5-16-2024

A Dual-Transformation Strategy for Team Recommendation

Zahra Kheiryashkuh
University of Windsor

A Dual-Transformation Strategy for Team Recommendation

By

Zahra Kheiryashkuh

A Thesis
Submitted to the Faculty of Graduate Studies
through the School of Computer Science
in Partial Fulfillment of the Requirements for
the Degree of Master of Science
at the University of Windsor

Windsor, Ontario, Canada

2024

A Dual-Transformation Strategy for Team Recommendation

by

Zahra Kheiryashkuh

APPROVED BY:

M. Hassanzadeh
Department of Electrical and Computer Engineering

S. Samet
School of Computer Science

Z. Kobti, Co-Advisor
School of Computer Science

K. Selvarajah, Co-Advisor
School of Computer Science

May 9, 2024

DECLARATION OF CO-AUTHORSHIP/PREVIOUS PUBLICATION

1. Co-Authorship

I hereby declare that this thesis incorporates material that is the result of research conducted under the supervision of Dr. Ziad Kobti and Dr. Kalyani Selvarajah. In all cases, the key ideas, primary contributions, experimental designs, data analysis, and interpretation were performed by the author, and the contribution of co-authors was primarily through the proofreading of the published manuscripts.

I am aware of the University of Windsor Senate Policy on Authorship, and I certify that I have properly acknowledged the contribution of other researchers to my thesis and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is a product of my own work.

2. Submitted Publication

This thesis includes the paper that has been submitted in **27th ECAI (European Conference on Artificial Intelligence)**, as follows:

Paper Title	Publication Status
Z. Kheiryashkuh, Z. Kobti, and K. Selvarajah. Unlocking Team Potential: Transfer Knowledge from Popular to Unpopular Experts.	Submitted (To appear in the proceedings of the ECAI conference, 19-24 October 2024)

I certify that I have obtained written permission from the copyright owner(s) to include the above-submitted material(s) in my thesis. I certify that the above material describes work completed during my registration as a graduate student at the University of Windsor.

3. General

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

In the field of collaborative work, effectively forming teams is crucial for achieving success. The challenge of constructing cohesive teams lies in selecting from a vast array of potential candidates, each with distinct skills, experiences, and personal attributes.Team recommender systems are designed to pinpoint the ideal combination of experts who collectively meet the skill requirements necessary to achieve a common goal. Recently, researchers have started to examine this problem through neural architectures that recommend the team of experts by learning a relationship between the skills and experts space. However, these models often exhibit a popularity bias, which refers to the tendency of the systems to recommend disproportionately more popular teams. We have introduced a dual transfer strategy to enhance team recommendation performance, which involves transferring knowledge from head instances to tail instances at both the model and instance levels. At the model level, the strategy creates a meta-mapping from few-shot to many-shot models, which indirectly improves data quality and enhances learning representations for teams that are less popular by implicitly augmenting data at the model level. The proposed dual knowledge transfer method at the instance level employs curriculum learning to bridge the gap between popular and not popular instances, ensuring a smooth transition of meta-mapping head teams to the tail ones. Our evaluation criteria is that we expect to improve team recommendation quality, particularly for teams that are in the tail of the distribution. We demonstrate how our proposed platform effectively addresses the issue of popularity bias prevalent in current team recommendation methodologies. Further more, we validate its effectiveness by comparing it with leading approaches, employing the DBLP dataset in our analysis.

ACKNOWLEDGEMENTS

I wish to genuinely convey my deepest appreciation to my co-supervisors, Dr. Kobti and Dr.Kalyani. They had a deep grasp of research challenges, which taught me a lot, and I appreciate their help and support throughout my research work. I extend special thanks to Dr. Saeed Samet, serving as the internal reader, and Dr. Mohammad Hassanzadeh, the external reader of my thesis committee, for their constructive feedback and invaluable recommendations. Lastly, I express my gratitude to the School of Computer Science and all those involved who supported me throughout this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

MRR	Mean Reciprocal Rank
MAP	Mean Average Precision
NDCG	Normalized Discounted Cumulative Gain

CHAPTER 1

Introduction

1.1 Background

1.1.1 Concept of Collaborative Work

Collaborative work refers to the process where individuals or groups work together to achieve shared goals. This concept is fundamental in various fields such as business, education, technology, and research, where the collective effort of a group can often produce outcomes that are superior to what could be achieved individually. Collaboration is a strategic asset across different fields which enhances productivity, innovation, and competitiveness.

The essence of collaborative work lies in harnessing diverse perspectives, expertise, and resources to address complex challenges and drive progress. The power of collaborative work lies in its ability to amalgamate individual strengths to achieve a common goal where synergy among team members can lead to innovative solutions and enhanced productivity.

A collaborative network defines the connections and interactions among various individuals, groups, organizations, and communities, all focused on collective objectives or complex problem-solving. Such interactions may stem from shared objectives, resources, skills, and formal relationships, including those among business partners, academic research groups, or interdisciplinary teams. Analyzing these networks helps to uncover patterns and dynamics within a community, highlighting how these con-

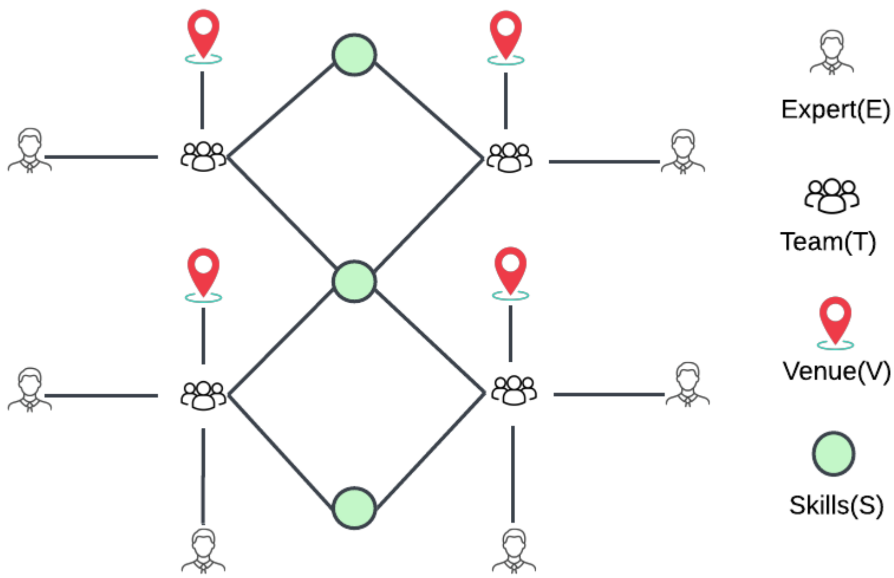

Fig. 1.1.1: DBLP Heterogeneous Collaborative Network

nections impact results, productivity, and creativity. This analysis proves valuable in numerous areas, including project coordination, fostering innovation, forming strategic partnerships, and beyond.

Collaborative network analysis and the Team Formation Problem are interconnected disciplines focused on analyzing relationships and interactions among individuals, groups, organizations, and communities. While collaborative network analysis delves into the structural nuances and characteristics of these networks, the Team Formation Problem focuses on the intricate task of curating the optimal mix of participants to form effective teams.

To capture the structural and semantic complexities within the DBLP dataset , we created a comprehensive collaboration network as Figure 1.1.1 illustrates. This network serves as a dynamic framework that intricately interconnects various nodes, including skills, experts, papers, and their corresponding venues. This network forms the foundation for insightful analysis and effective team formation strategies, contributing to our proposed team recommendation system. It represents a significant step forward in our quest to develop a robust team recommendation system. By leveraging the rich network of connections established within the DBLP dataset, our recommendation system aims to provide tailored suggestions for assembling high-performing teams.

These concepts are crucial across various sectors such as sports, business, and education, where constructing competent teams is pivotal to achieving success. In these areas, collaborative network analysis offers critical insights into the interconnections and dynamics among individuals, which is essential for the team formation process. It helps in pinpointing potential team members who not only have robust connections within the network but also have the essential skills, abilities, and preferences necessary for cohesive and effective teamwork.

The growing complexity of contemporary challenges across various sectors underscores the critical importance of effective collaboration. As projects expand in scope and ambition, the demand for diverse skills and perspectives intensifies. Consequently, the success of these collaborative endeavors largely depends on the careful analysis of the network's composition. Understanding the structure and dynamics of collaborative networks is essential for optimizing team performance, enhancing innovation, and ensuring the efficient achievement of shared goals. This analysis of collaborative networks becomes pivotal in orchestrating effective partnerships and strategic alignments within diverse teams.

1.1.2 Importance of team composition in achieving success in collaborative environments

The composition of a team plays a pivotal role in its success, particularly in collaborative environments. Teams that are well-composed, considering the right mix of skills, experience levels, and personality traits, tend to perform more efficiently and produce higher quality outcomes. Team composition plays a pivotal role in ensuring that the collaborative unit is well-equipped to meet and exceed project objectives. The right team composition enhances problem-solving capabilities, fosters creativity, and improves decision-making, contributing significantly to the success of collaborative projects.

The importance of team composition extends beyond the mere assembly of skills. It involves the integration of diverse cognitive abilities and emotional intelligences, which enhances the team's collective problem-solving capabilities. When team members bring different perspectives and approaches to the table, they collectively create a richer pool of resources to draw from in creativity and innovation. This diversity leads to more robust discussions, more creative solutions, and more effective decision-making.

Additionally, the right team composition significantly influences the motivation and engagement of its members. When individuals feel that their skills are being used effectively, and they are contributing to the success of the team, their satisfaction and productivity rise. Moreover, teams that include a balance of roles—leaders, strategists, facilitators, and executors—can navigate project challenges more smoothly and efficiently. Each member plays a critical part in driving the project forward, ensuring that all aspects of project management are covered.

Furthermore, team composition is crucial for the adaptive capacity of the team. In today's fast-paced and often unpredictable project environments, teams must be able to adjust quickly to new information or changing circumstances. Teams that possess a dynamic mix of adaptability, technical skills, and interpersonal abilities are better prepared to respond to challenges and seize opportunities as they arise.

In essence, the right team composition is foundational to fostering a productive and harmonious work environment. It underpins the team's ability to innovate, solve problems effectively, and achieve the goals set before them, thereby playing a pivotal role in the success of collaborative projects.

The importance of team composition cannot be overstated, as it fundamentally shapes the team's ability to innovate, solve problems effectively, and achieve project goals. A well-composed team is not only harmonious but also highly productive, benefiting from a rich blend of skills, experiences, and perspectives that enhance creativity and decision-making. However, achieving such an ideal composition is fraught with challenges.

1.1.3 Challenges in Team Formation

In the rapidly evolving realm of collaborative projects, the assembly of teams with a harmonious blend of expertise, diverse backgrounds, and complementary personality traits is crucial for project success. Effective team formation transcends mere group assembly; it requires strategic alignment of skills and personalities to achieve synergy and operational excellence. The challenge, however, lies in navigating the vast pool of potential candidates to curate the optimal team composition, a process complicated by the complexity and variability of human factors and professional skills.

As organizations and projects become more interdisciplinary and interconnected, the importance of crafting such balanced teams becomes even more pronounced. The diverse nature of modern projects, which often blend technical, creative, and strategic elements, demands a team that is not only skilled in multiple domains but also adept at communication and collaboration across these domains. Moreover, the dynamic nature of project requirements necessitates a flexible approach to team composition—one that can evolve as project objectives and environments change.

However, identifying the right mix of abilities and traits is only part of the challenge; equally critical is understanding and managing the interpersonal dynamics that can significantly impact team performance. Research indicates that teams with a strong relational foundation tend to perform better, particularly when they face complex and uncertain environments that require high levels of coordination and problem-solving. This underlines the need for systems that can effectively analyze and predict team success by considering the goals and required skills for a project.

1.1.4 Evolution of Team Formation Methods

Historically, team formation was predominantly a manual and intuitive process, heavily reliant on the personal judgment and experience of team leaders or managers. Before the advent of advanced computational tools, forming teams was often challenging, especially in large organizations with diverse talent pools. Leaders had to consider various factors such as expertise, availability, and interpersonal dynamics, which made the process time-consuming and prone to bias.

The limitations of traditional graph-based methods have paved the way for more sophisticated team recommender systems. Initially, researchers employed graph-based search methods to tackle the problem of team formation by extracting a subgraph from collaboration networks. These networks represent various experts, their skills, and their past collaborative efforts. Techniques such as heuristics and meta-heuristics were utilized to efficiently extract teams from the network [6, 7, 13, 14, 5].

Sozio et al. [16] further elaborate on this concept by describing how an optimal team is represented as a subgraph that maximizes a monotonic function while adhering to specific constraints. This underscores an analytical approach to team assembly that strives for optimal functional performance while meeting predefined criteria.

Despite their practical effectiveness, graph-based methods are specifically tailored to certain problems and often rely on strong assumptions about team dynamics. For example, Lappas et al. [6] proposed a model assessing team effectiveness based on communication costs within a social graph. They posited that lower communication costs generally lead to more successful collaboration. However, this is not universally true, as seen in the DBLP collaboration network where researchers with a record of successful publications might incur higher communication costs compared to less published individuals who still meet all task requirements. This observation indicates that minimizing communication costs does not consistently translate into successful outcomes. Furthermore, while graph-based methods might recommend teams with the necessary skills for a task, there is no guarantee these teams will collaborate successfully. This limitation primarily stems from the strong assumptions these methods make about successful teams. Therefore, although such methods can assemble teams with the appropriate skill mix, their foundational assumptions may not hold in practical scenarios, leading to less effective team performance.

Recent studies by Kargar et al. [17] and Bryson et al. [18] focus on strategies to minimize the sum of the weights in the induced subgraphs, a key technique in optimizing team structures within networks. However, a significant limitation of these methods is their computational intensity. This complexity mainly stems from the fact that optimizing these subgraphs relates closely to solving the Steiner tree problem, which is known to be NP-hard—a classification indicating that these problems are computationally demanding and challenging to solve efficiently. This computational complexity arises because the Steiner tree problem requires finding the smallest subset of nodes that connects a given set of terminals, a task that becomes exponentially more difficult as the size of the network increases.

Moreover, as networks grow and evolve, maintaining the efficiency of graph-based algorithms becomes increasingly difficult. These methods must frequently recalculate communication costs among individuals due to any changes in the network structure, which can severely limit their scalability. This issue is particularly problematic in dynamic environments where network changes are frequent, as highlighted by recent findings from Kaw et al. [19]. Each alteration in the network necessitates a complete reassessment of the optimal pathways and node connections, significantly increasing the processing time and computational resources required. This scalability challenge is a critical barrier to the practical deployment of these methods in large-scale, real-time applications where network dynamics are constantly evolving.

Given the limitations of traditional graph-based methods, researchers have increasingly turned to neural machine learning models that can learn the complex relationships between experts and their skills using advanced neural architectures. The primary objective of these studies is to create an efficient mapping from the skill set domain to the expert domain, thereby facilitating the swift formation of expert teams. A notable advancement in this area was the introduction of an autoencoder architecture by Sapienza et al.[20], which is designed to identify experts by learning from the collaborative patterns and skill distributions of team members.

However, the application of traditional autoencoder architectures in this context faces significant challenges, particularly due to the sparse distribution of skills across experts and teams. Sparsity can lead to models that do not generalize well beyond their training data, resulting in overfitting. Overfitting occurs when a model captures noise or random fluctuations in the training data instead of the underlying pattern, leading to poor performance on new, unseen data. This is a critical concern as it can severely compromise the model's ability to effectively recommend team compositions.

This problem primarily arises due to the sparse nature of input skills coupled with the long-tailed distribution of output experts, where a small number of experts account for the majority of successful collaborations, while the vast majority have only participated in a few, as illustrated in Figure 2.1.1. In response to the challenge of sparse skill distributions, Rad et al. [21] built upon the initial work by Sapienza et al., introducing a variational Bayesian neural framework. This approach was shown to manage sparsity issues more effectively, leading to the formation of higher-quality teams. While there are subtle differences across various models [22, 23], the overall improvement in evaluation metrics across these models does not significantly reflect the sophistication of their methodologies.

Furthermore, these models enhance their performance by incorporating contextual information from the DBLP heterogeneous network through techniques like metapath2vec [24], which captures the representations of nodes within the collaboration network. However, since metapath2vec depends on random walks [25] to generate node embeddings, it categorizes as a shallow graph embedding technique [26], which operates as a self-supervised learning method. Shallow graph embedding approaches face several challenges: the training complexity of these methods is directly proportional to the network size, $O(|V|)$, which can render them inefficient. Additionally, they are unable to generate embeddings for new nodes without re-training, posing challenges for dynamically evolving networks that exceed memory capacity. They also tend to overlook the importance of node attributes, crucial for understanding each node's role within the network.

To address these shortcomings, Kaw et al. [19] proposed the LANT network, which employs Graph Attention Networks (GAT) to learn graph structural features from the heterogeneous network and recommend experts for teams. Although this approach has enhanced the adaptability of network models post-training and reduced the time complexity of the learning process, the issue of popularity bias remains a significant challenge.

Popularity bias, largely driven by the long-tail distribution in data, tends to degrade model performance, especially when recommending experts from the less frequently represented tail segment of the data distribution. To mitigate this and enhance the quality of recommended teams, we propose a two-level knowledge transfer learning framework specifically tailored for team recommendations. This framework is particularly aimed at improving recommendations for less popular teams, which are often overlooked in typical recommendation systems.

Our two-level knowledge transfer learning framework is designed to predict the performance of teams based on a specific set of skills, utilizing a synergistic combination of model and skill level transfer learning. This approach is further augmented by meta-learning and curriculum learning techniques to effectively tackle the long tail distribution challenge. The architecture of the proposed team recommendation system consists of two primary components: model-level knowledge transfer (including the base learner and the meta-learner) and team-level knowledge transfer, which is facilitated through curriculum learning. Training of the base learner incorporates both many-shot and few-shot learning approaches, allowing the model to adapt and generate meaningful team embeddings, as illustrated in Figure 1.1.2. These embeddings provide a sophisticated representation that captures both explicit data points and the complex network of team relationships.

The contributions of this work are threefold:

- We introduce a two-level knowledge transfer learning framework that synergistically integrates knowledge transfer at both the model and team levels. This integration ensures that knowledge can be effectively utilized across teams, spanning from the most to the least popular in a long-tail distribution context.

- We develop a new curriculum designed to seamlessly transition the meta-mapping learned from more frequent (head) data instances to those in the long-tail, enhancing model performance across the entire distribution.

- Extensive experimental results demonstrate that our proposed framework significantly enhances recommendations for tail teams while also achieving robust improvements in overall and head team performances, as evaluated by standard

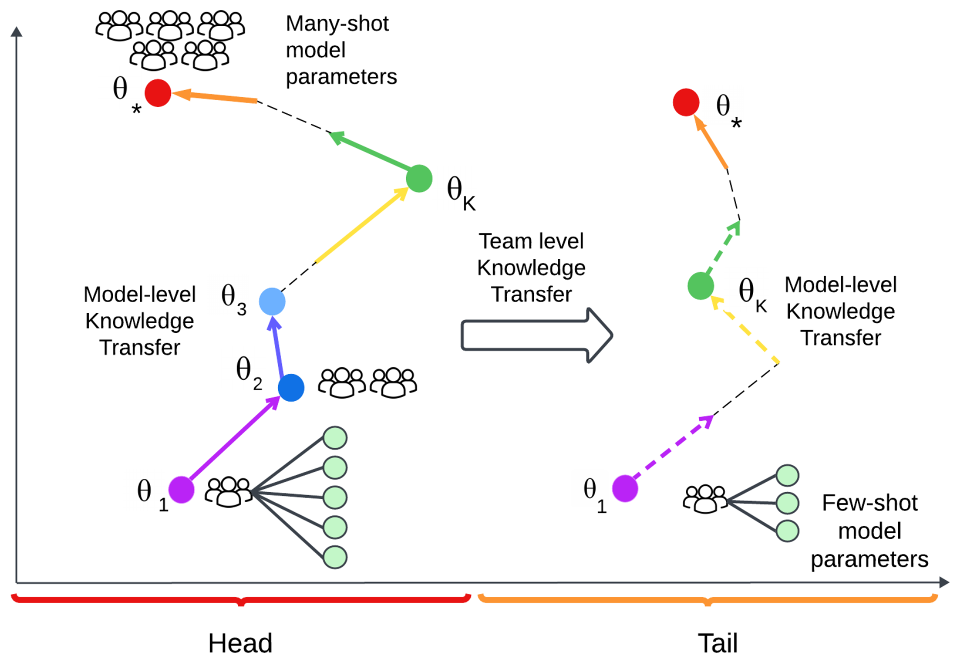

Fig. 1.1.2: It represents the model-level knowledge transfer that incorporates both many-shot and few-shot learning approaches and team-level knowledge transfer that facilitates through curriculum learning.

ranking metrics.

1.1.5 Role of Team Recommender Systems in Modern Team Formation

The process of forming teams in today's complex and dynamic environment necessitates the use of sophisticated tools that can navigate these complexities effectively. Team recommender systems have become indispensable in this context. These systems utilize a variety of strategies and algorithms, including graph-based methods, mathematical optimization, and advanced machine learning techniques, to tackle the Team Formation Problem. They are specifically designed to comb through large datasets to identify potential team members who not only possess the necessary technical skills but are also likely to synergize well with one another, thereby fostering a collaborative environment.

By leveraging data from previous collaborations, these recommender systems aim to minimize the uncertainties inherent in team formation and enhance the probability of project success. This approach not only improves the efficiency of forming teams but also contributes to more effective team dynamics and outcomes. Such systems are particularly valuable in settings where the alignment of skills and interpersonal compatibility are critical for the success of complex projects.

1.2 Thesis Motivation

The importance of team recommender systems in contemporary professional environments cannot be overstated. These systems are crucial for harnessing the full potential of an organization's human capital by efficiently assembling teams that are diverse, well-balanced, and highly effective. As projects across various sectors become increasingly complex and interdisciplinary, the ability to dynamically form teams that blend the right mix of skills, backgrounds, and personality traits becomes essential for driving innovation and achieving strategic objectives.

In today's rapidly evolving workplace, the success of projects often hinges on the ability to quickly adapt to new challenges and integrate diverse expertise. Team recommender systems play a pivotal role in this context by facilitating the swift assembly of teams with complementary skills and compatible working styles. This not only accelerates project initiation and execution but also enhances the quality of outcomes by leveraging a broader range of ideas and approaches.

Moreover, in an era where collaboration is key to organizational success, team recommender systems help break down silos and promote a culture of collaboration and knowledge sharing. By identifying the optimal groupings of individuals for specific projects, these systems ensure that all team members are positioned to contribute effectively, which maximizes productivity and fosters professional growth among employees.

The motivation for this research is driven by the pressing need to address the limitations of current team recommender systems in such dynamic and complex professional environments. As organizations increasingly undertake interdisciplinary projects that require a blend of diverse expertise, backgrounds, and personality traits, the capability to efficiently and fairly assemble effective teams becomes essential for success. However, existing team formation methodologies often face challenges such as data sparsity and popularity bias, where well-known experts are favored over potentially more suitable but less visible colleagues. This not only undermines the fairness of the recommendation process but also restricts the potential for discovering and leveraging hidden talents within an organization.

The motivation for this research is driven by the pressing need to address the limitations of current team recommender systems in such dynamic and complex professional environments. As organizations increasingly undertake interdisciplinary projects that require a blend of diverse expertise, backgrounds, and personality traits, the capability to efficiently and fairly assemble effective teams becomes essential for success. However, existing team formation methodologies often face challenges such as data sparsity and popularity bias, where popular team of experts are favored over potentially more suitable but less visible one. This leads to suboptimal team configurations that may not meet all project requirements effectively. This not only undermines the fairness of the recommendation process but also restricts the potential for discovering and leveraging hidden talents within an organization.

This research aims to overcome these challenges by developing a novel two-level knowledge transfer learning framework that enhances fairness and accuracy by effectively reducing popularity bias. Ultimately, this study seeks to imporve the way teams are formed in professional settings, making the process more fair, and effective in meeting the complex demands of modern projects.

1.3 Problem Statement

Consider a heterogeneous collaboration graph $G = (V, E, T)$, where V represents a set of nodes connected by edges $E \subseteq V \times V$. In $V \times V$ the product is used to describe all possible pairs of nodes within the graph. It means the set of edges E in the graph consists of these pairs of nodes. Each node $v \in V$ is associated with a type T_v. In the context of the DBLP network, these node types include entities such as papers (P), experts (E), venues (L), and skills (S), which are collectively categorized under $T = P, E, L, S$.

The core challenge addressed in this work is the recommendation of teams (s, e), where $e = e_1, e_2, ..., e_p \subseteq E$ represents a set of p experts, and $s = s_1, s_2, ..., s_q \subseteq S$ denotes a subset of q skills required for specific tasks. Our primary objective is to enhance the quality of team recommendations, especially for those teams characterized by having fewer skills that appear in the tail of the distribution. We aim to achieve this improvement while maintaining or even enhancing the overall performance of the system in recommending expert teams tailored to a given set of skills.

This problem is significant because it tackles the dual challenges of optimizing team composition in a way that not only meets specific task requirements but also addresses the inherent sparsity and distributional biases of skills and expertise within large-scale professional networks. By focusing on these less represented skills and the corresponding teams, we strive to bring more balance and fairness to the recommendation process, ensuring that all potential contributors are considered effectively, regardless of their typical visibility in the network.

1.4 Research Objectives and Questions

The primary Objective of this research is to design and implement a pioneering two-level knowledge transfer learning framework. This framework is engineered to facilitate the transfer of knowledge from popular to unpopular experts within professional networks. The objective is to bridge the gap between the abundance of data for some experts and the scarcity for others, enhancing the system's ability to offer balanced and fair recommendations across the board.

A crucial goal of this study is to enhance the quality of team recommendations by addressing key challenges inherent in existing recommender systems—specifically, data sparsity and popularity bias. By integrating meta-learning and curriculum learning techniques, this research seeks to refine the recommendation process, ensuring that it not only identifies the most skilled individuals for a given project but also fairly represents less visible experts in the network.

To establish the efficacy of the proposed framework, it is essential to conduct empirical validations against existing state-of-the-art methods using the DBLP dataset. This research will employ a robust set of evaluation metrics, including Recall, Mean Reciprocal Rank (MRR), Mean Average Precision (MAP), and Normalized Discounted Cumulative Gain (NDCG), to compare the performance of the newly developed two-level knowledge transfer learning framework with that of current leading approaches. The validation will focus on assessing improvements in performance and reduction of bias in team recommendations. We conducted a comprehensive evaluation by comparing various methods for top-K recommendation, assessing both head and tail recommendations, as well as overall performance on the DBLP dataset.

Three pivotal questions guide this research:

- First, how do current recommender systems handle the inherent challenges of data sparsity and popularity bias?
- Second, what role can transfer learning play in mitigating these effects and enhancing the fairness and accuracy of recommendations?
- Third, how do meta-learning and curriculum learning techniques contribute to improving the efficacy of team recommender systems?

1.5 Thesis Contributions

The contributions of this thesis to the field of team recommender systems are significant and multifaceted. This research addresses critical gaps in existing methodologies and introduces innovative approaches to enhance not only the overall performance of team recommender systems but also imporves the reocmmendation quality for experts who are located in tail segment of the data. The key contributions are as follows:

- Development of a Two-Level Knowledge Transfer Learning Framework

  This thesis introduces a novel two-level knowledge transfer learning framework that innovatively increasses the team recommender system performance by leveraging knowledge transfer from head to tail data instances and accordingly addressing the challenges of popularity bias caused by long tail distribution in training dataset in team recommender systems. By facilitating knowledge transfer from well-represented to poorly represented experts, the framework enhances the fairness and inclusivity of the recommendation process. This allows organizations to better utilize their entire talent pool, promoting diversity and innovation.

- Integration of Meta-Learning and Curriculum Learning Techniques
  The research pioneers the integration of meta-learning and curriculum learning techniques in the context of team recommender systems. Meta-learning is employed in model level knowledge transfer in order to capture the implicite change in base learner parameters between many shot and few shot learning while data augmentation is in process, while curriculum learning is used in instance level knowledge transfer in order to facilitate a smooth transition between head and tail data instances.

- Validation Using the DBLP Dataset
  The effectiveness of the proposed framework is rigorously validated using the DBLP dataset, a standard benchmark in the field of team recommender system research. The comparative analysis with state-of-the-art methods—specifically Paragraph Vectors'20, Metapath2Vec'21, and LANT'23—demonstrates significant improvements in recommendation accuracy, reduction of bias, and enhancement of team diversity. This validation provides concrete evidence of the framework's utility and effectiveness.

- Foundational Advances in Team Recommender Systems
  By addressing some of the most pressing challenges in team recommender systems, this thesis lays the groundwork for future research in the field. The methodologies and insights presented here can be further explored and expanded upon, paving the way for more sophisticated and effective team formation tech-

nologies.

1.6 Thesis Structure

The rest of the thesis/research work is organized in the following manner.
In chapter 2, We discuss the related works in the field of Team Formation Problem. The literature review comprises classical rule-based approaches including heuristics, meta-heuristics, etc., and the current state-of-the-art deep learning frameworks on team formation probem.
In chapter 3, We introduce our proposed framework, to solve the team formation problem. This chapter discusses step by step process of our approach and how it optimizes the state-of-art deep learning framework for team formaiton problem.
In chapter 4, We provide our experimental setup and environment, which includes the tools and libraries we used to implement our suggested framework, the system configuration, the dataset information, the hyper-parameters for training, the specifics of the evaluation metrics and the baselines we utilized to assess our model.
In chapter 5, We conducted experiments on the benchmark dataset DBLP. We compared our framework to the existing state-of-the-art methods team formation problem. learning in which our framework outperformed existing architecture in terms all the evaluation metrics in overall, head and tail recommendation performance.
In Chapter 6, We conclude our research, explain the insights we gained during our research work, and describe some of the opportunities for future work.

CHAPTER 2

Related Works

2.1 Introduction

The concept of team formation in professional networks was initially explored by Lappas et al. [6], who introduced a method using a minimum-cost spanning tree tailored to satisfy specific constraints, such as encompassing all required skills for a project. Their groundbreaking approach posited that minimizing communication costs among team members within a social network would enhance collaboration effectiveness. To navigate the complex expert collaboration graph, they developed heuristic methods for identifying pertinent sub-trees that could represent optimal team structures.

Building on this foundation, Lappas and colleagues identified communication costs as key to formulating the Team Formation Problem in social networks. They defined communication costs as the effort required for team members to interact and collaborate effectively. Recognizing the critical role of collaboration and communication between team members, they emphasized that minimizing these costs could lead to more efficient team configurations and, by extension, improve project success. Subsequent researchers like Kargar et al. [7] and Selvarajah et al. [42] further refined the optimization of communication costs. They employed advanced algorithms, encompassing sophisticated heuristics and meta-heuristics, treating the challenge as a single-objective optimization problem. This line of research aimed to streamline team assembly by reducing overheads associated with communication among team members. Example Objective Functions from [7] are indicated below. These objective functions aim to optimize the distances both among team members themselves and between the team members and the leader of the team, as represented by c_{s_i} denoting the skill of team member i and L presenting the leader.

$$\text{sumDistance} = \sum_{i=1}^{p} \sum_{j=i+1}^{p} d(c_{s_i}, c_{s_j}) \tag{1}$$

$$\text{leaderDistance} = \sum_{i=1}^{p} d(c_{s_i}, L) \tag{2}$$

Moreover, later studies by Ashnegar et al. [44], Han et al. [45], Chen et al. [14], and Selvarajah et al. [13] broadened the scope by integrating additional dimensions such as workload balance, personnel expenses, and geographical proximity into the optimization process. These enhancements transformed the problem into a multi-objective optimization dilemma, acknowledging that effective team formation is influenced by a variety of factors beyond simple communication efficiencies.

Despite these advances, most existing methodologies presuppose a static social network structure, relying on stable and predictable interactions among network members. The latest efforts by Kargar et al. [17], which focus on minimizing the sum of the weights of the induced subgraph, address some of these limitations but also reveal new challenges. These methods struggle with the inherent computational complexity of subgraph optimizations, which are similar in nature to a simplified form of the NP-hard Steiner tree problem [6]. Moreover, these strategies often fail to capture the dynamic and intricate relationships among experts, potentially leading to teams that do not fully leverage the diverse capabilities within the network. Table 2.1.1 illustrates the summary of research contributions in the context of various cost metrics.

To overcome the constraints of graph-based methods, recent research has shifted towards leveraging machine learning techniques to assemble teams of experts. Innovative neural machine learning models have been developed that discern relation-

Authors	**Communication Cost**	**Workload**	**Personnel Cost**	**Geo-proximity**
Lappas et al.[6]	✓			
Kargar et al.[7], Li et al.[11], Selvarajah et al.[42]	✓			
Majumdar et al.[43], Anagnostopoulos et al.[12]	✓	✓		
Ashenagar et al.[44]			✓	
Han et al[45], Chen at al[14]				✓
Selvarajah et al.[13]		✓		✓

Table 2.1.1: Summary of research contributions in the context of various cost metrics

ships among experts and their social attributes through advanced neural architectures [20, 21]. These models utilize historical data on successful team compositions as training samples, aiming to predict optimal teams for specific skill requirements. This approach enhances efficiency while preserving effectiveness, thanks to the iterative and online learning capabilities inherent in neural architectures.

One pioneering effort by Sapienza et al. [20] utilized a traditional autoencoder to formulate optimal teams within the context of the DOTA2 game dataset. However, the model faced challenges with uncertainty in training data and was particularly prone to overfitting when correlating specific skills with experts. The root of these challenges was the sparsity of input skills coupled with the long-tailed distribution of expert availability.

Addressing these issues, Rad et al. [21] introduced a Variational Bayesian Neural Network (VBNN). This model adopts a probabilistic approach to network weights, allowing it to represent and manage uncertainty more effectively. By employing variational inference, the VBNN improves upon traditional methods by learning the distribution of skills in relation to the availability of experts, thus mitigating some effects of the long-tail distribution problem. This approach demonstrated marked improvements in team recommendation accuracy over non-variational autoencoder models, as evidenced by various evaluation metrics. However, the research still overlooks other potentially valuable data, such as expert venues, and tends to treat the correlation between skills and experts in isolation.

Rad et al. [22] conducted a study in 2021 that delved into a diverse collaboration network within the DBLP dataset, incorporating different types of nodes such as skills, experts, papers, and venues. They utilized metapath2vec, a technique that employs meta-path-based random walks along with a skip-gram model to explore both the semantic and structural connections between various types of nodes. Initially, these studies leveraged a Variational Bayesian Neural Network (VBNN) to effectively rank experts for specific tasks.

Progressing further in 2022, Rad et al.[23] adopted subgraph representation learning techniques combined with metapath2vec. They focused on subgraphs encompassing both skills and groups of experts who had previously demonstrated successful collaboration. This approach was designed to provide a richer understanding of the relationships and interactions within the data, thereby delivering more nuanced insights into the connections between skills and experts. In this phase of research, the team chose to utilize a maximum similarity index to identify the closest expert subgraph representations to given skill subgraphs, moving away from using VBNN.

In a more recent initiative, Kaw et al.[19] unveiled LANT, a new model that marries transfer learning with neural team recommendations. This innovative approach has enhanced the flexibility of network modifications after training and has streamlined the learning process, especially in terms of reducing time complexity. Yet, despite these technological advancements, the challenge of popularity bias persists.

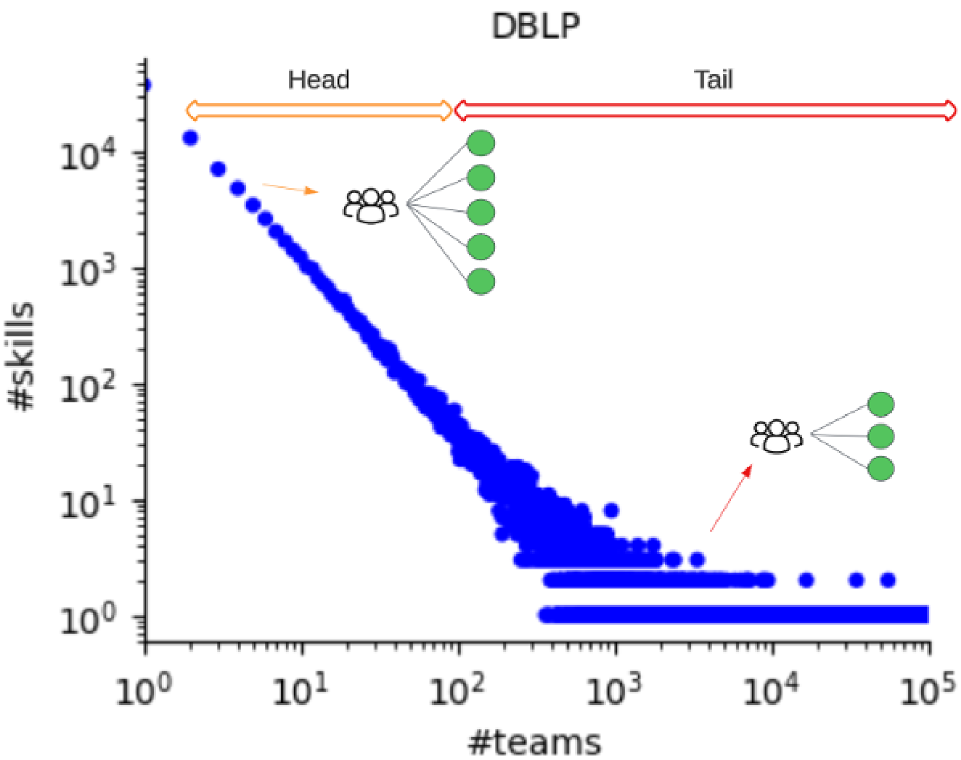

Fig. 2.1.1: Distribution of teams and skills in DBLP.

2.2 Long-tail Distribution Problem

The issue of long-tail distribution prominently affects task performance across various domains, including image classification [28], natural language processing [27], and recommendation systems [29]. This distribution pattern impacts the availability and performance of data in numerous real-world datasets.

Common strategies to tackle this issue include resampling techniques such as over-sampling and under-sampling [30, 33]. Over-sampling, which involves duplicating samples from minority classes, can lead to overfitting, while under-sampling, which reduces examples from the majority class, might eliminate crucial data, potentially degrading the model's performance. An alternative approach is to refine the loss function by incorporating different weights or regularization techniques for various classes or items [34, 28]. For example, recent advancements such as logQ corrections [37, 38] have been developed to introduce item frequency-aware regularization, aiming to balance the representation in datasets.

Moreover, some researchers explore innovative strategies like meta-learning [39] and decoupling the learning process to adjust the classifier separately [40], which are tailored to address the challenges of long-tail distributions more effectively. In the context of team recommender systems, the impact of long-tail distribution is evident from the varying distribution of teams across features such as the number of skills and experts within teams [19]. DBLP dataset displays a long-tail pattern in the distribution of its features. Our analysis is focused on the number of skills for different teams of experts. Figure 2.1.1 illustrates that a small fraction of teams, referred to as the head of the distribution, possess the highest number of skills, while the majority of teams fall into the tail segment, characterized by fewer skills. This variation significantly influences recommendation performance, particularly for teams that are in tail section of the data distribution.

2.3 Meta Learning

Meta-learning, commonly known as 'learning to learn,' has captured significant attention due to its versatility across diverse applications [47]. The primary goal of meta-learning is to develop models that can swiftly adapt to new tasks by acquiring a broad base of applicable knowledge. Typically, meta-learning approaches are categorized into three main types: metric-based, model-based, and optimization-based.

1. Metric-Based Methods: These techniques, such as those described by Snell et al. [48], focus on learning a standard prototype that can differentiate between tasks by understanding their distinctive features.

2. Model-Based Strategies: Model-based strategies, highlighted by Santoro et al. [49], design models capable of rapid learning from minimal data, enhancing their adaptability with few examples.

3. Optimization-Based Approaches: The Model-Agnostic Meta-Learning (MAML) framework, developed by Finn et al. [50], exemplifies optimization-based methods. It involves learning a generalized strategy that can quickly adapt to various tasks through iterative optimization.

Given the success of meta-learning in few-shot learning scenarios, researchers like Wang et al.[39] have begun applying these methodologies to tackle long-tail distribution challenges in fields such as image classification. However, our research identifies a gap in the application of meta-learning to the team recommendation problem.
In this study, we introduce a novel meta-learning approach designed to transfer knowledge from well-represented (head) instances to less-represented (tail) instances within team datasets. Our model comprises two primary components:

- **Base Learner**: This component is a single-tower, two-layer neural network tasked with generating meaningful team embeddings, capturing the essential characteristics of effective teams.
- **Meta-Learner**: The meta-learner operates at a higher level, adjusting the model parameters learned in few-shot scenarios to be applicable in many-shot contexts. It effectively maps or regresses from the parameters of a few-shot model to those of a more extensively trained many-shot model.

Additionally, we have integrated curriculum learning to refine the meta-mapping process further, especially targeting tail teams. This enhancement facilitates learning the intricate connections among teams, significantly improving the model's effectiveness in recommending tail teams in scenarios where they are substantially prevalent.

2.4 Curriculum Learning

Curriculum learning, initially proposed by Bengio et al. [54], has garnered attention for its effectiveness in diverse fields [55]. Inspired by human learning processes, which often benefit from structured progression, curriculum learning adapts this principle to machine learning contexts by organizing training data into a carefully designed sequence.
Recent innovations in curriculum learning have introduced various strategies, such as the teacher-student approach [56] and dynamic curriculum methods [57]. Theoretical

investigations by Weinshall and Cohen [58] have further emphasized the reliability and adaptability of curriculum learning across different conditions.

Of particular interest is curriculum learning's utility in scenarios marked by long-tail distribution patterns. By strategically arranging training examples, we utilize this approach to facilitate efficient knowledge transfer from head to tail instances, thereby enhancing the model's ability to generalize.

CHAPTER 3

Proposed Approach

3.1 Dual knowledge transfer Learning framework

We present a novel framework, drawing inspiration from recent work by Zhang et al. (2021), aimed at transferring knowledge from teams proficient in various skills (referred to as 'head instances') to teams composed of experts with a narrower skill set (termed 'tail instances'). This transfer is intended to enhance the recommendation of teams in scenarios characterized by a long-tail distribution. Our innovative two-level knowledge transfer learning framework capitalizes on the inherent distribution patterns in long-tail data, leveraging both meta-knowledge at the model level and feature connections at the instance level to facilitate effective knowledge transfer from head to tail.

To initiate the knowledge transfer process, our approach involves converting the dataset into a graph structure, which captures both explicit and implicit relationships among data points. Subsequently, we transform the node features of this graph into vectors, preparing them for the subsequent phase of model-level knowledge transfer. Here, these feature vectors are fed into a two-layer, fully connected neural network (referred to as the 'base learner,' denoted as $g(.)$), which generates embeddings for each data point. To address the challenge posed by the long-tail distribution within the dataset, we incorporate meta-learning during the embedding phase, ensuring that the embeddings produced for both head and tail instances are unbiased and accurate. Following this, we employ curriculum learning at the instance level, a technique that

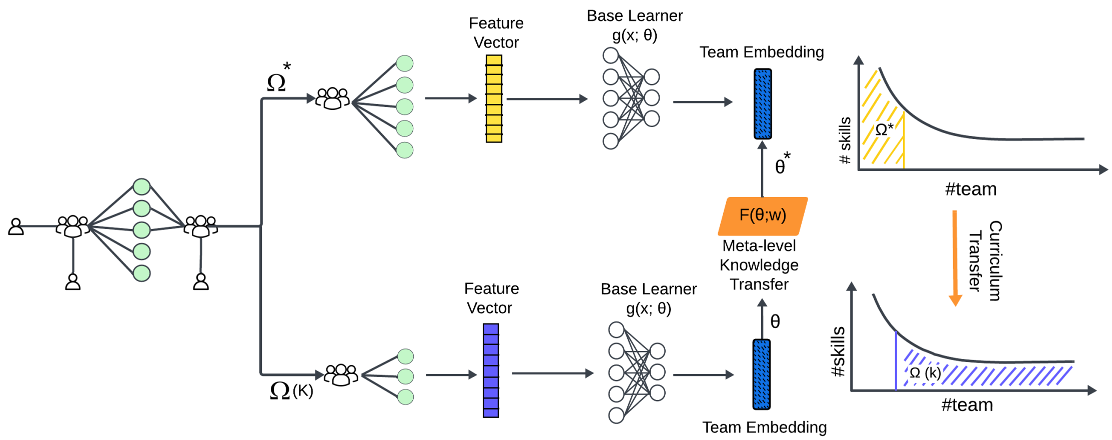

Fig. 3.1.1: Framework to Transfer Knowledge From Popular to Unpopular Experts: collaboration graph tail instances feed through meta-learner $F(\theta, w)$ first to learn the model parameter evolution from few-shot parameters θ to many-shot parameters θ^*, and then the curriculum transfer learn the relations between head and tail teams to transfer knowledge from head to tail.

involves reordering the output embeddings generated by the base learner. This step aims to enhance the effectiveness of model-level knowledge transfer by optimizing the performance of the neural network and refining the accuracy of team recommendations. The schematic depiction of this process is illustrated in Figure 3.1.1. Finally, the curated embeddings are passed through another neural network ($s(.)$), which predicts team performance based on the required skills, facilitating the recommendation of the top n teams.

3.2 Model Level Knowledge transfer

The challenge posed by the long-tail distribution in data significantly impacts the performance of team recommender systems, as highlighted by recent research findings (Kaw et al., 2023). An evident issue stemming from this distribution is popularity bias, where algorithms tend to favor a small subset of well-known or 'popular' experts. This bias often leads to an overrepresentation of recommendations for these experts, while neglecting the majority, particularly teams with fewer skills located at the distribution's tail. Consequently, such bias can severely impact the system's effectiveness, particularly in recommending less visible teams. Neural networks, for example, may tend to overfit to popular experts and struggle to recommend diverse, less common cases. Evaluating the effectiveness of these systems becomes challenging, as traditional metrics may not sufficiently capture their performance on rare items. Moreover, these models can inadvertently perpetuate existing biases by consistently favoring well-known teams over less popular ones.

To address this critical issue, we propose exploring the relationship between models trained on few data samples (few-shot) and those trained on a larger number of examples (many-shot), specifically through the use of a meta-learner. The underlying assumption is that the meta-learner can effectively capture implicit data augmentation. For instance, when considering a team with a diverse range of skills, the meta-learner is designed to infer and incorporate teams that may possess similar skill sets. Instead of directly including similar data points, the meta-learner assimilates the effects on the model parameters, a concept known as meta-level knowledge. This approach involves two key learners: the base learner, $s(.)$ and the meta-learner, denoted as $F(\theta; w)$.

3.2.1 Base Learner

In pursuit of creating insightful team embeddings that encapsulate not only the raw data points but also the intricate relationships encoded within the feature vectors, the base learner, denoted as $g(X_i; \theta)$ processes the feature vector X_i for team i to produce embeddings E_i for each team:

$$E_i = g(X_i; \theta) \tag{1}$$

By leveraging distinct training datasets, $g(X_i; \theta)$ is employed for both few-shot model learning and many-shot model learning. In many-shot learning, the base learner is trained on a dataset (Ω^*) consisting of teams with more than k skills. Concurrently, for few-shot learning, $g(\cdot)$ is trained on a dataset ($\Omega(k)$) focusing on teams with fewer skills (less than k). Consequently, for embedding optimization, the loss function for the base learner can be expressed as:

$$L(g, S; \Omega) = \frac{1}{|\Omega|} \sum_{(x_i, r_i) \in \Omega} (r_i - S(g(X_i; \theta); \phi))^2 \tag{2}$$

where ϕ denotes the parameters of the secondary neural network model $S(\cdot)$ responsible for interpreting embeddings, and r_i represents a reward (in our context, it signifies the number of citations).

Allowing the model to generate representations for learning spanning both the head and tail of the distribution exposes it to the inherent biases present in the long-tail distribution. To counteract this effect, the meta-learner assumes a pivotal role by mitigating the impact of the long-tail distribution through facilitating knowledge transfer between head and tail data instances.

3.2.2 Meta Learner

Once the base-learner $g(\cdot)$has trained both the few-shot and many-shot models, we employ the meta-learner $F(\cdot)$ to map these model parameters, capturing essential meta-level knowledge. This mapping allows the model parameters to evolve as additional training examples are incorporated. The meta-learner $F(\cdot)$ facilitates the transfer of knowledge from many-shot learning models to few-shot ones, ensuring a seamless integration of insights from data-rich (head) contexts into data-sparse (tail) scenarios. The objective function for the meta-learning process is defined as:

$$L(W, \theta | \Omega^{\cdot} \Omega(k)) = \|F(\theta; \omega) - \theta^*\|^2 + \lambda L(g(\theta) | \Omega(k)) \tag{3}$$

In this loss function, $L(W, \theta | \Omega^*, \Omega(k))$, the primary aim of the first term, $\|F(\theta; \omega) - \theta^*\|^2$, is to minimize the discrepancy between the meta-learner's parameter adjustments and the optimal parameter set θ^*. This minimization helps steer the base learner toward parameter configurations that are both effective and generalizable. The second term, $\lambda L(g(\theta)||\Omega(k))$, integrates a regularization factor or additional ob-

Algorithm 3.2.1 Meta Level Knowledge Transfer

Input: Feature matrix X
Feature matrix X derived from team, skill, and author information
Output: Optimized model parameters for team recommendations X

Initialize the neural network model parameters
Initialize base learner parameters θ
Initialize meta-learner parameters ω
Construct dataset Ω^* for many-shot learning
Construct dataset $\Omega(k)$ for few-shot learning
for each training epoch **do**
  for each batch in Ω^* **do**
    Update base learner using many-shot dataset.
    Update θ by minimizing loss on many-shot learning data
  end for
  for each batch in $\Omega(k)$ **do**
    Apply meta-learner transformation: $\theta \leftarrow F(\theta; \omega)$
    Update θ by minimizing loss on few-shot learning data
  end for
  Update ω by minimizing the meta-learning objective
end for
for each evaluation set (head, tail, overall) **do**
  Calculate Recall@K, MRR, MAP, NDCG
end for
return Optimized parameters θ, ω

jective aimed at enhancing performance on the few-shot tasks represented by $\Omega(k)$. The scalar λ serves as a balance between these two aims, ensuring that neither objective dominates the learning process unfairly.

Algorithm 3.2.1, Meta Level Knowledge Transfer, is designed to facilitate the transfer model level knowledge between head and tail data instances. Initially, the feature matrix X, derived from team, skill, and author information, serves as the primary input. The algorithm begins by initializing essential parameters, including the neural network model parameters, base learner parameters (θ), and meta-learner parameters (ω). Following initialization, two distinct datasets are constructed: Ω^*, tailored for many-shot learning tasks, and $\Omega(k)$, aimed at facilitating few-shot learning tasks. Throughout the training process, which unfolds over multiple epochs, the base learner undergoes updates using the many-shot dataset, refining (θ^*) to minimize loss. Subsequently, for each batch within the few-shot dataset, a meta-learner transformation is applied to (θ) using parameters (ω), followed by further optimization to minimize loss on the few-shot data. The algorithm also incorporates mechanisms for updating the meta-learner parameters based on a predefined meta-learning objective, ensuring equitable performance across various tasks. Additionally, the algorithm includes provisions for evaluating model performance using metrics such as Recall@K, MRR, MAP, and NDCG. Furthermore, to enhance the efficacy of training, curriculum learning techniques are integrated, sequentially introducing training examples based on their complexity or relevance.

3.3 Curriculum Learning

Curriculum learning is employed as a strategic approach to sequentially introduce training examples, which are organized based on their complexity or relevance. This technique has been integrated into our proposed framework with the goal of enhancing the system's capacity to mitigate the impacts of the long-tail distribution and boost the efficacy of the incorporated meta-learning techniques. To confront the issue of popularity bias pervasive in our team recommendation system, we have devised a curriculum that categorizes team skill sets by complexity.

In deploying curriculum learning within the team recommendation framework, a distinction is made between teams: those comprising a large number of skills (head of the distribution) are less frequent yet often preferred by the recommender, whereas teams with fewer skills (tail of the distribution) are more common but typically overlooked. We assess the complexity of an embedding inversely proportional to the team's skill count. Consequently, teams with fewer skills are deemed 'simpler,' while those with a broader skill set are considered 'more complex.' This inverse relationship strategically targets bias by initially focusing on the less complex, often neglected teams in the tail section of the data. This curriculum organizes the training data to first prioritize teams with fewer skills, gradually integrating more complex scenarios as the training progresses.

Scoring Function: A scoring function $f(E_i)$ has been formulated to assign scores to each team based on the inverse of their skill count. This scoring is defined as follows:

$$f(E_i) \quad = \quad \frac{1}{\text{number of skills in } E_i + \varepsilon} \tag{4}$$

where ε is a small constant included to prevent division by zero, accommodating instances where teams may lack recorded skills.

Sorting Method: The team embeddings $\{E_1, E_2, \ldots, E_n\}$ are arranged based on the scores computed using f, starting with embeddings representing 'simpler' teams and advancing towards those classified as 'more complex'. This structured approach ensures that the training data is introduced in a manner that respects the inherent complexity of each team's skill set, promoting a more balanced and inclusive recommendation process.

$$\text{Sorted\_Embeddings} \quad = \quad \text{sort}(E_1, E_2, \ldots, E_n, \text{key} = f) \tag{5}$$

3.3.1 Batch Processing and Internal Training

The methodically sorted embeddings were partitioned into practical batches and presented to the mode $s(\cdot)$ during the training phase. This organized approach to training begins with batches that primarily consist of teams characterized by fewer skills. Such a progressive, sequential introduction of data ensures that the model initially concentrates on delivering more accurate recommendations for teams that are often overlooked, thereby mitigating bias effectively. Fig. 3.3.1 indicates the incremental Training process.

This curriculum learning strategy significantly bolsters the model's capability to provide fair and equitable representations and recommendations of teams across a broad spectrum of skill distributions. By starting with simpler team configurations and progressively handling more complex ones, the system methodically reduces the prevalent biases favoring teams with extensive skill sets. This technique is crucially implemented in the later stages of our process, which strategically ensures a harmonious integration

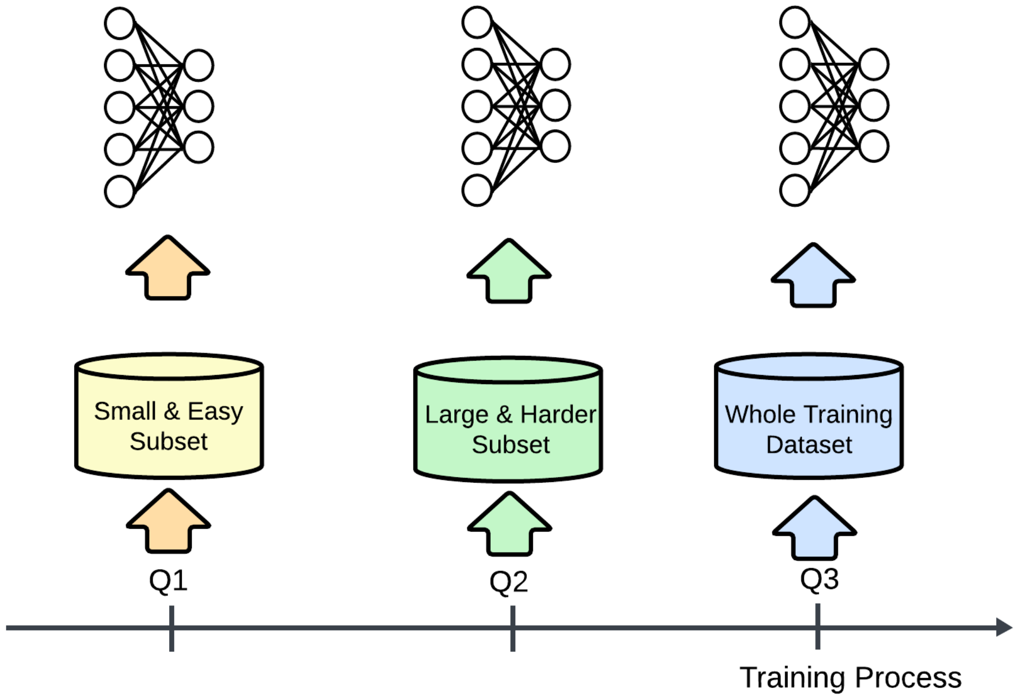

Fig. 3.3.1: Batch Processing and Incremental Training via Curriculum Learning: The structured embeddings were divided into manageable batches and delivered to the concluding stage of the neural network throughout the training phase. This method ensures systematic processing and incremental learning, aligning closely with the principles of curriculum learning for optimal effectiveness.

of both 'head' and 'tail' data embeddings.

Incorporating this approach into the training process profoundly enhances the fairness and accuracy of the team recommendation system. It allows the model to adapt gradually, improving its predictive performance and ensuring that all teams, particularly those in the tail segment that typically receive less attention, are represented fairly. This structured method not only enhances the overall reliability of the recommendations but also supports a more inclusive approach by acknowledging and addressing the disparities in visibility and preference that commonly afflict team formation systems.

CHAPTER 4

Experiments

4.1 Dataset

We employed the DBLP benchmark dataset [1], an exhaustive online repository that compiles bibliographic details of key publications within the computer science domain. This dataset provides a wide array of information on scientific papers, including the names of authors, titles of works, sources of publication (whether journals or conference proceedings), years of publication, and, in some cases, hyperlinks to digital libraries that host the complete texts. Although the DBLP dataset generally does not include the full texts of the documents, it offers a rich collection of citations and references to external resources.

The DBLP dataset is characterized by a long-tail distribution of features, making it particularly relevant for our study, which aims to enhance the precision of team recommendations, especially for groups located within the tail segment of the distribution. Our analysis focused on the distribution of skills among teams of experts who collaborate on academic papers. We observed that a minor fraction of the teams possesses a disproportionately large number of skills, representing the 'head' of the distribution, while the vast majority have fewer skills, categorized under the 'tail' segment of the distribution.

To systematically address the challenge of improving team recommendations in scenarios featuring a long-tail distribution, we structured our dataset into three distinct

[1] https://originalstatic.aminer.cn/misc/dblp.v12.7z

Table 4.1.1: DBLP Dataset Overview

Attribute	Value
# of Papers	4,107,340
# of Experts	2,464,404
# of Venues	2,050
# of Skills	810,734
# of Edges	7,384,478
# of Nodes	28,992,390

sections: training, testing, and evaluation, employing a leave-one-out methodology for each division. This approach ensured that the selection of data for each segment was random yet representative, maintaining the inherent long-tail nature of the dataset across different experimental groups. This methodological rigor supports our objective of developing a more nuanced understanding and capability in recommending teams, particularly enhancing the visibility and consideration of teams that traditionally fall into the overlooked 'tail' portion of data distributions.

4.2 Experimental Setup

4.2.1 Data Reprocessing

Data Preprocessing: In the initial stage of data preprocessing, a graph G is constructed to comprehensively represent entities and their interconnections within the dataset. This graph is populated with nodes corresponding to teams, authors, skills, and venues. Edges are created among these nodes to signify authorship and skill requirements, with edge weights reflecting the relevance of each skill to the team's objectives. Feature vectors for each team are then developed to encapsulate both team attributes, like impact (akin to citation count), and relational attributes, such as the weighted importance of skills and the interconnectedness of authors. To maintain data uniformity, these vectors are padded to match the maximum count of skills and authors recorded, incorporating elements such as citations, average skill weights, skill IDs, and author IDs into an exhaustive vector for each team. The structure of the feature vector for a team i is defined as:

$$X_i = [C_i, W_i] \oplus S_i \oplus A_i \tag{1}$$

Here, $\oplus$ indicates vector concatenation. After establishing the maximum limits for skills and authors to ensure consistent vector lengths, this prepared feature matrix is then applied as input to the base learner. This procedural approach is critical in ensuring that our model effectively utilizes both attribute and relational data, significantly enhancing its potential for generating nuanced recommendations.

These methodological steps are instrumental in fostering a thorough comprehension and utilization of the dataset, integrating both the basic data points and the complex network of relationships into the machine learning framework. This in-depth and meticulously detailed strategy equips our system to conduct advanced analyses, yielding profound insights and predictive capabilities that accurately mirror the intricate dynamics and interrelationships among teams, skills, and authors. Such a methodology significantly improves the production of meaningful team embeddings by the base learner, ensuring that the recommendations and analyses are deeply informed by the layered and multidimensional aspects of the data.

4.2.2 Evaluation Criteria

For the assessment of our model's performance, we utilize four established evaluation metrics: Recall, Mean Reciprocal Rank (MRR), Mean Average Precision (MAP), and Normalized Discounted Cumulative Gain (NDCG) at top K rankings. These metrics are instrumental in quantifying the effectiveness of the model in identifying and ranking relevant experts within a larger pool of candidates.

Recall is a measure that determines how well the model can identify relevant experts from the broader set of potential candidates within the dataset. It calculates the proportion of relevant experts that the model successfully identifies from all the experts in the dataset at the top K rankings. To calculate recall, the model is evaluated on a test set that includes examples of teams with known expert compositions. The model predicts a set of experts for each team, and the predicted set is compared to the true set of experts. **True positives** represent the number of correctly identified relevant experts, while **# of relevant experts represent** the true set of experts.

$$\text{Recall} \quad = \quad \frac{\text{true positives}}{\#\text{ of relevant experts}} \tag{2}$$

Mean Reciprocal Rank (MRR) evaluates the average of the reciprocal ranks of the first relevant expert identified in the model's predictions across all queries. It reflects the average position at which the first relevant expert appears, with a higher MRR indicating that relevant experts tend to appear earlier in the list generated by the model.

$$\text{MRR} \quad = \quad \frac{1}{|n|}\sum_{i=1}^{n}\frac{1}{\text{rank}_i} \tag{3}$$

In formula above n represents the total number of teams in the test set and $rank_i$ is the rank of the first relevant expert in the model's predictions for the i_{th} team and $\sum$ is the sum of all reciprocal ranks.

Mean Average Precision (MAP) is an extension of precision that considers the order of the predictions, giving more weight to relevant results that appear earlier. MAP is particularly useful for evaluating situations where the order of the outputs is significant.

$$P@K \quad = \quad \frac{\#\text{ of relevant experts in top } k \text{ predictions}}{k} \tag{4}$$

Where k is the current position in the ranking.

$$AP_i = \frac{1}{m}\sum_{i=1}^{l}\left(P@k * \mathrm{rel}(k)\right) \tag{5}$$

m is the total number of relevant experts in the team, l is the position where the last relevant expert appears, $P@K$ is the precision at the k_{th} position in the model's predictions, and $rel(k)$ is an indicator function that is 1 if the expert at the k_{th} position is relevant, and 0 otherwise.

$$\mathrm{MAP} = \frac{1}{n}\sum_{i=1}^{n} AP_i \tag{6}$$

n is the number of teams in the test set, AP_i is the average precision for the i_{th} team, and $\sum$ is the sum of all average precisions.

Finally, **Normalized Discounted Cumulative Gain (NDCG)** assesses the quality of the ranking by considering the relevance of the ranked experts, providing a measure of model performance across the ranking positions.

$$\mathrm{DCG} = \sum_{i=1}^{e}\left(\frac{\mathrm{rel}(i)}{\log_2(i+1)}\right) \tag{7}$$

$$\mathrm{NDCG} = \frac{\mathrm{DCG}}{\mathrm{iDCG}} \tag{8}$$

where e represents the total number of experts in recommended teams, $rel(i)$ is the graded relevance of the result at position i, and $iDCG$ is the ideal DCG, representing the best possible DCG value given the relevance scores.

These metrics are applied separately for the 'head' and 'tail' of the recommendation list, highlighting improvements specifically for teams segmented in the tail of the distribution. They also provide insights into the overall performance of the recommender system, gauging how well it improves the general accuracy and fairness of the recommendations.

4.2.3 Baselines

To thoroughly assess the efficacy of our two-level knowledge transfer learning framework, we conducted a comparative analysis with the most recent and high-performing models in team recommendation: Paragraph Vectors'20 [21], Metapath2Vec'22 [22], and LANT'23 [19]. ParagraphVectors'20 [9] leverages Variational Bayesian Neural Network to learn feature representations over a set of teams. Metapath2Vec'21 [11] utilizes a heterogeneous graph representation learning technique called metapath2vec for low-dimensional representation of one-hot encoded skill vectors, which are then used as inputs in the downstream task by a Variational Bayesian Neural Network model for ranking experts. Kaw et al.[19] introduced LANT, a novel model that combines transfer learning with neural team recommendations. This experiment was repeated five times to ensure reliability, and the results presented here are the averages of these trials.

In our experimental setup, we utilized a tower architecture commonly employed in deep learning, where each successive higher layer contains half the number of neurons of its preceding layer. The ReLU activation function was selected for its effectiveness in non-linear transformations. For those approaches incorporating curriculum learning, we conducted 100 training epochs per curriculum stage, with a total of two stages, to maintain consistency across comparisons. The division of data into 'head' and 'tail' segments influenced the configuration parameter k in $\Omega(k)$. We optimized regularization parameters through a grid search, exploring values from $0.3, 0.1, 0.01, 0.001, 0.0001, 0.00001$. Similarly, dropout rates and learning rates were fine-tuned via grid search, testing ranges of $0.1, 0.3, 0.5, 0.7, 0.9$ and $0.1, 0.01, 0.001, 0.0001, 0.00001$, respectively. The batch size for the training process was set at 1024 to balance computational efficiency with memory constraints. These methodological choices were aimed at optimizing model performance and ensuring that the comparison between the proposed framework and the benchmark models was as fair and informative as possible.

4.3 Results and Discussion

Our analysis undertook a detailed assessment of various top-K recommendation methods by systematically examining their performance on both the 'head' and 'tail' segments, as well as evaluating their overall effectiveness using the DBLP dataset. The results of this extensive evaluation are detailed in Table 4.3.1. The aggregate data indicate a remarkable progression, with our model registering an average improvement of around 20% across all metrics when juxtaposed with existing baseline models. This substantial enhancement underscores our model's ability to significantly better represent equity in team recommendations, primarily by adeptly tackling issues associated with the long-tail distribution of data. The baselines have not measured the performance of team recommendation for the head and tail segments of the data. We believe these methods are applicable to both the head and tail. "NA" in Table 4.3.1 indicates "not available." However, we anticipate that after applying these methods to the head and tail segments, our model will demonstrate improved performance compared to the baseline team recommender systems.

This marked improvement was not only observable in the generalized performance metrics but was also distinctly reflected in the precision of recommendations specifically aimed at the tail segment of the dataset. Such targeted advancements have contributed to a substantial mitigation of bias, ensuring a more balanced and fair recommendation system. The success of our model in these areas highlights its robustness and effectiveness in navigating the complexities of diverse data distributions and in rendering more accurate and equitable team recommendations across varied segments of the dataset. This comprehensive evaluation underscores the significant strides our approach has made towards improving the efficacy and fairness of recommendation systems in scholarly and professional settings.

Metrics	Models	Overall Top@3	Overall Top@5	Overall Top@10	Head Top@3	Head Top@5	Head Top@10	Tail Top@3	Tail Top@5	Tail Top@10
Recall(%)	Paragraph Vectors'20	2.20	3.29	5.46	NA	NA	NA	NA	NA	NA
	Metapath 2Vec'21	2.26	3.31	5.64	NA	NA	NA	NA	NA	NA
	LANT'23	2.25	3.31	5.72	NA	NA	NA	NA	NA	NA
	Our Method	**4.16**	**5.26**	**7.92**	**5.26**	**5.32**	**7.32**	**3.08**	**3.23**	**3.25**
MRR(%)	Paragraph Vectors'20	5.90	6.61	7.59	NA	NA	NA	NA	NA	NA
	Metapath 2Vec'21	6.17	6.81	7.70	NA	NA	NA	NA	NA	NA
	LANT'23	6.17	6.68	7.71	NA	NA	NA	NA	NA	NA
	Our Method	**8.32**	**8.68**	**8.90**	**8.24**	**9.03**	**8.95**	**3.55**	**3.23**	**3.02**
MAP(%)	Paragraph Vectors'20	1.64	1.90	2.28	NA	NA	NA	NA	NA	NA
	Metapath 2Vec'21	1.65	1.91	2.28	NA	NA	NA	NA	NA	NA
	LANT'23	1.66	1.91	2.29	NA	NA	NA	NA	NA	NA
	Our Method	**3.24**	**3.54**	**4.48**	**3.10**	**4.24**	**5.68**	**3.44**	**3.92**	**4.70**
NDCG(%)	Paragraph Vectors'20	3.09	3.41	4.41	NA	NA	NA	NA	NA	NA
	Metapath 2Vec'21	3.17	3.42	4.47	NA	NA	NA	NA	NA	NA
	LANT'23	3.20	3.45	4.50	NA	NA	NA	NA	NA	NA
	Our Method	**5.29**	**5.66**	**6.29**	**5.10**	**5.90**	**7.90**	**2.03**	**3.71**	**4.05**

Table 4.3.1: Comparison of evaluation metrics of our model against different models

CHAPTER 5

Conclusion, and Future Work

5.1 Conclusion

In summary, this thesis presented in "Unlocking Team Potential: Leveraging Transfer Knowledge from Popular to Unpopular Experts," tackles the prevalent issue of popularity bias within team recommendation systems, a problem that is often intensified by the long-tail distribution of data. the innovative two-level knowledge transfer learning framework, incorporating both model-level and instance-level knowledge transfers, has significantly improved the quality of recommendations, especially for teams typically underrepresented in data distributions. This method not only reduces bias but also enhances the model's capacity to accurately assess and predict team effectiveness across diverse scenarios. This thesis highlighted the effectiveness of proposed framework in leveraging deep data relationships through the strategic use of meta-learning and curriculum learning. The meta-learner in proposed framework plays a crucial role in facilitating substantial knowledge transfer from well-represented to less represented teams, enhancing equity and representation. Meanwhile, the implementation of curriculum learning ensures that proposed model incrementally addresses increasingly complex scenarios, thereby enhancing its predictive precision and reliability.

5.2 Limitations and Future Work

Looking ahead, this thesis aims to broaden its scope by applying the framework to other datasets and fields beyond the confines of the DBLP dataset, to further test and verify the robustness and adaptability of the methods across various contexts. Additionally, there is an interest in exploring the potential for real-time learning capabilities within the system. This would allow the model to continuously update and adjust to new data as it becomes available, thereby eliminating the need for comprehensive retraining.

To further enhance the model's performance and extend its adaptability, we propose several strategic modifications. Currently, our approach applies curriculum learning to the output embeddings from the base learner. To refine this process, we could apply curriculum learning directly to the constructed feature vectors derived from the graph representation before they are inputted to the base learner. This adjustment would allow for a more nuanced handling of data, potentially reducing the impact of popularity bias and improving the quality of team formation. Additionally, considering the complexity of the DBLP dataset, including long-tail distributions such as the number of team members and citations, the model could be expanded to address these aspects more comprehensively. Implementing more complex configurations and specific types of neural networks, instead of the fully connected two-layer neural network currently used by the base learner, could provide deeper insights and more precise outcomes.

Such advancements would not only augment the efficacy of the existing model but also expand its practical applicability and theoretical relevance across a broader spectrum of industries and organizational settings. The ultimate ambition is to develop a scalable, precise, and bias-conscious framework that is capable of facilitating the formation of optimal teams. These teams would not only be diverse and inclusive but also exceptionally functional, significantly boosting organizational and project outcomes across diverse industries and sectors.

REFERENCES

[1] M. Sozio and A. Gionis, "The community-search problem and how to plan a successful cocktail party," in Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2010, pp. 939–948.

[2] R. B. Bernabé, I. Álvarez Navia, and F. J. García-Peñalvo, "Faat: Freelance as a team," in Proceedings of the 3rd International Conference on Technological Ecosystems for Enhancing Multiculturality, 2015, pp. 687–694.

[3] K. M. Bursic, "Strategies and benefits of the successful use of teams in manufacturing organizations," in IEEE Transactions on Engineering Management, vol. 39, no. 3, 1992, pp. 277–289.

[4] M. Craig and D. McKeown, "How to build effective teams in healthcare," in Nursing Times, vol. 111, no. 14, 2015, pp. 16–18.

[5] K. Selvarajah, P. M. Zadeh, Z. Kobti, Y. Palanichamy, and M. Kargar, "A unified framework for effective team formation in social networks," in Expert Systems with Applications, vol. 177, 2021, pp. 114886.

[6] T. Lappas, K. Liu, and E. Terzi, "Finding a team of experts in social networks," in Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2009, pp. 467–476.

[7] M. Kargar and A. An, "Teamexp: Top-k team formation in social networks," in 2011 IEEE 11th International Conference on Data Mining Workshops, 2011, pp. 1231–1234.

[8] J. Magill-Evans, M. Hodge, and J. Darrah, "Establishing a transdisciplinary research team in academia," in Journal of Allied Health, vol. 31, no. 4, 2002, pp. 222–226.

[9] P. D. Sherer, "Leveraging human assets in law firms: Human capital structures and organizational capabilities," in ILR Review, vol. 48, no. 4, 1995, pp. 671–691.

[10] A. P. Thurow, C. W. Abdalla, J. Younglove-Webb, and B. Gray, "The dynamics of multidisciplinary research teams in academia," in The Review of Higher Education, vol. 22, no. 4, 1999, pp. 425–440.

[11] C.-T. Li, M.-K. Shan, and S.-D. Lin, "On team formation with expertise query in collaborative social networks," in Knowledge and Information Systems, vol. 42, 2015, pp. 441–463.

[12] A. Anagnostopoulos, L. Becchetti, C. Castillo, A. Gionis, and S. Leonardi, "Online team formation in social networks," in Proceedings of the 21st International Conference on World Wide Web, 2012, pp. 839–848.

[13] K. Selvarajah, P. M. Zadeh, Z. Kobti, M. Kargar, M. T. Ishraque, and K. Pfaff, "Team formation in community-based palliative care," in 2018 Innovations in Intelligent Systems and Applications (INISTA), 2018, pp. 1–7.

[14] L. Chen, Y. Ye, A. Zheng, F. Xie, Z. Zheng, and M. R. Lyu, "Incorporating geographical location for team formation in social coding sites," in World Wide Web, vol. 23, 2020, pp. 153–174.

[15] R. H. Rad, E. Bagheri, M. Kargar, D. Srivastava, and J. Szlichta, "Retrieving skill-based teams from collaboration networks," in Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2021, pp. 2015–2019.

[16] M. Sozio and A. Gionis, "The community-search problem and how to plan a successful cocktail party," in Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2010, pp. 939–948.

[17] M. Kargar, L. Golab, D. Srivastava, J. Szlichta, and M. Zihayat, "Effective keyword search over weighted graphs," in IEEE Transactions on Knowledge and Data Engineering, vol. 34, no. 2, 2020, pp. 601–616.

[18] S. Bryson, H. Davoudi, L. Golab, M. Kargar, Y. Lytvyn, P. Mierzejewski, J. Szlichta, and M. Zihayat, "Robust keyword search in large attributed graphs," in Information Retrieval Journal, vol. 23, 2020, pp. 502–524.

[19] S. Kaw, Z. Kobti, and K. Selvarajah, "Transfer learning with graph attention networks for team recommendation," in 2023 International Joint Conference on Neural Networks (IJCNN), 2023, pp. 1–8.

[20] A. Sapienza, P. Goyal, and E. Ferrara, "Deep neural networks for optimal team composition," in Frontiers in Big Data, vol. 2, 2019, pp. 14.

[21] R. H. Rad, H. Fani, M. Kargar, J. Szlichta, and E. Bagheri, "Learning to form skill-based teams of experts," in Proceedings of the 29th ACM International Conference on Information & Knowledge Management, 2020, pp. 2049–2052.

[22] R. H. Rad, E. Bagheri, M. Kargar, D. Srivastava, and J. Szlichta, "Retrieving skill-based teams from collaboration networks," in Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2021, pp. 2015–2019.

[23] R. H. Rad, E. Bagheri, M. Kargar, D. Srivastava, and J. Szlichta, ”Subgraph representation learning for team mining,” in Proceedings of the 14th ACM Web Science Conference 2022, 2022, pp. 148–153.

[24] Y. Dong, N. V. Chawla, and A. Swami, ”metapath2vec: Scalable representation learning for heterogeneous networks,” in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017, pp. 135–144.

[25] F. Schliski, J. Schlötterer, and M. Granitzer, ”Influence of random walk parametrization on graph embeddings,” in Advances in Information Retrieval: 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14–17, 2020, Proceedings, Part II 42, 2020, pp. 58–65.

[26] W. L. Hamilton, R. Ying, and J. Leskovec, ”Representation learning on graphs: Methods and applications,” arXiv preprint arXiv:1709.05584, 2017.

[27] P. Czarnowska, S. Ruder, E. Grave, R. Cotterell, and A. Copestake, ”Don't forget the long tail! a comprehensive analysis of morphological generalization in bilingual lexicon induction,” arXiv preprint arXiv:1909.02855, 2019.

[28] Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, ”Class-balanced loss based on effective number of samples,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 9268–9277.

[29] S. Liu and Y. Zheng, ”Long-tail session-based recommendation,” in Proceedings of the 14th ACM Conference on Recommender

Systems, 2020, pp. 509–514.

[30] J. Brownlee, ”Imbalanced classification with Python: better metrics, balance skewed classes, cost-sensitive learning,” Machine Learning Mastery, 2020.

[31] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, ”SMOTE: synthetic minority over-sampling technique,” in Journal of Artificial Intelligence Research, vol. 16, 2002, pp. 321–357.

[32] H. He, Y. Bai, E. A. Garcia, and S. Li, ”ADASYN: Adaptive synthetic sampling approach for imbalanced learning,” in 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), 2008, pp. 1322–1328.

[33] H. He and E. A. Garcia, ”Learning from imbalanced data,” in IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 9, 2009, pp. 1263–1284.

[34] A. Beutel, Ed H. Chi, Z. Cheng, H. Pham, and J. Anderson, ”Beyond globally optimal: Focused learning for improved recommendations,” in Proceedings of the 26th International Conference on World Wide Web, 2017, pp. 203–212.

[35] Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, "Class-balanced loss based on effective number of samples," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 9268–9277.

[36] Y. Bengio and J. Senécal, "Adaptive importance sampling to accelerate training of a neural probabilistic language model," in IEEE Transactions on Neural Networks, vol. 19, no. 4, 2008, pp. 713–722.

[37] A. K. Menon, S. Jayasumana, A. S. Rawat, H. Jain, A. Veit, and S. Kumar, "Long-tail learning via logit adjustment," arXiv preprint arXiv:2007.07314, 2020.

[38] X. Yi, J. Yang, L. Hong, D. Z. Cheng, L. Heldt, A. Kumthekar, Z. Zhao, L. Wei, and Ed H. Chi, "Sampling-bias-corrected neural modeling for large corpus item recommendations," in Proceedings of the 13th ACM Conference on Recommender Systems, 2019, pp. 269–277.

[39] Y.-X. Wang, D. Ramanan, and M. Hebert, "Learning to model the tail," in Advances in Neural Information Processing Systems, vol. 30, 2017.

[40] B. Kang, S. Xie, M. Rohrbach, Z. Yan, A. Gordo, J. Feng, and Y. Kalantidis, "Decoupling representation and classifier for long-tailed recognition," arXiv preprint arXiv:1910.09217, 2019.

[41] R. M. Karp, "Reducibility among combinatorial problems," Springer, 2010.

[42] K. Selvarajah, A. Bhullar, Z. Kobti, and M. Kargar, "Wscan-tfp: weighted scan clustering algorithm for team formation problem in social network," in The Thirty-First International Flairs Conference, 2018.

[43] A. Majumder, S. Datta, and KVM Naidu, "Capacitated team formation problem on social networks," in Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2012, pp. 1005–1013.

[44] B. Ashenagar, N. F. Eghlidi, A. Afshar, and A. Hamzeh, "Team formation in social networks based on local distance metric," in 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FS

KD), 2015, pp. 946–952.

[45] Y. Han, Y. Wan, L. Chen, G. Xu, and J. Wu, "Exploiting geographical location for team formation in social coding sites," in Advances in Knowledge Discovery and Data Mining: 21st Pacific-Asia Conference, PAKDD 2017, Jeju, South Korea, May 23-26, 2017, Proceedings, Part I 21, 2017, pp. 499–510.

[46] A. Dashti, S. Samet, and H. Fani, "Effective neural team formation via negative samples," in Proceedings of the 31st ACM International Conference on Information & Knowledge Management, 2022, pp. 3908–3912.

[47] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, "Meta-learning in neural networks: A survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 9, 2021, pp. 5149–5169.

[48] J. Snell, K. Swersky, and R. Zemel, "Prototypical networks for few-shot learning," in Advances in Neural Information Processing Systems, vol. 30, 2017.

[49] A. Santoro, S. Bartunov, M. Botvinick, D. Wierstra, and T. Lillicrap, "Meta-learning with memory-augmented neural networks," in International Conference on Machine Learning, 2016, pp. 1842–1850.

[50] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in International Conference on Machine Learning, 2017, pp. 1126–1135.

[51] H. Lee, J. Im, S. Jang, H. Cho, and S. Chung, "Melu: Meta-learned user preference estimator for cold-start recommendation," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 1073–1082.

[52] Y. Lu, Y. Fang, and C. Shi, "Meta-learning on heterogeneous information networks for cold-start recommendation," in Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2020, pp. 1563–1573.

[53] M. Dong, F. Yuan, L. Yao, X. Xu, and L. Zhu, "Mamo: Memory-augmented meta-optimization for cold-start recommendation," in Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2020, pp. 688–697.

[54] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," in Proceedings of the 26th Annual International Conference on Machine Learning, 2009, pp. 41–48.

[55] P. Soviany, R. T. Ionescu, P. Rota, and N. Sebe, "Curriculum learning: A survey," in International Journal of Computer Vision, vol. 130, no. 6, 2022, pp. 1526–1565.

[56] T. Matiisen, A. Oliver, T. Cohen, and J. Schulman, "Teacher–student curriculum learning," in IEEE Transactions on Neural Networks and Learning Systems, vol. 31, no. 9, 2019, pp. 3732–3740.

[57] Y. Wang, W. Gan, J. Yang, W. Wu, and J. Yan, "Dynamic curriculum learning for imbalanced data classification," in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 5017–5026.

[58] D. Weinshall, G. Cohen, and D. Amir, "Curriculum learning by transfer learning: Theory and experiments with deep networks," in International Conference on Machine Learning, 2018, pp. 5238–5246.

[59] Y. Zhang, D.

Z. Cheng, T. Yao, X. Yi, L. Hong, and Ed H. Chi, ”A model of two tales: Dual transfer learning framework for improved long-tail item recommendation,” in Proceedings of the Web Conference 2021, 2021, pp. 2220–2231.

VITA AUCTORIS

NAME:	Zahra Kheiryashkuh
PLACE OF BIRTH:	Alborz, Iran
YEAR OF BIRTH:	1996
EDUCATION:	University of Zanjan, B.Sc. in Software Engineering, Zanjan, Iran, 2022
	University of Windsor, M.Sc in Computer Science, Windsor, Ontario, Canada, 2024