國立臺灣大學管理學院資訊管理學研究所

碩士論文

Department of Information Management

College of Management

National Taiwan University

Master's Thesis

基於公司感知的遞迴式神經網路方法的人才流動預測

CAR-TFP: Company-aware RNN-based Modeling for Talent Flow Prediction

羅苡菱

Yi-Ling Lo

指導教授：魏志平博士

Advisor: Chih-Ping Wei, Ph.D.

中華民國 113 年 7 月

July 2024

doi:10.6342/NTU202403333

致謝

在碩士的兩年間，我得以從對學術研究毫無概念的菜鳥，變成可以研討論文、找出其他研究的優缺點、找出感興趣的問題、設計模型並完成論文。這一路上受到了許多支持，其中魏老師給予我極大的幫助。從問題定義、資料搜集、模型設計到最後的論文內容，老師都很有耐心的給予指導。尤其在使用的方法卡關或是模型的效果不如想像的時候，老師總能夠從我混亂的思緒中提取可行的想法帶我度過難關。有好幾次老師甚至撥空在半夜跟我一起討論把研究的結果修得盡善盡美。同時也感謝王佩琳學姊在一年前爭取到了 LinkedIn 的開放資料，讓我得以完成想研究的題目。

另外感謝三位口試委員楊老師、胡老師在口試時給予諸多論文寫作與研究上的意見，讓我受益良多。感謝實驗室的學長姐給予我研究上許多意見。感謝實驗室的夥伴們，在做研究並撰寫論文時一同在實驗室互相鼓勵、提供意見、一起抒發壓力。謝謝實驗室學弟妹在我們忙著趕論文的過程，協助我們許多事項，也幫我們照顧與維修實驗室的機器。謝謝虹鈞在這兩年中為我們處理相當多的行政事項。最後，非常感謝家人與朋友們忍受我在撰寫論文時焦躁不安的情緒，並給予我支持。

羅苡菱謹啟

于國立臺灣大學資訊管理學研究所

中華民國一一三年七月

doi:10.6342/NTU202403333

摘要

人才流動預測是協助企業人力資源團隊制定有效的人才管理策略的關鍵任務。此問題可以透過線上專業網絡的歷史工作轉換資料來分析與預測企業間未來的人才流動。過去的研究方法使用成對的公司資料結構進行特徵工程和預測。然而，成對的公司資料結構只能提供整體人才流動網絡的聚合特徵，並不能利用從焦點公司到其他公司的分佈。此外，先前的研究也使用股票數據作為額外的數據來源，但股票數據容易被眾多市場訊號影響，並不一定能反應一間公司對人才的吸引力。

為了解決這些限制，我們提出了一個深度學習模型，採用了公司列式的資料結構和公司評價網站的評分資料作為特徵，並使用遞迴式神經網路以擷取人才流動時間序列的特徵。此外，我們模型加入了公司感知結構和嵌入層去有效地捕捉了各焦點公司的人才流動模式。與現有模型相比，不同職位的預測表現提高了 3-4%。

關鍵字:人才管理、人才流動預測、深度學習、遞迴式神經網路

doi:10.6342/NTU202403333

Abstract

This thesis addresses the problem of talent flow prediction by leveraging historical job transition data from Online Professional Networks to forecast future talent movements, a crucial task for human resource teams in developing effective talent management strategies. Previous approaches used a pair-wise structure for feature engineering and prediction. However, the pair-wise structure can only provide aggregated features of the network and does not leverage the distribution from a focal company to other companies. Additionally, previous studies have used stock data as an additional data source, yet stock data can be sensitive to many market signals.

To address these limitations, we proposed a deep learning model employs a company list-wise structure and company rating data as features and feed into a RNN-based model to learn the time series features. After that, our model's company-aware structure and embedding layer effectively capture each focal company’s unique talent flow patterns. It demonstrates a 3-4% improvement in predictive performance over existing models across various positions.

Keywords: Talent Management, Talent Flow Prediction, Deep Learning, RNN-based Time Series Learning

doi:10.6342/NTU202403333

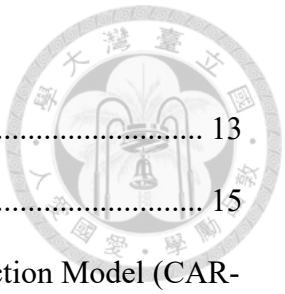
Table of Contents

doi:10.6342/NTU202403333

doi:10.6342/NTU202403333

List of Tables

doi:10.6342/NTU202403333



List of Figures

doi:10.6342/NTU202403333

Chapter 1 Introduction

1-1. Background

Talent stands out as a primary source of competitive advantage for contemporary companies (Hongal & Kinange, 2020). As the global workforce expands, becomes more diverse, and exhibits higher mobility, organizations must invest increased efforts in managing their workforce to acquire and sustain global competitive advantages (Tarique & Schuler, 2010). Consequently, talent management plays a pivotal role for organizations in today's fiercely competitive and dynamic environment.

According to a research article by McKinsey & Co (2021), the "Great Attrition" in 2019, during which more than 19 million U.S. workers resigned from their jobs since April, prompted many companies to adopt a talent-first culture to prevent similar occurrences. Additionally, with regards to company performance, a survey demonstrated that talent management ability, encompassing the attraction and retention of talent, exerts a positive influence on a company's total returns to shareholders (TRS) (McKinsey & Co, 2018). It is asserted that companies with effective talent management are six times more likely than those with ineffective talent management to achieve higher TRS than their competitors (McKinsey & Co, 2018). Given the compelling evidence, it is evident that, to maintain a competitive advantage and surpass competitors in this dynamic environment, talent management emerges as a crucial concern for companies worldwide.

doi:10.6342/NTU202403333

Within the realm of talent management, talent flow stands out as a significant challenge in its implementation (Carr et al., 2005). Originally, talent flow denoted a process wherein valuable workers moved between countries, motivated by factors such as seeking foreign work experience or returning to their home country to capitalize on economic development (Carr et al., 2005). Furthermore, talent flow manifests at the organizational level, signifying employees transitioning from one company to another (Xu et al., 2019), which can also be referred to as external job hops (Oentaryo et al., 2018).

Based on an article of Harvard Business Review, a real-world example of the talent flow challenge for organizations is when high-functioning groups within a company are headhunted by competitors, a phenomenon known as “lift outs” (Groysberg & Abrahams, 2006). One instance of a “lift out” is Conseco Capital Management, which lost its chief equity investment officer and several department members to a competitor, resulting in significant client loss. Another example is the investment bank HSBC, which was left with only a graduate trainee to handle media equities analysis after its entire team of media analysts departed for ABN AMRO. Such abrupt departures can lead to premature internal promotions (Groysberg & Abrahams, 2006).

Numerous studies have demonstrated that the analysis of organizational talent flow holds considerable impact in fields such as human resource planning, global brain drain analysis, and company recruitment. Talent flow reflects the dynamics of the workforce,

doi:10.6342/NTU202403333

job market, and employers (Oentaryo et al., 2018; Xu et al., 2019). Beyond these applications, the examination of talent flow across companies can serve as an indicator of a company's competitiveness, illustrating its allure to prospective job seekers (Zhang et al., 2020).

In summary, amid the escalating significance of talent management for companies seeking to retain, attract, and sustain their competitive advantages, studies of talent flow emerge as a viable approach to meet these objectives.

1-2. Overview of Previous Studies

In the field of talent flow studies, research can be categorized into two groups: talent flow analysis and talent flow prediction. Talent flow analysis aims to investigate the factors influencing talent retention or turnover within an organization or region, as well as to observe talent flow patterns (Qin et al., 2023). Previous studies exploring the factors of talent retention or turnover typically utilized surveys as their research method. These factors include economic, political, and cultural elements that drive high-level talent to relocate within a country or even migrate internationally (Carr et al., 2005; Zhou et al., 2018).

However, these studies exhibit three limitations due to their reliance on survey data (Qin et al., 2023). First, they are costly, as they require substantial time and financial resources to design, distribute, and analyze surveys. Second, survey studies often have a

doi:10.6342/NTU202403333

limited scope, focusing primarily on regional levels, which restricts the generalizability of the findings to broader populations or different geographic areas. Lastly, survey studies are non-retrospective, meaning they cannot be revisited to analyze additional factors or adjust the granularity of data once they are completed. This lack of flexibility hinders the ability to explore new insights or refine the study's focus based on initial findings (Qin et al., 2023). Consequently, a broader understanding of talent flow across regions may be obscured by the localized focus of these studies.

On the other hand, previous research on talent flow patterns has fully utilized Online Professional Networks (OPNs) as data sources. In recent years, OPNs such as LinkedIn have gained widespread popularity. Millions of job seekers globally have updated their digital resumes across 200 countries while actively job hunting, as indicated by the LinkedIn official website (LinkedIn, n.d.). Users willingly share their work experience, education, and accolades publicly on OPNs for job-seeking purposes and to expand professional connections. This extensive job transition data provides a unique opportunity for talent flow quantitative studies, allowing for insights on a larger scale of the talent pool and facilitating more precise analyses concerning location, organization, and time (Xu et al., 2019).

As a result, previous research on talent flow patterns has been able to utilize data mining techniques to extract features from OPN information in order to gain insights into

doi:10.6342/NTU202403333

talent flow patterns, such as regional or organizational job hopping and job-level talent flow network connectivity analysis (Qin et al., 2023). These techniques enable a more comprehensive understanding of how talent moves across different regions and organizations, providing valuable information for both academic research and practical applications in human resource management and recruitment strategies.

In addition to talent flow analysis studies, talent flow prediction primarily focuses on anticipating changes in the labor market by predicting the actual future talent flow amount or percentage at the organizational level (Xu et al., 2019; Zhang et al., 2019). This predictive approach provides valuable guidance for developing effective talent strategies (Qin et al., 2023). Latent variable modeling and time series techniques are employed to enhance the accuracy and flexibility of such predictions.

To date, only two studies have focused on this problem. Zhang et al. (2019) utilized normalized talent flow matrices and matrix factorization modeling techniques to predict future yearly normalized talent flow. In another study, Xu et al. (2019) aimed to use employed pair talent flow features and stock data to predict future monthly company pair-wise incremental talent flow amounts. These studies demonstrate the potential of advanced modeling techniques to forecast talent movements, thereby enabling organizations to better prepare for and respond to changes in the labor market.

1-3. Motivation

doi:10.6342/NTU202403333

Given the limitations of survey studies, including high costs in terms of time and money, limited research scope, and lack of retrospective analysis, we aim to leverage Online Professional Networks (OPN) data in our research. Although talent flow prediction is crucial for developing detailed talent strategies, there are few studies focused on this issue. The two existing studies on talent flow prediction have successfully modeled the problem, yet they present certain limitations.

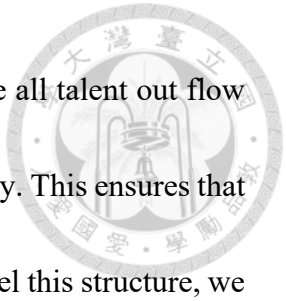
First, the inner product operation in matrix factorization constrains its ability to represent the complex relationships between companies, as the predicted value is based on a linear combination of features (Song & Wang, 2022). Second, pair-wise features only provide aggregated features of the network and do not leverage the distribution from a focal company to other companies. Lastly, stock prices are influenced by numerous market signals, making them sensitive and potentially unreliable for predicting talent flow.

1-4. Research Objective

Our proposed model Company-aware RNN-based Talent Flow Prediction Model (CAR-TFP) incorporates the following features:

- Deep Learning Approach: The deep learning model structure is capable of capturing complex interactions between companies and is extendable for multiple tasks. This approach allows for a more nuanced understanding of the factors influencing talent flow.

doi:10.6342/NTU202403333



- Company-aware Structure: For a focal company, we incorporate all talent out flow amounts as data features and predict its talent flow simultaneously. This ensures that the interactions of each talent flow value are considered. To model this structure, we use a company embedding awareness approach to capture the unique talent flow patterns of each target company.
- Company Ratings: Online company review websites such as Glassdoor and Indeed attract millions of unique job seekers who provide reviews. For instance, Glassdoor boasts over 55 million unique monthly visitors and hosts more than 180 million reviews shared by employees across 20 countries (Glassdoor n.d.). Compared to stock prices, company ratings directly reflect the sentiments of former or current employees about working at the company, making them more relevant to talent flow.

doi:10.6342/NTU202403333

Chapter 2 Literature Review

As stated in the previous chapter, we categorized related studies into two categories: talent flow analysis and talent flow prediction. In this chapter, we aim to review the research that falls within these two topics.

2-1. Talent Flow Analysis

Since talent flow analysis plays an important role in the field of talent management and human resource management, numerous studies have explored this area using various topics and approaches (Carr et al., 2005). Research in this domain can be broadly categorized into two primary areas: factors influencing talent retention and turnover, and talent flow patterns. These studies often focus on different levels, such as country, industry, or company.

2-1-1. Factors Influencing Talent Retention and Turnover

Research on factors influencing talent retention and turnover typically employs qualitative methodologies, using surveys to gather data from targeted groups of talents, such as immigrants or high-skilled professionals in a specific industry. The primary aim of these studies is to identify the factors that either retain valuable talent or drive them away, thereby addressing the issue of "brain drain". These factors include economic, psychological, and career-related aspects (Mao et al., 2009).

doi:10.6342/NTU202403333

For instance, two investigations conducted in New Zealand aimed to understand why high-skilled knowledge workers migrate or leave the targeted region. The studies revealed that lifestyle and family considerations serve as "pull" factors, encouraging knowledge workers to return home, while career and economic issues act as the main "push" factors for those choosing to stay overseas (Carr et al., 2005; Jackson et al., 2005). These factors can be further categorized into global features and local realities, where political and career opportunities fall under global features, while cultural values and family traditions are regarded as local realities (Carr et al., 2005).

Other research focuses on factors influencing the development of specific industries within cities, such as the challenges faced by Wuhan's automotive sector in China (Mao et al., 2009). In this case, talent shortages were identified as a significant challenge, primarily influenced by career and job-related aspects, including income, working environment, and industry cluster characteristics. Additionally, other crucial factors were linked to individual and urban environmental aspects (Mao et al., 2009). Another example is the study on Taipei's fashion industry, which revealed that the creative and cultural economy environment significantly influences talent retention in the city (Hu & Chen, 2014). This finding shows different factors with the results observed in Wuhan's automobile industry, illustrating differences attributable to industry and talent characteristics.

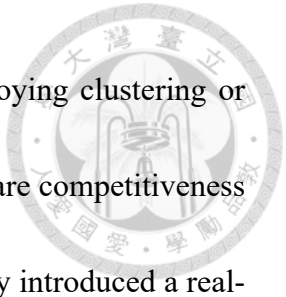
doi:10.6342/NTU202403333

2-1-2. Talent Flow Patterns

Studies focusing on identifying talent flow patterns have fully utilized information from Online Professional Networks (OPNs) by extracting data such as resume position titles, job duration, and working seniority. These studies aggregate information based on target granularity, such as country, city, or industry, allowing researchers to compare differences in culture or economic development levels (State et al., 2014, Oentaryo et al., 2018).

For instance, a study on professional migration in the United States utilized LinkedIn data to identify migration patterns between the United States and other regions (State et al., 2014). The findings indicate a decline in employment-based migrants to the United States, but an increase in students choosing the United States for overseas studies, with Asia emerging as a major destination for professional migration due to job opportunities. Another study compared job-hopping patterns within a region, specifically between Singapore, Hong Kong, and Switzerland (Oentaryo et al., 2018). By extracting attributes based on job titles, such as average working experience years and average job ages, the study revealed characteristics of job hops, distinguishing between promotions and demotions. Notably, workers in Singapore tend to achieve promotions through external job hops (joining other companies), while more Hong Kong employees experience promotions through internal job hops (within the same company).

doi:10.6342/NTU202403333



Other studies in this category view talent flow as a graph, employing clustering or graph learning techniques to identify job-hopping patterns and compare competitiveness between companies within the defined talent flow network. One study introduced a real-time system, JobMiner, designed to highlight the most influential companies and community information within the talent flow network by calculating their closeness and PageRank (Cheng et al., 2013). Another study developed a talent circle detection technique based on job transition networks, aiding human resource teams in identifying talent sources through clustering methods that maximize in-circle edge weights, representing the volume of talent flow (Xu et al., 2016). For graph learning techniques, one study adopted these methods to approximately calculate the Personalized PageRank of each company node in the talent flow graph. By learning the two attraction vectors of each company, the study demonstrated a comparison of competitiveness between each company pair (Zhang et al., 2020).

In summary, talent flow analysis of factors influencing talent retention and turnover predominantly concentrates on the regional level, aiming to identify the factors influencing high-skilled workers' decisions to stay or leave. The outcomes of these studies exhibit variations based on the specific industry and region under investigation, resulting in limited generalizability. Furthermore, these studies often lack large-scale market data, limiting their capacity to offer predictive insights into talent flow dynamics (Zhang et al.,

doi:10.6342/NTU202403333

2019). On the other hand, talent flow pattern studies fully utilize the large scale of OPN data for more quantitative analyses. The insights garnered from these studies provide a more precise understanding of regional and organizational talent flow networks, aiding in strategic planning for organizational human resources.

2-2. Talent Flow Prediction

As mentioned in the chapter 1, talent flow prediction studies focus on predicting the actual future talent flow amount or percentage at the organizational level. They also leverage OPN data to investigate historical labor market changes to predict future trends in the talent flow network. Only two studies have been found in the field of talent flow prediction, by Zhang et al. (2019) and Xu et al. (2019).

Zhang et al. (2019) leveraged normalized talent flow matrices and matrix factorization techniques to predict future yearly normalized talent flow. Figure 1 shows the structure of their approaches. For data preprocessing, they filtered companies that appear more than 1,000 times in their OPN dataset. They also categorized positions into 26 groups to predict company-wise talent flow percentages for each position group. The talent flow data was arranged to form a 3D talent flow adjacency matrix.

From this matrix, they derived two kinds of latent factors: origin company U_i^t and destination company V_j^t at time slice *t*, and a time-independent factor for each position W_k. The inner product combination of these three vectors represents the talent flow value

doi:10.6342/NTU202403333

from company *i* to company *j* for position *k*. The authors also applied an evolving tensor factorization technique to handle dynamic talent flow matrices. At time *t*, the latent vectors of company *i*, U_i^t and V_i^t, evolve through a combination of the previous vectors at time *t*-1 (U_i^{t-1}, V_i^{t-1}) and the vectors of neighboring companies (those with more than one transition) at time *t*-1. For initializing the latent factors, they assumed a zero-mean Gaussian distribution. To avoid overfitting in the matrix factorization method, the authors applied a company similarity regularizer in their loss function of the training model. This regularizer ensures that the components of the latent vectors are similar if their corresponding company attributes are similar. These attributes include the company's industry, scale, location, specialties, type, and age.

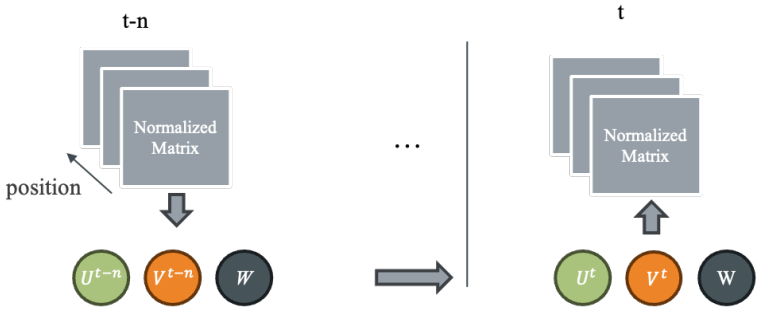


Figure 1 . Approach of Zhang et al. (2019)

On the other hand, the study by Xu et al. (2019) utilized company pair-wise talent flow features and stock data to predict future monthly pair-wise incremental talent flow amounts. The below Figure 2 shows the model structure of the study. For talent flow data preprocessing, they filtered public companies as their target and extracted talent flow data where the start time of the later job must be within ±2 months of the end time of the

doi:10.6342/NTU202403333

former job. To predict each talent flow from company c_i to company c_j in time slice *t*, the model features include historical monthly talent flow and historical monthly stock information. Historical monthly talent flow features include self-loops of c_i and c_j ($f_{ii}^{t-n}, f_{jj}^{t-n}$), in-out flow between the two ($f_{ij}^{t-n}, f_{ji}^{t-n}$), total in-out flow amount of c_i and c_j itself ($f_{i*}^{t-n}, f_{j*}^{t-n}, f_{*i}^{t-n}, f_{*j}^{t-n}$), and total in-out flow of c_i and c_j's industry, where *n* is a list of integers less than *t*. Additionally, historical monthly stock information includes c_i and c_j's stock prices p_i^{t-n}, p_j^{t-n} and trading volumes v_i^{t-n}, v_j^{t-n} of the month's last trading day.

The model structure contains two bi-directional LSTMs, which learn the time series trend of the talent flow features and stock price features separately. The hidden states of the two LSTMs at the same time slice are then concatenated for the prediction module. The attention-based decoder takes the concatenated hidden states to obtain the attention weights and combines the hidden states to form h^τ. After that, the company profiles of c_i and c_j are added along with the previous step's result $s^{\tau-1}$, h^τ, and the previously predicted talent flow $\hat{f}_{ij}^{\tau-1}$ to form the predicted talent flow $\hat{f}_{ij}^{\tau}$.

doi:10.6342/NTU202403333

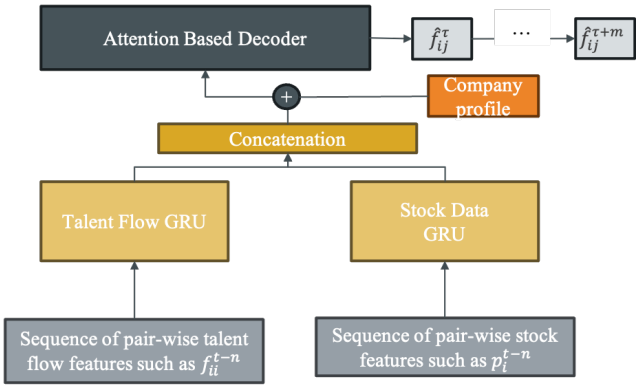


Figure 2. Approach of Xu et al. (2019)

The two studies on talent flow prediction employ different model structures to address the problem. Table 1 provides a general comparison between the two studies.

2-3. Research Gap

This research explores the application of deep neural networks to improve talent flow prediction. Compare with matrix factorization, deep learning models can effectively capture the complex non-linear relationships between talent flow values across companies and time windows but also leverage scalable and extendable features and tasks. This research aims to provide a more detailed understanding of talent dynamics.

Furthermore, previous study of Xu et al. (2019), have treated all job positions uniformly in their predictions. While this method provides a general overview of talent flow, it lacks the granularity necessary for detailed talent management strategies. Our study proposes grouping similar positions and modeling them separately to derive more actionable insights for future talent management.

doi:10.6342/NTU202403333

Additionally, Xu et al. (2019) employed a company pair-wise structure for modeling talent flow. This approach, however, fails to utilize trends from other companies' talent flow data. Our study aims to predict talent flow by leveraging the values of a focal source company in relation to other companies, thereby capturing the interactions and trends of outflow talent values more comprehensively.

Table 1 Comparison of Talent Flow Prediction Studies

	Prediction Target	Data Source	Prediction Method
Xu et al., 2019	Pair-wise monthly incremental talent flow amount of future m months	Historical OPN data Company stock price Company static profile	RNN based modeling with encoder-decoder prediction layer
Zhang et al., 2019	Normalized yearly talent flow rate grouped by position	Historical OPN data Company static profile	Latent factor-based Evolving Tensor Factorization model

Lastly, Xu et al. (2019) addressed the sparsity problem in talent flow prediction using stock price information. However, stock prices are often volatile and influenced by numerous external factors, potentially introducing noise into the model. We suggest incorporating company reviews, which reflect employee satisfaction and provide more stable indicators of a company's attractiveness to potential talent. This adjustment aims to enhance the accuracy of talent flow predictions by considering the direct experiences and ratings of current and former employees.

doi:10.6342/NTU202403333

Chapter 3 Methodology

3-1. Problem Formulation

Based on previous research on talent flow prediction, we can form a talent flow network as follows: At each time *t*, a talent flow network G^t is formed by company nodes C and the talent flow amounts between companies F^t. The target companies are static through time. The talent flow amount from company i (c_i) to company j (c_j) is denoted as f_{ij}^t. We also obtain company rating data as a model feature, where r_i^t is the average company rating of c_i at time *t*, and the difference between r_i^t and r_j^t is denoted as Δr_{ij}^t. Each position *p* has its independent sequence of talent flow networks, allowing us to model each position's network separately.

Given a sequence of talent flow networks from G^{t-n} to G^{t-1}, where $1 < n < t$, each G^{t-n} contains company talent flow amounts F^{t-n} and review ratings R^{t-n}. The goal of our model is to predict the future talent flow amount F^t by modeling the above features.

3-2. Model Structure

Our model consists of five modules: the time series learning module, the dimension reduction module, the learnable company embedding, the company embedding aware layer, and the prediction layer. The following Figure 3 shows a general structure of our model.

doi:10.6342/NTU202403333

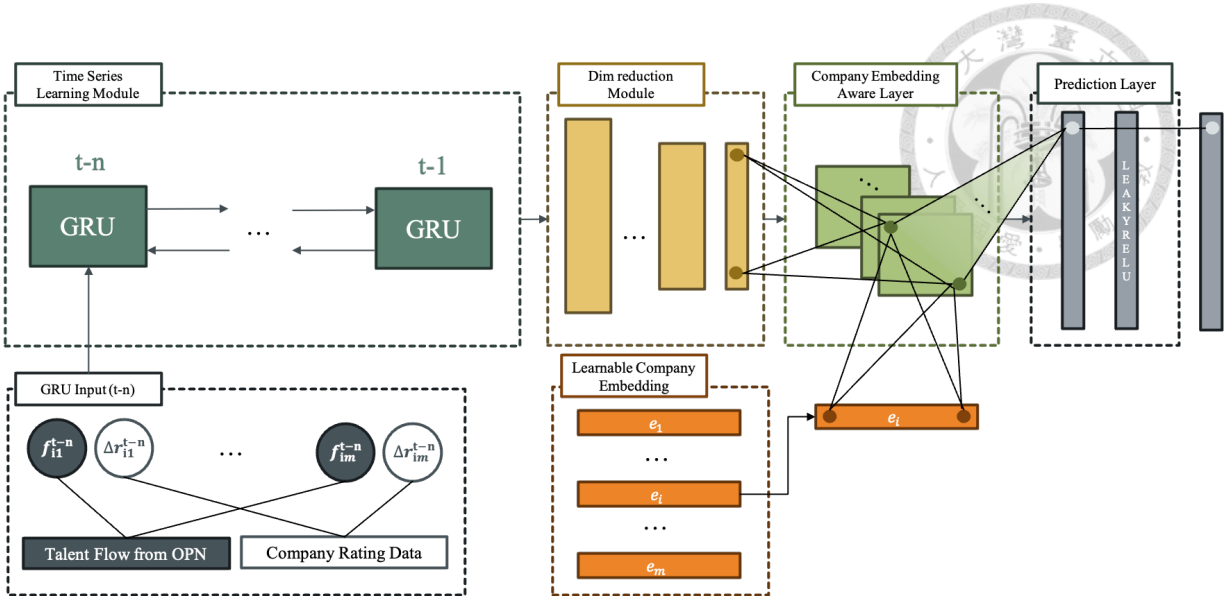


Figure 3. Structure of Company-aware RNN-based Talent Flow Prediction Model (CAR-TFP)

3-3. Input

As mentioned in previous chapters, we consider a company list-wise in-output structure for prediction, which also incorporates rating data. More precisely, for each time t and position p, a source company c_i will have two kinds of features: talent flow amount and rating difference, forming the input vector x_i^t.

- Talent Flow Amount (f_{ij}^t): This represents the talent outflow amount from c_i to c_j, where $j \in [1,2 \dots m]$. m is the number of target companies considered.
- Rating Difference (Δr_{ij}^t): This is the average rating difference of c_i compared to each c_j, where $j \in [1,2 \dots m]$.

Therefore, x_i^t will contain $2m$ features.

doi:10.6342/NTU202403333

3-4. Time Series Learning Layer

To predict the next talent flow value, we treat this as a time series problem. We choose the Gated Recurrent Unit (GRU), which is a Recurrent Neural Network (RNN)-based model, to learn the talent flow trend. The GRU model will process the input vectors x_i^t to capture temporal dependencies and trends in the talent flow data, facilitating accurate predictions for future talent flows.

In our model, the GRU layers are set bi-directional to enhance the prediction ability, since the model learn the talent flow trend from both sides. Formula 3.1 shows we put a sequence of input data with window length of n in the GRU layers. After getting the two final states of bi-directional GRU, we concatenate them together to form h_i.

$$\begin{aligned} \vec{h}_i &= GRU([x_i^{t-n} \dots x_i^{t-1}]), \\ \overleftarrow{h}_i &= GRU([x_i^{t-1} \dots x_i^{t-n}]), \\ h_i &= \begin{bmatrix} \vec{h}_i \\ \overleftarrow{h}_i \end{bmatrix}. \end{aligned} \tag{3.1}$$

3-5. Dimension Reduction Layer

The dimension reduction layer acts as a buffer between the bi-directional GRU module and the company embedding aware layer due to the significant dimension difference between them. We use multiple fully connected linear layers to reduce the high-dimensional output from the GRU module. The mechanism of a dimension reduction layer is shown in Formula 3.3.

$$d_i = W^r h_i + b^r. \tag{3.2}$$

doi:10.6342/NTU202403333

W^r is the learnable weight matrix, while b^r is the bias vector. $W^r \in \mathbb{R}^{n_d * n_g}$, where n_g is dimension of h_i and n_d is the dimension after dimension reduction. This dimension reduction process ensures that the data is in a suitable form for further processing in the company embedding aware layer, maintaining computational efficiency and improving model performance.

3-6. Learnable Company Embedding

In our model, the company embedding aims to capture the distinct talent flow patterns of different source companies after the shared time series learning and dimension reduction layers. While the shared layers can learn general talent flow trends over time, they may not capture the unique patterns of each source company. To address this, we introduce a time-independent learnable embedding vector e_i for each company c_i, where $e_i \in \mathbb{R}^k$ and k is a hyperparameter that defines the dimensionality of the embedding.

3-7. Company Embedding Aware Layer

This module is designed to capture the unique talent flow patterns of each source company by combining e_i with the final state of the shared time series learning and dimension reduction layer, d_i. We use a bilinear layer to perform this combination. A bilinear layer is a 3D matrix that learns the interaction between two input vectors by performing matrix multiplication with the layer's learnable weight. The bilinear layer has

doi:10.6342/NTU202403333

been used in vision problems, such as image classification tasks, where two CNN-based feature extractors need to capture pairwise feature interactions in a translationally invariant manner (Freeman & Tenenbaum, 1997). Although we considered using a linear layer, which would concatenate the two vectors and perform linear transformation, we found that it would only sum the weighted e_i and d_i to produce the prediction value. This approach does not capture the interactions between e_i and d_i. Since our goal is to model the interaction between these vectors to capture different patterns for each source company, we chose the bilinear layer.

The bilinear layer can be formulated as the following formula 3.3. W^b is the learnable weight of bilinear layer, b^b is the bias of bilinear layer, $\hat{y}_i^t$ is the layer's output at time *t*, where $d_i \in \mathbb{R}^{n_d}$, $e_i \in \mathbb{R}^k$, $\hat{y}_i^t \in \mathbb{R}^m$, $W^b \in \mathbb{R}^{m*n_d*k}$. Expanding Formula 3.3 to Formula 3.4 illustrates that the *n*-th matrix of the first dimension multiplies e_i with d_i to form the *n*-th value of the output vector, which *m* is the output layer's dimension.

$$\hat{y}_i^t = d_i^T W^b e_i + b^b. \tag{3.3}$$

$$\hat{f}_{in}^t = \Sigma_{j=1}^{b} \Sigma_{l=1}^{k} d_{ij} W_{njl}^b e_{il} + b_n^b, \tag{3.4}$$

$$n \in [1,2,\dots m].$$

3-8. Prediction Layer

For the prediction layer, we use the Leaky ReLU activation function to transform the results from the Company Embedding Aware Module. This choice helps avoid neurons

doi:10.6342/NTU202403333

being permanently inactive, which can occur with the standard ReLU activation when the output is less than zero. The Leaky ReLU function can be described by the Formula 3.6, where α is a very small number.

$$\begin{cases} \hat{f}_{ij}^{t}. & if\ \hat{f}_{ij}^{t} \geq 0 \\ \alpha\hat{f}_{ij}^{t}. & if\ \hat{f}_{ij}^{t} < 0 \end{cases} \tag{3.6}$$

The output $\hat{y}_i^t$ is a vector representing the talent outflow amounts of the input source company c_i at time *t*. This includes outflows from c_i to itself and to other target companies, so $\hat{y}_i^t$ can be expressed as $[\hat{f}_{i1}^t, \hat{f}_{i2}^t, \ldots\ \hat{f}_{im}^t]$.

3-9. Parameter Learning

The loss function for the model is the Mean Absolute Error (MAE) loss, which measures the difference between the predicted values $\hat{y}_i^t$ and the ground truth y_i^t. Given one company c_i, its loss shown as Formula 3.7.

$$Loss = \sum_{j=1}^{m} \left| f_{ij}^{t} - \hat{f}_{ij}^{t} \right| \tag{3.7}$$

The learnable parameters in the model include the layer weights and biases (W^*, b^*) as well as the company embeddings (e_*). These parameters are optimized using backpropagation through time. Additionally, hyperparameters such as the number of cells in the model, the dimensions of the company embeddings, and the learning rate are tuned through experimental procedures.

doi:10.6342/NTU202403333

Chapter 4 Data

Our data preprocessing consists of two separate processes due to the two data sources: OPN data and company review data. After these preprocessing steps, the two types of features, namely the talent flow amount and the rating score difference, are arranged into the model's input data and then split into training and testing datasets. A general process flow chart Figure 4 is shown below:

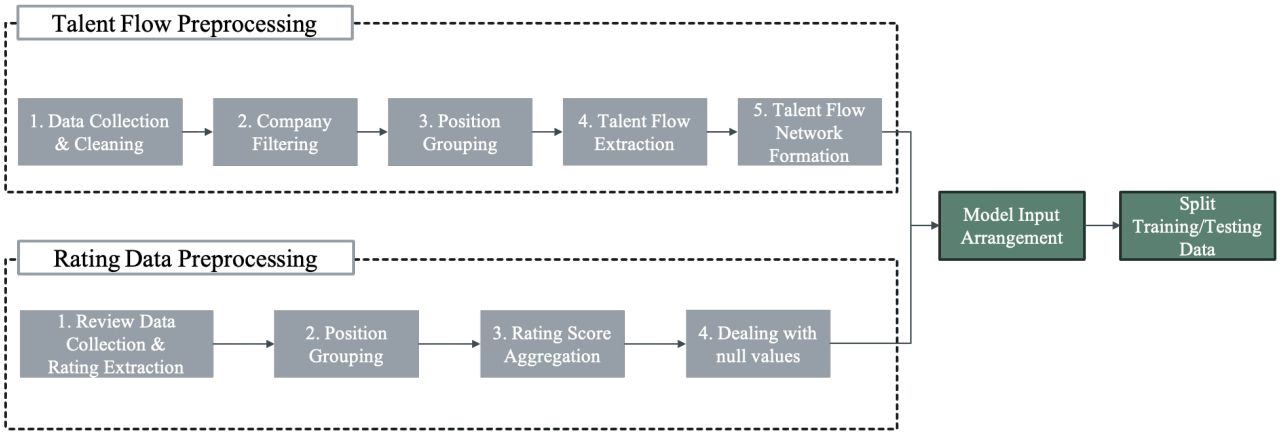


Figure 4 Data Preprocessing Process

4-1. Talent Flow Preprocessing

4-1-1. Data Collection & Cleaning

For OPN data collection, we chose LinkedIn as our target platform because it is one of the largest OPN platforms, with millions of job seekers updating their digital resumes. We obtained a dataset of LinkedIn users from The Bright Initiative, a global organization dedicated to promoting positive change by providing public bodies, non-profit organizations, and academic institutions with public web data. In this study, we are interested in the software industry, specifically profiles with the keyword "software

doi:10.6342/NTU202403333

engineer" in past work experiences. The Bright Initiative provided us with a dataset of 2,032,845 publicly available employee profiles from LinkedIn worldwide.

For the data cleaning process, we followed the procedures outlined by Wang (2023). Initially, we removed work experiences with abnormal job information, such as jobs where the end time was earlier than the start time. We also excluded records with important missing values, including start time, end time, company, and position title. Additionally, we filtered out jobs that did not meet our requirements, such as internships, part-time jobs, military jobs, academic experience, and volunteer experience.

4-1-2. Company Filtering

Next, we counted the appearance frequency of companies per resume, ensuring each company was counted only once per resume. We selected companies with an appearance frequency of no less than 100 times, resulting in 1,563 companies. Other companies were coded to the "out of bag company" (OOBC) category. After this transfer, we removed resumes containing only OOBC experiences, leaving us with 387,464 resumes.

4-1-3. Position Grouping

For position title cleaning, we removed stop words, punctuation, and converted all titles to lowercase. We then used a predefined dictionary to extract keywords from position names, allowing us to merge positions with the same keywords (e.g., SSE to senior software engineer). The predefined dictionary contained domain-specific Named Entity

doi:10.6342/NTU202403333

tags identifying responsibility or functional words (e.g., engineer, manager, system, finance) (Liu et al., 2020). For each merged position, the most frequent original position was selected as the standardized position name, resulting in 106,829 standardized positions. We selected positions appearing no less than 50 times, resulting in 2,227 positions.

To group positions with similar functions, we applied grouping rules and manual checks. The detailed grouping rules are shown in Table 2.

Table 2. Position Grouping Rules and Results

Group	Rules (position title w/ or w/o)	Count	Example
Software Developer Professional	w/ software, programmer, developer, engineer	1,116 (50%)	C# Developer, Principal Firmware Engineer, Mobile Software Engineer
	w/o consult		
Management Professional	w/ project, product	234 (10.5%)	Technical Project Lead, IT Project Coordinator, Project Analyst
	w/o engineer		
Consultant	w/ consult, account	231 (10.3%)	Sales Consultant, Support Consultant, Associate Business Consultant
	w/o software		
Data Professional	w/ data, etl, machine learning	126 (5.6%)	Technical Business Analyst, Principal Data Engineer, Data Scientist
	w/o database, dba		
Cloud & Architect	w/ cloud, architect, application	99 (4.4%)	Application Architect, Java Architect, Enterprise Cloud Architect
	w/o consult		
Infrastructure Professional	w/ network, infra, linux, os	76 (3.4%)	Network Consulting Engineer, Infrastructure Specialist
	w/o software		
UI/UX frontend Professional	w/ web, front, user, ui, ux	49 (2.2%)	Full Stack Web Developer, Senior UI Engineer, UI/UX Designer
	w/o consult		
Database Professional	w/ database, dba, tableau, sql, hadoop, sap, oracle	37 (1.6%)	Database Developer, SAP Developer, Oracle PL/SQL Developer
	w/o consult		

doi:10.6342/NTU202403333

The grouping was based on required keywords and the absence of other groups' required keywords. After rule-based grouping, we manually checked all position titles to avoid special cases and ensure they fit within the group requirements. This process resulted in 8 groups containing 1,968 positions, while positions not considered were transferred to the "out of bag position" (OOBP) category. After the position cleaning and filtering process, we removed resumes containing only OOBP experiences, resulting in a final dataset of 195,969 resumes.

4-1-4. Talent Flow Extraction

The remaining resumes are processed by sorting the jobs in each resume by their start date. If jobs have the same start date, the job with the later end date is placed first, as it tends to be more primary in an individual's experience. This step ensures the correct order when extracting talent flow. We then loop through each resume's experiences to extract talent flows, categorizing them into four types: short blank period, long blank period, partial overlap, and full overlap. These four types of talent flow can be shown in Figure 5, where prefix F means former job, prefix L means later job, F and L can be same in company.

For the short blank period, the later job’s start date is later than the former job’s end date but no more than a year after. We process this type of job transfer at the time of the former job’s end date. For the long blank period, the later job’s start date is more than a

doi:10.6342/NTU202403333

year after the former job's end date. This kind of job transfer has too long a blank period for a typical job transition, so we insert an "OOBC" experience between the two jobs.

In partial overlap situations, the later job's start date is earlier than the former job's end date, but the later job's end date is still later than the former job's end date. This situation suggests that some employees start a new job while continuing their former job. We process the talent flow to occur at the time the former job ends. Lastly, for contained overlap, the later job's start date is earlier than the former job's end date, and the later job's end date is not later than the former job's end date. We count only the former jobs that have longer durations, are ordered first in the resume, or are not OOBC.

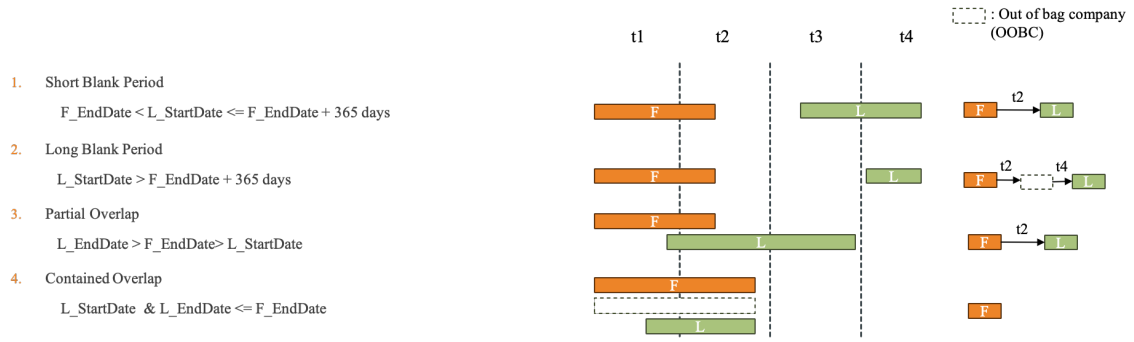


Figure 5. Talent Flow Extraction Types and Preprocessing Rules

4-1-5. Talent Flow Network Formation

To focus on talent flow between a limited number of companies, we select the top 100 companies that appear most frequently by resumes and transfer companies not in the list of target companies to OOBC. We also combine consecutive OOBC jobs into one job for easier processing.

doi:10.6342/NTU202403333

The final step of our talent flow preprocessing is to format our talent flow matrices. We set our time window size to 6 months, covering the period from January 2008 to December 2022, because our company review dataset starts from 2008. This results in 30 time windows in our dataset. We then separate valid talent flows into each time window and position, aggregating these flows within the same group to form an adjacency matrix. After aggregation, we obtain sequences of adjacency matrices for 8 positions. Each sequence has a length of 30, and each matrix is in the shape of 101 * 101 (100 target companies + OOBC).

4-2. Talent Flow Exploration

The following Figure 6 illustrates the aggregated talent flow amount grouped by position over time. Additionally, Table 3 presents the statistical information of the target companies' outflow amounts by position groups. The figure shows that, in general, all positions have increased consistently over time, reflecting the growing number of users on OPNs. Specifically, the "Software Developer Professional" group has the largest talent flow amount, followed by "Consultant", "Management Professional", and "Data Professional". This trend is likely due to our focus on resumes containing software engineer experiences. The steady increase in talent flow across all positions indicates the expanding user base and the dynamic nature of job transitions within the software industry. The detailed statistics in Table 3 further highlight the prominence of software-related

doi:10.6342/NTU202403333

roles in the dataset, underscoring the relevance of our focus on software engineer experiences in understanding talent flow patterns.

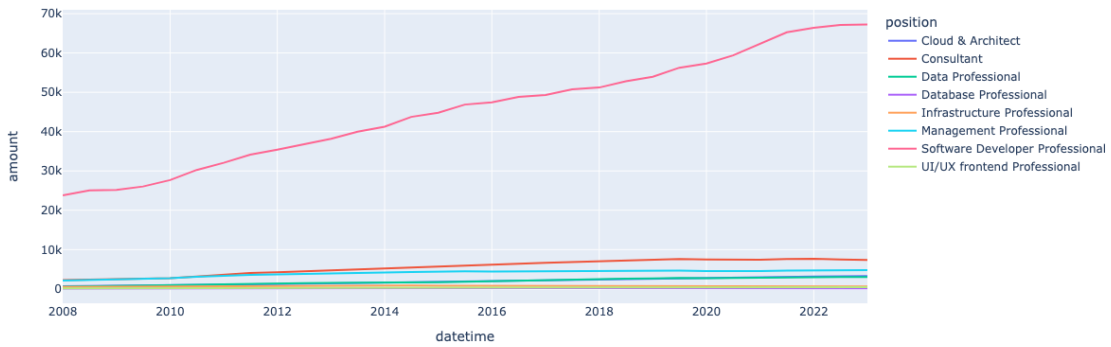


Figure 6. Talent flow size (sum of cells in each adjacency matrix)

Table 3. Statistics of sum of outflows across time windows by companies and positions

Position Group	mean	std	min	max
Software Developer Professional	14,069.61	19,746.19	1,197	134,427
Management Professional	1,242.30	2,578.34	20	20,025
Consultant	1,715.62	4,341.53	3	31,246
Data Professional	594.53	2,081.14	3	20,563
Cloud & Architect	593.00	1,128.17	3	7,028
Infrastructure Professional	221.29	455.10	0	2,959
UI/UX frontend Professional	91.14	209.74	0	1,886
Database Professional	48.09	86.21	0	554

We also present the average talent outflow rate for the company along with the outflow rates to the top 1 and top 3 competitors. For instance, Table 4 displays the top 10 companies in terms of talent flow size within the "Software Developer Professional" category and their respective average talent outflow rates. From the data, it is evident that Google and Microsoft have relatively low outflow rates, indicating that most employees

doi:10.6342/NTU202403333

tend to remain with these companies. In contrast, Wipro and Tech Mahindra exhibit higher outflow rates, approximately 6-7% greater than those of the companies with the lowest outflow rates mentioned previously. Moreover, Tech Mahindra's primary competitor for talent accounts for 1% of its outflow.

Table 4. Talent outflow rates of the top 10 "Software Developer Professional" companies

Company	Average Outflow Rate	Average Outflow Rate to Top1	Average Outflow Rate to Top3
Accenture	9.36%	0.43%	1.00%
Infosys	11.24%	0.63%	1.38%
IBM	8.92%	0.39%	0.86%
Microsoft	5.40%	0.38%	0.82%
Tata-consultancy-services	11.54%	0.60%	1.30%
Wipro	12.50%	0.80%	1.97%
Tech-mahindra	12.75%	1.02%	2.38%
Hcltech	11.81%	0.81%	1.80%
Capgemini	8.57%	0.53%	1.17%
Google	4.75%	0.43%	0.78%

We also found the talent flow matrices to be sparse. Table 5 shows the sparsity of company talent flow matrices by position. We observed that the average sparsity of talent flow matrices for all positions exceeds 90%. The "Software Developer Professional" group is less sparse, followed by "Consultant", "Management Professional", and "Data Professional". Due to the high sparsity of other positions, we will focus on these four positions for our experiments. Figure 7 illustrates the sparsity of the top 30 companies by appearance frequency. The trend for each company is similar, with the "Software Developer Professional" group being less sparse. Additionally, some companies exhibit

doi:10.6342/NTU202403333

less sparsity in specific professions. For instance, the French consulting firm Capgemini has a lower sparsity in the "Consultant" category than in the "Software Developer Professional" category.

This analysis highlights the variability in sparsity across different positions and companies, guiding our focus towards the most relevant and densely populated categories for further experiments.

Table 5. Sparsity of company talent flow matrix by position

Position Group	Sparsity
Software Developer Professional	90.52%
Consultant	97.11%
Management Professional	97.73%
Data Professional	98.21%
Cloud & Architect	98.41%
Infrastructure Professional	98.84%
UI/UX frontend Professional	99.19%
Database Professional	99.39%

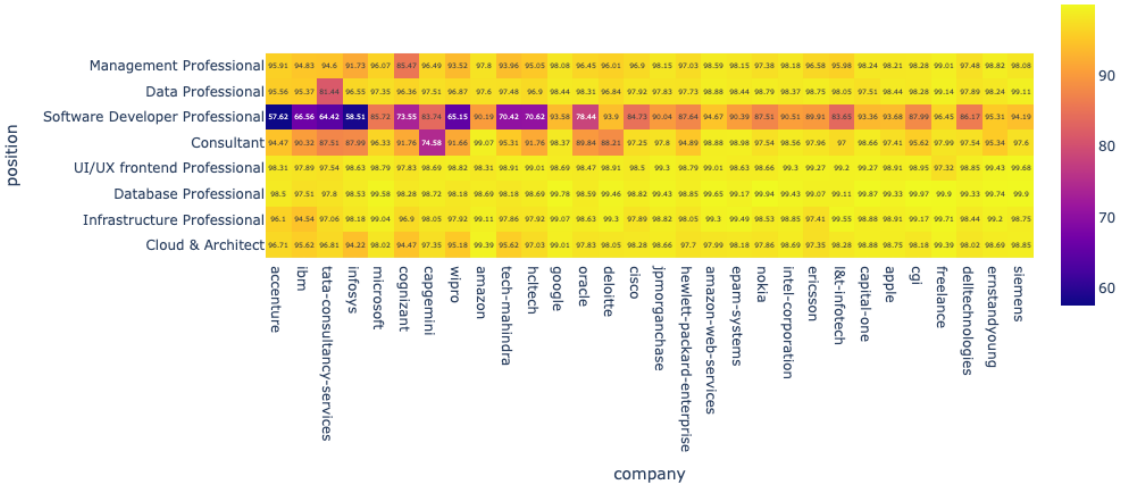


Figure 7. Sparsity of top 30 companies

doi:10.6342/NTU202403333

4-3. Rating Data Preprocessing

We sourced data from Glassdoor, the largest company review platform, founded in 2008. Glassdoor hosts more than 180 million reviews, salary details, and insights shared by employees across 20 countries. Our data crawler retrieved reviews for all 100 target companies, extracting rating scores, employee positions, and review dates from these reviews. Figure 8 shows an example of a company review on Glassdoor, with rating scores ranging from 1 to 5.

For the position cleaning and grouping process, we followed the same procedure used for talent flow positions to align the two data sources. Consequently, ratings associated with positions not considered (OOBP) were removed. All review dates were then classified into each time window to align with the talent flow network. After grouping ratings by position groups and time windows, we aggregated the rating scores based on these components and their respective companies. We calculated the accumulated mean score for each key group, resulting in a rating score r_i^t for each position p of company c_i at time t.

doi:10.6342/NTU202403333

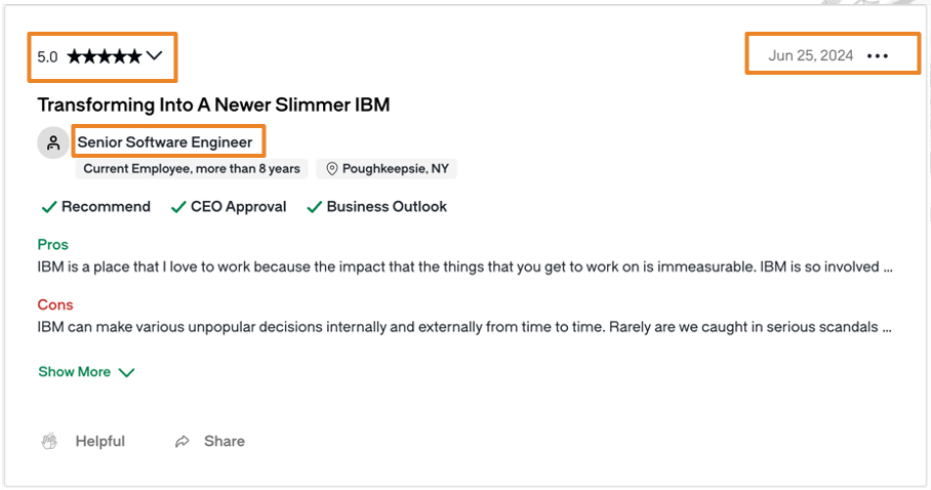


Figure 8. A Company Review Example from Glassdoor

We observed that some groups had missing values due to a lack of data in certain time windows and position groups. For instance, Amazon Web Services was not separated from its parent company Amazon on Glassdoor until 2021, resulting in null values in earlier time windows. We filled these blanks with the previous time window's value. For the first time window with a null value, we assigned a score of the first value that appears in this company dataset. For OOBC's rating scores in each position group and time window, we averaged the ratings of all other companies in the group.

4-4. Rating Data Exploration

Table 6 shows the statistical information on the total review counts for each company, while Table 7 presents the average rating information. As Table 6 indicates, the top four groups have the most rating data counts, similar to our talent flow data. A small difference between the trends in talent flow and rating data is seen in the "Database Professional" category, where we found a higher percentage of review data on the company review

doi:10.6342/NTU202403333

website compared to the talent flow amount. This discrepancy may be due to the focus on software engineers when requesting OPN data. Table 7 provides a general distribution of company ratings, with average ratings by position ranging from 3.52 to 3.75.

Table 6. Statistic Information of Company's Total Review Count by Position

Position Group	mean	std	min	max
Software Developer Professional	3,361.74	7,146.84	38	47,167
Management Professional	586.88	1,035.72	5	8,108
Consultant	851.97	1,706.83	9	9,002
Data Professional	492.76	1,215.38	2	9,565
Cloud & Architect	134.78	234.58	1	1,438
Infrastructure Professional	178.95	350.68	1	2,660
UI/UX frontend Professional	83.46	207.66	1	1,438
Database Professional	362.78	1,132.84	3	5,910

Table 7. Statistic Information of Company's Average Rating by Position

Position Group	mean	std	min	max
Software Developer Professional	3.52	0.38	2.75	4.88
Management Professional	3.58	0.45	2.46	4.91
Consultant	3.53	0.54	1.58	4.91
Data Professional	3.56	0.51	2.42	5.0
Cloud & Architect	3.54	0.64	1.48	5.0
Infrastructure Professional	3.54	0.58	2.14	5.0
UI/UX frontend Professional	3.75	0.80	1.0	5.0
Database Professional	3.60	0.73	1.87	5.0

doi:10.6342/NTU202403333

Chapter 5. Experiment

5-1. Training and Testing Data

As mentioned in the chapter 4, the two types of preprocessed data is used to form the model input data. For each time *t* and position *p*, a source company c_i has two kinds of features. One is the talent flow, denoted as f_{ij}^t, where $j \in [1,2 \dots 101]$. These values are obtained from the talent flow adjacency matrix. The other feature is the rating score difference, denoted as Δr_{ij}^t, where $j \in [1,2 \dots 101]$. These values are calculated by determining the rating difference between c_i and c_j at time *t*. Therefore, each data instance has 202 features (101 talent flow values+ 101 rating score differences). We then remove data where the source company belongs to OOBC. For each position *p*, the dataset contains 3,000 instances of talent flow and rating information. For splitting the training and testing data, the dataset is split into approximately 75% training data and 25% testing data based on the time window. Therefore, the predicted target time of the testing data will be later than January 2020. This structured approach avoids that the training model data from information leaking. Additionally, we think that it is not reasonable to apply temporal validation since the dataset has only 30 time windows, it will lack of training data for our model in the first or second validation rounds.

5-2. Evaluation matrices

We set three kinds of evaluation metrics to assess our model's performance:

doi:10.6342/NTU202403333

- Mean Absolute Error (MAE): MAE provides a measure of the average magnitude of the prediction errors without considering their direction. It is calculated as formula 5.1, where i means for each source company c_i, and m means number of target company.

$$MAE = \frac{1}{m}\sum_{j=1}^{m}\left|f_{ij}^{t} - \hat{f}_{ij}^{t}\right| \tag{5.1}$$

- Mean Absolute Percentage Error (MAPE): MAPE provides a measure of prediction accuracy in percentage terms. It is calculated as formula 5.2, where ε is used to avoid division by zero when $f_{ij}^{t} = 0$ (set to 2 in our case).

$$MAPE = \frac{100\%}{m}\sum_{j=1}^{m}\frac{\left|f_{ij}^{t} - \hat{f}_{ij}^{t}\right|}{f_{ij}^{t} + \varepsilon} \tag{5.2}$$

- Precision @ k (where k = [2, 3, 5]): Precision @ k measures the accuracy of a model's predictions by considering only the top k predicted values. It is calculated as formula 5.3.

$$precision@k = \frac{Number\ of\ true\ positive\ predictions\ in\ top\ k}{k} \tag{5.3}$$

Note that for precision @ k we do not consider the value of talent flow to OOBC since it has less managerial significance. Therefore, we sort the other 100 talent flow values to get the top k companies. If the count of ground truth vector's zero values is more than k, we only count the companies of the top non-zero values as our ground truth.

doi:10.6342/NTU202403333

The purpose of choosing these three metrics is to observe the model's numeric accuracy using MAE and MAPE, indicating how close the regression model's results are to the ground truth amounts. Precision @ k shows the general order accuracy of the benchmark models, indicating whether the model can identify the trend of the source company's talent flow target. Both MAE and MAPE are separately averaged and calculated for three kinds of prediction targets: self-loopf_{ii}^t, outflow amount f_{ij}^t, and outflow to OOBC's talent flow amount f_{i101}^t.

5-3. Hyperparameter Settings

To ensure optimal performance, we tune several hyperparameters, including the number of cells in the model, the dimensions of the company embeddings, and the learning rate. These hyperparameters are adjusted through experimental procedures to achieve the best possible model performance. We apply Adam as the model training's optimizer, and the learning rate is set to 0.00005. Other hyperparameter in our model such as the hidden dimension of the bi-directional GRU is set to be 2048, so the output dimension of our time series learning module will be 4096. The dimension reduction layer number is set to 2 and the corresponding output dimension is set to 1024 and 512. The learnable company embedding's vector dimension is set to 200. All the model layer

doi:10.6342/NTU202403333

dimension and the learning rate is tuned by grid search. Lastly, for the GRU input sequence length we set to 4.

5-4. Benchmarks

For our experiment's benchmark models, we chose the following four approaches to compare with our model:

- Auto Regression (AR)

An AR model predicts future values based on a linear combination of past values. To align with our model, we set the AR model time window length to 4.

- Vector Auto Regression (VAR)

Unlike univariate autoregressive models, which predict a single time series based on its own past values, VAR models predict multiple time series simultaneously, considering the interrelationships between them. For data preprocessing for the VAR model, we flatten our talent flow adjacency matrix to fit the model. Similar to the AR model, we also set the VAR model time window length to 4.

- Pair-wise RNN-based Model (Pair-RNN)

Inspired by Xu et al. (2019), we designed a model structure and features using a pair-wise approach for prediction. Figure 9 shows the model structure. To predict every talent flow amount from c_i to c_j (f_{ij}^t), the model includes pair-wise talent flow and rating features from time window *t*-*n* to *t*-1. The talent flow features of the model include self-

doi:10.6342/NTU202403333

loops of c_i and c_j $(f_{ii}^{t-n}, f_{jj}^{t-n})$, in-out flow between the two $(f_{ij}^{t-n}, f_{ji}^{t-n})$, and total in-out flow amount of c_i and c_j $(f_{i*}^{t-n}, f_{j*}^{t-n}, f_{*i}^{t-n}, f_{*j}^{t-n})$. For company rating features, we include self-rating (r_i^{t-n}, r_j^{t-n}), rating difference for each pair (Δr_{ij}^{t-n}), review counts, and average self-rating for all positions. These 15 features are then fed into a Bi-directional GRU model and a fully connected layer to predict the future talent flow value f_{ij}^t. The model time window length is also set to 4.

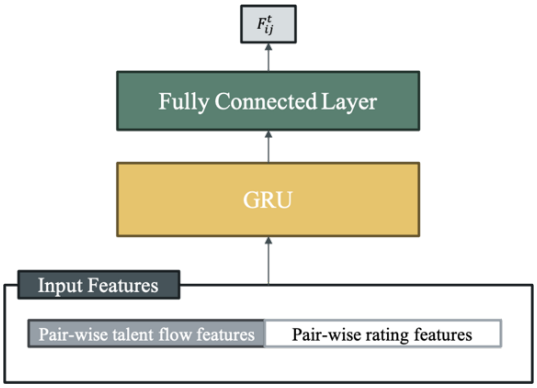


Figure 9. Model Structure of Pair-RNN

- Company Distance-aware Time Series Prediction Model (CDA-RNN)

This additional model considers the relationship or distance between companies when predicting talent flow between them. Figure 10 illustrates the model structure. We first utilize Global Vectors for Word Representation (GloVe) to obtain a company embedding vector for each company (Pennington et al., 2014). GloVe is initially designed to capture semantic relationships between words by analyzing word co-occurrence statistics from a large corpus and obtaining vector representations for words. In this model, we treat each company as a word and the co-occurrence of words corresponds to the talent flow in and

doi:10.6342/NTU202403333

out flow amounts. Companies with higher talent flow amounts are considered to have closer relationships.

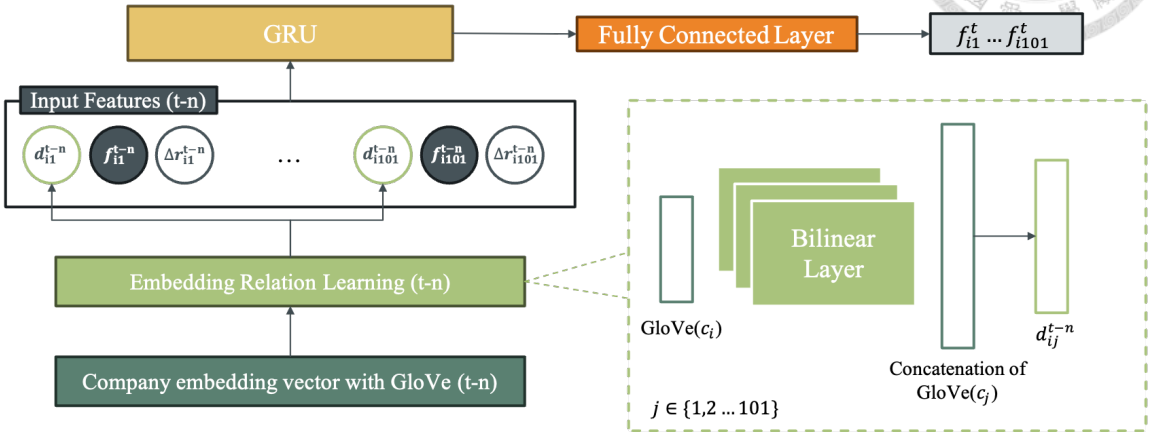


Figure 10. Model Structure of CDA-RNN

Technically, we first transform each time t's talent flow adjacency matrix into a non-directional adjacency matrix, treating it as a co-occurrence matrix input for GloVe to learn each company's embedding. After obtaining the embedding of each company, we use a bilinear layer to capture the interaction between the source company embedding and other companies' embeddings, learning the relationships between each company pair. The relationship result is integrated with each talent flow and rating pair, denoted as d_{i1}^{t-n} in Figure 10. The combined result is then fed into a Bi-directional GRU model and a fully connected layer to predict all future talent outflow values of c_i. The model time window length is set to 4.

5-5. Results

The following tables show the evaluation results for each benchmark model and our model (CAR-TFP) for the position groups "Software Developer Professional" (Table 8),

doi:10.6342/NTU202403333

"Consultant" (Table 9), "Management Professional"(Table 10) and "Data Professional" (Table 11). The results marked in bold represent the best performance for each evaluation metric, while those underlined indicate the second-best results.

According to the results, we can see that our model outperforms other benchmarks in most indicators. Specifically, in the “Software Developer Professional” group (Table 8), our model shows superior performance in all metrics except for precision@5, where it is slightly lower than the best result. This indicates that our model can effectively predict talent flow amounts.

Table 8. Results of “Software Developer Professional”

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
AR	88.4425	33.05%	0.8512	24.11%	22.7065	100.06%	58.66%	49.69%	**48.03%**
VAR	257.1529	112.66%	0.2734	10.29%	16.145	91.23%	55.50%	46.86%	44.04%
Pair-RNN	172.0582	24.32%	3.5084	161.58%	17.4642	88.52%	53.75%	40.86%	32.54%
CDARNN	134.9162	28.06%	0.7026	31.54%	11.8999	69.82%	58.93%	44.81%	38.52%
CAR-TFP	**73.3302**	**14.07%**	**0.2426**	**9.40%**	**10.5559**	**56.91%**	**61.63%**	**50.01%**	46.17%

When comparing the four position groups together, the self-loop’s MAE and MAPE outperform the second-best benchmark by 11-12%. Similarly, the MAE and MAPE for talent flow to OOBC are slightly higher than the second-best benchmark by 5%. However, the MAE and MAPE of the average talent flow to other companies are slightly lower than the linear regression models in the “Consultant” (Table 9), “Management Professional” (Table 10) and “Data Professional” (Table 11) position groups. We attribute this to the

doi:10.6342/NTU202403333

higher sparsity of the talent flow matrices in these position groups. Simpler models tend to predict values near 0 if the former values in the observed time windows are all 0, while more complex models might predict higher values due to the influence of other features or data, increasing the MAE and MAPE of the average talent flow to other companies. Despite this, our model's precision @ k indicators mostly outperform other benchmarks, indicating that our model can still identify the top talent flow values accurately.

Table 9. Results of "Consultant"

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
AR	15.7123	39.69%	**0.0465**	**1.70%**	3.867	56.89%	81.58%	78.38%	77.33%
VAR	25.6199	68.09%	0.0584	2.47%	2.9554	55.80%	73.50%	70.97%	70.23%
Pair-RNN	25.2963	35.44%	0.2244	10.35%	2.5922	28.22%	68.50%	65.11%	63.44%
CDA-RNN	8.6591	24.72%	0.1015	4.65%	1.6966	38.13%	80.27%	76.93%	76.34%
CAR-TFP	**3.9191**	**12.96%**	0.0632	2.72%	**1.3397**	**26.91%**	**83.96%**	**81.30%**	**79.79%**

Table 10. Results of "Management Professional"

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
AR	10.1722	39.47%	0.0554	2.57%	1.8218	47.67%	53.91%	78.77%	78.40%
VAR	16.726	73.53%	**0.0297**	**1.31%**	1.7163	44.01%	75.00%	73.36%	72.90%
Pair-RNN	10.0221	26.61%	0.2626	12.73%	2.5806	57.25%	78.08%	75.75%	74.75%
CDA-RNN	7.153	29.72%	0.0964	7.37%	1.3447	44.27%	79.70%	77.96%	77.27%
CAR-TFP	**3.3279**	**14.86%**	0.0444	2.08%	**0.9734**	**27.18%**	**84.93%**	**82.91%**	**82.20%**

doi:10.6342/NTU202403333

Table 11. Results of "Data Professional"

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
AR	11.5564	51.30%	**0.0195**	**0.73%**	39.078	186.25%	56.66%	76.63%	75.84%
VAR	12.5042	89.80%	0.0205	0.89%	1.5415	39.98%	73.33%	72.22%	71.68%
Pair-RNN	11.4793	29.52%	2.3067	81.75%	2.5213	44.05%	39.66%	37.22%	36.30%
CDA-RNN	5.8071	28.10%	0.0823	3.96%	1.3294	42.30%	70.15%	67.75%	67.45%
CAR-TFP	**2.9327**	**17.94%**	0.0415	1.92%	**1.060**	**29.99%**	**80.19%**	**77.78%**	**77.32%**

Comparing other benchmarks, we find that linear combination models, including AR and VAR, perform better in precision @ k indicators than non-linear benchmark models. However, in larger numeric predictions such as self-loop and OOBC's MAE and MAPE, the non-linear models show higher performance. This suggests that linear models better capture the general trend of talent flow, while non-linear benchmarks are more effective at identifying and predicting peak values in a talent flow vector.

5-6. Sensitivity Test

We conducted a sensitivity test over the time window length. The following Table 12, Table 13, Table 14, Table 15 presents the results for the 4 position groups that we tested in the chapter 5-5. We found that a window length of 2 provides insufficient information about the talent flow trend, resulting in worse performance across most indicators. Conversely, a window length of 6 includes too much noise when predicting talent flow,

doi:10.6342/NTU202403333

leading to decreased numerical accuracy. However, it maintains the ability to predict the general trend, as shown by the results of precision @ k.

Table 12. Sensitivity Test of window size by "Software Developer Professional"

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
Length=2	90.6365	14.93%	0.2867	10.72%	10.9783	**51.69%**	61.64%	48.66%	42.76%
Length=3	74.1926	14.18%	0.2736	10.48%	**10.2959**	55.54%	**62.64%**	**50.21%**	44.29%
Length=4	**73.3302**	14.07%	**0.2426**	**9.40%**	10.5559	56.91%	61.63%	50.01%	**46.17%**
Length=5	73.6780	14.16%	0.2506	9.74%	10.5689	57.78%	61.00%	49.95%	44.75%
Length=6	74.7320	**14.00%**	0.2570	9.98%	10.7480	58.98%	61.77%	50.03%	44.79%

Table 13 Sensitivity Test of window size by "Consultant"

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
Length=2	4.4686	14.08%	0.0747	3.15%	1.4430	**25.49%**	84.29%	81.19%	79.91%
Length=3	3.9824	**12.72%**	0.0703	3.01%	1.4310	27.49%	83.47%	81.02%	79.40%
Length=4	**3.9191**	12.96%	0.0632	2.72%	**1.3397**	26.91%	83.96%	81.30%	79.79%
Length=5	3.9917	13.24%	**0.0629**	**2.71%**	1.4074	27.46%	**84.57%**	81.50%	80.35%
Length=6	4.144	13.55%	0.0641	2.76%	1.4201	27.38%	83.90%	**81.52%**	**80.40%**

Table 14. Sensitivity Test of window size by "Management Professional"

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
Length=2	3.6566	**13.43%**	0.0512	2.38%	1.1406	29.34%	83.97%	81.56%	80.41%
Length=3	3.3483	14.69%	0.0473	2.21%	0.9832	**26.91%**	**85.08%**	**82.94%**	82.06%
Length=4	**3.3279**	14.86%	**0.0444**	**2.08%**	**0.9734**	27.18%	84.93%	82.91%	**82.20%**
Length=5	3.3535	15.21%	0.0452	2.11%	0.9992	28.58%	85.03%	82.88%	82.19%
Length=6	3.4465	15.67%	0.0456	2.13%	0.9901	28.19%	84.93%	82.73%	82.09%

doi:10.6342/NTU202403333

Table 15. Sensitivity Test of window size by "Data Professional"

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
Length=2	3.4625	16.77%	0.0434	2.00%	1.1254	**28.80**%	**81.47%**	**78.91%**	**78.36%**
Length=3	3.1734	18.40%	**0.0407**	**1.87%**	1.0850	29.72%	80.34%	77.94%	77.32%
Length=4	**2.9327**	**17.94%**	0.0415	1.92%	**1.0603**	29.99%	80.19%	77.78%	77.32%
Length=5	3.2092	18.53%	0.0422	1.95%	1.0746	30.39%	80.19%	77.81%	77.46%
Length=6	3.4597	19.20%	0.0427	1.98%	1.0678	30.46%	80.34%	77.78%	77.42%

5-7. Ablation Test

We also conducted an ablation test to evaluate the effectiveness of each module in our model. The following Table 16, Table 17, Table 18, Table 19 presents the ablation test results for the 4 position groups. The experiments included "w/o rating", "w/o dimension reduction", and "w/o bilinear layer & w/ linear layer".

- Experiment: w/o Rating

This experiment removed the company rating data. We observed that the MAE and MAPE for self-loop, as well as precision@2 and precision@3, were slightly better when the rating data was included.

- Experiment: w/o Dimension Reduction

In this experiment, we removed the dimension reduction module, causing the hidden state of the time series learning module to directly feed into the company-aware layer with the company embedding. This resulted in a significant increase in model parameter

doi:10.6342/NTU202403333

weight numbers. The results showed that the absence of the dimension reduction layer caused an average decline of 7% in MAE and MAPE indicators.

- Experiment: w/o Bilinear Layer & w/ Linear Layer

Here, we replaced the bilinear layer in the company-aware layer with a linear layer. The output from the dimension reduction layer was concatenated with the source company embedding and fed into the linear layer. All performance indicators worsened by approximately 10%, with the biggest difference being up to 22% worse than the original model. This highlights the importance of the interaction between the company embedding and the shared hidden state results.

Table 16. Ablation Test of "Software Developer Professional"

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
CAR-TFP	73.3302	14.07%	0.2426	9.40%	10.5559	56.91%	61.63%	50.01%	**46.17%**
w/o rating	**70.1013**	**13.56%**	**0.2367**	**8.89%**	**10.3467**	**51.58%**	**61.75%**	**50.19%**	45.87%
w/o dimension reduction	105.4989	17.79%	0.5963	27.28%	11.1084	61.68%	61.49%	49.53%	43.47%
w/o bilinear layer	115.6652	28.37%	0.5274	21.92%	13.3872	79.13%	58.27%	46.59%	39.91%

doi:10.6342/NTU202403333

Table 17 Ablation Test of "Consultant"

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
CAR-TFP	**3.9422**	**13.06%**	0.0687	3.61%	1.4526	32.39%	84.43%	81.47%	80.52%
w/o rating	4.1206	13.09%	**0.0398**	**1.56%**	**1.3814**	**26.49%**	**85.44%**	**82.40%**	**80.98%**
w/o dimension reduction	7.7799	20.90%	0.4153	20.37%	1.5991	37.71%	72.44%	69.40%	68.18%
w/o bilinear layer	6.6878	29.39%	0.0829	3.62%	2.3642	61.01%	79.01%	76.00%	74.90%

Table 18 Ablation Test of "Management Professional"

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
CAR-TFP	**3.3431**	15.25%	0.0550	2.60%	0.9830	27.87%	84.26%	82.28%	81.52%
w/o rating	3.3494	**15.00%**	**0.0233**	**1.01%**	**0.9741**	**26.71%**	**85.07%**	**83.18%**	**82.25%**
w/o dimension reduction	4.8061	21.49%	0.4278	21.24%	1.1251	34.24%	75.44%	73.37%	72.49%
w/o bilinear layer	7.481	32.99%	0.0559	2.63%	1.3111	41.44%	81.92%	79.79%	79.08%

Table 19 Ablation Test of "Data Professional"

	Self-loop		Avg to other companies		To OOBC		Precision		
	MAE	MAPE	MAE	MAPE	MAE	MAPE	Precision @2	Precision @3	Precision @5
CAR-TFP	**3.1503**	**16.54%**	0.0457	2.30%	1.0511	32.55%	**80.48%**	77.83%	**77.43%**
w/o rating	3.3527	18.47%	**0.0212**	**0.91%**	**1.049**	**28.49%**	80.24%	**77.86%**	77.32%
w/o dimension reduction	4.1649	24.08%	0.4192	20.81%	1.2111	37.59%	71.97%	69.20%	68.79%
w/o bilinear layer	5.3625	30.92%	0.0414	1.90%	1.4848	46.16%	72.93%	70.60%	70.30%

doi:10.6342/NTU202403333

Chapter 6. Conclusion

6-1. Summary

In this study, we focus on the talent flow prediction problem, aiming to utilize historical company talent flow information to predict future talent flow amounts or percentages. This problem is crucial for supporting human resource teams in enhancing companies' talent management strategies and related implications. Despite its importance, few studies have concentrated on this area. We designed a company list-wise structure with a company-aware mechanism in a deep learning approach and leveraged company rating data as model features. Compared to existing models, our approach improved the predictive performance of each position group by an average of 3-4%.

6-2. Contributions

For model design, the company-aware structure in our model significantly enhances performance compared to benchmark models. By using company list-wise features and output structures, our model provides a more detailed distribution of outflow information. Additionally, the company embedding aware layer captures the unique talent flow patterns of different source companies after the shared time series learning layer. Notably, we are the first to incorporate company ratings as a model feature, which reflects employee satisfaction levels over time.

doi:10.6342/NTU202403333

As for managerial contributions, our talent flow prediction model offers valuable guidance for developing effective talent strategies, including recruitment, retention, and turnover prevention. It also serves as a competitiveness detector, identifying potential competitors based on human resource dynamics. By accurately predicting talent flow, companies can better plan and implement strategies to maintain a competitive edge in the labor market.

6-3. Future Research Directions

- Optimizing the Current Model

To enhance our current model, increasing the variety of data, particularly for positions beyond software engineers, is crucial. This expansion will improve generalization and provide deeper insights across various roles. Another approach is to develop a multi-task learning framework that simultaneously predicts talent flows for multiple positions within a company. This framework can leverage shared information across different job positions, improving prediction accuracy. Additionally, considering the interactions and dependencies between different job position groups, such as the impact of talent flow between software professionals and consultant roles, could refine predictions. A deeper analysis of company reviews is also recommended. While we used company rating scores, their influence was limited. Analyzing the textual content of reviews through sentiment analysis could extract new features, such as underlying sentiments and factors influencing

doi:10.6342/NTU202403333

employee satisfaction. This could involve forming arrays of mentioning factors and performing cosine similarity between companies.

- Extension Related Topics

Future research could explore the connection between talent flow information and other company actions, such as predicting new product lines or market entries based on talent recruitment patterns. This approach could support tasks related to predicting company actions and understanding market dynamics. By linking talent flow data with strategic company decisions, we can gain a more comprehensive understanding of how talent dynamics influence overall business strategies and competitiveness.

doi:10.6342/NTU202403333

References

Carr, S. C., Inkson, K., & Thorn, K. (2005). From global careers to talent flow: Reinterpreting 'brain drain'. *Journal of World Business*, *40*(4), 386-398.

Cheng, Y., Xie, Y., Chen, Z., Agrawal, A., Choudhary, A., & Guo, S. (2013). Jobminer: A real-time system for mining job-related patterns from social media. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1450-1453).

Glassdoor (n.d.). About Glassdoor. Retrieved July 16, 2024, from https://www.glassdoor.com/about/

Groysberg, B. & Abrahams, R., (2006). Lift outs: How to acquire a high-functioning team. *Harvard Business Review, 84*(12), 133-140.

Freeman, W. T., & Tenenbaum, J. B. (1997). Learning bilinear models for two-factor problems in vision. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 554-560.

Hongal, P., & Kinange, U. (2020). A study on talent management and its impact on organization performance-an empirical review. *International Journal of Engineering and Management Research*, *10*, 64-71.

doi:10.6342/NTU202403333

Hu, T. S., & Chen, K. C. (2014). Creative talent drive transformation of professionals' constitution in the modern city: A case study of fashion talent flow in Taipei. *European Planning Studies*, *22*(5), 1081-1105.

Jackson, D. J., Carr, S. C., Edwards, M., Thorn, K., Allfree, N., Hooks, J., & Inkson, K. (2005). Exploring the dynamics of New Zealand's talent flow. *New Zealand Journal of Psychology*, *34*(2), 110-117.

Liu, J., Ng, Y. C., Wood, K. L., & Lim, K. H. (2020). IPOD: A large-scale industrial and professional occupation dataset. In *Companion Publication of the 2020 Conference on Computer Supported Cooperative Work and Social Computing*, pp. 323-328.

LinkedIn (n.d.). About LinkedIn. Retrieved July 17, 2024, from https://about.linkedin.com/zh-tw?lr=1

Mao, G., Hu, B., & Song, H. (2009). Exploring talent flow in Wuhan automotive industry cluster at China. *International Journal of Production Economics*, *122*(1), 395-402.

McKinsey & Co. (2018, August 7). Winning with your talent-management strategy. https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/winning-with-your-talent-management-strategy

doi:10.6342/NTU202403333

McKinsey & Co. (2021, September 8). ‘Great Attrition’ or ‘Great Attraction’? The choice is yours. https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/great-attrition-or-great-attraction-the-choice-is-yours

Oentaryo, R. J., Lim, E.-P., Ashok, X. J. S., Prasetyo, P. K., Ong, K. H., & Lau, Z. Q. (2018). Talent flow analytics in online professional network. *Data Science and Engineering*, *3*, 199-220.

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532-1543.

Qin, C., Zhang, L., Zha, R., Shen, D., Zhang, Q., Sun, Y., Zhu, C., Zhu, H., & Xiong, H. (2023). A comprehensive survey of artificial intelligence techniques for talent analytics. *arXiv preprint arXiv:2307.03195*.

Song, W., & Wang, C. (2022). Hybrid recommendation based on matrix factorization and deep learning. In *Proceedings of the 4th International Conference on Big Data Engineering*, pp. 81-85.

State, B., Rodriguez, M., Helbing, D., & Zagheni, E. (2014). Migration of professionals to the US: Evidence from LinkedIn data. In *Proceedings of 6th International Conference on Social Informatics (SocInfo 2014)*, Barcelona, Spain, pp. 531-543.

doi:10.6342/NTU202403333

Tarique, I., & Schuler, R. S. (2010). Global talent management: Literature review, integrative framework, and suggestions for further research. *Journal of World Business*, *45*(2), 122-133.

Wang, P. L. (2023). What's next after quitting? Predicting the next job using time-series embeddings. Unpublished Master's Thesis, Department of Information Management, National Taiwan University, Taipei, Taiwan, ROC.

Xu, H., Yu, Z., Yang, J., Xiong, H., & Zhu, H. (2016). Talent circle detection in job transition networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.655-664.

Xu, H., Yu, Z., Yang, J., Xiong, H., & Zhu, H. (2019). Dynamic talent flow analysis with deep sequence prediction modeling. *IEEE Transactions on Knowledge and Data Engineering*, *31*(10), 1926-1939.

Zhang, L., Xu, T., Zhu, H., Qin, C., Meng, Q., Xiong, H., & Chen, E. (2020). Large-scale talent flow embedding for company competitive analysis. In *Proceedings of the World Wide Web Conference* (pp. 2354-2364).

Zhang, L., Zhu, H., Xu, T., Zhu, C., Qin, C., Xiong, H., & Chen, E. (2019). Large-scale talent flow forecast with dynamic latent factor model? In *Proceedings of the World Wide Web Conference*, pp. 2312-2322.

doi:10.6342/NTU202403333

Zhou, Y., Guo, Y., & Liu, Y. (2018). High-level talent flow and its influence on regional unbalanced development in China. *Applied Geography*, *91*, 89-98.

doi:10.6342/NTU202403333