

Article

A Discriminative-Based Geometric Deep Learning Model for Cross Domain Recommender Systems

John Kingsley Arthur ^{*}, Conghua Zhou ^{*}, Eric Appiah Mantey , Jeremiah Osei-Kwakye and Yaru Chen

Computer Science Department, Jiangsu University, Zhenjiang 212013, China; ericphd@yahoo.com (E.A.M.); 5103170335@stmail.ujs.edu.cn (J.O.-K.); cyr793512702@outlook.com (Y.C.)

^{*} Correspondence: jkarthur@vvu.edu.gh (J.K.A.); chzhou@ujs.edu.cn (C.Z.)

Abstract: Recommender systems (RS) have been widely deployed in many real-world applications, but usually suffer from the long-standing user/item cold-start problem. As a promising approach, cross-domain recommendation (CDR), which has attracted a surge of interest, aims to transfer the user preferences observed in the source domain to make recommendations in the target domain. Traditional machine learning and deep learning methods are not designed to learn from complex data representations such as graphs, manifolds and 3D objects. However, current trends in data generation include these complex data representations. In addition, existing research works do not consider the complex dimensions and the locality structure of items, which however, contain more discriminative information essential for improving the performance accuracy of the recommender system. Furthermore, similar outcomes between test samples and their neighboring training data restrained in the kernel space are not fully realized from the recommended objects belonging to the same object category to capture the embedded discriminative information effectively. These challenges leave the problem of sparsity and the cold-start of items/users unsolved and hence impede the performance of the cross-domain recommender system, causing it to suggest less relevant and undistinguished items to the user. To handle these challenges, we propose a novel deep learning (DL) method, Discriminative Geometric Deep Learning (D-GDL) for cross-domain recommender systems. In the proposed D-GDL, a discriminative function based on sparse local sensitivity is introduced into the structure of the DL network. In the D-GDL, a local representation learning (i.e., a local sensitivity-based deep convolutional belief network) is introduced into the structure of the DL network to effectively capture the local geometric and visual information from the structure of the recommended 3D objects. A kernel-based method (i.e., a local sensitivity deep belief network) is also incorporated into the structure of the DL framework to map the complex structure of recommended objects into high dimensional feature space and achieve an effective recognition result. An improved kernel density estimator is created to serve as a weighing function in building a high dimensional feature space, which makes it more resistant to geometric noise and computation performance. The experiment results show that the proposed D-GDL significantly outperforms the state-of-the-art methods in both sparse and dense settings for cross-domain recommendation tasks.

Keywords: recommender systems; deep learning; cross domain; geometric deep learning; non-Euclidean domain



Citation: Arthur, J.K.; Zhou, C.; Mantey, E.A.; Osei-Kwakye, J.; Chen, Y. A Discriminative-Based Geometric Deep Learning Model for Cross Domain Recommender Systems. *Appl. Sci.* **2022**, *12*, 5202. <https://doi.org/10.3390/app12105202>

Academic Editor: Ángel González-Prieto

Received: 28 March 2022

Accepted: 14 May 2022

Published: 20 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Due to the proliferation of social networks coupled with the advancement of sophisticated electronic devices, there are very large amounts of multimedia content (including audio, videos and articles) on the Internet [1,2]. As a result, Internet users are able to access vast resources with just the click of a button. However, it is very challenging to retrieve information or multimedia content that is relevant to the user's interest. This problem has attracted much interest in research and application in recent times. Hence, recommender systems (RS) provide a machinery to effectively rank and suggest items for users based on

their past clicks on items they may be interested in. The suggested items usually improve and facilitates user's loyalty to a particular multimedia or social website through recommendations. Recommender systems can be categorized into specific (those that pertain to specific areas such as jobs, news and taxi or ride sharing) and general recommendations.

The Problem: Traditional machine learning and deep learning methods are not designed to learn from complex data representations such as graphs, manifolds and 3D objects. However, current trends in data generation include these complex data representations. In addition, existing research works do not consider the complex dimensions and the locality structure of items which, however, contain more discriminative information essential for improving the performance accuracy of the RS. Furthermore, similar outcomes between test samples and their neighboring training data restrained in the kernel space are not fully realized from the recommended objects belonging to the same object category to capture the embedded discriminative information effectively. These challenges leave the problem of sparsity and cold-start of items/users unsolved and hence impede the performance of the cross-domain RS, causing it suggest less relevant and undistinguished items to the user.

Gap analysis: Recent works such as Wang et al. (2021), Wang et al. (2020), Cao et al. (2022), Veeramachaneni et al. (2022), Zhu et al. (2022), Vikas et al. (2022), and Liao et al. (2022) have attempted to solve the cold-start and sparsity problems using machine learning and deep learning methods. Wang et al. (2021) proposed a framework for building cross-domain RSs that extracts personality trait information from an auxiliary domain using a probabilistic matrix factorization. Additionally, Wang et al. (2020) proposed a cross-domain RS based on a shared user's information from an e-commerce site and an advertising site with the aim of alleviating the cold-start and sparsity problems. In their work, a deep learning method, a generalized Matrix Factorization and a word2vec is applied to turn textual information on users and items into latent vectors as their representations. The word2vec technique faces the challenge of handling unknown or out-of-vocabulary words that it is unable to interpret or build vectors for.

Cao et al. (2022) proposed an information bottleneck (IB) principle to enforce user/item representations in a source domain to a target domain in capturing domain-shared information aimed at solving the cold-start problem. Veeramachaneni et al. (2022) proposed a transfer learning approach for cross-domain recommendation, wherein the cluster-level rating pattern (codebook) of the source domain is obtained via a co-clustering technique. The authors thereafter apply the Maximum Margin Matrix factorization (MMMMF) technique on the codebook in order to learn the user and item latent features of codebook. Zhu et al. (2022) also attempted to solve the cold-start problem by proposing a personalized transfer of user preferences for a cross-domain recommender system. Specifically, a meta network fed with users' characteristic embeddings is learned to generate personalized bridge functions to achieve personalized transfer of preferences for each user.

Vikas et al. (2022) utilizes Latent Dirichlet Allocation (LDA) and ontology methods to extract additional information from user reviews to create an ontological profile for each user by mapping genres to user profiles using a dictionary. In the work of Liao et al. (2022), domain-shared and domain-specific features are extracted to enable knowledge transfer between multiple heterogeneous source and target domains. To ensure positive transfer, the domain-shared subspaces from multiple domains are maximally matched by a multiclass domain discriminator in an adversarial learning process. The recommendation in the target domain is completed by a matrix factorization module with aligned latent features from both the user and the item side.

Moreover, the matrix factorization-based techniques suffer lower accuracy performances when the data is extremely sparse. Additionally, without significant modification to their models, they cannot be applied directly to non-Euclidean data feature extraction.

The traditional deep neural network was not originally designed to capture new data trends (non-Euclidean), such as graphs, manifolds, and 3D (Peng et al., 2021). Hence, they cannot correctly represent the local information of these complex data. Furthermore, most traditional deep networks are based on convolutions that do not include a generalization

operator, making them linearly incompatible with the representation of non-Euclidean data (Hu et al., 2022). On the other hand, geometric deep learning (GDL) is an emerging field of research that employs deep learning on non-Euclidean data. GDL has seen several successes in classification [3], drug discovery [4], biochemistry [5]; however, it has not been exploited in the area of recommender systems (either single-domain or cross-domain). However, these techniques could not fully exploit the local information existing in the complex and nonlinear structure of items/users concerning the magnitude of items and users as indicated in the following works [6–10].

Therefore, in this paper, considering the intrinsic and extrinsic feature capabilities of recommender systems from the manifolds domain, we propose a novel geometric deep-learning approach for cross-domain RSs based on different modalities to address the item/user cold-start and sparsity problems. The proposed method combines the advantages of geometric DL techniques, considering geometric information and complex visual information of cross recommender models using varied viewing angles. In summary, the geometrically based modalities and the visual features are learned using geometric deep local convolutional neural networks with a local fusing technique based on the restricted Boltzmann machine (LRBM) and a kernel-based density estimator for the selection of relevant recommender systems. The proposed approach uses an adopted local sensitivity based deep convolutional belief network (LDCBN) and a local-sensitivity convolutional neural network (LCNN) to learn or extract geometric information and visual-based information for cross domain recommender systems. A more discriminative result is obtained by fusing the geometric and the visual information using LRBM. Figure 1 depicts the framework of the proposed model.

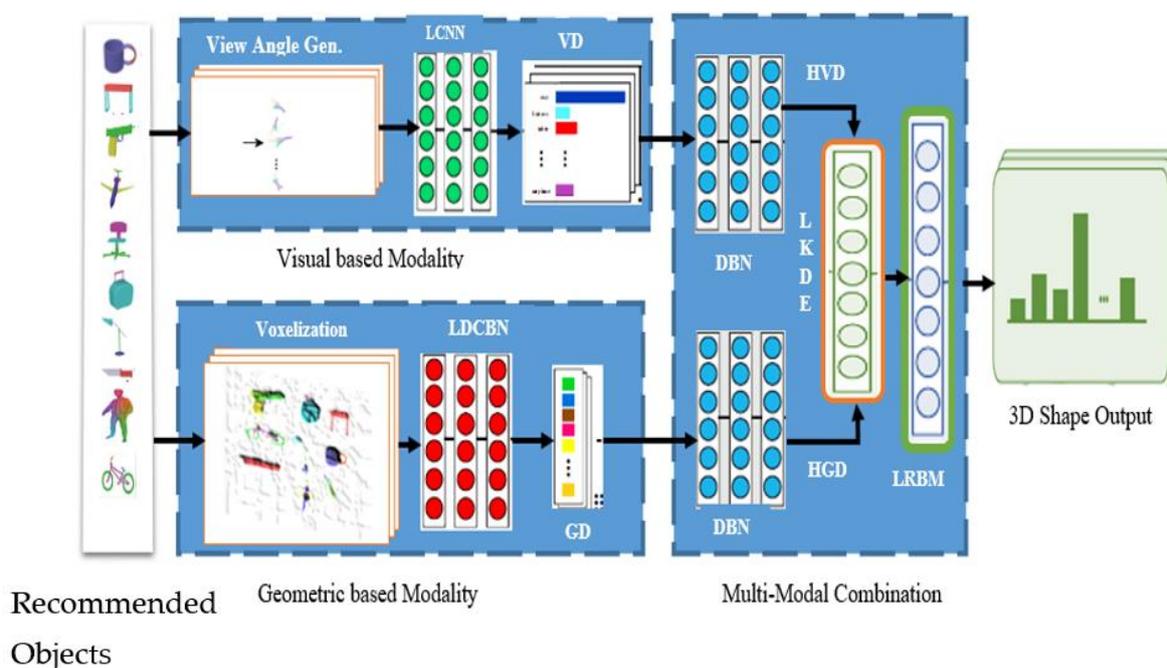


Figure 1. The Framework of the proposed model.

Contributions: The prime contributions made by this work are as follows:

1. We modelled and implemented a novel cross-domain recommendation network capable of learning from complex data representations effectively by incorporating a sparse local sensitivity mechanism into a geometric deep learning algorithm to resolve the problem of sparsity.
2. We introduce a local sensitive adaptor to enforce discriminability by capturing fully the essential local geometric information, which maybe hidden in the structure of recommender systems, for efficient and effective recognition results.

We performed a comprehensive experiment on the proposed model to ascertain its performance on multiple datasets in the sparse and dense settings. We modelled and implemented a novel cross-domain recommendation network capable of learning from complex data representations effectively by incorporating a sparse local sensitivity mechanism into a geometric deep learning algorithm to resolve the problem of sparsity.

2. Review of Related Works

2.1. The Conception of Geometric Deep Learning

Many current studies in various scientific fields concern data belonging to the non-Euclidean domain rather than the Euclidean domain [11]. Examples of these include, but are not limited to, multimedia technology, social networks in computational sciences, functional networks, regulatory networks, sensor networks and meshed surfaces in computer graphics. The structure of geometric data is known to be very complex and non-linear in nature; hence, many scientific researchers are focusing on this area. For instance, deep learning is one of the techniques being exploited and is an important tool in realizing solutions for RSs [11,12]. However, most of the existing approaches have been implemented in the Euclidean domain, with a few of them paying attention to data belonging to the non-Euclidean domain. As a result, there is a need for a representation solution to effectively explore data repositories of non-Euclidean data. This led to a generalized learning algorithm referred to as geometric deep learning (GDL) for the non-Euclidean domains, such as manifolds and graphs.

The aim of the following sections concerning GDL (an umbrella term used to describe emerging technology aiming to structure or generalize techniques in the non-Euclidean domain) is to discuss some of the conceptions of the technique and also to present some available solutions, challenges. Additionally, research directions and how the proposed approach will be utilized in addressing challenges will be discussed. DL is the process of hierarchically learning complex concepts from simple ones in a multi-layer fashion. Examples of such learning approaches include, but are not limited to, neural networks (NN), artificial NN, feedforward NN and Convolution Neural Network (CNN), which have the ability to train large sets of data and have resulted in significant breakthroughs in multimedia applications, computer vision, and speech recognition.

In a broader sense, GDL issues can be categorized into two groups. In the first group, we intend to associate characteristics with the structure of the data, and in the second group, we analyze the functions based on which a given non-Euclidean domain is defined [11]. The two groups are interrelated in the sense that because particular information is transmitted and understanding the properties of functions defined on a domain is essential. The same can also be said about the properties forced on the function by the structure of the domain. As an example of the first group of problems, an assumption could be made that a group of data points are given, where low-dimensional constructions are entrenched into a high-dimensional Euclidean space. If that is the case, then recuperating that low-dimensional construction from the high dimensional space is mostly termed as nonlinear dimensionality reduction or manifold learning [12]. This is an instance of un-supervised learning, and most such methods consist of two phases. The first phase begins with the structuring of a representation solution based on the local affinity of the data point, and in the second phase, the data points are entrenched into a low-dimensional feature space with the aim of preserving originality.

Some examples of studies examining manifold learning include, but are not limited to, the ones discussed in [13–15] that attempt to preserve global geometric information such as the geodesic coverage of the graph. Other studies examine decomposing the graphs into minute sub-graphs referred to as graph-lets or motifs, rather than embedding the vertices [16–19]. Graphs or manifolds are well-known and utilized for constructing representation solutions for social network analysis [20,21], natural language processing [22–24] and many other applications by capturing local geometric information.

2.2. Deep Learning Technology in the Euclidean Domain

To date, most machine-learning-based techniques are used to address issues bordering on the Euclidean domain. Assume multi-dimensional data in the Euclidean domain where $\mathcal{U} = [0, 1]^d \subset \mathbb{R}^d$, based on which functions such as image analysis applications are explicated $f \in L^2(\mathcal{U})$. In a general supervised learning domain, an unknown function derived on a training set may be considered as $L^2(\mathcal{U}) \rightarrow y$. Such a function is defined in Formula (1):

$$\left\{ f \in L^2(\mathcal{U}), y_i = y(f_i) \right\}_{i \in \mathcal{I}'} \quad (1)$$

where y represents the target space, \mathcal{U} denotes an arbitrary domain, and the square integrable function on \mathcal{U} is represented by f . In many shape recognition systems, the y could be replaced by multi-dimensional simplex, representing class probabilities, $p(y|x)$. Regression analysis could also be defined as $y = R^m$. There are several other assumptions that could be considered for unknown functions in many speech analysis and computer vision related tasks. These assumptions may be subjugated by structures of the CNN as can be seen in Formula (2):

$$\mathcal{I}_v f(x) = f(x - v), \quad x, v \in \mathcal{U}, \quad (2)$$

where \mathcal{I}_v represents a translation operator. The function y is assumed to be either equivalent with regards to translation or invariant, depending on the task at hand.

2.3. Data Representation in the Non-Euclidean Domain

Functions or problems defined based on the non-Euclidean domain can be further subdivided into sub-classes [11,25,26]. Firstly, we can consider situations consisting of multiple domains, such as identifying similarities and correspondences among items/user recommendations with computer vision and graphics applications. Functions defined in these settings are similar to CNNs based on the spatial domain and are more appropriate. Secondly, situations with fixed domains could also be considered. An example could be the identification and prediction of global positioning of persons belonging to a social network based on their previous behaviors. In this example, the domain, which is the social network, is considered to be fixed. Functions defined in the domain are similar to CNNs based on spectral domains.

The focus of this study is to examine functions based on the non-Euclidean domain, especially those that attempt to generalize or extend their solutions by exploiting their implementations with CNN techniques. Graphs and manifolds are examples of data in the non-Euclidean domain, and several studies [27–29] have attempted to generalize the application of NNs to graphs by a combination of deep learning models. The foremost application of a CNN to graphs in the spectral domain was proposed by Bruna [30]. However, the approach faced computational difficulties, which were resolved in [29,31,32], resulting in some state-of-the-art outcomes. The first application of a CNN on a mesh surface in the spatial domain centered on local intrinsic patches was proposed in [33]. These applications were found to have resulted in an up-to-the-minute performance on deformable recommender systems, whereas other approaches were based on intrinsic patches on point clouds and general graphs [11].

There has been an explosive increase in the application of DL to manifolds and graphs in recent times. This can be seen in the numerous such techniques that have attempted to address issues pertaining to recommender systems and biochemistry.

By way summary, the main objective of the current DL techniques, is to extend the application of the technique in the non-Euclidean domain. The non-Euclidean domain is represented by two (2) structures considered prototypically as graphs and manifolds. Graphs such as social networks consist of structured network data made up of edges and nodes. Manifolds on the other hand describe geometric-like applications of recommender systems like spatial coordinates based on the surface of an entity brought forth by a scan

(LiDAR). These structures mathematically represent graph theory and differential geometry, respectively, and share numerous known characteristics.

2.4. Convolutional Neural Network (CNN)

A CNN is a DL technique used in object recognition and is developed to train neural networks in the same manner as human beings learn. CNNs are particularly suitable for training large datasets and have achieved a lot of recognition in image classification and a number of computer vision tasks, such as object detection, classification and scene detection. Hence it would not be out of place to extend the technique to recommendations [34]. CNNs are more stable and possess stationarity to local translation and consist of many convolutional layers, which are of the form $f = C\tau(g)$ being applied to $g(x) = (g_1(x), \dots, g_p(x))$, which is a p -dimensional input with bank filters $\tau = (\gamma_i, l), i = 1, \dots, q, l = 1, \dots, p$ and element-wise non-linearity ξ :

$$f_i(x) = \xi \left(\sum_{l=1}^p (g_l * \gamma_{i,l})(x) \right), \tag{3}$$

where $f_i(x)$ results in a q -dimensional outcome ($f(x) = f_1(x), \dots, f_q(x)$) denoting feature maps, and the standard convolution is represented by Equation (4).

$$(g * \gamma)(x) = \int_{\Omega} g(x - x') \gamma(x') dx' \tag{4}$$

The filters τ have a compact spatial support. In addition, a pooling layer or down-sampling $f = P(g)$ may be utilized as stated in Equation (5):

$$f_i(x) = P((g_i(x') : x' \in N(x))), i = 1, \dots, q, \tag{5}$$

where the nearness neighbor around x is represented by $N(x) \subset \Omega$, and the permutation invariant function, such as L_p -norm, is denoted by P . A CNN may be engineered by combining several convolutional and pooling layers selectively, resulting in a generic resultant hierarchical representation as indicated in Equation (6):

$$U_{\ominus}(g) = (C\tau^K \dots P \dots \circ C\tau^{(2)} \circ C\tau^{(1)})(g), \tag{6}$$

where the filters coefficients referred to as the hyper-vector of the network regularizations are denoted by $\ominus = (\tau^1, \dots, \tau^K)$. The models of this nature consisting of multiple layers are said to be deep. Although this notion is ambiguous, one can still find CNN-based techniques utilizing this approach with as many as one hundred layers. For supervised learning tasks, one could find parameters of a CNN by minimizing the cost function on the training set.

Appreciating the nature of the optimization function and obtaining a solution through the adaptation of effective strategies has led to an area of research known as deep learning [35–38]. Another factor in the successful implementation of CNN is that their learning results in the avoidance of the curse of dimensionality due to their learning complexity. Due to NP hard solutions or shift invariance, the convolution operation in the Euclidean domain is seen as passing a template to each other and keeping track of the points for each template. This makes it possible for us to learn specialized features from samples to achieve a ground-breaking results recognition performance in various applications such as image segmentation, detection, classification and annotation [39]. Now taking into consideration the construction of CNNs in the non-Euclidean domain, we will investigate the two typical geometric data: graphs and manifolds. Graphs consist of edges and nodes, whereas manifolds are mostly used to designate geometric 3D shapes, particularly in the spatial domain.

3. Proposed Method

Many large companies offer diversified products or services to customers. For instance, Google provides customers with mobile applications, web searches, and news services, and books, electronics, and clothes can be bought from Amazon. Single-domain recommender systems only focus on one domain while ignoring user interests in other domains, which also exacerbates sparsity and cold-start problems [40,41]. A cross-domain recommender system, which assists target domain recommendation with the knowledge learned from source domains, provides a desirable solution for these problems. One of the most widely studied topics in cross-domain recommendation is transfer learning, which aims to improve learning tasks in one domain by using knowledge transferred from other domains [41]. Deep learning is well suited to transfer learning as it learns high-level abstractions that disentangle the variations of different domains. Several existing works, as reviewed in [41], demonstrate the efficacy of deep learning in catching the generalizations and differences across different domains and generating better recommendations on cross-domain platforms. Therefore, this is a promising but largely underexplored area where more studies are expected, hence the need for this study.

The proposed discriminative-based geometric deep learning (D-GDL) for cross-domain recommender systems, particularly in the non-Euclidean domain, is presented in this section. Most of the objects that are recommended in recent times are 3D. A 3D object means a three-dimensional shape that can be defined as a solid figure with properties of depth (length), width and height. For example, virtual shopping malls, which integrate the advantages of 3D virtual environments and traditional online websites, accommodate users in a life-like shopping environment [42]. In these environments, users can effectively interact with virtual products by viewing them from different angles, zooming in and out, touching the surface and even trying them on. The virtual product experiences enable users to form a direct, intuitive and concrete understanding of the quality and performance of products and make better informed purchase decisions [43]. Examples of 3D datasets are the Flixter, Netflix and CiteUlike datasets. Hence, the recommender system must be modeled in a way to efficiently support such objects. The main objective of the approach is to introduce a general framework for geometric DL on non-Euclidean domain for cross-domain recommender systems. In this approach, a local representation learning is introduced into the structure of deep learning techniques to effectively capture the local geometric and visual information from the structure of the recommended 3D objects, which serves as the basis for 3D and recommended searches. A kernel-based technique is also adopted and introduced into the structure of the DL technique to map the complex structure of recommended objects into high dimensional feature space to achieve an effective recognition result. The kernels serve as a weighing function in building a high dimensional feature space, which makes it more robust to geometric noise and results in a good computation outcome.

The proposed method could achieve an optimal dictionary to further enhance the power of discriminability of recommended objects and enables the realization of geometric DL representation of the 3D non-linear features. The D-GDL extracts visual and geometric features from recommended objects using LCNN and LDCBN models for optimal and efficient recognition performance. These models are pre-trained by depth image generation and voxelization instead of using down-sampling techniques. This is achieved by passing the extracted features through an improved kernel density estimator. With the extraction of geometric features, recommended objects are transformed into voxel-like forms from mesh and are almost like the original recommended object, which reduces the need for down-sampling. The depth images are used as input features in the visual feature extraction process, which also does not require down-sampling due to the multi-image representation of recommended objects by way of projection. Most of the conversional approaches are based on the Euclidean domain rather than the non-Euclidean domain. Even those that are based on the non-Euclidean domain do not fully capture the local discriminative and geometric information, which is very important for an optimal recognition result of recommended objects. Kernel-based models achieve the desired result when the training

samples are sufficient. Therefore, due to the complex nature of recommended objects, it is necessary to work with some very large datasets since deep learning algorithms perform well with such datasets. These give optimal outcomes in the capture of nonlinear information for recognition of recommended objects.

D-GDL for cross-domain recommender systems, particularly in the non-Euclidean domain, is proposed to address the challenges of recognition accuracy, efficiency and scalability, geometric information recognition and discriminability of recommended objects in the Euclidean domain. This technique is expected to demonstrate an outstanding efficient recognition performance, taking into account the linear structure existing amongst recommended objects with similar relationship category. To date, simulated results across a number of datasets indicate a significant recognition accuracy result for recommended objects content recognition. The objective function for our proposed approach, which is based on a generalized CNNs for non-Euclidean domain, is as stated in Equation (7):

$$(f * g)(x_i^j) = \sum_{j=1}^J \sum_{i=1}^{N_j} (g_i^j) D_j(x_i^j) f + \sum_{j=1}^J \sum_{i=1}^{N_j} \|B_i^j \odot x_i^j\|_2^2, \tag{7}$$

$$s.t. \mathbf{1}^T x_i^j = 1 \quad \forall i = 1, \dots, N_j, j = 1, \dots, J$$

where x_i^j is any point i on the manifold belonging to the group j , and the patch operator is denoted by $D_j(x_i^j) f$, which is mapped into high-dimensional feature space. The local sensitivity adaptor is represented by B_i^j , the dimensionality of the extracted patch and the element-wise multiplicative operator is represented by \odot . The local sensitive adaptor is used to enforce discriminability by helping in the capturing of essential local geometric information, which may be hidden in the structure of recommended objects, for efficient and effective recognition results.

LDCBN is an essential deep learning tool utilized to learn prevailing discriminative features automatically as it belongs to a DL network and is unsupervised. Recommended objects are very complex in nature topologically and have geometric variations. It is very challenging to analyze them directly so they are firstly discretized into a regular grid using its voxelized form as the input recommended objects to extract the geometric descriptor using the LDCBN. Voxelization is the process of transforming recommended 3D objects into voxel representations from the mesh form, which could be likened to the original form of the recommended 3D object. This does not only detail the information about the model's surface but also the internal representations of the recommended objects. This discretizes and lessens the complexity of the recommended model because of the spatial relationship properties of recommended objects, leading to a substantial amount of geometric information. This makes it simple for intrinsic geometric features to be extracted with LDCBN. A recommended 3D matrix representation is used to realize the geometric representations of the 3D shape with possibility dissemination. If a single voxel is placed into the 3D mesh, its analogous matrix is then set to 1 otherwise zero. The 3D matrix then serves as input in extracting the geometric descriptors. In our approach, the local-sensitivity adaptor is introduced into the structure of the original convolutional deep belief Network (CDBN) as an extension to support recommended objects.

Deep belief network is a suitable probabilistic principle for modelling joint probabilistic distribution over pixels and labels. Nonetheless, it is a challenging task to extend the model from 2D pixel data to 3D voxel data. This is because the volume size of a 3D voxel will be much greater in size than an original image. In addition, a fully connected DBN involves a lot of parameters, which makes the training of the model very challenging. As a result, we propose the LDCBN by weight sharing to lessen the number of parameters used. Again, unlike the conventional DL techniques, the pooling layers are ignored in our approach due to uncertainties that may arise when features are being generated. The local-sensitivity convolutional layer's energy in the proposed model is as in Equation (8):

$$E(f, g) = - \sum_{j=1}^J \sum_{i=1}^{N_j} ((g_i^j) D_j(x_i^j) f + c^j g_i^j) - \sum_l b_l f_l + \sum_{j=1}^J \sum_{i=1}^{N_j} \|B_i^j \odot x_i^j\|_2^2, \quad (8)$$

where x_i^j is any point (hidden unit) i on the manifold belonging to the group or feature space j , the patch operator is denoted by $D_j(x_i^j)f$, which is mapped into high dimensional feature space, and the index of the visible units is denoted by l . The local sensitivity adaptor is represented by B_i^j , the dimensionality of the extracted patch and the element-wise multiplicative operator is represented by \odot . Each hidden unit in feature space j is denoted by g_i^j , the visible unit in the 3D voxel input is denoted by f_l , the convolutional filter is represented by D_j , c^j and b_l represents bias terms in hidden and visible units, respectively. The local sensitive adaptor is used to enforce discriminability by helping in the capturing of essential local geometric information, which may be hiding in the structure of recommended objects, for efficient and effective recognition results. The proposed approach is modeled around the theories of [44].

We could set a voxel grid of $45 \times 45 \times 45$ to a 3D shape with three extra padding cells in all dimensions to lessen the boundaries of the convolution objects. The labels are then put forward to ascertain the standards of the variables. There are two (2) steps followed in the training process of the proposed 3D LDCBN. These are layer by layer before-training and generative refined-tuning processes. The initial four (4) layers are separately trained during the before-training phase with a divergence algorithm described in [45], and a fast persistence divergence is used to train the topmost layer [46]. The underlining hidden activations serve as inputs into the next layer once the bottom layer is learned with their weights settled. In the refined-tuning step, methods from past studies such as the wake-sleep algorithm were adopted, where the wake phase utilizes the activations for learning of positive signals and by propagating the data bottom-up. For the sleep phase, a tenacious chain on the uppermost layer is maintained, and the data top-down section is propagated to assemble the deleterious learning signal. This refined-tuning technique copycats the generation and recognition performance of the model and functions well in principle. As long as the weights of the entire networks have been learned, forward computation is utilized to engender the geometric descriptor applying the input data of voxelization.

In analyzing recommended objects from the angle of view-based descriptors, the recommended model is transformed into varied images through various perspectives. In principle, the images encompass as much information as possible from the recommended model. During the course of generating our visual descriptor, 30 directions were used for the translation of the recommended objects and further extract the visual features with the proposed LCNN technique. The two (2) steps involved in the proposed LCNN are the before-treatment of the recommended model and the assemblage of image depths. In the first step, the highest polar distance is measured from the origin point to the surface of the point after setting the origin point of the recommended object model. In the second stage, depth images are assembled. From 30 vertices of a regular 3D shape, a type of 2D image is rendered with its center as the origin.

The regular object is then revolved 15 times in order to make the features vigorous against gyration. The gyration view is carefully set to ensure even distribution of the cameras so that different view angles could be captured for the recommended 3D model. The 30 vertices of the object generate considerable data able to provide significant information but at a large computational cost. Each image is resized by rendering into 124×124 to remove with the out-layers and make the data more compact, which further augments the feature learning with LCNN since it has an input range of 0 to 1, each dimension is normalized using (0,1).

The images obtained from the descriptions thus far include suitable visual information of the recommended 3D model and are used to extract visual features from recommended objects for each image. After the features are extracted, we adopt the local-sensitivity density estimator (LKDE) technique to reduce feature redundancy by selecting salient

features and make our approach more discriminative. Our LCNN approach is as depicted in Figure 1 and the layer-by-layer approach is as stated in Equation (9):

$$F_l = \text{pool}(\text{sigmoid}(D_l * F_{l-1} + b_l)) + g_l, \quad (9)$$

where l denotes the layers, the bias parameter of the l -th is represented by b_l , LKDE is represented by g_l and D_l represents the kernel of the convolution with the initial feature map being F_0 in the 2D images. The threshold function used here is the sigmoid function and the pooling operator results in an activation of every nearest neighbor being considered. The optimal pooling operator is seen as the pooling expression that derives the optimal activation in the nearest neighbor with an invariance. The weight of the approach is determined from back-propagation with input depth images of recommended objects and their corresponding label information. LCNN features are derived for every depth image after fully training the model.

In this section, we propose a novel feature fusion approach that combines extracted multimodal data. The visual and geometric descriptors denote visual and spatial characteristics, respectively, of recommended objects. Hence, for recommended objects, the two descriptors complement each other and both geometric and visual feature information are complex and extremely non-linear in nature. In our approach, to analogously associate the two, high-level descriptor information is captured from feature descriptors. This is done to visualize or bring to prominence, the high-level features with the properties of the recommended model. This means the model becomes more discriminative to recommended 3D models rather than with some specific modalities. Consequently, both the high-level visual and geometric descriptors are extracted using DBNs just as described in [44]. This is done through a greedy learning, layer-by-layer bottom-up approach, which is acknowledged to be very effective. LRBM is then employed to analogously associate both descriptor modalities in learning multi-modal feature fusion for recommended objects. The function of LRBM is as stated in Equation (10):

$$E(f, g; \theta) = - \sum_{i=1}^F \sum_{j=1}^G ((g_i) D_{ij} (x_i^j) f_j - \sum_{i=1}^F b_i f_i - \sum_{j=1}^G a_j f_j + \sum_{j=1}^J \sum_{i=1}^{N_j} \|B_i^j \odot x_i^j\|_2^2), \quad (10)$$

where a_j and b_i both denote biases, F and G represent numbers of visible and hidden layers, respectively, and the interactions between the hidden g_i and the visible f_i layers are denoted by D_{ij} , with the 3D model parameter being denoted by $\theta = (w, a, b)$.

The LRBM is used to inherently combine the visual relationships of the 3D shape and the spatial characteristics of the recommended 3D model, which makes it more robust and discriminative due to the introduction of the local-sensitivity adaptor across the models. In addition, the model effectively captures relevant local geometric features that may be hidden with the structure of the complex non-linear recommended objects. This demonstrates its more robust features, efficiency and accurate recognition performance. The experimental results are discussed in detail in the next section of the paper.

4. Experimental Setup and Results Analysis

In this section, extensive experiments were conducted on three different real-world datasets to demonstrate the recommendation effectiveness of our proposed D-GDL approach in comparison with some state-of-the-art baseline approaches.

4.1. Database Selection and Experiment Evaluation

To verify the effectiveness of the proposed D-GDL method, the performance of the discriminative geometric deep learning approach for cross-domain recommender systems is evaluated by comparison with the following six baseline approaches: A hinge-loss based codebook transfer for cross-domain recommendation with non-overlapping data (TCH) [2], A cross-domain recommendation approach based on topic modeling and ontology (TMO) [47], personalized transfer of user preferences for cross-domain recommendation (PTUPCDR) [48], cross-domain recommendation to cold-start users via variational

information bottleneck (CDRIB) [49], deep learning-based matrix factorization (DLMF) [50] and heterogeneous multidomain recommender system through adversarial learning (HM-Rec) [51]. All the experimental results were obtained by twenty-fold cross-validation. The training samples were randomly selected from various video datasets, and the rest set as test samples. In our experiments, four datasets were utilized from different real-world settings as follows: Flixter, Netflix, CiteULike. These datasets were chosen taking into consideration their different degrees of sparsity and scale to mimic the different practical traits. Flixter is a social network platform, where users can rate movies and users can also add friends or other users to create a social network. Two datasets, namely CiteULike-t and CiteULike-a, were chosen from CiteULike [52]. CiteULike-t was selected independently of CiteULike-a, which is mostly from [53]. The Netflix dataset also consists of movie titles and ratings from Netflix challenge datasets. The detailed statistics of the datasets are as presented in Table 1.

Table 1. Statistics of the selected datasets.

	Netflix	Flixster	CiteULike-a	CiteULike-t
Users	407,261	1,049,445	5551	7947
Items	9228	492,359	16,980	25,975
Ratings	15,348,808	8,238,597	0.22%	0.07%

As in the case of [34], users with fewer than three articles were excluded. It can be seen in Table 1 that CiteULike-t has many more users and items than CiteULike-a. Additionally, CiteULike-t is comparatively sparser as only 0.07% of its user-item matrix entries contain rating compared with 0.22% for CiteULike-a. To maintain consistency with the implicit feedback mechanisms based on the CiteULike datasets, only positive ratings of 5 were extracted for testing and training with the Netflix dataset, and users with fewer than three positive ratings were not considered either.

The preprocessing text information, which was the content of the items including the plots of the movies, was extracted from the abstracts and the titles of the articles. The top S discriminative words were selected according to their tf-idf values after eliminating the stop words. The tf-idf values selected to represent the vocabularies S are respectively 8000, 20,000, 20,000 and 15,000 for the four datasets. Again, it can be seen in Table 1 that the user-item-rating matrices of the CiteULike datasets are much sparser than those of Netflix, and that the Netflix matrices are much sparser than those of Flixster. Hence, CiteULike datasets can better evaluate the performance with high sparsity of the proposed model. The datasets used have evolved over the years; however, there is no substantial difference between the content and structure of these datasets.

Furthermore, to ascertain the effectiveness of the proposed approach, we also considered cross-domain datasets such as the Amazon and the Imhonet datasets. The Imhonet dataset was sourced from an online Russian recommender site at Imhonet.ru. These give users the opportunity to review and rate a large number of items, some of which are 3D and range across several domains including movies, architectural monuments, mobile phones and books. The dataset also contains some elements of social networks with blogs, friendship networks and comments. It is a unique dataset and is available across many cross-domain recommendations because it gives explicit user feedback (ratings) with varied domains. The Imhonet dataset considered against this research contains movies, books, games and perfumes and it consists of the complete set of user ratings across the four domains. The statistics of the Imhonet dataset are shown in Table 2 below. The last dataset considered is the Amazon dataset. It contains two movie and book data items. It consists of users and their ratings for items on a scale from 0 to 5 with varying degrees of sparsity and scales. All the experimental results were realized by tenfold cross validation in which training samples were selected randomly and the remaining data used as testing samples for the various datasets. The first dataset has Amazon books as its source domain and Amazon movies as its target domain. The second dataset uses Amazon movies as its

source and Amazon books as its target domain. To better evaluate the proposed model, sparse users with ratings for fewer than 20 items were removed from different domains. The statistics of the two datasets are shown in Table 3.

Table 2. Statistics of the Imhonet datasets used.

Statistics	Movies	Books	Games	Perfumes
Users	426,897	3,622,448	72,307	19,717
Items	90,793	167,384	12,768	3640
Density	0.00073	0.00022	0.00140	0.00350
Number of records	28,281,946	13,438,520	1,324,945	253,948
Average no. of ratings per user	66.30	37.0771	18.2339	12.8796
Average no. of ratings per item	311.4992	80.2856	103.7708	69.7659

Table 3. Statistics of the Amazon Datasets Used.

Statistics	Movie	Book
Users	9043	38,032
Items	30,279	105,651
Ratings	1,264,244	3,637,313

All the rating information of entities in the target domain are randomly removed and are regarded as cross-domain cold start users for making recommendations. The predictive ratings of the candidate articles are then ranked and the first M items are recommended to the target user according to the formulation of Equation (11).

4.2. Evaluation Scheme and Settings

Similar to [35,36], P items related to each user were randomly chosen to form a training set with the rest used as the test sample for each dataset. To conduct a comparative analysis of the models under both dense and sparse settings, P was set to 1 and 10 in our experiments, and for each value of P , the process was repeated randomly ten times with different training sets with their average values reported. As has been discussed extensively in [25,34,54–57], recall was utilized as the measure for performance since the rating information implicitly is obtained as feedback [9,19]. In situations where zero entries are obtained, it suggests that either the users are not aware of the item's existence or are not interested. Moreover, such precision is not appropriate for measuring performance. As in many recommender systems, the predicted ratings of the candidate items are sorted and the top M items are recommended to the target user. As a result, the recall M for each user is then measured as:

$$\text{Recall} = \frac{\text{Number of items that the user likes among the Top M}}{\text{total no of items that the user likes}} \quad (11)$$

Finally, the average recall value over all users is reported. The cutoff point of the mean average precision (mAP) is also set at 500 for each user.

In the experiments, the optimal parameters were determined for all the baseline approaches (TCH, TMO, PTUPCDR, CDRIB, DLMMF and GDL) using a validation set. The parameter was set to 10 for all the base models. With reference to TCH, TMO, PTUPCDR and CDRIB, the optimal performance of these models is achieved using a two-layer architecture. For DLMMF and GDL, the values of a , b and K were set to 1, 0.01 and 50, respectively. A grid search was performed on the parameters with the training dataset split using ten-fold cross validation. For the proposed D-GDL approach, the values of w , a , and b were set to 1, 0.01, and 0.01, respectively, for all experiments. With regards to the Amazon and the Imhonet datasets, 80% of the users were chosen to represent the training set while the remaining 20% used as the test user set with ten-fold user stratified cross-validation settings. The best set of parameters was chosen using 20% of users for validation and 80% as their ratings

from the training set. This procedure was repeated 10 times with the average performance of the algorithm reported.

4.3. Experimental Results and Discussions

Figures A1–A8 in the Appendix A show the comparative results of the proposed D-GDL approach and the baseline approaches against CiteULike-a, CiteULike-t, Netflix and Flixster datasets, respectively, in both sparse and dense settings. The experiments were conducted under sparse ($p = 1$) and dense ($p = 10$) settings. Figures A1–A4 indicate the results in the sparse settings and those of the dense settings are shown in Figures A5–A8. It can be seen from the results that, the proposed D-GDL approach gave the best results in all experiments.

It can also be seen from Figures A1–A4 that, PTUPCDR is the best comparative approach, which came close to the results of the proposed method and was superior to DLMF, TCH, TMO, CDRIB and GDL in all experiments despite the fact that all the comparative approaches have deep architecture. GDL, DLMF and DMF outperforms DeepMusic in the sparse settings for all datasets as shown in Figures A1–A4. The poor performance of DeepMusic is due to overfitting and a lack of ratings. DeepMusic performs better than DMF and GDL in the dense settings for all datasets as can be seen in Figures A5–A8. However, DeepMusic is outperformed by DLMF, CDL, CVAE and the proposed D-GDL for all the four datasets. The proposed approach gave the best performance results compared with all the comparative approaches for all datasets. Furthermore, the performance of the models increases with an increase in K values both in the sparse and dense settings for all algorithms. However, as can be observed in the figures, a larger K has a greater influence in the dense settings than the sparse settings. This is because much guidance inference neural networks for variational inference are offered for dense ratings than for sparse counterparts. Thus, it requires much larger representation abilities to learn, which is the hallmark of any deep learning-based approach.

Table 4 indicates the mAP results for the four datasets. It can be seen that the proposed D-GDL approach gave the highest result against the other baseline approaches for all datasets.

Table 4. mAP Results for the Four Datasets.

	CiteULike-a	CiteULike-t	Netflix	Flixster
D-GDL	0.0714	0.0651	0.0531	0.0492
CVAE	0.0662	0.0545	0.0454	0.0334
CDL	0.0526	0.0465	0.0326	0.0375
DLMF	0.0312	0.0267	0.0189	0.0249
DMF	0.0159	0.0176	0.0167	0.0183
GDL	0.0274	0.0104	0.0158	0.0273
DeepMusic	0.0160	0.0103	0.0187	0.0157

Figure A9 indicates variations in the recall rate with varying values for both Amazon books and movies datasets. The black lines represent the Amazon movie domain as the source domain and the books domain as the target domain, whilst the red line indicates the Amazon books domain being used as the source with Amazon movies used as the target domain. When the value 0 is chosen, it means the user preference of the target domain is solely dependent on the user preference without the use of the source domain. When 1 is chosen, it means that, the user preference depends entirely on the knowledge of the source domain without taking into consideration the user preference information of the target domain. Therefore, a compromise is chosen between the target and the source domain because the user behavior characteristics of the target domain may not be well utilized. As seen in Figure A9, the highest recall value when the movies domain is used as the source is 0.7 and is 0.6 when the books domain is used as the source.

The evaluation results of the proposed D-GDL approach against the baseline approaches for the Amazon dataset are shown in Figure A10. The Amazon book dataset was used in this case as the source domain with the Amazon Movie dataset used as the target domain. As can be seen from the figure, the CVAE approach achieves the best performance results amongst the baseline approaches, followed closely by the CDL technique. This means that the structure of CVAE learns a better representation solution compared with the other approaches. The proposed D-GDL had optimal performance against all the baseline approaches. This could be due to the geometric deep learning structure of the proposed approach considering the 3D items that the datasets may have contained. The proposed D-GDL approach is able to outperform all the comparative approaches by a margin of approximately 9.8% to 11.5% when a generative network of target users in a probabilistic propagative latent variable is modelled.

The evaluation results of the proposed D-GDL approach against the baseline approaches for the Amazon dataset are shown in Figure A11. The Amazon Movie dataset was used in this case as the source domain, with the Amazon Book dataset used as the target domain. As can be seen from the figure, the CVAE approach achieves the best performance amongst the baseline approaches, followed closely by the CDL technique. This means that the structure of CVAE learns a better representation solution compared with the other approaches. The proposed D-GDL had optimal performance against all the baseline approaches. This could be due to the geometric deep learning structure of the proposed approach considering the 3D items that the datasets may have contained. The proposed D-GDL approach is able to outperform all the comparative approaches by a margin of approximately 7.6% to 9.5% when a generative network of target users in a probabilistic propagative latent variable is modelled. It could also be observed that, when M is less, the performance of CVAE is very close to the proposed approach and widens until it reaches $M = 25$ and widens again from there under sparse data target settings.

The average RMSE of the comparative approaches and the proposed D-GDL approach with two different percentages of training data on the four datasets of Imhonet are presented in Table 5. It can be seen from Table 5 that, the proposed D-GDL approach returned the lowest RMSE values in comparison with the baseline approaches with all the four datasets. This demonstrates the effectiveness of the proposed D-GDL approach due to the incorporation of discriminative information coupled with geometric deep learning into the structure of the proposed D-GDL model. In addition, the results affirm the strength of the proposed approach with RMSE metric. The CVAE, CDL and DLMF also achieved relatively good performance results.

Table 5. Average RMSE of the comparative approaches using different percentages of training data on the four datasets of Imhonet.

Models	Movies		Books		Games		Perfumes	
	70%	95%	70%	95%	70%	95%	70%	95%
DMF	0.5883	0.5354	0.9735	0.9612	0.5895	0.5456	0.5515	0.5425
DLMF	0.5698	0.5215	0.9674	0.9554	0.5745	0.5318		0.5342
DeepMusic	0.5989	0.5673	0.9845	0.9786	0.5992	0.5693	0.5772	0.5682
GDL	0.5982	0.5513	0.9834	0.9765	0.6024	0.5623	0.5579	0.5496
CDL	0.5668	0.5198	0.9664	0.9478	0.5782	0.5298	0.5467	0.5335
CVAE	0.5435	0.5136	0.9579	0.9448	0.5535	0.5212	0.5367	0.5232
D-GDL	0.5212	0.5057	0.9244	0.8979	0.5323	0.5176	0.5136	0.5011

5. Conclusions and Recommendations for Future Work

In this paper, a discriminative geometric deep learning (D-GDL) algorithm for cross-domain recommender systems is proposed. This research was conducted to handle efficiently user recommendations in both single and cross-domain settings, particularly for

objects in the non-Euclidean domain. The proposed D-GDL method achieves optimal dictionary and further enhances the power of discriminability of recommended systems. This thus enables the realization of geometric DL representation of the 3D non-linear features in both sparse and dense settings to effectively deal with cold-start and sparsity issues. The D-GDL extracts visual and geometric features from recommended objects using LCNN and LDCBN models for an optimal and efficient recognition performance in both single and cross-domain recommender systems. The models are pre-trained by depth image generation and voxelization instead of using down-sampling techniques. This is achieved by passing the extracted features through an improved kernel density estimator. With the extraction of geometric features, the recommended objects are transformed into voxel-like forms from mesh similar to the original recommended object, which eliminates the need for down-sampling. The depth images are used as input features in the visual feature extraction process, which also undergoes down-sampling due to the multi-image representation of recommended objects by way of projection. It can be seen from the experimental results that incorporating kernel and geometric deep learning into the structure of representation solutions for recommender systems inherently improves the performance accuracy compared with other baseline deep learning approaches. Furthermore, the proposed method is more computationally effective as a result, which increases its stability. Thus, close-form solutions are derived from the learning network for both sparse and dense settings. We, however, propose to incorporate weighted K-neariness neighbor into the structure of the kernel discriminative sparse representation to enhance the power of classification and recognition accuracy of recommender systems.

Author Contributions: Conceptualization, J.K.A., C.Z. and E.A.M.; methodology, J.K.A. and E.A.M.; software, J.K.A. and Y.C.; validation, J.K.A., C.Z., J.O.-K. and E.A.M.; writing—original draft preparation, J.K.A. and E.A.M.; writing—review and editing, J.K.A., J.O.-K., C.Z., E.A.M. and Y.C.; supervision, C.Z.; funding acquisition, C.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was supported by the key research and development plan (social development) projects BE2016630 and BE2017628 of Jiangsu province, the scientific research project Z201603 of Wuxi health and family planning commission.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

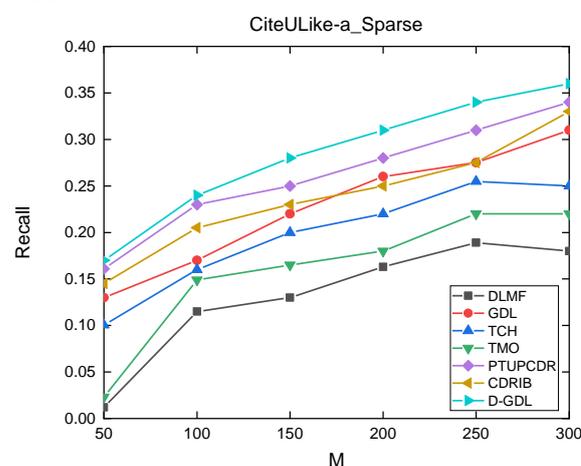


Figure A1. Performance of D-GDL against the comparative approaches based on Recall vs. M for CiteULike-a Dataset in the Sparse Setting.

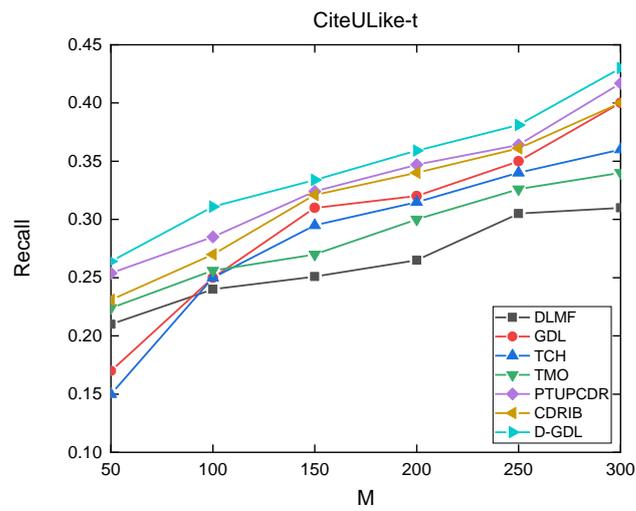


Figure A2. Performance of D-GDL against the comparative approaches based on Recall vs. M for CiteULike-t dataset in the sparse setting.

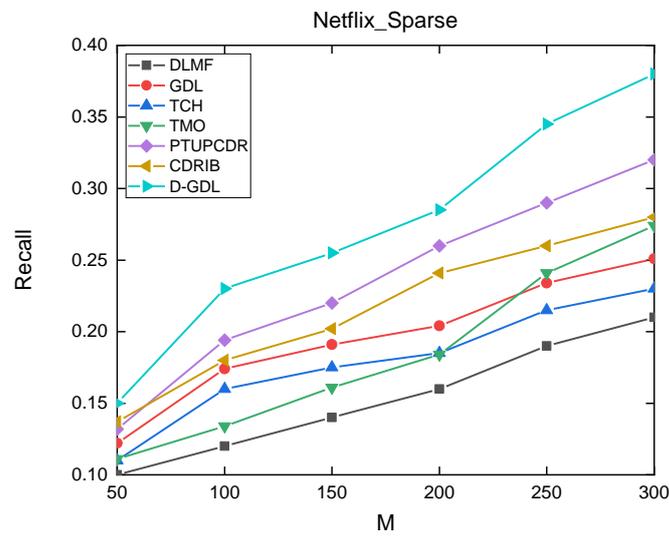


Figure A3. Performance of D-GDL against the comparative approaches based on Recall vs. M for netflix dataset in the sparse setting.

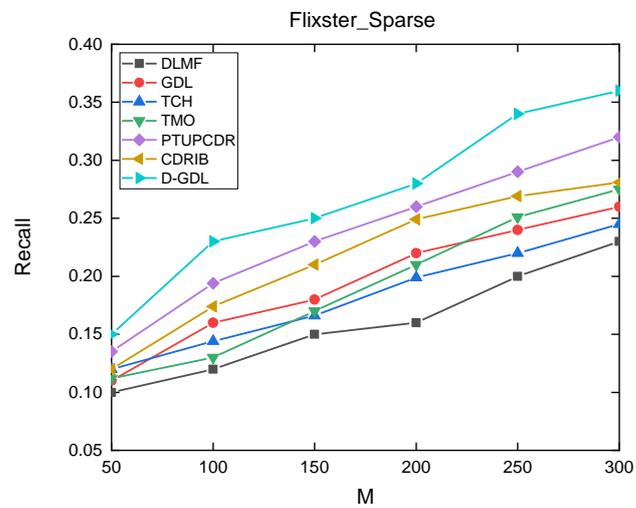


Figure A4. Performance of D-GDL against the comparative approaches based on Recall vs. M for Flixster dataset in the sparse setting.

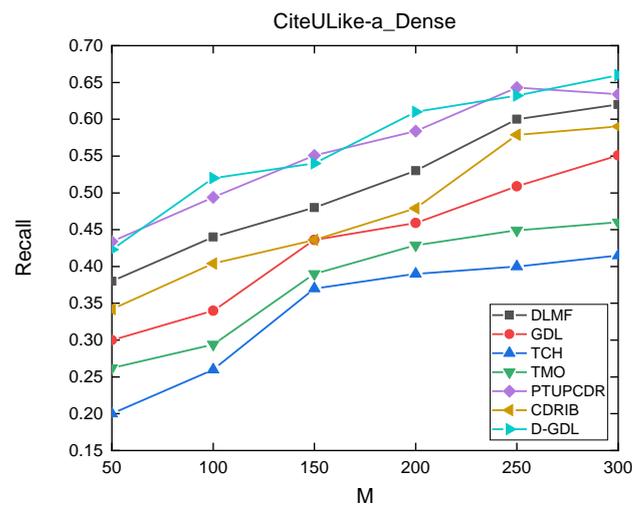


Figure A5. Performance of D-GDL against the comparative approaches based on recall vs. M for CiteULike-a dataset in the dense setting.

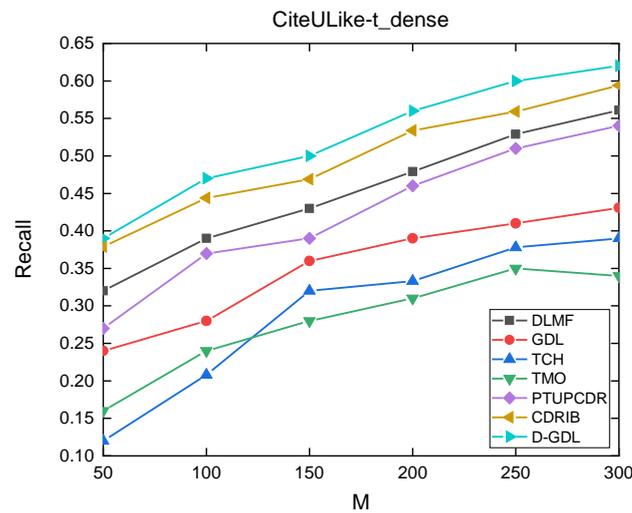


Figure A6. Performance of D-GDL against the comparative approaches based on Recall vs. M for CiteULike-t dataset in the dense setting.

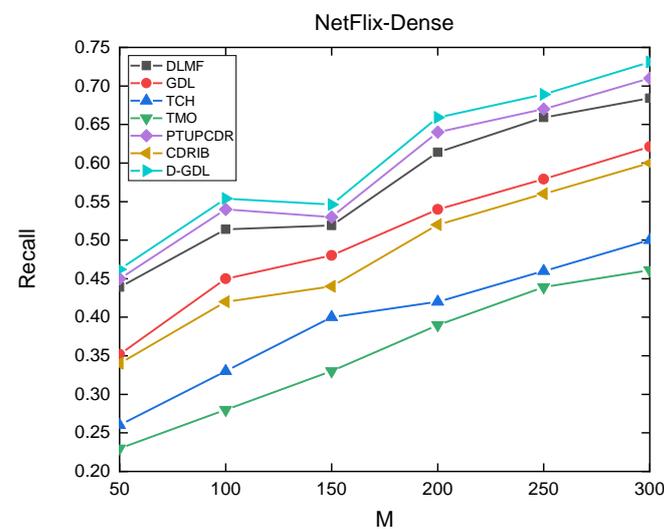


Figure A7. Performance of D-GDL against the comparative approaches based on Recall vs. M for Netflix dataset in the dense setting.

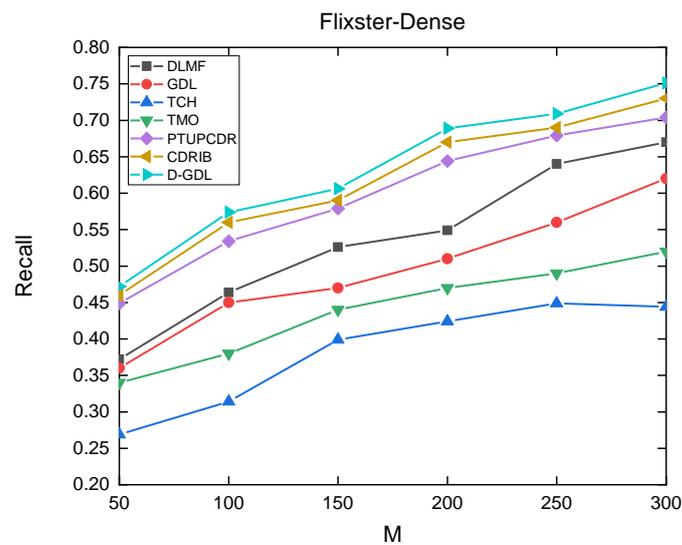


Figure A8. Performance of D-GDL against the comparative approaches based on Recall vs. M for Flixster dataset in the dense setting.

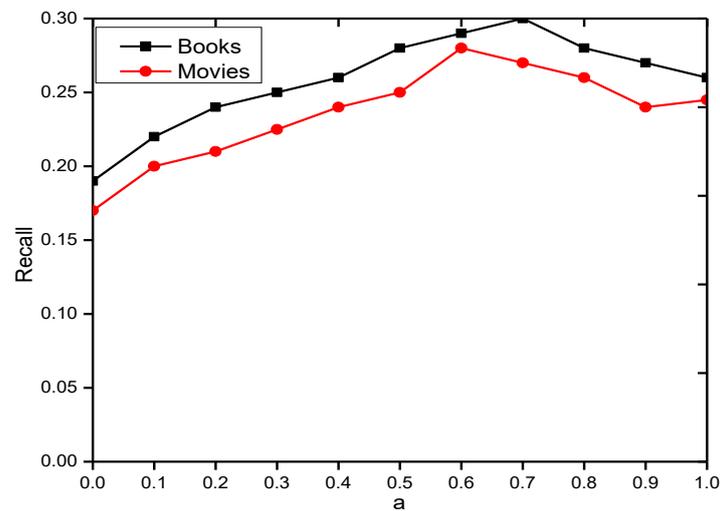


Figure A9. Performance comparison of D-GDL for different values of α based on recall for the Amazon books and movies datasets.

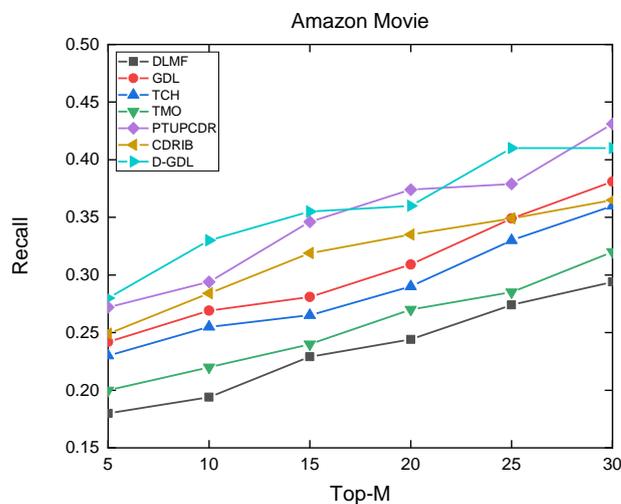


Figure A10. Performance results of the proposed method against the comparative approaches based on Recall for the Amazon movie dataset.

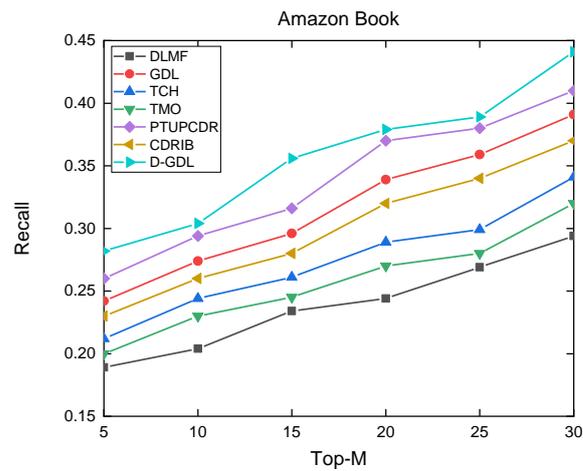


Figure A11. Performance Results of the proposed method against the comparative approaches based on Recall for the Amazon book dataset.

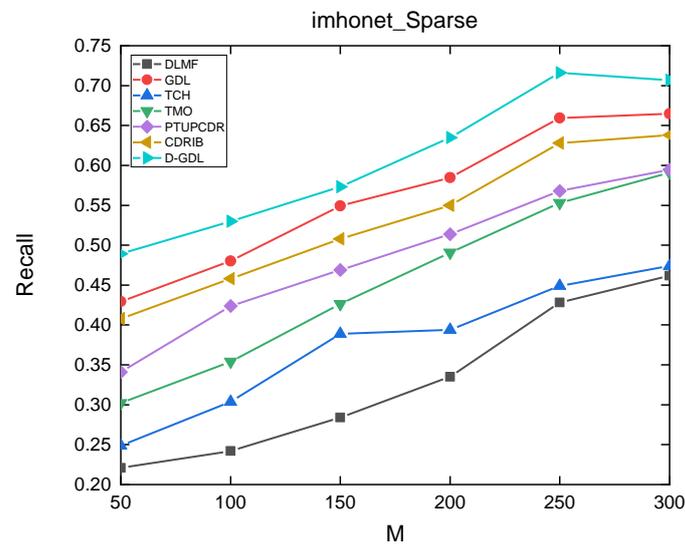


Figure A12. Performance of D-GDL against the comparative approaches based on Recall vs. M for Imhonet dataset in the sparse setting.

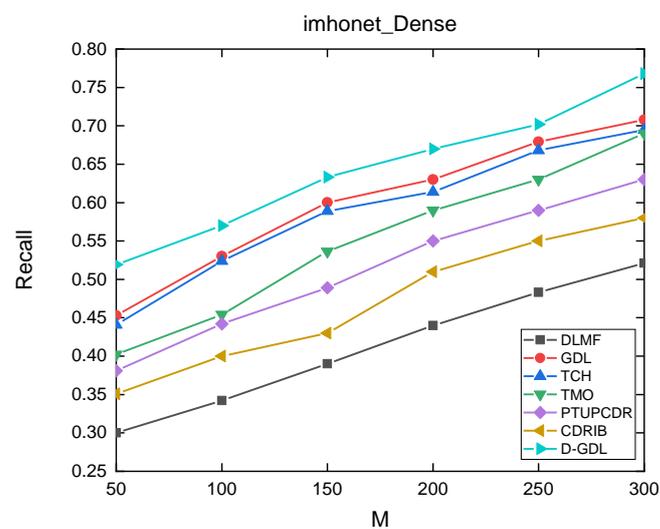


Figure A13. Performance of D-GDL against the comparative approaches based on Recall vs. M for Imhonet dataset in the dense setting.

References

1. Li, X.; She, J. Collaborative variational autoencoder for recommender systems. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017.
2. Veeramachaneni, S.D.; Pujari, A.K.; Padmanabhan, V.; Kumar, V. A hinge-loss based codebook transfer for cross-domain recommendation with non-overlapping data. *Inf. Syst.* **2022**, *107*, 102002. [CrossRef]
3. Wang, Z.; Dong, Q.; Guo, W.; Li, D.; Zhang, J.; Du, W. Geometric imbalanced deep learning with feature scaling and boundary sample mining. *Pattern Recognit.* **2022**, *126*, 108564. [CrossRef]
4. Stärk, H.; Ganea, O.E.; Pattanaik, L.; Barzilay, R.; Jaakkola, T. Equibind: Geometric deep learning for drug binding structure prediction. *arXiv* **2022**, arXiv:2202.05146.
5. Powers, A.; Yu, H.; Suriana, P.; Dror, R. Fragment-Based Ligand Generation Guided by Geometric Deep Learning on Protein-Ligand Structure. *bioRxiv* **2022**. [CrossRef]
6. Monti, F.; Bronstein, M.M.; Bresson, X. Deep geometric matrix completion: A new way for recommender systems. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018.
7. Candès, E.J.; Recht, B. Exact matrix completion via convex optimization. *Found. Comput. Math.* **2009**, *9*, 717–772. [CrossRef]
8. Kalofolias, V.; Bresson, X.; Bronstein, M.; Vandergheynst, P. Matrix completion on graphs. *arXiv* **2014**, arXiv:1408.1717.
9. Levie, R.; Monti, F.; Bresson, X.; Bronstein, M.M. Caylennets: Graph convolutional neural networks with complex rational spectral filters. *IEEE Trans. Signal Process.* **2018**, *67*, 97–109. [CrossRef]
10. Berg, R.V.D.; Kipf, T.N.; Welling, M. Graph convolutional matrix completion. *arXiv* **2017**, arXiv:1706.02263.
11. Bronstein, M.M.; Bruna, J.; LeCun, Y.; Szlam, A.; Vandergheynst, P. Geometric Deep Learning: Going beyond Euclidean data. *IEEE Signal Process. Mag.* **2017**, *34*, 18–42. [CrossRef]
12. Monti, F.; Boscaini, D.; Masci, J.; Rodolà, E.; Svoboda, J.; Bronstein, M.M. Geometric deep learning on graphs and manifolds using mixture model cnns. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
13. Gajamannage, K.; Paffenroth, R.; Bollt, E.M. A nonlinear dimensionality reduction framework using smooth geodesics. *Pattern Recognit.* **2019**, *87*, 226–236. [CrossRef]
14. Luciano, L.; Ben Hamza, A. A global geometric framework for 3D shape retrieval using deep learning. *Comput. Graph.* **2019**, *79*, 14–23. [CrossRef]
15. Arnaudon, A.; Holm, D.D.; Sommer, S. A Geometric Framework for Stochastic Shape Analysis. *Math. Ann.* **2019**, *19*, 653–701. [CrossRef]
16. Chu, Y.; Feng, C.; Guo, C. Social-guided representation learning for images via deep heterogeneous hypergraph embedding. In Proceedings of the 2018 IEEE International Conference on Multimedia and Expo (ICME), San Diego, CA, USA, 23–27 July 2018.
17. Xu, Y. *An Empirical Study of Locally Updated Large-Scale Information Network Embedding (Line)*; UCLA: Los Angeles, CA, USA, 2017.
18. Zhang, J.; Zhang, H.; Xia, C.; Sun, L. Graph-Bert: Only Attention is Needed for Learning Graph Representations. *arXiv* **2020**, arXiv:2001.05140.
19. Felmler, D.; McMillan, C.; Towsley, D.; Whitaker, R. Social network motifs: A comparison of building blocks across multiple social networks. In Proceedings of the Annual Meetings of the American Sociological Association, Philadelphia, PA, USA, 11–14 August 2018.
20. Souravlas, S.; Anastasiadou, S.; Katsavounis, S. A Survey on the Recent Advances of Deep Community Detection. *Appl. Sci.* **2021**, *11*, 7179. [CrossRef]
21. Chan, J.; Wang, Z.; Xie, Y.; Meisel, C.; Meisel, J.; Solano, P.; Murillo, H. Identifying Potential Managerial Personnel Using PageRank and Social Network Analysis: The Case Study of a European IT Company. *Appl. Sci.* **2021**, *11*, 6985. [CrossRef]
22. Ying, R.; He, R.; Chen, K.; Eksombatchai, P.; Hamilton, W.L.; Leskovec, J. Graph convolutional neural networks for web-scale recommender systems. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 August 2018.
23. Lee, C.; Han, D.; Han, K.; Yi, M. Improving Graph-Based Movie Recommender System Using Cinematic Experience. *Appl. Sci.* **2022**, *12*, 1493. [CrossRef]
24. AlBadani, B.; Shi, R.; Dong, J.; Al-Sabri, R.; Moctard, O.B. Transformer-Based Graph Convolutional Network for Sentiment Analysis. *Appl. Sci.* **2022**, *12*, 1316. [CrossRef]
25. Boscaini, D.; Masci, J.; Rodolà, E.; Bronstein, M. Learning shape correspondence with anisotropic convolutional neural networks. In *Advances in Neural Information Processing Systems*. 2016. Available online: <https://proceedings.neurips.cc/paper/2016/hash/228499b55310264a8ea0e27b6e7c6ab6-Abstract.html> (accessed on 1 April 2022).
26. Atwood, J.; Towsley, D. Diffusion-convolutional neural networks. In *Advances in Neural Information Processing Systems*. 2016. Available online: <https://proceedings.neurips.cc/paper/2016/hash/390e982518a50e280d8e2b535462ec1f-Abstract.html> (accessed on 1 April 2022).
27. Beck, D.; Haffari, G.; Cohn, T. Graph-to-Sequence Learning using Gated Graph Neural Networks. *arXiv* **2018**, arXiv:1806.09835.
28. Khademi, M.; Schulte, O. Dynamic gated graph neural networks for scene graph generation. In *Asian Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2018.

29. Sukhbaatar, S.; Fergus, R. Learning multiagent communication with backpropagation. In *Advances in Neural Information Processing Systems*. 2016. Available online: <https://proceedings.neurips.cc/paper/2016/hash/55b1927fdafef39c48e5b73b5d61ea60-Abstract.html> (accessed on 1 April 2022).
30. Bruna, J.; Zaremba, W.; Szlam, A.; LeCun, Y. Spectral networks and locally connected networks on graphs. *arXiv* **2013**, arXiv:1312.6203.
31. Defferrard, M.; Bresson, X.; Vandergheynst, P. Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in Neural Information Processing Systems*. *arXiv* **2016**, arXiv:1606.09375.
32. Henaff, M.; Bruna, J.; Lecun, Y. Deep convolutional networks on graph-structured data. *arXiv* **2015**, arXiv:1506.05163.
33. Masci, J.; Boscaini, D.; Bronstein, M.M.; Vandergheynst, P. Geodesic convolutional neural networks on riemannian manifolds. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, Santiago, Chile, 7–13 December 2015.
34. Su, H.; Maji, S.; Kalogerakis, E.; Learned-Miller, E. Multi-view convolutional neural networks for 3D shape recognition. In *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, 7–13 December 2015.
35. Liu, B.; Shen, M. Some Geometrical and Topological Properties of DNNs' Decision Boundaries. *arXiv* **2020**, arXiv:2003.03687.
36. Freeman, C.D.; Bruna, J. Topology and geometry of half-rectified network optimization. *arXiv* **2016**, arXiv:1611.01540.
37. Chen, T.; Goodfellow, I.; Shlens, J. Net2net: Accelerating learning via knowledge transfer. *arXiv* **2015**, arXiv:1511.05641, 2015.
38. Kawaguchi, K. Deep learning without poor local minima. In *Advances in Neural Information Processing Systems*. 2016. Available online: <https://proceedings.neurips.cc/paper/2016/hash/f2fc990265c712c49d51a18a32b39f0c-Abstract.html> (accessed on 1 April 2022).
39. Cao, W.; Yan, Z.; He, Z.; He, Z. A Comprehensive Survey on Geometric Deep Learning. *IEEE Access* **2020**, *8*, 35929–35949. [[CrossRef](#)]
40. Khan, M.M.; Ibrahim, R.; Ghani, I. Cross domain recommender systems: A systematic literature review. *ACM Comput. Surv. (CSUR)* **2017**, *50*, 1–34. [[CrossRef](#)]
41. Zhang, S.; Yao, L.; Sun, A.; Tay, Y. Deep learning based recommender system: A survey and new perspectives. *ACM Comput. Surv. (CSUR)* **2019**, *52*, 1–38. [[CrossRef](#)]
42. Xu, B.; Yu, Y. A personalized assistant in 3D virtual shopping environment. In *2010 Second International Conference on Intelligent Human-Machine Systems and Cybernetics, Nanjing, China, 26–28 August 2010*; IEEE: Washington, DC, USA, 2010.
43. Guo, G.; Elgendi, M. A New Recommender System for 3D E-Commerce: An EEG Based Approach. *J. Adv. Manag. Sci.* **2013**, *1*, 61–65. [[CrossRef](#)]
44. Bu, S.; Wang, L.; Han, P.; Liu, Z.; Li, K. 3D shape recognition and retrieval based on multi-modality deep learning. *Neurocomputing* **2017**, *259*, 183–193. [[CrossRef](#)]
45. Ma, X.; Wang, X. Average Contrastive Divergence for Training Restricted Boltzmann Machines. *Entropy* **2016**, *18*, 35. [[CrossRef](#)]
46. Jang, H.; Choi, H.; Yi, Y.; Shin, J. Adiabatic persistent contrastive divergence learning. In *Proceedings of the 2017 IEEE International Symposium on Information Theory (ISIT)*, Aachen, Germany, 25–30 June 2017.
47. Tyagi, B.; Kumar, V.; Sharma, P. Cross-Domain Recommendation Approach Based on Topic Modeling and Ontology. In *Soft Computing: Theories and Applications*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 397–406.
48. Zhu, Y.; Tang, Z.; Liu, Y.; Zhuang, F.; Xie, R.; Zhang, X.; Lin, L.; He, Q. Personalized transfer of user preferences for cross-domain recommendation. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, Tempe, AZ, USA, 21–25 February 2022.
49. Cao, J.; Sheng, J.; Cong, X.; Liu, T.; Wang, B. Cross-Domain Recommendation to Cold-Start Users via Variational Information Bottleneck. *arXiv* **2022**, arXiv:2203.16863.
50. Deng, S.; Huang, L.; Xu, G.; Wu, X.; Wu, Z. On Deep Learning for Trust-Aware Recommendations in Social Networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *28*, 1164–1177. [[CrossRef](#)] [[PubMed](#)]
51. Liao, W.; Zhang, Q.; Yuan, B.; Zhang, G.; Lu, J. Heterogeneous Multidomain Recommender System Through Adversarial Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**. [[CrossRef](#)] [[PubMed](#)]
52. Wang, H.; Chen, B.; Li, W.-J. Collaborative topic regression with social regularization for tag recommendation. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, Beijing, China, 3–9 August 2013.
53. Wang, C.; Blei, D.M. Collaborative topic modeling for recommending scientific articles. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, CA, USA, 21–24 August 2011.
54. Kanagawa, H.; Kobayashi, H.; Shimizu, N.; Tagami, Y.; Suzuki, T. Cross-domain recommendation via deep domain adaptation. In *European Conference on Information Retrieval*; Springer: Berlin/Heidelberg, Germany, 2019.
55. Sahebi, S.; Brusilovsky, P.; Bobrov, V. Cross-domain recommendation for large-scale data. In *Proceedings of the CEUR Workshop Proceedings*, 1 January 2017; Available online: <http://d-scholarship.pitt.edu/id/eprint/35050> (accessed on 1 April 2022).
56. Hong, W.; Zheng, N.; Xiong, Z.; Hu, Z. A Parallel Deep Neural Network Using Reviews and Item Metadata for Cross-Domain Recommendation. *IEEE Access* **2020**, *8*, 41774–41783. [[CrossRef](#)]
57. Liu, B.; Shen, M. Some geometrical and topological properties of DNNs' decision boundaries. *Theor. Comput. Sci.* **2022**, *908*, 64–75. [[CrossRef](#)]