

## Trust-based collaborative filtering for cross-domain recommendations using ANFIS and FAFCM algorithm

G. Spoorthy <sup>1</sup> and S. G. Sanjeevi <sup>2</sup>

<sup>1</sup>Assistant Professor, Department of CSE- (CyS, DS) and AI&DS, VNRVJIET, Telangana, India

<sup>1,2</sup>Department of CSE, NIT Warangal, Warangal, INDIA

gspoorthy6@gmail.com, sgs@nitw.ac.in

### Abstract

Recommendation systems are very useful in domains such as e-commerce, news portals and software requirement analysis. Collaborative filtering models have been used widely, these models often suffer from data-sparsity, interpretability and cold start problems. To solve these issues, various machine-learning, deep-learning and kernel-based models have been employed. Among these, trust-based collaborative filtering and cross-domain recommendations have successfully solved the issues to some extent. However, in the recent literature, cross-domain recommendations (CDRs) are made by taking common user ratings. In our paper, we introduce a novel approach that combines CDRs with trust-aware collaborative filtering, which employs, a partial item overlap' scenario. Our model operates in two phases: an offline phase calculates trust between source and target users, and coarse rating prediction using an Adaptive Neuro-Fuzzy Inference System (ANFIS), and clustering via Firefly Fuzzy C-Means (FAFCM); and an online phase, where cluster information and item similarities are used to generate personalized recommendations. Evaluated on the Douban and MovieLens datasets using MAE and RMSE metrics, our approach demonstrates improved performance compared to existing methods, effectively mitigating common limitations in recommendation systems..

**Keywords:** Cross-domain recommendation systems, trust-aware cross-domain recommendations, ANFIS, fuzzy c-means clustering.

## 1 Introduction

Enormous and ever-growing complex and heterogeneous digital information available on many e-commerce sites, news portal sites overload human ability to choose. To make humans cope with such information overloads, recommendation systems came into existence. Recommendation Systems (RS) main objective is to provide recommendations of the products/items according to the taste/preferences of customers/users. Recommendation system (RS) models derive recommendations by leveraging users' historical data, which may include content (used in content-based RS), ratings (used in collaborative filtering), or a combination of both content and ratings (used in hybrid approaches). Many approaches have been implemented to improve the accuracy with which the system is predicting the recommendations.

Content-based RS (CBF) [26] tries to recommend items/products that share similar attributes concerning historical item information. The attribute information is not sparse, so content based RS doesn't suffer from data-sparsity. Content-based filtering (CBF) can struggle with performance, especially when it comes to recommending items to new users, since it doesn't have enough information about their preferences yet. Collaborative filtering (CF) [5, 10] recommends items/products based on the historical ratings given by the user on items. As the rating matrix is sparse with respect to the vast set of items with few rated items. CF suffers from data-sparsity issues. CF also suffers from cold-start issues. CF performance depends on user ratings which makes the model less effective in predicting recommendations

Corresponding Author: G. Spoorthy

Received: August 2024; Revised: September 2025; Accepted: November 2025.

<https://doi.org/10.22111/ijfs.2025.49547.8745>

whenever a new user or item is introduced. To address the issues of CBF and CF, studies opted for hybrid approaches.

Cross-domain recommendation systems(CDR) [4] can solve the data sparsity and cold-start problems by capturing relevant information like user-preferences and item information from the source domain to improve recommendations in the target domain. The source domain is the domain where there are sufficient user-interactions available to derive possible insights. The target domain refers to the domain where there are limited or sparse user-interactions. The target domain goes through data sparsity and cold-start issues, which lead to inaccurate recommendations. For instance, in the Amazon dataset, if a user likes/ purchases books of machine-learning then Amazon can recommend a kindle subscription to read relevant books or CDs. Another example is suggesting hotels as the users travel destinations. These recommendations can enhance domains with minimal user interactions to help the organization with more customer engagements. If data sparsity and cold-start issues are effectively managed, it can lead to a better user experience and greater customer satisfaction.

CDRs are mainly categorized into two types: content-based and transfer-based models. Content-based methods [23] try to build similarity between different domains based on user and item attributes (such as tags, and reviews) [6, 7, 12, 30]. Transfer-based methods aim to map the latent representations [13] or rating patterns from one domain where data is available to another where it's limited. Transfer-based models to transfer knowledge employ machine learning and neural networks [16, 25]. The accuracy of these methods relies on the precise mapping of latent representations. Transfer-based models are further divided into three categories:

(i) Inductive transfer-based models handle scenarios where the task in the source domain varies from the task learned in the target domain;

(ii) Unsupervised transfer-based models, focus on solving tasks related to clustering, density estimation, and dimensionality reduction in the target domain;

(iii) Transductive transfer-based models address the cases where learning tasks in both domains are the same but occur in different domains. The traditional transfer-based models often fail to attain accurate latent factor mappings between source and target domains. For instance, [13] focuses on mapping only the latent factors of users or items that exist in both domains, whereas [20] explores the intricate relationships among latent factors across two domains.

In our model, we include trust computations between source and target users, making the traditional collaborative filtering based on the transfer model more attentive towards the trusted users in the source domain. We use ANFIS and FAFCM algorithm model to build recommendations. Our distinct contributions and key innovations are summarized below:

- We introduce a novel method to foresee ratings given by users from the source domain for items in the target domain under a partial item overlap scenario. By leveraging trust-based relationships across domains, our method solves issues related to having too little data and handling new users or items.
- A unique trust-aware recommendation framework is developed, combining trust relationships and ratings using Adaptive Neuro-Fuzzy Inference Systems (ANFIS).
- To improve accuracy, we employ FAFCM for clustering source domain user ratings, ensuring better grouping and recommendation quality. This clustering technique provides superior granularity and relevance in user segmentation.
- Our model is evaluated using MAE and RMSE metrics, and its performance is compared to state-of-the-art methods, such as MVDNN [11], GCBAN [15], and NeuCDCF [25]. The findings highlight the effectiveness of our method compared to others demonstrating significant improvements in prediction accuracy and recommendation quality.

In Section 2, we describe the related works, followed by our proposed method in Section 3. Evaluation and discussion are provided in Section 4, while conclusions and future directions are discussed in Sections 5 and 6.

## 2 Related work

### 2.1 Cross-domain recommendation systems

Cross-domain recommendation systems (CDRs) gained attention in the recommendation research due to their tolerance towards data-sparsity and cold start problems. Cross-domain [18] refers to the practice of using knowledge from the source domain to enhance the quality and produce accurate recommendations in the target domain. Based on the paper [32] the data overlap in CDRs is divided into four categories:

- (i) No Overlap: Here we have no shared users or items between the domains. ( $u_{st} = u_s \cap u_t = \emptyset; I_{st} = I_s \cap I_t = \emptyset$ ).
- (ii) Users Overlap: Common users exist in source and target domains. ( $u_{st} = u_s \cap u_t \neq \emptyset$ ).
- (iii) Items Overlap: Common items exist in source and target domains. ( $I_{st} = I_s \cap I_t \neq \emptyset$ ).
- (iv) Users and Items Overlap: There exist common users and items between domains. ( $u_{st} = u_s \cap u_t \neq \emptyset; I_{st} = I_s \cap I_t \neq \emptyset$ ).

They are further split into two categories based on the overlapping that is, either overlap is partial/fully overlap at all [22] or there is no-overlap [16, 28]. In partial/fully overlap scenarios, latent features of users or items are shared by source and target domains. In partial item overlap scenario, the source and target domain have common items between source and target domains. It is represented as  $I_{ST}$ . In a no-overlap case, there is no overlap and no sharing of latent features. TUCross model [28] builds recommendations without any common ratings by using user context and trust information for cross-domain recommendations.

### 2.2 Deep learning based cross-domain recommendation systems

In the current research, deep learning models have been in demand for building recommendations. Recommendation system models [19] are used to represent data in low-dimensional representations, which enhances the recommendation systems performance. MVDNN [11] proposes a deep learning framework for CDRs to map users and items ‘auxiliary information’ to the latent space. GCBAN [15] model uses two Bayesian neural networks, one learns fewer dimension representations of ratings and the other one projects the source domain information to different latent spaces. Later on both the networks are shared to different target domains. NeuCDCF [25] model proposes a neural network framework for CDRs without ‘auxiliary information’. TLRec [7], uses transfer knowledge between source and target domains using user overlap and item overlap. To learn latent vectors [7] used bipartite graphs which represent implicit user-item interactions. MINDL model [14] divides the rating matrix into codebooks (multiple-cluster level matrices). Using these codebooks reconstructs the matrix to learn ratings of different target domains. The above stated transfer- based models, GCBAN [15] do not take collaborative filtering approaches. NeuCDCF [25] performance depends on the weight optimization used and does not include trust to enhance the performance. [28] can be enhanced using neural network models. TLRec [7] and MINDL [16] performance depends on overlapped user-items and codebooks. Thus, our models employ a trust-based neural network within a cross-domain framework to generate accurate recommendations. In the subsection 2.3, we will discuss the strengths and weaknesses of the literature.

### 2.3 Literature review

Reference	Strengths	Weaknesses
TLRec: Transfer Learning for Cross-Domain Recommendation. [7]	Uses transfer learning to leverage knowledge from source domains, improving recommendations in the target domain. Provides solid experimental validation	Transfer learning approaches might struggle with highly different domains. May face challenges with low-quality or incomplete source domain data.
Cross-Domain Sequential Recommendation (CDSR) [31]	Effectively combines single-domain and cross-domain user preferences. Provides a comprehensive framework for sequential recommendations.	May struggle with domains having significantly different user behaviors. Computational complexity increases with the number of domains.
Review-Based on Cross-Domain Recommendation via Hyperbolic Embedding [8, 27]	Introduces hyperbolic embedding to capture hierarchical relationships. Focuses on domain disentanglement for improved recommendation accuracy.	Hyperbolic embedding requires specialized knowledge to implement. May not scale well with very large datasets.

Reference	Strengths	Weaknesses
Exploiting Trust and Usage Context for Cross-Domain Recommendation. [28]	Focuses on combination of trust and usage context, which enhances recommendation quality. It explores real-world applicability in cross-domain	Limited explanation of scalability across diverse domains. May not be effective in cases with extremely sparse data.
A multi-view deep learning approach for cross-domain user modeling in recommendation systems. [11]	Introduces multi-view deep learning, which enhances user modeling by integrating diverse data sources. Strong focus on cross-domain user modeling.	May be computation expensive due to deep learning methods. Complexity increases as more data views are added.
A General Cross Domain Recommendation Framework via Bayesian Neural Networks [15]	Utilizes a Bayesian neural network, which provides a probabilistic framework, making the model robust to uncertainties. Offers a generalized recommendation framework.	Bayesian neural networks can be complex and require substantial computational resources. - May face difficulties in highly dynamic domains
Neural Cross-Domain Collaborative Filtering with Shared Entities [25] & Focuses on neural networks for collaborative filtering, offering high potential for accurate recommendations. Incorporates shared entities for cross-domain recommendation	Requires large amounts of labeled data for training. Potential overfitting due to neural network complexity.	
A Cross-Domain Recommender System Based on Common-Sense Knowledge Bases.[24]	Leverages common-sense knowledge, which enhances the model's ability to understand contextual relationships. Effective in domains requiring contextual awareness	May not be applicable in domains lacking structured knowledge bases. Knowledge base dependency may limit adaptability

## 2.4 Research prerequisites

### 2.4.1 Description of ANFIS

The fuzzy logic (FL) concept was derived from the theory of fuzzy sets. FL is designed in 1975. FL is the mathematical representation of human reasoning. A fuzzy set  $F$  is defined as  $F \rightarrow R[0, 1]$  where  $R$  is represented by the membership function. A fuzzy set is a collection of elements with various degrees of membership. Membership Function (MF) defines how each input is mapped to a membership value between 0 and 1. The MF is denoted as  $\mu_F$  and the function is defined as  $\mu_F : R[0, 1]$ . The degree of membership of an element 'r' (where  $r \in R$ ) in the fuzzy set  $R$ . Fuzzy C-Means Algorithm is used to solve big data clustering issues too [9] Adaptive Neuro Fuzzy Interface System (ANFIS) was originally designed in 1990 to enhance the models by fuzzy interface and Neural networks. ANFIS is a hybrid model that integrates the advantages of fuzzy logic and neural networks. It combines the interpretability of fuzzy systems with the learning capability of neural networks. ANFIS is used in various applications like pattern recognition [2], predictions, Energy Congestion Control Scheme for Wireless Sensor Networks [1], and system model. ANFIS basically consists of five layers. The five layers are described below:

Layer 1: Fuzzification Layer: Using fuzzy linguistic variables and fuzzy linguistic terms, calculate the membership function and convert a crisp set to a fuzzy set. For example, in our model, linguistic variables are highly distrusted (HD), Distrust (D), UnTrust (UT), Trust (T), Highly Trusted (HT), Very Low (VL), Low (L), Medium (M), High (H), and Very High (VH). The membership function represents the degree of satisfaction a user has with a particular item.

Layer 2: Antecedent Layer: Here, the firing strength of fuzzy rules is derived. This gives the weight factor for the layer 1 input and automatically generates fuzzy rules.

Layer 3: Consequent Layer: In this layer, normalization of firing strength is calculated.

Layer 4: Fuzzy interface layer: combines the independent variables with the dependent variables. It is represented as IF - Antecedent - THEN - Consequent.

### 2.4.2 Description of FAFCM

Firefly Fuzzy C-Means (FAFCM) was developed in 2014 [21]. FAFCM is used to improve cluster quality by minimizing the objective function. FAFCM works in two stages: in the first stage, cluster centers using the firefly algorithm are

identified. These cluster centers serve as inputs for the second stage, where the clusters are refined using Fuzzy C-Means.

Stage 1: Firefly Algorithm [29] is proposed as a nature-inspired meta-heuristic approach using the concept of bioluminescent fireflies. Initially, the fireflies are distributed in the search space. The movement of the fireflies from one position to another depends on the light intensity and attractiveness. The intensity and attractiveness are calculated using mathematical formulae given below:

$$I_i = I_0 e^{-\gamma r_{ij}^2}, \quad (1)$$

where  $I_i$  is the light intensity at 'i',  $I_0$  is the light intensity at  $r_{ij} = 0$ , and  $\gamma$  is the absorption coefficient.

$$\beta(r) = \beta_0 e^{-\gamma r^2}, \quad (2)$$

The attractiveness  $\beta_{ij}$  of 'i' firefly on 'j' firefly is influenced by the light intensity seen and the distance  $r_{ij}$  between them and attraction at  $\beta_0$ .

$$r_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\| = \sqrt{\sum_{k=1}^D (x_{i,k} - x_{j,k})^2}, \quad (3)$$

$x_i, x_j$  are positions of fireflies.

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \beta_0 e^{-\gamma r_{ij}^2} (\mathbf{x}_j^t - \mathbf{x}_i^t) + \alpha(\text{rand} - 0.5). \quad (4)$$

The movement of firefly is extremely based on the attraction and distance between the fireflies. Using the above equations, we iterate the algorithm till we reach the stopping criteria.

Stage 2: FCM is proposed in [3] and has been used as a clustering approach in [17]. We initialize 'k' cluster centers attained in stage 1 as input to FCM. A membership matrix (fuzzy partition matrix) which defines the fuzzy membership function matrix. It is denoted as  $MF_c$ .

$$MF_c = \left\{ \begin{array}{l} W \in R^{cn} | w_{ik} \in [0, 1], \forall i, k; \\ \sum_{i=1}^c w_{ik} = 1, \forall k; 0 \quad lt; \sum_{k=1}^n w_{ik} \quad lt; n, \forall i \end{array} \right\}. \quad (5)$$

By using the Euclidean distance metric and membership value  $MF_c$  we calculate the objective function.

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, \quad (6)$$

where  $1 \leq k \leq n; 0 \leq i \leq c; m \in (1, +\infty)$

$$d_{ik} = \|x_k - p_i\|. \quad (7)$$

Where m gives the fuzziness of resulting clusters, and  $d_{ik}$  is the distance from  $x_k$  to cluster center  $p_i$ .

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m x_i}{\sum_{i=1}^N u_{ij}^m}. \quad (8)$$

$$\mu_{ij} = \frac{1}{\sum_{k=1}^C \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}. \quad (9)$$

## 3 Method

### 3.1 Problem formulation

In our model, we will be using a partial overlap scenario. Our reason to do so is that no overlapping or no sharing is seldom seen according to our research. Let  $U_S = U_{S1}, U_{S2}, \dots, U_{Sm}$  be the users and  $I_S = I_{S1}, I_{S2}, \dots, I_{Sn}$  be items of a source domain.  $U_T = U_{T1}, U_{T2}, \dots, U_{Tm}$  and  $I_T = I_{T1}, I_{T2}, \dots, I_{Tn}$  be the users and items of a target domain. The rating matrix for the source domain is  $R_S = U_S I_S$ , and for the domain is  $R_T = U_T I_T$ . Let the partial common 'c' items between source and target domains be  $I_{ST} = I_{ST1}, I_{ST2}, \dots, I_{STc}$ . We predict the coarse predicted ratings using ANFIS on  $T_{U_S U_T}$  and  $R_T$ , where  $T_{U_S U_T}$  is the trust between source and target users.  $R_T$  is the target domain user-item matrix. After obtaining the predicted ratings, we generate the co-occurrence matrix of items in source and target domains, and the similarity between them is calculated. The source domain users are clustered using FAFCM. For an active user, their respective cluster is identified. Subsequently, ratings are predicted, and the items are recommended. Figure 1 illustrates the architecture of our model.

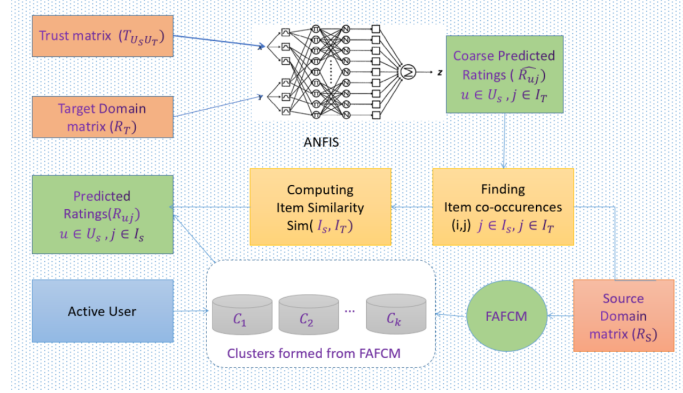


Figure 1: Trust-based collaborative filtering for Cross-Domain Recommendations using ANFIS and FAFCM Algorithm (TCFCDRS)

### 3.2 Proposed method

The proposed model consists mainly of two phases: the offline phase and the online phase. Phase 1: The offline phase is subdivided into two components: the coarse rating phase and the rating prediction phase. (I) Coarse prediction of ratings: Based on the trust between the source and target domains. We calculate the coarse ratings of the users of the source domain on the target items. Step 1: First, we build trust between users in different domains using Pearson's correlation coefficient (PCC). The mathematical formula is given below:

$$T_{U_s U_v} = \frac{\sum_i (r_{U_s i} - \bar{r}_{U_s})(r_{U_T i} - \bar{r}_{U_T})}{\sqrt{\sum_i (r_{U_s i} - \bar{r}_{U_s})^2} \sqrt{\sum_i (r_{U_T i} - \bar{r}_{U_T})^2}}, \quad (10)$$

where  $T_{U_s U_v}$  is the trust between the source and the target users,  $r_{U_s i}$ ,  $r_{U_T i}$  are ratings at 'i' and  $\bar{r}_{U_s}$  and  $\bar{r}_{U_T}$  are means. Step 2: We predict the coarse ratings of source users on target items.  $T(U_s U_v)$  and  $R_T$  are given as an input to ANFIS to predict trust-aware recommendations. ANFIS is further divided into 5 layers.

- (i) Fuzzification layer: This layer converts crisp data to fuzzy sets. In our recommendation model, we used 'trust' and 'ratings' as input linguistic variables. We used a trapezoidal membership function (TMF), which let us fix a range. TMF, denoted as  $\mu(x)$  for trust is defined below:

$$\mu(x) = \begin{cases} 1 & a_2 \leq x \leq a_3 \\ 0 & x = a_1 \text{ or } x = a_4 \\ \frac{x-a_1}{a_2-a_1} & a_1 < x < a_2 \\ \frac{x-a_4}{a_3-a_4} & a_3 < x < a_4. \end{cases} \quad (11)$$

We defined trust levels as highly distrusted (HD), Distrust (D), UnTrust (UT), Trust (T) and Highly Trusted (HT). Table II values of trust levels. TMF, denoted as  $\mu(x)$  for target ratings is defined below: We defined rating

Table 1: values of trust levels.

Trust Level	TMF
HD	[-1,-0.7,-0.5]
D	[-0.7,-0.5,0]
UT	[0,0,0,0.2]
T	[0,0.2,0.5,0.7]
HT	[0.5,0.7,1]

levels as Very Low (VL), Low (L), Medium (M), High (H) and Very High (VH). Table III values of rating levels.

Table 2: values of rating levels.

Rating Level	TMF
VL	[0,0,1]
L	[1,1,2]
M	[1,1,2,3]
H	[2,3,3,4]
VH	[3,4,5]

- (ii) Antecedent Rule layer: This layer defines the linguistic rules used by the rule evaluator. Rules are in the form of if-then. For the  $l^{th}$  rule, representation can be written as:

$$Rule_l : IF x_i \text{ is } X_i^l \text{ AND } x_1 \text{ is } X_1^l, \text{ THEN } y \text{ is } Y_l, \forall l = 1, 2, \dots, l, \quad (12)$$

where  $X_i^l, X_1^l$  are antecedent sets,  $x_i, x_1$  are inputs,  $y$  is the output and  $Y_l$  are consequent sets. Antecedent Rule will use two inputs “Trust” and “ratings” and give “trusted target ratings”. The trusted target ratings will be used for coarse rating generation.

$$d_j = \prod_{i=1}^m \mu_{X_j^i}(x_i^0), \quad (13)$$

where membership degree  $d$  of  $j^{th}$  rule and  $x_i^0$  is the input value to the trust value.

$$w_j = \prod_{i=1}^m \mu_{r_T^i}(r_i^0), \quad (14)$$

where membership degree  $w$  of  $j^{th}$  rule and  $r_i^0$  is the input value to the rating value.

- (iii) Consequent Layer: This layer gives the firing strengths.

$$\bar{D} = \frac{d_i}{\sum_{i=1}^x d_i}, \quad (15)$$

where  $\bar{D}$  is firing strength of Trust value.

$$\bar{W} = \frac{w_i}{\sum_{i=1}^r w_i}, \quad (16)$$

where  $\bar{W}$  is firing strength of rating value.

- (iv) Fuzzy interface model: we use Mamdani Fuzzy Logic principles to combine all rules mapping from the antecedent layer to the output layer. In this layer, using Max-Min composition all the antecedents are combined. If Rule uses OR then data uses S-norm else if it is AND data uses T-norm. In our model we used AND so the data takes T-norm.

$$Rule^l : f^l = \overline{\mu_{Trust}(x)} * \overline{\mu_{Rating}(y)}, \overline{\mu_{Trust}(x)} * \overline{\mu_{Rating}(y)}, \quad (17)$$

where  $*$  is a min t-norm

$$D = \frac{\sum_{i=1}^m d_i y_i}{\sum_{i=1}^x d_i}. \quad (18)$$

The result of the trust value

$$W = \frac{\sum_{i=1}^m w_i y_i}{\sum_{i=1}^x w_i}. \quad (19)$$

The result of rating values Step3- Calculating Coarse prediction of ratings

$$\hat{R}_{uj} = \frac{\sum_{u \in U_S, v \in U_T, j \in I_{ST}} (D_{u,v} \times W_{v,j})}{\sum W_{u,v}}, \quad (20)$$

where rating predicted  $\hat{R}_{uj}$  for source user  $u$  on item  $j$  of the target user and  $D(u, v)$  trust between source and target users,  $W(v, j)$  is the target user rating for the overlapped item.  $W(u, v)$  is the source and target user

ratings. Rating prediction: The rating history of both domains we calculate item similarities. Based on the coarse predictions obtained in phase 1 and item similarities we build prediction for cross-domain. Step 4: Build the co-occurrence matrix between items of Coarse ratings predicted in step 3 and source items. Step 5: Calculate the similarity between source  $V_{I_s}$  and target items  $V_{I_T}$  using Cosine Similarity. The mathematical formula is given below:

$$\text{Sim}_{I_s I_T} = \frac{V_{I_s} \cdot V_{I_T}}{|V_{I_s}| \cdot |V_{I_T}|}. \quad (21)$$

Step 6: Using FAFCM shown in Figure 2, we form the clusters. When an active user comes we give the cluster to which the active user belongs as input for predicting ratings. The FAFCM algorithm is a clustering technique that It begins by initializing fuzzy memberships matrix and cluster centers shown in equation 5. Iteratively memberships are updated based on weighted distances shown in equation 6 and 7. It recalculates cluster centers using fuzzy memberships. Weights adjustment are adjusted the influence of each feature dynamically using equation 8 and 9, making the algorithm robust to sparse data. The process continues until convergence. The FAFCM is shown below:

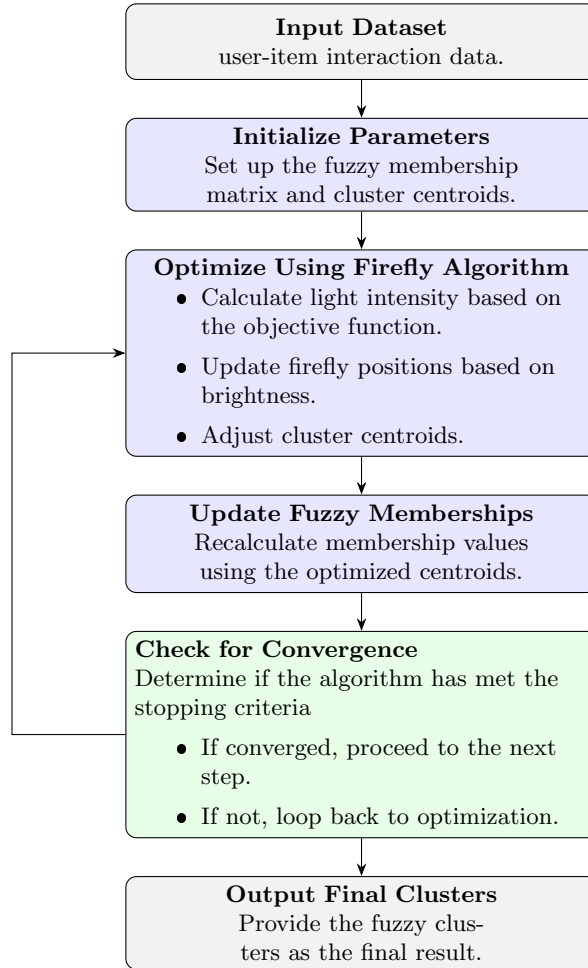


Figure 2: FAFCM

This ends our offline phase. Phase 2: In the online phase, we identify the corresponding cluster of an active source domain user. This cluster, combined with the item similarity matrix, is used to predict ratings and generate recommendations for target domain items. Step 7: Predict ratings  $P_{(AU_ratings, C_x, j)}$  for an active user of the source domain. The mathematical formula is given below:

$$P_{AU_{ratings} C_x, j} = \overline{R_{AU}} + \frac{\sum_{y \in c_{au}, u \in U_s, i \in I_s, j \in I_r} \text{sim}_{i,j}(AU, y) \times (R_{u,i})}{\sum_{y \in c_{au}} (|\text{sim}_{i,j}(AU, y)|)}. \quad (22)$$

Where  $\overline{R_{AU}}$  is the average ratings of active user, similar users  $\text{sim}_{i,j}(AU, y)$  with respect to active users,  $R_{u,i}$  is the ratings of user for items in source domain

## 4 Evaluation and discussion

### 4.1 Dataset description

There are 3 types of datasets: Movie100k [14], Movie1m and Douban [32]. Movie100k and Movie1m datasets were given by GroupLens. The ratings range from 1-5. The dataset has user information of users' age, gender and occupation; item information of movie name, IMDB Url and genres. We split Movie100k and Movie1m into 2 domains making sure items with high similarity belong to the same domain. D1 Contains movies with high similarity in certain genres or features. For example, movies predominantly in the action or adventure genres. User preferences might align more with these genres in this domain. D2 contains movies with high similarity in other genres or features that are distinct from D1. For example, movies predominantly in the romance or drama genres.

Douban dataset includes 3 tasks for recommendations, namely Douban-Music, Douban-Movies and Douban-Books. The dataset has user information such as age, gender, place, tags and occupation; Douban-Music has item information of genres, singer name and publisher. Douban-Movie has item information of movie name, genres, director, etc. Douban-Books has item information of book name, publisher, category, country, etc.

To perform our CDRs model, we choose eight tasks similar to [25]. The eight tasks are  $Movie100k - D1 \longleftrightarrow Movie100k - D2$ ,  $Movie1m - D1 \longleftrightarrow Movie1m - D2$ ,  $Douban - Movie \longleftrightarrow Douban - Music$  and  $Douban - Music \longleftrightarrow Douban - Books$ . The overlapped users are taken as target users of the testing data to test our model. The statistics of the dataset are similar to those of GCBAN [15].

### 4.2 Experiment evaluation

We used MAE and RMSE for comparing with three base models- MVDNN [11], GCBAN [15] and NeuCDCF [25] with our model TCFCDRS based on over model performance and Performance of the model when Cold-start issue arises.

### 4.3 Performance comparison

Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) performance metrics are used for comparing with three base models- MVDNN [11], GCBAN [15] and NeuCDCF [25] the basis of over model's performance, the average performance of the model in five trails are recorded. The definition of MAE and RMSE is given below:

- (i) MAE: Average of the difference between predicted ratings and actual ratings.

$$\frac{\sum_{i=1}^n |y_i - x_i|}{n}. \quad (23)$$

- (ii) RMSE: Square root of Mean Squared Error.

$$\sqrt{\frac{\sum_{i=1}^n (\hat{y}_n - y_i)^2}{n}}. \quad (24)$$

RMSE and MAE comparison of our model w.r.t. related research on different datasets. The training and test data are divided into 80% and 20% respectively. Only the target domain is subjected to the splitting of data into train-test sets. Therefore, only target domains get to be analyzed for the performance analysis. Table IV shows the relevant RMSE and MAE comparisons on MovieLens datasets. From the above comparisons, we can see that the effect of including neural networks in the Cross-Domain recommendations makes the model perform more efficiently compared to other deep learning based approaches like MVDNN. The improvement of GCBAN over NeuCDCF may be due to the Bayesian framework, which makes the model learn uncertainty along with which GCBAN learns approximate posteriors. Both GCBAN and NeuCDCF takes auxiliary information to enhance the recommendations. Our model's performance may be due to using a partial-overlap of items, so we concentrate more on the data rather than the latent feature representations. Though this model takes more time to build, the trust relationships between source and target domains enable our model to learn more flexible recommendations. To optimize the reduced time and resources needed for recommendations, we introduced a clustering approach. For calculating recommendations for an active user, instead of taking the entire source domain ratings matrix as input, we find to which the active user belongs. Only that cluster information in the source domains builds the recommendations. This point highlights how important cluster formation of role. To achieve the optimized clusters, we utilized the FAFCM model. Table IV shows the performance of our model with respect to related work on the movielens100k data with 2 domains, D1 and D2, and movielens1m data with 2 domains D1 and D2. Table V shows the performance of our model with respect to related work on the Douban dataset with movies, music and books as the target domain. Table V shows the relevant RMSE and MAE comparisons on Douban dataset.

Table 3: Comparative Performance Results

Dataset	Model	Measure	D1 as TD	D2 as TD
Movie 100k	MVDNN	RMSE	1.1135	1.1076
		MAE	0.9257	0.9162
	GCBAN	RMSE	0.9042	0.9291
		MAE	0.7137	0.7333
	NeuCDCF	RMSE	0.9209	0.9616
		MAE	0.7183	0.7572
	TCFCDRS	RMSE	0.8761	0.8949
		MAE	0.6845	0.6994
Movie 1m	MVDNN	RMSE	1.0783	1.0981
		MAE	0.8759	0.9047
	GCBAN	RMSE	0.8502	0.8675
		MAE	0.6668	0.6815
	NeuCDCF	RMSE	0.8825	1.0740
		MAE	0.9115	0.8747
	TCFCDRS	RMSE	0.8178	0.8203
		MAE	0.6241	0.6334

#### 4.4 Cold-start issues

When a model has to generate recommendations for a new user, it leads to new user cold-start issues. Similarly, when new items are added to the system, how efficiently a model can recommend this new item is a really challenging task. A model’s ability to perform accurate predictions when faced with cold-start situations sets it apart from other related models. To evaluate the model’s performance when cold-start issues are encountered, we conducted some cold-start issues experiments. We have taken the users with only given 5 ratings in the test data (20%) for new user cold-start issues. We have taken the items with only 5 ratings in the test data (20%) for new item cold-start issues. The dataset used for the evaluation is the MovieLens 100k dataset. Figure 3 shows our model addresses new user cold-start issues more efficiently than the related research MVDNN, GCBAN and NeuCDCF. This superior performance can be attributed to our model’s ability to establish trust relationships between users before predicting recommendations, enabling it to excel in this scenario.

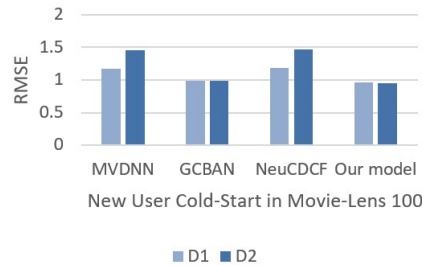


Figure 2: New User Cold-start in Movie-Lens

Similarly, Figure 4 shows our model is more efficient than the related research MVDNN, GCBAN and NeuCDCF to deal with the new item cold-start issue. This is likely because our model calculates item similarities between the source and target domains beforehand, ensuring more accurate and reliable recommendations. Figure 3 and Figure 4 show the efficiency of our model compared to other models when the cold-start situation.

Table 4: Performance Measures for Different Models

Dataset	Model	Measure	TD
Movie as Target Domain	MVDNN	RMSE	0.902
		MAE	0.6746
	GCBAN	RMSE	0.70
		MAE	0.5462
	NeuCDCF	RMSE	0.7496
		MAE	0.5689
TCFCDRS	RMSE	0.681	
	MAE	0.5188	
Music as Target Domain	MVDNN	RMSE	0.7963
		MAE	0.6317
	GCBAN	RMSE	0.6531
		MAE	0.5167
	NeuCDCF	RMSE	0.7823
		MAE	0.5485
TCFCDRS	RMSE	0.6179	
	MAE	0.4984	
Books as Target Domain	MVDNN	RMSE	0.8079
		MAE	0.666
	GCBAN	RMSE	0.7461
		MAE	0.5713
	NeuCDCF	RMSE	1.0365
		MAE	0.7842
TCFCDRS	RMSE	0.7086	
	MAE	0.5508	

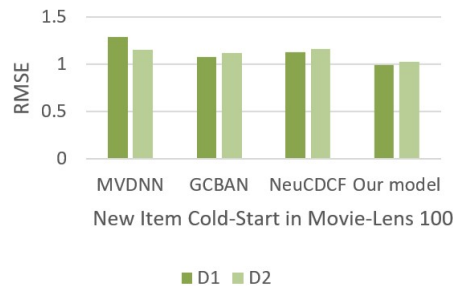


Figure 3: New User Cold-start in Movie-Lens 100k dataset when test set is 20%

## 5 Conclusion

In this research, we proposed a cross-domain recommendation system to capture user interests by exploring user preferences by elevating trust between source domain and target domain users. In this paper, our focus is on solving data sparsity, scalability, cold start and accuracy issues. Our model is a ‘partial item-overlap’ cross-domain recommendations, it tries to build a general model by exploiting the preferences of users in other domains to make interesting recommendations to users. Our model has four advantages, mainly compared to related research, MVDNN, GCBAN and NeuCDCF.

- (i) Our model uses trust between the users instead of concentrating on common users between two domains, making our model, more adaptable and very near to a realistic view.
- (ii) Our model uses an ANFIS model which gives coarse predicted ratings that are more transparent, as ANFIS possesses both numerical linguistic knowledge
- (iii) our model finds the item similarities between the source and target domains before predicting the recommendation.

- (iv) Our model using FAFCM clusters reduces the time and resources needed for recommendations. Finally, we compared our model with MVDNN, GCBAN and NeuCDCF with respect to measures of MAE, RMSE and cold start issues and showed that using our model gave better results compared to previous research works, MVDNN [11], GCBAN [15] and NeuCDCF [25].
- (v) While ANFIS improves interpretability compared to some neural networks, the integration with FAFCM and trust-based CF can still result in a model that is challenging to fully understand and justify to end-users.

The trust-aware mechanism notably enhanced recommendation quality, especially in sparse domains, by effectively utilizing reliable connections across domains. FAFCM clustering contributed to more personalized and accurate recommendations by identifying meaningful user groupings based on trust and behavioral patterns. Additionally, the adaptive inference system (ANFIS) captured complex, non-linear relationships among trust, user similarity, and cross-domain influences—patterns that traditional models often fail to recognize. This combination not only improved the system’s ability to handle new users and items by transferring trust-based insights across domains but also enhanced the overall interpretability and adaptability of the recommendations through the use of fuzzy logic.

## 6 Future scope

As a future scope, Cross-Domain recommendations with “Partial item overlap” are very sparse in the literature. We suggest to build a new CDR model of this type. Including trust and user and item auxiliary information can make more accurate recommendations. While the model addresses cold-start issues to some extent, it still relies on partial item overlap and trust relationships. If a user or item lacks sufficient interactions or trust information in both domains, recommendation quality may degrade. These issues have to be addressed in future research.

## References

- [1] A. Ali, M. Vadivel, *Efficient energy congestion control scheme for wireless sensor networks using adaptive neuro fuzzy inference system with black widow optimization*, Iranian Journal of Fuzzy Systems, **20**(6) (2023), 187-202. <https://doi.org/10.22111/ijfs.2023.43990.7748>
- [2] V. Behbood, J. Lu, G. Zhang, *Fuzzy refinement domain adaptation for long term prediction in banking ecosystem*, IEEE Transactions on Industrial Informatics, **10**(2) (2014), 1637-1646. <https://doi.org/10.1109/TII.2013.2291942>
- [3] J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, New York, (1981), 95-107. <https://doi.org/10.1007/978-1-4757-0450-1>
- [4] I. Cantador, I. Fernández-Tobías, S. Berkovsky, P. Cremonesi, *Crossdomain recommender systems*, in Recommender Systems Handbook. Boston, MA, USA: Springer, (2015), 919-959. [https://doi.org/10.1007/978-1-4939-1614-2\\_29](https://doi.org/10.1007/978-1-4939-1614-2_29)
- [5] R. Chen, Q. Hua, Y. S. Chang, B. Wang, L. Zhang, X. Kong, *A survey of collaborative filtering-based recommender systems: From traditional methods to hybrid methods based on social networks*, IEEE Access, **6** (2018), 64301-64320. <https://doi.org/10.1109/ACCESS.2018.2877543>
- [6] L. Chen, Y. Zhang, W. Li, *Large language models for domain-agnostic recommendation systems*, Proceedings of the 2023 IEEE International Conference on Data Mining (ICDM), (2023), 345-355. <https://doi.org/10.1145/3705727>
- [7] L. Chen, J. Zheng, M. Gao, A. Zhou, W. Zeng, H. Chen, *TLRec: Transfer learning for cross-domain recommendation*, Proceedings - 2017 IEEE International Conference on Big Knowledge (ICBK), (2017), 167-172. <https://doi.org/10.1109/ICBK.2017.30>
- [8] Y. Choi, *Review-based cross-domain recommendation via hyperbolic embedding and hierarchy aware domain disentanglement*, Arxiv, (2024). <https://doi.org/10.1145/3618349.3636594>
- [9] L. S. Dandugala, K. S. Vani, *Advancing big data clustering with fuzzy logic-based IMV-FCA and ensemble approach*, Iranian Journal of Fuzzy Systems, **21**(2) (2024), 141-160. <https://doi.org/10.22111/ijfs.2024.44621.7855>

- [10] X. Deng, F. Huangfu, *Collaborative variational deep learning for healthcare recommendation*, IEEE Access, **7** (2019), 55679-55688. <https://doi.org/10.1109/ACCESS.2019.2913468>
- [11] A. M. Elkahky, Y. Song, X. He, *A multi-view deep learning approach for cross domain user modeling in recommendation systems*, in Proc. 24th International Conference World Wide Web (WWW), (2015), 278-288. <https://doi.org/10.1145/2736277.2741667>
- [12] I. Fernández-Tobías, I. Cantador, *Exploiting social tags in matrix factorization models for cross-domain collaborative filtering*, In CBRRecSys RecSys, (2014), 34-41. <https://doi.org/10.1145/2645710.2645719>
- [13] S. Gao, H. Luo, D. Chen, S. Li, P. Gallinari, J. Guo. *Cross-domain recommendation via cluster-level latent factor model*, In ECML and PKDD, Springer, (2013), 161-176. [https://doi.org/10.1007/978-3-642-40994-3\\_11](https://doi.org/10.1007/978-3-642-40994-3_11)
- [14] F. M. Harper, J. A. Konstan, *The movieLens datasets: History and context*, ACM Transactions on Interactive Intelligent Systems (TiiS), **5**(4) (2015), 1-9. <https://doi.org/10.1145/2827872>
- [15] J. He, R. Liu, F. Zhuang, F. Lin, C. Niu, Q. He, *A general cross-domain recommendation framework via Bayesian neural network*, Proceedings - IEEE International Conference on Data Mining (ICDM), (2018), 1001-1006. <https://doi.org/10.1109/ICDM.2018.00130>
- [16] M. He, J. Zhang, J. Zhang, *MINDTL: Multiple incomplete domains transfer learning for information recommendation*, China Communications, **14**(11) (2017), 218-236. <https://doi.org/10.1109/CC.2017.8233662>
- [17] H. Izakian, A. Abraham, *Fuzzy c-means and fuzzy swarm for fuzzy clustering problem*, Expert Systems with Applications, **38** (2011), 1835-1838. <https://doi.org/10.1016/j.eswa.2010.07.104>
- [18] H. Li, J. Cao, Y. Gao, Y. Liu, S. Pang, *Cross-domain sequential recommendation via neural process*, Arxiv, (2024). <https://doi.org/10.48550/arXiv.2410.13588>
- [19] D. Liang, R. G. Krishnan, M. D. Hoffman, T. Jebara, *Variational autoencoders for collaborative filtering*, WWW '18: Proceedings of the 2018 World Wide Web Conference, (2018), 689-698. <https://doi.org/10.1145/3178876.3186150>
- [20] T. Man, H. Shen, X. Jin, X. Cheng, *Cross-domain recommendation: An embedding and mapping approach*, Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence Main track, (2017), 2464-2470. <https://doi.org/10.24963/ijcai.2017/343>
- [21] J. Nayak, M. Nanda, K. Nayak, B. Naik, H. S. Behera, *An improved firefly fuzzy c-means (FAFCM) algorithm for clustering real world data sets*, Advanced Computing, Networking and Informatics, **1** (2014), 339-348. [https://doi.org/10.1007/978-3-319-07353-8\\_35](https://doi.org/10.1007/978-3-319-07353-8_35)
- [22] W. Pan, Z. Ming, *Interaction-rich transfer learning for collaborative filtering with heterogeneous user feedback*, IEEE Intelligent Systems, **29**(6) (2014), 48-54. <https://doi.org/10.1109/MIS.2014.59>
- [23] S. Tan, J. Bu, X. Qin, C. Chen, D. Cai. *Cross domain recommendation based on multi-type media fusion*, Neurocomputing, **127** (2014), 124-134. <https://doi.org/10.1016/j.neucom.2013.06.035>
- [24] Y. T. Tsai, C. S. Wuy, H. L. Hsuy, T. Liuy, P. L. Cheny, W. H. Chen, K. T. Liao, *A cross-domain recommender system based on common-sense knowledge bases*, IEEE Conference on Technologies and Applications of Artificial Intelligence (TAAI), (2017), 80-83. <https://doi.org/10.1109/TAAI.2017.8277258>
- [25] M. Vijaikumar, et al., *Neural cross-domain collaborative filtering with shared entities*, Arxiv, (2020). [https://doi.org/10.1007/978-3-030-46130-9\\_17](https://doi.org/10.1007/978-3-030-46130-9_17)
- [26] D. Wang, Y. Liang, D. Xu, X. Feng, R. Guan, *A content based recommender system for computer science publications*, Knowledge-Based Systems, **157** (2018), 1-9. <https://doi.org/10.1016/j.knosys.2018.05.001>
- [27] H. Wang, X. Liu, T. Zhang, *Review-based cross domain recommendation via hyperbolic embedding*, Proceedings of the 2024 IEEE Conference on Artificial Intelligence (AI), (2024), 2345-2356. <https://doi.org/10.1145/3701551.3703486>
- [28] Z. Xu, F. Zhang, W. Wang, H. Liu, X. Kong, *Exploiting trust and usage context for cross-domain recommendation*, IEEE Access, **4** (2016), 2398-2407. <https://doi.org/10.1109/ACCESS.2016.2559530>

- [29] X. S. Yang, *Firefly algorithms for multimodal optimization*, In: Watanabe O., Zeugmann T., (eds.) Stochastic algorithms: Foundations and applications, SAGA 2009, LNCS, **5792** (2009), 169-78. [https://doi.org/10.1007/978-3-642-04944-6\\_16](https://doi.org/10.1007/978-3-642-04944-6_16)
- [30] L. Zhao, S. J. Pan, Q. Yang, *A unified framework of active transfer learning for cross-system recommendation*, Artificial Intelligence, **245** (2017), 38-55. <https://doi.org/10.1016/j.artint.2016.12.001>
- [31] T. Zhou, Y. Wang, X. Chen, *Cross-domain sequential recommendation*, Proceedings of the 2023 ACM Conference on Recommender Systems (RecSys), (2023), 123-130. <https://doi.org/10.1145/3604915.3608752>
- [32] F. Zhu, C. Chen, Y. Wang, G. Liu, X. Zheng, *DTCDR: A framework for dual-target cross-domain recommendation*, Proceedings of the 28th ACM International Conference on Information and Knowledge Management, (2019), 1533-1542. <https://doi.org/10.1145/3357384.3357904>