



# Human-in-the-loop latent space learning for biblio-record-based literature management

Shingo Watanabe<sup>1</sup> · Hiroyoshi Ito<sup>1</sup> · Masaki Matsubara<sup>1</sup> · Atsuyuki Morishima<sup>1</sup>

Received: 25 July 2023 / Revised: 3 November 2023 / Accepted: 4 November 2023 / Published online: 4 January 2024  
© The Author(s) 2024

## Abstract

Every researcher must conduct a literature review, and the document management needs of researchers working on various research topics vary. However, there are two major challenges. First, traditional methods such as the tree hierarchy of document folders and tag-based management are no longer effective with the enormous volume of publications. Second, although their bibliographic information is available to everyone, many papers can only be accessed through paid services. This study attempts to develop an interactive tool for personal literature management based solely on their bibliographic records. To make such a tool possible, we developed a principled “human-in-the-loop latent space learning” method that estimates the management criteria of each researcher based on his or her feedback to calculate the positions of documents in a two-dimensional space on the screen. As a set of bibliographic records forms a graph, our model is naturally designed as a graph-based encoder–decoder model that connects the graph and the space. In addition, we also devised an active learning framework using uncertainty sampling for it. The challenge here is to define the uncertainty in a problem setting. Experiments with ten researchers from the humanities, science, and engineering domains show that the proposed framework provides superior results to a typical graph convolutional encoder–decoder model. In addition, we found that our active learning framework was effective in selecting good samples.

**Keywords** Graph neural network · Human-in-the-loop system · Document management

## 1 Introduction

Every researcher must conduct a literature review, and there is a personalized need for researchers working on various research topics in their document management. They must organize publications according to their criteria to find relevant research and understand their field trends.

However, there are two significant challenges faced in personalized literature management. First, researchers must

manage much research. Fire [6] found that recently, more than seven million new scholarly studies are published annually.

Therefore, the traditional approaches such as the tree hierarchy of document folders and tag-based management, are no longer effective. There is a need for automated literature management techniques.

Second, accessing the content of a paper is challenging. Although their bibliographic information is available to everyone, many papers can only be accessed through paid services. According to Nicolson et al. [15], 65% out of the 100 most cited papers were paywalled. This is a major barrier to researchers accessing relevant.

Therefore, methods for automatic literature management that use the literature contents [24, 31, 40] have limited applicability.

With this background, this study attempts to develop an interactive tool for personal literature management based on bibliographic records *without the need to access the con-*

---

Hiroyoshi Ito, Masaki Matsubara and Atsuyuki Morishima have contributed equally to this work.

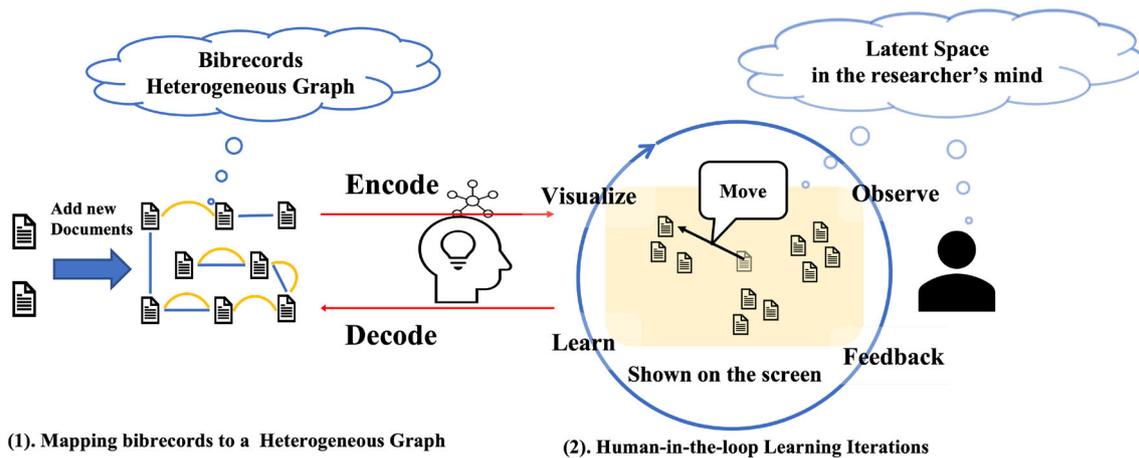
✉ Shingo Watanabe  
watanabe.shingo.ss@alumni.tsukuba.ac.jp

Hiroyoshi Ito  
ito@slis.tsukuba.ac.jp

Masaki Matsubara  
masaki@slis.tsukuba.ac.jp

Atsuyuki Morishima  
mori@slis.tsukuba.ac.jp

<sup>1</sup> University of Tsukuba, Tsukuba, Japan



**Fig. 1** Overview of the framework. (1) A graph of biblio-records is constructed for the arrived documents; documents are connected if they mention the same terms, are written by the same author, etc. (2) Next,

we have a human-in-the-loop learning iteration for learning the latent space in the user's mind. Then, the learner *predicts* the positions of newly arrived papers in the space

*tents of papers*<sup>1</sup> The tool asks the researcher to place icons corresponding to papers in a two-dimensional space on the screen using their own criteria, and then *predicts the positions of newly arrived papers that the user would place*. Figure 1 illustrates this process. First, since the relationships among bibliographic records are naturally modeled as a graph, the set of biblio-records is represented as a heterogeneous graph of biblio-records whose nodes correspond to papers, authors, conference names, years, etc. (Fig. 1(1)). The graph connects papers that share the same authors, the same years, and so on. Subsequently, the machine learner that implements our human-in-the-loop latent space learning method (Sect. 4), computes and visualizes the positions in a *two-dimensional space on the screen* that corresponds to the space for papers that exists in the researcher's mind (see Fig. 1(2)). Next, the researcher provides feedback on the suggested positions by *moving papers from incorrect positions according to their criteria to the correct position*. In the feedback phase, the researchers are provided details about the literature, including the title, authors, publication location, and year. Then, the learner receives the feedback and updates the criteria in the space so that it can correctly predict the positions of newly arrived papers.

The *interactive* nature not only captures the current latent space of papers in each researcher's mind, but also allows the system to follow the researcher's criteria that are evolving over time [1].

Thus, our problem can be considered as *latent space learning* with a graph convolutional encoder–decoder model [18]. Here, the encoder and decoder map the paper nodes in the graph to points in the latent space and vice versa, and

the objective function is the cross-entropy loss for generating adjacency matrices for document clusters in the space. However, existing models do not support our *human-in-the-loop* approach; that is, they do not allow the user to provide interactive feedback to the latent space. Therefore, we developed a principled “human-in-the-loop latent space learning” method that estimates the management criteria of each researcher based on their feedback on the estimated positions of documents in a two-dimensional space on the screen. Our challenge is how to make the model capture the characteristics of the latent space for literature management.

## 1.1 Challenges and contributions

(1) We present a *principled* framework for interactive latent space learning in literature management. It is based on a common graph convolutional encoder–decoder model, in which the criteria for individual literature management are represented by the weights of a set of *meta-paths* (i.e., sequences of attributes at the schema of bib-records data), which are a popular means of capturing the semantics of heterogeneous graph [24, 32]. Our model is unique in that it is based on the following two assumptions. First, the user's criteria in the latent space are consistent *only locally*. This was inspired by the results in psychology such as [35]. Thus, our first research question (RQ1) is whether each researcher has different criteria for different sub-spaces in the latent space or not.

Second, the two papers are connected through paths on the graph if they are close to each other in the latent space. Therefore, unlike other popular graph convolutional encoder–decoder models, our decoder is based on the Euclidean distance between the latent vectors. Thus, our second question (RQ2) is whether our decoder is effective or not.

<sup>1</sup> This paper is an extended version of our ICADL paper [39].

(2) We show the experimental results of ten academic researchers from the *science, engineering, and the humanities domains*. The results answer the two research questions positively and show that the approach is much superior to a typical graph convolutional model. The resulting quality is practically good in that it can place the new paper in a position close to the correct one although it does not necessarily exact one. This implies that our tool can help researchers manage relevant publications based on their own criteria.

(3) Based on the above experimental results, a natural question is whether we can devise an active learning method to improve the learning efficiency. We devised an active learning approach using uncertainty sampling. The challenge here is to define “uncertainty” in our problem setting—what the system learns is the importance of each meta-path for each cluster of each user. Thus, our third research question (RQ3) is whether we can develop an effective active learning for the setting or not. We formally define our uncertainty in the setting and the framework is then evaluated experimentally. The result shows that the uncertainty sampling strategy allows the system to boost the performance compared to random sampling, with a statistical significance.

## 1.2 Limitations

This study does not intend to find the best feature set or the best performance in learning the latent space using bibliographic data that are potentially available to the public.

## 2 Related work

### 2.1 Literature management tools

Tools to assist researchers in organizing related papers are widely used, and studies have been conducted on such tools. Francese [7] conducted a survey at the University of Turin to determine the manner in which students and researchers manage their bibliographies. The results of the survey showed that EndNote was the most popular bibliography management software for researchers to manage their electronic literature online, used by 49% of the respondents, followed by BibTex (11%) and Mendeley (9%). In general, such tools can automatically classify the documents with *objective criteria* such as years and authors and require explicit inputs from users (such as tags given to each paper) to manage them using the users’ criteria for document management. By contrast, our system automatically estimates the *user’s* document management criteria and can map new documents onto the space so that the user can easily grasp how they are related to other papers.

### 2.2 Document classification, clustering, recommendation

Document classification, clustering and recommendations are of increasing interest because of the increasing number of academic papers that researchers must manage. Various methods have been proposed, such as hierarchical Bayesian clustering [14] and metric learning [25, 37]; however, almost all these approaches use natural language processing methods [12, 30]. Unlike our method, most existing methods classify, cluster, and recommend documents by analyzing the abstracts and content of papers assuming that the *document contents* can be accessed, which limits their applicability in the current digital library situation.

Studies have been conducted on personalized paper recommendation methods that do not require document contents [16, 21, 24, 36, 38]. Paper recommendation is orthogonal to the latent space learning problem in that the former does not identify any criteria for how researchers manage the papers, and our method does not address the problem of identifying papers to recommend. Combining these two approaches is an interesting topic for future research.

### 2.3 Active learning

Active learning is utilized in various machine learning techniques, including the latent space learning [4, 26]. Typically, sampling strategies for active learning are designed to increase classifications and regressions in terms of their evaluation measures. In contrast, our feedback system on the placement of documents in the latent space serves as an oracle for latent space learning while allowing the criteria for organizing documents in each cluster to *evolve* with the interactive interface. In addition, some studies have used active learning for graph convolutional encoder–decoder models [2, 3]. In these studies, the information entropy and the graph structure are used to select the most informative nodes for the next iteration. Whereas the methods in these studies asks users for nodes with uncertain labels, our method determines the data to be asked based on the uncertainty regarding the importance of the meta-path.

### 2.4 Latent space learning

Latent space learning has been used to learn data features and comprehending data patterns and/or structural similarities in various contexts. For example, PTE [33] is a semi-supervised latent space learning technique used for textual data. In addition, doc2vec [20] creates representations for each document using latent space learning.

Network-embedding techniques that consider latent semantics have attracted considerable attention for graphs [8, 11, 13, 28, 29]. Some of the techniques such as Deepwalk

**Table 1** Notations used in this paper

Symbol	Description
$\mathcal{D}$	A set of documents $\{d_1, \dots, d_{ \mathcal{D} }\}$
$\mathcal{P}$	A set of meta-path $\{p_1, \dots, p_{ \mathcal{P} }\}$
$\mathcal{V}$	A set of nodes
$\mathcal{E}$	A set of edges
$G(\mathcal{D}, \mathcal{E})$	Undirected graph
$\mathcal{A}$	A set of object types of nodes
$\mathcal{R}$	A set of relation types of edges
$L$	Embedding dimension
$\mathcal{W}_p$	The weight of a meta-path $p$
$\mathbf{A}^p \in \mathbb{R}^{\mathcal{D} \times \mathcal{D}}$	Adjacency matrix of a meta-path $p$
$C_k \subseteq \mathcal{D}$	$k$ -th cluster
$\tilde{\mu}_k \in \mathbb{R}^L$	A centroid of $k$ -th cluster
$\tilde{\mathbf{A}} \in \mathbb{R}^{\mathcal{D} \times \mathcal{D}}$	Adjacency matrix weighted with $W$
$\tilde{\mathbf{A}}_{d_i, d_j}$	The $(i, j)$ element of $\tilde{\mathbf{A}}$
$\mathbf{X} \in \mathbb{R}^{\mathcal{D} \times \mathcal{D}}$	Features
$\tilde{z}_{d_i} \in \mathbb{R}^L$	Latent vector of $d_i$
$\mathbf{Z} \in \mathbb{R}^{\mathcal{D} \times L}$	Matrix of $\tilde{z}$
$\hat{\mathcal{Z}}$	Feedback $\{(\tilde{z}_{d_i}, \hat{z}_{d_i}), \dots, (\tilde{z}_{d_j}, \hat{z}_{d_j})\}$
$\theta$	Decoder network parameters
$\phi$	Encoder network parameters
$\mathcal{Q}$	Unknown documents $\{q_1, \dots, q_{ \mathcal{Q} }\}$

[27] and Node2vec [9] rely on random walks to produce a distributed representation of nodes; LINE [34] consider and embed nodes that indirectly have edges attached to one another; the Kipf and Welling GCN [18] method learns the latent vectors of the nodes while considering the network structure. In addition, to fit autoencoders [17] to network data, the GCN was used in graph autoencoders (GAE) and variational graph autoencoders (VGAE) [19]. Both methods involve a two-layer graph convolutional network and reconstruction of the adjacency matrix using an encoder–decoder algorithm. Our model is unique in that it addresses the local consistency of criteria in the latent space and adopts the distance-based decoder tailored for literature management.

### 3 Definitions and the problem

We discuss our problem using the notations listed in Table 1. First, we define the important concepts we used in the discussion, and then define our problem.

#### 3.1 Heterogeneous information network

Real-world systems, such as bibliographic information networks, are structured into HINs [5, 32]. A heterogeneous

information network (HIN) is a special type of network structure that has multiple types of nodes and edges.

**Definition 1 (Heterogeneous Information Network)** An HIN is defined as a directed graph  $G(\mathcal{V}, \mathcal{E})$  with an object-type mapping function  $\tau : \mathcal{V} \rightarrow \mathcal{A}$  and relation-type mapping function  $\phi : \mathcal{E} \rightarrow \mathcal{R}$ , where  $\mathcal{V}$  and  $\mathcal{E}$  represent set of the nodes and edges, and  $\mathcal{A}$  and  $\mathcal{R}$  are the set of the object types and the relation types, respectively. In general,  $|\mathcal{A}| + |\mathcal{R}| > 2$ . For example, in a bibliographic information network, there are object types, such as paper (P), author (A), term (T), year (Y), and relation types, for example, published a paper (A-P) or a paper is published in a venue (P-V). By constructing a schema of paths called a *meta-path* from these types of objects and relations, we can explain the rich semantics of HIN.

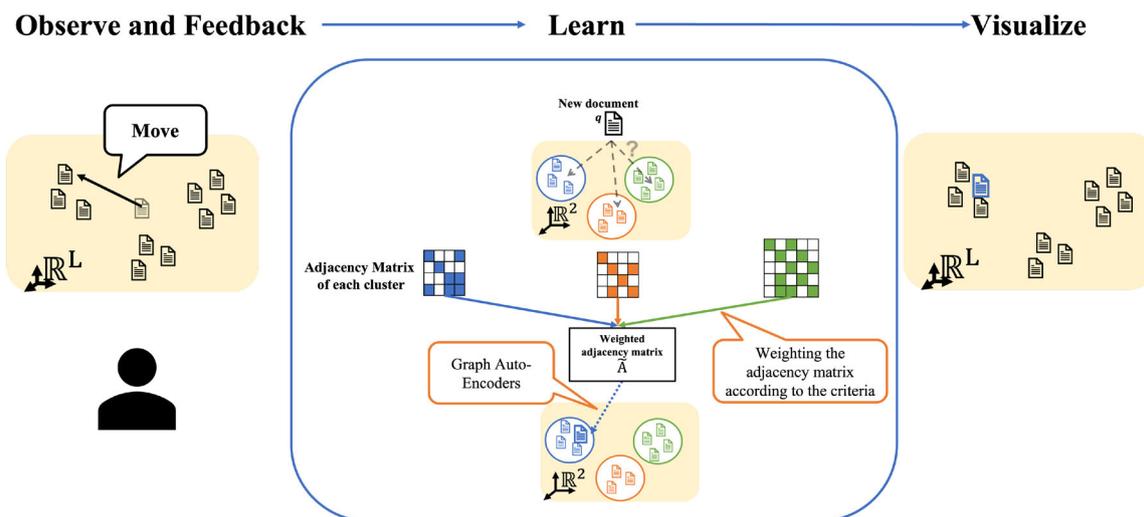
#### 3.2 Meta-path

Intuitively, a *meta-path* is a sequence of object-types that can have an instance in the graph. For example, A-P and PAP are meta-paths. Meta-paths are commonly used to capture rich semantics of [24, 32].

**Definition 2 (Meta-Path)** The meta-path  $P$  is defined as  $\mathcal{A}_1 \xrightarrow{R_1} \mathcal{A}_2 \xrightarrow{R_2} \dots \xrightarrow{R_l} \mathcal{A}_{l+1}$  and defines a composite relation  $\mathcal{R} = \mathcal{R}_1 \circ \mathcal{R}_2 \circ \dots \circ \mathcal{R}_l$  between types  $\mathcal{A}_1$  and  $\mathcal{A}_{l+1}$  where  $\circ$  denotes the composition operator of the relations. As this study focuses on the relationships between papers, we consider a meta-path in which both the starting and ending points of the meta-path are papers (P). For example, the meta-path “Paper (P)– Author (A)– Paper (P)” indicates the relationship between papers written by the same author.

#### 3.3 Problem

We assume that a set of documents represents an HIN and that each document has features. In this study, we assumed that an attribute is the index of a document, which is represented as a matrix  $\mathbf{X} \in \mathbb{R}^{|\mathcal{D}| \times |\mathcal{D}|}$ . We construct adjacency matrices  $\{\mathbf{A}^p \in \mathbb{R}^{|\mathcal{D}| \times |\mathcal{D}|}\}_{p \in \mathcal{P}}$ , each of which represents the relationships between documents in a meta-path  $p$ . Additionally, the user interaction processes are provided to estimate the user’s document management criteria. This interaction is denoted as a set of tuples  $(\tilde{z}, \hat{z}) \in \hat{\mathcal{Z}}$ , where  $\tilde{z}_{d_i}$  represents the initial point of the  $d_i$ ’s latent vector and  $\hat{z}_{d_i}$  represents the point of the vector after the interaction. We formally define our research as follows. Given a set of adjacency matrices  $\{\mathbf{A}^p\}_{p \in \mathcal{P}}$ , the feature of documents  $X$ , and a set of interactions  $\hat{\mathcal{Z}}$ , we find  $Z_{\mathcal{Q}}$ , which is a set of latent vectors of a set of unknown documents  $\mathcal{Q}$ .



**Fig. 2** One iteration of Observe-Feedback-Learn-Visualize. Given the visualized latent space of papers, the user gives the feedback by moving papers from incorrect positions to the correct ones, then the learning process updates the space based on the feedback

### 4 Proposed learning method

This section explains our proposed learning method, called ISLE (Interactive latent Space Learning). Algorithm 1 illustrates the structure of ISLE. The components of the algorithm are explained below.

To enable the model to capture the problem of identifying the positions of documents in the latent space in the user’s mind, our method was designed based on the two assumptions: First, there is some locality of the criteria for managing documents in the space in mind; when the researcher moves papers to a place near some of the other papers, there is a consistent criteria in the neighborhood, but the consistency is not guaranteed in other places. Second, two papers are connected through many paths some way on the graph if they are close to each other in the latent space. Therefore, unlike other popular graph convolutional network-based encoder–decoder models, our model’s decoder is based on the Euclidean distance of the latent vectors.

The learning phase of our proposed framework comprises three steps:

1. Clustering the latent vectors;
2. Estimating document management criteria in each cluster;
3. Learning the latent vectors of documents based on graph autoencoders and obtaining the latent vector for the new document.

These steps were included in the iterations of our human-in-the-loop framework. Each time a user provides feedback, this step to updates the clusters and fine-tunes the models. Figure 2 illustrates the learning phase in one “move-learn-display” iteration in our framework in Fig. 1.

#### 4.1 Clustering the latent vectors

The first step in our proposed method is to cluster the latent space in which the user provides feedback. (This corresponds to Line 22–23 in Algorithm 1.) The k-means clustering method is used in [10]. Clustering by k-means results in an adjacency matrix and center of mass for each cluster. The k-means optimization problem is expressed as follows:

$$\{\mathbf{r}_{d_i}\}_{d_i \in \mathcal{D}}, \{\bar{\mu}_k\}_{k \in [n_c]} = \arg \min_{\{\mathbf{r}_{d_i}\}, \{\bar{\mu}_k\}} J, \tag{1}$$

where  $\mathbf{r}_{d_i} = (r_{d_i,1}, \dots, r_{d_i,k})^\top$  represents the cluster assignment vector of the document  $d_i$ . Each element  $r_{d_i,j}$  is one if document  $d_i$  belongs to cluster  $j$  and zero otherwise.  $\bar{\mu}_k \in \mathbb{R}^L$  is the centroid vector of cluster  $k$ . The objective function  $J$  is defined as follows:

$$J = \sum_{d_i \in \mathcal{D}} \sum_{k \in [n_c]} r_{d_i,k} \|\bar{z}_{d_i} - \bar{\mu}_k\|_2^2, \tag{2}$$

where  $\bar{z}_{d_i}$  is a latent vector for a document  $d_i$ . After solving the K-means clustering, we obtain the  $k$ -th cluster

$$\mathcal{C}_k = \{d_i \in \mathcal{D} \mid r_{d_i,k} = 1\}. \tag{3}$$

#### 4.2 Estimation of document management criteria in a given cluster

The second step in our proposed method is to estimate the user’s document management criteria in each cluster (corresponding to Line 24–25 of Algorithm 1). We used meta-paths as the management criteria for documents and weighed the meta-paths from the user’s space.

We assume that the management criteria is unique to each cluster created by a user. The fundamental concept for determining the weight of the meta-path is that when two documents in a cluster are related to a meta-path, the user manages the cluster by considering the meta-path. From this insight, we calculate the weight of meta-path  $p$  for  $k$ -th cluster from the adjacency matrix as follows:

$$w_k^p = \frac{n_k^p}{\sum_{p \in \mathcal{P}} n_k^p} \quad (4)$$

where  $n_k^p$  is the number of paths which are assigned to the cluster  $k$  and have the meta-path  $p$  and is defined as follows:

$$n_k^p = \left| \left\{ d_i \in \mathcal{C}_k \mid \exists d_j \in \mathcal{C}_k : \mathbf{A}_{d_i, d_j}^p = 1 \right\} \right|. \quad (5)$$

Once the weights of the meta-paths within a cluster  $k$  are determined, the adjacency matrices are weighted accordingly. The weighted adjacency matrix is defined as follows:

$$\tilde{\mathbf{A}} = \sum_{k \in [n_c]} \sum_{p \in \mathcal{P}} w_k^p \mathbf{A}_{\mathcal{C}_k}^p, \quad (6)$$

### 4.3 Learning the latent vector of documents based on graph autoencoders

The third step in our proposed method is to learn the latent vector of documents based on graph autoencoders. (This corresponds to Line 26–30 of Algorithm 1.) We used the weighted adjacency matrix  $\tilde{\mathbf{A}}_k$  to obtain a latent vector in the latent space for each document. To this end, we constructed a graph convolutional network(GCN)-based encoder–decoder model with supervision from the user’s interactions.

#### 4.3.1 Encoder

Our encoder is GCN [18] with two layers.

Particularly, the latent vectors are calculated using the following equations:

$$\mathbf{Z} = GCN_\phi(\mathbf{X}, \tilde{\mathbf{A}}), \quad (7)$$

where  $\mathbf{X}$  is the feature matrix.  $GCN$  is defined as

$$GCN_\phi(\mathbf{X}, \tilde{\mathbf{A}}) = \hat{\mathbf{A}}ReLU(\hat{\mathbf{A}}\mathbf{X}\mathbf{W}^{(0)})\mathbf{W}^{(1)} \quad (8)$$

with the  $GCN$  parameter set  $\phi = \{\mathbf{W}^{(0)}, \mathbf{W}^{(1)}\}$  where  $\mathbf{W}^{(0)} \in \mathbb{R}^{|\mathcal{D}| \times h_1}$  is the weight of first layer and  $\mathbf{W}^{(1)} \in \mathbb{R}^{h_1 \times L}$  is the weight of the second layer.  $\hat{\mathbf{A}}$  is defined as

$$\hat{\mathbf{A}} = \mathbf{D}^{-\frac{1}{2}} \tilde{\mathbf{A}} \mathbf{D}^{-\frac{1}{2}}. \quad (9)$$

The decoder reconstructs the adjacency matrix  $\tilde{\mathbf{A}}$  by computing the probability  $p_\theta(\tilde{\mathbf{A}}|\mathbf{Z}_\mathcal{D})$  of the edge generation based on the latent vector of each document where

$$p_\theta(\tilde{\mathbf{A}}|\mathbf{Z}) = \prod_{d_i \in \mathcal{D}} \prod_{d_j \in \mathcal{D}} p_\theta(\tilde{\mathbf{A}}_{d_i, d_j} | \vec{z}_{d_i}, \vec{z}_{d_j}). \quad (10)$$

The decoder in the generative model was configured using the Euclidean distance between the latent vectors. This is intended to increase the probability of generating edges between documents that are placed closer together because the user provides feedback to the system based on the distance between documents. The decoder is expressed as follows:

$$p_\theta(\tilde{\mathbf{A}}_{d_i, d_j} | \vec{z}_{d_i}, \vec{z}_{d_j}) = \sigma \left( \frac{a}{\|\vec{z}_{d_i} - \vec{z}_{d_j}\|_2^2} + b \right), \quad (11)$$

where  $\sigma(\cdot)$  denotes a sigmoid function and  $\theta = \{a, b\}$  denotes a set of parameters used in the decoder.

#### 4.3.2 Objective function

The objective function consists of cross-entropy loss for generating the adjacency matrix and supervision from the interaction by the user. Parameters  $\phi = \{\phi\}$  and  $\theta = \{\theta\}$  are learned to maximize them.

The cross-entropy used to generate the adjacency matrix is defined as follows:

$$\mathcal{L}_{GAE} = \sum_{d_i \in \mathcal{D}} \sum_{d_j \in \mathcal{D}} \tilde{\mathbf{A}}_{d_i, d_j} \log p_{\theta_k}(\tilde{\mathbf{A}}_{d_i, d_j} | \vec{z}_{d_i}, \vec{z}_{d_j}) \quad (12)$$

$$= \mathbb{E}_{GCN_\phi(\mathbf{X}, \tilde{\mathbf{A}})}[\log p_\theta(\tilde{\mathbf{A}}|\mathbf{Z})] \quad (13)$$

Moreover, we define a loss function that measures the difference between the user’s feedback and the learned latent vectors to minimize disagreement. We measured this disagreement using the conditional probability that given the user feedback and the generation probability of the latent vector. The objective function is defined as follows:

$$\mathcal{L}_{feedback} = \log p(GCN_\phi(\mathbf{X}, \tilde{\mathbf{A}}) | \hat{\mathcal{Z}}) \quad (14)$$

$$= \sum_{(\vec{z}_{d_i}, \hat{z}_{d_i}) \in \hat{\mathcal{Z}}} \log \mathcal{N}(\vec{z}_{d_i} | \hat{z}_{d_i}, \sigma^2 \mathbf{I}) \quad (15)$$

$$= - \sum_{(\vec{z}_{d_i}, \hat{z}_{d_i}) \in \hat{\mathcal{Z}}} \left\| \vec{z}_{d_i} - \hat{z}_{d_i} \right\|_2^2 + const., \quad (16)$$

where  $\mathcal{N}(\vec{x} | \vec{\mu}, \Sigma)$  denotes the multivariate normal distribution and  $\vec{z}_{d_i}^k$  represents a latent vector generated by the encoder. The overall optimization problem is defined as follows:

$$\phi, \theta = \arg \max_{\phi, \theta} \mathcal{L}_{GAE} + \alpha \mathcal{L}_{feedback}, \tag{17}$$

where  $\alpha$  denotes the hyper-parameter.

### 4.4 Sampling strategy for active learning

In this section, we introduce an active learning sampling strategy for ISLE. (This step corresponds to Line 20–21 in Algorithm 1) Our strategy is a type of *uncertainty sampling* [22], which selects a sample that provides us one of the most informative answers to improve the model. The challenge here is to formalize the notion of uncertainty in our setting. To efficiently obtain users’ literature management criteria, we ask users for more important data in acquiring a user’s literature management criteria. Therefore, we define our uncertainty as the uncertainty of the meta-path weight in the clusters, where the weights are represented as a Dirichlet distribution, and measure the uncertainty as the information entropy of the probability distribution. The difference between the entropies of the prior and posterior distributions is defined as an increase in information acquired by asking the user. In our query strategy, our system is designed to ask a user for the most informative data, regardless of the cluster in which the requested data will be placed.

The prior distribution of the meta-path weight in the  $k$ -th cluster is a Dirichlet distribution, which is defined as follows:

$$p(\pi | \mathbf{n}_k) = Dir(\pi | \mathbf{n}_k) = C_D(\mathbf{n}_k) \pi_p^{n_k^p - 1} \tag{18}$$

where  $C_D(\mathbf{n}_k)$  is normalizing constant;  $C_D(\mathbf{n}_k) = \frac{\Gamma(\sum_p n_k^p)}{\prod_p \Gamma(n_k^p)}$   $\prod_{p \in \mathcal{P}} \mathbf{n}_k = (n_k^1, n_k^2, \dots, n_k^{|P|})$ , and  $n_k^p$  is defined as Eq. (5). After the user provides feedback on the position of a new paper, the posterior distribution is calculated as follows:

$$p(\pi | \mathbf{n}_k, \mathbf{n}_{new}) = \frac{p(\pi, \mathbf{n}_{new} | \mathbf{n}_k)}{p(\mathbf{n}_{new} | \mathbf{n}_k)} \propto p(\pi, \mathbf{n}_{new} | \mathbf{n}_k) \tag{19}$$

$$= Multi(\mathbf{n}_{new} | \pi) Dir(\pi | \mathbf{n}_k) \tag{20}$$

$$\propto \prod_{p \in \mathcal{P}} \pi_p^{n_{new}^p} \prod_{p \in \mathcal{P}} \pi_p^{n_k^p - 1} \tag{21}$$

$$= \prod_{p \in \mathcal{P}} \pi_p^{n_k^p + n_{new}^p - 1} \tag{22}$$

$$\propto Dir(\pi | \mathbf{n}_k + \mathbf{n}_{new}) \tag{23}$$

where  $\mathbf{n}_{new} = (n_{new}^1, n_{new}^2, \dots, n_{new}^{|P|})$  is the number of new paths in the meta-path  $p$  in a cluster when a new document  $d_{new}$  is added to the cluster:

$$n_{new}^p = \left| \left\{ d_i \in \mathcal{C}_k \mid \mathbf{A}_{d_i, d_{new}}^p = 1 \right\} \right| \tag{24}$$

Based on the prior and posterior distribution, the information gain in a cluster  $k$  is defined as follows:

$$\Delta H_k = \mathbb{E}[-\log Dir(\pi | \mathbf{n}_k)] - \mathbb{E}[-\log Dir(\pi | \mathbf{n}_k + \mathbf{n}_{new})] \tag{25}$$

where the entropy of the Dirichlet distribution is calculated as follows:

$$\begin{aligned} &\mathbb{E}[-\log Dir(\pi | \alpha)] \\ &= -\sum_{k=1}^K (\alpha_k - 1) (\psi(\alpha_k) - \psi(\sum_{i=1}^K \alpha_i)) - \ln C_D(\alpha) \end{aligned} \tag{26}$$

$\psi$  is the digamma function. Based on the decrease of the entropy in a cluster  $k$ , we define the overall information gain of the asked data as the sum of them:

$$\phi_{gain} = \sum_{k \in [n_c]} \Delta H_k \tag{27}$$

Based on the above criteria, we ask for the most informative data for the user and get feedback.

## 5 Experiment

We conducted an experiment to answer our three research questions and determine the effectiveness of the method. For RQ1, we compared our method with its variation that assumes the consistency of the criteria across the latent space. For RQ2, we compared our framework with a popular encoder–decoder model for graphs as a baseline that uses an inner-product-based decoder. For RQ3, we compared the effectiveness of the proposed method using the active query strategy using randomly selected queries.

### 5.1 Settings

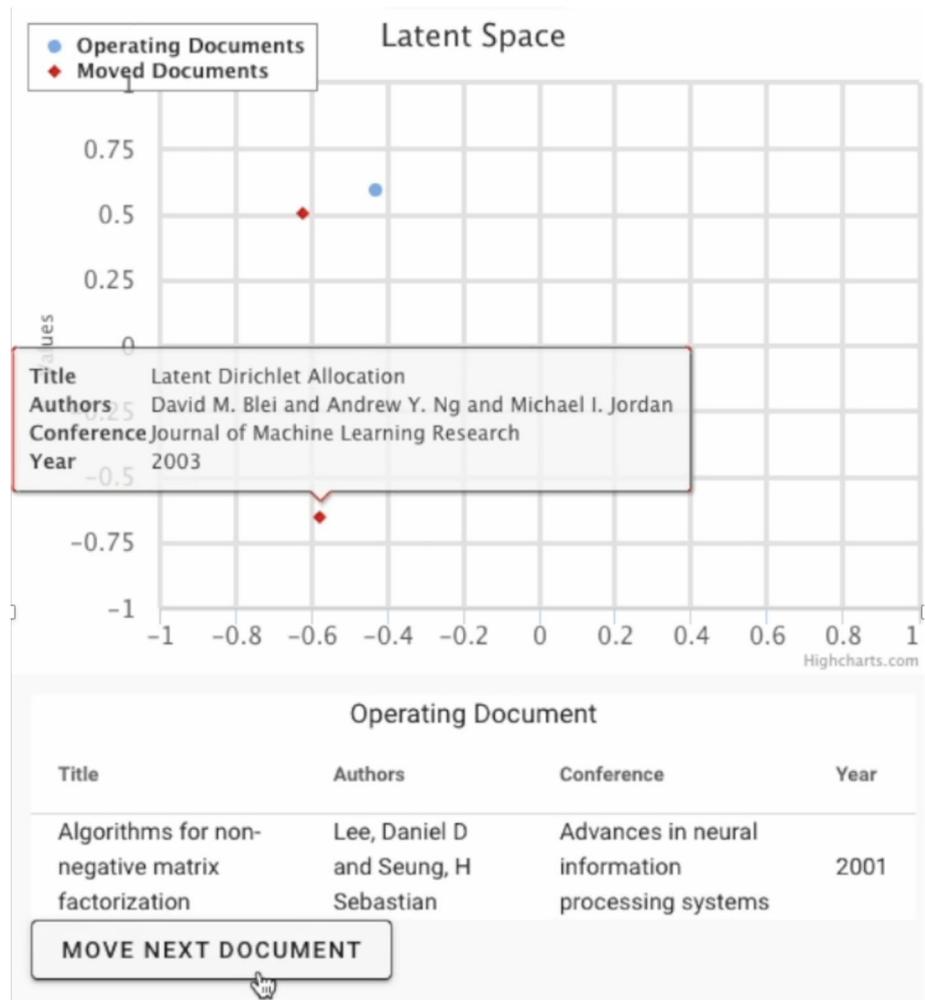
#### 5.1.1 Interface

We developed a two-dimensional literature management tool prototype to conduct our experiments. Figure 3 shows the actual interface used in the experiments. The interface displays a two-dimensional space to place the icon of the papers, and the subjects place the papers onto the space.

#### 5.1.2 Participants

We recruited ten researchers (a humanities domain researcher, two data engineering domains, three HCI(Human Computer Interaction) domains and four ML(Machine Learning) domain researchers).

**Fig. 3** Interface we developed to conduct the experiments. The interface enables us to put the papers into a two-dimensional space. The red and blue dots represent the put papers and an operating paper, respectively. While operating the tool, the biblio-record of the operating paper is displayed at the bottom of the interface, and when the mouse floats on the icon of the paper, the biblio-record of the paper is displayed on the tool



### 5.1.3 Data collection

First, we asked each of the participants to send us the BibTeX records of any 50 papers related to his or her research. Second, we asked them to use our tool, in such a way that the tool shows the biblio-records in random order and the user puts each into the two-dimensional space. As a result, we obtained the history of how they behaved in the 50 iterations, that is, how they moved their papers to incrementally place all 50 papers in their spaces. The user sees the title, author, conference, and year of publication of each paper in the phase. Figure 4 shows the spaces created by three participants. The distribution of the icon of the papers and the placement scheme of the subjects differ from each other, indicating that there are various document management criteria for each subject.

### 5.1.4 Evaluation

First, we randomly selected 10 data points and let them be the unknown documents,  $Q$ . The remaining 40 data were

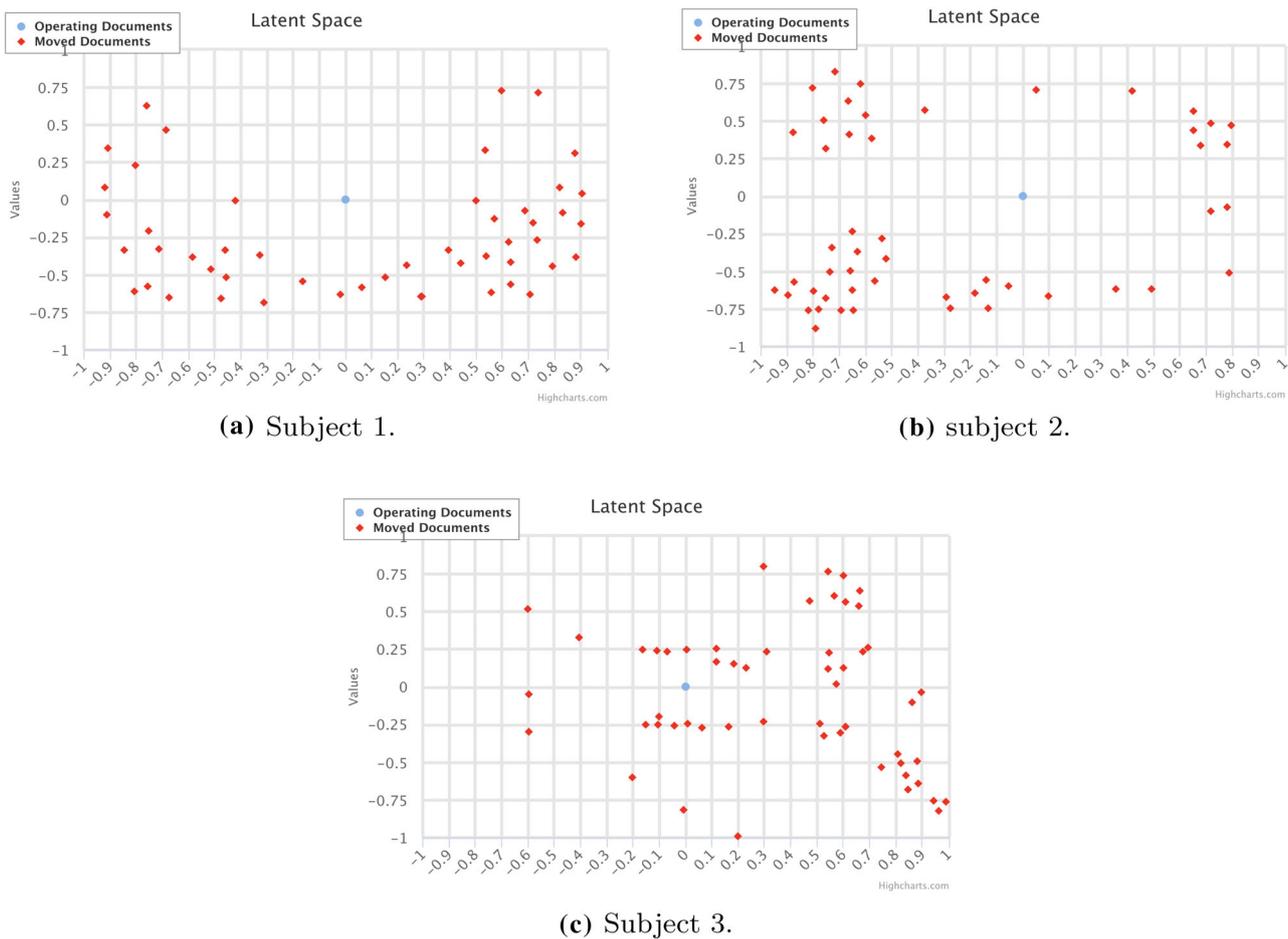
used as training data, and when the user interacted with them one by one, we simulated whether  $Q$  could be placed in the position expected by the user. In the experiments, we set the hyper-parameters to  $L = 2$ ,  $\alpha = 100$ ,  $n_c = 6$ , and  $h_1 = 4$ .

### 5.1.5 Metrics

We used Recall@k and nDCG@k [23] where k=6. Recall@k is expressed using the following equation:

$$Recall@k = \frac{|\mathcal{U} \cap \mathcal{P}_k|}{|\mathcal{U}|}, \quad (28)$$

where  $\mathcal{U}$  denotes the set of the closest (with the Euclidean distance)  $k$  documents in the latent space to the test data placed by the user, and  $\mathcal{P}_k$  is the set of the closest  $k$  latent vectors to the position of the test data predicted by the model. nDCG@k [23] is obtained by dividing the value of DCG@k by the most ideal value of DCG@k, that is, if all model predictions are correct. The inverse of the distance from the correct position was used as the relevance value.



**Fig. 4** Spaces that each subject created. The distribution of icons and putting scheme is different in each subject, which clearly indicates the management criteria differ from each other

### 5.1.6 Active learning

Since our active learning strategy (Sect. 4.4) required a seed to choose the next query, we randomly choose the first query in the experiment.

## 5.2 Baselines and variations

**(1) VGAE.** VGAE is a popular encoder–decoder model for graphs [19]. In our experiment, VGAE is a variant of ISLE, in which the decoder is replaced by the decoder used in ordinary VGAE. That is, the decoder expressed in Eq. (11) in Sect. 5 is replaced with the inner product of each latent representation, which is expressed as follows:

$$p_{\theta}(\mathbf{A}_{d_i, d_j} | \vec{z}_{d_i}, \vec{z}_{d_j}) = \sigma(\vec{z}_{d_i} \cdot \vec{z}_{d_j}), \quad (29)$$

where  $\sigma$  denotes the sigmoid function. Note that the meta-paths are considered to create the adjacency matrix, and Step 1 (clustering) (Sect. 4.1) is applied.

**Table 2** Meaning of each meta-path

Meta-path	Meaning
PAP	Papers which the same author wrote
PTP	Papers that mention the same word in the title
PYP	Papers published in the same year
PVP	Papers presented at the same conference

**(2) ISLE and VAGE without clustering.** We used VGAE and ISLE which omit Step 1 (clustering) (Sect. 4.1), to address RQ1.

**(3) ISLE with different sets of meta-paths.**

The meta-paths used are listed in Table 2. We compared the following five cases for ISLE, while we used all meta-paths for VGAE.

(a) **ALL:** The adjacency matrix comprises the PAP, PTP, PYP, and PVP meta-paths. (b) **PAP Only:** The adjacency matrix is composed of PAP only. (c) **PTP Only:** The adjacency matrix is composed of PTP only. (d) **PYP Only:** The

**Algorithm 1** The active ISLE process for obtaining a new document  $q$ 's latent vector

**Input:** The node feature matrix  $\mathbf{X}$ , the adjacency matrix  $\mathbf{A}^p$  consisting with each meta path  $p$ , user's feedback  $\hat{\mathcal{Z}}$ , and a new document  $q$   
**Output:** a new document  $q$ 's latent vector  $\vec{z}_q$

- 1: **function** CLUSTERING( $\hat{\mathcal{Z}}$ )
- 2:  $\mathbf{r}_{d_i} \leftarrow \arg \min_{\mathbf{r}_{d_i}} \sum_{d_i, k} r_{d_i, k} \|\hat{\mathbf{z}}_{d_i} - \vec{\mu}_k\|_2^2$
- 3:  $\mathcal{C}_k \leftarrow \{d_i \in \mathcal{D} \mid r_{d_i, k} = 1\}$
- 4: **return**  $\mathcal{C}_k$
- 5: **end function**
- 6:
- 7: **function** WEIGHTING ADJACENCY MATRIX
- 8:  $w_k^p \leftarrow \frac{n_k^p}{\sum_{p \in \mathcal{P}} n_k^p}$
- 9:  $\tilde{\mathbf{A}} \leftarrow \sum_{k \in [n_c]} \sum_{p \in \mathcal{P}} w_k^p \mathbf{A}_{\mathcal{C}_k}^p$
- 10: **return**  $\tilde{\mathbf{A}}$
- 11: **end function**
- 12:
- 13: **function** TRAINING( $\mathbf{X}$ ,  $\tilde{\mathbf{A}}$ ,  $\hat{\mathcal{Z}}$ )
- 14:  $\mathcal{L}_{GAE} \leftarrow \log p_{\theta}(\tilde{\mathbf{A}} \mid GCN_{\phi}(\mathbf{X}, \tilde{\mathbf{A}}))$
- 15:  $\mathcal{L}_{feedback} \leftarrow \log p(GCN_{\phi}(\mathbf{X}, \tilde{\mathbf{A}}) \mid \hat{\mathcal{Z}})$
- 16:  $\mathcal{L} \leftarrow \mathcal{L}_{GAE} + \alpha \mathcal{L}_{feedback}$
- 17: **return**  $\mathcal{L}$
- 18: **end function**
- 19:
- 20: **Step 0 (Section 4.4):** Choosing a query for asking a user by using the sampling strategy for active learning (This step is required for the active learning setting.)
- 21: Choosing a query for asking a user by using the sampling strategy for active learning 27.
- 22: **Step 1 (Section 4.1):** Clustering the latent vectors by using the user's feedback.
- 23:  $\mathcal{C}_k \leftarrow \text{CLUSTERING}(\hat{\mathcal{Z}})$
- 24: **Step 2 (Section 4.2):** Estimating the document management criteria in each cluster.
- 25:  $\tilde{\mathbf{A}} \leftarrow \text{WEIGHTING ADJACENCY MATRIX}$
- 26: **Step 3 (Section 4.3):** Learning the latent vectors of documents based on Graph Auto-Encoders.
- 27: **for** epoch in  $1, \dots, n_{epoch}$  **do**
- 28:  $\mathcal{L} \leftarrow \text{TRAINING}(\mathbf{X}, \tilde{\mathbf{A}}, \hat{\mathcal{Z}})$
- 29:  $\phi, \theta \leftarrow \arg \max_{\phi, \theta} \mathcal{L}$
- 30: **end for**

adjacency matrix is composed of PYP only. (e) PVP Only: The adjacency matrix is composed of PVP only.

**(4) Active ISLE.** ISLE implements the sampling strategy introduced in Sect. 4.4

### 5.3 Results: passive setting

Figures 5a–6b show the result. The solid line in each figure represents the mean, and the shaded area represents the 95% confidence interval. The red lines in each figure indicate the results of our proposed method when the adjacency matrix provided as the input consists of ALL, as described in

Sect. 5.2. The blue lines in Fig. 5 indicate the results of VGAE when the adjacency matrix provided as the input consists of ALL, as described in Sect. 5.2. The yellow and olive lines in the Fig. 5 indicate the results of estimating the document management criteria without using clusters in the proposed method. The green, peach, purple, and gray lines in Fig. 6 depict the limited types of meta-paths given as inputs in the proposed method. The figures demonstrate that the ISLE outperformed all the methods and that the accuracy improved as the number of feedbacks increased.

Note that in our context, recall@k indicates how close the predicted position is to the correct position, whereas nDCG@k indicates how it maintains the order of distances. Unlike in the ordinary information retrieval context, Recall@k is more critical for our problem because the order of distances can dramatically change, even if the position is slightly moved.

Figure 6a compares for the results with different sets of meta-paths. The results show that ISLE performs the best when we use all of the four meta-paths. As we noted in the limitation part, finding the best feature set was not our research question. However, this implies that researchers are aware of multiple criteria when managing papers and that the proposed method can flexibly express these criteria by using multiple meta-paths.

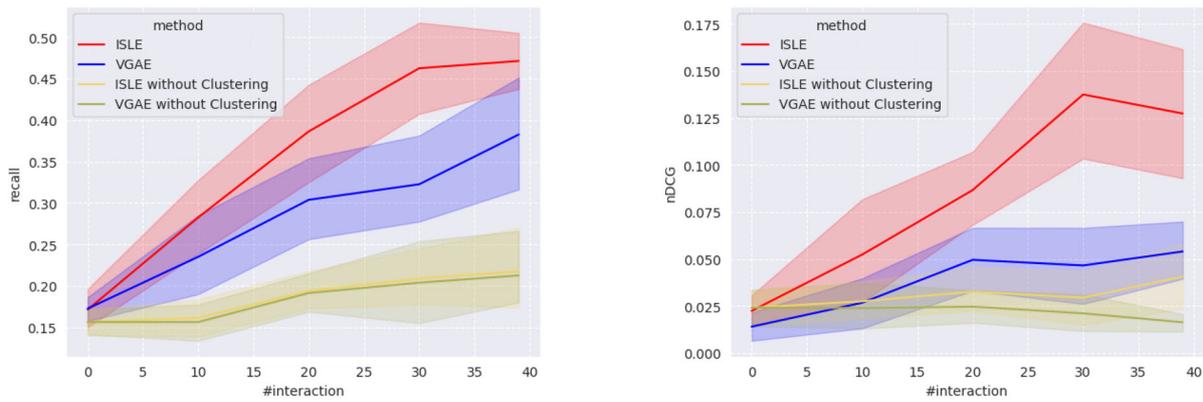
## 5.4 Results: active setting

Figure 8 compares the results of ISLE and Active ISLE, where the blue and orange lines show the results for Active ISLE and ISLE, respectively. As the Fig. 8 shows, Active ISLE exhibits a higher recall value with fewer interactions. This implies that the sampling strategy works well for quickly identifying the criteria for each document cluster.

## 6 Discussion

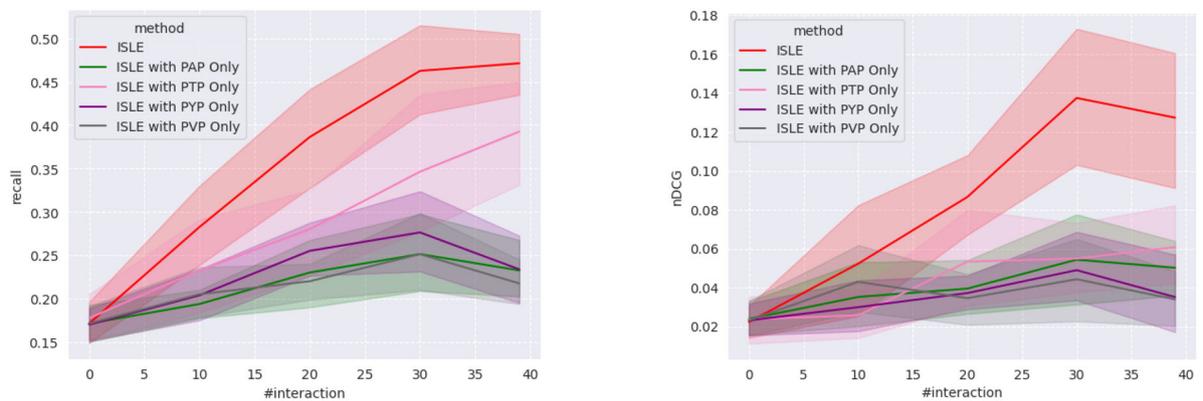
### 6.1 The locality of criteria in the latent space (RQ1)

The results shown in Fig. 5a and b clearly indicate that methods with a clustering phase are superior to those without clustering. This shows that the clusters of each researcher have a different set of weights for meta-paths, which means that researchers use different criteria in the sub-spaces in their latent space. Figure 7 shows the normalized distribution of the meta-path weights in each cluster for three of the ten subjects. Although PTP accounts for a large proportion of their distributions, their weights are often considerably different to each cluster even for the same researcher.



(a) Value of recall@k for the proposed method and VGAE (b) Value of nDCG@k for the proposed method and VGAE

Fig. 5 Horizontal axis represents the number of times feedback is received from the user and the vertical axis represents the values of recall and nDCG. The results of ISLE outperformed other methods



(a) Value of recall@k for the proposed method and when the type of meta-path is limited (b) Value of nDCG@k for the proposed method and when the type of meta-path is limited

Fig. 6 Horizontal axis represents the number of times feedback is received from the user the vertical axis represents the values of recall and nDCG. The results of ISLE using multiple meta-paths were more accurate

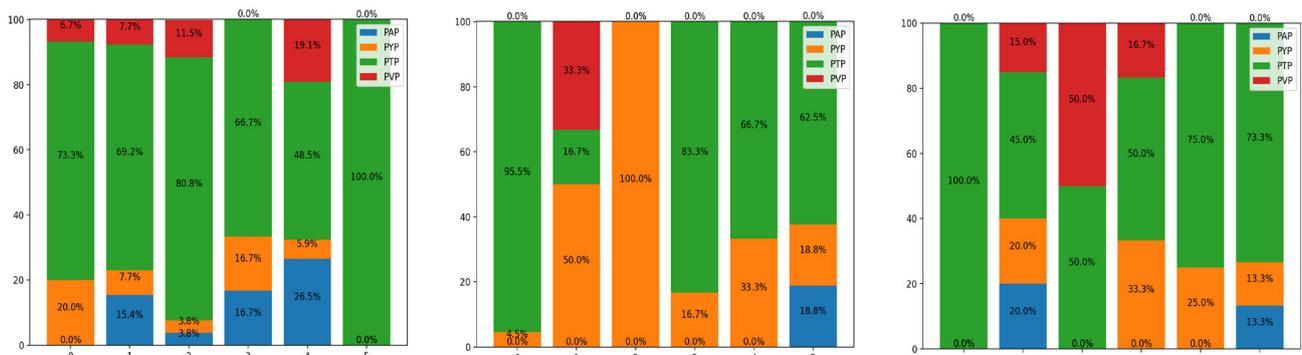
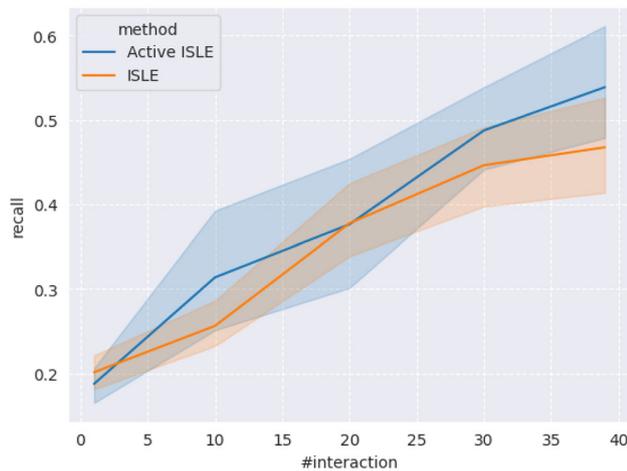


Fig. 7 Normalized distribution of meta-path weights in each cluster. These figures show the ratio of the meta-paths in each cluster in the two-dimensional space that each subject made. These figures indicate the management criteria are very different in each cluster, even in the same subject



**Fig. 8** Comparison of the proposed method with and without active learning. The horizontal axis represents the number of times feedback is received from the user, and the vertical axis represents the values of recall@k. The results of the proposed method using active learning were more accurate

### 6.2 Effectiveness of the Euclidean distance-based decoder (RQ2)

The idea behind our second assumption is that the user provides feedback in the latent space based on the Euclidean distance rather than the angle between documents (which is the principle of the VGAE's decoder). Therefore, the ISLE decoder, which calculates the generation probability of edges based on the Euclidean distance between documents, is more accurate. The results presented in Fig. 5a and b clearly support this assumption.

### 6.3 Effectiveness of the active learning (RQ3)

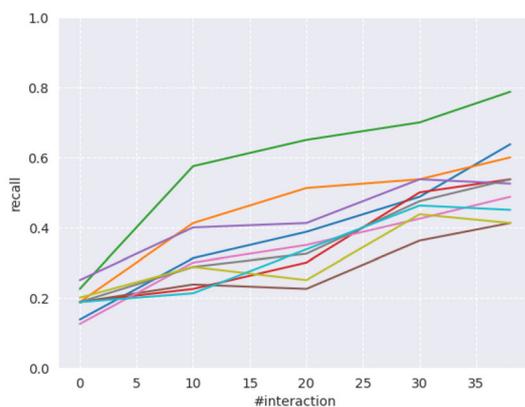
Figure 8 shows that the accuracy of the active learning method is generally higher than that of the random sampling

method. The main objective of introducing active learning is to achieve high accuracy with a small number of interactions by asking the user for an informative node. Figure 8 shows that this objective was achieved by introducing active learning. When comparing the recall values at the 10th interaction, a statistically significance was observed. This implies the potential effectiveness of query strategies that focus on the uncertainty of meta-path importance. In this experiment, the first dataset was randomly selected. An effective way to select the first data point is a subject for future studies.

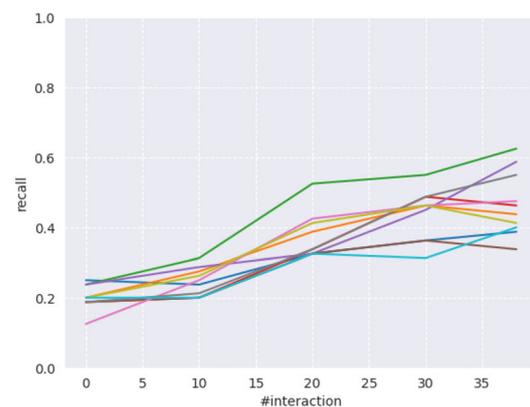
### 6.4 Individual difference

We collected data from ten researchers, and Fig. 9 shows how each researcher's feedback affected the accuracy of the data. Figure 9a, b shows that the accuracy generally improves as the number of feedback cycles increases. Our findings show that the management criteria of each researcher can be captured using meta-paths, although there are individual differences. In addition, the accuracy of the active learning method was higher than that of the random sampling method in almost all the researchers.

The degree of accuracy improvement through interaction varies from user to user. We assume that this is due to the manner in which users create their latent space. Figure 4 illustrates how each subject creates a latent space. If the user has cluster regions that are clearly divided in the latent space, we can estimate the criteria for managing the literature. However, if the clusters were ambiguous and could not be clustered according to a user's expected document management criteria, we believe that the increase in interactions did not dramatically improve the accuracy. The following are possible reasons for the decrease in accuracy during the experiment: (1) The criteria changed during the experiment and the cluster was



**(a)** Value of recall@k of Active Learning by each researcher



**(b)** Value of recall@k of Random Sampling by each researcher

**Fig. 9** Comparison of recall by each researcher

reconstructed. (2) The English paper is mixed with papers in another language.

## 7 Conclusion and future work

In this study, we proposed a method that estimates a user's document management criteria based on human-in-the-loop latent space learning.

The experimental results showed that the proposed method accurately placed unknown documents at the user's desired position compared with the baseline method. In addition, experiments with multiple and a limited number of meta-paths showed that the proposed method (ISLE) is more accurate when multiple meta-paths were used, indicating that ISLE is effective even when users manage documents according to various criteria. Based on the above results, we added an active learning framework to estimate a user's document management criteria with a fewer number of interactions. The experimental results demonstrated the effectiveness of active learning.

In the future, we intend to study (1) develop a document management system based on ISLE that can be used in the real-world (2) consider using longer meta-paths. The realization of these goals will not only provide a better method for human-in-the-loop latent space learning but will also provide support for researchers in literature management.

**Acknowledgements** This work was partially supported by Kumagai Gumi Co., Ltd., JSPS KAKENHI Grant Number 22H00508, 22K17944, 21H03552, and JST CREST. This study was approved by the IRB of the University of Tsukuba. We thank to Masao Takaku for his valuable comments.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bates, M.J.: The design of browsing and berrypicking techniques for the online search interface. *Online Rev.* (1989). <https://doi.org/10.1108/eb024320>
- Cai, H., Zheng, V.W., Chang, K.C.C.: Active learning for graph embedding (2017). arXiv preprint [arXiv:1705.05085](https://arxiv.org/abs/1705.05085)
- Chen, X., Yu, G., Wang, J., Domeniconi, C., Li, Z., Zhang, X.: Activehne: active heterogeneous network embedding (2019). arXiv preprint [arXiv:1905.05659](https://arxiv.org/abs/1905.05659)
- Deng, Y., Yuan, Y., Fu, H., Qu, A.: Query-augmented active metric learning. *J. Am. Stat. Assoc.* (2022). <https://doi.org/10.1080/01621459.2021.2019045>
- Dong, Y., Chawla, N.V., Swami, A.: metapath2vec: scalable representation learning for heterogeneous networks. In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 135–144 (2017). <https://doi.org/10.1145/3097983.3098036>
- Fire, M., Guestrin, C.: Over-optimization of academic publishing metrics: observing Goodhart's law in action. *GigaScience* **8**(6), giz053 (2019). <https://doi.org/10.1093/gigascience/giz053>
- Francesca, E.: Usage of reference management software at the University of Torino. In: *Usage of Reference Management Software at the University of Torino*, pp. 145–174 (2013). <https://doi.org/10.4403/jlis.it-8679>
- Fu, X., Zhang, J., Meng, Z., King, I.: Magnn: metapath aggregated graph neural network for heterogeneous graph embedding. In: *Proceedings of the Web Conference 2020*, pp. 2331–2341 (2020). <https://doi.org/10.1145/3366423.3380297>
- Grover, A., Leskovec, J.: node2vec: scalable feature learning for networks. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 855–864 (2016). <https://doi.org/10.1145/2939672.2939754>
- Hartigan, J.A., Wong, M.A.: Algorithm as 136: a k-means clustering algorithm. *J. R. Stat. Soc. Ser. C (Applied Statistics)* **28**(1), 100–108 (1979). <https://doi.org/10.2307/2346830>
- Hsu, Y.L., Tsai, Y.C., Li, C.T.: Fingat: financial graph attention networks for recommending top-k profitable stocks. *IEEE Trans. Knowl. Data Eng.* (2021). <https://doi.org/10.1109/TKDE.2021.3079496>
- Hu, X., Yoo, I.: A comprehensive comparison study of document clustering for a biomedical digital library medline. In: *Proceedings of the 6th ACM/IEEE-CS joint conference on digital libraries (JCDL'06)*, pp. 220–229. IEEE (2006). <https://doi.org/10.1145/1141753.1141802>
- Huang, X., Qian, S., Fang, Q., Sang, J., Xu, C.: Meta-path augmented sequential recommendation with contextual co-attention network. *ACM Trans. Multimed. Comput. Commun. Appl. (TOMM)* **16**(2), 1–24 (2020). <https://doi.org/10.1145/3382180>
- Iwayama, M., Tokunaga, T.: Hierarchical Bayesian clustering for automatic text classification. In: *Proceedings of the 14th International Joint Conference on Artificial Intelligence-Volume 2*, pp. 1322–1327 (1995)
- Josh, N., Pepe, A.: 65 out of the 100 most cited papers are paywalled (2019). <https://www.authorea.com/users/8850/articles/125400-65-out-of-the-100-most-cited-papers-arepaywalled>. Accessed 30 June 2022
- Kang, Y., Hou, A., Zhao, Z., Gan, D.: A hybrid approach for paper recommendation. *IEICE Trans. Inf. Syst.* **104**(8), 1222–1231 (2021). <https://doi.org/10.1587/transinf.2020BDP0008>
- Kingma, D.P., Welling, M.: Auto-encoding variational Bayes (2013). arXiv preprint [arXiv:1312.6114](https://arxiv.org/abs/1312.6114)
- Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks (2016). arXiv preprint [arXiv:1609.02907](https://arxiv.org/abs/1609.02907)
- Kipf, T.N., Welling, M.: Variational graph auto-encoders (2016). arXiv preprint [arXiv:1611.07308](https://arxiv.org/abs/1611.07308)
- Le, Q., Mikolov, T.: Distributed representations of sentences and documents. In: *International Conference on Machine Learning*, pp. 1188–1196. PMLR (2014)
- Lee, J., Lee, K., Kim, J.G.: Personalized academic research paper recommendation system (2013). arXiv preprint [arXiv:1304.5457](https://arxiv.org/abs/1304.5457)
- Lewis, D.D.: A sequential algorithm for training text classifiers: corrigendum and additional data. In: *Acm Sigir Forum*, vol. 29, pp.

- 13–19. ACM New York (1995). <https://doi.org/10.1145/219587.219592>
23. Liang, D., Krishnan, R.G., Hoffman, M.D., Jebara, T.: Variational autoencoders for collaborative filtering. In: Proceedings of the 2018 World Wide Web Conference, pp. 689–698 (2018). <https://doi.org/10.1145/3178876.3186150>
  24. Ma, X., Wang, R.: Personalized scientific paper recommendation based on heterogeneous graph representation. *IEEE Access* **7**, 79887–79894 (2019). <https://doi.org/10.1109/ACCESS.2019.2923293>
  25. Mikawa, K., Goto, M.: Regularized distance metric learning for document classification and its application. *J. Jpn. Ind. Manag. Assoc.* **66**(2E), 190–203 (2015). <https://doi.org/10.11221/jima.66.190> <https://doi.org/10.11221/jima.66.190>
  26. Nadagouda, N., Xu, A., Davenport, M.A.: Active metric learning and classification using similarity queries. In: Uncertainty in Artificial Intelligence, pp. 1478–1488. PMLR (2023)
  27. Perozzi, B., Al-Rfou, R., Skiena, S.: Deepwalk: online learning of social representations. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 701–710 (2014). <https://doi.org/10.1145/2623330.2623732>
  28. Salehi, A., Davulcu, H.: Graph attention auto-encoders (2019). arXiv preprint [arXiv:1905.10715](https://arxiv.org/abs/1905.10715).
  29. Salha-Galvan, G., Hennequin, R., Chapus, B., Tran, V.A., Vazirgiannis, M.: Cold start similar artists ranking with gravity-inspired graph autoencoders. In: Fifteenth ACM Conference on Recommender Systems, pp. 443–452 (2021). <https://doi.org/10.1145/3460231.3474252>
  30. Scharpf, P., Schubotz, M., Youssef, A., Hamburg, F., Meuschke, N., Gipp, B.: Classification and clustering of arxiv documents, sections, and abstracts, comparing encodings of natural and mathematical language. In: Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020, pp. 137–146 (2020). <https://doi.org/10.1145/3383583.3398529>
  31. Sherkat, E., Nourashrafeddin, S., Milios, E.E., Minghim, R.: Interactive document clustering revisited: a visual analytics approach. In: 23rd International Conference on Intelligent User Interfaces, pp. 281–292 (2018). <https://doi.org/10.1145/3172944.3172964>
  32. Sun, Y., Han, J.: Meta-path-based search and mining in heterogeneous information networks. *Tsinghua Sci. Technol.* **18**(4), 329–338 (2013). <https://doi.org/10.1109/TST.2013.6574671>
  33. Tang, J., Qu, M., Mei, Q.: Pte: predictive text embedding through large-scale heterogeneous text networks. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1165–1174 (2015). <https://doi.org/10.1145/2783258.2783307>
  34. Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., Mei, Q.: Line: large-scale information network embedding. In: Proceedings of the 24th International Conference on World Wide Web, pp. 1067–1077 (2015). <https://doi.org/10.1145/2736277.2741093>
  35. Vlaev, I.: Local choices: rationality and the contextuality of decision-making. *Brain Sci.* **8**(1), 8 (2018). <https://doi.org/10.3390/brainsci8010008>
  36. Waheed, W., Imran, M., Raza, B., Malik, A.K., Khattak, H.A.: A hybrid approach toward research paper recommendation using centrality measures and author ranking. *IEEE Access* **7**, 33145–33158 (2019). <https://doi.org/10.1109/ACCESS.2019.2900520>
  37. Wang, J., Wu, S., Vu, H.Q., Li, G.: Text document clustering with metric learning. In: Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 783–784 (2010). <https://doi.org/10.1145/1835449.1835614>
  38. Wang, W., Tang, T., Xia, F., Gong, Z., Chen, Z., Liu, H.: Collaborative filtering with network representation learning for citation recommendation. *IEEE Trans. Big Data* **8**(5), 1233–1246 (2020). <https://doi.org/10.1109/TBDATA.2020.3034976>
  39. Watanabe, S., Ito, H., Matsubara, M., Morishima, A.: Bibrecord-based literature management with interactive latent space learning. In: Proceeding of 24th International Conference on Asian Digital Libraries, ICADL, pp. 155–171. Springer (2022). [https://doi.org/10.1007/978-3-031-21756-2\\_13](https://doi.org/10.1007/978-3-031-21756-2_13)
  40. Wei, C.P., Chiang, R.H., Wu, C.C.: Accommodating individual preferences in the categorization of documents: a personalized clustering approach. *J. Manag. Inf. Syst.* **23**(2), 173–201 (2006). <https://doi.org/10.2753/MIS0742-1222230208>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.