



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
UNIVERSITY OF PIRAEUS



A Study of Knowledge Transfer techniques in Recommender Systems

by

Angeliki Kostopoulou

Submitted

in partial fulfilment of the requirements for the degree of

Master of Artificial Intelligence

at the

UNIVERSITY OF PIRAEUS

June 2026

University of Piraeus, NCSR "Demokritos". All rights reserved.

Author Angeliki Kostopoulou

II-MSc “Artificial Intelligence”
Month 06, 2026

Certified by Maria Halkidi, Professor

Thesis Supervisor

Certified by Ilias Maglogiannis, Professor

Member of Examination
Committee

Certified by Michalis Filippakis, Professor

Member of Examination
Committee

A Study of Knowledge Transfer techniques in Recommendation Systems

by

Angeliki Kostopoulou

Submitted to the II-MSc “Artificial Intelligence” on June 2026, in partial fulfillment of the requirements for the MSc degree

Abstract

Recommender systems play a fundamental role in modern digital platforms by helping users discover relevant items among vast amounts of available content. Applications such as e-commerce, streaming services, and online media increasingly rely on recommendation algorithms to improve user experience and engagement. However, the effectiveness of traditional recommendation methods is often limited by data sparsity and cold-start problems, particularly when insufficient user-item interactions are available for learning reliable preference patterns. Transfer Learning has emerged as a promising approach for addressing these limitations by enabling knowledge acquired in one domain to be reused in another. In particular, Cross-Domain Recommender Systems aim to transfer useful information from a source domain to improve recommendation quality in a target domain. Most existing approaches, however, assume the existence of overlapping users or items between domains, an assumption that is often unrealistic in practical applications.

This thesis investigates a genre-based cross-domain recommendation framework designed for fully non-overlapping domains. The proposed approach uses the Book-Crossing dataset as the source domain and the MovieLens dataset as the target domain. Since the two domains share neither users nor items, knowledge transfer is achieved through semantic genre information. Book metadata are enriched using external sources, semantically aligned with movie genres using SentenceTransformer embeddings and cosine similarity, and subsequently integrated into a Genre-Aware Matrix Factorisation model. The framework introduces genre embeddings as shared semantic representations that capture relationships between user preferences and content categories. These embeddings are first learned in the source domain and then transferred to the target domain through pretrained initialization, enabling cross-domain knowledge transfer without requiring entity overlap. The transferred representations are subsequently fine-tuned on movie ratings and evaluated under progressively increasing sparsity conditions. Experimental evaluation is conducted using Singular Value Decomposition (SVD), Matrix Factorisation (MF), Genre-Aware Matrix Factorisation, and the proposed Cross-Domain Transfer Learning model. Performance is assessed using Mean Absolute Error (MAE) under 10-fold cross-validation and incremental training availability. The results demonstrate that incorporating semantic genre information significantly improves recommendation accuracy compared with traditional collaborative filtering methods. Furthermore, the findings indicate that semantic representations learned from an auxiliary domain can be successfully transferred to support recommendation tasks in a target domain. The proposed framework contributes to the development of more flexible recommendation systems capable of operating effectively under sparse-data conditions, while demonstrating the potential of semantic transfer learning in fully non-overlapping recommendation environments.

Thesis Supervisor: Maria Halkidi

Title: Professor

Μελέτη τεχνικών μεταφοράς γνώσης σε συστήματα συστάσεων

από

Αγγελική Κωστοπούλου

Υποβλήθηκε στο ΔΠΜΣ «Τεχνητή Νοημοσύνη» τον Ιούνιο 2026 ως υποχρέωση για την λήψη Μεταπτυχιακού Διπλώματος Σπουδών

Περίληψη

Τα συστήματα συστάσεων αποτελούν βασικό στοιχείο σύγχρονων ψηφιακών πλατφορμών, καθώς υποστηρίζουν τους χρήστες στην αναζήτηση σχετικού περιεχομένου μέσα σε μεγάλο όγκο προϊόντων, υπηρεσιών και πληροφοριών. Εφαρμογές ηλεκτρονικού εμπορίου, υπηρεσίες streaming και διαδικτυακά μέσα αξιοποιούν αλγορίθμους συστάσεων για την εξατομίκευση της εμπειρίας χρήστη. Ωστόσο, η αποτελεσματικότητα των παραδοσιακών μεθόδων περιορίζεται σημαντικά σε περιβάλλοντα αραιών δεδομένων και ψυχρής εκκίνησης, όπου οι διαθέσιμες αλληλεπιδράσεις χρηστών-αντικειμένων δεν επαρκούν για την αξιόπιστη εκμάθηση προτιμήσεων.

Η παρούσα διπλωματική εργασία διερευνά ένα πλαίσιο διατομεακών συστάσεων βασισμένο στη Μεταφορά Μάθησης, με στόχο τη βελτίωση των προβλέψεων σε συνθήκες περιορισμένης διαθεσιμότητας δεδομένων. Το προτεινόμενο μοντέλο χρησιμοποιεί το σύνολο δεδομένων Book-Crossing ως πεδίο πηγής και το MovieLens ως πεδίο στόχο. Η βασική πρόκληση είναι ότι τα δύο πεδία δεν μοιράζονται ούτε κοινούς χρήστες ούτε κοινά αντικείμενα, γεγονός που δεν επιτρέπει την άμεση μεταφορά παραγόντων χρηστών ή αντικειμένων.

Για την αντιμετώπιση του προβλήματος, η εργασία αξιοποιεί τα είδη περιεχομένου ως κοινή σημασιολογική γέφυρα μεταξύ βιβλίων και ταινιών. Αρχικά, τα μεταδεδομένα των βιβλίων εμπλουτίζονται μέσω εξωτερικών πηγών και στη συνέχεια αντιστοιχίζονται σημασιολογικά με τα είδη των ταινιών, με χρήση SentenceTransformer embeddings και cosine similarity. Οι αντιστοιχισμένες κατηγορίες ενσωματώνονται σε ένα Genre-Aware Matrix Factorisation μοντέλο, το οποίο μαθαίνει αναπαραστάσεις χρηστών, αντικειμένων και ειδών. Οι genre embeddings εκπαιδεύονται αρχικά στο πεδίο των βιβλίων και μεταφέρονται ως προεκπαιδευμένες παράμετροι στο πεδίο των ταινιών, όπου προσαρμόζονται μέσω fine-tuning.

Η αξιολόγηση περιλαμβάνει τέσσερα μοντέλα: Singular Value Decomposition, Matrix Factorisation με Stochastic Gradient Descent, Genre-Aware Matrix Factorisation και το προτεινόμενο Cross-Domain Transfer Learning μοντέλο. Η σύγκριση πραγματοποιείται με Mean Absolute Error, μέσω 10-fold cross-validation και σταδιακής αύξησης της διαθέσιμης εκπαίδευσης. Τα αποτελέσματα δείχνουν ότι η ενσωμάτωση σημασιολογικής πληροφορίας βελτιώνει ουσιαστικά την ακρίβεια σε σχέση με τις κλασικές συνεργατικές μεθόδους, ενώ το προτεινόμενο διατομεακό μοντέλο παρουσιάζει την καλύτερη συνολική συμπεριφορά υπό συνθήκες αυξανόμενης αραιότητας.

Η εργασία καταδεικνύει ότι η σημασιολογική μεταφορά γνώσης μέσω ειδών μπορεί να υποστηρίξει αποτελεσματικά συστήματα συστάσεων σε πλήρως μη επικαλυπτόμενα πεδία.

Επιβλέπων Καθηγητής: Μαρία Χαλκίδη

Τίτλος: Καθηγήτρια

Acknowledgments

I would like to express my sincere gratitude to Associate Professor Maria Halkidi of the Department of Digital Systems at the University of Piraeus for her valuable guidance, constructive feedback, and continuous support throughout the development of this thesis. Her scientific expertise, insightful comments, and encouragement were instrumental in shaping both the direction and quality of this research. Her willingness to provide guidance whenever needed and her commitment to academic excellence greatly contributed to the successful completion of this work.

I would also like to thank the faculty members and academic staff of the Master's Program in Artificial Intelligence, jointly offered by the University of Piraeus and the National Centre for Scientific Research "Demokritos". Their teaching, knowledge, and support have provided the academic foundation necessary for conducting this research.

Finally, I would like to express my deepest appreciation to my family and close friends for their patience, understanding, and unwavering encouragement throughout my studies. Their continuous support and confidence in my efforts have been a constant source of motivation during this demanding yet rewarding journey.

Contents

LIST OF FIGURES.....	8
LIST OF ALGORITHMS	9
LIST OF TABLES.....	10
Chapter 1: Introduction to Recommender Systems and Transfer Learning.....	11
1.1 Overview of Recommender Systems.....	11
1.2 Motivation and Problem Statement.....	12
1.3 Thesis Aim and Organization.....	13
Chapter 2 Transfer Learning in Recommender Systems	15
2.1 Recommender Systems: Concepts and Taxonomy	15
2.2 Fundamentals of Transfer Learning.....	16
2.3 Transfer Learning in Recommender Systems.....	17
Chapter 3 Theoretical Foundation of the Proposed Model.....	20
3.1 Problem Formulation and Scope	20
3.2 Theoretical Foundation of the Proposed Genre-Based Transfer Model	23
3.2.1 Semantic transfer through shared genre structure	23
3.2.2 Latent representation of users, items, and genres.....	24
3.2.3 Genre-aware rating prediction model	25
3.2.4 Cross-domain transfer at the parameter level	25
3.2.5 Theoretical assumptions, expected benefits, and limitations.....	26
3.3 Mathematical Formulation of the Model	27
3.3.1 Notation and Problem Definition	27
3.3.2 Latent Factor Representation.....	28
3.3.3 Rating Prediction Function	28
3.3.4 Optimisation Objective	28
3.3.5 Source and Target Domain Learning.....	29
3.3.6 Interpretation of the Model	29
3.4 Cross-Domain Transfer Mechanism	30
3.4.1 Transfer of Semantic Representations.....	30
3.4.2 Role of Genre Embeddings as Priors.....	30
3.4.3 Adaptation to the Target Domain.....	31
3.5 Theoretical Advantages and Limitations of the Proposed Approach	31
3.5.1 Theoretical Advantages.....	31
3.5.2 Limitations and Assumptions	32
Chapter 4 Proposed Framework, Implementation and Evaluation	33
4.1 System Overview and Architecture	33
4.2 Data Preparation and Semantic Alignment	34
4.3 Recommendation Models.....	36
4.3.1 Baseline Collaborative Filtering Models.....	37
4.3.2 Genre-Aware Recommendation Model	37

4.3.3 Cross-Domain Transfer Learning Model	39
4.4 Experimental Methodology	40
4.4.1 Evaluation Process	40
4.4.2 Hyperparameter Configuration	41
4.4.3 Evaluation Metric.....	42
Chapter 5 Experimental Evaluation	43
5.1 Data Preparation and Processing	43
5.1.1 Dataset Overview	43
5.1.2 Cleaning and Validation of Raw Data.....	43
5.1.3 Genre Enrichment of the Book Domain.....	44
5.1.4 Semantic Alignment Between Domains.....	45
5.1.5 Integration of Ratings and Genre Information	46
5.1.6 Movie Genre Processing	47
5.1.7 Indexing and Matrix Representation	48
5.1.8 Construction of the User–Item Matrix.....	48
5.1.9 Preparation for Singular Value Decomposition	49
5.1.10 Evaluation strategy	49
5.2 Model Implementation and Deep Analytical Evaluation	50
5.2.1 Singular Value Decomposition (SVD)	50
5.2.2 Matrix Factorisation with Stochastic Gradient Descent	51
5.2.3 Genre-Aware Matrix Factorisation	53
5.2.4 Cross-Domain Transfer Learning Model	56
5.2.5 Comparative Interpretation	59
Chapter 6 Conclusions and Future Work.....	61
6.1 Conclusions	61
6.2 Limitations.....	61
6.3 Future Work	62
REFERENCES.....	63

LIST OF FIGURES

Figure 1.1: Process of general recommendation system models, including collaborative filtering, content-based filtering, and hybrid recommendation[20].

Figure 1.2 : Figure 1.2: Levels of cross-domain recommendation[19].

Figure 2.1: Matrix factorization of a user–item interaction matrix into latent user and item representations[21].

Figure 2.2: Classification of Transfer Learning Methods[22].

Figure 2.3: Cross-domain recommendation using auxiliary domain knowledge to improve recommendations in a sparse target domain.[21]

Figure 3.1.1 : Cold-Start Problem in Collaborative Filtering[21].

Figure 3.1.2 :Cross-Domain Recommendation with Auxiliary Domain Ratings[21].

Figure 3.1.3 : Semantic Knowledge Transfer via Shared Attributes[21]

Figure 3.2 : Sparse target domain supported by a richer auxiliary domain[21].

Figure 3.2.1 : Incremental Training and Repetition in Extended Cross-Validation[21].

Figure 4.1: Overall architecture of the proposed genre-based cross-domain transfer learning framework.

Figure 4.2: Data preparation and semantic alignment process for the proposed cross-domain recommendation framework.

Figure 4.3.2: Training workflow of the Genre-Aware Matrix Factorisation model.

Figure 4.3.3: Transfer of genre embeddings from the Book-Crossing domain to the MovieLens domain.

Figure 4.4.1: Evaluation procedure based on 10-fold cross-validation and incremental training availability.

Figure 5.2.2.1: illustrates the evolution of MAE as the percentage of available training data increases.

Figure 5.2.3.1: MAE performance of the Genre-Aware Matrix Factorisation model under increasing training availability.

Figure 5.2.4.1: MAE performance of the Cross-Domain Transfer Learning model under increasing training availability.

Figure 5.2.5: Comparative MAE performance of all recommendation models under increasing training data availability.

LIST OF ALGORITHMS

Algorithm 5.1 : Genre Enrichment Procedure

Algorithm 5.2 : Semantic Genre Mapping[23,24]

Algorithm 5.3 : Cross-Domain Transfer Learning Procedure[25,26]

LIST OF TABLES

Table 4.2.1: Examples of genre labels retrieved during the genre enrichment process.

Table 4.2.2: Examples of semantic mappings between book genres and movie genres.

Table 4.3: Recommendation models used in the experimental evaluation.

Table 4.4.3: Illustrative example of Mean Absolute Error (MAE) calculation.

Table 5.1 Dataset Statistics

Table 5.2 Genre Enrichment Results

Table 5.3 Distribution of Mapped Movie Genres

Table 5.4 Example of the Unified Dataset

Table 5.6 SVD Performance under Increasing Training Availability

Table 5.7 Matrix Factorisation Performance under Increasing Training Availability

Table 5.8 Genre-Aware Matrix Factorisation Performance

Table 5.9 Cross-Domain Transfer Learning Performance

Chapter 1: Introduction to Recommender Systems and Transfer Learning

1.1 Overview of Recommender Systems

In modern digital environments, users are constantly exposed to vast amounts of content, products, and services. From e-commerce platforms to streaming applications and online media, the number of available choices has grown to a level that makes manual exploration increasingly difficult. In this context, recommender systems play a crucial role by guiding users toward content that matches their preferences, thereby reducing information overload and supporting decision-making[1].

Recommender systems operate by analyzing patterns in user behavior, preferences, and past interactions in order to predict items that a user is likely to find relevant[1,2]. Their output appears in many forms, including product recommendations in online stores, personalized playlists in music platforms, and suggested videos in streaming services. The widespread adoption of these systems highlights their importance in shaping user experience and increasing engagement[2].

Major technology companies such as Netflix, Amazon, and Spotify rely heavily on recommendation algorithms to personalize their services. A large proportion of user activity on these platforms is driven by automated suggestions, demonstrating the practical and economic significance of recommender systems[3]. In real-world applications, these systems often combine multiple techniques in order to achieve better performance. Hybrid approaches, which integrate collaborative and content-based methods, are commonly used to improve accuracy and robustness while overcoming the limitations of individual techniques[2].

Despite their success, recommender systems face important challenges when applied in realistic settings. In particular, the availability of user interaction data is often limited, which affects the ability of models to learn reliable preference patterns[4]. These limitations motivate the need for more advanced approaches capable of operating effectively under data-scarce conditions. The present thesis builds on this motivation by focusing on methods that improve recommendation performance through the use of cross-domain information.

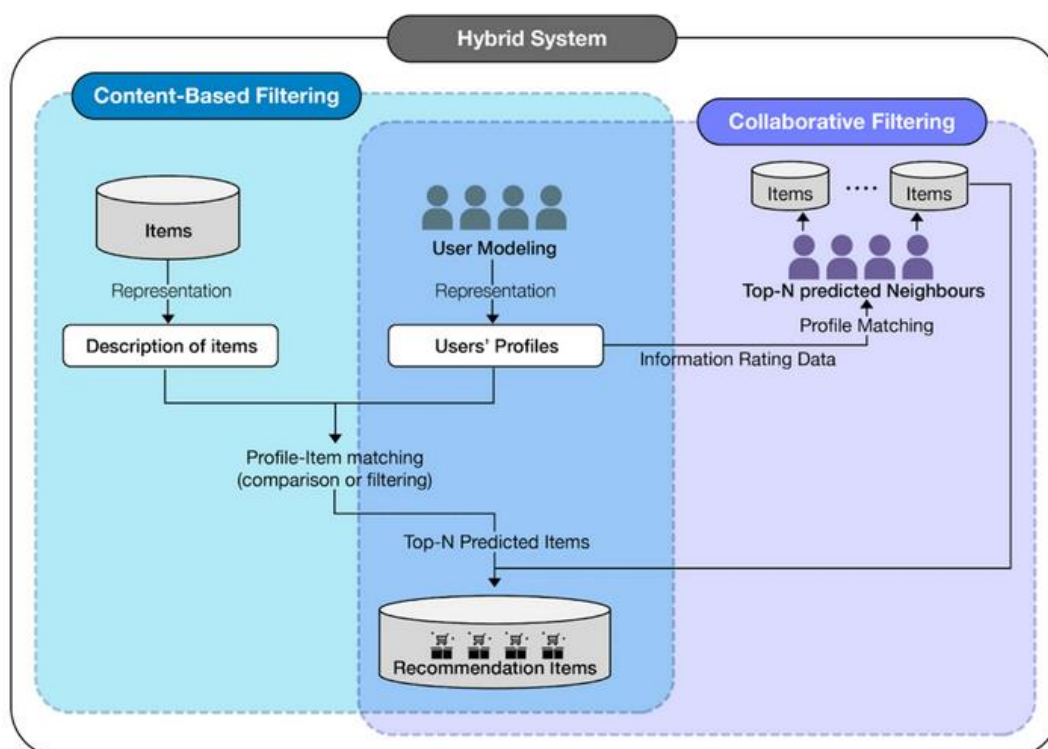


Figure 1.1: Process of general recommendation system models, including collaborative filtering, content-based filtering, and hybrid recommendation.

1.2 Motivation and Problem Statement

Traditional recommendation algorithms, while effective in well populated systems, degrade sharply in scenarios where user–item interactions are minimal or non existent. In newly launched platforms or for users/items that are new to an existing platform, the lack of historical data significantly reduces recommendation quality and the overall utility of the system. This limitation has motivated research into techniques that can transfer knowledge from auxiliary data sources to improve recommendations in the data sparse target scenario[5]. In particular, cross domain recommender systems have emerged as a promising approach to address sparsity and cold-start issues[5,6]. The core idea is to leverage information from a different but related domain to inform recommendations in the target domain where data is scarce. For example, a user’s preferences observed in a domain like music or books could be used to make better recommendations in another domain like movies or games, assuming there are latent correspondences between the domains[6].

Most existing cross-domain recommendation methods, however, assume that the source and target domains share some common reference points typically, overlapping users or overlapping items[6,7]. In practice, such overlap is often unavailable due to privacy or context separation; users of a book recommendation service might not be users of a movie streaming service, for instance, and the item catalogs naturally differ. When no users or items are shared, directly transferring learned user or item factors from one domain to another is not feasible because the latent feature spaces have no anchor for alignment[7]. This thesis is motivated by this fully non overlapping cross domain scenario, which is both realistic and challenging of how can we improve recommendations in a target domain (movies) by using data from a source domain (books) when the two domains do not share users or items.

Several studies have investigated the use of auxiliary information from external domains to alleviate cold-start conditions in recommender systems. Enrich et al. (2013) demonstrated that cross-domain knowledge can improve recommendation quality when limited information is available in the target domain.

Our proposed solution is to exploit shared high level semantic information as the bridge between domains, instead of relying on shared users or items. In particular, we focus on genres as a form of domain independent side information that carries meaning across domains. The underlying intuition is that while individual users and specific items differ between a book platform and a movie platform, the genre categories have a consistent semantic interpretation in both contexts. A user who enjoys science fiction novels is likely to enjoy science fiction films as well, even if they are a new user in the movie domain. By modeling a user’s genre preferences in the source domain and transferring that knowledge to the target domain, we can obtain informative priors for new users or items in the target system. In other words, genre affinity serves as a latent conduit for transfer learning across domains.

Concretely, the research investigates a cross domain recommendation framework where genre based embeddings are learned from an auxiliary domain and then used to initialize and inform a model in the target domain. The approach does not require any overlapping entities. Instead it assumes the domains share a defined set of genre or category labels. During training on the source domain (book), the system learns latent factor representations not only for users and items but also for each genre. These genre vectors capture how different genres relate to user preferences and item characteristics as observed in the source domain’s data. We then transfer the learned genre representations into the target domain’s (movie) model as prior knowledge. Intuitively, if the model has learned what the genre means in the context of book ratings, it starts with an informed notion of the genre when making movie recommendations, even for cold start users or movies. This thesis posits that such semantic transfer can significantly improve the quality of recommendations under extreme sparsity.

By using genres as a semantic bridge, we address the cold-start problem in a novel way. Instead of waiting to accumulate user interactions in the target domain, we prepare the recommender with cross domain knowledge about content categories. The expectation is that this additional information will guide the recommender’s predictions until sufficient target specific data accrues. In summary, the problem addressed in this work is how to design a recommender system that remains effective in the face of sparse data and no overlapping entities, by transferring abstract knowledge through shared semantic features. Solving this problem contributes to more flexible and generalizable recommendation frameworks applicable to real-world multi domain environments, where leveraging any available side information can make the difference between a cold-start failure and successful personalization.

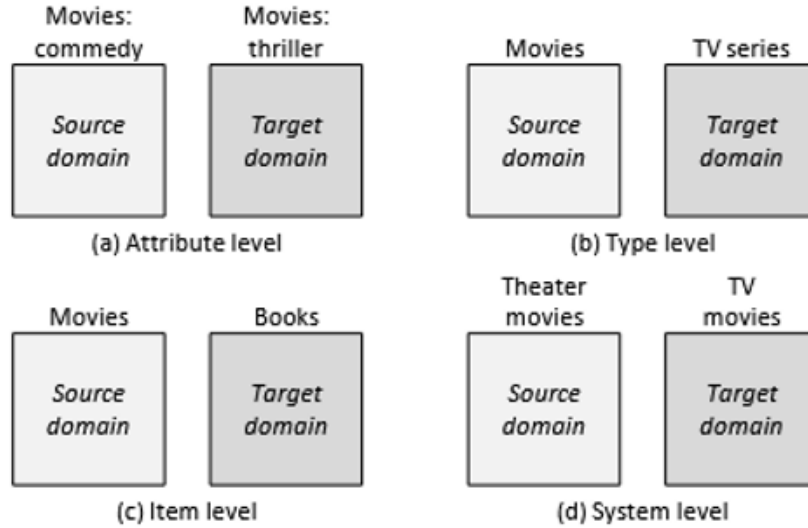


Figure 1.2: Levels of cross-domain recommendation.

1.3 Thesis Aim and Organization

The principal aim of this thesis is to investigate and develop a cross domain recommender system that applies transfer learning via genre based embeddings to enhance personalization in data sparse and cold-start environments. In essence, the study asks whether abstract, high level knowledge captured in the form of genre preference patterns can be learned in one domain and effectively transferred to another domain to alleviate sparsity. The target scenario involves a source domain (books) and a target domain (movies) that do not share users or items. By learning semantic correspondences between the two domains, the proposed approach seeks to overcome one of the fundamental limitations of traditional single domain recommenders, the dependency on large amounts of domain specific interaction data for effective personalization.

The thesis is organized into seven chapters, each addressing a distinct aspect of the research problem.

Chapter 1 introduces the research area of recommender systems and transfer learning. It discusses the challenges of data sparsity and cold-start conditions, presents the motivation behind cross-domain recommendation, and defines the objectives and scope of the thesis.

Chapter 2 provides the theoretical background and literature review. It introduces the fundamental concepts of recommender systems, including collaborative filtering, content-based filtering, and hybrid approaches, and reviews transfer learning techniques and their application in recommender systems. The chapter also

discusses existing cross-domain recommendation methods and identifies the research gap addressed in this work.

Chapter 3 presents the proposed genre-based cross-domain transfer learning framework. It formulates the recommendation problem, introduces the theoretical foundations of semantic transfer through genre information, and develops the mathematical formulation of the proposed model. The chapter also explains the transfer mechanism, theoretical assumptions, advantages, and limitations of the approach.

Chapter 4 describes the overall system architecture and experimental methodology. It explains the datasets used in the study, the semantic alignment process between domains, the recommendation models considered, and the evaluation framework adopted for the experimental analysis.

Chapter 5 presents the experimental evaluation of the proposed framework. The chapter describes the data preparation process, genre enrichment and semantic mapping procedures, model implementation, and experimental results. It also provides a comparative analysis of the evaluated recommendation models under different levels of data sparsity.

Chapter 6 concludes the thesis by summarising the main findings and contributions of the research. It discusses the limitations of the proposed framework and outlines potential directions for future work in cross-domain recommender systems and transfer learning.

Chapter 2 Transfer Learning in Recommender Systems

2.1 Recommender Systems: Concepts and Taxonomy

Recommender systems are software tools that suggest items (movies, books, music, etc.) to users based on their preferences and behavior. They operate by analyzing past user–item interactions (ratings, clicks) to predict how a user might rate or enjoy new items. Three classic paradigms are widely used in modern recommenders: collaborative filtering, content based filtering, and hybrid approaches[1,2].

Collaborative Filtering (CF): This approach assumes that users with similar past tastes will exhibit similar future preferences[2]. Collaborative filtering relies only on user–item interaction data like ratings and ignores item descriptions. Two main CF methods exist: *memory-based* (neighborhood) models and *model-based* models. In memory-based CF, the system directly computes similarities between users or items in the rating matrix to make predictions. For example, in **user-based CF**, a target user’s neighborhood of most similar users is found (using Pearson or cosine similarity), and the target’s rating for an item is predicted as a weighted average of those neighbors’ ratings. Conversely, **item-based CF** finds items similar to those the user has rated and bases recommendations on those similarities. Memory-based CF is simple and easy to interpret but can struggle with sparse or very large datasets [2].

Model-based CF learns a **latent factor model** from the rating data, with **matrix factorization** being a prominent example. Matrix factorization (e.g. using Singular Value Decomposition) decomposes the large user–item rating matrix into the product of a user-factor matrix and an item-factor matrix[4]. Each user and item is thus represented by a dense vector in a low-dimensional latent feature space, such that the **predicted rating** is the dot product of a user’s and item’s feature vectors. This captures hidden characteristics: for instance, one latent dimension may correspond to a user’s affinity for the “action” genre, and another to “romance.” The model infers these factors from patterns in the data. Matrix factorization proved highly effective in the Netflix Prize, outperforming classic nearest-neighbor methods.

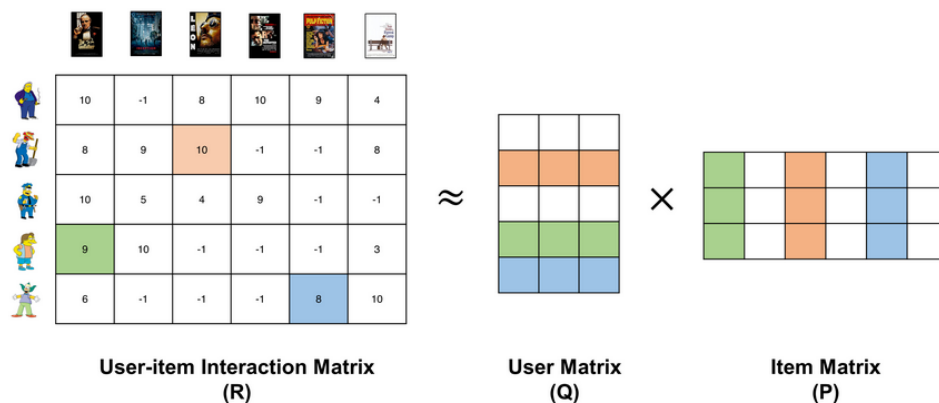


Figure 2.1: Matrix factorization of a user–item interaction matrix into latent user and item representations.

Content-Based Filtering (CBF): Content-based recommenders utilize descriptive **features of items** and a user’s own interaction history to generate suggestions. The system builds a **user profile**, a structured representation of the user’s preferences, typically by analyzing the attributes of items the user liked. For example, in a movie recommender, item features might include genre, director, cast, or keywords. If a user has highly rated several sci-fi movies, the profile will emphasize features like “science fiction” or “space setting.” New candidate movies are then scored by similarity to this profile. Commonly, content-based algorithms represent items and user profiles in a feature vector space and use similarity metrics (e.g. cosine similarity) to recommend items most similar to the user’s profile. An advantage of CBF is that it can recommend *new items* that have no ratings yet (the “cold start” item problem), as long as descriptive content is available. However, content-based methods may overly specialize to a

user's known interests, failing to introduce novel or diverse items. They also require rich item metadata.

Hybrid Approaches: To overcome the limitations of pure CF or CBF, many real-world systems use **hybrid recommender systems** that combine both techniques. Hybrids can merge CF and CBF in various ways: for example, by **implementing both and blending their outputs** (score combination or rank aggregation), or by **augmenting one with components of the other** (e.g. a primarily CF model that uses item content to address cold-start cases). Another strategy is a **unified model** that integrates collaborative and content signals into a single learning algorithm. Hybrid systems can achieve improved accuracy and coverage – leveraging CF's ability to find community trends as well as CBF's ability to use item descriptors. They tend to mitigate problems like *cold start* (CF struggles when a new user or item lacks interaction data, which CBF can handle using attributes) and *sparsity* (CBF can still operate when few users have rated an item, by relying on item content). For instance, Netflix's recommender incorporates both collaborative signals (viewing patterns) and content (genre, title) to personalize results. Such a system can use a CF module to draw on **community preferences** and a CBF module to ensure recommendations align with the user's stated likes or item metadata [2].

2.2 Fundamentals of Transfer Learning

Transfer learning is a machine learning paradigm in which knowledge obtained in one context is applied to improve learning in another, related context[5]. In traditional machine learning, we assume training and test data come from the same distribution and feature space. Transfer learning relaxes this assumption: the source domain (where a model is initially learned) and the target domain (where it is applied) can differ in data distribution, features, or task. By transferring the patterns, representations, or parameters learned from a source task with abundant data to a target task with limited data, transfer learning aims to overcome problems like data sparsity and costly labeled data in the target domain [5].

Conventional machine learning techniques typically assume that training and testing data are drawn from the same feature space and follow the same distribution. However, this assumption often does not hold in real world applications. For instance, a sentiment analysis model trained on product reviews might perform poorly when directly applied to movie reviews due to differences in vocabulary and context. Transfer learning mitigates this limitation by allowing models to apply knowledge gained from one domain (source) to improve performance in another (target).

A foundational classification of transfer learning includes three major types: inductive, transductive, and unsupervised [5,8]. In inductive transfer learning, the source and target tasks are different, and the target domain has at least some labeled data. Transductive transfer learning refers to scenarios where the source and target domains differ, but the tasks remain the same, typically without labeled data in the target domain. Unsupervised transfer learning, by contrast, applies when neither domain has labels, and the tasks are unsupervised, such as clustering or dimensionality reduction .

Beyond these distinctions, transfer learning strategies are also categorized based on what is transferred. These include instance based approaches like reusing or reweighting source data, feature representation transfer like learning shared latent spaces, parameter transfer like sharing model components like weights or priors, and relational knowledge transfer like leveraging structured interdependencies between entities across domains[8] . Each strategy targets different challenges in data alignment, representation, or generalization.

In recent years, transfer learning has become increasingly prevalent in domains such as natural language processing, computer vision, and recommender systems. In recommender systems, it is particularly useful for handling cold-start problems or sparse interaction data, as it enables the model to leverage auxiliary knowledge from other domains[6,7]. For example, data from user

interactions with books can inform recommendations for movies, provided the system can identify shared patterns such as genre preferences or semantic similarities .

The rising importance of transfer learning in machine learning research is also evidenced by the growing body of literature and its incorporation into modern frameworks. In recommender systems, it enables more robust personalization across platforms and facilitates generalization even in disconnected environments making it a cornerstone for contemporary cross domain recommendation strategies [5,8]

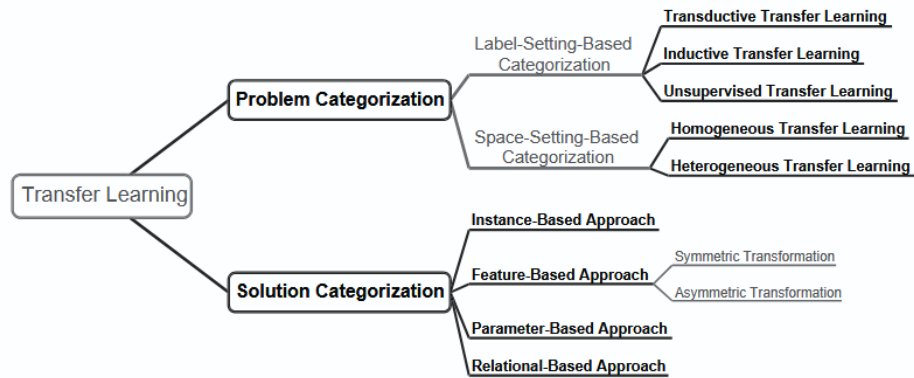


Figure 2.2: Classification of Transfer Learning Methods

2.3 Transfer Learning in Recommender Systems

Among the early approaches addressing sparsity through cross-domain recommendation, Enrich et al. (2013) proposed a framework combining cross-domain collaborative filtering with tag information for cold-start management. Their results demonstrated that information originating from an auxiliary domain can significantly improve recommendation quality in sparse environments.

Integrating transfer learning into recommender systems has given rise to cross domain recommendation, where knowledge from one domain (the source) is used to improve recommendations in another domain (the target)[7,9]. This approach directly addresses the data sparsity and cold-start problem. For example, a user new to a music service might still get good recommendations if the system can leverage that user's history from a movie service . Broadly, transfer learning based recommenders seek to find common latent features or mappings that allow transferring preferences across domains[7,9]. Recent surveys categorize cross domain recommender systems by domain relatedness, overlap scenario, and knowledge transfer technique. Here we outline key concepts and methods in transfer-learning recommenders, focusing especially on cross domain collaborative filtering and the use of genre information as transferable knowledge.

Cross-Domain Recommendation (CDR): In cross domain recommendation, we use one domain as an auxiliary source to help generate suggestions in another domain. The fundamental assumption is that there exist latent connections between user preferences in different domains. For example, a person's taste in books may be related to their taste in movies , a user who enjoys novels might also enjoy sci-fi films. CDR techniques aim to exploit these connections. A common scenario is the cold-start user: a user with no history in the target domain can still receive recommendations by importing their behavior from a source domain. Approaches to CDR vary based on what is shared between domains[9]. If the domains share users (the same people use both services) or share items (the same products sold in two shops), algorithms can directly link entities. If there is no overlap, one must rely on higher level feature bridges like common genres, tags, or attributes. CDR methods can be grouped by the amount of overlap:

- **complete overlap** (same user IDs in both domains),
- **partial overlap** (some users in common),
- **no overlap** (distinct user sets, requiring a pure feature-based transfer).

An example of a cross-domain system is one that recommends *books* to users based on their movie ratings, using genre as the bridge. If a user rates many “fantasy” movies highly, the system can suggest fantasy novels.

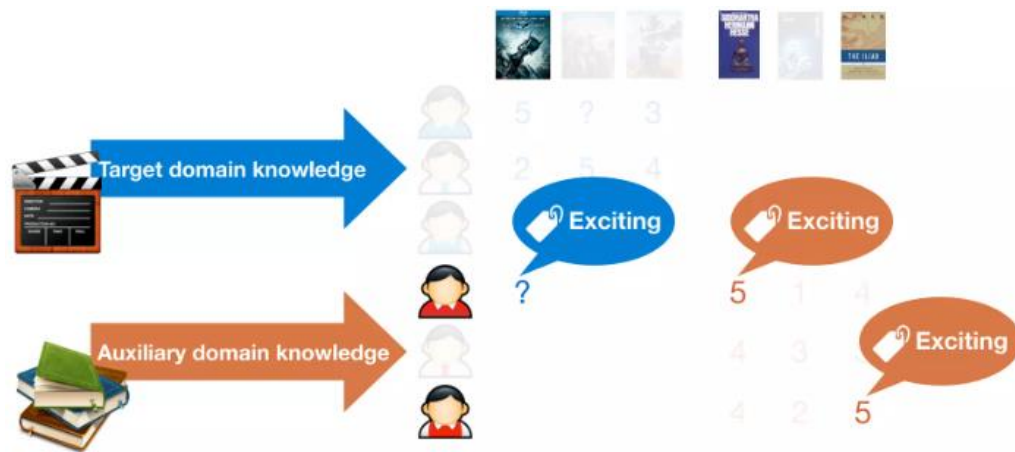


Figure 2.3: Cross-domain recommendation using auxiliary domain knowledge to improve recommendations in a sparse target domain.

Latent Factor Transfer: Many cross domain recommenders extend matrix factorization to multiple domains. A representative approach is to perform a collective matrix factorization jointly factorizing the user-item rating matrices of two domains with some shared factors[10]. For instance, suppose we have a user-movie matrix and a user-book matrix. A collective factorization might force the user latent feature matrix to be the same in both decompositions, while allowing separate item feature matrices for movies and books. This way, a user’s learned preferences (latent factors) transfer from the movie domain to inform recommendations in the book domain. Early work by Singh and Gordon (2008) on collective matrix factorization did exactly this: sharing user factors across recommendation domains. Similarly, tensor factorization approaches can incorporate domain as an additional dimension and discover latent factors that span domains. These factorization techniques essentially create a common latent space aligning users (or items) across domains.

Mapping Functions and Domain Adaptation: An important class of cross-domain recommendation techniques focuses on learning mapping functions that can transfer knowledge between domains by aligning their latent feature spaces[11]. These methods are particularly relevant when the domains do not share items but have a subset of overlapping users, which can be used as anchors for learning the transformation.

Zero-Shot and Few-Shot Recommendation: A growing interest is in zero-shot recommenders, where an entire domain is new (no prior interactions) and one must rely completely on transferred knowledge. This intersects with cold-start problems and often leverages content based transfer. For example, recent works use pretrained language models to encode item descriptions in multiple domains, thereby allowing a system to recommend an item in a new domain by understanding its description relative to a user’s profile learned elsewhere. In general, as foundation models become available, zero-shot cross-domain recommendation (where a model is trained on many domains and tested on a new one) has become feasible[12].

Incorporating Side Information (Genres and Tags): One effective transfer strategy is to use side information as the bridge between domains[13]. For example, genre preferences are a high-

level summary of a user's taste that can be shared across domains. A user who loves "romance" might exhibit that preference in movies, books, and music. If we know a user's favorite movie genres, we can cold-start recommend books in those genres. This was demonstrated by early cross-domain studies: a user's genre vector from one domain can improve recommendations in another. The drawback is that if tags/genres are too general (many items share them), the recommendations might be obvious or not specific, but if they are too sparse, they may not sufficiently bridge the gap. Still, semantic information like genres is invaluable for an intuitive transfer: it provides a human-interpretable feature for why an item in domain A might appeal to a user based on their domain.

Example – Genre-Based Cross-Domain Transfer: Suppose we have a Movie recommender and a Book recommender, and we notice that users who enjoy fantasy and adventure movies also tend to rate fantasy novels highly. A genre-based cross-domain system would transfer a user's genre affinity profile from the movie domain to the book domain. If a new user has no book ratings but has rated many fantasy movies, the system can recommend popular fantasy books to that user. This approach was explored by combining user-tag matrices: building a user-genre matrix for movies and for books, and using the movie matrix to fill in the user's preferences in the book matrix. One challenge is that genres between domains might not align one to one (the granularity or definitions can differ), but a curated mapping (like matching "fantasy" books to "fantasy adventure" films) or a vector space embedding of genres can help. Empirical studies confirm that cross domain collaborative filtering can significantly improve accuracy when user overlaps are present, and even when there are no overlapping users, a shared feature like genre can enable knowledge transfer that a single-domain model would[13].

Recent Trends: Modern cross-domain recommender research has embraced deep learning and graph-based methods[14,15]. For example, graph neural networks have been used to connect domains by constructing a giant graph of users and items across domains and learning embeddings that naturally encode multi-domain connectivity. Another trend is using large pretrained models (e.g. BERT for texts, or image models) to create domain agnostic representations of items, so that items from different domains can be compared in a shared feature space. This aligns with the success of foundation models: e.g., a CLIP like model could embed both movie posters and book covers in the same space, helping identify similar themes.

In summary, transfer learning has become a cornerstone of advanced recommender systems, enabling what we call multi domain personalization. By transferring user preferences across domains (especially through latent factors or content features like genres), systems can provide better recommendations in data-scarce situations and give newly launched services a "warm start." However, implementing cross-domain recommenders also brings challenges: how to effectively align different domains' data, how to avoid negative transfer (where source knowledge misleads the target), and how to interpret recommendations across domains. These open issues define an active research frontier, which we highlight in the following section when discussing the gaps that our work aims to address.

Chapter 3 Theoretical Foundation of the Proposed Model

3.1 Problem Formulation and Scope

Although previous studies have demonstrated the effectiveness of cross-domain information for cold-start management, most approaches rely on user-generated tags, overlapping entities, or additional metadata. For example, Enrich et al. (2013) utilised collaborative filtering and tags to transfer knowledge across domains. In contrast, the framework proposed in this thesis investigates whether semantic genre representations alone can support knowledge transfer between domains that share neither users nor items.

Modern recommender systems constitute essential components of digital platforms, enabling personalised content discovery within vast catalogues of movies, books, music, and products. Their core mechanism relies on analysing historical user-item interactions to infer latent preferences and to make predictions about future behaviour. Collaborative Filtering and latent factor models, such as matrix factorisation, have demonstrated substantial effectiveness in modelling these patterns [16]. Their performance, however, depends critically on the density of the available data. When user-item interactions are limited, the system struggles to learn reliable representations, resulting in high prediction error. This phenomenon, commonly referred to as the data sparsity problem, remains one of the most persistent obstacles in recommender systems [17].

Data sparsity is not an exceptional case but rather an inherent characteristic of real world recommendation systems. Users typically rate only a small fraction of available items, while new users and new items continuously enter the system. Consequently, recommender models are frequently required to operate under cold-start or near cold-start conditions, where limited data prevents reliable learning. This challenge is particularly evident in the movie recommendation domain, where even large scale datasets such as MovieLens contain only a tiny fraction of all possible user-item interactions.



Figure 3.1.1 : Cold-Start Problem in Collaborative Filtering

Cross domain recommendation has emerged as a promising strategy to alleviate sparsity by leveraging auxiliary information from an external domain. The central assumption is that user preferences exhibit some form of consistency across domains, allowing knowledge learned in a source domain to improve recommendations in a target domain. However, cross-domain recommendation becomes significantly more challenging when the two domains do not share

users or items. In such non-overlapping scenarios, direct alignment of collaborative signals is impossible, and knowledge transfer must rely on higher level semantic structures.

Cross domain recommendation must address two core questions: what knowledge should be transferred and how this transfer can be achieved.

A crucial distinction among Cross domain recommendation scenarios lies in the relational structure between the domains. Some works assume that the two domains share either users or items, enabling direct alignment of latent factors. Others rely on shared interaction structures, co occurrence patterns, or common behavioural signals. However, the scenario studied in this diploma corresponds to one of the most challenging categories identified in the Cross domain recommendation taxonomy, the full non overlap scenario, in which there are no shared users and no shared items between the source and target domains. This is considered one of the most difficult types of cross-domain transfer, because it eliminates any possibility of direct collaborative alignment, requiring instead the identification of higher-level transferable structures [18].

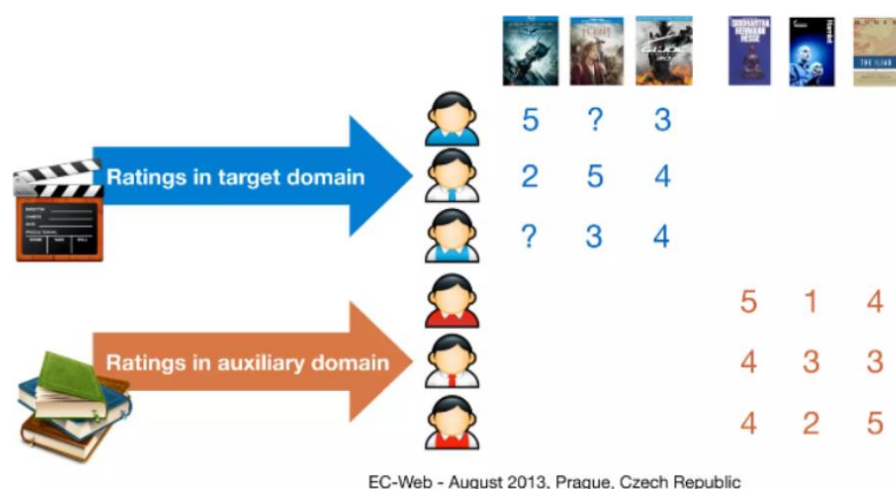


Figure 3.1.2 :Cross-Domain Recommendation with Auxiliary Domain Ratings

This thesis addresses one of the most challenging settings in cross-domain recommendation: a fully non-overlapping scenario in which the source domain (Book-Crossing) and the target domain (MovieLens) share neither users nor items. The two datasets differ substantially in identifiers, rating scales, interaction densities, and content characteristics, and even semantic metadata alignment is non-trivial, requiring enrichment through external information sources to recover genre labels for books. As a result, traditional cross-domain recommendation techniques based on user overlap, item overlap, or joint matrix factorisation are not applicable.

The central objective of this work is therefore to improve rating prediction accuracy in a sparse movie recommendation environment by transferring knowledge from an auxiliary book-rating domain using genres as a semantic bridge. The underlying assumption is that, although users and items differ across domains, both books and movies share a common semantic structure in the form of genres, which capture high-level content properties and reflect systematic patterns in user preferences.

To exploit this shared structure, the thesis proposes a genre-based semantic transfer framework built on latent factor models. A genre aware matrix factorisation model is first trained on the Book Crossing dataset, learning latent representations for users, items, and genres. The resulting genre embeddings encode domain-level semantic regularities derived from large scale book rating behaviour. These embeddings are then transferred at the parameter level to initialise the corresponding genre representations in a MovieLens model, which is subsequently trained and fine tuned using only movie ratings. In this way, genre embeddings act as informative priors that

guide learning in the target domain, particularly under severe data sparsity and cold start conditions.

The scope of the thesis is deliberately constrained to ensure interpretability and methodological clarity. The final prediction task is single domain rating prediction in the movie domain, while the book domain is used exclusively as a source of auxiliary knowledge. Only explicit ratings are considered, each item is assigned a single dominant genre, and the models are limited to latent factor approaches without deep neural architectures. Evaluation is performed offline using an extended cross validation protocol that systematically varies the amount of available target-domain data, allowing a detailed analysis of how genre information and cross-domain transfer affect prediction accuracy across different sparsity regimes.

Formally, the problem can be defined as follows. Let R_b denote the sparse user-book rating matrix in the Book Crossing domain, and R_m the target user-movie matrix in MovieLens. Let G denote a unified genre space constructed through semantic alignment between the source and target domains.

$$\hat{r}_{ui} = f(u, i, g_i)$$

that estimates how a movie i will be rated by a user u , where g_i is the dominant genre of the movie. The function f is parameterised by latent user vectors P , item vectors Q , and genre vectors Y . The defining challenge of this setting is that, due to the complete absence of overlapping users and items between the two domains, the user and item latent spaces cannot be aligned or shared. Consequently, genre embeddings constitute the only feasible conduit for cross-domain knowledge transfer.

Additionally, due to differences in genre vocabularies between books and movies, a semantic alignment process is applied to construct a unified genre space across domains using embedding-based similarity matching.

Under this constraint, the learning objective is structured as a two stage optimisation process. First, genre embeddings $Y(B)$ are learned in the source domain by minimising the reconstruction error of observed book ratings. Second, these learned genre representations are transferred to the target domain and used to initialise the corresponding genre parameters $Y(M)$ in the movie model. User and item latent vectors in the target domain are then learned exclusively from available MovieLens ratings. Finally, the effectiveness of this transferred semantic knowledge is evaluated by measuring prediction accuracy under varying levels of target-domain data sparsity.

This formulation directly addresses the cold-start and sparsity challenges identified both in the literature and in the empirical behaviour observed in this work. As evidenced by the experimental results, models trained on limited fractions of MovieLens data (10%, 20%, ..., 100%) exhibit substantial degradation in prediction accuracy when neither genre information nor transferred knowledge is incorporated. Introducing genre-aware latent factors improves performance across all sparsity levels, while transferring genre embeddings from the book domain yields the lowest Mean Absolute Error (MAE), particularly in extreme low data regimes. These findings confirm that semantically informed transfer provides a strong inductive bias that stabilises learning when collaborative signals are weak.

Within this framework, the scope of the thesis is clearly delimited. The study considers a fully non overlapping cross domain setting, transferring knowledge from books to movies without shared users or items. The prediction task is restricted to single domain rating prediction in the movie domain, with the book domain serving solely as an auxiliary source of semantic knowledge. All models belong to the class of latent factor approaches and explicitly incorporate genre embeddings as semantic components. Knowledge transfer is performed at the parameter level by reusing genre latent representations learned in the source domain. Evaluation is conducted through a multi fold, sparsity aware experimental protocol based on incremental training percentages. To preserve interpretability and computational feasibility, the analysis is

limited to explicit ratings, assigns each item a single dominant genre, and excludes deep neural architectures.

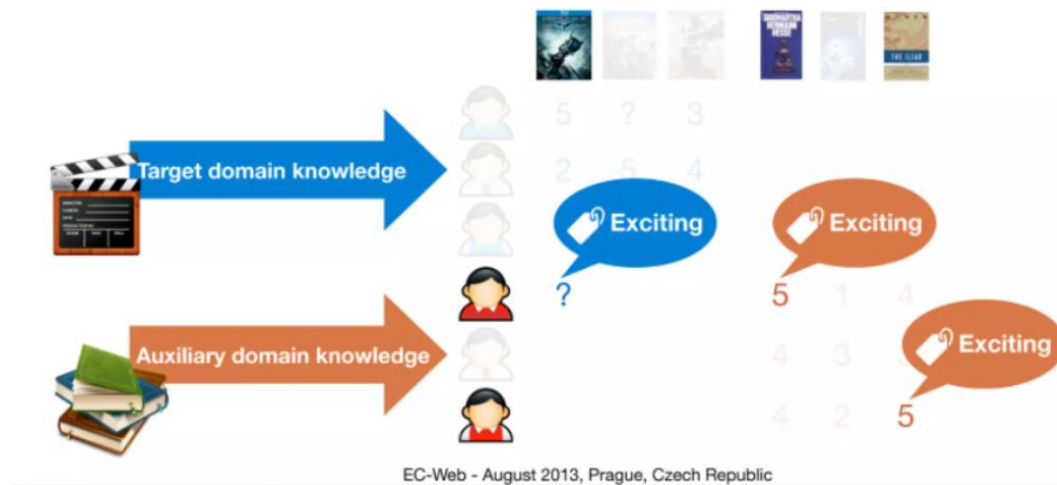


Figure 3.1.3 : Semantic Knowledge Transfer via Shared Attributes

3.2 Theoretical Foundation of the Proposed Genre-Based Transfer Model

This section develops the theoretical basis of the proposed approach for cross-domain recommendation under full non-overlap between domains[18]. As established in the previous section, the problem addressed in this thesis concerns a setting in which the source domain (Book-Crossing) and the target domain (MovieLens) share neither users nor items. Under such conditions, conventional cross-domain recommendation methods based on shared identities or joint interaction structures cannot be directly applied. The challenge, therefore, is to identify a form of knowledge that is both semantically meaningful and transferable across domains.

The approach adopted in this thesis is based on the assumption that, although books and movies are distinct types of items, they can still be described through a partially shared semantic space. This common space is defined through genres. Genres such as *Drama*, *Fantasy*, *Romance*, or *Science Fiction* do not depend on specific user identities or item identifiers; instead, they express higher-level thematic and stylistic properties that can be interpreted across domains. For this reason, genres can serve as semantic anchors through which preference-related knowledge may be transferred from one domain to another.



Figure 3.2 : Sparse target domain supported by a richer auxiliary domain.

3.2.1 Semantic transfer through shared genre structure

The theoretical motivation of the proposed model lies in the distinction between collaborative similarity and semantic similarity. Collaborative methods infer preference patterns from rating behaviour alone, relying on user–item interaction overlap in order to discover latent relations. In

fully non-overlapping domains, however, no such overlap exists, and direct collaborative transfer becomes impossible. Semantic transfer offers an alternative by shifting the focus from identities to attributes.

In the present work, genres are treated as transferable semantic units. The central hypothesis is that genre-related preference patterns learned from book ratings can provide useful prior knowledge for movie recommendation. Although the users of the two domains are different and the items themselves are unrelated, the semantic interpretation of genres remains sufficiently stable to allow knowledge reuse. A latent representation of the genre *Fantasy*, for example, may encode preference tendencies associated with imagination, adventure, or speculative narratives. These tendencies may emerge in both books and movies, even if they are expressed through different media forms.

This reasoning places the proposed approach within the broader family of semantic transfer methods in recommender systems. Rather than attempting to align user vectors or item vectors across incompatible domains, the method transfers only those components of the model that correspond to shared semantic concepts. As a result, the model remains theoretically coherent even in the absence of any overlapping entities.

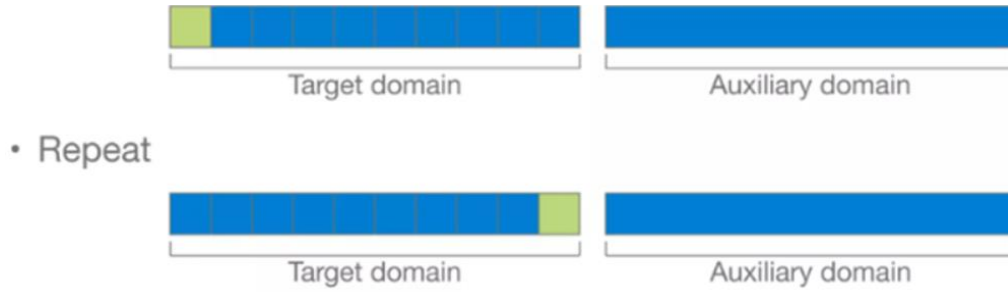


Figure 3.2.1 : Incremental Training and Repetition in Extended Cross-Validation

3.2.2 Latent representation of users, items, and genres

The proposed model is grounded in latent factor theory[16]. In matrix factorisation-based recommender systems, users and items are represented as vectors in a low-dimensional latent space. The interaction between these vectors determines the predicted rating. This representation allows the model to capture hidden patterns that are not explicitly observed in the raw rating matrix.

To extend this framework toward semantic transfer, the present thesis introduces a third type of latent representation: genre vectors. In this formulation, each user u is associated with a latent vector p_u , each item i with a latent vector q_i , and each genre g with a latent vector y_g . The genre vector is intended to capture content-related information that is shared among items belonging to the same semantic category.

This extension changes the interpretation of the latent space. Item preference is no longer determined solely by the interaction between user-specific and item-specific factors, but also by a semantic component associated with the item's genre. Consequently, the model can generalise more effectively in sparse conditions, because information is partially shared across items that belong to the same genre class. If two movies belong to the same genre, they are no longer treated as completely unrelated when little rating data are available.

Within the scope of this thesis, each item is assigned a single dominant genre. This simplification makes the model easier to interpret and keeps the semantic representation compact. Although real items may belong to multiple genres, the dominant-genre assumption provides a clear one-to-one link between items and transferable semantic factors, which is appropriate for the theoretical formulation adopted here.

3.2.3 Genre-aware rating prediction model

Based on the above representation, the predicted rating of user u for item i is defined through a genre-aware latent interaction model. Let $g(i)$ denote the dominant genre of item i . Then the predicted rating can be expressed as:

$$\hat{r}_{ui} = \mu + b_u + b_i + p_u^\top (q_i + y_{g(i)})$$

where:

- μ is the global mean rating,
- b_u is the user bias,
- b_i is the item bias,
- p_u is the latent vector of user u ,
- q_i is the latent vector of item i ,
- $y_{g(i)}$ is the latent vector of the dominant genre of item i .

This formulation combines collaborative and semantic information in a unified prediction function. The term $p_u^\top q_i$ captures the classical collaborative interaction between a user and an item, while the term $p_u^\top y_{g(i)}$ captures the user's affinity toward the item's genre. In this sense, the genre vector functions as a semantic correction or enrichment of the item representation.

If genre information is omitted, the model reduces to the standard biased matrix factorisation form:

$$\hat{r}_{ui} = \mu + b_u + b_i + p_u^\top q_i$$

The comparison between these two formulations is central to the logic of the thesis. It allows the contribution of genre information to be examined independently from the contribution of collaborative factors. More importantly, it establishes the theoretical mechanism through which semantic information can influence prediction quality under sparse-data conditions.

3.2.4 Cross-domain transfer at the parameter level

The cross-domain component of the proposed method is implemented at the level of genre parameters. Let $Y(B)$ denote the set of genre embeddings learned in the book domain and $Y(M)$ the corresponding genre embeddings in the movie domain. Because users and items are entirely disjoint across domains, user vectors and item vectors cannot be shared, reused, or aligned. Genre embeddings, however, belong to the shared semantic vocabulary and therefore constitute the only transferable component of the model..

The idea of transferring knowledge between recommendation domains has been previously explored in cross-domain collaborative filtering studies (Enrich et al., 2013). However, unlike

tag-based transfer approaches, the proposed framework relies on semantic genre embeddings as the primary transfer mechanism.

The transfer mechanism is defined as follows. First, a genre-aware latent factor model is learned in the source domain using book ratings. Through this process, the genre vectors $Y(B)$ acquire a semantic structure shaped by the rating behaviour observed in the auxiliary domain. These learned vectors are then used to initialise the corresponding genre vectors in the target domain:

$$Y(M) \leftarrow Y(B)$$

This form of transfer is best understood as parameter-level semantic initialisation. The movie-domain model does not inherit any user-specific or item-specific knowledge from the book domain. Instead, it starts from an informed representation of genres before learning from movie ratings. In theoretical terms, the transferred genre vectors act as informative priors. They constrain the early stages of learning in the target domain and provide a meaningful inductive bias when collaborative evidence is weak.

This is especially important in cold-start and sparse-data regimes. In such cases, user and item vectors estimated only from limited movie ratings are unstable and noisy. By contrast, pre-trained genre vectors introduce an additional structure that guides the optimisation process toward semantically plausible solutions. Thus, the transfer mechanism does not replace target-domain learning; rather, it supports and stabilises it.

3.2.5 Theoretical assumptions, expected benefits, and limitations

The proposed model is based on several theoretical assumptions. The first is that genres are sufficiently aligned across books and movies to function as a common semantic vocabulary. The second is that genre-related preference patterns learned in one domain remain at least partially informative in the other. The third is that this transferred knowledge is most useful when target-domain collaborative information is limited.

Under these assumptions, the model is expected to offer three main benefits. First, it should improve prediction quality in sparse-data settings by providing semantically meaningful priors. Second, it should reduce the severity of the cold-start problem, especially during the early stages of learning in the target domain. Third, it should preserve interpretability, since the transferred knowledge is associated with explicit semantic categories rather than opaque end-to-end representations.

At the same time, the theoretical framework also imposes limitations. Genre alignment across domains is not perfect, and the meaning of a genre label may differ between books and movies. The dominant-genre assumption also simplifies the semantic richness of real items, many of which belong to multiple categories simultaneously. In addition, genres provide only coarse-grained information and cannot capture finer semantic nuances such as narrative tone, writing style, pacing, or aesthetic form. These limitations do not invalidate the model, but they define its scope and clarify the type of knowledge it can realistically transfer.

Overall, the theoretical contribution of this section is to establish a coherent cross-domain recommendation model for the fully non-overlapping case. The model does not depend on unrealistic assumptions of shared users or shared items. Instead, it builds on a transferable semantic layer represented by genre embeddings, integrating collaborative filtering principles with parameter-level transfer learning. In this way, the proposed framework offers a theoretically grounded solution to recommendation under sparsity, while remaining sufficiently simple and interpretable for rigorous experimental evaluation.

3.3 Mathematical Formulation of the Model

This section formalises the proposed genre-based cross-domain recommendation model within the latent factor framework. The objective is to define the representation of users, items, and genres, as well as the prediction function and optimisation objective that govern the learning process.

3.3.1 Notation and Problem Definition

Let the source domain (books) be denoted as $\mathcal{D}_B$ and the target domain (movies) as $\mathcal{D}_M$. Each domain is defined by a set of users, a set of items, and a sparse rating matrix:

$$\mathcal{D}_B = (U_B, I_B, R_B), \quad \mathcal{D}_M = (U_M, I_M, R_M)$$

where:

- U denotes the set of users,
- I denotes the set of items,
- $R \in \mathbb{R}^{|U| \times |I|}$ is the partially observed rating matrix.

By definition of the problem setting:

$$U_B \cap U_M = \emptyset, \quad I_B \cap I_M = \emptyset$$

indicating that there are no shared users or items between the two domains.

Let G denote the set of genres shared across both domains. Each item $i \in I$ is associated with a dominant genre $g(i) \in G$. This shared genre set constitutes the only common structure between domains and serves as the basis for knowledge transfer.

The goal of the model is to learn a prediction function:

$$f : (u, i) \rightarrow \hat{r}_{ui}$$

that estimates the unknown rating r_{ui} of user $u \in U_M$ for item $i \in I_M$, leveraging both target-domain data and transferred semantic knowledge from the source domain.

3.3.2 Latent Factor Representation

The proposed model is based on matrix factorisation, where users, items, and genres are embedded into a shared latent space of dimensionality K .

Specifically:

- Each user u is represented by a latent vector $p_u \in \mathbb{R}^K$,
- Each item i is represented by a latent vector $q_i \in \mathbb{R}^K$,
- Each genre $g \in G$ is represented by a latent vector $y_g \in \mathbb{R}^K$.

These latent vectors capture hidden factors that explain user preferences and item characteristics. The introduction of genre vectors extends the classical matrix factorisation model by incorporating semantic information into the latent space.

3.3.3 Rating Prediction Function

In the standard latent factor model, the predicted rating is computed as the inner product between user and item vectors, optionally augmented with bias terms. In the proposed genre-aware formulation, the item representation is enriched by incorporating its associated genre embedding.

The predicted rating is defined as:

$$\hat{r}_{ui} = \mu + b_u + b_i + p_u^\top (q_i + y_{g(i)})$$

where:

- μ is the global mean rating,
- b_u is the bias term for user u ,
- b_i is the bias term for item i ,
- $p_u^\top q_i$ captures collaborative interaction,
- $p_u^\top y_{g(i)}$ captures semantic preference for the item's genre.

This formulation allows the model to combine collaborative filtering with content-based information in a unified framework. The genre vector acts as a semantic adjustment to the item representation, enabling information sharing across items that belong to the same genre.

In the absence of genre information, the model reduces to the classical biased matrix factorisation form:

$$\hat{r}_{ui} = \mu + b_u + b_i + p_u^\top q_i$$

which serves as a baseline for evaluating the contribution of semantic features.

3.3.4 Optimisation Objective

Model parameters are learned by minimising the discrepancy between observed ratings and predicted ratings over the set of known interactions. The optimisation objective is defined as a regularised squared error loss:

$$\min_{P, Q, Y, b_u, b_i} \sum_{(u,i) \in \mathcal{K}} (r_{ui} - \hat{r}_{ui})^2 + \lambda (\|p_u\|^2 + \|q_i\|^2 + \|y_{g(i)}\|^2 + b_u^2 + b_i^2)$$

where:

- $\mathcal{K}$ is the set of observed user–item interactions,
- $\lambda > 0$ is the regularisation parameter controlling model complexity.

The first term minimises prediction error, while the second term prevents overfitting by penalising large parameter values. Regularisation is particularly important in sparse-data settings, where the number of observations is insufficient to reliably estimate high-dimensional latent vectors.

3.3.5 Source and Target Domain Learning

The optimisation process is applied separately in the source and target domains, with a key difference in the role of genre embeddings.

In the **source domain**, all model parameters are learned from scratch by minimising the reconstruction error of book ratings. This process yields a set of genre embeddings $Y(B)$ that capture semantic preference patterns derived from the auxiliary domain.

In the **target domain**, the objective function remains the same, but the initialisation of genre vectors differs due to cross-domain transfer. Instead of random initialisation, genre embeddings are initialised using the values learned in the source domain:

$$y_g(M) \leftarrow y_g(B)$$

User and item vectors in the target domain are learned exclusively from movie ratings, while genre vectors act as informed initial parameters.

This formulation preserves the independence of user and item latent spaces across domains while enabling knowledge transfer through shared semantic components.

3.3.6 Interpretation of the Model

The proposed formulation can be interpreted as a hybrid latent factor model with an explicit semantic component. The prediction function decomposes user preference into two complementary parts:

1. **Collaborative component** ($p_u^T q_i$)
Captures interaction patterns derived from observed ratings.
2. **Semantic component** ($p_u^T y_g(i)$)
Captures preference tendencies associated with item genres.

In cross-domain settings, the semantic component becomes particularly important, as it provides a mechanism for transferring knowledge when collaborative signals cannot be shared. The genre embeddings effectively encode domain-independent preference structures, which can guide learning in the target domain.

3.4 Cross-Domain Transfer Mechanism

The core idea of this thesis lies in transferring useful knowledge from the source domain to the target domain, despite the complete absence of shared users or items. As discussed in the previous sections, this setting makes traditional cross-domain approaches inapplicable, since there is no direct way to align user behaviour or item characteristics across domains. For this reason, the transfer mechanism must rely exclusively on information that is common at a higher, semantic level.

In the proposed approach, this role is fulfilled by genres. Genres constitute the only shared structure between the two domains and therefore act as the bridge through which knowledge can be transferred. The transfer mechanism is designed around this observation and operates at the level of genre representations.

3.4.1 Transfer of Semantic Representations

The transfer process begins in the source domain, where a genre-aware latent factor model is trained using book ratings. During this stage, the model learns latent representations not only for users and items, but also for each genre. These genre embeddings capture how different types of content are perceived and rated by users in the book domain. In other words, they encode patterns such as which genres tend to receive higher ratings and how user preferences vary across genre categories.

Once training in the source domain is complete, the learned genre embeddings are extracted and transferred to the target domain. Importantly, only the genre vectors are transferred. User and item representations are not reused, since there is no correspondence between users or items across domains. This ensures that the transfer mechanism remains consistent with the fully non-overlapping setting considered in this work.

Formally, if $Y(B)$ denotes the set of genre embeddings learned in the book domain, then the initial genre representations in the movie domain are set as:

$$Y(M) \leftarrow Y(B)$$

This step can be interpreted as providing the target-domain model with prior knowledge about the semantic structure of genres before any movie-specific training takes place.

3.4.2 Role of Genre Embeddings as Priors

From a theoretical perspective, the transferred genre embeddings act as informative priors. Instead of starting from random initial values, the model begins with representations that already reflect meaningful preference patterns learned from a richer dataset. This is particularly beneficial in sparse-data scenarios, where the target-domain model would otherwise struggle to learn stable representations from limited observations.

The presence of these priors influences the learning process in the target domain. Early in training, when user–item interactions are scarce, the model relies more heavily on the semantic information encoded in the genre vectors. As more movie ratings become available, the model gradually adjusts these representations to better fit the characteristics of the target domain. In this way, the transferred knowledge guides the learning process without constraining it.

This behaviour is consistent with general principles of transfer learning, where prior knowledge is most influential when target-domain data are limited and becomes less dominant as more data are observed.

3.4.3 Adaptation to the Target Domain

Although genres provide a shared semantic structure, their interpretation may not be identical across domains. For example, the concept of “Drama” in literature may emphasise different aspects compared to its interpretation in film. To account for such differences, the model allows genre embeddings to be updated during training in the target domain.

This adaptation process ensures that the transferred representations remain relevant. Instead of being fixed, genre vectors are refined using movie-rating data, enabling the model to correct any discrepancies between the source and target domains. This step is essential for avoiding negative transfer, where knowledge from the source domain could otherwise degrade performance in the target domain.

At the same time, the initialisation from the source domain ensures that the model does not start from an uninformed state. The combination of prior knowledge and subsequent adaptation provides a balance between transfer and domain-specific learning.

3.5 Theoretical Advantages and Limitations of the Proposed Approach

The proposed genre-based transfer model is designed to address one of the most challenging settings in recommender systems: cross-domain recommendation without shared users or items. By relying on semantic information rather than direct interaction overlap, the approach offers a number of conceptual advantages. At the same time, it is important to recognise the assumptions and limitations that define its applicability.

3.5.1 Theoretical Advantages

A primary advantage of the proposed method is its ability to operate under full non-overlap conditions. Many cross-domain recommendation techniques assume that at least some users or items are shared between domains, allowing latent factors to be aligned directly. In contrast, the present approach does not require any such overlap. By transferring only genre representations, it remains valid even when the two domains are completely disjoint. This makes the model more flexible and applicable to real-world scenarios where data integration across platforms is limited or infeasible.

Another important strength lies in its handling of data sparsity and cold-start conditions. In traditional collaborative filtering models, performance depends heavily on the availability of user–item interactions. When such data are scarce, the model struggles to learn meaningful representations. In the proposed framework, genre embeddings introduce an additional source of information that is shared across items. As a result, the model can generalise more effectively, even when only a small number of ratings are available. The transferred genre representations further enhance this effect by providing an informed starting point for learning in the target domain.

The approach also benefits from its interpretability. Unlike more complex models, particularly those based on deep learning, the structure of the proposed model remains relatively transparent. Each component has a clear role: user vectors represent individual preferences, item vectors capture item-specific characteristics, and genre vectors encode semantic information. This separation makes it easier to understand how predictions are formed and to analyse the contribution of each component. In an academic context, this interpretability is particularly valuable, as it allows for clearer reasoning about model behaviour.

In addition, the method maintains conceptual simplicity while extending a well-established framework. By building on matrix factorisation, the model leverages a technique that is widely understood and empirically effective. The introduction of genre embeddings does not fundamentally alter the underlying structure but rather enriches it in a controlled way. This makes the approach easier to analyse and compare with baseline methods, while still capturing meaningful semantic relationships.

Finally, the use of semantic transfer aligns with broader trends in recommender systems research. There is increasing recognition that relying solely on interaction data is often insufficient, especially in sparse environments. By incorporating high-level features such as genres, the proposed model reflects a hybrid perspective that combines collaborative and content-based reasoning. This integration allows the system to benefit from both types of information without requiring complex architectures.

3.5.2 Limitations and Assumptions

Despite its advantages, the proposed approach is subject to several limitations that should be considered when interpreting its results.

One key limitation is the reliance on genre information as the sole transfer mechanism. Genres provide a high-level description of items, but they are inherently coarse and may not capture finer distinctions between them. For example, two movies classified under the same genre may differ significantly in style, tone, or narrative structure. As a result, the semantic information provided by genres may be insufficient to fully represent user preferences, particularly in more nuanced recommendation tasks.

A related issue is the assumption that genre semantics are consistent across domains. Although books and movies share common genre labels, the way these genres are expressed can differ between media. For instance, the notion of “Action” in literature may not correspond exactly to its cinematic counterpart. These differences introduce a degree of semantic misalignment, which may affect the quality of transferred knowledge. While the model partially addresses this issue through adaptation in the target domain, complete alignment cannot be guaranteed.

Another limitation arises from the simplified modelling assumptions adopted in this work. Each item is assigned a single dominant genre, even though many items naturally belong to multiple categories. This simplification reduces the complexity of the model but also limits its expressive power. A more realistic representation would allow items to be associated with multiple genres, potentially capturing richer semantic relationships.

Furthermore, the transfer mechanism itself is restricted to item-side information. Since no users are shared between domains, it is not possible to transfer user-level knowledge. This means that the model cannot directly capture cross-domain user behaviour patterns, which could otherwise provide valuable signals. As a result, the effectiveness of the approach depends entirely on the quality and relevance of the genre representations.

Finally, the model assumes the availability of reliable and consistent genre metadata in both domains. In practice, such information may be incomplete, noisy, or inconsistently defined, particularly in datasets such as Book-Crossing. Additional preprocessing and external data sources may be required to construct a usable genre vocabulary. These practical considerations highlight that the success of the method depends not only on the model itself but also on the quality of the underlying data.

Chapter 4 Proposed Framework, Implementation and Evaluation

4.1 System Overview and Architecture

The proposed framework combines collaborative filtering, semantic representation learning, and transfer learning within a unified cross-domain recommendation system. The objective of the architecture is to improve recommendation performance under sparse-data conditions by transferring semantic genre knowledge from books to movies.

The framework uses the Book-Crossing dataset as the source domain and the MovieLens dataset as the target domain. Since the two domains do not share users or items, the transfer process relies entirely on semantic genre information rather than collaborative overlap.

The overall workflow of the proposed system is illustrated below:



Figure 4.1: Overall architecture of the proposed genre-based cross-domain transfer learning framework.

The first stage of the framework focuses on preparing and semantically aligning the two datasets. While the MovieLens dataset already contains genre information, the Book-Crossing dataset lacks consistent genre metadata. To address this limitation, external metadata sources were used to retrieve genre labels for books.

After genre enrichment, semantic mapping was performed using SentenceTransformer [19] embeddings and cosine similarity. This process allowed semantically related genres from the two domains to be aligned within a shared semantic space. For example, labels such as “Fantasy fiction” in books could be associated with the “Fantasy” genre used in movies.

Once semantic alignment was completed, a genre-aware Matrix Factorization model was trained on the source-domain book ratings. During training, the model learned latent representations for users, items, and genres simultaneously. The learned genre embeddings captured semantic preference patterns associated with different content categories.

The learned source-domain embeddings were then transferred to the movie recommendation model through pretrained initialization:

```
G_movies = G_books.copy()
```

After transfer, the movie recommendation model was fine-tuned using MovieLens ratings. During this stage, the transferred embeddings adapted to the characteristics of the movie domain while preserving the semantic information learned from the source domain.

The framework was implemented in Python using pandas and NumPy for preprocessing and matrix operations, SentenceTransformers for semantic embedding generation, and scikit-learn for evaluation and cross-validation. Experimental results were visualized using Matplotlib.

All recommendation models were trained using stochastic gradient descent optimisation, while recommendation performance was evaluated using Mean Absolute Error (MAE) under progressively increasing sparsity conditions through 10-fold cross-validation.

4.2 Data Preparation and Semantic Alignment

Before model training could begin, both datasets had to be transformed into a compatible semantic representation. Since the proposed framework performs transfer learning between two heterogeneous domains, preprocessing and semantic alignment were necessary in order to construct a shared genre space between books and movies.

The MovieLens dataset already contains genre information for movies. In contrast, the Book-Crossing dataset does not provide complete or consistent genre metadata. For this reason, an additional enrichment stage was required for the source-domain dataset.

The preprocessing workflow followed the pipeline shown below:

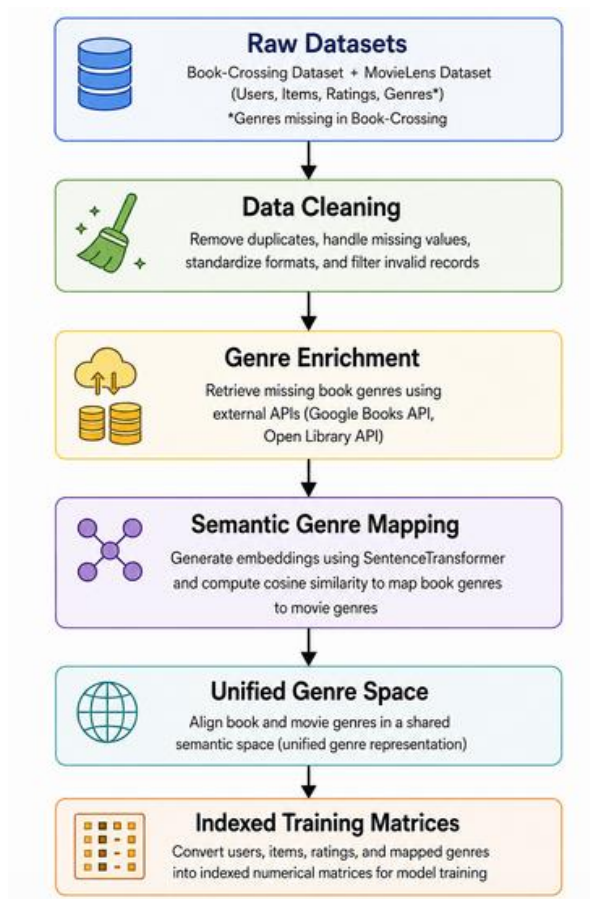


Figure 4.2: Data preparation and semantic alignment process for the proposed cross-domain recommendation framework.

Initially, both datasets were cleaned and transformed into structured tabular form using pandas. User identifiers and item identifiers were converted into indexed numerical representations in order to support efficient latent factor computation during matrix factorization training.

The MovieLens dataset already includes predefined movie genres. However, since movies may belong to multiple categories simultaneously, a simplified representation was adopted in which each movie was assigned a single dominant genre.

The Book-Crossing dataset required additional preprocessing because genre information was missing for a large number of books. To address this limitation, genre metadata were retrieved using the Google Books API and Open Library API. Queries were generated using book titles and author information, and the retrieved genre labels were incorporated into the dataset.

Examples of retrieved genre information include:

Table 4.2.1: Examples of genre labels retrieved during the genre enrichment process.

Book	Retrieved Genre
Harry Potter	Fantasy fiction
Sherlock Holmes	Detective stories
Dune	Science fiction

After genre enrichment, semantic mapping was performed between the book and movie genres. Direct string matching was not sufficient because many categories differed linguistically while still representing similar semantic concepts.

For example:

Table 4.2.2: Examples of semantic mappings between book genres and movie genres.

Book Genre	Movie Genre
Fantasy fiction	Fantasy
Detective stories	Mystery
Science fiction	Sci-Fi

To identify these relationships, semantic embeddings were generated using SentenceTransformer models. Each genre label was converted into a dense semantic vector representation within a shared embedding space.

Cosine similarity was then used to measure semantic similarity between genre embeddings:

$$\cos(\theta) = (A \cdot B) / (\|A\| \|B\|)$$

where:

- A and B represent semantic embedding vectors,
- and $\cos(\theta)$ measures their similarity.

The semantic mapping process was implemented by comparing the similarity between book-genre embeddings and movie-genre embeddings:

$$\text{similarity} = \text{cosine_similarity}(\text{book_embedding}, \text{movie_embedding})$$

Genres with high similarity scores were mapped to the same semantic category, creating a unified genre representation across both domains.

Once semantic alignment was completed, the datasets were transformed into indexed latent-factor structures suitable for recommendation training. Each interaction contained:

- a user index,
- an item index,
- a rating value,
- and a mapped genre index.

The final output of the preprocessing stage consisted of two semantically aligned datasets sharing a common genre representation. This unified semantic space became the foundation of the genre-aware recommendation models and the cross-domain transfer learning process described in the following sections

4.3 Recommendation Models

This section presents the recommendation models implemented in the experimental evaluation of the proposed framework. The objective of the comparison is to examine how semantic genre information and cross-domain transfer learning affect recommendation performance under sparse-data conditions.

Four models were implemented and evaluated progressively, beginning with standard collaborative filtering approaches and extending toward semantic genre modelling and transfer learning.

The evaluated models are summarised below:

Table 4.3: Recommendation models used in the experimental evaluation.

Model	Description
SVD	Singular Value Decomposition collaborative filtering
Matrix Factorization	Latent factor collaborative filtering with SGD
Genre-Aware MF	Matrix Factorization with genre embeddings
Transfer Learning Model	Cross-domain genre embedding transfer

4.3.1 Baseline Collaborative Filtering Models

The first two models serve as collaborative filtering baselines. These approaches rely exclusively on user–item interaction data and do not incorporate semantic information.

The first baseline model is based on Singular Value Decomposition (SVD), where the sparse user–item rating matrix is decomposed into lower-dimensional latent matrices representing hidden user and item preference factors.

The second baseline model uses Matrix Factorization trained with stochastic gradient descent optimisation. In this formulation:

- each user is represented by a latent vector,
- and each item is represented by a latent vector within the same latent space.

The predicted rating is computed as:

$$\text{Prediction} = P[u] \cdot Q[i]$$

where:

- $P[u]$ represents the latent vector of user u ,
- and $Q[i]$ represents the latent vector of item i .

Initially, all latent vectors are randomly initialized with small values. During training, the model iteratively generates predictions, computes prediction error, and updates the latent vectors using stochastic gradient descent.

The prediction error is defined as:

$$\text{error} = \text{real rating} - \text{predicted rating}$$

Although these collaborative filtering approaches are effective in dense recommendation environments, they rely entirely on observed interactions. Consequently, their performance deteriorates under sparse-data conditions where limited ratings are available for learning stable latent representations.

4.3.2 Genre-Aware Recommendation Model

To address the limitations of purely collaborative filtering approaches, the third model extends Matrix Factorization by incorporating semantic genre information directly into the recommendation process.

In addition to user and item latent vectors, the model also learns latent genre embeddings:

$P[u] \rightarrow$ user latent vector
 $Q[i] \rightarrow$ item latent vector
 $G[g] \rightarrow$ genre embedding

These embeddings capture semantic relationships associated with categories such as Fantasy, Mystery, Sci-Fi, and Romance.

Instead of predicting ratings using only user and item vectors, the proposed model enriches the item representation with semantic genre information:

$$\text{Prediction} = P[u] \cdot (Q[i] + G[g])$$

This formulation allows semantically related items to share information through common genre embeddings. As a result, the model can generalize more effectively under sparse-data conditions where direct collaborative signals may be limited.

The learning process follows the optimisation pipeline illustrated below:

During optimisation, the genre embeddings gradually evolve into meaningful semantic representations based on user rating behaviour.

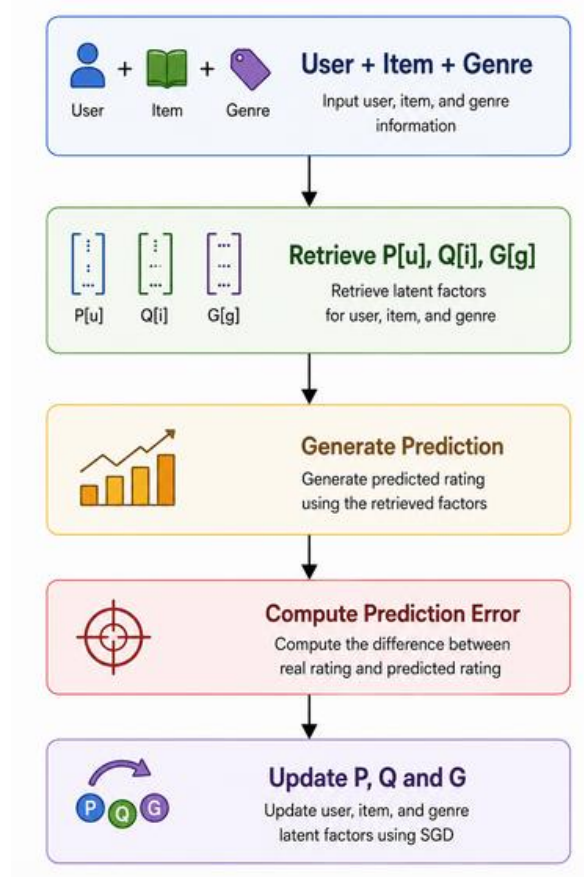


Figure 4.3.2: Training workflow of the Genre-Aware Matrix Factorisation model.

As training progresses, semantically related genres tend to develop nearby latent representations within the embedding space.

These learned genre embeddings later become the transferable semantic component used in the cross-domain transfer learning model.

4.3.3 Cross-Domain Transfer Learning Model

The final model represents the main contribution of the thesis. This architecture transfers semantic genre embeddings learned from the source-domain book dataset into the target-domain movie recommendation model.

The transfer learning process consists of three main stages:

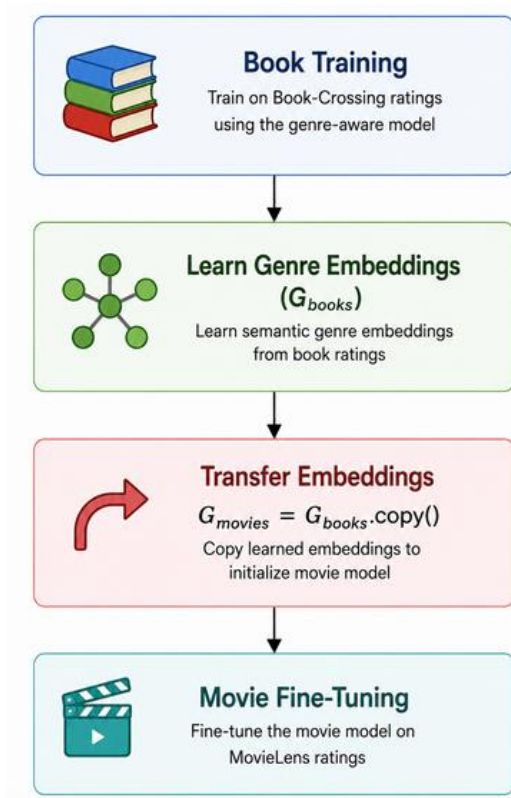


Figure 4.3.3: Transfer of genre embeddings from the Book-Crossing domain to the MovieLens domain.

During source-domain training, the genre-aware recommendation model learns semantic genre embeddings from book-rating behaviour. These embeddings capture latent semantic preference patterns associated with different categories.

Once training is completed, the learned embeddings are transferred to the movie recommendation model through pretrained initialization:

```
G_movies = G_books.copy()
```

This means that the movie model does not begin training with random semantic representations. Instead, it starts with pretrained genre embeddings already containing semantic knowledge learned from the source domain.

After transfer, the model undergoes fine-tuning using MovieLens ratings. During this stage:

- predictions are generated,
- prediction errors are computed,
- and all latent vectors continue to be updated using stochastic gradient descent.

The prediction function remains identical to the genre-aware formulation:

$$\text{Prediction} = P_movies[u] \cdot (Q_movies[i] + G_movies[g])$$

However, unlike the previous model, the genre embeddings are initialized using pretrained semantic representations rather than random values.

This pretrained initialization provides additional semantic guidance during optimisation, particularly under sparse-data conditions where collaborative information alone may be insufficient for learning stable latent representations.

4.4 Experimental Methodology

The proposed framework was evaluated under progressively increasing sparsity conditions in order to examine the effect of semantic genre modelling and cross-domain transfer learning on recommendation performance.

The experimental evaluation compared: standard collaborative filtering models, the genre-aware recommendation model, and the proposed transfer learning architecture

under identical training and evaluation conditions.

The evaluation process combined:

- 10-fold cross-validation,
- incremental training percentages,
- and Mean Absolute Error (MAE) evaluation.

4.4.1 Evaluation Process

The MovieLens dataset was divided using 10-fold cross-validation. In each iteration: one fold was used as the test set, while the remaining folds formed the available training pool.

Unlike standard cross-validation approaches, the experiments did not use the full training pool immediately. Instead, the amount of available training data increased progressively in order to simulate different sparsity levels.

The evaluation workflow is illustrated below:



Figure 4.4.1: Evaluation procedure based on 10-fold cross-validation and incremental training availability.

For each fold, the recommendation models were trained repeatedly using progressively larger portions of the available training data:

- 10%,
- 20%,
- 30%,
- up to 100%.

This process simulated recommendation environments ranging from extremely sparse conditions to fully observed interaction settings.

At each sparsity level, the model was trained, predictions were generated for the test set, and MAE was computed.

The final evaluation scores were obtained by averaging the results across all folds.

4.4.2 Hyperparameter Configuration

All recommendation models were trained using stochastic gradient descent optimisation.

The latent dimensionality was set to:

$$K = 50$$

while the learning rate and regularisation parameters were defined as:

$$\begin{aligned} \alpha &= 0.005 \\ \lambda &= 0.02 \end{aligned}$$

Training was performed for 10 epochs with random shuffling of training samples during each iteration.

User, item, and genre embeddings were initialized using small random values. However, in the transfer learning model, the genre embeddings were initialized using pretrained vectors learned from the source-domain book dataset.

This pretrained initialization constitutes the core transfer mechanism of the proposed framework.

4.4.3 Evaluation Metric

Recommendation performance was evaluated using Mean Absolute Error (MAE), which measures the average absolute difference between predicted ratings and real user ratings.

The metric is defined as:

$$MAE = (1 / |T|) \sum |r_{ui} - \hat{r}_{ui}|$$

where:

- r_{ui} represents the real user rating,
- $\hat{r}_{ui}$ represents the predicted rating,
- and T represents the set of test interactions.

Lower MAE values indicate better recommendation accuracy.

An example of MAE computation is shown below:

Table 4.4.3: Illustrative example of Mean Absolute Error (MAE) calculation.

Real Rating	Predicted Rating	Absolute Error
5	4.2	0.8
4	3.7	0.3
2	2.5	0.5

The final MAE value corresponds to the average prediction error across all test interactions.

Using the same evaluation procedure for all models ensured consistent comparison between: collaborative filtering baselines, genre-aware recommendation and cross-domain transfer learning approaches.

Chapter 5 Experimental Evaluation

5.1 Data Preparation and Processing

The data preparation stage constitutes a critical component of the proposed recommendation framework, as it enables the integration of two heterogeneous domains, namely books and movies. These datasets differ significantly in terms of structure, metadata availability, and rating distributions. Therefore, preprocessing is not simply a technical requirement but a fundamental step that ensures the feasibility of cross-domain knowledge transfer, as described in Chapter 3.

5.1.1 Dataset Overview

The experimental evaluation was conducted using two publicly available recommendation datasets originating from different domains. The Book-Crossing dataset was selected as the source domain, while the MovieLens 1M dataset was used as the target domain.

The Book-Crossing dataset contains user ratings for books and provides information regarding users, books, authors, publishers, and publication years. However, the dataset does not contain structured genre information, which constitutes a major limitation for cross-domain recommendation. Consequently, an additional genre enrichment process was required before the dataset could be used within the proposed framework.

The MovieLens 1M dataset contains movie ratings together with predefined genre information. Unlike the book dataset, genre metadata are already available and can therefore be directly incorporated into the recommendation process.

Table 5.1 Dataset Statistics

Dataset	Users	Items	Ratings
Book-Crossing	105,283	35,220	105,283*
MovieLens 1M	6,040	3,706	1,000,209

* The Book-Crossing representation used in the experiments corresponds to the processed dataset after genre enrichment and semantic alignment.

The two datasets differ considerably in both scale and structure. Most importantly, they do not share users or items. Consequently, traditional cross-domain recommendation approaches based on overlapping entities cannot be applied. Instead, the proposed framework relies on semantic information in the form of genre representations, which constitute the only transferable component between domains.

5.1.2 Cleaning and Validation of Raw Data

Before any recommendation model could be trained, both datasets were subjected to a cleaning and validation process. The objective of this stage was to ensure that the experimental evaluation relied exclusively on valid and consistent user–item interactions.

Records containing missing values in critical attributes, including user identifiers, item identifiers, ratings, and genre information, were removed. Furthermore, only valid numerical ratings were retained. Entries containing incomplete metadata or corrupted values were excluded from further analysis.

This filtering process is particularly important in recommendation systems because latent factor models are highly sensitive to noisy observations. Incorrect or inconsistent ratings may distort the latent representation learned by the model and consequently reduce prediction accuracy. By removing problematic records prior to training, the reliability of the subsequent experimental results is improved.

The cleaning stage also ensured consistency between the source and target domains before semantic integration. Since the proposed framework relies on transferring knowledge between heterogeneous datasets, maintaining a clean and well-structured representation of both domains is essential.

5.1.3 Genre Enrichment of the Book Domain

One of the primary limitations of the Book-Crossing dataset is the absence of structured genre information. Since genres constitute the semantic bridge between the source and target domains, the dataset cannot be directly used for cross-domain transfer learning without additional processing.

To overcome this limitation, an external genre enrichment procedure was implemented. Metadata were retrieved using the Google Books API and the Open Library API. Book titles and authors were used as search queries, while the returned category information was converted into genre labels suitable for recommendation modelling.

Whenever genre information was unavailable from one source, the alternative source was queried in order to maximise coverage. This approach significantly increased the number of books associated with valid genre labels and enabled the construction of a semantically enriched source domain.

ALGORITHM 5.1	
GENRE ENRICHMENT PROCEDURE	
Input:	Book titles and authors
Output:	Enriched book dataset with genre labels
1:	for each book b in Book-Crossing do
2:	Query Google Books API using title and author
3:	if genre information is returned then
4:	Extract genre categories
5:	else
6:	Query Open Library API
7:	Extract genre categories
8:	end if
9:	Store retrieved genres for book b
10:	end for
11:	return Enriched book dataset

The enrichment process was applied to all books included in the experimental dataset. As shown in Table 5.2, genre information was successfully retrieved for the entire collection used in the experiments.

Table 5.2 Genre Enrichment Results

Metric	Value
Books Processed	35,220
Books Successfully Enriched	35,220
Coverage (%)	100%

5.1.4 Semantic Alignment Between Domains

Although genre information was successfully retrieved for the book dataset, the resulting genre labels were not directly compatible with those used in the MovieLens dataset. The two domains employ different genre taxonomies and naming conventions. Consequently, a semantic alignment process was required in order to establish a common representation between books and movies.

To achieve this objective, semantic embeddings were generated for both book genres and movie genres using the SentenceTransformer framework. The similarity between genre labels was subsequently measured using cosine similarity. Each book genre was then mapped to the most semantically similar movie genre, creating a unified genre space shared across the two domains.

This alignment process is fundamental to the proposed transfer learning framework. Since no users or items overlap between domains, genre information becomes the only transferable component capable of carrying knowledge from the source domain to the target domain.

The complete semantic alignment procedure is summarised in Algorithm 5.2.

ALGORITHM 5.2 SEMANTIC GENRE MAPPING	
Input: Book genres G^B , MovieLens genres G^M	
Output: Mapping from book genres to movie genres	
1:	for each book genre g_i^B in G^B do
2:	Compute embedding e_i^B using SentenceTransformer
3:	for each movie genre g_j^M in G^M do
4:	Compute embedding e_j^M
5:	Compute cosine similarity $s_{ij} = \cos(e_i^B, e_j^M)$
6:	end for
7:	Let $g_{best}^M = \arg \max_j(s_{ij})$
8:	Map g_i^B to g_{best}^M
9:	end for
10:	return Mapping M

Algorithm 5.2 describes the semantic alignment procedure used to transform heterogeneous genre labels into a common representation suitable for cross-domain learning.

The resulting genre distribution after semantic alignment is presented in Table 5.3.

Table 5.3 Distribution of Mapped Movie Genres

Genre	Frequency
Fantasy	114,397
Action	24,324
Animation	23,892
Adventure	21,587
Documentary	18,425
Sci-Fi	16,348
Romance	13,480
Children	12,684
Mystery	9,645
Musical	3,125
Comedy	2,673
Drama	1,896
Crime	1,846
Western	1,103
Horror	577
War	532
Film-Noir	510
Thriller	384
IMAX	4

Table 5.3 reveals a highly uneven genre distribution. A small number of categories, particularly Fantasy, Action, Animation, and Adventure, account for a substantial proportion of all mapped observations. In contrast, genres such as IMAX, Thriller, Film-Noir, and War appear only rarely.

This imbalance is expected given the characteristics of the source dataset and the semantic mapping procedure. However, it may influence the learning process by providing substantially more training information for dominant genres. Consequently, genre embeddings associated with highly represented categories are expected to be learned more accurately than those corresponding to rare genres.

Despite this limitation, the mapping process successfully establishes a shared semantic space that enables knowledge transfer between domains and forms the foundation of the proposed cross-domain recommendation framework.

5.1.5 Integration of Ratings and Genre Information

After semantic alignment, the mapped genre information was integrated with the corresponding rating records. This step combines collaborative information, represented by user ratings, with semantic information derived from genre labels.

The resulting dataset contains both user–item interactions and the associated genre representation for each item. Consequently, every observation includes not only a rating value but also a semantic descriptor that can be incorporated into the learning process.

This integration is essential for the proposed recommendation framework because it allows semantic information to influence latent factor learning. While traditional

collaborative filtering methods rely exclusively on user–item interactions, the proposed approach incorporates genre information as an additional source of knowledge.

Table 5.4 presents a sample of the integrated dataset after semantic alignment.

Table 5.4 Example of the Unified Dataset

User-ID	ISBN	Rating	Book Genre	Mapped Movie Genre
276725	034545104X	0	Fiction	Fantasy
276726	0155061224	5	Religion	Action
276727	0446520802	0	Fiction	Fantasy
276729	052165615X	3	Foreign Language Study	Romance
276729	0521795028	6	Foreign Language Study	Romance

The unified representation forms the foundation for all subsequent recommendation models. In particular, the genre-aware and transfer learning models directly exploit this information when constructing latent representations.

5.1.6 Movie Genre Processing

Unlike books, movies may belong to multiple genres simultaneously. For example, a movie can be classified as both *Action* and *Thriller* or *Adventure* and *Fantasy*.

Examples from the MovieLens dataset include:

Movie Genres
Adventure Animation Children Comedy Fantasy
Comedy Romance
Action Crime Thriller
Mystery Thriller
Crime Mystery Thriller

To simplify the representation and maintain consistency with the proposed framework, a single dominant genre was assigned to each movie. Specifically, the first genre listed in the MovieLens metadata was selected as the representative category.

This design choice establishes a one-to-one relationship between items and genre embeddings, reducing model complexity and simplifying the transfer learning process. Although some genre information is discarded, the resulting representation

improves computational efficiency and facilitates interpretation of the learned semantic embeddings.

5.1.7 Indexing and Matrix Representation

Before model training, all user and item identifiers were transformed into numerical indices. Recommendation algorithms based on matrix factorisation operate on numerical matrix representations rather than raw identifiers. Consequently, each user and each item must be assigned a unique index.

This transformation creates a consistent representation that is reused across all recommendation models considered in this thesis, including Singular Value Decomposition, Matrix Factorisation, Genre-Aware Matrix Factorisation, and the proposed Cross-Domain Transfer Learning model.

The indexing process also enables efficient access to user and item latent factors during optimisation.

5.1.8 Construction of the User–Item Matrix

After indexing, user–item interaction matrices were constructed for both domains.

In this representation:

- Rows correspond to users.
- Columns correspond to items.
- Matrix entries correspond to rating values.

The resulting matrices constitute the primary input for latent factor recommendation models.

Table 5.5 summarises the dimensions of the constructed matrices.

Table 5.5 User–Item Matrix Dimensions

Dataset	Matrix Dimensions
Books	(105,283 × 19)
Movies	(6,040 × 3,706)

The substantial difference in matrix dimensions illustrates the heterogeneous nature of the two domains. Furthermore, both matrices remain highly sparse, reinforcing the motivation for applying transfer learning techniques capable of operating under limited data availability.

5.1.9 Preparation for Singular Value Decomposition

To establish a baseline recommendation model, Singular Value Decomposition (SVD) was applied to the MovieLens rating matrix.

SVD decomposes the original user–item matrix into a set of lower-dimensional latent representations:

$$R \approx U\Sigma V^T$$

where:

- R denotes the original rating matrix,
- U contains user latent factors,
- Σ contains singular values,
- V^T contains item latent factors.

This decomposition enables the reconstruction of the original rating matrix using a reduced number of latent dimensions while preserving the most significant interaction patterns.

5.1.10 Evaluation strategy

The objective of the experimental evaluation is to examine the behaviour of the proposed framework under varying levels of data availability. Since the primary motivation of this thesis is to address sparsity and cold-start conditions, the evaluation procedure was specifically designed to simulate progressively richer training environments.

A 10-fold cross-validation strategy was employed to obtain reliable performance estimates and reduce evaluation bias. During each iteration, one fold was reserved for testing, while the remaining folds formed the training pool.

Rather than training the recommendation models using the entire training pool from the beginning, an incremental training strategy was adopted. The amount of available training data was gradually increased from 10% to 100% in order to analyse how recommendation accuracy evolves as additional information becomes available.

The evaluation process is summarised in Algorithm 5.3.

5.2 Model Implementation and Deep Analytical Evaluation

This section presents the implementation and evaluation of the proposed recommendation models. The analysis focuses on how each model behaves under **different levels of data availability**, using the progressive training strategy (10% to 100%) .

Performance is evaluated using **Mean Absolute Error (MAE)**, which measures the average deviation between predicted and actual ratings.

5.2.1 Singular Value Decomposition (SVD)

The first model is based on Singular Value Decomposition, which approximates the user–item rating matrix through a low-rank factorisation. The implementation constructs a dense matrix representation and applies mean-centering prior to decomposition.

The Singular Value Decomposition model approximates the user–item rating matrix through a low-rank factorisation. After mean-centering the observed ratings, the rating matrix is decomposed into latent user and item representations according to:

$$R \approx U \Sigma V^T$$

where R denotes the rating matrix, U and V represent latent user and item factors respectively, and Σ contains the singular values. Predicted ratings are reconstructed from the reduced latent representation and subsequently transformed back to the original rating scale.

Results

Table 5.6 SVD Performance under Increasing Training Availability

Training Data (%)	Average MAE
70	0.8478
80	0.8331
90	0.8187
100	0.8045

Analysis

The experimental results demonstrate a clear and consistent decrease in prediction error as the percentage of available training data increases. Specifically, the MAE decreases from approximately 0.92–0.93 at low training levels to approximately 0.80 when the full training dataset is utilised. This behaviour is observed consistently across multiple cross-validation folds, indicating stable model performance.

The results confirm that SVD benefits from increased data availability, as additional user–item interactions allow a more accurate approximation of the underlying rating structure. Nevertheless, despite the observed improvement, the overall predictive performance remains relatively limited compared with more advanced recommendation approaches.

One explanation for this limitation lies in the highly sparse nature of recommendation datasets. Under sparse conditions, the decomposition process has access to only a small fraction of the complete user–item interaction space, making it difficult to recover reliable latent representations. Consequently, the model may fail to capture meaningful preference patterns, particularly when only limited training data are available.

This effect is most evident at low training percentages, where the MAE remains above 0.90. In these scenarios, the available interactions are insufficient to accurately estimate user and item latent factors, resulting in reduced recommendation accuracy.

Even when trained using the full dataset, the model achieves an MAE of approximately 0.8045, which remains inferior to the performance of the more sophisticated models evaluated later in this chapter. These findings suggest that although SVD successfully captures global interaction patterns, its ability to handle severe sparsity is limited.

Overall, the results indicate that while SVD provides a useful baseline and improves steadily as additional data become available, it is not ideally suited for sparse recommendation environments. This limitation motivates the exploration of more advanced approaches that optimise directly on observed ratings and incorporate semantic information into the learning process.

5.2.2 Matrix Factorisation with Stochastic Gradient Descent

The second recommendation model is based on Matrix Factorisation (MF) trained using Stochastic Gradient Descent (SGD). Unlike Singular Value Decomposition, which reconstructs the rating matrix through matrix decomposition, Matrix Factorisation directly learns latent user and item representations by optimising the prediction error on observed ratings.

The underlying assumption of the model is that each user and each item can be represented by a low-dimensional latent vector. User preferences and item characteristics are therefore captured through latent factors that are learned during training.

The predicted rating is computed as:

$$\hat{r}_{ui} = p_u^T q_i$$

where:

- $\hat{r}_{ui}$ is the predicted rating of user u for item i ,
- p_u is the latent factor vector associated with user u ,
- q_i is the latent factor vector associated with item i .

Model parameters are estimated by minimising the following objective function:

$$\min_{P,Q} \sum_{(u,i) \in \Omega} (r_{ui} - p_u^T q_i)^2 + \lambda (\|p_u\|^2 + \|q_i\|^2)$$

where:

- r_{ui} denotes the observed rating,
- Ω represents the set of observed user–item interactions,
- λ is the regularisation parameter used to reduce overfitting.

The optimisation process is performed using stochastic gradient descent, which iteratively updates the latent representations until convergence.

Results

Table 5.7 Matrix Factorisation Performance under Increasing Training Availability

Training Data (%)	Average MAE
70	0.7397
80	0.7310
90	0.7260
100	0.7211

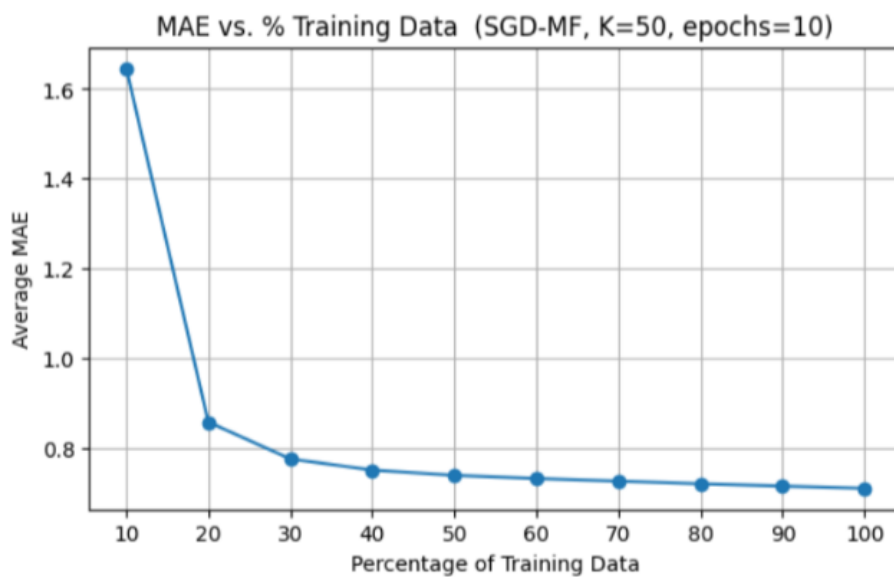


Figure 5.2.2.1: illustrates the evolution of MAE as the percentage of available training data increases.

Analysis

The experimental results clearly demonstrate a strong dependence of the model on data availability. As shown in Figure 5.2.21, the MAE decreases significantly as the percentage of training data increases.

At very low data availability (10% training data), the model performs poorly, with MAE values ranging between approximately 1.63 and 1.65 across multiple cross-validation folds. This indicates that the model is unable to learn reliable latent representations when user–item interactions are extremely sparse.

A substantial improvement is observed when the amount of training data increases from 10% to 20%, where the MAE decreases sharply to approximately 0.85–0.86. This behaviour suggests that even a relatively small increase in observed interactions significantly stabilises the learning process.

As additional training data become available, the improvement becomes more gradual. From 30% to 100%, the MAE steadily decreases from approximately 0.77 to 0.72, eventually reaching an average value of 0.7211 when the full training dataset is used. This trend is also clearly visible in Figure 5.2.2.1.

The observed behaviour reflects the fundamental characteristics of latent factor recommendation models. Since Matrix Factorisation optimises prediction error directly on observed ratings, it avoids some of the limitations associated with matrix reconstruction methods. However, its effectiveness remains strongly dependent on the availability of sufficient interaction data.

At low training percentages, users and items are associated with only a limited number of observations, resulting in unstable latent factor estimates and high prediction error. As more interactions become available, the model gradually learns more reliable preference patterns, leading to improved recommendation accuracy.

Compared with the SVD baseline, Matrix Factorisation achieves substantially better performance at higher training levels, reducing the MAE from approximately 0.80 to approximately 0.72. This result confirms the effectiveness of direct optimisation approaches for recommendation tasks.

Nevertheless, the poor performance observed under severe sparsity highlights an important limitation of the model. In cold-start environments, where very limited interaction data are available, Matrix Factorisation struggles to estimate reliable latent representations and therefore produces less accurate recommendations.

Overall, the results demonstrate that Matrix Factorisation provides a significant improvement over SVD once sufficient training data become available. However, its sensitivity to sparse interactions motivates the incorporation of additional semantic information, which is explored in the following section through the Genre-Aware Matrix Factorisation model.

5.2.3 Genre-Aware Matrix Factorisation

To address the limitations of purely collaborative recommendation models, semantic information is incorporated into the learning process through genre embeddings. Unlike standard Matrix Factorisation, where predictions depend solely on user and item latent factors, the proposed model enriches item representations with genre-based semantic information.

The rating prediction function is defined as:

$$\hat{r}_{ui} = p_u^T (q_i + g_i)$$

where:

- $\hat{r}_{ui}$ is the predicted rating of user u for item i ,
- p_u denotes the latent representation of user u ,
- q_i denotes the latent representation of item i ,
- g_i represents the genre embedding associated with item i .

In this formulation, item representations are augmented with semantic information derived from genre categories. Consequently, items belonging to the same genre are able to share information during training, allowing the model to learn more robust representations under sparse-data conditions.

The latent user factors, item factors, and genre embeddings are learned simultaneously through stochastic gradient descent by minimising the prediction error on observed ratings.

Results

Table 5.8 Genre-Aware Matrix Factorisation Performance

Training Data (%)	Average MAE
70	0.7269
80	0.7233
90	0.7169
100	0.7108

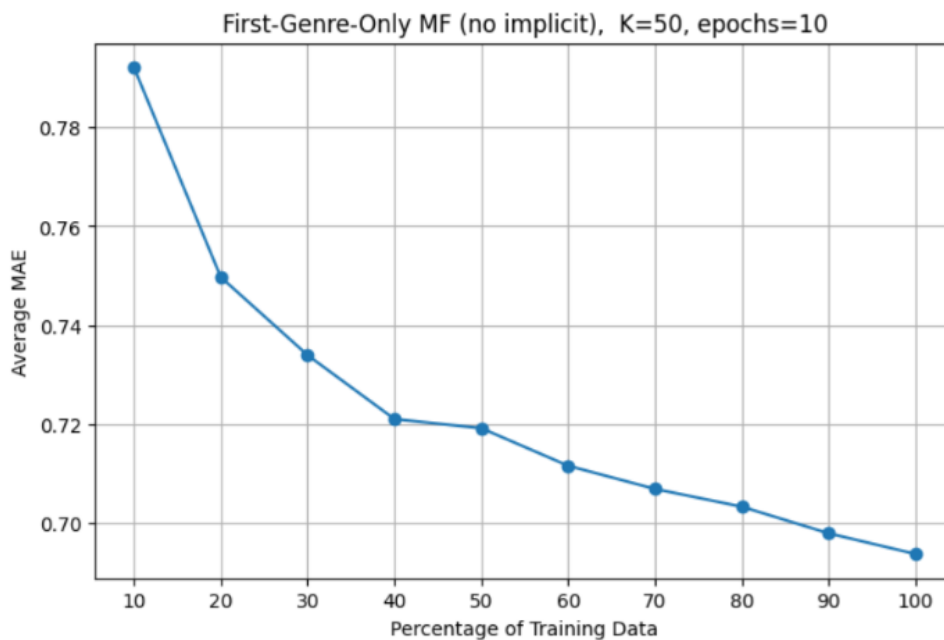


Figure 5.2.3.1: MAE performance of the Genre-Aware Matrix Factorisation model under increasing training availability.

Analysis

The experimental results demonstrate a clear and consistent improvement in recommendation performance when semantic information is incorporated into the learning process. As shown in Figure 5.2.3.1, the MAE decreases steadily as the percentage of training data increases, falling from approximately 0.79 at 10% training availability to 0.7108 when the full dataset is used.

Compared with the standard Matrix Factorisation model, the improvement is particularly pronounced under sparse-data conditions. At 10% training availability, the MAE decreases from approximately 1.64 for the baseline Matrix Factorisation model to approximately 0.79 for the Genre-Aware model, representing a substantial improvement. This result indicates that semantic genre information effectively mitigates the sparsity problem by providing an additional source of information beyond observed ratings.

The learning curve is also noticeably smoother than that of the standard Matrix Factorisation model. Whereas the baseline model exhibits a dramatic improvement between 10% and 20% training data, the Genre-Aware model demonstrates a more gradual and stable reduction in prediction error. This behaviour suggests that semantic information provides useful guidance even when only a limited number of user–item interactions are available.

The observed improvement can be explained by the role of genre embeddings in the prediction function. Rather than relying exclusively on individual item representations, the model incorporates shared semantic representations associated with genre categories. Consequently, items belonging to the same genre contribute to the learning of common latent structures, allowing information to be shared across semantically related items.

From a theoretical perspective, genre embeddings introduce an inductive bias that guides the optimisation process toward semantically meaningful relationships. This reduces the model’s dependence on direct user–item interactions and improves generalisation in sparse recommendation environments.

As the amount of training data increases, the performance gains become more gradual, with MAE decreasing from approximately 0.75 at 20% training availability to approximately 0.71 at full training availability. This observation suggests that semantic information is particularly valuable in low-data scenarios, although it continues to provide measurable benefits even when sufficient interaction data are available.

Overall, the results confirm that the incorporation of semantic genre information constitutes one of the most effective improvements introduced in the proposed framework. The Genre-Aware Matrix Factorisation model consistently outperforms both SVD and standard Matrix Factorisation across all training percentages, demonstrating its effectiveness in addressing data sparsity and improving recommendation accuracy.

5.2.4 Cross-Domain Transfer Learning Model

The final recommendation model extends the Genre-Aware Matrix Factorisation framework by incorporating transfer learning from an external source domain. Unlike the previous model, where genre embeddings are learned exclusively from movie ratings, the proposed approach exploits semantic knowledge acquired from the book domain before training begins in the target domain.

The transfer process starts by training the Genre-Aware Matrix Factorisation model on the Book-Crossing dataset. During this stage, latent user factors, item factors, and genre embeddings are learned from book ratings. The resulting genre embeddings capture semantic relationships between content categories and user preferences within the source domain.

Once training on books is completed, the learned genre embeddings are transferred to the movie domain and used as an informed initialisation for the corresponding movie genre representations. The movie recommendation model is then fine-tuned using MovieLens ratings while preserving the semantic knowledge acquired from the source domain.

The overall transfer learning procedure is illustrated in Algorithm 5.3

ALGORITHM 5.3
CROSS-DOMAIN TRANSFER LEARNING PROCEDURE
Input: Book-Crossing ratings, MovieLens ratings, shared genre space Output: Fine-tuned movie recommendation model
1: Train Genre-Aware MF on Book-Crossing 2: Learn genre embeddings G_{books} 3: Transfer embeddings to the movie domain $G_{movies} \leftarrow G_{books}$ 4: Initialise the movie model using G_{movies} 5: Fine-tune the model on MovieLens ratings 6: Generate predictions for the test set 7: Compute MAE on the test set 8: Return the trained movie model

Results

Table 5.9 Cross-Domain Transfer Learning Performance

Training Data (%)	Average MAE
70	0.7069
80	0.7014
90	0.6983
100	0.6938

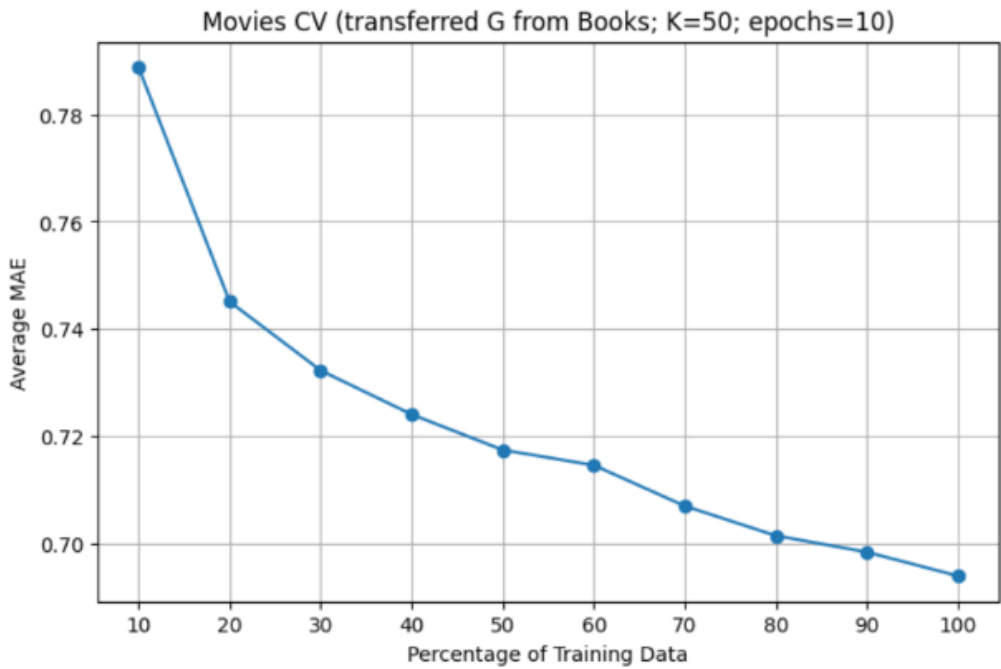


Figure 5.2.4.1: MAE performance of the Cross-Domain Transfer Learning model under increasing training availability.

Analysis

The experimental results indicate that transferring semantic knowledge from the book domain provides a further improvement over the genre-aware recommendation model. As shown in Figure 5.2.4.1, the MAE decreases steadily as the percentage of available training data increases, reaching a final value of approximately 0.6938 when the full training dataset is used.

Unlike the Genre-Aware Matrix Factorisation model, which learns genre embeddings exclusively from movie ratings, the proposed transfer learning framework begins with genre representations that have already been trained in the source domain. Consequently, the optimisation process starts from a more informative initial state, reducing the amount of target-domain information required to learn meaningful semantic relationships.

The benefits of transfer learning are particularly evident under sparse-data conditions. At low training percentages, the transferred embeddings provide useful prior knowledge regarding genre preferences and content relationships. This reduces uncertainty during optimisation and enables the model to achieve stable performance even when only a limited number of movie ratings are available.

The learning curve also exhibits a high degree of stability throughout the evaluation process. The gradual reduction in MAE suggests that the transferred semantic representations remain useful across all levels of data availability and continue to support learning even as more target-domain interactions become available.

An important characteristic of the proposed framework is that knowledge transfer is achieved without requiring overlapping users or overlapping items between domains. The transfer mechanism relies exclusively on semantic genre information, demonstrating that shared content representations can serve as an effective bridge between otherwise disconnected recommendation environments.

Compared with both SVD and standard Matrix Factorisation, the proposed approach consistently achieves lower prediction errors. Furthermore, the model slightly improves upon the Genre-Aware Matrix Factorisation baseline, indicating that semantic information becomes even more valuable when enriched with knowledge learned from an external domain.

Overall, the results confirm the central hypothesis of this thesis: semantic genre representations learned in one domain can be successfully transferred to another domain and contribute to improved recommendation accuracy. The findings demonstrate that cross-domain transfer learning constitutes an effective strategy for addressing sparsity and cold-start problems in recommendation systems, even when no users or items are shared between domains.

5.2.5 Comparative Interpretation

The comparative evaluation of the implemented models demonstrates a clear progression in recommendation performance as additional semantic and cross-domain information is incorporated into the framework. Figure X illustrates the evolution of MAE across different percentages of training data for all evaluated approaches.

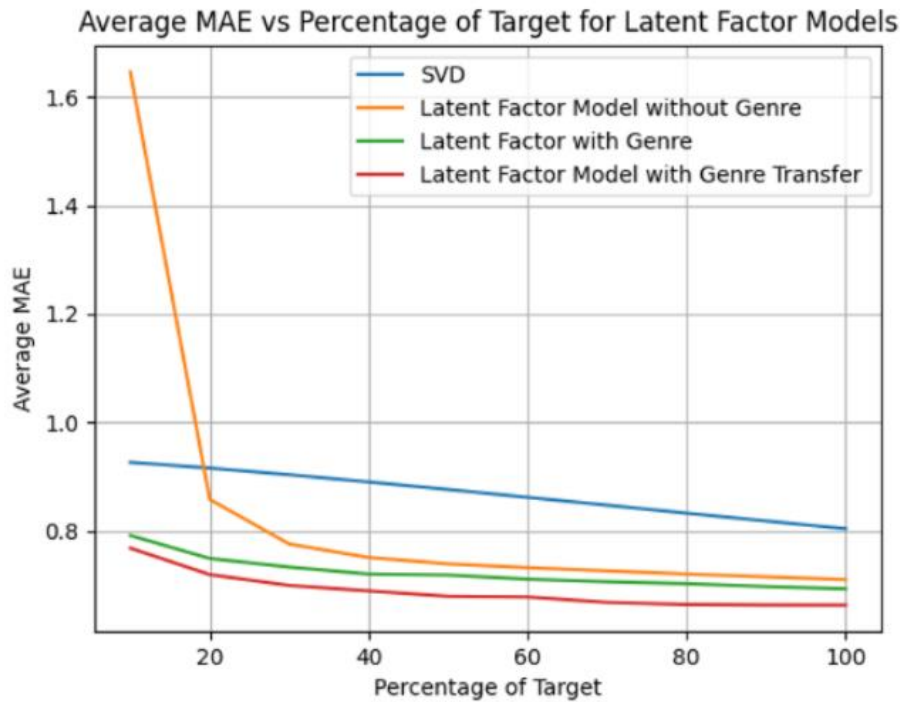


Figure 5.2.5: Comparative MAE performance of all recommendation models under increasing training data availability.

The SVD model exhibits the weakest overall performance throughout the experiments. Although the MAE gradually decreases as more training data become available, the model consistently remains significantly above the performance of the other approaches. This behaviour confirms the limitations of dense matrix reconstruction methods in sparse recommendation environments, where missing values dominate the interaction matrix.

The standard latent factor model trained with stochastic gradient descent performs substantially better at higher training percentages, eventually achieving MAE values close to 0.71. However, its behaviour under sparse conditions is highly unstable, with extremely poor performance at 10% training data ($\text{MAE} \approx 1.63$). This result demonstrates that traditional collaborative filtering methods are highly dependent on sufficient interaction data and struggle severely in cold-start scenarios.

The introduction of semantic genre information produces the most substantial improvement in the entire experimental framework. The genre-aware latent factor model significantly reduces prediction error across all training configurations, demonstrating that semantic representations allow the model to generalise more effectively even when direct collaborative signals are limited. By enabling information sharing across semantically related items, the model becomes considerably more robust against sparsity.

The best overall performance is achieved by the proposed cross-domain transfer learning model. As shown in Figure X, the transfer-based approach consistently produces the lowest MAE values across all percentages of training data, outperforming every baseline model throughout the experiments. The model achieves approximately **0.788 MAE at 10% training data** and improves further to approximately **0.693 MAE at full training data**.

This result is particularly important because it directly validates the central objective of this thesis: improving recommendation quality in sparse environments through semantic transfer learning across heterogeneous domains.

Unlike conventional recommendation systems, the proposed framework successfully transfers semantic knowledge learned from the book domain into the movie domain through shared genre embeddings. Even though the two domains contain no overlapping users or items, the model is still able to exploit transferable semantic structures to improve prediction accuracy. This demonstrates that semantic representations can function as an effective bridge between otherwise disconnected recommendation domains.

The experimental results further indicate that transfer learning is especially valuable under low-data conditions, where traditional collaborative filtering methods suffer from insufficient information. In sparse settings, the transferred genre embeddings provide semantically informed initialisation, reducing uncertainty during optimisation and improving model stability from the earliest stages of training.

Another important observation is that the transfer learning model not only achieves the lowest overall MAE, but also produces the most stable learning behaviour across increasing training percentages. While other models exhibit sharper performance fluctuations, the proposed approach maintains a smoother and consistently decreasing error curve, indicating improved generalisation capability.

Overall, the comparative analysis confirms that the proposed transfer learning framework successfully addresses the sparsity problem that motivated this research. The integration of semantic genre embeddings together with cross-domain knowledge transfer enables the recommendation system to achieve superior predictive performance compared to both traditional collaborative filtering methods and non-transfer semantic models.

The observed improvements under sparse-data conditions are consistent with previous findings in the cross-domain recommendation literature. Similar to the conclusions reported by Enrich et al. (2013), the results indicate that auxiliary knowledge originating from a source domain can contribute to improved recommendation quality in cold-start scenarios.

Chapter 6 Conclusions and Future Work

6.1 Conclusions

This thesis investigated the use of semantic transfer learning for cross-domain recommendation in fully non-overlapping environments. The proposed framework aimed to improve movie recommendation performance under sparse-data conditions by transferring semantic genre knowledge learned from a book recommendation domain.

Unlike many traditional cross-domain recommendation approaches, the proposed architecture does not rely on overlapping users or shared items between domains. Instead, knowledge transfer is achieved exclusively through semantic genre embeddings. This design allows transfer learning to remain feasible even when collaborative overlap between domains is completely absent.

To support this objective, the framework combined: collaborative filtering, semantic genre modelling, and transfer learning

within a unified recommendation architecture.

The system first enriched and semantically aligned genre information across the Book-Crossing and MovieLens datasets. A genre-aware Matrix Factorization model was then trained on the source-domain book ratings in order to learn semantic genre embeddings. These embeddings were subsequently transferred to the movie recommendation model and fine-tuned using target-domain ratings.

The experimental evaluation compared: standard collaborative filtering baselines, genre-aware recommendation models, and the proposed transfer learning architecture under progressively increasing sparsity conditions.

The results demonstrated that semantic genre information improves recommendation robustness in sparse environments. More importantly, the transfer learning model consistently achieved the best performance under low-data conditions by using pretrained genre embeddings as informative semantic initialization.

The findings therefore support the central hypothesis of the thesis: semantic transfer learning through genre embeddings can improve recommendation performance in fully non-overlapping recommendation environments.

Overall, the proposed framework demonstrates that transferable semantic representations can partially compensate for the absence of collaborative overlap between domains and improve latent representation learning when target-domain interaction data are limited.

6.2 Limitations

Although the proposed framework achieved encouraging results, several limitations remain.

First, the recommendation process relies on a simplified single-genre representation for each item. In practice, many books and movies belong to multiple genres simultaneously, and richer multi-label semantic representations could potentially improve recommendation quality.

Second, semantic genre alignment depends on external metadata retrieved through third-party APIs. Consequently, the quality of the semantic mapping process is influenced by the completeness and consistency of the retrieved metadata.

Third, the framework transfers only semantic genre embeddings and does not incorporate additional contextual information such as textual descriptions, user demographics, or temporal interaction behaviour.

Another limitation concerns the use of Matrix Factorization as the underlying recommendation architecture. Although latent factor models remain highly effective and interpretable, more advanced deep learning recommendation models may capture more complex semantic relationships between users and items.

Finally, the evaluation focused specifically on book-to-movie transfer learning. Different domain combinations may exhibit different semantic transfer characteristics and require alternative alignment strategies.

6.3 Future Work

Several directions may extend and improve the proposed framework in future research.

One possible extension involves the use of multi-genre semantic representations instead of single dominant genre assignments. This would allow the recommendation model to capture richer semantic relationships between items.

Another potential improvement is the integration of textual information such as plot summaries, reviews, or descriptions through transformer-based language models. Combining textual semantic representations with collaborative filtering could further improve recommendation quality under sparse conditions.

Future work may also explore more advanced neural recommendation architectures, including:

- deep collaborative filtering,
- graph neural networks,
- and attention-based recommendation models.

Additionally, future research could investigate transfer learning across multiple recommendation domains simultaneously rather than using a single source-target relationship.

Finally, alternative semantic transfer mechanisms beyond genre embeddings could be explored, including multimodal semantic representations derived from text, images, or user-generated content.

These directions could further improve recommendation robustness and expand the applicability of transfer learning in sparse and fully non-overlapping recommendation environments.

REFERENCES

- [1] Adomavicius, G. and Tuzhilin, A. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), pp.734–749.
- [2] Ricci, F., Rokach, L. and Shapira, B. (2015). *Recommender Systems Handbook*. 2nd ed. New York: Springer.
- [3] Gomez-Uribe, C.A. and Hunt, N. (2015). The Netflix Recommender System: Algorithms, Business Value, and Innovation. *ACM Transactions on Management Information Systems*, 6(4), pp.1–19.
- [4] Koren, Y., Bell, R. and Volinsky, C. (2009). Matrix Factorization Techniques for Recommender Systems. *Computer*, 42(8), pp.30–37.
- [5] Pan, S.J. and Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), pp.1345–1359.
- [6] Zhang, Y., Yao, Q. and Sun, A. (2019). A Survey of Deep Learning Based Recommender Systems. *ACM Computing Surveys*, 52(1), pp.1–38.
- [7] Cantador, I., Fernández-Tobías, I., Berkovsky, S. and Cremonesi, P. (2015). Cross-Domain Recommender Systems. In: *Recommender Systems Handbook*. Springer, pp.919–959.
- [8] Weiss, K., Khoshgoftaar, T.M. and Wang, D. (2016). A Survey of Transfer Learning. *Journal of Big Data*, 3(1), pp.1–40.
- [9] Fernández-Tobías, I., Cantador, I., Kaminskis, M. and Ricci, F. (2012). Cross-Domain Recommender Systems: A Survey of the State of the Art.
- [10] Singh, A.P. and Gordon, G.J. (2008). Relational Learning via Collective Matrix Factorization. *Proceedings of KDD 2008*, pp.650–658.
- [11] Man, T., Shen, H., Jin, X. and Cheng, X. (2017). Cross-Domain Recommendation: An Embedding and Mapping Approach. *IJCAI 2017*.
- [12] Ruder, S. (2017). An Overview of Multi-Task Learning in Deep Neural Networks. *arXiv:1706.05098*.
- [13] Fernández-Tobías, I. and Cantador, I. (2014). Exploiting Social Tags in Cross-Domain Recommendation. *UMAP 2014*.
- [14] Wu, L., He, X., Wang, X. et al. (2022). Graph Neural Networks in Recommender Systems: A Survey. *ACM Computing Surveys*, 55(5).
- [15] Devlin, J., Chang, M.W., Lee, K. and Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL-HLT 2019*.
- [16] Koren, Y., Bell, R. and Volinsky, C. (2009). Matrix Factorization Techniques for Recommender Systems. *Computer*, 42(8), pp.30–37.
- [17] Schafer, J.B., Frankowski, D., Herlocker, J. and Sen, S. (2007).

- [18] Cantador, I., Fernández-Tobías, I., Berkovsky, S. and Cremonesi, P. (2015).
- [19] Reimers, N. and Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT Networks.
- [20] H. Ko, S. Lee, Y. Park, and A. Choi, “A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields,” *Electronics*, vol. 11, no. 1, Article 141, 2022.
- [21] M. Enrich, M. Braunhofer, and F. Ricci, “Cold-Start Management with Cross-Domain Collaborative Filtering and Tags,” *ResearchGate*, Aug. 2013.
- Available: https://www.researchgate.net/publication/289916807_Cold-Start_Management_with_Cross-Domain_Collaborative_Filtering_and_Tags
- [22] Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2019). *A comprehensive survey on transfer learning*. arXiv:1911.02685.
- [23] Reimers, N. and Gurevych, I., 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks.
- [24] Manning, C.D., Raghavan, P. and Schütze, H., 2008. *Introduction to Information Retrieval*. Cambridge University Press.
- [25] Pan, S.J. and Yang, Q., 2010. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), pp.1345–1359.
- [26] Li, B., Yang, Q. and Xue, X., 2009. Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction. *Proceedings of IJCAI*.
- [27] Koren, Y., Bell, R. and Volinsky, C., 2009. *Matrix Factorization Techniques for Recommender Systems*.