



HAL
open science

Few-shot and LightGCN Learning for Multi-Label Estimation of Lesser-Known Tourist Sites Using Tweets

Landy Rajaonarivo, Tsunenori Mine, Yutaka Arakawa

► **To cite this version:**

Landy Rajaonarivo, Tsunenori Mine, Yutaka Arakawa. Few-shot and LightGCN Learning for Multi-Label Estimation of Lesser-Known Tourist Sites Using Tweets. 2023 IEEE International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), Oct 2023, Venice, France. pp.103-110, 10.1109/WI-IAT59888.2023.00020 . hal-04436419

HAL Id: hal-04436419

<https://hal.science/hal-04436419v1>

Submitted on 3 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Few-shot and LightGCN learning for multi-label estimation of lesser-known tourist sites using tweets

Landy Rajaonarivo
Kyushu University
Japan
h.l.rajaonarivo.a03@m.kyushu-u.ac.jp

Tsunenori Mine
Kyushu University
Japan
mine@ait.kyushu-u.ac.jp

Yutaka Arakawa
Kyushu University
Japan
arakawa@ait.kyushu-u.ac.jp

Abstract—When a tourist wants to visit a site, he/she first asks about the category of the site. The availability of detailed information about the site allows him/her to tailor his/her visit to his/her preferences. This information is therefore essential for Point of Interest (POI) recommendation. However, it is rarely or never available for lesser-known POIs. Lesser-known POIs can be considered as places or events that local people know well, and that may be important to them, but that others do not know about. We propose an approach to estimate the categories of lesser-known POIs based on information from social media. The originality of this approach lies in the extraction of information and links between them, the encoding of the POIs, the representation of the data, and the combination of machine learning techniques such as *Few Shot Learning*, *LightGCN*, and *Clustering* for the estimation of POI categories. The results of the experiments would allow us to confirm that our approach can estimate POI categories and thus discover information about POIs that may be relevant. This approach would be useful for our future work on POI recommendations.

Index Terms—Multi-label estimation, Few Shot Learning, LightGCN, syntactic study, social media data

I. INTRODUCTION

Points of Interest (POIs) are locations that are of special or particular interest for a particular purpose¹. It can be a tourist attraction, a cultural, architectural, or recreational site, a scenic spot, or a site with historical or archaeological interest. A POI is considered lesser-known if it has features that are primarily of interest to tourists such as the culture of the local people, and categories of places that tourists need to visit such as restaurants, and accommodations, but if it is rarely mentioned on well-known tourism review sites such as Foursquare, TripAdvisor, Yelp, Booking, Google My Business, etc. Some examples of lesser-known POIs are presented in Section III-F. When a tourist wants to visit a place, he/she first learns about the type or category of POI in question. The category of information is then the minimum information that a tourist wants to know before deciding whether or not to visit the POI. This information is therefore essential in the field of POI and itinerary recommendation. It is therefore important to estimate the categories of POI if this information is not available. If a POI is lesser-known, it may take longer to find information about it than for POIs that are well or moderately known. This requires, for example, consulting

multiple information from different sources or people to obtain more reliable information. This is costly in terms of search and consultation time, but also in terms of the work involved in comparing and synthesizing the information.

Conventional learning uses a large amount of data, which plays an important role in the quality of learning performance. In our context, lesser-known POIs imply less available data. The challenge then is to propose a POI label estimation model that can learn with little data and is more efficient than conventional approaches. Recently, researchers have been interested in using minimal data for learning in the classification domain, such as few-shot, one-shot, and zero-shot learnings (FSL [1], OSL [2], ZSL [3]). These techniques allow a machine learning model to make predictions for new classes with limited labeled data. The choice of techniques depends on the specific problem and the amount of labeled data available for new categories or labels. We are interested in using Few Shot Learning (FSL) in our approach. This does not require us to have a large amount of data to do the learning, so it is less costly compared to the conventional approaches in terms of data collection, pre- and post-processing, and especially in terms of learning time and complexity. As mentioned earlier, lesser-known POIs are rarely present on well-known web sites for tourism, so we adopt the use of social network data. The use of this data could be interesting, as tourists are currently used to sharing information on social networks, such as the places they visited, their appreciation, information about the places, etc.

To encode the POI, we propose an approach based on the syntactic study of tweets by retrieving verbs, nouns, and adjectives and the semantic relations between them. We believe that these three parameters are important to characterize a POI. For example, even if we do not know the category of a given POI if the information about it repeatedly contains the verbs “visit” and “take pictures”, we can estimate that the POI might be a tourist place. The same is true for the adjectives “delicious”, which may characterize restaurant-type places. This approach reduces the complexity of the data used since we do not consider all the words in the texts, and the normalization at the level of verbs and adjectives, as well as nouns, reduces the variability of the words to be processed.

But how to estimate the category of POIs with this kind of information? We propose to use the LightGCN approach [4] to maximize the similarity of feature vectors of POIs belonging

This work was supported by Japan Society for the Promotion of Science.

¹<https://locationiq.com/glossary/points-of-interest>

to the same category and also to maximize the divergence of feature vectors of POIs belonging to different categories. This will help in estimating the categories of POIs not yet seen. We then propose a POI category estimation approach by implementing FSL and LightGCN and using tweets.

II. PROPOSED APPROACH

A. Problem definition

This section explains the data distribution method for learning and validation using the FSL technique and the learning models for estimating POI categories such as data embedding using LightGCN, clustering and probability estimation.

1) *Few Shot Learning*: FSL is defined as a machine learning model that is able to generalize from few training examples [5]. This approach are usually formed by N-way-K-shot classification and meta-learning. It allows us to determine the number of tasks to be set. Suppose we have C categories of POIs. To form a task, we randomly select N categories of C and K POIs from the POIs that have labels in the N selected categories. Suppose we have a set of training tasks \mathcal{T}_{train} and testing tasks \mathcal{T}_{test} , where $\mathcal{T}_{train} = \{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_m\}$ and $\mathcal{T}_{test} = \{\mathcal{T}_{m+1}, \dots, \mathcal{T}_{m+t}\}$. Each task consists of a data set for training called the “support set” and a data set for testing called the “query set”. Let us denote the data set used in task \mathcal{T}_i as $D_i = \{D_i^{supp}, D_i^{query}\}$ and the categories as $C_{\mathcal{T}_i}$. Therefore, the data set of the \mathcal{T}_{train} will be $D_{train} = \{D_1, D_2, \dots, D_m\}$ and the data test in \mathcal{T}_{test} will be $D_{test} = \{D_{m+1}, \dots, D_{m+t}\}$ where $|D_1^{supp}| = |D_2^{supp}| = \dots = |D_{m+t}^{supp}| = K$. Note that there are no common elements between the categories:

$$\bigcap_{\mathcal{T}_i \in \mathcal{T}} C_{\mathcal{T}_i} = \emptyset$$

Usually K is a small number (e.g. 1, 5, 10). Each dataset consists of a set of POIs, which we can define as follows $D_i = \{(x_j, y_j)\}$ where x_j is the j^{th} POI in D_j and y_j is the set of labels of x_j .

The meta-learning problem then consists in approximating the function f with parameters θ as follows for each task \mathcal{T}_i :

$$y \approx f(D_i^{supp}, x, \theta) \text{ where } (x, y) \in D_i^{query} \quad (1)$$

The result value of θ should minimise the sum of the loss function \mathcal{L} on each task, which is defined as follows:

$$\theta = \arg \min \sum_{D_i \in D_{train}} \sum_{(x,y) \in D_i^{query}} \mathcal{L}(f(D_i, x, \theta), y) \quad (2)$$

2) *Multi-label estimation*: The multi-label estimation approach consists of estimating the labels of a given POI that is not one of the POIs in either D^{supp} or D^{query} . Suppose that during the meta-learning of FSL, the process in each task generates a model of multi-label estimation that is specific to each task. Let $\mathcal{M}_i = model(\mathcal{T}_i, f, \theta_i)$ be the model generated by learning on task \mathcal{T}_i , where f is the multi-label estimation function and θ_i is the parameter that optimizes function f on task \mathcal{T}_i . Since a given POI can have several labels that can be found in different tasks, we propose to generate some models that are a combination of models. Suppose \mathcal{M}_{ij} is the

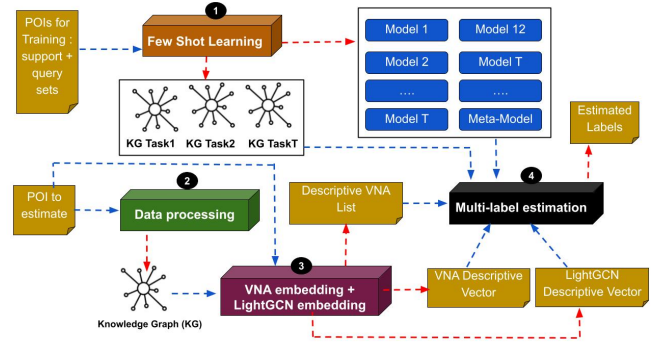


Fig. 1. Global approach to multi-label estimation

model generated from the learning using the dataset of the combination of tasks \mathcal{T}_i and \mathcal{T}_j , \mathcal{M}_{ij} is defined as follows: $\mathcal{M}_{ij} = model(\mathcal{T}_i \cup \mathcal{T}_j, f, \theta)$.

In order to estimate the labels of a given POI x , the first step is to select the most appropriate model that allows to obtain the best value of the estimation rate. Let g be the function allowing to evaluate the correspondence of a given model to a given POI, \mathcal{M} be the set of all available models, x_f be the features of x and X_f^i be the common features of the POIs in D_i^{supp} . The function g is defined as follows:

$$m = \arg \max_{\mathcal{M}_i \in \mathcal{M}} g(\mathcal{M}_i, X_f^i, x_f) \quad (3)$$

The next step is to estimate the label of the POI x using the estimated appropriate model. The label estimation is defined as follows:

$$\hat{y} = f(D_m^{supp}, x, \theta_m) \quad (4)$$

B. Proposed multi-label classification models for POI category estimation

We propose an approach based on Few Shot Learning (FSL) for multi-label estimation. Fig. 1 illustrates the global architecture of our proposed multi-label estimation approach. It consists of four main engines: data processing, embedding generation, FSL and multi-label estimation.

1) *Data processing*: The data processing is carried out by four main following treatments: (i) tweet collection, (ii) syntactical analysis of collection, (iii) collection filtering, and (iv) syntactical knowledge graph (KG) generation or updating. These treatments are applied to each POI. For the tweet collection, we used the Twitter API² dedicated to academic research.

The syntactic analysis study consists of analysing the syntactic dependencies in the tweets. Since our study area is in Japan and 90% of the tweets related to lesser-known POIs are written in Japanese, we are interested in studying syntactic dependencies on Japanese texts. Fig. 2 shows an example of dependency parsing of a sentence. The input text is “I am eating delicious steak and drinking craft beer at ABC”. We can see that the sentence is tokenized and each token has a type

²<https://developer.twitter.com/en/products/twitter-api/academic-research>

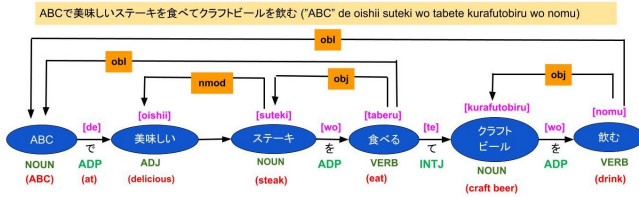


Fig. 2. Dependency parsing

(NOUN, VERB, ADJ, ADP). The tokens are related to each other according to their syntactic relations in the sentence. Two tokens are linked if they occur one after the other in a sentence, or if there are syntactic dependencies between them (*nmod*, *obj*, *obl*³, etc.), or if they are linked by a particle (“*suteki*” and “*taberu*” are linked by a particle “*wo*”). The syntactic dependencies in this example are *nmod*(nominal modifier), *obj*(object), *obl*(indirect nominal). There are libraries for performing syntactic analysis of Japanese texts. We use GiNZA⁴ (Japanese Universal Dependencies Models), which is based on spaCy⁵. The choice was arbitrary.

The purpose of the collection filtering system is to remove tweets that do not present any information. For a given tweet, we check for dependency structures by referring to the universal dependencies for Japanese presented in [6]. Japanese is a SOV language, which means that the order of words in a sentence is S (subject) - O (object) - V (verb) (e.g. kare (S) wa biru (O) wo nomimasu (V): he drinks beer). The structure of Japanese sentences is flexible. For example, the subject can be omitted if it is obvious in the context or situation. Also, the order of the subject and object(s) can vary. Since tweets are not well structured, we consider a tweet if at least one basic dependency structure is present (e.g. OV, SV, SOV, etc.), otherwise we ignore it.

Knowledge graph generation consists of automatically generating or updating syntactic knowledge graphs from tweets. These graphs are formed by nodes and edges, where the nodes represent the word (verbs, nouns, or adjectives) detected during the syntactic analysis, while the edges represent two group types of relationships between words, such as syntactic relationships (e.g. *nmod*, *obj*, etc.) and proximity relationships between two words (words that follow each other in a sentence). Let $G = (\mathcal{N}, \mathcal{E})$ denote a syntactic KG, \mathcal{N} a set of nodes and \mathcal{E} a set of edges that make up the graph. Let \mathcal{N}_{vb} , \mathcal{N}_{nn} , \mathcal{N}_{adj} be the respective sets of verb, noun and adjective type nodes, where $\mathcal{N} = \mathcal{N}_{vb} \cup \mathcal{N}_{nn} \cup \mathcal{N}_{adj}$. \mathcal{N} is defined as follows: $\mathcal{N} = \{v = (v_{name}, v_{type}, v_{occ}, v_{tweet_ids})\}$ where v_{name} , v_{type} , v_{occ} and v_{tweet_ids} are respectively the name of the node, its type, its frequency of occurrence and the list of tweets containing it. $\mathcal{E} = \{e = (e_{type}, v_i, v_j, w_{ij}, pr_{ij}), v_i, v_j \in \mathcal{N}, w_{ij} \in \mathbb{N}\}$ where e_{type} is the type of relation, w_{ij} is the number of times v_i and v_j are syntactically linked, and pr_{ij}

is the list of particles linking v_i and v_j according to the tweet collection.

2) *Embedding generation*: Data embedding consists of encoding a POI with a descriptive vector. In our approach, we propose to embed POIs by verbs, nouns, adjectives (VNA). Taking the example shown in Fig. 2, the POI named ABC can be characterised by verbs (eat and drink), nouns (steak and craft beer) and adjective (delicious). Data embedding involves three steps, such as generating descriptive VNAs (verbs, nouns, adjectives), determining dimensions and embedding POIs according to the dimensions. The selection of VNAs involves the TF-IDF technique [7], the aim of which is to be able to prioritise the most specific VNAs for each POI. Note that the TF-IDF technique is applied here to VNAs generated from syntactic KGs and not from tokenization as in standard TF-IDF applications.

Let us denote $F_x = \{(v, \sigma_v^x), v \in \mathcal{N}, \sigma_v^x \in \mathbb{R}\}$ as the set of features that allow to describe the POI x , where σ_v^x is the weight of the node v related to the POI x . F_x is given by the following equation:

$$F_x^i = \Phi(G_i, x, p) \quad (5)$$

where Φ is a function that queries the knowledge graph G_i and collects VNA nodes that have syntactic relations with the node x within a path distance less than p from x . The weight σ of each collected node v is assigned according to three parameters: the weight of the edge w_{vx} , the type of relationship (dependency or proximity) and the frequency of appearance of the node v in all collected paths.

To generate the descriptive VNA embedding vectors, we need to determine the dimensions. We assume that the POIs in the same task of FSL should have the same dimensions and they are generated from the POIs in the “support set”. The dimensions are defined as the common features of a set of POIs. The dimensions of a task \mathcal{T}_i marked $F_{\mathcal{T}_i}$ are defined as follows:

$$F_{\mathcal{T}_i} = \bigcup_{x \in D_i^{supp}} \Phi(G_{\mathcal{T}_i}, x, p)$$

Assuming that d is the desired dimension size, the duplicate elements in $F_{\mathcal{T}_i}$ are removed by aggregating their weight values σ to the unique element, and then the elements are ordered according to σ , the dimension set is defined as follows $F_{\mathcal{T}_i}^* = \{\psi_i, \psi_i \in F_{\mathcal{T}_i}, i \in [1..d]\}$ where $\forall i < j, \sigma_{\psi_i} > \sigma_{\psi_j}$. The descriptive VNA vector of a POI x according to the dimension $F_{\mathcal{T}_i}^*$ is defined as follows: $x = (u_1, u_2, \dots, u_d)$ where

$$u_i = \begin{cases} \sigma_{\psi_i}^x & \psi_i \in F_{\mathcal{T}_i}^* \\ 0 & \text{otherwise.} \end{cases}$$

The values of u_i are normalised where $\sum_{i=1}^d u_i = U$.

3) *Few Shot Learning-based model*: Fig. 3 illustrates our proposal of FSL for multi-label estimation. POIs are assigned to tasks according to their categories. We have adopted the use of *Prototypical Networks* [8] as a type of meta-learning that belongs to the family of approaches using the prior of similarity. This meta-learning consists of applying a clustering

³<https://universaldependencies.org/u/dep/>

⁴<https://megagonlabs.github.io/ginza/>

⁵<https://spacy.io/>

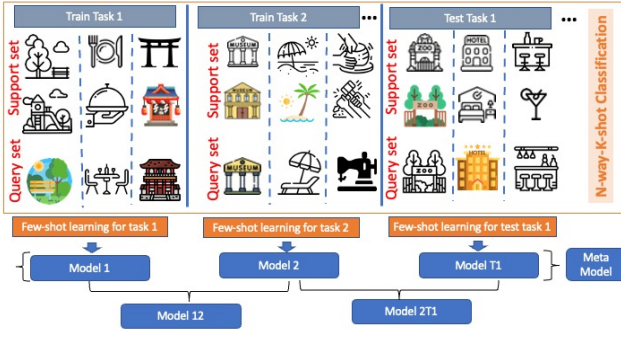


Fig. 3. Few Shot Learning for multi-label estimation

system on the data and estimating based on the similarity of the input POI and the representative of each cluster. We propose to use LightGCN learning, before applying clustering to bring POIs belonging to the same category closer together and to move POIs belonging to different categories away from each other. Each POI in a given task \mathcal{T}_i is embedded as follows: Let $R_{\mathcal{T}_i} \in \mathbb{R}^{K \times L}$ be the POI-label matrix in the task \mathcal{T}_i , where K and L denote the number of POIs and labels respectively, and each entry r_{uv} in $R_{\mathcal{T}_i}$ is 1 if the POI x_u has a label l_v otherwise 0. The adjacency matrix of the POI-label graph is $\begin{pmatrix} 0 & R_{\mathcal{T}_i} \\ R_{\mathcal{T}_i}^T & 0 \end{pmatrix}$.

Let the first layer embedding matrix be $X_{\mathcal{T}_i}^{(0)} \in \mathbb{R}^{(K+L) \times d}$, where d is the embedding size.

$X_{\mathcal{T}_i}^{(n+1)} = (D_{\mathcal{T}_i}^{-1/2} * A_{\mathcal{T}_i} * D_{\mathcal{T}_i}^{-1/2}) * X_{\mathcal{T}_i}^{(n)}$ where $D_{\mathcal{T}_i}$ is a degree matrix in the task \mathcal{T}_i , $A_{\mathcal{T}_i}$ is the adjacent matrix and $X_{\mathcal{T}_i}^{(n)}$ is the combination of POI in D_i^{supp} and labels in $C_{\mathcal{T}_i}$, both encoded with $F_{\mathcal{T}_i}^*$. In order to encode an label l with the VNA, we will retrieve the most representative VNA of the POIs with l as a label.

A row in the K first elements of the matrix $X_{\mathcal{T}_i}^{(n)}$ represents the LightGCN descriptor vector of a POI in D_i^{supp} . Clustering is applied to the LightGCN embedding vectors and DBSCAN [9] is used to classify the POIs in training. The cosine measure is used for clustering. Two similarity measures are proposed to estimate the labels of a given POI: the proposed similarity measure related to the original embedding noted sim_{prop} and the cosine similarity related to the LightGCN embedding. The similarity between two POIs x_1 and x_2 is defined as follows:

$$sim(x_1, x_2) = \alpha * sim_{prop}(x_1^O, x_2^O) + \beta * cosine(x_1^{LG}, x_2^{LG}) \quad (6)$$

where $\alpha + \beta = 1$, x^O the original VNA embedding vector and x^{LG} the LightGCN embedding vector (α is 1 and β is 0 if LightGCN is ignored). Let x be a POI, the estimation of the labels of the POI x is defined as follows:

- embed POI according to the VNAs that have syntactical links with it in the KG, then embed this generated embedding vector by using LightGCN.

- obtain the closest clusters by comparing the similarity value between x and the representatives of the clusters using cosine

and LightGCN embedding vectors. The cluster representative vector is the average of the values of the POIs that make it up. This generates a learning parameter that is the size of the nearest clusters to be considered that we have noted z .

- recover all POIs in the closest clusters and computes the similarity between x and each POI in the closest clusters using the similarity measure defined in the equation 6. This will generate another parameter value of α or β .

- retrieve all labels appearing in the nearest clusters and assign a value ω to each label based on the similarity of the POIs in the nearest clusters and the POI x . Suppose x_1 and x_2 in the nearest clusters are both labelled by label l^i , $\omega_x^i = sim(x_1, x) + sim(x_2, x)$.

- calculate the probability that a label l^i can be a label of the POI x , which is defined as follows:

Suppose $d_x^i = 1 - \omega_x^i$.

$$p(y = l^i | x) = \frac{exp(-d_x^i)}{\sum_{j=1}^L exp(-d_x^j)} \quad (7)$$

The loss function is defined as follows:

$$\mathcal{L} = \frac{1}{|Q|} \sum_{x \in Q} \sum_{i=1}^L -l^i \log p(y = l^i | x) \quad (8)$$

where Q is the query set

4) *Multi-label estimation model*: Multi-label estimation consists of selecting the more appropriate model and estimating the labels using the parameter values of this proposed model. Suppose we want to estimate the labels of a POI x , the approach is to select a model based on the features F_x and the common features or dimension $F_{\mathcal{T}_i}^*$ of each task \mathcal{T}_i .

$$g(x) = m, \gamma_m = \max \gamma_i \quad \forall i \in [1..|\mathcal{M}|] \quad (9)$$

γ is defined by two parameters: γ^{inter} and γ^{weight} where $\gamma^{inter} = |F_x \cap F_{\mathcal{T}_i}^*|$, $\gamma^{weight} = \sum \sigma_{\psi_z}^x \quad \forall \psi_z \in F_x \cap F_{\mathcal{T}_i}^*$

III. EXPERIMENT AND ABLATION STUDY

A. Data set

We have applied the FSL technique to the data set related to the city of Fukuoka, Japan. In order to have labelled data for learning, we decided to use the POIs available on Foursquare and the categories it offers. To do this, we used a Foursquare API. We collected POIs by category. Nine categories were chosen (park, craft store, rail station, museum, shopping mall, shrine, zoo, bar, hotel) and they belong to six parent categories (arts and entertainment, landmarks and outdoor, community and government, travel and transport, retail, dining and drink). After applying the 3-way-3-shot classification by random selection, three tasks are generated: \mathcal{T}_1 , \mathcal{T}_2 , \mathcal{T}_3 . The categories of POIs treated in \mathcal{T}_1 are park, museum and shopping mall, those treated in \mathcal{T}_2 are bar, shrine and zoo and those treated in \mathcal{T}_3 are hotel, rail station and craft store. For each task and category, three POIs are used as support and two POIs are used as query. So we use 45 POIs for FSL. \mathcal{T}_1 , \mathcal{T}_2 are used as training tasks and \mathcal{T}_3 is used as test task. After applying meta-learning, we obtain the values of our learning parameters. We

have generated seven models: \mathcal{M}_1 , \mathcal{M}_2 , \mathcal{M}_3 , \mathcal{M}_{12} , \mathcal{M}_{13} , \mathcal{M}_{32} and \mathcal{M}_{meta} .

To evaluate our multi-label estimation approach, we collected POIs from different prefectures with different characteristics. The prefectures considered are Hokkaido, Kyoto, Tokyo, Kagawa, Fukui, Shimane and Okinawa. Hokkaido is Japan’s northernmost prefecture and its climate is between temperate and polar, with milder summers and harsher winters. The city of Kyoto is known for the presence of at least 17 major historical monuments recognized by UNESCO, such as temples, shrines and castles. Tokyo is located in the center of the country. The city of Tokyo is one of the world’s largest cities and the largest of Japan’s 47 prefectures. It is the main tourist destination in Japan. Kagawa Prefecture is best known for its size (it is the smallest prefecture in Japan) and its gastronomy. Fukui and Shimane are both small prefectures. They are located on the coast of the Sea of Japan. Okinawa Prefecture is the southernmost prefecture in Japan. It has a subtropical oceanic climate and a warm climate throughout the year. Okinawan culture is quite different from other parts of Japan. These prefectures were chosen to see if our model was able to estimate the categories of POIs in small and large cities, as well as those of less known, moderately known and very known POIs.

B. Baseline approaches

We compare our model to three approaches that use our data embedding but do not use FSL. In these baseline approaches, instead of dividing data and processes into tasks, all data is used as a whole for training. Table IV illustrates the results of these baseline approaches (orange part: line 1, 2 and 3).

C. Model setup

We generated a syntactic knowledge graph for each city, so we have eight knowledge graphs, one for training (Fukuoka) and seven for testing (other cities). We used Neo4j and Python to generate and store the knowledge graphs. The choice of POIs is based on their popularity, most of the POIs are not well-known. The criteria used to determine whether a POI is well-known or less known is to check its existence in Google, Yelp (Yp), Foursquare (FQ), TripAdvisor (TA) and Jalan (JI). For Google, 4 parameters are used: the number of search results (GR), the presence of the POI in Maps (M) and the number of photos (P) and reviews (R). For the other platforms, we use the 3 parameters M, P and R. We can consider a POI as less-known if the number of GR values is less than 300,000 and the value of P and R is less than 20. The POIs for the tests were searched manually. Table II shows the distribution of the test data by category. For each city and category, we tried to select 4 POIs: 1 very well-known, 1 moderately well-known and 2 lesser-known. For small cities such as Kagawa, Fukui and Shimane, most of the POIs are lesser-known. For each POI, we have collected the related tweets for 3 years (2020, 2021 and 2022). As the number of results for each query is limited to 500, we split the queries into one query per

3 months, so we have 12 queries for 3 years. If a POI is well-known, we can collect 500 tweets per query, so 6000 tweets for 3 years and 18000 tweets for 3 POIs in the training database (support sets). We can see from Table I that the numbers of tweets for lesser-known POIs in the support and request sets are low compared to the number of tweets for well-known POIs.

D. Experimental results

Table III shows the experiment results for the POIs in seven cities. The average accuracy values for the five categories (park, museum, shrine, zoo and arts/craft store) are between 0.80 and 0.89, while the average F1-score values are between 0.49 and 0.77. The minimum and maximum accuracy values are respectively 0.67 and 1 while the minimum and maximum score values for F1 are respectively 0 and 1. The zoo category is very well estimated. The zoos of three prefectures out of the seven are perfectly estimated with an F1 of 1 (Tokyo, Kagawa and Fukui). Most of the zoos are park zoos, so they have two categories (park and zoo), in the case of a value of 1, the system has exactly estimated both correct categories. The second well-estimated category is the shrine, with an average accuracy of 0.85 and an F1 of 0.61.

E. Ablation and limitation of this study

Table IV shows the results of the baseline approaches (orange lines: 1, 2 and 3) and those of the proposed approach using FSL. It also allows us to see the impact of using or not using LightGCN and TF-IDF in the approach (blue lines: 5, 6, 7). We can see that none of the lines in the baselines contain maximum accuracy or F1 score values. The lines with the maximum accuracy and F1 values for all categories are lines 4, 6 and 7, where the FSL was taken into account. The one with the maximum average values for all categories is line 4, which is our proposed approach. If we compare the line 4 and line 5 where the consideration of LightGCN differs them, we can see that almost of the accuracy and F1 score values for the line 4 are higher than those in the line 5. However, there are some cases where the evaluation score for line 5 is higher than those of line 4 and the differences are not insignificant. The use of LightGCN can improve the estimation results in almost all cases, but there are some cases where the results of the models without LightGCN became the same or higher than those with LightGCN (case of museum and craft store). If we focus on lines 2 and 3, the estimates for line 3, where the TF-IDF is considered, are higher than those for line 2. This means that the TF-IDF can improve the estimation slightly. In conclusion, approaches that incorporate FSL perform better than baseline approaches and the differences are significant. It is important to determine in which situation the LightGCN should or should not be used to obtain optimal performance. The search for optimal values of α and β in equation 6 should be improved. The use of the TF-IDF technique can slightly improve the performance of the system.

TABLE I
TRAIN DATA DISTRIBUTION

Set/#tweets	T1			T2			T3		
	Park	Museum	Shopping-mall	bar	shrine	Zoo	education	train station	Craft store
Support (3 POIs)	6821	10624	3166	407	2277	3969	1943	546	329
Query (2 POIs)	6116	2393	774	159	2338	115	90	130	280

TABLE II
TEST DATA DISTRIBUTION

City/categories	Park	Museum	Shrine	Zoo	Craft store
Hokkaido	5	4	4	4	3
Kyoto	2	2	5	2	4
Tokyo	4	4	4	3	2
Kagawa	3	3	4	2	1
Fukui	2	2	3	1	3
Shimane	3	3	3	4	2
Okinawa	3	4	3	2	4
Total	22	22	26	18	19
	107				

F. Example of lesser-known POIs

Table V shows some examples of lesser-known POIs. “Ryuusen hamono” is a handicraft factory and shop where you can buy or sharpen the knives you normally use while listening to a lecture by a craftsman. “Kametani yougyou” is a boutique for the sale and manufacture of handmade ceramics using traditional techniques. It also offers apprentice workshops. “Noguchi senpo” is an artisan factory that works on fabrics, especially on operations related to kimono, dyeing, processing, manufacturing and cleaning. They also propose some workshops such as the coloring of fabrics. Their numbers of tweets are very low for a period of 3 years (2020, 2021, 2020). We asked ChatGPT⁶ the categories of the POIs in Table V by asking the following question: “What is the category of the place called?”. Normally, when we ask ChatGPT, it gives not only the answer related to our question, but also some details or explanations. This is the case when we ask this question for well-known POIs. We can see from the first POI (Ryuusen hamono) in Table V that ChatGPT relies only on the name of the POI to determine its type, there is no detail or explanation. As the word “hamono” (刀物) means cutlery, it guesses that the category of this place is cutlery or blades. ChatGPT is not able to guess the categories of the last two POIs. There is no response at all for these POIs. We can conclude that ChatGPT does not have enough information for these POIs to be able to estimate their categories. The accuracy and F1 values of our approach are quite good. It detected 2 categories, while in the ground truth there is only one category, which is retail. Looking closer at the second POI, we saw that the shop also serves food, which is why our approach detects “Dining and drinking” as a category. Since people also talked in tweets about how to get to these places, the first and third POIs are assigned the category “Travel and Transport”. We also query our system to get some details about what we can

⁶<https://chat.openai.com/>

do at these places. The answers are as follows (the numbers in brackets indicate the number of tweets):

- Ryuusen hamono: Ryuusen cutlery teaches you how to make cutlery (88), Ryuusen cutlery sells for half the price (7).
- Kametani yougyou: Making a bell at Kametani Pottery (55).
- Noguchi senpo: An event is being held at Noguchi Senpo (2), a dyeing experience at Noguchi senpo shop (1).

We can affirm that our approach is useful not only for estimating categories of lesser-known POIs, but also for providing information about them, based on questions that tourists may ask. This aspect will be studied in greater depth in our future work. The examples shown in Table V are POIs that may be of interest to tourists and that may be specific to the cities in which they are located. Our approach makes it possible to promote these types of POIs by including them in recommendation systems and having them discovered by tourists. The fact that ChatGPT cannot provide information about them shows that this information is difficult to find. Our approach therefore offers advantages in terms of complexity, where search and page consultation times, as well as information assembly and analysis tasks, are significantly reduced.

IV. RELATED WORK

There are approaches of POI category estimation that are based on textual data such as: names of POIs [10], reviews [11], descriptions provided by institutions or from web sources [12] [13], information from social networks like tweets [14] [15] [16]. Other approaches are interested in exploiting data related to the historical trace of tourist visits, such as syntactic patterns [12], traveller behaviour [14], places of publication, trajectories [17] [18]. Some of them also use spatio-temporal information such as the location of POIs, opening hours or periods when places are much busier [15]. As the data on the labelled POIs to be used for learning and validation are still not complete or available, depending on application field and the objectives of the studies, some studies are working on the manual labelling of information before working on the estimation of information categories [19] [13]. These approaches for estimating POI labels have been proposed for different purposes, such as estimating the category of the next POI to recommend [11], studying traveller behaviour [14], detecting leisure activities from texts [13].

These approaches use different techniques. Those using textual data propose data encoding techniques based on tokenisation, bag-of-words, syntactic or semantic studies of the data, representations in the form of graphs. With regard to learning and estimating categories, the following techniques are proposed SVM [11], TF-IDF [14], Long Short-

TABLE III
EVALUATION OF THE MULTI-LABEL ESTIMATION APPROACH

City/F1	Park		Museum		Shrine		Zoo		Craft store	
	A	F1	A	F1	A	F1	A	F1	A	F1
Hokkaido	0.77	0.33	0.79	0.41	0.83	0.5	0.75	0.34	0.83	0.67
Kyoto	0.75	0.33	0.83	0.75	0.7	0.4	0.92	0.83	0.83	0.5
Tokyo	0.92	0.75	0.71	0.17	0.96	0.96	1	1	1	1
Kagawa	0.83	0.67	0.89	0.67	0.83	0.5	1	1	0.67	0
Fukui	0.83	0.5	1	1	0.89	0.667	1	1	0.72	0.22
Shimane	0.78	0.5	0.67	0.2	0.89	0.67	0.83	0.59	0.66	0.33
Okinawa	0.78	0.5	0.75	0.25	0.83	0.56	0.75	0.65	0.92	0.75
Average	0.81	0.51	0.81	0.49	0.85	0.61	0.89	0.77	0.80	0.5

TABLE IV
ABLATION OF THE STUDY

	Approaches/F1			Park		Museum		Shrine		Zoo		Craft store		Average	
	FSL	LGCN	TF-IDF	A	F1	A	F1	A	F1	A	F1	A	F1	A	F1
1	N	Y	Y	0.41	0.10	0.53	0.36	0.42	0.27	0.41	0.46	0.64	0.37	0.48	0.31
2	N	N	Y	0.73	0.43	0.79	0.52	0.80	0.58	0.70	0.49	0.82	0.64	0.77	0.53
3	N	N	N	0.63	0.13	0.73	0.42	0.80	0.57	0.63	0.42	0.80	0.58	0.72	0.42
4	Y	Y	Y	0.81	0.51	0.81	0.49	0.85	0.61	0.89	0.77	0.80	0.50	0.83	0.58
5	Y	N	Y	0.67	0.36	0.81	0.53	0.81	0.61	0.78	0.59	0.83	0.64	0.78	0.55
6	Y	N	N	0.66	0.24	0.82	0.61	0.85	0.70	0.74	0.55	0.85	0.64	0.78	0.55
7	Y	Y	N	0.74	0.29	0.77	0.44	0.88	0.71	0.77	0.47	0.79	0.43	0.79	0.47

■ Without FSL ■ Proposed Approach ■ With/Without LightGCN / TF-IDF

TABLE V
EXAMPLES OF LESSER-KNOWN POIS
FOR YP, GQ, TQ, JL, GR, M, P, R, PLEASE REFER TO SECTION III-C.
C (CATEGORIES), D (DETAILS), E (EXPLANATION), Y (YES), N (NO), #TW (NUMBER OF TWEETS), A (ACCURACY), F1 (F1 SCORE).

POI/Sources	Google		Yp	FQ	TA	Jl	ChatGPT	Our approach			
	GR	M,P,R	M,P,R	M,P,R	M,P,R	M,P,R	C,D,E	# TW	Estimated categories	A	F1
Ryūsen hamono (龍泉刀物), Fukui	219K	(Y,14,12)	(Y,0,0)	(Y,0,0)	N	(Y,9,0)	(Cutlery or Blades, N, N)	349	Retail, Travel and transportation	0.83	0.67
Kametani yōgyō (亀谷窯業), Shimane	8K	(Y,16,10)	(Y,0,0)	N	N	(Y,7,1)	N	41	Retail Dining and drinking	0.83	0.67
Noguchi senpo (野口染舗), Hokkaido	16K	(Y,10,7)	N	N	N	(Y,15,2)	N	63	Retail Travel and transportation	0.83	0.67

Term Memory (LSTM) [15], Bi-LSTM [16], Bidirectional Encoder Representations from Transformers (BERT) [20], Linear Discriminant Analysis (LDA) [13], Graph Convolution Networks with Attention (GCN-A) [21], Non-negative Matrix Factorization (NMF) [10]. The descriptive texts of the POIs are interesting to exploit, but they are not often available and their manual production is very costly in terms of time and tasks, it is the same for the construction of ontologies. The other problem is that the complexity of their processing increases with the size of the texts to be processed and the number of POIs available, which is the case for bag-of-words, ngram, SVM and BERT. In our study, we propose to collect information from social networks, especially Twitter, in order to extract characteristic information of POIs. Instead of using all the words of the basic texts, we focus our work on the syntactic study of the information, retrieving verbs, nouns, adjectives and the syntactic links between them. These

three elements make it possible to describe POIs and their normalisation would make it possible to reduce the size of the data to be processed.

Recently, studies have focused on using small data sets, which reduces the cost and complexity of processing and does not require a large amount of data for learning. We are interested in this type of approach, more specifically in FSL. This kind of approach is widely used in image [22] [23] and video [24] classification, but so far less so in text processing. Approaches consisting in estimating the labels of the images using FSL and ZSL are respectively proposed in [22] [23]. The use of FSL with textual data is often aimed at matching text with labels. [20] propose an approach to match tweets to information categories such as news, advertising or business information. This approach uses BERT and information from web sources. Another approach uses medical information and consists of tagging documents with predefined labels [25].

Each label is accompanied by a descriptive text or document. The labels and input documents are encoded using keywords extracted from the texts. This approach uses ngram, GCNN and FSL for training and estimation.

ChatGPT can also be used to request the category of a given POI. ChatGPT which is a chatbot incorporating artificial intelligence, is developed by OpenAI. The data used for learning comes from different sources. In our context, it tries to determine the category(ies) of the POI as well as its location from the collected information and learning.

V. CONCLUSION AND FUTURE WORK

We propose a multi-label estimation approach for lesser-known POIs using the techniques of FSL, LightGCN, and TF-IDF. Our approach is mainly based on encoding POIs from verbs, nouns, and adjectives extracted via syntactic knowledge graphs. These knowledge graphs are generated from tweets. This approach not only outperforms baseline approaches but also significantly reduces processing time by working with a small number of learning POIs. Since standard embedding approaches that directly embed documents require a lot of information, they do not necessarily work well with FSL. Our proposed VNA-based approach, which extracts a small number of salient information from tweets for better embedding, is suitable for FSL and therefore provides better performance for estimating categories of lesser-known POIs. The limitation of this approach is at the data source level. If there are no tweets related to a given POI, our system is not able to estimate its categories. In future work, we plan to extend the work by considering other data sources as well as integrating this approach into a recommendation system dedicated to tourists or tourist offices. We also plan to consider other languages to evaluate the genericity of our approach.

ACKNOWLEDGMENT

This work was supported in part by KAKENHI No. JP21F21377, JP22KF0288, and JP21H00907.

REFERENCES

- [1] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, "Generalizing from a few examples: A survey on few-shot learning," *ACM computing surveys (csur)*, vol. 53, no. 3, pp. 1–34, 2020.
- [2] L. Fei-Fei, R. Fergus, and P. Perona, "One-shot learning of object categories," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 4, pp. 594–611, 2006.
- [3] B. Romera-Paredes and P. Torr, "An embarrassingly simple approach to zero-shot learning," in *International conference on machine learning*. PMLR, 2015, pp. 2152–2161.
- [4] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang, "Lightgcn: Simplifying and powering graph convolution network for recommendation," in *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*, 2020, pp. 639–648.
- [5] A. Parnami and M. Lee, "Learning from few examples: A summary of approaches to few-shot learning," *arXiv preprint arXiv:2203.04291*, 2022.
- [6] T. Tanaka, Y. Miyao, M. Asahara, S. Uematsu, H. Kanayama, S. Mori, and Y. Matsumoto, "Universal dependencies for japanese," in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (Irec'16)*, 2016, pp. 1651–1658.

- [7] J. Ramos *et al.*, "Using tf-idf to determine word relevance in document queries," in *Proceedings of the first instructional conference on machine learning*, vol. 242, no. 1. Citeseer, 2003, pp. 29–48.
- [8] J. Snell, K. Swersky, and R. Zemel, "Prototypical networks for few-shot learning," *Advances in neural information processing systems*, vol. 30, 2017.
- [9] K. Khan, S. U. Rehman, K. Aziz, S. Fong, and S. Sarasvady, "Dbscan: Past, present and future," in *The fifth international conference on the applications of digital information and web technologies (ICADIWT 2014)*. IEEE, 2014, pp. 232–238.
- [10] J. Zhou, S. Gou, R. Hu, D. Zhang, J. Xu, A. Jiang, Y. Li, and H. Xiong, "A collaborative learning framework to tag refinement for points of interest," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 1752–1761.
- [11] M. Li, W. Zheng, Y. Xiao, K. Zhu, and W. Huang, "Exploring temporal and spatial features for next poi recommendation in lbsns," *IEEE Access*, vol. 9, pp. 35 997–36 007, 2021.
- [12] D. Zhang, D. Wang, H. Zheng, X. Mu, Q. Li, and Y. Zhang, "Large-scale point-of-interest category prediction using natural language processing models," in *2017 IEEE International Conference on Big Data (Big Data)*. IEEE, 2017, pp. 1027–1032.
- [13] D. van Weerdenburg, S. Scheider, B. Adams, B. Spierings, and E. van der Zee, "Where to go and what to do: Extracting leisure activity potentials from web data on urban space," *Computers, Environment and Urban Systems*, vol. 73, pp. 143–156, 2019.
- [14] K. Takahashi, D. Kato, M. Endo, T. Araki, M. Hirota, and H. Ishikawa, "Analyzing travel behavior using multi-label classification from twitter," in *Proceedings of the 9th International Conference on Management of Digital EcoSystems*, ser. MEDES '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 50–56. [Online]. Available: <https://doi.org/10.1145/3167020.3167028>
- [15] N. Lagos, S. Ait-Mokhtar, and I. Calapodescu, "Point-of-interest semantic tag completion in a global crowdsourced search-and-discovery database," in *ECAI 2020*. IOS Press, 2020, pp. 2993–3000.
- [16] H. Guo, X. Li, L. Zhang, J. Liu, and W. Chen, "Label-aware text representation for multi-label text classification," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 7728–7732.
- [17] Rimjhim and S. K. Dandapat, "Tagging multi-label categories to points of interest from check-in data," *IEEE Transactions on Emerging Topics in Computational Intelligence*, pp. 1–14, 2023.
- [18] Y. Wang, M. Chen, X. Yu, and Y. Liu, "Lce: A location category embedding model for predicting the category labels of pois," in *Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14–18, 2017, Proceedings, Part V 24*. Springer, 2017, pp. 710–720.
- [19] H. Hu, Y. Zheng, Z. Bao, G. Li, J. Feng, and R. Cheng, "Crowdsourced poi labelling: Location-aware result inference and task assignment," in *2016 IEEE 32nd International Conference on Data Engineering (ICDE)*, 2016, pp. 61–72.
- [20] D. Das, R. Chenchu, M. Abdollahi, J. Huh, and J. Srivastava, "Adbert: An effective few shot learning framework for aligning tweets to superbowl advertisements," in *Proceedings of the Eighth Workshop on Noisy User-generated Text (W-NUT 2022)*, 2022, pp. 171–179.
- [21] K. Xie, Z. Wei, L. Huang, Q. Qin, and W. Zhang, "Graph convolutional networks with attention for multi-label weather recognition," *Neural Computing and Applications*, vol. 33, pp. 11 107–11 123, 2021.
- [22] A. Alfassy, L. Karlinsky, A. Aides, J. Shtok, S. Harary, R. Feris, R. Giryes, and A. M. Bronstein, "Laso: Label-set operations networks for multi-label few-shot learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 6548–6557.
- [23] G. Ou, G. Yu, C. Domeniconi, X. Lu, and X. Zhang, "Multi-label zero-shot learning with graph convolutional networks," *Neural Networks*, vol. 132, pp. 333–341, 2020.
- [24] K. Cao, J. Ji, Z. Cao, C.-Y. Chang, and J. C. Niebles, "Few-shot video classification via temporal alignment," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 10 618–10 627.
- [25] A. Rios and R. Kavuluru, "Few-shot and zero-shot multi-label learning for structured label spaces," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, vol. 2018. NIH Public Access, 2018, p. 3132.