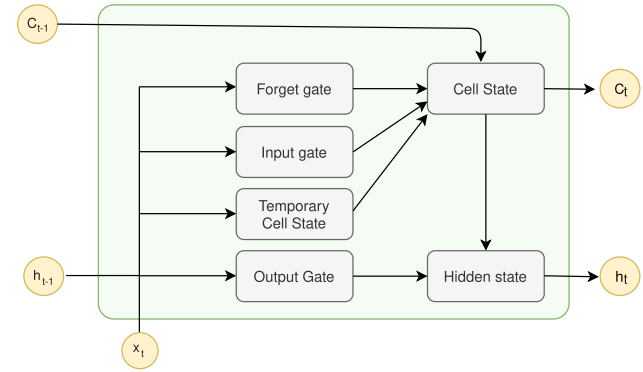

Fig. 2. RNN model.

Fig. 3. Basic LSTM Ccell.

networks compete against each other with different goals. One of them is called "Generator" and the other is called "Discriminator". The generator network generates candidates whereas the discriminator network tries to assess those candidates. The generator tries to "fool" the discriminator by creating novel sample candidates whereas the discriminator tries to distinguish those samples from the true data distribution. A generic GAN is shown in Fig. 4. In recent years, GANs have shown impressive performance in image synthesis, video game resolution upscaling, art generation, and so on.

5. Dataset description

Prior works in Point-of-Interest recommendation have used check-in data collected from a wide range of LBSNs that include Foursquare, Gowalla, Yelp, Twitter, Facebook, Brightkite, Instagram, WeChat, and Baidu Map. Most of these datasets consist of tabular data that records the user-POI and the user-user relationship in LBSNs. User-POI data typically contains user check-in information including timestamps, location, and semantic features. Here POI Semantic features include the categories of the POIs and tags included in user LBSN posts, creation date of the POIs, geolocations (latitude, longitude), check-in counts, number of users checked-in, radius, etc. On the other hand, user semantic features contain the number of posts, friends, check-ins, etc. In order to keep the social influence in context, datasets like Foursquare, Gowalla, Weeplace also contain user-user relationships as a many-to-many schema, where each user is connected to all his friends. Because of all these data and a huge number of check-ins, some of these datasets become exceedingly large. Consequently, most of the prior works focused on a specific region or a country to keep the size tractable. A brief discussion on these LBSN datasets is given below:

5.1. Foursquare

Founded in 2009, Foursquare[1] has worked with world-wide collection and distribution of location data to facilitate technological corporations and brands. Most of the POI recommendation models discussed in this review use the datasets of Foursquare from a time range of 2010 to 2014. The datasets contain check-in data collected mostly from the USA and Tokyo. This dataset also contains the list of all friends of each user in the LBSN.

5.2. Gowalla

Gowalla is a location-based social media platform dedicated to location check-ins. The platform was founded in 2007 and acquired
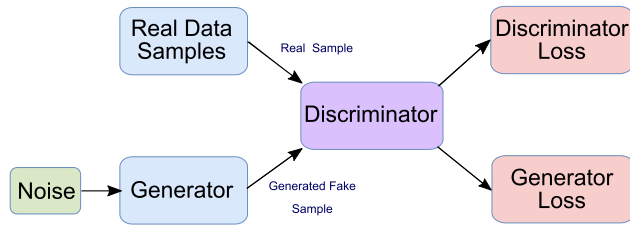
---

[1] https://foursquare.com/.

Fig. 4. Generative adversarial network.

by Facebook in 2012. Gowalla was primarily a mobile application that allowed users to check into locations that they visited using their mobile devices. The datasets from the functioning period of Gowalla were available via the Gowalla API and currently, there are no official distributors for the datasets. Gowalla is the second most used dataset in the POI recommendation models discussed in this paper. Most prior works discussed have used check-in data from February 2009 to October 2010. Like Foursquare, the Gowalla dataset also contains the list of friends of every user in the dataset. Besides, a detailed description of each POI and user profiles are also available in this dataset.

5.3. Brightkite

Brightkite was a location-based social media network that launched in 2007 and got dissolved in 2012. The platform provided the ability to check-in through text messaging or a mobile application after visiting a location. Datasets of Brightkite are no longer officially available but still can be accessed from various research archives around the globe.

5.4. Yelp

Yelp[2] is another popular LBSN platform. As users tend to check-in different business locations, Yelp provides reviews and ratings from customers who shared their own experiences both for personal and research purposes. Yelp was founded in 2004 and is still operational as a reviewing company for business establishments. POI recommendation models are greatly benefited from the textual reviews of Yelp because reviews provide semantic information of POIs.

5.5. Weeplaces

WeePlaces is a service that visually maps users' check-ins on location-based services. Weeplaces has been integrated with Gowalla and Facebook, giving users the ability to visualize where users have announced their locations to friends across Foursquare, Gowalla, and Facebook Places. Like the Foursquare and Gowalla dataset, the Weeplace dataset also contains the list of friends connected to a user which can be leveraged for capturing social influence in POI recommendation.

5.6. Instagram

Instagram[3] is a social media platform initially released in 2010 and currently owned by Facebook. Instagram allows users to post and share photos and videos online. Users can browse other users' content by tags and locations. Instagram dataset primarily contains user check-in data which has been used in some recent works on POI recommendation.

5.7. Twitter

Twitter[4] is a highly popular social media platform where a user can post, see, and share short messages known as 'tweets. Twitter was founded in 2006 and is currently the most popular micro-blogging service around the world. The check-in functionality of Twitter enables users to record visits to locations. The datasets are available through Twitter public API.

5.8. Other datasets

Some other used datasets are collected from WeChat,[5] Baidu Maps,[6] Facebook.[7] Functionalities of WeChat include texting, voice messaging, video conferencing, and location sharing. Baidu Maps provide street maps and views, satellite views of terrains, and route planners for traveling. Facebook is currently the largest social media site where users can post text, photos, and multimedia to share information about themselves. Posts can include check-in information about a visited place and thus datasets of Facebook are officially available for POI prediction research.

For detailed information on some of the most used datasets used in POI recommendation, see Table 1. And to see the statistics of the datasets associated with each of the discussed papers, see Table 4.

6. POI recommendation models

Thanks to the astounding growth of the user base in LBSNs, the amount of check-in data collected from these platforms have increased rapidly in recent years. This large volume of data has fueled the adaptation of deep learning techniques in the field of POI recommendation. While earlier works used conventional machine learning models, recent deep learning-based models have mostly replaced them due to the significantly higher performance with abundant potential to further improve the performance. Thus, in this survey, we mainly focus on Deep Neural Network (DNN) based POI recommendations. We categorize all of the proposed models used thus far into six major categories. They are RNN based models (Section 6.1), LSTM models (Section 6.2), GRU models (Section 6.3), Graph Embedding models (Section 6.4) GAN models (Section 6.5), and other models (Section 6.6). These categories are described in the following subsections. We briefly highlight these models in Table 2, and the dataset used in each model are depicted in Table 4. A concise summary of these models, evaluation metrics, and their performance across different datasets are given in Table 3.

6.1. RNN based models

Recurrent Neural Networks (RNN) are renowned for their high effectiveness in NLP problems. As POI recommendation problems show similar properties that resemble NLP tasks, many recent POI recommendation models use RNN as their base architecture. In this subsection, we discuss the basic RNN based POI recommendation models that can map a POI-sequence to another POI-sequence (successive POI recommendation), or only one POI (next POI recommendation). Here we discuss the next POI recommendation only.

Liu et al. [40] proposed a model, called **Spatial-Temporal Recurrent Neural Networks (ST-RNN)**, for POI recommendations. The ST-RNN model extends the RNN model for capturing spatial

---

Table 1
Summary of datasets.

Dataset name	Attributes	Description of the table
Foursquare	`userID, Time(GMT), VenueId, VenueName, VenueLocation, VenueCategory`	Describes the user-POI relationship. Contains all user-checkins with date-time of checkin; name, location and type of POI
	`userID, friendID`	Describes the user-user relationship. Contains the list of all friends of each user in the LBSN
Gowalla	`userId, timestamp, latitude, longitude`	Describes the user-POI relationship. Contains all check-in information of each user such as location and time of the visit
	`userId, friendId`	Describes the user-user relationship. Contains the list of all friends of each user in the LBSN
	`id, name, created_at, lng, lat, photos_count, checkins_count, users_count, radius_meters, highlights_count, items_count, max_items_count, spot_categories, city_state`	Description of each POI of the LBSN including the counts of user, checkins, photos, items and highlights. The name, city, date-time, category and radius on the map of each POI are also included.
	`id, bookmarked_spots_count, challenge_pin_count, country_pin_count, highlights_count, items_count, photos_count, pins_count, province_pin_count, region_pin_count, state_pin_count, trips_count, friends_count, stamps_count, checkin_num, places_num`	Details of the profile of each user of the LBSN
Brightkite	`user, check in time, latitude, longitude, location id`	Describes the user-POI relationship. Each row contains time and location information of check-ins made by one user
	`userid1, userid2`	Describes the user-user relationship. The friendship network of Brightkite users is described in this table.
Yelp	`business_id, name, address, city, state, postal_code, latitude, longitude, stars, review_count, is_open, attributes, categories, hours`	Contains POI business data including location data, attributes, and categories. Attributes include whether the restaurants accept takeouts and have business parkings
	`review_id, user_id, business_id, stars, date, text, useful, funny, cool`	Contains full review text data including the user_id that wrote the review and the business_id the review is written for.
	`user_id, name, review_count, yelping_since, friends, useful, funny, cool, fans, elite, average_stars, compliment_hot, compliment_more, compliment_profile, compliment_cute, compliment_list, compliment_note, compliment_plain, compliment_cool, compliment_funny, compliment_writer, compliment_photos`	User data including the user's friend mapping and all the metadata associated with the user.
	`business_id, date`	Checkins on a POI of all the users
	`text, date, compliment_count, business_id, user_id`	Tips are written by a user on a POI business location. Tips are shorter than reviews and tend to convey quick suggestions.
	`photo_id, business_id, caption, label`	Contains photo data including the caption and classification
Weeplace	`userid, placeid, datetime, lat, lon, city, category`	Describes the user-POI relationship. Each row contains check-in information of a user, date and time of the visit. The row also contains the location, category, subcategory, and name of the city of the POI.
	`userid1, userid2`	Describes the user-user relationship. Contains the list of all friends of each user in the LBSN
Instagram	`user_id, latitude, longitude, timestamp`	Details of the user-POI relationship. Each row contains check-in information of a user, the date and time of the visit, and the location of the POI
Twitter	`userID, tweetID, latitude, longitude, time, placeID, contentInfo`	Contains information about a tweet having a check-in. Each row represents a tweet, a user, location and identifier of the POI, and tags associated with the POI

and temporal effects. This model adds a time-specific and distance-specific transition matrix for capturing temporal cyclic effect and geographical influence, respectively. The model also applies linear interpolation for the training of the transition matrix. In another recent work, Yang et al. [41] proposed a model called **Flashback** in which they use Basic RNN. The model uses sparse user mobility data by focusing on rich spatio-temporal contexts and doing flash-backs on hidden states in RNNs. Furthermore, the model uses the weighted average of historical hidden states for better capturing the spatio-temporal effects. Additionally, the paper also uses user embedding for considering user preferences. Zhao et al. [42] proposed **Adaptive Sequence Partitioner with Power-law Attention (ASPPA)** model to learn the latent structures of the check-in sequences. The idea is a blend of Adaptive Sequence Partitioner (ASP) for texts by Griffiths et al. [69] and the stacked RNN architecture of El et al. [70]. The paper aims to automatically detect and identify each semantic subsequence of POIs and discover their sequential patterns. The model is designed to be a stacked RNN and it adopts a binary boundary detector to control the pattern of cell update. This model uses the Power-law Attention (PA) mechanism to integrate spatial and temporal contexts of each check-in into the model. The output layer consists of two fully con-

nected layers and a drop-out layer. Since most of the POI recommendations are designed upon a cloud-based paradigm, there are many disadvantages including privacy concerns. To alleviate these problems, Wang et al. [43] proposed **Light Location Recommender System (LLRec)**. Here the authors introduced teacher and student models. The teacher model is deployed on the server whereas the student model is deployed on the user mobile devices. The student model will fetch the pre-trained model from the server. And a user can get the next POI recommendation by using the pre-trained model without sending data to the teacher model. However, the storage and computational capabilities in mobile devices are very limited. Consequently, the student model must be efficient, lightweight, and fast. For getting a lightweight and fast model, LLRec uses FastGRNN [71]. The parameters of this model are further compressed by using the tensor-train format [72]. Furthermore, the knowledge distillation framework [73] is used to improve prediction quality with very limited data. Time-specific and distance-specific transition matrices in vanilla FastGRNN are used for capturing spatio-temporal correlations between two adjacent check-ins. On the other hand, a powerful and computationally intensive teacher model can be deployed in the cloud. In addition, the teacher model uses an attention mechanism to learn user pref-

Table 2
Categorization of papers.

Category	Subcategory	Year	Reference
Basic RNN	-	2016	ST-RNN[40]
		2018	MCI-DNN[76]
		2020	Flashback[41], ASPPA[42], LLRec[43], DRLM[75]
LSTM	Basic LSTM	2018	TMCA[45], SLSTM[85]
		2019	LSPL[77], ASTEN[79], ATST-LSTM[82], MMR[83], SGBA[84]
		2020	PLSPL[78], LSTPM[44], iMTL[46], CatDM[47], ARNN[80], STAR[81]
	Bi-LSTM	2020	GT-HAN[86], t-LocPred[88], CAPRE[89]
	Modified LSTM	2018	HST-LSTM[90]
		2020	STGN[91]
	Self-Attention	2020	GeoSAN[51], SANST[52]
GRU	-	2018	DeepMove[48], CARA[49]
		2019	MGRU[50]
Graph Embedding	-	2016	GE[94]
		2017	STA[53]
		2018	JLGE[96]
		2019	RELINE[54]
		2020	DYSTAL[55], HCT[100], UP2VEC[103], HMRM[102]
GAN	-	2019	Geo-ALM[104], APOIR[105]
		2020	AdattTUL[106]
Others	-	2017	Geo-Teaser[26], LCE[108]
		2018	CAPE[109], ST-DME[110]
		2019	Bi-STDDP[111], TEMN[112], SSSER [38]
		2020	HME[113], MPR[114], PGIM[115], CEM [39]

erences and converts textual content of POIs into low dimensional embeddings via Word2Vec [74]. Besides, Huang et al. [75] and Liao et al. [76] also extensively used the RNN architecture in their POI recommendation model.

6.2. LSTM models

Since RNN models cannot capture long-term dependencies, Long-Short Term Memory (LSTM) [63] has been extensively used in the recent works in POI recommendations. Most of the models use the basic unaltered LSTM for their predictions. Some of the models modify the basic LSTM model or use the bidirectional variant for better capturing the POI domain-specific factors. To further improve long-term dependency modeling, attention mechanisms have also been used alongside LSTM. A detailed discussion of different kinds of LSTM based POI recommendation models are given below.

6.2.1. Basic LSTM

Here we discuss the state-of-the-art models that use the basic LSTM model in their novel POI recommendation model. We start our discussion on the Basic LSTM model with the model proposed by Li et al. [45] called **Temporal and Multi-level Context Attention (TMCA)** that uses LSTM based encoder-decoder network and three types of attention: multi-level context attention (micro, macro) and temporal attention. The paper introduces two attention mechanisms to select relevant historical and contextual factors. The model also uses embedding to incorporate heterogeneous contextual factors in a unified way.

For capturing both long-term and short-term preference Wu et al. [77] proposed **Long- and short-term preference learning model (LSPL)**. LSPL has two modules i.e., (1) the long-term module consists of embedding layer and attention layer and captures the long-term preference of a user by learning contextual features of POIs; (2) the short-term module uses two separate LSTM modules: one for location level and other for category level and captures sequential behavior of a user. Then all of them are combined to predict the next POI.

The authors of LSPL [77] extended their work by proposing **Personalized Long- and Short-term Preference Learning (PLSPL)**

[78] model. PLSPL adds a user-based linear combination unit with their existing LSPL model, which captures the personalized preferences for different users by learning the personalized weights over long- and short-term modules.

Based on embedding, LSTM, and an attention mechanism, Doan et al. [79] proposed **Attentive Spatio TEmporal Neural (ASTEN)** model. ASTEN embeds the POIs and represents a check-in efficiently. LSTM uses the POI and check-in representation and along with the attention mechanism which captures the sequential, temporal, and geographical influence. The model addresses the noise in user-trajectory data by attention mechanism.

Sun et al. [44] proposed **Long and Short-Term Preference Modeling (LSTPM)** by modifying basic LSTM models. LSTPM divides all check-ins into several trajectories. The model actually develops three modules including the long-term preference modeling, the short-term preference modeling, and the prediction module. Long-term preference modeling uses all the trajectories, short-term preference modeling uses the last trajectory and combinedly predicts the next POI. Another important aspect of the next POI recommendation is that the next POI does not depend only on the recent check-in; however, it can depend on any earlier check-in. But RNN/LSTM based approaches have the drawback of being unable to model the relations between two nonconsecutive POIs. For capturing this effect the model uses a **geo-dilated LSTM** scheme along with basic LSTM in short-term preference modeling.

Zhang et al. [46] proposed **Interactive multi-task learning (iMTL)**, which uses a two-channel encoder and a task-specific decoder. The two-channel encoder (temporal-aware activity and spatial-aware location preference encoders) aims to capture the sequential correlations of activities and location preferences. The representations encoded by the LSTM are utilized in the task-specific decoder to interactively perform the prediction tasks. A novel contribution of this paper is that they focus on the collective POIs. Suppose, l_1 is a building and each building contains many individual apartments i.e., l_2, l_3, l_4, l_5, then the model denotes l_1 as collective POI and l_2, l_3, l_4, l_5 as individual POIs. This paper proposes a fuzzy characterization strategy for better prediction of individual POI from a collective POI.

Most of the POI recommendation models only predict the location of POI ignoring the timestamp but Yu et al. [47] proposed

Md. Ashraful Islam, Mir Mahathir Mohammad, Sarkar Snigdha Sarathi Das et al.

Neurocomputing 472 (2022) 306–325

Table 3
Model summery and performance of the state-of-the-art method.

Model, Venue year [ref]	Method summary	Performance
ST-RNN, *AAAI 2016* [40]	Use time-specific and distance-specific transition matrix for capturing temporal cyclic effect and geographical influence respectively.	Gowalla: Rec@5[1]=0.1524, Rec@10=0.2714. GTD: Rec@5=0.4986, Rec@10=0.6812
Flashback, *IJCAI 2020* [41]	Model sparse user mobility data by doing flashbacks on hidden states in RNNs and uses the weighted average of historical hidden states for better capturing spatio-temporal effects.	Foursquare: Acc[2]@5=0.5399, Acc@10=0.6236. Gowalla: Acc@5=0.2754, Acc@10=0.3479.
ASPPA, *IJCAI 2020* [42]	Automatically identify the semantic subsequnce of POIs and discovers their sequential patterns by hierarchically learning the latent structure from check-in list and power-law attention mechanism.	Foursquare (US): Acc@10=0.3371, Acc@20=0.3950. Gowalla: Acc@10=0.2947, Acc@20=0.3573.
LLRec (Teacher), *WWW 2020* [43]	Capture long-term, short-term preferences, textual feature of POIs and complex dependencies among user preferences by using embedding, recurrent component and attention mechanism.	Foursquare: Acc@10=0.3542, Acc@20=0.4594 Gowalla: Acc@10=0.3874, Acc@20=0.4781
TMCA, *ICDM 2018* [45]	Capture complex spatial and temporal dependencies among historical check-in activities by using LSTM based encoder-decoder model, attention mechanism and embedding method.	Gowalla: Rec@5=0.21926, Rec@10=0.27725. Foursquare: Rec@5=0.02870, Rec@10=0.04809.
LSPL, *CIKM 2019* [77]	Capture both sequential and contextual information via long-term and short-term preference learning.	Foursquare (NYC): Prec[3]@10=0.3901, Prec@20=0.4461 Foursquare (TKY): Prec@10=0.3986, Prec@20=0.4596.
PLSPL, *TKDE 2020* [78]	Extend their previous work LSPL [77] by introducing user-based linear combination unit which better captures user preferences.	Foursquare (NYC): Prec@10=0.3953, Prec@20=0.4475 Foursquare (TKY): Prec@10=0.4020, Prec@20=0.4664.
ASTEN, *PAKDD 2019* [79]	Capture the sequential, temporal and geographical influence by using LSTM and attention mechanism.	Foursquare (US): Rec@5=0.328, Rec@10=0.414 Foursquare (EU): Rec@5=0.281, Rec@10=0.35 Gowalla: Rec@5=0.152, Rec@10=0.266
LSTPM, *AAAI 2020* [44]	Capture long-term preference modeling by using a non-local network and short-term preference modeling by using geo-dialated LSTM.	Foursquare (NY): Rec@5=0.3372, Rec@10=0.4091 Gowalla: Rec@5=0.2021, Rec@10=0.2510
iMTL, *IJCAI 2020* [46]	Use a two-channel encoder and a task-specific decoder for capturing the sequential correlations of activities and location preferences.	POI Prediction: Foursquare (CLT): Rec@10=0.0534, Map[4]@10=0.0238 Foursquare (CAL): Rec@10=0.0691, Map@10=0.0443 Foursquare (PHO): Rec@10=0.0769, Map@10=0.0352
CatDM, *WWW 2020* [47]	capture temporal influence, geographical influence and overcome data sparsity by using two LSTM based deep encoder, two filter, metric embedding and attention mechanism.	Foursquare (NYC): Rec@5=0.2407, Rec@10=0.3113 Foursquare (TKY): Rec@5=0.2148, Rec@10=0.2739.
ARNN, *AAAI 2020* [43]	Capture data sparsity by using new concept called transition regularity. Also capture sequential, spatial, temporal, semantic influence by using embedding, knowledge graph, LSTM and attention mechanism.	Foursquare (NY): Acc@10=0.4162, Acc@20=0.4393 Foursquare (TK): Acc@10=0.4285, Acc@20=0.4864 Gowalla (SF): Acc@10=0.2336, Acc@20=0.2530
GT-HAN, *Neurocomputing 2020* [86]	Capture great variation in geographical co-influence across POIs, temporal dependency and sequence dependency in check-in list by using embedding layer, Bi-LSTM and attention mechanism.	Foursquare: AUC8=0.9661, acc@5: 0.13-0.15, acc@10: 0.17-0.19, acc@20: 0.23-0.25 (depending on latent dimensionality)
t-LocPred, *TKDE 2020* [88]	capture a users' coarse-grained spatiotemporal movement pattern by using CNN and ConvLSTM and fine-grained POI check-in information by using spatial-aware memory-augmented LSTM with time-aware attention.	Gowalla: MRR[5]=0.247 (C=6, all), Weeplaces: MRR=0.277 (C=6, all), Brightkite: MRR=0.388 (C=4, all)
CAPRE, *SDM 2020* [89]	Capture the various perspectives of user about POIs along with content-aware and geographical user behavior pattern by using character-level CNN, multi-head attention, Bi-LSTM and MLP.	Foursquare: Rec@5=0.1724, Rec@10=0.2084 Instagram: Rec@5=0.2934, Rec@10=0.3588
HST-LSTM, *IJCAI 2018* [90]	Propose a new Spatial-Temporal LSTM (ST-LSTM) model for resolving data sparsity and also use hierarchical model using st-LSTM to encode the periodicity of people's movement.	Baidu Map: Acc@10=0.4847, Acc@20=0.5657
STGCN, *AAAI 2019* [91]	Modify the basic LSTM model slightly by introducing new gates and cell to capture short-term and long-term preference easily.	Foursquare (CA): Acc@5=0.1308, Acc@10=0.1612. Foursquare (SIN): Acc@5=0.2737, Acc@10=0.3017. Gowalla: Acc@5=0.1644, Acc@10=0.2020. Brightkite: Acc@5=0.4953, Acc@10=0.5231.
GeoSAN, *KDD 2020* [51]	Resolve the sparsity issue by introducing a new loss function and represent the hierarchical gridding of each GPS point with a self-attention based geography encoder for better use of geographical information.	Foursquare: Acc@5=0.3735, Acc@10=0.4867. Gowalla: Acc@5=0.4951, Acc@10=0.6028. Brightkite: Acc@5=0.5258, Acc@10=0.6425.
SANST, *arXiv 2020* [52]	Capture the spatial-temporal and sequential patterns by using embedding, self-attention network (transformer network), Bi-LSTM.	Gowalla: Acc@10=0.2273, Los Angeles: Acc@10=0.3941, Singapore: Acc@10=0.2417
DeepMove, *WWW 2018* [48]	Capture the complex dependencies and multi-level periodicity nature of human by using embedding, GRU and attention mechanism.	Foursquare (NY): Rec@5=0.3372, Rec@10=0.4091. Gowalla: Rec@5=0.2021, Rec@10=0.2510
CARA, *SIGIR 2018* [49]	Capture the different types of impact of different contextual information by using embedding, GRU and two gating mechanism.	Foursquare: Acc@10=0.8851, Yelp: Acc@10=0.5587, Brightkite: Acc@10=0.7385
MGRU, *JAIHC 2019* [50]	Capture dynamic and transition context using Multi-GRU (Two special gate are added with GRU).	Foursquare: Rec@10=0.9214, Rec@15=0.9214 Gowalla: Rec@10=0.8512, Rec@15=0.8765
GE, *CIKM 2016* [94]	Capture data sparsity, context awareness, cold start, dynamic of personal preference by using the embedding of four graphs into a shared low dimensional space.	Foursquare: Acc@10=0.372, Acc@20=0.435. Gowalla: Acc@10=0.462, Acc@20=0.533.
STA, *arXiv 2017* [53]	This paper generalizes the knowledge Graph Embedding and takes location and time as a spatiotemporal pair for connecting users and POIs.	Foursquare: Acc@10=0.439, Acc@20=0.486. Gowalla: Acc@10=0.488, Acc@20=0.540.
JLGE, *DSAA 2018* [96]	Jointly learn the embeddings of the users and the POIs into the same latent space from the six informational graphs using LINE model[97].	Foursquare: Acc@10=0.410, Acc@20=0.462. Weeplaces: Acc@10=0.488, Acc@20=0.536.
RELINE, *arXiv 2019* [54]	This paper extends the previously discussed paper [96] by adding two new networks: i.e., stay points and routes.	Foursquare: Acc@10=0.410, Acc@20=0.462. Weeplaces: Acc@10=0.488, Acc@20=0.536. Gowalla: Acc@10=0.518,

Table 3 (*continued*)

Model, Venue year [ref]	Method summary	Performance
Geo-Teaser, *WWW 2017* [26]	Use Skip-Gram model for temporal POI embedding and Bayesian Personalized Ranking for pairwise ranking of POIs. A unified framework combines the temporal POI embedding and pairwise ranking model.	Acc@20=0.556. Foursquare: Prec@5=0.13, Prec@10=0.1, Rec@5=0.15, Rec@10=0.2 Gowalla: Prec@5=0.16, Prec@10=0.13, Rec@5=0.07, Rec@10=0.12
DYSTAL, *Information Processing and Management 2020* [55]	Capture complex spatio-temporal patterns of visiting behaviors by jointly learning the effects of users' social relationships, textual reviews, and POIs' geographical proximity using a network embedding method and dynamic factor graph model.	Foursquare (SIN): Acc@10=0.232, Rec@10=0.152, Yelp: Acc@10=0.206, Rec@10=0.098
HCT, *Information Sciences, Elsevier 2020* [100]	Utilize Skip-Gram model to model the categorical transitions at different layers of categorical hierarchies as well as the hierarchical dependencies between POIs and categories	Foursquare(SIN): Prec@5=0.613 Rec@5=0.0403 Foursquare(NYC): Prec@5=0.0585, Rec@5=0.0352 Foursquare(LA): Prec@5=0.0653, Rec@5=0.0305
APOIR, *2019* [105]	Use matrix factorization and GRU to learn user preferences and train two competitive component: recommender and discriminator to generate prediction.	Yelp: Prec@5=0.1, Rec@5=0.16, MAP@5=0.233, NDCG[6]@5=0.094
CAPE, *IJCAI 2018* [109]	Use text content layer and check-in content layer for embeds the POIs and generates their own dataset.	With STELLAR: Rec@5=0.2384, Rec@10=0.2989 With LSTM: Rec@5=0.2412, Rec@10=0.3054 With GRU: Rec@5=0.2433, Rec@10=0.3079 With ST-RNN: Rec@5=0.2239, Rec@10=0.2601
TEMN, *KDD 2019* [112]	Capture both neighbourhood-based and global preferences by using a combination of supervised and unsupervised learning.	WeChat (GPR): TEMN (GPR): Acc@5=0.70389, Acc@10=0.81752. TEMN (CPR): Acc@5=0.72876, Acc@10=0.83398.
HME, *SIGIR 2020* [113]	Capture POI sequential transitions, geographical, semantic and user preferences by using hyperbolic metric embedding along with Poincaré ball and Einstein midpoint aggregation method.	Foursquare (NYC): Rec@5=0.0962, Rec@10=0.1371 Foursquare (TKY): Rec@5=0.1527, Rec@10=0.2172. Gowalla (Houston): Rec@5=0.1533, Rec@10=0.2318

[1] Recall@K is the presence of the correct POI among the top K recommended POIs [44].

[2] Acc@k is 1 if the visited POI appears in the set of top-K recommendation POIs and 0 otherwise [91]. The overall Acc@K is calculated as the average value of all testing instances. Also known as Accouracy@K or Hit Rate@k or Hit Ratio@k or HR@k.

[3] Precision@K indicates that whether the ground truth POI appears in the top-k recommended POIs [77].

[4] MAP (Mean Average Precision) measures the order of our recommendation list [77].

[5] MRR is the average reciprocal rank of positive examples. This metric reflects the overall ranking ability of the model [42].

[6] NDCG (Normalized Discounted Cumulative Gain) measures the quality of top-K ranking list [44].

Category-aware Deep Model (CatDM) which predicts POIs that are likely to be visited by users in the next 24 hours. CatDM contains (1) metric embedding that learns the latent features of a user, POI, POI category, and time; (2) first deep encoder for capturing user preferences in POI categories; (3) two filters for reducing search space to generate candidates; (4) another deep encoder for user preferences in POIs, and (5) a module for ranking the candidate set. For ranking candidate set this model considers four correlations simultaneously i.e., the correlation between user and POI, the correlation between the user and POI category, the correlation between POI and temporal influence, and the correlation between POI and user's current preferences. The model uses POI categories and geographical influence for overcoming data sparsity. The model also uses an attention mechanism for getting better results.

Let us assume that l_1, l_2, l_3 are three locations and l_2 is nearby to l_1. Now, if a check-in from l_2 to l_3 presents in a user check-in list then after visiting l_1 it will most likely to visit l_3. This type of situation is called transition regularity. Most of the existing POI recommendation methods capture sequential regularity only. On the other hand, Gua et al. [80] proposed **Attentional Recurrent Neural Network (ARNN)** model which captures both sequential and transition regularities for resolving sparsity problem. ARNN consists of several layers i.e., (1) Neighbor discovery layer: neighbors are extracted from heterogeneous data by using knowledge graph (KG) and meta-path; (2) Embedding Layer: transforms the sparse features of check-in sequence into dense representation and learns the spatio-temporal features, semantic context by using multimodal embedding layer; (3) Attention layer: calculates the similarity between the current location and each neighbor and capture the transition regularities of the neighbors; (4) Recurrent layer: captures higher-order sequential regularity by using LSTM. Besides these models, Wang et al. [81], Huang et al. [82] and Li et al. [83] used LSTM and attention layers for POI prediction. Yang et al. [84] used LSTM and embedding layers for the next POI recommendation. Another model called **SLSTM** [85] used stacked LSTM and embedding layers for sequential check-in prediction.

6.2.2. Bi-LSTM

Vanilla LSTM processes input only in one direction sequentially. While it helps the models getting sequential information from the previous inputs, information from later parts of the input cannot be captured. Bi-directional LSTM (Bi-LSTM) solves this problem by considering both directions of inputs. Some recent POI recommendation models also use Bi-LSTM to capture sequence features from both directions to achieve better performance.

From the model GeoIE [15] we find that POI has two propensities i.e., (1) Geo-influence: directs its visitors to other POIs and (2) Geo-susceptibility: the receipt of visitors from other POIs. By considering the above-mentioned properties, Liu et al. [86] proposed a model called **Geographical-Temporal Awareness Hierarchical Attention Network (GT-HAN)**. For better capturing the great variation in geographical co-influence across POIs GT-HAN uses three factors i.e., the geo-influence of POIs, the geo-susceptibility of POIs, and the distance between POIs. The main part of GT-HAN is an embedding layer, a geographical-temporal attention network layer, and a context-specific co-attention network layer. The embedding layer captures geo-influence, geo-susceptibility, and semantic effects. The geographical-temporal part explores the geographical relations between POIs and the temporal dependency of a check-in list and uses the Bi-LSTM model to capture the sequence dependence of a user's check-in list. The context-specific co-attention network captures dynamic user preferences. GT-HAN [86] is actually the improved version of a previous model that is also known as GT-HAN [87] and proposed by the same authors.

Liu et al. [88] proposed **time-aware Location Prediction (t-LocPred)** model. This model has two basic parts i.e., ConvAoI and mem-attLSTM. ConvAoI uses the CNN layer and ConvLSTM layer [12] to find the correlations among adjacent AoIs (Area-of-Interest) and time periods within a day and a week respectively. So, CNN and ConvLSTM work as short-term and long-term coarse-grained spatial-temporal modeling respectively. On the other hand, mem-attLSTM captures complex long-term correlation using a spatial-aware memory augmented LSTM and time-aware

Md. Ashraful Islam, Mir Mahathir Mohammad, Sarkar Snigdha Sarathi Das et al.

Neurocomputing 472 (2022) 306–325

attention mechanism. So, mem-attLSTM works as a fine-grained filter that selects the most likely POIs a user will visit.

Chang et al. [89] proposed **Content-aware successive POI recommendation (CAPRE)** which is a complete POI recommendation model that uses user-generated textual content. CAPRE has four modules: (1) input layer: takes check-in history as input; (2) content encoder layer: uses character-level CNN (Convolutional Neural Network), multi-head attention mechanism, and POI embedding for capturing various perspectives of user interests about POIs; (3) user behavior pattern: captures content-aware and geographical user behavior patterns using Bi-LSTM; (4) Output layer: multi-layer perceptron (MLP) to capture users' general preferences for POIs.

6.2.3. Modified LSTM

Some models modify the basic LSTM [63] model for enhanced POI prediction. The underlying idea is to better capture user short-term and long-term preferences by modifying the basic LSTM.

Kong et al. proposed a new LSTM based model [90], namely **Hierarchical Spatial-Temporal Long-Short Term Memory (HST-LSTM)**. HST-LSTM uses a hierarchical model to encode the periodicity of people's movement. This model captures users' historical visiting sequences in an encoder and decoder manner which improves the performance of POI recommendation. They add spatial-temporal factors into the gates of the basic LSTM model which naturally combines spatial-temporal influence into the basic LSTM model for resolving data sparsity issues.

Zhao et al. [91] proposed **Spatio-Temporal Gated Network (STGN)** that modifies the basic LSTM [63] to capture short-term and long-term preference easily. This model adds four new gates i.e., two for long-term preferences and the other two for short-term preferences. This model also adds a new cell state. So, in the proposed model, there is one cell state for short-term preferences and one cell state for long-term preferences. STGN model is further improved by using coupled input and forget gates called STCGN (Spatio-Temporal Coupled Gated Network). STCGN reduces the number of parameters and thus this model can be trained easily and improves efficiency. This model is further improved by the authors, which is named as STGN [92].

6.2.4. Self-attention

Following the success of self-attention in language modeling, state-of-the-art POI recommendation models have leveraged this powerful approach to achieve best in class performance. Among them, Lian et al. [51] proposed **Geography-aware sequential recommender based on the Self-Attention Network (GeoSAN)**, which uses a geography-aware self-attention network and geography encoder. The attention network consists of an embedding layer, a self-attention encoder, a target-aware attention decoder, and a matching function. The geography encoder uses map gridding and GPS mapping for encoding GPS location as quad keys. To address the sparsity challenge, the paper proposes a weighted binary cross-entropy loss function based on importance sampling, so that informative negative samples are more weighted.

In another work, Guo et al. [52] proposed a model – **self-attentive networks along with spatial and temporal pattern learning for next POI recommendation (SANST)**. SANST is actually a modification of the model SASRec [93] which uses a two-layer *transformer network* [67]. SASRec does not capture spatial and temporal patterns. So, for capturing spatial patterns, SANST updates the embedding of check-ins by adding the location of the checked-in POI. Since the coordinate representation of a POI location is sparse, the model discretizes the whole data space with a grid such that POI locations are represented by the grid cell IDs. Grid Cell IDs are learned using *GeoHash* encoding and Bi-LSTM net-

work. For capturing temporal effect SANST adds a parameter in the self-attention network which is computed using the time difference between two check-ins.

Analyzing LSTM Based models:

By deeply analyzing all the proposed models of this section we develop a generalized model in this LSTM paradigm, which is shown in Fig. 5. A detailed description of different key parts of this generalized LSTM model with the commonality and disparity among other models is given below.-

- **Input:** The input of all the models is a check-in list. Some papers [44,80,87,90] divide the check-in list into trajectories. Each trajectory consists of a sequence of POIs visited by the user. This type of modeling helps the model to capture the temporal and short-term user preferences.

- **Encoding:** One of the main challenges in the next POI recommendation problem is data scarcity. To resolve the data scarcity issue, existing techniques use embeddings that map a higher-dimensional sparse data space into a lower-dimensional dense data space. For this reason, most of the papers [45,77–79,44,46,47,80,87,89] use embedding layers preprocess the check-ins. Along with the Fig. 5: Generalization of LSTM Models embedding layers, TMCA [45] uses multi-level context attention to identify relevant historical check-ins and contextual factors. ARNN [80] utilizes knowledge graphs to find relevant neigh-
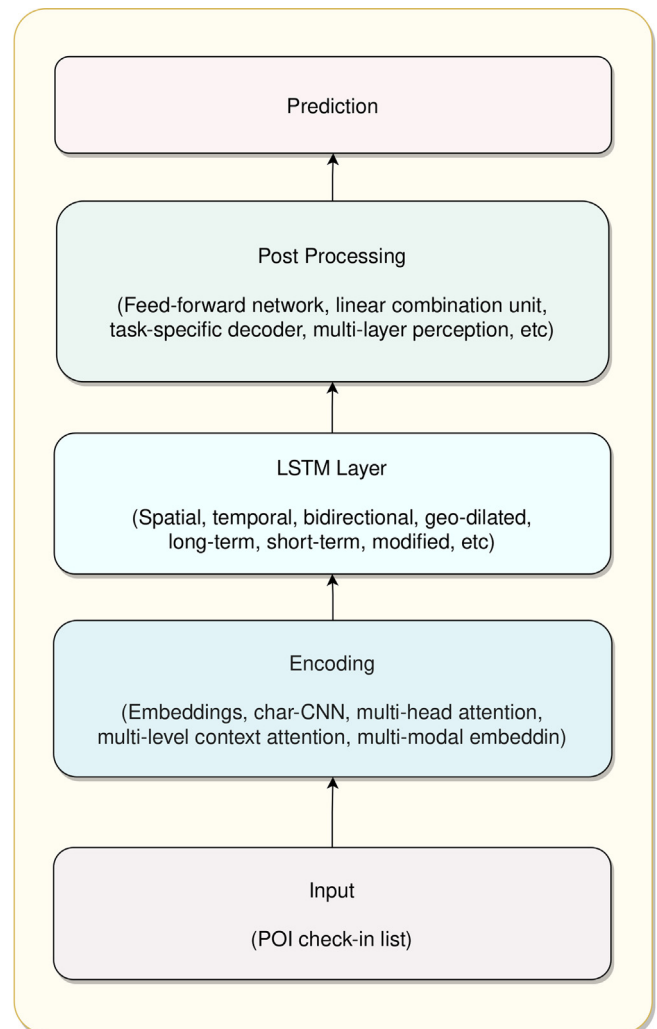

Fig. 5. Generalization of LSTM models.

Md. Ashraful Islam, Mir Mahathir Mohammad, Sarkar Snigdha Sarathi Das et al.

Neurocomputing 472 (2022) 306–325

bors. CAPRE [89] makes use of character level CNN and multi-head attention mechanism for capturing various perspectives of users' interests about POIs.

- **LSTM Layer:** LSTM layers are highly effective in capturing both long and short-term preferences. Leveraging this property, Wu et al. [77,78] use LSTM for capturing shortterm preferences. Sun et al. [44] incorporate LSTM for capturing long-term preferences and geo-dilated LSTM for capturing both geographical and temporal factors. Liu et al. [87] and Chang et al. [89] utilize Bi-LSTM for better capturing of user behavior pattern. Some researchers modified the basic LSTM cell to get better predictions. HSTLSTM [90] of Kong et al. introduces spatio-temporal factors into the existing three gates of the basic LSTM. Zhao et al. modified the basic LSTM in STGN [91] where their model has two time gates and two distance gates for capturing both short-long term interests of interval and distance. So, both of the models use modified versions of the basic LSTM for capturing the aspects of LBSNs.

- **Post Processing:** Similar to the meaning of the word 'attention in natural language processing, the attention mechanism in DNN pays attention to the portion of the data that is more relevant. So, in this post-processing layer, [77–79,44,46,47,80,87] use attention for eliminating noise and better POI recommendation. Along with the attention mechanism, the model [44] uses some non-local networks for capturing long-term preferences. However, [89] uses multi-layer perceptron for capturing users' general preferences and [46] uses a task-specific decoder for performing the next activity, POI type and location prediction tasks.

- **Prediction:** This final layer is used to recommend the next POIs.

6.3. GRU

RNN shows a great promise for analyzing time series data and LSTM helps in capturing short and long-term effects of visiting POIs. However, both RNN and LSTMs tend to suffer from cold start problems. Most RNN models rely on the last hidden layer which limits the learning of user information from the hidden layers. To solve such cases, a modified form of LSTM- GRU [64] has been introduced in many POI recommendation models. The GRU has fewer parameters to learn comparing to LSTM cells but it has additional gates such as forget gate to compensate for the problems mentioned above alongside solving exploding and vanishing gradient problems.

A popular model called **DeepMove** is proposed by Feng et al. [48] . DeepMove actually predicts human mobility which is very similar to our POI recommendation. DeepMove has two modules i.e., (1) Multi-modal Recurrent Prediction Framework: extract features by jointly embedding spatiotemporal and personal feature into a dense representation, which is then fed into GRU unit to model long-range and complex dependencies in a trajectory sequence; (2) Historical Attention Module: captures multi-level periodicity of human mobility.

Most of the model considers identical impact from different types of contexts on the users' preferences. But their impacts are not identical. To solve this problem Manotumruksa et al. [49] proposed **Contextual Attention Recurrent Architecture (CARA)** model. CARA has four layers i.e., input layer, embedding, recurrent layer, output layer. In the recurrent layer, this model uses GRU. For capturing different contextual impact on the users' preferences this model uses two types of gating mechanisms i.e., (1) Contextual Attention Gate (CAG): controls the influence of ordinary and transition contexts on the users' dynamic preferences and (2) Time-and Spatial-based Gate (TSG): considers the time intervals and geographical distances between successive check-ins to control the influence of the hidden state of previous GRU units.

Kala et al. proposed [50] **Multi-GRU (MGRU)** which modifies the basic GRU unit by adding two additional gates for a better recommendation. The first added gate is Dynamic Contextual-Attention-Gate (DCAG-α) which captures the effect of dynamic contexts like - time of the day, companion, user's mood, etc. The other gate is Transition-Contextual-Attention-Gate (TCAG-β) captures the effect of transition contexts like - time interval and geographical distance from past POI to future POI. MGRU has three layers i.e., (1) input layer: pre-process and embeds check-in sequence; (2) recurrent layer: captures sequential patterns using MGRU; (3) output layer: recommend the next POI.

6.4. Graph embedding

Some recent techniques leverage the potentials of Graph Embeddings (GE) that learn low-dimensional key features of the dataspace modeled as different forms of graphs such as POI-POI, user-POI, and POI-time.

Xie et al. [94] proposed **GE** that uses graph embedding for recommending next POI. GE jointly captures the sequential effect, geographical influence, temporal cyclic effect, and semantic effect in a unified way using four bipartite graphs. POI-POI captures sequential effect, POI-Region graph captures geographical influence, POI-Time graph captures temporal cyclic effect and POI-Word graph captures semantic effect. The model embeds these four relational graphs into a shared low-dimensional space. Then this model computes the similarity between a users' query (users' embedding, query time, and location) and the POIs that are not visited by that user. Most similar POIs are taken for the recommendation.

Liu et al. [53] proposed **SpatioTemporal Aware (STA)** which generalizes knowledge Graph Embedding [95] in their model. GE [94] embeds both users and POIs in a common latent space. The users and POIs are inherently different objects so that approach is unnatural. On the other hand, STA takes location and time as a spatiotemporal pair ⟨$time, location$⟩ and uses the embedding of this pair as a relationship for connecting users and POIs.

Christoforidis et al. [96] proposed **Jointly Learn the Graph Embeddings (JLGE)** which uses six informational graphs. They are two unipartite (user-user and POI-POI) and four bipartite (user-POI, user-time, POI-user, and POI-time). This model consists of a three-step process. In the first step, the model builds information graphs and weights the edges. In the second step, the model jointly learns the embeddings of the users and the POIs into the same latent space from these six informational graphs using the LINE model [97]. Finally, in the third step, the model personalizes the POI recommendations for each user by tuning the influence of the participation networks for the final suggestions of the target user.

The authors of the JLGE [96] model extended their work by introducing a new model called **Recommendations with multiple Network Embeddings (RELINE)** [54]. The model introduces two new networks: *i*) Stay Points, which represents the locations of the user stayed the most, and *ii*) Routes, the path followed when visiting POIs. Additionally two new bipartite graphs i.e., user-route and POI-stay points are added with the previous JLGE [96] model for better capturing the users' preference dynamics.

In another work, Xiong et al. [55] proposed a semi-supervised learning framework called **Dynamic Spatio-temporal POI recommendation (DYSTAL)**. DYSTAL has two key components: a network embedding method and a dynamic factor graph model. The network embedding method jointly learns the embedding vectors of users and POIs of three subgraphs i.e., POI-POI, user-POI, and user-user to excavate complex spatio-temporal patterns of visiting behaviors. The Dynamic factor graph model captures different fac-

tors including the correlation of users' vectors and POIs' vectors from the previous embedding layer via the Factor Graph Model (FGM) [98]. This model also considers the textual reviews of users by using SentiStrength [99] tool.

Zhang et al. [100] propose **Hierarchical Category Transition (HCT)** which extends the Skip-Gram [74] model to learn the hierarchical dependencies between POIs and categories and the hierarchical category transition. HCT models the dynamic user preference by considering recently visited POIs and the associated hierarchical categorical sets. They formulated the dynamic user representation by incorporating the representations of POIs as well as the associated hierarchical category sets. Besides these models, Liu et al. [101] propose a model based on bi-weighted low-rank graph for providing POI recommendations for a specific time period. Chen et al. [102] use embeddings and context filtering for modeling spatial trajectories. Furthermore, Qiao et al. [103] utilize embeddings for predicting next POIs.

Analyzing Graph Embedding Based Models:

From the discussion of the aforementioned graph embedding based models, we find that all the models basically follow an identical structure which is shown in Fig. 6.

The graph embedding based models use different types of graphs that include POI-POI [94,96,54,55], POI-region [94], POI-time [94,96,54], POI-word [94], User-User [96,54,55], user-POI [96,54,55], POI-user [96,54], user-time [96,54], user-route [54] and POI-stay point [54] graphs. By using the embeddings of these above mentioned graphs, the models enhanced their recommendation performance for recommending the next POI.

6.5. GAN

Generative Adversarial Networks (GAN) [68] is a popular model in deep learning where two Neural Network models compete with

Fig. 6. Generalization of graph embedding models.

each other for giving better predictions. Few papers use GAN in POI recommendation because POI recommendation problems do not necessarily fall under the solvable problem domain of GAN.

Liu et al. [104] proposed **Geographical information-based adversarial learning model (Geo-ALM)** model that uses two modules: discriminator and generator which are essentially different but inspired from the conventional GAN. The pairwise ranking is regarded as a discriminator that tries to predict the ranking relationship between generated sample pairs and is trained to maximize ranking samples' likelihood. The generator continually generates critical negative samples, which are then coupled with positive samples, forming training instances. The framework interchangeably learns the parameters between two different modules.

Zhou et al. [105] proposed **Adversarial POI Recommendation (APOIR)** which combines GAN, GRU, and Matrix factorization for POI recommendation. GRU and MF combinedly learn both temporal and sequential preferences of users. Two competitive components: recommender and discriminator are alternatively optimized by training both of them through an objective function using the learned preferences of users. The discriminator tries to maximize the probability of correctly distinguishing the true check-in locations from the generated recommended POIs by the recommender. Gao et al. [106] also uses GAN networks for identifying individuals by exploiting their trajectories.

Another state-of-the-art method proposed by Li et al. [107] uses GAN into transfer learning to transfer knowledge across domains. The model consists of four components: (1) an embedding layer to learn latent factors, (2) a cross-domain transfer layer for capturing shared private feature extractor, (3) the cross-domain recommendation part for an item recommendation, and (4) the domain discriminator for preventing domain-specific features to be transferred.

6.6. Other models

There are some other deep learning POI recommendation models that use hybrid architectures. In this section, we cover those POI recommendation models.

Zhao et al. [26] propose **Geo-Temporal sequential embedding rank (Geo-Teaser)**. Geo-Teaser is based on the Skip-Gram [74] model which learns the representations of context POIs given a target POI. The model attempts to learn the temporal POI embeddings by maximizing an objective function. The geographically hierarchical pairwise preference ranking model uses Bayesian Personalized ranking to learn the user preference on POIs. The core Geo-Teaser model is a unified framework that combines the temporal embedding model and the pairwise ranking method.

Chang et al. [109] proposed a somewhat different approach called **Content-Aware hierarchical POI Embedding (CAPE)** for POI recommendation. Most models do not use the text content of POIs, since most of the datasets do not contain such textual content. The authors thus generated a new dataset from Instagram which contains a textual description of POI written by the users. CAPE actually is a POI embedding model that consists of a check-in context layer and text content layer. The check-in content layer captures the geographical influence of POIs, while the text content layer captures the characteristics of POIs.

A different type of work on missing POI check-in identification by Xi et al. [111] is also notable in this context. Here, the authors proposed a model called **Bi-directional Spatio-Temporal Dependence and users' Dynamic Preferences (Bi-STDDP)** to capture complex global spatial information, local temporal dependency relationships, and users' dynamic preferences. Bi-STDDP takes two check-in lists as input; one list for before the missing check-in, and one for after. Besides, this model also uses user-embedding, POI-embedding, and temporal patterns in the model.
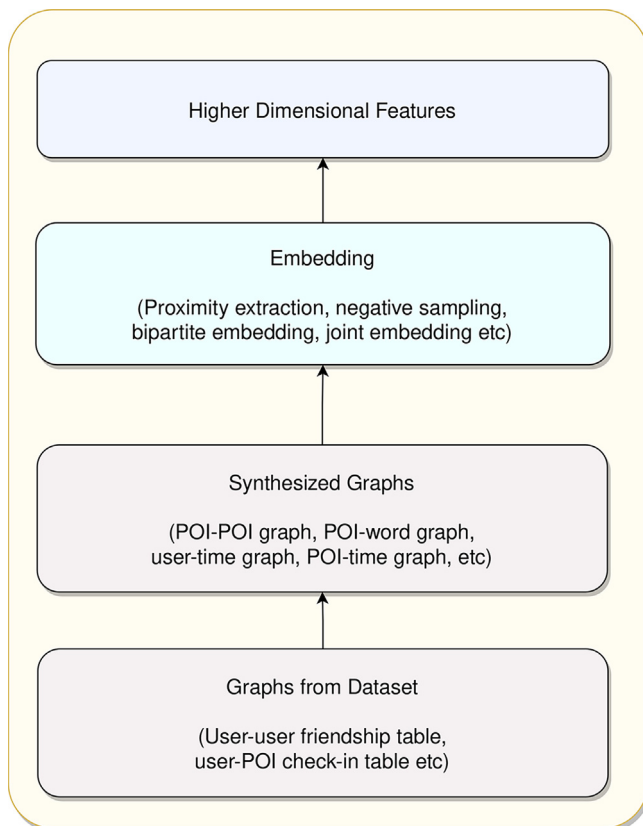
Md. Ashraful Islam, Mir Mahathir Mohammad, Sarkar Snigdha Sarathi Das et al.

Neurocomputing 472 (2022) 306–325

Zhou et al. [112] proposed a hybrid architecture called **Topic-Enhanced Memory Network (TEMN)**. TEMN consists of three key parts: a Memory Network (MN), Temporal Latent Dirichlet Allocation (TLDA) [116] and geographical modeling. MN learns the complex interaction between user and POIs and captures the neighborhood-based interests of a user. TLDA is an unsupervised generative probabilistic model which captures temporal preferences and the inner interest of users. This module provides the pattern-user probability distribution. The distributions of venues and time slots associated with each pattern can also be estimated through TLDA. TLDA and MN can jointly learn the characteristics of users and POIs. The Geographical model captures the geographical influence.

Recent researches show that hierarchical structures can be modeled using hyperbolic representation methods [117–119]. For that reason, Feng et al. [113] proposed a novel **Hyperbolic Metric Embedding (HME)** approach for the next-POI recommendation task. HME can be divided into two part i.e., (1) Hyperbolic Metric Embedding: uses Poincaré ball model [118] to learn four different relationships (POI-POI, POI-User, POI-Category, and POI-Region) by projecting them in a shared hyperbolic space; (2) Recommending with Hyperbolic Embeddings: combines the user preferences and POI sequential transitions in the Poincaré ball model an Einstein midpoint aggregation method [120,121]. The geographical distance is also considered in this model because users tend to visit the POI that is close to them [122].

Along with the aforementioned models, recent works have proposed some advanced models that take advantage of multiple techniques to improve performance. Chen et al. [123] proposed pivot-based principles for space transformation and partitioning so that high dimensional semantic vectors can be rationally mapped to low dimensional coordinates, Zhang et al. [115] used embeddings for geographical influence modeling. Wang et al. [108] used embeddings for predicting the next POIs. Ding et al. [110] used DNN for time-specific POI recommendation. The CNN models have also been used in POI recommendation in some prior works [38,39]. Massimo et al. [124] experimented with Inverse Reinforcement Learning (IRL) [125–127] to analyze the performance of IRL in POI recommendation. Besides these models Cai et al. [128] proposed a novel holistic influence diffusion model that takes into account both cyber and physical user interactions effectively and practically. Li et al. [129] proposed few-shot learning for new user recommendation.

Summary of different paradigms of deep learning models:
By analyzing all the aforementioned models, we find a comprehensive picture of how different deep learning paradigms have been utilized to handle different aspects of POI recommendation models. Since successful recommendation fundamentally depends on historical POI information, sequential models (i.e., RNN, LSTM) have been primarily used in recent POI recommendation models. Among them, LSTM is the most popular approach due to its long-term sequential information capturing capability. A large body of works has experimented with minor modifications of LSTMs to improve long and short-term modeling preferences. Thus, sequence information capture has been the center of interest in POI recommendation models. Recent state-of-the-art models are also leveraging the self-attention transformer mechanism for POI recommendation, which greatly suppresses the problems associated with very long sequences. They are also computationally parallelizable. Besides capturing historical sequence information, researchers have also analyzed the spatial influence in this regard. Due to the nature of check-in data, users and POIs form a relationship graph that can be highly useful to model spatial dependency. Thus, graph embedding methods have garnered attention in the past couple of years in POI recommendations. LSTM coupled with these graph embedding models thus opens up the opportunity to capture both sequential and spatiotemporal features from a given dataset. Advanced models have also utilized adversarial learning models and specialized embedding methods to enhance the performance of the recommendation.

7. POI sequence recommendation

So far we've primarily discussed the next POI recommendation models. However, some state-of-the-art methods also recommend a sequence of POIs that are more likely to be visited by the user in the future. In POI sequence recommendation the input is a sequence of check-ins or check-in list and the output also a check-in list. So, we can think of this recommendation as a sequence to sequence (Seq2Seq) task. Several state-of-the-art methods have been proposed in recent years for solving the POI sequence recommendation task. Here we briefly discuss some of them.

Baral et al. [130] proposed **Contextualized Location Sequence Recommender (CLoSe)** which incorporates different contexts (e.g., social, temporal, categorical, and spatial) into the hidden and output layer. This model uses either the simple RNN or the LSTM model. The results show that CLoSe-LSTM performs better than CLoSe-RNN.

Huang et al. [131] proposed **Dynamic Recommendation of POI Sequence (DRPS)** which is based on DNN. This model consists of an encoder and a decoder module and for getting better performance, this model takes into account the POI embedding feature, the geographical and categorical influences of historical trajectories, and the positional encoding. This paper also proposed two new evaluation metrics for better performance evaluation.

In another work, Lu et al. [132] proposed **Graph-based Latent Representation model (GLR)** which can capture geographical influence, temporal influence, user preference, etc. GLR learns the latent vectors based on word2vec [133] technique. Here, the authors added user preference, temporal successive transition influence, geographic influence, and LSTM [63] with GLR model and propose a new model GLR_GT_LSTM which can capture users' complex successive transition behavior.

Alongside these models, Wang et al. [134], Baral et al. (HiCaPS) [135], Lin et al. [136] also proposed different POI sequence recommendation models. On the other hand, Li et al. [137] worked with a variation of this task, missing POI check-ins prediction by leveraging an attention-based seq2seq generative model.

8. Influential factors

In Section 6, we have highlighted existing research works in the domain of POI recommendation. In these works, researchers have tried to figure out the most influential factors that affect POI recommendation. It is quite difficult to conclude the factors of choosing the next POI, as POI recommendation is greatly affected by human behavior which changes over time. However, most researchers agree with some common factors like sequential effect, geographical influence, semantic effect, social influence, temporal influence, etc. that affect POI recommendation. These factors are generally derived from the patterns of the users' decisions. In this section, we elaborately discuss these influential factors. In Table 5, we also summarize how different POI recommendation models cover these influential factors.

8.1. Sequential effect

The sequential effect of POI recommendation is similar to the analogy of NLP problems- such as making a sentence where the next word depends on the previous words.

Md. Ashraful Islam, Mir Mahathir Mohammad, Sarkar Snigdha Sarathi Das et al.

Neurocomputing 472 (2022) 306–325

Table 4
Descriptions of datasets.

Dataset name	Reference	Date	Region	#User	#POI	#Check-in
Foursquare	Flashback [41]	Apr 2012–Jan 2014	World	46065	69005	9450342
	ASPPA [42]	Apr 2012–Sep 2013	US (except Alaska, Hawaii)	49005	206097	425691
	LSPL [77], PLSPL [78]	Apr 2012–Feb 2013	New York	1083	38333	227428
	CatDM [47], ARNN [80], HME [113], Bi-STDDP [111]	Apr 2012–Feb 2013	Tokyo	2293	61858	573703
	ASTEN [79]	-	USA	21878	21651	569091
			EU	15387	115567	3227845
	LSTPM [44]	Feb 2010–Jan 2011	New York	934	9296	52983
	iMTL [46]	Apr 2012–Sep 2013	Charlotte	1580	1791	20940
			Calgary	301	985	13954
			Phoenix	1623	2441	22620
	GT-HAN [86], APOIR [105]	Apr 2012–Sep 2013	USA	24941	28593	1196248
	CAPRE [89]	–	–	4163	121142	483813
	STGN [91]	Jan 2010–Feb 2011	California	49005	206097	425691
		Aug 2010–Jul 2011	Singapore	30887	18995	860888
	GeoSAN [51]	Apr 2012–Jan 2014	World	12695	37344	1941959
	DeepMove [48]	Feb 2010–Jan 2011	New York	15639	43379	293559
	CARA [49]	–	World	10766	10695	1336278
	MGRU [50]	Aug 2010–Jul 2011	Singapore	4630	6176	201525
	GE [94]	Sep 2010–Jan 2011	World	114508	114,508	1434668
	STA [53], JLGE [96], RELINE [54]	Sep 2010–Jan 2011	USA	114508	62462	1434668
	Geo-ALM [104]	Aug 2010–Jul 2011	Singapore	2321	5596	194108
	DYSTAL [55]	–	Singapore	74250	–	–
	HCT [100]	–	New York City	1982	2454	187750
	Geo-Teaser [26]	Jan 2011–Jul 2011	World	10034	16561	865647
Gowalla	ST-RNN [40]	Feb 2009–Oct 2010	World	10997	–	6400000
	Flashback [41]	Feb 2009–Oct 2010	World	52979	121851	3300986
	ASPPA [42]	Feb 2009–Oct 2010	World	4996	6871	245157
	LLRec [43], RELINE [54]	Jan 2009–Aug 2011	World	319063	2844076	36001959
	TMCA [45]	Feb 2009–Oct 2010	World	22209	50569	1493799
	ASTEN [79]	Feb 2009–Oct 2010	World	52484	115567	3227845
	LSTPM [44]	Feb 2009–Oct 2010	World	5802	40868	301080
	ARNN [80]	Feb 2009–Oct 2010	San Fransisco	170	7340	32058
	GT-HAN [86], STGN [91], APOIR [105]	Feb 2009–Oct 2010	World	18737	32510	1278274
	t-LocPred [88]	Jan 2009–Oct 2010	Goteborg (Sweden)	5342	12229	103787
	GeoSAN [51]	–	–	31708	131329	2963373
	SANST [52], Geo-ALM [104]	Feb 2009–Oct 2010	California, Nevada	10162	24250	456988
	MGRU [50]	Nov 2009–Oct 2010	Austin	2321	5596	194108
	GE [94], STA [53], Bi-STDDP [111]	Feb 2009–Oct 2010	World	107092	1280969	6442892
	Geo-Teaser [26]	Jan 2011–Jul 2011	World	3240	33578	556453
	HME [113]	Nov 2010–Jun 2011	Houston	4627	15135	362783
Brightkite	t-LocPred [88]	May 2008–Oct 2010	Tokyo	2263	37196	183298
	STGN [91]	May 2008–Oct 2010	World	51406	772967	4747288
	GeoSAN [51]	Apr 2008–Oct 2010	World	5247	48181	1699579
	CARA [49]	Apr 2008–Oct 2010	World	14374	5050	681024
WeChat	TEMN [112]	Sep 2016–Aug 2017	Beijing	28566	13826	509589
				75973	28183	5644965
Yelp	TMCA [45]	Jan 2014–Jun 2017	World	11564	18683	492489
	CARA [49]	–	World	38945	34245	981379
	DYSTAL [55]	–	Las Vegas	337084	26809	1605396
	APOIR [105]	–	World	30887	18995	860888
Baidu	HST-LSTM [90]	Dec 2015–Dec 2015	Peking	–	–	–
Instagram	CAPRE [89], CAPE [109]	–	New York City	78233	13187	2216631
Weeplaces	LLRec [43], JLGE [96], RELINE [54]	Nov 2003–Jun 2011	World	15799	971309	7658368
	t-LocPred [88]	Nov 2003–Jun 2011	New York City	4855	38537	900906

[2,4,94,54] also show that sequential effect puts a major impact on POI recommendation. Zhang et al. [4] extracted sequential patterns from the location sequences of all users and model them as a concise location-location transition graph. Then they determine the transition probabilities in terms of transition counts and outgoing counts. Finally, the model processes the check-in locations according to their arrival order and incrementally updates the constructed location-location transition graph. The sequential probabilities are derived with additive Markov Chain applied on the location-location graph. Xie et al. [94] designed a fully connected deep LSTM network for skeleton-based action recognition. This architecture enables fully exploit the inherent correlations among skeleton joints to capture sequential effect. Christoforidis et al. [54] jointly learned the graph embeddings of different information networks in the same latent space. The model is optimized using negative sampling. All the embeddings of the input bipartite graphs are integrated into the model.

8.2. Geographical influence

Recent researches show that people tend to visit places that are close to him/her or are close to the places already visited by that person. Users who check in a location within a region have a relatively larger probability to visit the places in close proximity. Users tend to go to stores, marketplaces, or visit picnic spots that are close to where they live. Also, after visiting a certain tourist spot, people tend to go to nearby restaurants or malls. Thus, spatial proximity is a worthy concern to predict users' next location.

Md. Ashraful Islam, Mir Mahathir Mohammad, Sarkar Snigdha Sarathi Das et al.

Neurocomputing 472 (2022) 306–325

Table 5
Descriptions of influential factors.

Reference	Sequential Effect	Geographical Influence	Semantic Effect	Social Influence	Temporal Influence
ST-RNN [40], Flashback [41], LSTPM [44], TMCA [45], DeepMove [48], SANST [52], ASTEN [79], GT-HAN [87], HST-LSTM [90], STGN [91], Bi-STDDP [111], TEMN [112]	✔	✔	×	×	✔
ASPPA [42], iMTL [46], CatDM [47], CARA [49], MGRU [50], LSPL [77], PLSPL [78], ARNN [80], GT-HAN [86], t-LocPred [88], GE [94]	✔	✔	✔	×	✔
CAPE [109]	×	✔	✔	✔	×
Geo-Teaser [26], GeoSAN [51]	✔	✔	×	×	×
CAPRE [89]	✔	✔	✔	✔	×
STA [53]	×	✔	✔	×	✔
RELINE [54], JLGE [96], APOIR [105]	✔	✔	×	✔	✔
LLRec (Teacher) [43]	✔	✔	✔	✔	✔
HME [113]	✔	✔	✔	×	×
DYSTAL [55]	×	✔	✔	✔	✔
HCT [100]	×	✔	✔	×	×

[32,94,44,138,139,109] show great interest in analyzing the geographical influence for predicting POIs. Ye et al. [32] perform spatial analysis on real datasets of Foursquare and Whrrl. The study finds the implication of distance on user check-in behavior by measuring the probability of a pair of check-ins being within a certain distance. The study confirmed the above-mentioned implications of the proximity of POI in predicting the next POI. The model introduces a collaborative recommendation method based on the naive Bayes to realize the POI recommendation. The paper proposes a unified framework to perform collaborative recommendations that fuse user preference, social influence, and geographical influence. Furthermore, the model uses a linear fusion framework to integrate ranked lists provided by the three recommender systems. Kefalas et al. [138] use contextual pre-filtering of the information to select the most relevant proximate users for the POI recommendations. The spatial influence of users' reviews represents the impact of proximate users who reviewed similar businesses to the target user. The proposed model extends the item-based contextual filtering in two ways, (i) by leveraging the proximity factor when computing the similarity of two users and (ii) by considering the history of proximate user reviews. Liu et al. [139] introduce a geographical probabilistic factor analysis framework for POI recommendation. To learn geographical user preferences, the model encodes the spatial influence and user mobility into the user check-in process. Furthermore, [139] adopts a Bayesian probabilistic non-negative latent factor model for encoding both the spatial influence and personalized preferences.

8.3. Semantic effect

Every POI has some properties and two POIs are semantically close to each other if they have similar properties. Every human also has his/her own preferences and a person wants to visit those POIs which are matched to his/her preferences [140]. So, from the check-in list of a user, we can capture the user preferences and try to predict those POIs that are semantically very similar to previous check-ins.

[140,138,77,45,109] are some of the papers that use the semantic information from the datasets to predict the next POI. Kefalas et al. [138] utilize the textual influence among the reviews that refer to the similarity between the reviews. Ye et al. [140] use a semantic annotation technique for POI networks to automatically annotate all places with category tags. The annotation algorithm learns a binary SVM classifier for each tag in the tag space to support multi-label classification. This algorithm extracts features and handles semantic annotation from places with the same tag and the relatedness among places. Wu et al. [77] learn the short and long-term contextual features of POIs and leverage attention

mechanisms to capture users' preference. Li et al. [45] propose an encoder-decoder neural network model that leverages the embedding method to incorporate heterogeneous contextual factors associated with each check-in activity, to populate the semantics of check-ins. The paper embeds check-in user and time, numerical factors, and categorical factors in contexts. Chang et al. [109] utilize the text content that provides information about the characteristics of a POI. They also measure the correlation between words by calculating the Jaccard similarity of POIs in their text content. The text content layer treats text content as a sentence and trains the word embedding vector using Word2Vec [74]. All these works employed different techniques to capture the semantic features from the POIs for getting better recommendation performance.

8.4. Social influence

Humans are social beings. So, the decisions of a person greatly depend on his/her social status, friends, neighbors, culture, etc. These social influences affect a person's interest in visiting a POI. The probability of a user visiting a POI is increased when his/her friends give good reviews about that POI [109]. To tackle the cold start problems, the social circle of a new user can be heavily beneficial, and thus, models can learn a user's preferences by suggesting the same POIs of his/her social circle.

[54,109,138] use social influence to improve their POI predictions. Christoforidis et al. [54] incorporate social influence alongside spatial and temporal context and combined the graphs into a unified prediction model. Kefalas et al. [138] try to capture the social influence using users' reviews. The users having similar vocabularies are considered to be related. Here, the social influence corresponds to the correlation between the target user and others concerning the lexical analysis of their reviews. The abundance of works leveraging social influence proves the importance of social effect in POI recommendation.

8.5. Temporal influence

Human-lives consist of fixed time patterns. A user shows distinct check-in preferences at different hours of the day and tends to have similar preferences in consecutive hours than non-consecutive hours [141]. User activities are influenced by time. A user will go to a restaurant rather than a bar at noon and people may tend to go to visit places when a holiday appears.

[142,141,94,54,79] considered the temporal effects in POI recommendation systems. Gao et al. [141] introduced a temporal state to represent hours of the day. Then they defined the time-dependent user check-in preferences using the temporal state.

The paper proposes a temporal regularization to minimize an objective function using temporal coefficients. Their proposed framework, LRT, consists of temporal division, temporal factorization, and temporal aggregation. Yuan et al. [142] perform collaborative filtering by exploiting other user's temporal preferences to POIs. To capture the fixed routine of users' daily mobility, the model splits time into hourly slots and models the temporal preference to POIs of a user in a time slot by the POIs visited by the user in this time slot. They leverage a time factor when computing the similarity between two users and consider the historical check-ins at a time in the repository. Doan et al. [79] use an attention mechanism designed for utilizing spatio-temporal information.

9. Future directions

9.1. Preparing benchmark datasets

While popular check-in datasets like Foursquare, Gowalla, Yelp, Weeplaces are largely used in the previous works, the sheer scale of these datasets make them impossible to work with all at once. Consequently, most of the prior works take only a subset (e.g. check-ins of a city within a specific timeframe) of one dataset for evaluation. Furthermore, recent advancements in POI recommendation show that besides check-ins, data from other modalities such as textual description, user reviews, etc. can improve the recommendation performance. However, popular datasets do not contain these data which results in the proposal of new datasets (e.g. Instagram dataset) that contain these data. The variety of datasets used in different models makes it very difficult to compare the performance against different state-of-the-art models. A benchmark dataset containing data from various modalities can resolve these issues creating a solid ground to assess the performance across different models.

9.2. Developing data anonymization techniques for LBSN

Most of the datasets used by the models mentioned in our paper are from deprecated LBSN platforms such as Gowalla, Brightkite, and Weeplaces; and upon the dissolution of such platforms, the authority of those platforms distributed their datasets for research purposes. Thus existing datasets used in POI research are quite old. Such older datasets lack the features that reflect modern user-behavior and as a result, models trained using such datasets may not perform as expected on modern social media platforms. Furthermore, modern social media platforms tend to restrict their datasets from public usage because of privacy issues. Researchers need to put more emphasis on collecting anonymous user data by obfuscating user identity so that the privacy of the users can be preserved.

9.3. Incorporating online learning

After analyzing different models in this review paper, we have observed that most of the POI recommendation models in the literatures use offline learning techniques i.e. techniques that allow models to be trained only once with the available datasets before deployment. In a real-world scenario, this strategy is not optimal because users are continuously generating thousands of new check-in data every day that carry crucial information about changes in user preferences. An online learning strategy that can update the models as new data gets available is thus of significant importance so that the model can provide optimal recommendation performance over time even in changing circumstances.

9.4. Developing privacy preserving POI recommendation techniques

Like many other location-based services, user privacy is a major bottleneck for the proliferation of POI recommendation systems. In many cases, users tend to not share their GPS traces because adversaries can reveal many sensitive and private information of the user from those location traces. Thus there is an increasing need to devise a privacy-preserved POI recommendation system. A couple of non-deep learning approaches- Liu et al. [143] and Chen et al. [144] preserve the privacy of user data. However, since these approaches also require aggregating data from users in a central location, a lot of privacy concerns still exist. We envision that POI recommendation techniques can exploit a new domain of privacy-preserving learning, namely federated learning, that does not require accumulating user data in any central site.

Essentially, federated learning is a machine learning technique that trains an algorithm across multiple devices using their local data samples, without exchanging them. In traditional distributed learning, all data are distributed across multiple centralized servers which does not ensure the privacy of user data and data security. But in federated learning, a user does not need to share their data. They can train a model using their local data and share the model parameters. So, federated learning ensures data privacy and data sparsity issues. In this context, Wang et al. [43] recently proposed a model LLRec by generating teacher and student models. While the idea is close to the notion of federated learning, it does not fully take advantage of parameter sharing like federated learning. Thus, developing federated learning-friendly models is an interesting research direction in this domain.

9.5. Recommending POI for social groups

Previous works mostly focus on personalized POI recommendations where users' historical check-ins, as well as other attributes, are taken into account. However, POI recommendation for a group of users has mostly been out of focus in the literatures. Recommendation for a social group is significantly different from personalized POI recommendations, since each group member may have different preferences for choosing POIs. The social aspect becomes particularly important when recommending POIs for a group of users, which most of the present personalized POI recommendation models are unable to handle properly. Wang et al. [145] used matrix factorization and clustering techniques for group POI recommendation. But a simple model like this fails to utilize external features which largely limits the model performance. Consequently, a deep learning-based POI recommendation model for social groups can be particularly helpful for group tourism and developing such a model can be worth of research interest.

10. Conclusion

In this paper, we have presented a comprehensive survey on deep learning-based POI recommendation systems. We have presented insightful findings of a plethora of recent research papers in this emerging area of research. We have categorized the POI recommendation models based on different deep learning paradigms and compare their competitive (dis)advantages. We have also presented the performance results of these techniques with respect to different performance metrics for different real datasets. We have identified different factors that impact the POI recommendations and provided a tabular analysis of each factor. Finally, we have discussed a series of future works on POI recommendations that provide a guideline for new researchers in this domain. To the best of our knowledge, this work is the first comprehensive review of deep learning-based POI recommendations.

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.

Acknowledgements

References

[1] S. Rendle, C. Freudenthaler, L. Schmidt-Thieme, Factorizing personalized markov chains for next-basket recommendation, in: Proceedings of the 19th International Conference on World Wide Web, 2010, pp. 811–820.

[2] C. Cheng, H. Yang, M.R. Lyu, I. King, Where you like to go next: successive point-of-interest recommendation, in: Twenty-Third International Joint Conference on Artificial Intelligence, 2013.

[3] M. Chen, Y. Liu, X. Yu, Nlpmm: a next location predictor with markov modeling, in: Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer, 2014, pp. 186–197. .

[4] J.-D. Zhang, C.-Y. Chow, Y. Li, Lore: exploiting sequential influence for location recommendations, in: Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2014, pp. 103–112. .

[5] J. Ye, Z. Zhu, H. Cheng, What's your next move: user activity prediction in location-based social networks, in: Proceedings of the 2013 SIAM International Conference on Data Mining, SIAM, 2013, pp. 171–179.

[6] R. Yue-Qiang, W. Ze, S. Xiao-Na, S. Shi-Min, A multi-element hybrid location recommendation algorithm for location based social networks, IEEE Access 7 (2019) 100416–100427.

[7] Y. Wen, J. Zhang, Q. Zeng, X. Chen, F. Zhang, Loc2vec-based cluster-level transition behavior mining for successive poi recommendation, IEEE Access 7 (2019) 109311–109319.

[8] R. Mao, Z. Han, Z. Liu, Y. Liu, X. Lv, P. Xuan, Prme-gts: a new successive poi recommendation model with temporal and social influences, in: International Conference on Advanced Data Mining and Applications, Springer, 2019, pp. 266–274. .

[9] X. Xiong, S. Qiao, Y. Li, N. Han, G. Yuan, Y. Zhang, A point-of-interest suggestion algorithm in multi-source geo-social networks, Eng. Appl. Artif. Intell. 88 (2020) 103374.

[10] Y. Koren, R. Bell, C. Volinsky, Matrix factorization techniques for recommender systems, Computer 42 (8) (2009) 30–37.

[11] D. Lian, C. Zhao, X. Xie, G. Sun, E. Chen, Y. Rui, Geomf: joint geographical modeling and matrix factorization for point-of-interest recommendation, in: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014, pp. 831–840.

[12] Y. Liu, W. Wei, A. Sun, C. Miao, Exploiting geographical neighborhood characteristics for location recommendation, in: Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, 2014, pp. 739–748.

[13] Z. Yao, Y. Fu, B. Liu, Y. Liu, H. Xiong, Poi recommendation: a temporal matching between poi popularity and user regularity, in: 2016 IEEE 16th International Conference on Data Mining (ICDM), 2016, pp. 549–558. .

[14] P. Zhao, X. Xu, Y. Liu, Z. Zhou, K. Zheng, V.S. Sheng, H. Xiong, Exploiting hierarchical structures for poi recommendation, in: 2017 IEEE International Conference on Data Mining (ICDM), IEEE, 2017, pp. 655–664.

[15] H. Wang, H. Shen, W. Ouyang, X. Cheng, Exploiting poi-specific geographical influence for point-of-interest recommendation, IJCAI (2018) 3877–3883.

[16] L. Cai, J. Xu, J. Liu, T. Pei, Integrating spatial and temporal contexts into a factorization model for poi recommendation, Int. J. Geogr. Inf. Sci. 32 (3) (2018) 524–546.

[17] H. Ying, J. Wu, G. Xu, Y. Liu, T. Liang, X. Zhang, H. Xiong, Time-aware metric embedding with asymmetric projection for successive poi recommendation, World Wide Web 22 (5) (2019) 2209–2224.

[18] Z. Pan, L. Cui, X. Wu, Z. Zhang, X. Li, G. Chen, Deep potential geo-social relationship mining for point-of-interest recommendation, IEEE Access 7 (2019) 99496–99507.

[19] Y. Su, X. Li, B. Liu, D. Zha, J. Xiang, W. Tang, N. Gao, Fgcrec: fine-grained geographical characteristics modeling for point-of-interest recommendation, in: ICC 2020-2020 IEEE International Conference on Communications (ICC), IEEE, 2020, pp. 1–6. .

[20] H.A. Rahmani, M. Aliannejadi, M. Baratchi, F. Crestani, Joint geographical and temporal modeling based on matrix factorization for point-of-interest recommendation, in: European Conference on Information Retrieval, Springer, 2020, pp. 205–219.

[21] S. Rendle, C. Freudenthaler, Z. Gantner, L. Schmidt-Thieme, Bpr: Bayesian personalized ranking from implicit feedback, arXiv preprint arXiv:1205.2618 (2012)..

[22] J. He, X. Li, L. Liao, Category-aware next point-of-interest recommendation via listwise bayesian personalized ranking., in: IJCAI, vol. 17, 2017, pp. 1837–1843..

[23] J. He, X. Li, L. Liao, M. Wang, Inferring continuous latent preference on transition intervals for next point-of-interest recommendation, in: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer, 2018, pp. 741–756.

[24] Q. Cui, Y. Tang, S. Wu, L. Wang, Distance2pre: personalized spatial preference for next point-of-interest prediction, in: Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer, 2019, pp. 289–301.

[25] S. Zhao, T. Zhao, H. Yang, M.R. Lyu, I. King, STELLAR: spatial-temporal latent ranking for successive point-of-interest recommendation, in: D. Schuurmans, M.P. Wellman (Eds.), Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12–17, 2016, Phoenix, Arizona, USA, 2016, pp. 315–322. .

[26] S. Zhao, T. Zhao, I. King, M.R. Lyu, Geo-teaser: geo-temporal sequential embedding rank for point-of-interest recommendation, in: Proceedings of the 26th International Conference on World Wide Web Companion, 2017, pp. 153–162. .

[27] S. Zhao, I. King, M.R. Lyu, Geo-pairwise ranking matrix factorization model for point-of-interest recommendation, in: International Conference on Neural Information Processing, Springer, 2017, pp. 368–377.

[28] S. Zhao, M.R. Lyu, I. King, Stellar: spatial-temporal latent ranking model for successive poi recommendation, in: Point-of-Interest Recommendation in Location-Based Social Networks, Springer, 2018, pp. 79–94..

[29] R. Gao, J. Li, X. Li, C. Song, Y. Zhou, A personalized point-of-interest recommendation model via fusion of geo-social information, Neurocomputing 273 (2018) 159–170.

[30] W. Liu, Z.-J. Wang, B. Yao, M. Nie, J. Wang, R. Mao, J. Yin, Geographical relevance model for long tail point-of-interest recommendation, in: International Conference on Database Systems for Advanced Applications, Springer, 2018, pp. 67–82.

[31] M. Ye, P. Yin, W.-C. Lee, Location recommendation for location-based social networks, in: Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2010, pp. 458–461.

[32] M. Ye, P. Yin, W.-C. Lee, D.-L. Lee, Exploiting geographical influence for collaborative point-of-interest recommendation, in: Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2011, pp. 325–334.

[33] C. Yang, L. Bai, C. Zhang, Q. Yuan, J. Han, Bridging collaborative filtering and semi-supervised learning: a neural approach for poi recommendation, in: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017, pp. 1245–1254.

[34] J. Li, G. Liu, C. Yan, C. Jiang, Lori: a learning-to-rank-based integration method of location recommendation, IEEE Trans. Comput. Social Syst. 6 (3) (2019) 430–440.

[35] S. Qiao, N. Han, J. Zhou, R.-H. Li, C. Jin, L.A. Gutierrez, Socialmix: a familiarity-based and preference-aware location suggestion approach, Eng. Appl. Artif. Intell. 68 (2018) 192–204.

[36] W. Wang, J. Chen, J. Wang, J. Chen, Z. Gong, Geography-aware inductive matrix completion for personalized point-of-interest recommendation in smart cities, IEEE Internet Things J. 7 (5) (2019) 4361–4370.

[37] R. Baral, S.S. Iyengar, X. Zhu, T. Li, P. Sniatala, Hirecs: a hierarchical contextual location recommendation system, IEEE Trans. Comput. Social Syst. 6 (5) (2019) 1020–1037.

[38] Y. Xu, X. Li, J. Li, C. Wang, R. Gao, Y. Yu, Ssser: spatiotemporal sequential and social embedding rank for successive point-of-interest recommendation, IEEE Access 7 (2019) 156804–156823.

[39] M. Chen, Y. Zuo, X. Jia, Y. Liu, X. Yu, K. Zheng, Cem: a convolutional embedding model for predicting next locations, IEEE Trans. Intell. Transp. Syst. (2020).

[40] Q. Liu, S. Wu, L. Wang, T. Tan, Predicting the next location: a recurrent model with spatial and temporal contexts, in: Thirtieth AAAI Conference on Artificial Intelligence, 2016.

[41] D. Yang, B. Fankhauser, P. Rosso, P. Cudre-Mauroux, Location prediction over sparse user mobility traces using rnns: flashback in hidden states!, in: in: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20, 2020, pp 2184–2190.

[42] K. Zhao, Y. Zhang, H. Yin, J. Wang, K. Zheng, X. Zhou, C. Xing, Discovering subsequence patterns for next poi recommendation, in: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, 2020, pp. 3216–3222. .

[43] Q. Wang, H. Yin, T. Chen, Z. Huang, H. Wang, Y. Zhao, N.Q. Viet Hung, Next point-of-interest recommendation on resource-constrained mobile devices, Proceedings of The Web Conference 2020 (2020) 906–916.

[44] K. Sun, T. Qian, T. Chen, Y. Liang, Q.V.H. Nguyen, H. Yin, Where to go next: modeling long-and short-term user preferences for point-of-interest recommendation, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, 2020, pp. 214–221..

[45] R. Li, Y. Shen, Y. Zhu, Next point-of-interest recommendation with temporal and multi-level context attention, in: 2018 IEEE International Conference on Data Mining (ICDM), IEEE, 2018, pp. 1110–1115.

[46] L. Zhang, Z. Sun, J. Zhang, Y. Lei, C. Li, Z. Wu, H. Kloeden, F. Klanner, An interactive multi-task learning framework for next poi recommendation with uncertain check-ins, CAL 301 (985) (2020) 13954.

[47] F. Yu, L. Cui, W. Guo, X. Lu, Q. Li, H. Lu, A category-aware deep model for successive poi recommendation on sparse check-in data, Proceedings of The Web Conference 2020 (2020) 1264–1274.

[48] J. Feng, Y. Li, C. Zhang, F. Sun, F. Meng, A. Guo, D. Jin, Deepmove: predicting human mobility with attentional recurrent networks, in: Proceedings of the 2018 World Wide Web Conference, 2018, pp. 1459–1468.

[49] J. Manotumruksa, C. Macdonald, I. Ounis, A contextual attention recurrent architecture for context-aware venue recommendation, in: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, 2018, pp. 555–564. .

[50] K. Kala, M. Nandhini, Context-category specific sequence aware point-of-interest recommender system with multi-gated recurrent unit, J. Ambient Intell. Hum. Comput., Springer, 2019..

[51] D. Lian, Y. Wu, Y. Ge, X. Xie, E. Chen, Geography-aware sequential location recommendation, in: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2020, pp. 2009–2019.

[52] Q. Guo, J. Qi, Sanst: a self-attentive network for next point-of-interest recommendation, arXiv preprint arXiv:2001.10379 (2020)..

[53] B. Liu, T. Qian, B. Liu, L. Hong, Z. You, Y. Li, Learning spatiotemporal-aware representation for poi recommendation, arXiv preprint arXiv:1704.08853 (2017)..

[54] G. Christoforidis, P. Kefalas, A.N. Papadopoulos, Y. Manolopoulos, Reline: point-of-interest recommendations using multiple network embeddings, arXiv preprint arXiv:1902.00773 (2019)..

[55] X. Xiong, F. Xiong, J. Zhao, S. Qiao, Y. Li, Y. Zhao, Dynamic discovery of favorite locations in spatio-temporal social networks, Inf. Process. Manage. 57 (6) (2020) 102337.

[56] J. Bao, Y. Zheng, D. Wilkie, M. Mokbel, Recommendations in location-based social networks: a survey, GeoInformatica 19 (3) (2015) 525–565.

[57] Y. Liu, T.-A.N. Pham, G. Cong, Q. Yuan, An experimental evaluation of point-of-interest recommendation in location-based social networks (2017)..

[58] S. Zhao, I. King, M.R. Lyu, A survey of point-of-interest recommendation in location-based social networks, arXiv preprint arXiv:1607.00647 (2016)..

[59] X. Zheng, J. Han, A. Sun, A survey of location prediction on twitter, IEEE Trans. Knowl. Data Eng. 30 (9) (2018) 1652–1671.

[60] S. Liu, User modeling for point-of-interest recommendations in location-based social networks: the state of the art, Mobile Inf. Syst. 2018 (2018).

[61] S. Wang, J. Cao, P. Yu, Deep learning for spatio-temporal data mining: a survey, IEEE Trans. Knowl. Data Eng. (2020).

[62] S. Xu, X. Fu, J. Cao, B. Liu, Z. Wang, Survey on user location prediction based on geo-social networking data, World Wide Web (2020) 1–44.

[63] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997), https://doi.org/10.1162/neco.1997.9.8.1735, 1735–80.

[64] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using rnn encoder-decoder for statistical machine translation, arXiv preprint arXiv:1406.1078 (2014)..

[65] D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate, arXiv preprint arXiv:1409.0473 (2014)..

[66] I. Sutskever, O. Vinyals, Q.V. Le, Sequence to sequence learning with neural networks, in: Advances in Neural Information Processing Systems, 2014, pp. 3104–3112..

[67] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, in: Advances in Neural Information Processing Systems, 2017, pp. 5998–6008..

[68] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: Advances in Neural Information Processing Systems, 2014, pp. 2672–2680..

[69] T.L. Griffiths, M.I. Jordan, J.B. Tenenbaum, D.M. Blei, Hierarchical topic models and the nested chinese restaurant process, in: Advances in Neural Information Processing Systems, 2004, pp. 17–24..

[70] S. El Hihi, Y. Bengio, Hierarchical recurrent neural networks for long-term dependencies, in: Advances in Neural Information Processing Systems, 1996, pp. 493–499..

[71] A. Kusupati, M. Singh, K. Bhatia, A. Kumar, P. Jain, M. Varma, Fastgrnn: a fast, accurate, stable and tiny kilobyte sized gated recurrent neural network, in: Advances in Neural Information Processing Systems, 2018, pp. 9017–9028. .

[72] X. Ma, P. Zhang, S. Zhang, N. Duan, Y. Hou, M. Zhou, D. Song, A tensorized transformer for language modeling, Adv. Neural Inf. Process. Syst. (2019) 2232–2242.

[73] G. Hinton, O. Vinyals, J. Dean, Distilling the knowledge in a neural network, arXiv preprint arXiv:1503.02531 (2015)..

[74] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, arXiv preprint arXiv:1301.3781 (2013)..

[75] Z. Huang, X. Lin, H. Liu, B. Zhang, Y. Chen, Y. Tang, Deep representation learning for location-based recommendation, IEEE Trans. Comput. Social Syst. (2020).

[76] J. Liao, T. Liu, M. Liu, J. Wang, Y. Wang, H. Sun, Multi-context integrated deep neural network model for next location prediction, IEEE Access 6 (2018) 21980–21990.

[77] Y. Wu, K. Li, G. Zhao, X. Qian, Long-and short-term preference learning for next poi recommendation, in: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, 2019, pp. 2301–2304.

[78] Y. Wu, K. Li, G. Zhao, Q. Xueming, Personalized long-and short-term preference learning for next poi recommendation, IEEE Trans. Knowl. Data Eng. (2020).

[79] K.D. Doan, G. Yang, C.K. Reddy, An attentive spatio-temporal neural model for successive point of interest recommendation, in: Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer, 2019, pp. 346–358.

[80] Q. Guo, Z. Sun, J. Zhang, Y.-L. Theng, An attentional recurrent neural network for personalized next location recommendation, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, 2020, pp. 83–90..

[81] H. Wang, H. Shen, X. Cheng, Modeling poi-specific spatial-temporal context for point-of-interest recommendation, in: Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer, 2020, pp. 130–141.

[82] L. Huang, Y. Ma, S. Wang, Y. Liu, An attention-based spatiotemporal lstm network for next poi recommendation, IEEE Trans. Serv. Comput. (2019).

[83] L. Li, Y. Liu, J. Wu, L. He, G. Ren, Multi-modal representation learning for successive poi recommendation, in: Asian Conference on Machine Learning, 2019, pp. 441–456..

[84] K. Yang, J. Zhu, Next poi recommendation via graph embedding representation from h-deepwalk on hybrid network, IEEE Access 7 (2019) 171105–171113.

[85] S. Zhao, X. Chen, I. King, M.R. Lyu, Personalized sequential check-in prediction: Beyond geographical and temporal contexts, in: in: 2018 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2018, pp. 1–6.

[86] T. Liu, J. Liao, Z. Wu, Y. Wang, J. Wang, Exploiting geographical-temporal awareness attention for next point-of-interest recommendation, Neurocomputing (2020).

[87] T. Liu, J. Liao, Z. Wu, Y. Wang, J. Wang, A geographical-temporal awareness hierarchical attention network for next point-of-interest recommendation, in: in: Proceedings of the 2019 on International Conference on Multimedia Retrieval, 2019, pp. 7–15.

[88] C.H. Liu, Y. Wang, C. Piao, Z. Dai, Y. Yuan, G. Wang, D. Wu, Time-aware location prediction by convolutional area-of-interest modeling and memory-augmented attentive lstm, IEEE Trans. Knowl. Data Eng. (2020).

[89] B. Chang, Y. Koh, D. Park, J. Kang, Content-aware successive point-of-interest recommendation, in: Proceedings of the 2020 SIAM International Conference on Data Mining, SIAM, 2020, pp. 100–108.

[90] D. Kong, F. Wu, Hst-lstm: A hierarchical spatial-temporal long-short term memory network for location prediction., in: IJCAI, Vol. 18, 2018, pp. 2341–2347..

[91] P. Zhao, A. Luo, Y. Liu, F. Zhuang, J. Xu, Z. Li, V.S. Sheng, X. Zhou, Where to go next: A spatio-temporal gated network for next poi recommendation, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, 2019, pp. 5877–5884..

[92] P. Zhao, A. Luo, Y. Liu, F. Zhuang, J. Xu, Z. Li, V.S. Sheng, X. Zhou, Where to go next: A spatio-temporal gated network for next poi recommendation, IEEE Trans. Knowl. Data Eng. (2020).

[93] W.-C. Kang, J. McAuley, Self-attentive sequential recommendation, in: in: 2018 IEEE International Conference on Data Mining (ICDM), IEEE, 2018, pp. 197–206.

[94] M. Xie, H. Yin, H. Wang, F. Xu, W. Chen, S. Wang, Learning graph-based poi embedding for location-based recommendation, in: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, 2016, pp. 15–24.

[95] Y. Lin, Z. Liu, M. Sun, Y. Liu, X. Zhu, Learning entity and relation embeddings for knowledge graph completion, AAAI (2015) 2181–2187.

[96] G. Christoforidis, P. Kefalas, A. Papadopoulos, Y. Manolopoulos, Recommendation of points-of-interest using graph embeddings, in: in: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), IEEE, 2018, pp. 31–40.

[97] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, Q. Mei, Line: Large-scale information network embedding, in: Proceedings of the 24th international conference on world wide web, 2015, pp. 1067–1077.

[98] H. Lin, J. Jia, J. Qiu, Y. Zhang, G. Shen, L. Xie, J. Tang, L. Feng, T.-S. Chua, Detecting stress based on social interactions in social networks, IEEE Trans. Knowl. Data Eng. 29 (9) (2017) 1820–1833.

[99] X. Xiong, Y. Li, S. Qiao, N. Han, Y. Wu, J. Peng, B. Li, An emotional contagion model for heterogeneous social media with multiple behaviors, Physica A 490 (2018) 185–202.

[100] L. Zhang, Z. Sun, J. Zhang, H. Kloeden, F. Klanner, Modeling hierarchical category transition for next poi recommendation with uncertain check-ins, Inf. Sci. 515 (2020) 169–190.

[101] Y. Liu, C. Liu, B. Liu, M. Qu, H. Xiong, Unified point-of-interest recommendation with temporal interval assessment, in: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 1015–1024.

[102] M. Chen, Y. Zhao, Y. Liu, X. Yu, K. Zheng, Modeling spatial trajectories with attribute representation learning, IEEE Trans. Knowl. Data Eng. (2020).

[103] Y. Qiao, X. Luo, C. Li, H. Tian, J. Ma, Heterogeneous graph-based joint representation learning for users and pois in location-based social network, Information Processing & Management 57 (2) (2020) 102151.

[104] W. Liu, Z.-J. Wang, B. Yao, J. Yin, Geo-alm: Poi recommendation by fusing geographical information and adversarial learning mechanism., in: IJCAI, 2019, pp. 1807–1813. .

[105] F. Zhou, R. Yin, K. Zhang, G. Trajcevski, T. Zhong, J. Wu, Adversarial point-of-interest recommendation, The World Wide Web Conference (2019) 3462–34618.

[106] Q. Gao, F. Zhang, F. Yao, A. Li, L. Mei, F. Zhou, Adversarial mobility learning for human trajectory classification, IEEE Access 8 (2020) 20563–20576.

[107] Y. Li, J.-J. Xu, P.-P. Zhao, J.-H. Fang, W. Chen, L. Zhao, Atlrec: An attentional adversarial transfer learning network for cross-domain recommendation, Journal of Computer Science and Technology 35 (4) (2020) 794–808.

Md. Ashraful Islam, Mir Mahathir Mohammad, Sarkar Snigdha Sarathi Das et al.

Neurocomputing 472 (2022) 306–325

[108] Y. Wang, M. Chen, X. Yu, Y. Liu, Lce: A location category embedding model for predicting the category labels of pois, in: International Conference on Neural Information Processing, Springer, 2017, pp. 710–720. .

[109] B. Chang, Y. Park, D. Park, S. Kim, J. Kang, Content-aware hierarchical point-of-interest embedding model for successive poi recommendation., IJCAI (2018) 3301–3307.

[110] R. Ding, Z. Chen, X. Li, Spatial-temporal distance metric embedding for time-specific poi recommendation, IEEE Access 6 (2018) 67035–67045.

[111] D. Xi, F. Zhuang, Y. Liu, J. Gu, H. Xiong, Q. He, Modelling of bi-directional spatio-temporal dependence and users' dynamic preferences for missing poi check-in identification, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, 2019, pp. 5458–5465..

[112] X. Zhou, C. Mascolo, Z. Zhao, Topic-enhanced memory networks for personalised point-of-interest recommendation, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 3018–3028.

[113] S. Feng, L.V. Tran, G. Cong, L. Chen, J. Li, F. Li, Hme: A hyperbolic metric embedding approach for next-poi recommendation, in: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2020, pp. 1429–1438. .

[114] H. Luo, J. Zhou, Z. Bao, S. Li, J.S. Culpepper, H. Ying, H. Liu, H. Xiong, Spatial object recommendation with hints: When spatial granularity matters, in: in: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2020, pp. 781–790.

[115] Y. Zhang, G. Liu, A. Liu, Y. Zhang, Z. Li, X. Zhang, Q. Li, Personalized geographical influence modeling for poi recommendation, IEEE Intell. Syst. (2020).

[116] X. Zhou, A. Noulas, C. Mascolo, Z. Zhao, Discovering latent patterns of urban cultural interactions in wechat for modern city planning, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018, pp. 1069–1078.

[117] C. De Sa, A. Gu, C. Ré, F. Sala, Representation tradeoffs for hyperbolic embeddings, Proceedings of machine learning research 80 (2018) 4460.

[118] M. Nickel, D. Kiela, Poincaré embeddings for learning hierarchical representations, in: Advances in neural information processing systems, 2017, pp. 6338–6347..

[119] M. Nickel, D. Kiela, Learning continuous hierarchies in the lorentz model of hyperbolic geometry, arXiv preprint arXiv:1806.03417 (2018)..

[120] C. Gulcehre, M. Denil, M. Malinowski, A. Razavi, R. Pascanu, K.M. Hermann, P. Battaglia, V. Bapst, D. Raposo, A. Santoro, et al., Hyperbolic attention networks, arXiv preprint arXiv:1805.09786 (2018)..

[121] A.A. Ungar, Analytic hyperbolic geometry: Mathematical foundations and applications, World Scientific, 2005.

[122] S. Feng, X. Li, Y. Zeng, G. Cong, Y.M. Chee, Personalized ranking metric embedding for next new poi recommendation, in: IJCAI'15 Proceedings of the 24th International Conference on Artificial Intelligence, ACM, 2015, pp. 2069–2075.

[123] X. Chen, J. Xu, R. Zhou, P. Zhao, C. Liu, J. Fang, L. Zhao, S2r-tree: a pivot-based indexing structure for semantic-aware spatial keyword search, GeoInformatica 24 (1) (2020) 3–25.

[124] D. Massimo, F. Ricci, Harnessing a generalised user behaviour model for next-poi recommendation, in: Proceedings of the 12th ACM Conference on Recommender Systems, 2018, pp. 402–406.

[125] D. Ramachandran, E. Amir, Bayesian inverse reinforcement learning., in: IJCAI, Vol. 7, 2007, pp. 2586–2591..

[126] G. Neu, C. Szepesvári, Training parsers by inverse reinforcement learning, Machine learning 77 (2–3) (2009) 303.

[127] B.D. Ziebart, A.L. Maas, J.A. Bagnell, A.K. Dey, Maximum entropy inverse reinforcement learning., in: Aaai, Vol. 8, Chicago, IL, USA, 2008, pp. 1433–1438..

[128] T. Cai, J. Li, A.S. Mian, T. Sellis, J.X. Yu, et al., Target-aware holistic influence maximization in spatial social networks, IEEE Trans. Knowl. Data Eng. (2020).

[129] R. Li, X. Wu, X. Wu, W. Wang, Few-shot learning for new user recommendation in location-based social networks, Proceedings of The Web Conference (2020, 2020,) 2472–2478.

[130] R. Baral, S. Iyengar, T. Li, N. Balakrishnan, Close: contextualized location sequence recommender, in: Proceedings of the 12th ACM conference on recommender systems, 2018, pp. 470–474.

[131] J. Huang, Y. Liu, Y. Chen, C. Jia, Dynamic recommendation of poi sequence responding to historical trajectory, ISPRS International Journal of Geo-Information 8 (10) (2019) 433.

[132] Y.-S. Lu, J.-L. Huang, Glr: A graph-based latent representation model for successive poi recommendation, Future Generation Computer Systems 102 (2020) 230–244.

[133] Q. Le, T. Mikolov, Distributed representations of sentences and documents, in: in: International conference on machine learning, 2014, pp. 1188–1196.

[134] M.-F. Wang, Y.-S. Lu, J.-L. Huang, Spent: A successive poi recommendation method using similarity-based poi embedding and recurrent neural network with temporal influence, in: 2019 IEEE International Conference on Big Data and Smart Computing (BigComp), IEEE, 2019, pp. 1–8. .

[135] R. Baral, S. Iyengar, T. Li, X. Zhu, Hicaps: hierarchical contextual poi sequence recommender, in: Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2018, pp. 436–439.

[136] I.-C. Lin, Y.-S. Lu, W.-Y. Shih, J.-L. Huang, Successive poi recommendation with category transition and temporal influence, in: 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), Vol. 2, IEEE, 2018, pp. 57–62. .

[137] Y. Li, Y. Luo, Z. Zhang, S. Sadiq, P. Cui, Context-aware attention-based data augmentation for poi recommendation, in: in: 2019 IEEE 35th International Conference on Data Engineering Workshops (ICDEW), IEEE, 2019, pp. 177–184.

[138] P. Kefalas, Y. Manolopoulos, A time-aware spatio-textual recommender system, Expert Syst. Appl. 78 (2017) 396–406.

[139] B. Liu, Y. Fu, Z. Yao, H. Xiong, Learning geographical preferences for point-of-interest recommendation, in: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 2013, pp. 1043–1051.

[140] M. Ye, D. Shou, W.-C. Lee, P. Yin, K. Janowicz, On the semantic annotation of places in location-based social networks, in: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, 2011, pp. 520–528.

[141] H. Gao, J. Tang, X. Hu, H. Liu, Exploring temporal effects for location recommendation on location-based social networks, in: in: Proceedings of the 7th ACM conference on Recommender systems, 2013, pp. 93–100.

[142] Q. Yuan, G. Cong, Z. Ma, A. Sun, N.M. Thalmann, Time-aware point-of-interest recommendation, in: Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, 2013, pp. 363–372.

[143] A. Liu, W. Wang, Z. Li, G. Liu, Q. Li, X. Zhou, X. Zhang, A privacy-preserving framework for trust-oriented point-of-interest recommendation, IEEE Access 6 (2017) 393–404.

[144] C. Chen, Z. Liu, P. Zhao, J. Zhou, X. Li, Privacy preserving point-of-interest recommendation using decentralized matrix factorization, arXiv preprint arXiv:2003.05610 (2020)..

[145] X. Wang, M. Nguyen, J. Carr, L. Cui, K. Lim, A group preference-based privacy-preserving poi recommender system, ICT Express (2020).

Md. Ashraful Islam is a final year undergraduate student of the Department of Computer Science and Engineering (CSE) at Bangladesh University of Engineering and Technology (BUET). His research interests are Natural Language Processing (NLP), working with low resource native language (Bangla), Reinforcement Learning (RL) and Generative Adversarial Networks (GAN).

Mir Mahathir Mohammad is a final year undergraduate student of the Department of Computer Science and Engineering (CSE) at Bangladesh University of Engineering and Technology (BUET). His research interests are image processing, Natural Language Processing (NLP), working with automated sign language generation, Generative Adversarial Networks (GAN) and Reinforcement Learning (RL).

Sarkar Snigdha Sarathi Das has completed his B.Sc. Engg in CSE from the Department of Computer Science and Engineering (CSE) at Bangladesh University of Engineering and Technology (BUET). He is currently working as a full time research assistant in the same department.

Dr. Mohammed Eunus Ali is a Professor in the Department of Computer Science and Engineering (CSE) at Bangladesh University of Engineering and Technology (BUET), Dhaka since May 2014. He is the group leader of Data Science and Engineering Research Lab (DataLab) in the department of Computer Science and Engineering at BUET. He also worked as a Research Fellow and Visiting Research Scholars in the University of Melbourne, Monash University, Swinburne University of Technology, and RMIT University. Dr. Eunus is the recipient of the prestigious UGC Award in the year 2012 for his outstanding research contribution. Dr. Eunus's research falls in the intersection of data management and machine learning. His research areas cover a wide range of topics that include spatial databases, practical machine learning, Bangla NLP, and big data analytics. Dr. Eunus's research papers have been published in top ranking journals and conferences such as the VLDB Journal, Information Systems, WWWJ, DKE, ICDE, CIKM, EDBT, PVLDB, and UbiComp. He served as a Program Committee Members of many prestigious conferences that include SIGMOD, AAAI, and SIGSPATIAL. Dr. Eunus was the founding Chair of BUET ACM Chapter.