Your Career Path Matters in Person-Job Fit

Zhuocheng Gong[1], **Yang Song**[2], **Tao Zhang**[2], **Ji-Rong Wen**[3], **Dongyan Zhao**[1,4,5†], **Rui Yan** [3†]
[1] Wangxuan Institute of Computer Technology, Peking University
[2] BOSS Zhipin
[3] Gaoling School of Artificial Intelligence, Renmin University of China
[4] National Key Laboratory of General Artificial Intelligence
[5] Beijing Institute for General Artificial Intelligence
{gzhch,zhaody}@pku.edu.cn, {ruiyan,jrwen}@ruc.edu.cn, {songyang,kylen.zhang}@kanzhun.com

Abstract

We are again confronted with one of the most vexing aspects of the advancement of technology: automation and AI technology cause the devaluation of human labor, resulting in unemployment. With this background, automatic person-job fit systems are promising solutions to promote the employment rate. The purpose of person-job fit is to calculate a matching score between the job seeker's resume and the job posting, determining whether the job seeker is suitable for the position. In this paper, we propose a new approach to person-job fit that characterizes the hidden preference derived from the job seeker's career path. We categorize and utilize three types of preferences in the career path: **consistency**, **likeness**, and **continuity**. We prove that understanding the career path enables us to provide more appropriate career suggestions to job seekers. To demonstrate the practical value of our proposed model, we conduct extensive experiments on real-world data extracted from an online recruitment platform and then present detailed cases to show how the career path matters in person-job fit.

Introduction

The job market is experiencing a rapid transformation in hiring practices due to the emergence of online recruitment services. The ongoing AI revolution has further accelerated the urgent need for an effective online recruitment system. Recent breakthroughs in NLP (e.g. ChatGPT*) and CV (e.g. Diffusion Model) have given us a glimpse into a future where human labor may no longer be necessary for many industries, potentially leading to a decline in employment. As we stand on the brink of another technological revolution that will reshape human society, we believe that employment is now more crucial than ever as a social issue. Therefore, the task of person-job fit, which seeks to automatically match job seekers with suitable employers, has garnered significant attention.

Numerous researchers have delved into the concept of person-job fit using vast amounts of online recruitment data. A common approach among these studies is to frame the task as a supervised text-matching problem, where resumes and job postings are treated as text pieces and various text mining and deep learning techniques are applied to the task. Based on this approach, different perspectives have been explored to improve performance, including job-oriented ability modeling (Qin et al. 2018), user preference mining (Yan et al. 2019), psychological motivation modeling (Le et al. 2019), and two-way relationship modeling (Yang et al. 2022). These methods have demonstrated significant success and achieved promising milestones in real-world applications with commercial value.

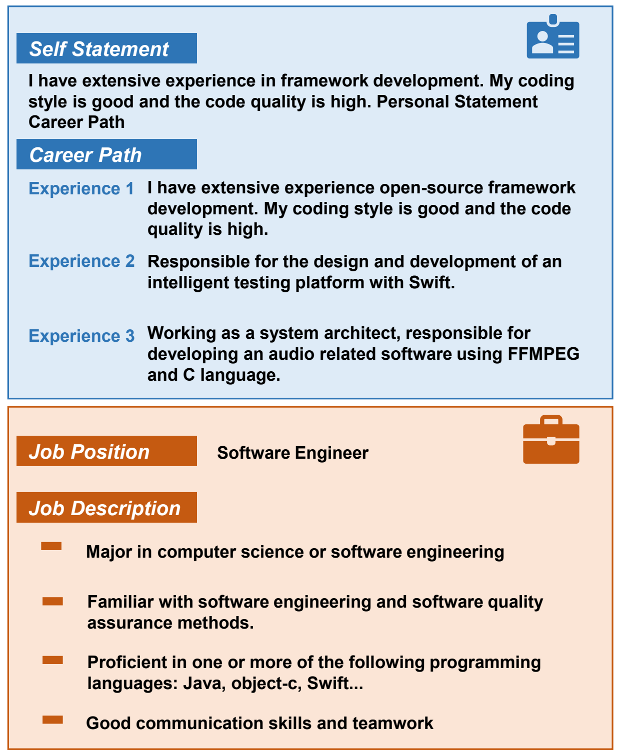

Figure 1: An illustrative example for person-job fit in our scenario. The upper part indicates a candidate's resume and the lower part denotes a job posting.

†Corresponding authors: Dongyan Zhao (zhaody@pku.edu.cn) and Rui Yan (ruiyan@ruc.edu.cn).

*https://openai.com/blog/chatgpt/

Although significant progress has been made in automatic online recruitment, there is an important aspect that has been overlooked - career path information. Specifically, a resume typically includes a brief *self statement* section and a *career path* section that outlines the job seeker's work experiences from their first employment to their current job situation. We believe that the career path is a critical factor in job seeking, and analyzing the career path of each job seeker can reveal hidden preferences that can enhance person-job fit.

We have identified three types of preferences inferred from the career path: **consistency**, **likeness**, and **continuity**. Consistency measures whether the job seeker's work experiences are consistent with each other and with their overall career path. For instance, in Figure 1, the IT practitioner's work experiences are all IT-related and therefore consistent with each other. Likeness measures how closely a job posting aligns with a candidate's work experiences, allowing us to identify their job preferences. By analyzing work experiences, we can determine a candidate's job preferences and understand whether they are interested in a particular type of job or not. Continuity involves the evolution of work experiences over time, including the accumulation of job skills and new responsibilities. By tracking the evolving nature of work experiences along the career path, we can filter out positions that a candidate is overqualified for but still a good fit. For example, in Figure 1, a job seeker who started as a software development engineer (SDE) and then became a senior software development engineer (Senior SDE) may be a good fit for a software architect position.

Our aim is to enhance the person-job matching task by incorporating previous work experiences of job-seekers within their resumes. We recognize the benefits of characterizing talent career paths and propose to use them to better understand job-seekers' preferences, working levels, skill status, and future career development. The career path is considered a sequence of work experiences in chronological order. To utilize this information, we propose two self-supervised auxiliary objectives and a contrastive objective to model the consistency, continuity, and likeness of the career path, respectively. We then integrate the captured preferences into our proposed person-job matching framework, which we call the **W**ork **E**xperience enhanced **P**erson-**J**ob **M**atching model (**WEPJM**).

To sum up, our contributions are manifold as follows:

• We investigate the resume information in fine-grain and capture and utilize different aspects of career path preference. We incorporate techniques of self-supervised learning and contrastive learning to model the career path.

• We conduct experiments on real-world data. Experimental validation confirms that our framework matches job seekers with jobs that better align with their preferences and experience, indicating that AI technology helps to address employment issues.

Related Work

There is an increasing number of studies on recruitment-oriented talent science that covers a lot of topics, including person-job matching (Zhu et al. 2018; Qin et al. 2018; Bian et al. 2019; Le et al. 2019; Yan et al. 2019; Luo et al. 2019; Bian et al. 2020; Jiang et al. 2020), job mobility prediction (Meng et al. 2019; Zhang et al. 2019; Xu et al. 2015; Li et al. 2017), person-organization fit (Sun et al. 2019), job skill mining (Qin et al. 2019; Wu et al. 2019; Xu et al. 2018) and organization analysis (Lin et al. 2017).

Our study is highly related to the topic of person-job fit. Pioneering work of (Malinowski et al. 2006) learns the representation of jobs and talents with approaches based on latent factors. Along this line of research, (Zhu et al. 2018) proposes to encode the job and resume with two convolutional neural networks (CNN) respectively and calculate matching scores by cosine similarity. (Qin et al. 2018) leverages hierarchical recurrent neural networks (RNN) to encode the documents and incorporates the attention mechanism to model job abilities and skills. More recently, (Yan et al. 2019) incorporates interview history from both the job-seeker and the recruiter, and (Luo et al. 2019) introduces an adversarial training method in the representation learning of the job posting which constrains the representation to be indistinguishable from a prior Laplace distribution. (Bian et al. 2020) proposes a co-teaching network to handle both text-based matching information and relation-based information in a single framework. Recently, (Yang et al. 2022) proposes to model two-way selection preference for person-job fit. (Wang et al. 2022) uses co-attention and GNN to model the related recruitment history, achieving promising performance. In this work, we study the effect of modeling career path preference.

Recently, contrastive learning has attracted much attention for its strong performance on sentence representation learning (Fang et al. 2020; Giorgi et al. 2021). In this work, we propose a contrastive learning strategy to align the representations of job postings and their related work experiences.

Notations and Task Formulation

In this paper, we aim to tackle the person-job matching task, which measures whether a particular job position is suitable for the background of a candidate job-seeker. To formulate the task, we use $\mathrm{j}=\{\mathrm{j}_1, \mathrm{j}_2, \ldots, \mathrm{j}_{|\mathrm{j}|}\}$ to denote a job description. The job description j contains $|\mathrm{j}|$ sentences where each sentence j_i describes the job requirements and/or job responsibilities for the work position. Similarly, we use r to represent the resume of a candidate's talent. As mentioned in Section , we decompose the resume r into two parts, i.e., $\mathrm{r} = \mathrm{s} \cup \mathrm{c}$ where s stands for the statement part and c indicates the career path of work experiences. We denote $\mathrm{s} = \{\mathrm{s}_1, \mathrm{s}_2, \ldots, \mathrm{s}_{|\mathrm{s}|}\}$. A career path c consists of a sequence of work experiences as $\{\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_{|\mathrm{c}|}\}$ where c_i indicates a particular work experience as $\mathrm{w}_i = \{\mathrm{w}_{i1}, \mathrm{w}_{i2}, \ldots, \mathrm{w}_{i|\mathrm{w}_i|}\}$. Unlike the job part and the statement part, the career path is organized as a hierarchical structure from the sentence level to the career level.

The task of person-job fitting is formally defined as a classification problem that predicts the matching degree given the resume and the job description. For each pair of (j, r), we have the corresponding recruitment label $y \in \{0, 1\}$, which indicates whether the selected candidate and the job position

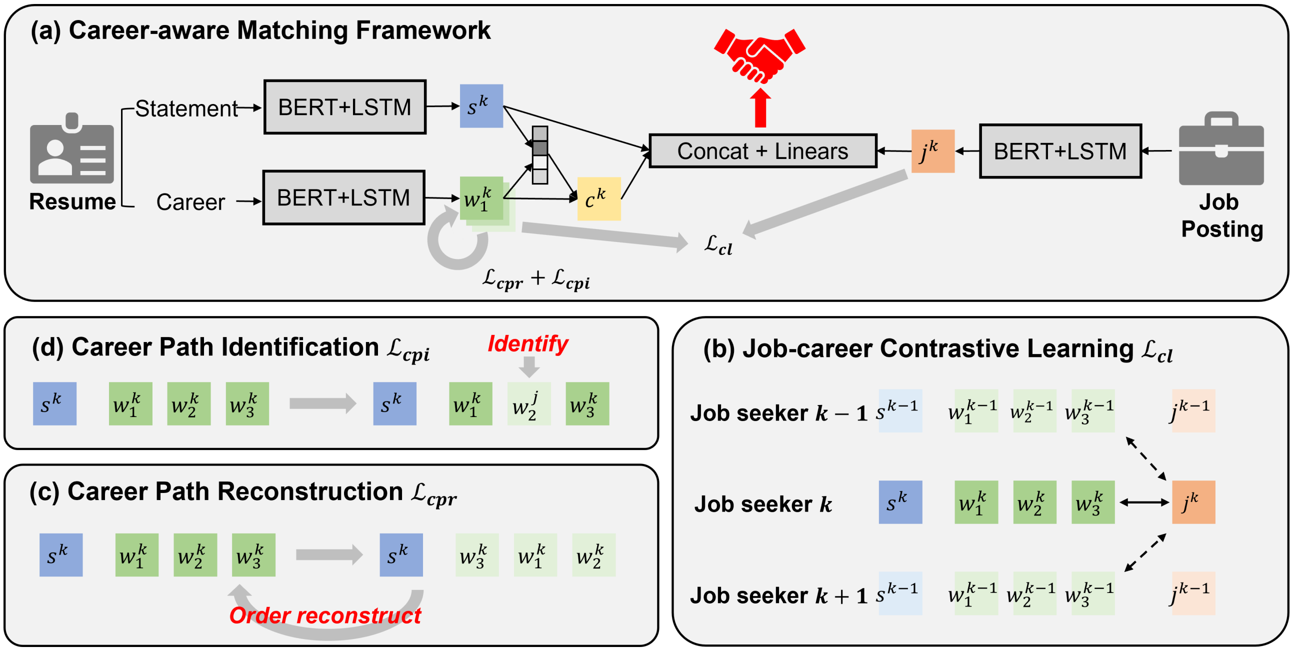

Figure 2: Model overview. WEPJM learns representations for input resumes and job postings with hierarchical encoders. The overall objective can be decomposed into the main task of person-job fit and three auxiliary tasks. The auxiliary tasks include (b) *Career Path identification*, (c) *Career Path Reconstruction*, as well as (d) *Job-career Contrastive Learning*.

form a good matching pair (i.e., y=1) or not (i.e., y=0). Our objective is to learn a matching function $\mathrm{M}(\mathrm{r}, \mathrm{j})$ that maximizes the probability of predicting the right matching decision given the candidate's resume and the job description. Formally,

$$\hat{\mathrm{M}} = \arg\min_{\mathrm{M}} loss(\mathrm{M}(\mathrm{r}, \mathrm{j}), y) \tag{1}$$

Proposed Method

In this section, we focus on the details of our proposed **WEPJF** as shown in Figure 2. We first introduce the overall architecture design and then elaborate on the auxiliary objectives introduced to model the career path preference.

Work Experience Enhanced Person-Job Matching Framework

A person-job fit system deals with inputs from two sides, namely the resume of the job seeker and the job posting. As described before, we decompose the resume into the statement part and the career path part, processing them with separate text encoders. The text encoder is a Bi-LSTM on top of the pre-trained BERT sentence encoder, where BERT is used to capture sentence-level semantics while LSTM captures document-level semantics. After obtaining the statement representation **s** and a series of work representations $\{\mathbf{w}_i\}_{i=1}^{|\mathrm{c}|}$, we apply the attention mechanism to inject the statement semantic into the representation of work experiences. Then we take the weighted sum of the work experience representations as the career representation, denoted as **c**, formally,

$$\begin{aligned} \mathbf{e}_i &= \tanh(\mathbf{W}_1\mathbf{s} + \mathbf{W}_2\mathbf{w}_i), \\ \alpha_i &= \frac{\exp(\mathbf{v}^T\mathbf{e}_i)}{\sum_{j=1}^{|\mathrm{c}|}\exp(\mathbf{v}^T\mathbf{e}_j)}, \\ \mathbf{c} &= \sum_{i=1}^{|\mathrm{c}|}\alpha_i\mathbf{w}_i \end{aligned} \tag{2}$$

The encoding process for the job posting side is the same. After extracting the statement representation **s**, the career representation **c** and the job representation **j**, then we concatenate them together and feed to a multi-layer perception to calculate the matching score. The main objective for person-job fit is defined as:

$$\mathcal{L}_{\mathrm{main}} = -y\log(\mathrm{M}(\mathrm{r}, \mathrm{j})) - (1-y)\log(1-\mathrm{M}(\mathrm{r}, \mathrm{j})). \tag{3}$$

Mining Career Path Preference

As described above, we categorize the preference inferred from the career path into three types, namely **consistency**, **likeness**, and **continuity**. In this section, we elaborate on the details of how we model such preference. For the consistency of the career path, we introduce the task of *career path identification*. For maintaining the continuity of the career path, we proposed the task of *career path reconstruction*. Finally, for the likeness aspect, we propose a *job-career contrastive learning* objective.

Career Path Identification In the auxiliary task of Career Path Identification (CPI), we replace some of the work experiences in the career path with random work experiences sampled from the pool of work experiences. We train the model to identify whether they are original or replaced work experiences. Concretely, we assign each work experience w_i with a label $y_i^{\mathrm{CPI}} \in \{0, 1\}$, where $y_i^{\mathrm{CPI}} = 1$ represents the i-th work experience remains unchanged, and $y_i^{\mathrm{CPI}} = 0$ indicates that the i-th work experience is randomly replaced.

We design the module for career path identification with joint supervision from the statement part in the resume. This auxiliary task is built upon an MLP classifier to output a score indicating whether the work experience belongs to the same job seeker without any changes. The module is formulated as $\mathrm{M}_{\mathrm{CPI}}(\mathrm{s}, \mathrm{c}, \mathrm{w}_i)$, where w_i is the work experience to be examined. For each job seeker, we replace at most one work experience, ensuring the replacement has minimal impact on the main task.

The CPI task actually characterizes the latent preference of the job-seeker, indicating whether a particular work experience belongs to the talent or not. The task is optimized with cross-entropy loss, which can be formulated as:

$$\begin{aligned} \mathcal{L}_{\mathrm{CPI}} = &- y_i^{\mathrm{CPI}} \log(\mathrm{M}_{\mathrm{CPI}}(\mathrm{s}, \mathrm{c}, \mathrm{w}_i)) \\ &- (1 - y_i^{\mathrm{CPI}}) \log(1 - \mathrm{M}_{\mathrm{CPI}}(\mathrm{s}, \mathrm{c}, \mathrm{w}_i)). \end{aligned} \quad (4)$$

Career Path Reconstruction As we have mentioned, the career path is not just a simple collection of work experiences but an evolving work experience sequence over time, so we model the "path" information with this auxiliary task. The development of the career path is (to some extent) order-preserving. With the accomplishment of junior-level positions, the job-seeker will be eligible for senior-level choices.

In order to characterize such intuition, we shuffle the order of the job seekers' work experiences and regard the disordered career path as a negative training instance in contrast to the original ones as positive instances. Again, we apply the MLP classifier to distinguish the positive/negative samples, namely $\mathrm{M}_{\mathrm{CPR}}(\mathrm{s}, \mathrm{c})$. For this task, half of the training samples are shuffled. Those with the wrong order are masked from the main task to avoid the negative transfer.

Also, the CPR task is optimized with cross-entropy loss, which can be formulated as:

$$\begin{aligned} \mathcal{L}_{\mathrm{CPR}} = &- y^{\mathrm{CPR}} \log(\mathrm{M}_{\mathrm{CPR}}(\mathrm{s}, \mathrm{c})) \\ &- (1 - y^{\mathrm{CPR}}) \log(1 - \mathrm{M}_{\mathrm{CPR}}(\mathrm{s}, \mathrm{c})). \end{aligned} \quad (5)$$

Job-career Contrastive Learning There are similarities between work experience and job descriptions, as both describe content related to work. To enhance the relationship modeling between work experiences and job descriptions of job postings, we propose a job-career contrastive learning strategy to close the continuous representation gap between work experiences **w** and job postings **j**. Specifically, for a given job posting, we treat the last work experience from the positive person-job pair as a positive example. Work experiences from other job seekers and other job postings in the same batch as negative examples. Therefore, the model is optimized by minimizing the objective function:

$$\mathcal{L}_{\mathrm{CL}} = -\log \frac{e^{cos(\mathbf{w}^k_{|\mathrm{c}^k|}, \mathbf{j}^k)/\tau}}{\sum_{i \neq k} (e^{cos(\mathbf{w}^i_{|\mathrm{c}^k|}, \mathbf{j}^i)/\tau} + e^{cos(\mathbf{j}^i, \mathbf{j}^k)/\tau})}, \quad (6)$$

where τ is the temperature hyperparameter.

Statistics	Values
# of job postings	82,362
# of resumes	33,285
# of work experiences in resumes	117,780
avg # of work experiences per resume	3.57
avg # of sentences per job posting/resume	11.83/19.75
avg # of words per job posting/resume	114.68/258.50
# of positive person-job pairs	119,031
# of negative person-job pairs	359,721

Table 1: The statistics of the dataset.

Training WEPJF

We apply a multi-task learning manner for **WEPJF**. The general objective function can be formulated as follows:

$$\begin{aligned} \mathcal{L}_{overall} &= \mathcal{L}_{main} + \mathcal{L}_{auxiliary}, \\ \mathcal{L}_{auxiliary} &= \lambda_{\mathrm{CPI}}\mathcal{L}_{\mathrm{CPI}} + \lambda_{\mathrm{CPR}}\mathcal{L}_{\mathrm{CPR}} + \lambda_{\mathrm{CL}}\mathcal{L}_{\mathrm{CL}}, \end{aligned} \quad (7)$$

where λ_{CPI}, λ_{CPR} and λ_{CL} are hyperparameters that adjust the weight of corresponding auxiliary objective functions.

Experiments

In this section, we conduct extensive experiments on a real-world dataset to evaluate our proposed model. We first describe the dataset, experimental setups, and baseline methods. Then we compare our proposed model with the state-of-the-art neural network based person-job matching baselines in terms of *accuracy*, *precision*, *recall*, $F1$, and AUC metrics. Moreover, we explore the impact of career path modeling, indicating auxiliary tasks through ablation studies.

Experiment Setup

Dataset We build a dataset by collecting data from a real-world online recruiting platform. † We anonymize all identity information to protect the privacy of job-seekers and recruiters. We collect both positive samples and negative samples of person-job pairs annotated by the system log: if a candidate's talent chats with a job recruiter on the platform, the person-job pair will be labeled as positive. If the talent views the job profile without taking further action, it will be labeled as a negative instance. We summarize the statistics of the dataset in Table 1.

Comparison Methods We compare our model with both classic classification methods and the latest neural network approaches.

We include Logistic Regression (**LR**) (Galton 1886), Decision Tree (**DT**) (Quinlan 1987), Naive Bayes(**NB**) (Rish et al. 2001), Random Forests (**RF**) (Pal 2005), and Gradient Boosting Decision Tree (**GBDT**) (Friedman 2001) as traditional classification methods. For these methods, we use BERT as the feature extractor that outputs semantic representations. Then the extracted representations are concatenated and fed to the classical baseline methods.

We include the following neural baselines:

†anonymouswebsite

Methods	Acc.	Prec.	Recall	F1	AUC
LR	0.525	0.580	0.523	0.550	0.545
DT	0.624	0.602	0.629	0.615	0.627
NB	0.522	0.472	0.524	0.497	0.530
RF	0.648	0.722	0.629	0.673	0.696
GBDT	0.557	0.612	0.552	0.580	0.594
PJFNN	0.743	0.780	0.693	0.737	0.836
APJFNN	0.760	0.790	0.722	0.750	0.847
JRMPM	0.780	0.799	0.743	0.773	0.870
ResumeGAN	0.774	0.794	0.743	0.767	0.863
PJFCANN	0.813	0.824	0.807	0.816	0.894
BERT	0.768	0.794	0.728	0.760	0.855
BERT+TAPT	0.804	0.819	0.780	0.798	0.887
WEPJM	**0.850***	**0.844***	**0.829***	**0.837***	**0.929***

Table 2: Overall performance of all methods. '$\star$' indicates that we accept the improvement hypothesis of our model over the best baseline at a significant test level of 0.01.

• **PJFNN** (Zhu et al. 2018) leverages CNN to extract text representations and then calculates the cosine similarity to model the resume-job relations.

• **APJFNN** (Qin et al. 2018) extracts ability-aware representations for resumes by incorporating an attention mechanism to align key information in job postings to the resume documents.

• **JRMPM** (Yan et al. 2019) considers the preference of both the job-seeker and the recruiter by leveraging two memory modules to "remember" records in application history.

• **ResumeGAN** (Luo et al. 2019). This approach integrates different types of information in a sophisticated way and introduces adversarial learning to learn more expressive representation.

• **PJFCANN** (Wang et al. 2022) uses co-attention and GNN to model the related recruitment history.

• **BERT** (Devlin et al. 2018). For this approach, we use BERT-base as the backbone model and fine-tune the pre-trained model.

• **BERT+TAPT** (Gururangan et al. 2020). This approach performs task-adaptive pre-training (TAPT) on the task dataset to adapt pre-trained BERT to the person-job matching task.

Implementation Details We use BERT-base to learn sentence-level representations. ‡ After encoding with BERT, the dimension of hidden states is set to 200. The batch size is set to 16. λ_{CPI}, λ_{CPR} and λ_{CL} are search in $\{0.1, 1\}$. The temperature hyperparameter is set to 1. The model is trained with Adam optimizer (Kingma and Ba 2014) with the learning rate initialized as 5e-4. Both the size of the validation set and the testing set are set to 3840 with 1920 positive samples and 1920 negative samples. The training will be early stopped if the evaluation results do not increase for 3 successive epochs.

‡https://github.com/huggingface/transformers

	Acc.	Prec.	Recall	F1	AUC
All	0.850	0.844	0.829	0.837	0.929
No	0.813	0.818	0.794	0.803	0.889
+CPI	0.827	0.832	0.823	0.828	0.917
+CPR	0.819	0.821	0.815	0.818	0.898
+CL	0.821	0.825	0.821	0.824	0.904
-CPI	0.831	0.833	0.814	0.823	0.913
-CPR	0.842	0.833	0.822	0.827	0.917
-CL	0.834	0.833	0.816	0.825	0.912

Table 3: Ablation studies of auxiliary tasks. We run experiments to test different combinations of the 3 auxiliary tasks, i.e., CPI, CPR, and CL. '+' indicates to use the single auxiliary task only while '-' denotes to exempt the auxiliary task from all three tasks. We enumerate all possible combinations of the auxiliary tasks.

Experiment Results

Overall Performance We discuss the overall performance of WEPJM and comparison methods. The evaluation results are reported in Table 2. We observe that neural network based methods generally outperform traditional classification algorithms by a large margin, showing the advantage of neural networks in capturing deep semantic information from text, which is consistent with previous studies (Yan et al. 2019; Qin et al. 2018; Bian et al. 2020).

By comparison, we can see that APJNN outperforms PJFNN, indicating that the quality of learned representations is a strong factor that has a sufficient impact on the final matching performance. Besides, JRMPM incorporates historical information to enhance the learning of representations for both resumes and jobs, resulting in better performance Different from the previously mentioned methods, ResumeGAN integrates different types of information to learn more expressive representations. The performance of ResumeGAN is compatible with JRMPM while higher than the rest of the non-pre-trained baseline methods by a large margin, which demonstrates that explicitly modeling different information flow—instead of merging everything in a coarse grain—can improve the performance. PJFCANN performs best in baselines. Without sophisticated architecture design, directly fine-tuning a pre-trained BERT can achieve a decent performance when compared with other baselines. After adding the task-adaptive pre-training, the model can better capture the domain knowledge, resulting in a boost in performance.

We observe that the results of our proposed method have overall advantages over the performance of all the baselines on all the metrics and the improvements have passed the significance test, i.e., t-test with p-value¡0.01. The results indicate that our hypothesis to explicitly model the career path to improve the person-job matching performance has been verified. Next, we proceed to investigate how the different model components contribute to the overall performance by conducting ablation studies.

Self Statement
I am a software develop that have worked in the development of six software. I am familiar with the whole process of software development and have a deep understanding of software design principles. I am proficient in Java and C# and familiar with most mainstream programming languages.

Career Path
Work Experience 1 - Junior Programmer
- Participated in the development of software applications under the guidance of experienced co-workers.
- Guided by team leader to practice job-required skills including Java…

Work Experience 2 - Experienced Software Developer
- Participated as a core programmer in the development of software applications, collaborating with other teammates.
- Designed and implemented software features and functionality.
- Be proficient in certain programming languages.

Work Experience 3 - Experienced Software Developer
- Participated as a core programmer in the development of software applications.
- Discuss with the project manager.

Work Experience 4 - Senior Software Developer
- Led the design of a software application.
- Be familiar with the complete process of software development.
- Be proficient in program language such as Java, C# …

Job Description (PJFNN)
- Work with a team to develop software applications
- Self-motivated and passionate and be good at communication
- Be familiar with Java

Job Description (APJFNN)
- Have experience in software development.
- Expertise in one or more programming languages, such as Java, Python.

Job Description (WEPJF)
- Minimum of 5 years of experience in software development, participated in at least 3 development projects.
- Lead the development of software and collaborate with other software developer.
- Experience with project management.
- Have good problem-solving abilities

Self Statement
I am a graduate of animation major and proficient in hand drawing. I am skillful in PS, AI and other painting software. I can adapt to a variety of painting styles while maintaining steady preference over different styles. I have experience in creating IP illustrations and designing derivative products. I intend to develop myself in the IP industry.

Career Path
Work Experience 1 - Junior Character Art Designer
- Conducted the character painting assigned by the editor and senior designers
- Participated in modifying and polishing character artwork according to the comments given by the editor

Work Experience 2 – Creative Advertisement Designer
- Designed game advertising images, logos and icons.
- Be responsible for designing posters, product images, comic pictures, emoji gifs, banner images, etc.
- Be responsible for the design of daily images, posters of "XXX" Weibo and post-editing video

Work Experience 4 – Art Director
- Initiated the creation of a new IP and its related products.
- Hand-painted the characters of the hottest anime and make related products.
- Cooperated with other artists to accomplish other design work, giving revision opinions and suggestions.
- Worked as a group leader, design and develop new IP and run the Weibo account.

Job Description (PJFNN)
- Have solid art foundation on poster design, and be creative
- Have creative thinking and unique design innovation capabilities.
- Proficient in graphic design software.

Job Description (APJFNN)
- Design posters according to requirements.
- Be skillful in PS and AI.
- Major in art and be self-motivated.
- Have a team spirit

Job Description (WEPJF)
- Responsible for the overall design and planning, manage the work plan of the design department.
- Organize design meetings and propose insightful opinions.
- Responsible for the company's brand designing and UI designing.

Figure 3: Two cases to show the advantage of our model over other methods by capturing the evolving career development over time. The blue parts are the job seekers' resumes while the red parts are the job postings selected by different methods. Due to space limits, we remove unrelated content and simplify some detailed content in both resumes and job postings.

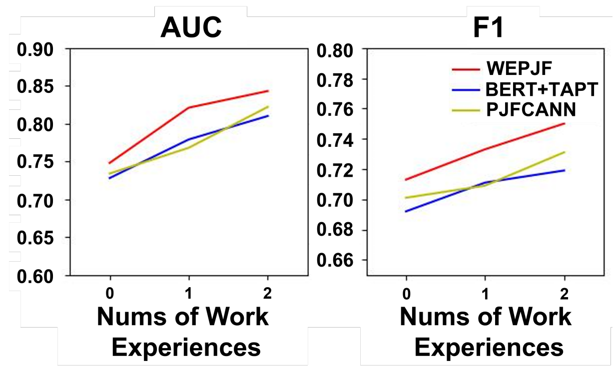

Figure 4: Performance of fresh job seeker (job seekers with less than 3 work experiences).

Does three auxiliary tasks really work? As we include three auxiliary tasks to model the career path of the job-seekers, now we study the contribution of each auxiliary task. We consider all possible combinations of the three auxiliary tasks where each auxiliary task can either be included or exempted.

Now we analyze the effects of different model variants on the person-job matching performance. The results are reported in Table 3. Firstly, removing all three tasks (denoted as 'No Auxiliary') causes a significant decrease in performance: all the other variants perform better compared with "No Auxiliary", indicating that all auxiliary tasks can bring benefits for the person-job matching task.

Then we consider cases with only one auxiliary task, i.e., "+CPI", "+CPR", and "+CL". By comparing the results, we can roughly see the different contributions of various auxiliary tasks. Specifically, career path identification (CPI) performs best out of three, which leads to a 2.5% improvement, followed by adding career-job contrastive learning (CL). The model learns the talent's job preference from CPI by identifying whether a work experience belongs to a particular talent, and learns to align the representations of work experiences and jobs with the contrastive objective. Similar observations have been identified from the "leave-one-out" examinations, i.e., '-CPI', '-CPR', and '-CL', which concur with our conclusion that all auxiliary tasks can help.

What if the job seeker has no/few work experiences? A basic fact about the employment market is that every job seeker has their first job. Therefore, we are curious about how well WEPJF performs for fresh job seekers with few or zero previous work experiences. To test this, we filter out the talents with less than 3 work experiences and evaluate the model's performance. As shown in Figure 4, the model's performance degrades with fewer work experiences, as it becomes more challenging to infer career path preferences accurately. However, there is still a significant performance gain over other text-matching baselines. We attribute this advantage to the sophisticated modeling of the career path, from which the model gains "virtual" career path preference by aligning the resume of the fresh job seeker with that of experienced ones.

Case Study

Finally, we use a case study to demonstrate how career path information matters in person-job fit. The first case is a software developer whose career path clearly reveals a progressive development. The job seeker grows from a junior programmer to an experienced developer, accumulating practical development experiences and improving job skills during his rich experiences. If looking into his career path, we can see that his third and fourth work experiences already relate to more complex responsibilities other than programming. He definitely is overqualified for the job positions proposed by PJFNN and APJFNN which require no more than software development experience.

The second case in Figure 3 shows the resume of an art designer, from which we can see an evolving path in the previous works. The candidate starts from a junior position and can only take orders from editors. Afterward, this talent starts to design advertising images and logos and is responsible for poster designs. The latest work experience indicates that the candidate now becomes a group leader in charge of new IP initiation. Our proposed model identifies career development and matches with a senior position that requires leadership and teamwork. For a talent like this, recommending job positions as ordinary art designers may not be the best option.

It is worth noticing that the purpose of the case study is not to depreciate other person-job fit methods in order to elevate ours, Actually, the job positions present in Figure 3 are all reasonable matching in the sense of qualification. However, there are still differences in these "good" matching: whether they conform to the preference revealed in the career path. We believe that the ideal next job follows the direction of career advancement for talents who are self-motivated and seeking advancements and perfection in career development.

Conclusions

In this paper, we propose a new perspective of person-job fit to emphasize the preference of career path. We introduce three auxiliary tasks: career path identification, career path reconstruction, and career-job contrastive learning to investigate and extract the career path information and propose an effective architecture WEPJF that successfully fuse the extracted career path preference into the objective of person-job fit. Through experiments and case studies, we demonstrate that WEPJF is a competitive method and career path really matters in person-job fit. Given that employment is under the strike of a new wave of automation and AI technology, more and more people will rely on online recruitment platforms for the foreseeable future. How to provide job seekers with better job-matching services is becoming an increasingly urgent problem. In this paper, we alleviate this problem by utilizing career path information.

Acknowledgments

This work is supported by National Key R&D Program of China (No. 2022YFC3301900) and National Natural Science Foundation of China (NSFC Grant No. 62122089). We sincerely thank all reviewers for their valuable comments and suggestions, which are crucial for improving our work.

References

Bian, S.; Chen, X.; Zhao, W. X.; Zhou, K.; Hou, Y.; Song, Y.; Zhang, T.; and Wen, J.-R. 2020. Learning to Match Jobs with Resumes from Sparse Interaction Data using Multi-View Co-Teaching Network. In *CIKM*, 65–74.

Bian, S.; Zhao, W. X.; Song, Y.; Zhang, T.; and Wen, J.-R. 2019. Domain Adaptation for Person-Job Fit with Transferable Deep Global Match Network. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 4812–4822.

Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Fang, H.; Wang, S.; Zhou, M.; Ding, J.; and Xie, P. 2020. Cert: Contrastive self-supervised learning for language understanding. *arXiv preprint arXiv:2005.12766*.

Friedman, J. H. 2001. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189–1232.

Galton, F. 1886. Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15: 246–263.

Giorgi, J.; Nitski, O.; Wang, B.; and Bader, G. 2021. DeCLUTR: Deep Contrastive Learning for Unsupervised Textual Representations. In *ACL-IJCNLP*, 879–895.

Gururangan, S.; Marasović, A.; Swayamdipta, S.; Lo, K.; Beltagy, I.; Downey, D.; and Smith, N. A. 2020. Don't Stop Pretraining: Adapt Language Models to Domains and Tasks. In *ACL*, 8342–8360.

Jiang, J.; Ye, S.; Wang, W.; Xu, J.; and Luo, X. 2020. Learning Effective Representations for Person-Job Fit by Feature Fusion. In *CIKM*, 2549–2556.

Kingma, D. P.; and Ba, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Le, R.; Hu, W.; Song, Y.; Zhang, T.; Zhao, D.; and Yan, R. 2019. Towards Effective and Interpretable Person-Job Fitting. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 1883–1892.

Li, H.; Ge, Y.; Zhu, H.; Xiong, H.; and Zhao, H. 2017. Prospecting the career development of talents: A survival analysis perspective. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 917–925.

Lin, H.; Zhu, H.; Zuo, Y.; Zhu, C.; Wu, J.; and Xiong, H. 2017. Collaborative company profiling: Insights from an employee's perspective. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Luo, Y.; Zhang, H.; Wen, Y.; and Zhang, X. 2019. ResumeGAN: An Optimized Deep Representation Learning Framework for Talent-Job Fit via Adversarial Learning. In *CIKM*, 1101–1110.

Malinowski, J.; Keim, T.; Wendt, O.; and Weitzel, T. 2006. Matching people and jobs: A bilateral recommendation approach. In *HICSS'06*, volume 6, 137c–137c. IEEE.

Meng, Q.; Zhu, H.; Xiao, K.; Zhang, L.; and Xiong, H. 2019. A Hierarchical Career-Path-Aware Neural Network for Job Mobility Prediction. In *SIGKDD*, 14–24.

Pal, M. 2005. Random forest classifier for remote sensing classification. *International journal of remote sensing*, 26(1): 217–222.

Qin, C.; Zhu, H.; Xu, T.; Zhu, C.; Jiang, L.; Chen, E.; and Xiong, H. 2018. Enhancing person-job fit for talent recruitment: An ability-aware neural network approach. In *SIGIR*, 25–34.

Qin, C.; Zhu, H.; Zhu, C.; Xu, T.; Zhuang, F.; Ma, C.; Zhang, J.; and Xiong, H. 2019. DuerQuiz: A Personalized Question Recommender System for Intelligent Job Interview. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2165–2173.

Quinlan, J. R. 1987. Simplifying decision trees. *International journal of man-machine studies*, 27(3): 221–234.

Rish, I.; et al. 2001. An empirical study of the naive Bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, volume 3, 41–46.

Sun, Y.; Zhuang, F.; Zhu, H.; Song, X.; He, Q.; and Xiong, H. 2019. The impact of person-organization fit on talent management: A structure-aware convolutional neural network approach. In *SIGKDD*, 1625–1633.

Wang, Z.; Wei, W.; Xu, C.; Xu, J.; and Mao, X.-L. 2022. Person-job fit estimation from candidate profile and related recruitment history with co-attention neural networks. *Neurocomputing*, 501: 14–24.

Wu, X.; Xu, T.; Zhu, H.; Zhang, L.; Chen, E.; and Xiong, H. 2019. Trend-aware tensor factorization for job skill demand analysis. In *IJCAI*, 3891–3897. AAAI Press.

Xu, H.; Yu, Z.; Xiong, H.; Guo, B.; and Zhu, H. 2015. Learning career mobility and human activity patterns for job change analysis. In *2015 IEEE International Conference on Data Mining*, 1057–1062. IEEE.

Xu, T.; Zhu, H.; Zhu, C.; Li, P.; and Xiong, H. 2018. Measuring the popularity of job skills in recruitment market: A multi-criteria approach. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Yan, R.; Le, R.; Song, Y.; Zhang, T.; Zhang, X.; and Zhao, D. 2019. Interview Choice Reveals Your Preference on the Market: To Improve Job-Resume Matching through Profiling Memories. In *SIGKDD*, 914–922.

Yang, C.; Hou, Y.; Song, Y.; Zhang, T.; Wen, J.-R.; and Zhao, W. X. 2022. Modeling Two-Way Selection Preference for Person-Job Fit. In *Proceedings of the 16th ACM Conference on Recommender Systems*, 102–112.

Zhang, L.; Zhu, H.; Xu, T.; Zhu, C.; Qin, C.; Xiong, H.; and Chen, E. 2019. Large-Scale Talent Flow Forecast with Dynamic Latent Factor Model? In *The World Wide Web Conference*, 2312–2322.

Zhu, C.; Zhu, H.; Xiong, H.; Ma, C.; Xie, F.; Ding, P.; and Li, P. 2018. Person-job fit: Adapting the right talent for the right job with joint representation learning. *ACM Transactions on Management Information Systems (TMIS)*, 9(3): 1–17.