Job-SDF: A Multi-Granularity Dataset for Job Skill Demand Forecasting and Benchmarking

Xi Chen[1,*] **Chuan Qin**[2,*] **Chuyu Fang**[3], **Chao Wang**[1],
Chen Zhu[1], **Fuzhen Zhuang**[4,5], **Hengshu Zhu**[2,†] **Hui Xiong**[6,7,†]
[1]University of Science and Technology of China
[2]Computer Network Information Center, Chinese Academy of Sciences
[3]Baidu Inc. [4]Institute of Artificial Intelligence, Beihang University
[5]SKLSDE, School of Computer Science, Beihang University
[6]AI Thrust, The Hong Kong University of Science and Technology (Guangzhou)
[7]Department of Computer Science and Engineering,
The HongKong University of Science and Technology, Hong Kong SAR
`chenxi0401@mail.ustc.edu.cn,`
`{chuanqin0426, fangchuyu2022, chadwang2012, zc3930155}@gmail.com,`
`zhuangfuzhen@buaa.edu.cn, zhuhengshu@gmail.com, xionghui@ust.hk`

Abstract

In a rapidly evolving job market, skill demand forecasting is crucial as it enables policymakers and businesses to anticipate and adapt to changes, ensuring that workforce skills align with market needs, thereby enhancing productivity and competitiveness. Additionally, by identifying emerging skill requirements, it directs individuals towards relevant training and education opportunities, promoting continuous self-learning and development. However, the absence of comprehensive datasets presents a significant challenge, impeding research and the advancement of this field. To bridge this gap, we present *Job-SDF*, a dataset designed to train and benchmark job-skill demand forecasting models. Based on millions of public job advertisements collected from online recruitment platforms, this dataset encompasses monthly recruitment demand. Our dataset uniquely enables evaluating skill demand forecasting models at various granularities, including occupation, company, and regional levels. We benchmark a range of models on this dataset, evaluating their performance in standard scenarios, in predictions focused on lower value ranges, and in the presence of structural breaks, providing new insights for further research. Our code and dataset are publicly accessible via the `https://github.com/Job-SDF/benchmark`.

1 Introduction

Job skills encompass a range of abilities and competencies essential for performing tasks effectively in the workplace. These skills are broadly categorized into hard skills, such as technical and analytical abilities, and soft skills, including communication, teamwork, and adaptability [1]. Accurate forecasting of skill demand helps businesses and policymakers anticipate and address skill shortages and mismatches, and promotes skill development in high-demand areas, thereby supporting economic growth and stability [2, 3]. By identifying emerging skill requirements, individuals are directed towards relevant training and education opportunities, fostering continuous self-learning and development to stay competitive in the labor market [4–10]. By tracking skill demand trends,

---

[*]Equal contributions
[†]Corresponding Authors

38th Conference on Neural Information Processing Systems (NeurIPS 2024) Track on Datasets and Benchmarks.

employers gain deeper insight into recruits' priorities, enhancing person-job fit. [11–21]. Moreover, forecasting informs educational and training programs, ensuring that curricula align with the labor market's evolving needs [22–24].

Traditionally, skill demand analysis has relied on labor-intensive, survey-based methods limited to specific companies or occupations [25–27]. However, over the past decade, the rapid evolution of the internet has spurred the emergence of online recruitment platforms. These platforms have become the primary channels for job advertisements for numerous enterprises and organizations, accumulating vast amounts of job advertisement data. By leveraging this data, researchers have formulated skill demand forecasting as a time series task, utilizing various machine learning models such as autoregressive integrated moving average (ARIMA) [28], recurrent neural networks (RNNs) [29], and dynamic graph autoencoders (DyGAEs) [30], to predict future skill needs.

A major challenge impeding progress in this field is the lack of comprehensive and publicly accessible datasets. Existing studies do not provide open-source datasets, making it difficult for researchers to replicate experimental results and identify bottlenecks in current research. Furthermore, these datasets primarily focus on predicting skill demand variations across different occupations, with a notable lack of modeling and prediction at other granularities, such as companies or regions. This limitation hinders comprehensive comparisons between different models and impedes the exploration of potential downstream applications, such as human capital strategy development and regional policy formulation. Additionally, the significant variations in skill demand present further challenges. Existing studies, which rely on metrics such as Mean Squared Error (MSE), struggle to evaluate the performance of skill demand forecasting models for low-frequency skill terms. For instance, some emerging skills, such as large language models (LLMs), may initially show low demand but are crucial for the job market due to their potential to reshape existing occupations.

To this end, in this paper, we propose *Job-SDF*, a multi-granularity dataset designed for job skill demand forecasting research. Specifically, we collected millions of public job advertisements from online recruitment platforms. By extracting skill terms from job advertisement texts, we quantified the monthly skill demand at various granularities, including occupations, companies, and regions, to construct our dataset. This dataset encompasses 2,324 types of skills, 52 occupations, 521 companies, and 7 regions. We then use the Job-SDF dataset to benchmark a wide range of models for job skill demand forecasting tasks at various granularities. These models include statistical time series models (e.g., ARIMA [31]), deep learning-based methods such as RNN-based models [32, 30], Transformer-based models [33–37], MLP-based models [38, 39], as well as several state-of-the-art time-series forecasters [40, 41]. Performance is evaluated using regression metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Additionally, we use the Symmetric Mean Absolute Percentage Error (SMAPE) [42] and Relative Root Mean Squared Error (RRMSE) [43] metrics to account for the significantly varying nature of skill demand values, which is particularly useful for evaluating model performance in predicting lower value ranges. Moreover, we further investigate the impact of structural breaks in job skill demand time series data on model performance. The Job-SDF dataset, along with data loaders, example codes for different models, and evaluation setup, are publicly available in our GitHub repository: `https://github.com/Job-SDF/benchmark`.

2 Related Work

Skill demand forecasting can analyze how skills evolve over time, aiding experts in evaluating technological advancements [44–46], assessing wage inequality [47–49], and generating employment opportunities [50]. Furthermore, the skills required in the 21st-century workplace will differ significantly from those in previous eras [51]. Predicting skill demands benefits personal career transitions and corporate management strategies.

Recently, with the rapid accumulation of data and continuous advancements in technology, skill demand forecasting has demonstrated significant vitality. *Das et al.* proposed a method for dynamic task allocation to investigate the evolution of job task requirements over a decade of AI innovation across different salary levels [28]. Given the effectiveness of RNN in multi-step prediction, some researchers have integrated skill demand forecasting with RNN algorithms, achieving promising results [29, 32]. In addition, considering the supply-demand dynamics of the labor market concurrently, CHGH designed a joint prediction model based on the encoder-decoder architecture to achieve trend prediction for both skill supply and demand sides [30]. Moreover, to capture the dynamic information

2

of occupations, a pretraining-enhanced dynamic graph autoencoder has been developed to efficiently forecast skill demand at the occupational granularity [52].

However, the predominance of closed-source datasets has significantly elevated the barrier of researchers and constrained the pace of methodological advancements. While open-source skill-related datasets such as O*NET [53] and ESCO [54] provide skill taxonomies, they do not quantify skill demand. Furthermore, the current research data focuses either on macro-market skill demand predictions or analyses at a specific granularity, neglecting multi-level labor market analysis. This limitation generally hampers the transferability of the modeling approaches.

3 Job-SDF Dataset

The Job-SDF dataset is built from job advertisements collected on online recruitment platforms, encompassing dynamic job skill demand time series data at various granularities, recorded monthly. The dataset is CC BY-NC-SA 4.0 licensed, accessible via the URL `https://github.com/Job-SDF/benchmark`. We summarize the dataset construction process, task description, and dataset analysis below, with supplementary details provided in the Appendix F.

3.1 Data Collection and Processing

Job Advertisement Collection. We collected public job advertisements for 52 occupations from 521 companies on online recruitment platforms. We obtained unique records after removing identical job advertisements posted simultaneously by different companies on various platforms. Each record contains five types of information: (1) *Job Requirement*, which is a text segment that outlines the specific skills required of candidates applying for the job; (2) *Company*, which identifies the company that posted the job advertisement; (3) *Occupation*, which specifies the job advertisement's category. Our dataset encompasses 52 detailed occupations (L2-level), such as front-end development engineer and financial investment analyst. Additionally, these 52 occupations are grouped into 14 broader categories (L1-level); (4) *Region*, which indicates the primary geographic divisions in China where the job postings are located. These regions are classified based on their geographical orientation; (5) *Posting Time*, which records the date when the job was posted, including the year, month, and day.

Job Skill Extraction. After acquiring the job advertisement data, we utilized a Named Entity Recognition (NER) model, as referenced in [55–58], to explicitly extract skill requirements from the *Job Requirement* of each advertisement. Specifically, we first annotated a dataset for training the NER model by identifying skill terms within the job requirement texts. To achieve this, we devised a set of regular expressions tailored to the characteristics of skill descriptions and used these to match skill words in job advertisements. Subsequently, we merged all matched skill words to formulate a raw skill dictionary, including their corresponding frequencies across job advertisements. We then filtered out low-frequency words and manually annotated the raw skill dictionary to create a refined skill dictionary. Along this line, we excluded unreasonable skill words matched by the regular expressions that did not appear in the refined skill dictionary, establishing an initial correspondence between the *Job Requirement* and the skill requirements.

Based on this annotated data, we trained an NER model to extract required skills from the *Job Requirement* section for all job advertisements. Experts then aggregated the skills extracted by the NER model based on their meaning and content, grouping together those with similar meanings or repeated expressions. This process resulted in a skill dictionary $\mathcal{S}$ of 2,324 standardized skill words, mapping original skill word descriptions to standardized skill words. The skill dictionary was then used to filter and map the skill words extracted by the NER model, ultimately obtaining standardized skill requirements for each job requirement. These standardized requirements were added to the job advertisement data as a new field, *Skill Requirements*.

Job Skill Demand Estimation. Generally, the demand for different skills in the job market can be estimated by the volume of job advertisements listing these specific skills as requirements within a given time period [30]. Formally, given job advertisement data $\mathcal{P} = \{\mathcal{P}_1, ..., \mathcal{P}_t, ..., \mathcal{P}_T\}$, where each $\mathcal{P}_t$ represents the job advertisements posted at timestamp t, we use $D_{s,t} = \sum_{p \in \mathcal{P}_t} \mathbf{1}(s \in p)$ to estimate the demand for skill $s \in \mathcal{S}$ at time t. $s \in p$ indicates that job advertisement p requires skill s.

Along this line, we can calculate skill demand at various granularities, such as occupation and company levels. We define the sets of L1-level occupations, L2-level occupations, companies, and

3

(a) The skill demands under two occupations.

(b) Appying the Chow test to two skill demand time series.
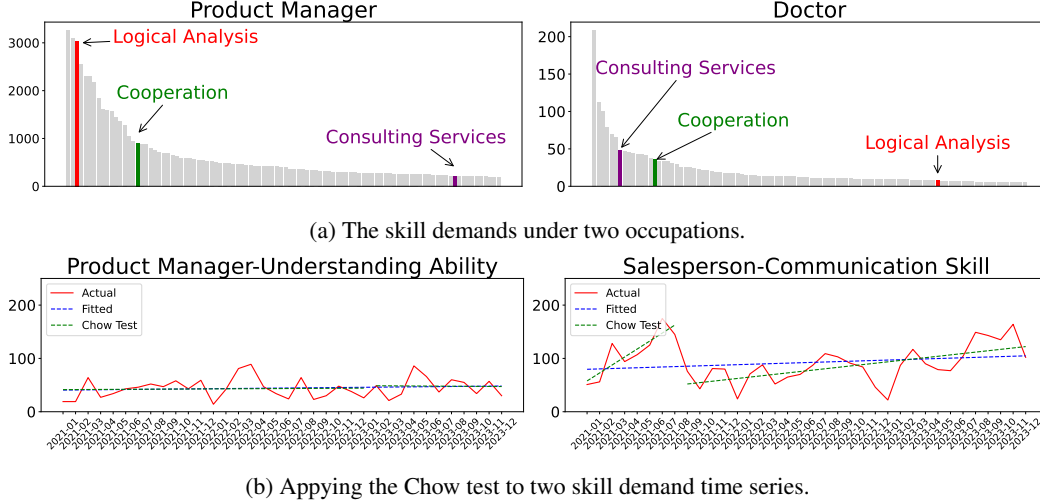
Figure 1: Data analysis on *Job-SDF*. (a) illustrates the long-tail phenomenon of skill demands under the product manager and doctor occupations. (b) illustrates the results under the Chow test for the absence (left) and presence (right) of structural breaks.

regions as $\mathcal{A}^{o_1}$, $\mathcal{A}^{o_2}$, $\mathcal{A}^c$, and $\mathcal{A}^r$, respectively. The demands for skill s at time t under granularity $i \in \{o_1, o_2, c, r\}$ is then defined as follows:

$$D_{s,t}^i = [D_{s,t,a^i}^i]_{a^i \in \mathcal{A}^i}, \quad D_{s,t,a^i}^i = \sum_{p \in \mathcal{P}_t} \mathbf{1}(s \in p) \cdot \mathbf{1}(a^i \in p), \tag{1}$$

where $a^i \in p$ represents a job advertisement p containing the attribute a^i under granularity i. Similarly, we can further define skill demands $D_{s,t}^{i,j,...,k}$ across multiple granularities $\{i, j, ..., k\}$ by calculating:

$$D_{s,t,\overline{a}}^{i,j,...,k} = \sum_{p \in \mathcal{P}_t} \mathbf{1}(s \in p) \cdot \mathbf{1}(a^i \in p \wedge a^j \in p \wedge ... \wedge a^k \in p), \tag{2}$$

where $\overline{a} = \{a^i, a^j, ..., a^k\}$, $a^i \in \mathcal{A}^i, a^j \in \mathcal{A}^j, ..., a^k \in \mathcal{A}^k$, and $D_{s,t}^{i,j,...,k} \in \mathbb{R}^{|\mathcal{A}^i||\mathcal{A}^j|...|\mathcal{A}^k|}$.

3.2 Job Skill Demand Forecasting Tasks

We study model performance through job skill demand forecasting tasks at different granularities, including single and multiple levels. The primary goal of these tasks is to predict future job skill demands based on historical time series data of various skills. Formally, we have:

Definition 1 (Job Skill Demand Forecasting) *Given a granularity or a set of granularities g and the observed job skill demand series from the previous K timestamps, i.e., $\{D_{:,t-K+1}^g, ..., D_{:,t}^g\}$, the goal of job skill demand forecasting is to learn a forecasting model $\mathcal{M}$ to predict the demand values for the next H timestamps, denoted by $\{\hat{D}_{:,t+1}^g, \ldots, \hat{D}_{:,t+H}^g\}$.*

Our dataset includes skill demand time series data for L1-level occupations, L2-level occupations, companies, regions, and their combinations. We follow a standard protocol [59] that categorizes all time-series data into training, validation, and test sets in chronological order with a ratio of 9:1:2. In the main text, we demonstrate results with K set to 6 months and consider H as 3 months to evaluate the performance of different forecasting models. More settings and results can be found in the Appendix D and our project repository. Based on the Job-SDF dataset, other researchers can easily adjust the parameters to suit their research objectives.

3.3 Dataset Analysis

Varying Nature of Skill Demand. The values of skill demand exhibit significant differences and generally follow a long-tail distribution. This indicates that, at a specific granularity, only a few skills have high demand, while a wide range of skills are required by a limited number of jobs. For instance,

4

Figure 1a presents the skill demands under the product manager and doctor occupations in December 2022. The results clearly demonstrate the varying nature of skill demand values. This suggests that relying solely on metrics like RMSE to evaluate forecasting models' performance may overlook the prediction accuracy for low-frequency skills.

Structural Break Phenomenon. As the labor market evolves, job skills that are not widely required today may become crucial in the future, while those currently in high demand may be supplanted by others. This dynamic can induce significant changes in the statistical properties of skill demand time series at various points in time. These changes may be reflected in the mean, variance, trend, or autocorrelation structure of the series. This phenomenon is known as structural breaks. A common method for detecting structural breaks is the Chow test, which evaluates whether there are significant differences in the regression coefficients across different periods [60]. Figure 1b illustrates the application of the Chow test in detecting structural breaks in various skill demand time series. The presence of structural breaks can impact the predictive accuracy of forecasting models. Further discussion will be provided in the experimental section.

Inter-Series Correlation. Intuitively, the proposed job skill demand forecasting tasks can be categorized as multivariate time-series forecasting tasks [61]. Figure 2 shows the absolute values of the Pearson correlation coefficients of different skill demand series for the backend development engineer, salesperson, and product manager occupations. We found that the time series data for some skills exhibit significant correlation within the same occupation (i.e., product design and market analysis in product manager), as well as for the same skills across different occupations (i.e., product design in backend development engineer and product manager). This demonstrates the necessity of considering all variables as inputs for job skill demand forecasting models, as it captures the interrelationships among variables, preventing the loss of critical information when variables are considered in isolation.
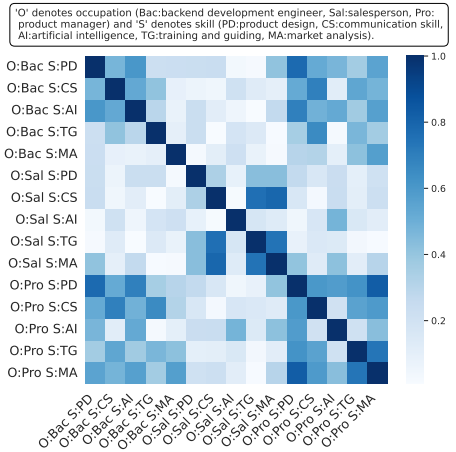
Figure 2: Pearson Correlation Coefficients.

Dataset Limitation. Recent studies [62] suggest that incorporating relationships between different variants can enhance the performance of multivariate time-series forecasting. However, due to the lack of prior knowledge, our Job-SDF dataset does not yet include a graph of relationships between skills, such as predecessor-successor relationships. Instead, we constructed a skill graph based on the co-occurrence of skills in job advertisements from the training data. This graph is included in our dataset. For detailed construction methods, please refer to the Appendix F.

4 Benchmark

4.1 Benchmark Models

We evaluated several SOTA time-series learning models using our proposed Job-SDF dataset. These models are categorized into six groups based on their underlying architectures: statistical time series models, RNN-based models, Transformer-based models, MLP-based models, Graph-based models, and Fourier-based models. The implementation details for each model are provided in the Appendix D, and the open-source model implementations are available on our GitHub repository.

Statistical Time Series Model. We first consider two statistical methodologies, namely *ARIMA* [63] and *Prophet* [64], both of which have been widely used in various contexts. The ARIMA model, which integrates differencing and moving averages within autoregression, has proven effective in forecasting occupational task demands [28]. *Prophet* decomposes time series data into trend, seasonality, and holiday components, allowing it to handle both linear and nonlinear trends with changepoints. However, these models often struggle to capture complex nonlinear relationships and exhibit suboptimal performance in large-scale data scenarios.

RNN-based Model. RNN-based methods are effective in capturing temporal state transitions through their recurrent structures, making them widely used in various time series forecasting tasks [65–

5

69]. Notably, *LSTM* have demonstrated their effectiveness in predicting changes in skill shares over time [29]. However, conventional RNNs often encounter performance degradation when handling excessively long look-back windows and forecast horizons. To address this challenge, *SegRNN* [32] introduces segment-wise iterations, which reduce the recurrence count within RNNs, thereby significantly enhancing performance in time series forecasting tasks.

Transformer-based Model. Recently, Transformer-based models [70] have gained widespread recognition in long-term time series forecasting due to their global modeling capabilities. Leveraging the attention mechanism, *Reformer* [37] introduces locally sensitive hashing to approximate attention by grouping similar queries. *Informer* [33] incorporates low-rank matrices in self-attention mechanisms to accelerate computation. *Autoformer*[34] employs block decomposition and autocorrelation mechanisms to more effectively capture the intrinsic features of time series data. *FedFormer* [36] utilizes DFT-based frequency-enhanced attention, obtaining attentive weights through the spectrums of queries and keys and calculating the weighted sum in the frequency domain. To address the challenges of non-stationary time series, the *Non-stationary Transformer (NStransformer)* [35] introduces a sequence stabilization module and proposes a de-stationary attention mechanism. Additionally, *PatchTST* [71] is a channel-independent patch time series transformer model that features patching and channel-independence as its key design elements.

MLP-based Model. Multiple Layer Projection (MLP) has been introduced in time series forecasting, demonstrating superior performance compared to transformer-based models in both accuracy and efficiency [38]. Specifically, *DLinear* [38] uses series decomposition as a pre-processing step before linear regression. *FreTS* [72] explores a novel approach by applying MLPs in the frequency domain for time series forecasting. *TSMixer* [39] employs MLPMixer blocks, segments input time series into fixed windows, and applies gated MLP transformations and permutations to enhance accuracy.

Graph-based Models. Graph Neural Networks (GNNs) can learn non-Euclidean relationships, making them effective for identifying associations in structured data and generating joint representations from different perspectives [73–76]. CHGH [30] uses an adaptive graph enhanced by skill co-occurrence relationships to link skill supply and demand sequences. This fusion of representations across views improves the performance of joint skill supply and demand prediction tasks. Pre-DyGAE [52] targets skill demand prediction from an occupational perspective. It builds an occupation-skill bipartite graph based on the skill demands of occupations and captures the dynamic changes in these relationships. This method allows for predicting both potential occupational skills and skill demands, leveraging a dynamic graph perspective.

Fourier-based Models. By utilizing Fourier projection, *FiLM* [40] not only captures long-term time dependencies but also effectively reduces noise in forecasting. To address the challenge of non-stationary time-series forecasting, *Koopa* [41] disentangles time-variant and time-invariant components from complex non-stationary series using a Fourier Filter and designs the Koopman Predictor to forecast dynamics.

4.2 Evaluation Metrics

To evaluate the performance of various benchmark models in job skill demand forecasting tasks, we selected two commonly used regression metrics: MAE and RMSE. MAE is calculated over H observations using the formula: $\frac{1}{H} \sum_{i=1}^{H} |y_i - \hat{y}_i|$, where y_i represents the ground truth value and $\hat{y}_i$ is the predicted value. RMSE is calculated as: $\sqrt{\frac{1}{H} \sum_{i=1}^{H} (y_i - \hat{y}_i)^2}$. Both MAE and RMSE are scale-dependent metrics, which makes them unsuitable for comparison across different granularities. Additionally, these metrics are less sensitive to prediction errors at lower skill demand values. Therefore, we additionally applied SMAPE [42] and RRMSE [77] to assess the performance of various forecasting models. SMAPE considers both the magnitude and direction of errors, making it suitable for comparing forecasts across different scales. RRMSE measures the square root of the average of the squared percentage errors.

$$SMAPE = \frac{2}{H} * \sum_{i=1}^{H} \frac{|y_i - \hat{y}_i|}{|y_i| + |\hat{y}_i|}, \quad RRMSE = \sqrt{\frac{\frac{1}{H} \sum_{i=1}^{H} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{H} (\hat{y}_i)^2}}. \tag{3}$$

6

Table 1: Performance comparison on MAE and RMSE.

Model	L1-Occupation		L2-Occupation		Region&L1-O		Region&L2-O		Company	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
ARIMA	20.27	256.89	6.46	115.79	3.98	58.65	1.31	27.42	1.31	38.88
Prophet	29.15	356.67	8.95	161.01	5.08	72.21	1.62	33.02	1.55	41.19
LSTM	19.05	194.67	7.09	116.36	3.92	51.59	1.29	23.31	1.35	26.47
SegRNN	12.28	108.28	5.01	68.83	3.14	34.26	1.05	15.96	1.01	16.03
CHGH	22.09	261.49	7.09	116.58	3.91	51.46	1.28	23.24	1.34	26.52
Pre-DyGAE	22.98	187.90	7.04	82.97	4.24	38.62	1.37	17.39	1.24	18.24
Transformer	22.06	215.09	7.58	118.21	4.01	52.04	1.35	23.44	1.26	24.99
Autoformer	23.06	186.76	8.22	100.02	6.45	57.77	2.41	24.10	3.31	38.55
Informer	22.21	205.24	7.43	117.38	3.88	50.13	1.30	23.07	1.26	24.92
Reformer	22.11	204.35	7.46	116.60	3.91	50.95	1.25	22.81	1.54	27.37
FEDformer	22.87	181.93	7.46	88.97	4.63	43.21	1.98	21.73	2.43	26.92
NStransformer	17.36	149.46	5.75	86.24	3.45	37.09	1.15	17.45	2.13	34.83
PatchTST	14.91	141.06	5.15	78.86	3.10	35.38	1.04	16.57	1.01	19.09
DLinear	16.61	154.88	5.44	81.61	3.24	36.67	1.07	16.79	1.05	18.85
TSMixer	21.34	192.85	8.14	106.65	5.81	62.14	5.95	68.26	13.96	144.96
FreTS	16.47	167.61	6.52	106.39	3.65	47.81	1.22	21.92	1.26	25.39
FiLM	12.95	117.17	5.08	65.65	3.24	29.90	1.14	14.01	1.17	15.87
Koopa	19.91	179.30	6.05	91.87	3.53	40.73	1.15	18.71	1.08	20.18

4.3 Benchmark Results

Overall Performance. In Table 1, we present the performance of various models evaluated using two metrics: MAE and RMSE. The following conclusions can be drawn: (1) The traditional statistical method, Prophet, demonstrates relatively poor predictive performance. This may be due to seasonal and holiday factors not being the primary influencers in skill demand prediction. (2) Most Transformer-based models, including Transformer, Autoformer, Informer, and Reformer, exhibit subpar overall predictive performance. This is likely because these models are designed to address long-range temporal dependencies, which are not well-suited for the current shorter time series context. (3) In contrast, PatchTST, unlike these Transformer-based models that perform point-wise modeling of time series, segments the time series into patches and inputs them into the Transformer. This allows the model to focus on more local information. A similar idea is also employed in the SegRNN. This strategy significantly enhances the performance of these models in predicting job skill demand. (4) The performance of different linear models on our dataset varies significantly. For instance, DLinear outperforms most Transformer-based models, while TSMixer performs poorly. This discrepancy may be due to the tendency of more complex MLP-based models to overfit our dataset. (5) CHGH and Pre-DyGAE exhibit poor performance in the separate skill demand forecasting scenario, likely due to a mismatch between their model design and the context of our dataset. Specifically, CHGH relies on sequential data from the supply side of skills, which is lacking in our dataset. Conversely, Pre-DyGAE focuses more on predicting whether a skill will be required by an occupation in the future. (6) Finally, FiLM achieved the best performance in most cases, demonstrating the robustness of the denoising-based model.

Low-Demand Skill Prediction Performance. Considering the varying nature of skill demand values, we further employed SMAPE and RRMSE metrics to focus on the predictive performance of different models for low-demand skills. As shown in Table 2, the experimental results indicate the following: (1) PatchTST achieved the best SMAPE performance in most cases, validating its ability to more accurately predict the trends of low-demand skills. (2) Based on scale-independent metrics, we can compare the performance of models at different granularities. It can be observed that RRMSE exhibits a significant trend of variation across different granularities; specifically, as the granularity becomes finer, the RRMSE performance deteriorates. This indicates that predicting skill demand at finer granularities is more challenging. Additionally, FiLM shows the least variation across multiple granularities, further validating its ability to provide stable and reliable predictions under varying granularities and demand value ranges. (3) Although Koopa performs averagely on MAE and RMSE metrics, it excels in predicting low-demand skills, particularly in terms of SMAPE. Similarly, NStransformer also performs well in scenarios focusing on low-demand skill predictions. This success can be attributed to both methods being designed to handle non-stationary time series. They effectively filter noise from historical sequences and restore intrinsic non-stationary information

7

Table 2: Performance comparison on SMAPE and RRMSE.

Model	L1-Occupation (%)		L2-Occupation (%)		Region&L1-O (%)		Region&L2-O (%)		Company (%)	
	SMAPE	RRMSE	SMAPE	RRMSE	SMAPE	RRMSE	SMAPE	RRMSE	SMAPE	RRMSE
ARIMA	35.72	47.89	25.00	58.87	23.86	58.07	13.58	73.57	20.17	147.94
Prophet	41.22	67.78	28.35	88.47	26.75	71.60	15.07	93.04	22.31	167.77
LSTM	41.38	57.90	32.85	83.70	31.58	68.40	22.93	87.36	30.26	174.40
SegRNN	39.81	37.58	33.35	50.53	35.30	48.53	23.84	61.90	33.07	86.27
CHGH	40.27	66.05	29.60	84.10	28.11	68.42	17.42	87.45	26.72	176.70
PreDyGAE	49.87	83.67	60.54	83.60	59.32	66.56	72.67	98.09	26.21	145.73
Transformer	55.59	64.25	44.23	84.27	31.15	76.16	33.04	86.87	27.61	164.36
Autoformer	70.28	53.75	74.37	63.40	90.14	65.57	91.51	74.46	107.05	99.60
Informer	56.85	58.18	44.04	88.72	34.75	69.59	29.29	90.15	32.41	164.37
Reformer	56.58	61.35	40.58	83.70	32.21	72.87	20.86	90.85	45.25	169.87
FEDformer	69.30	54.03	69.29	60.00	73.17	52.69	81.73	70.06	94.19	97.97
NStransformer	38.11	47.19	26.30	60.73	24.98	48.89	14.55	63.29	24.20	100.78
PatchTST	34.70	51.17	24.52	58.80	25.15	44.96	13.50	67.48	19.89	115.34
DLinear	41.84	52.89	34.35	60.22	33.47	51.05	25.77	64.65	30.71	108.66
TSMixer	56.59	61.17	72.29	99.35	82.48	87.29	120.85	96.49	155.20	102.14
FreTS	39.76	54.42	30.18	80.44	28.58	66.11	17.62	85.04	27.24	174.56
FiLM	39.51	37.55	29.65	43.86	28.79	37.66	17.24	47.75	25.72	76.92
Koopa	37.84	58.30	25.72	65.34	24.41	57.81	13.98	74.00	20.43	123.96

Table 3: Performance comparison on data with structural breaks on MAE and RMSE.

Model	L1-Occupation		L2-Occupation		Region&L1-O		Region&L2-O		Company	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
LSTM	87.30	554.46	57.95	400.22	18.99	149.53	7.91	52.38	24.40	159.02
SegRNN	61.92	390.54	43.97	276.57	15.85	114.04	6.56	37.84	17.98	112.13
CHGH	94.30	629.32	58.06	401.45	19.00	149.75	7.90	52.50	24.37	159.44
PreGyGAE	78.35	493.83	48.69	336.15	17.49	136.66	7.31	38.88	19.76	164.43
Transformer	98.66	580.58	61.73	404.17	19.37	151.12	8.45	55.46	22.41	152.27
Autoformer	107.22	533.06	67.66	350.97	26.84	156.50	12.19	63.04	44.10	208.96
Informer	98.89	570.35	59.95	402.75	19.03	146.91	7.72	49.15	22.37	151.87
Reformer	98.14	569.83	60.71	401.21	19.25	149.91	7.52	49.10	25.65	160.69
FEDformer	105.43	532.24	62.10	325.10	20.49	128.45	10.37	55.47	34.09	155.28
NStransformer	82.43	462.24	49.30	318.44	16.59	119.91	6.85	37.56	40.05	196.03
PatchTST	77.44	474.86	45.02	303.76	14.88	111.01	6.56	38.60	18.03	127.72
DLinear	81.17	485.25	46.67	307.34	15.94	118.94	6.50	37.72	18.18	124.32
TSMixer	107.47	614.93	83.60	479.39	29.99	187.08	25.83	190.29	155.10	766.58
FreTS	82.45	537.12	56.54	393.38	18.55	148.33	7.88	52.87	24.21	160.01
FiLM	62.86	404.82	42.63	260.99	14.31	101.23	6.37	32.28	18.78	110.65
Koopa	91.26	516.75	50.44	324.15	17.43	128.39	7.07	41.29	19.04	133.26

into time-dependent relationships, making them more adept at handling the fluctuating nature of low-demand skill time series data.

Performance on Skill Demand Series with Structural Breaks. As described in Section 3.3, in the dynamically changing job market, skill demand time series data exhibit structural breaks. To assess the impact of this phenomenon on different models in the skill demand forecasting task, we used the Chow test to detect structural breaks in the skill demand time series. The corresponding predictive performance of different models is presented in Tables 3 and 4. We observe the following phenomena: (1) Compared to the predictive performance on the full dataset, the performance on time series data with structural breaks is significantly worse. This finding underscores the complexity and unpredictability of skill trends that experience structural breaks. (2) FiLM has achieved results close to the overall skill demand prediction in terms of SMAPE and RRMSE metrics. This validates that FiLM can effectively mitigate the disruptive impact of structural breaks on skill demand forecasting. (3) Furthermore, while the overall predictive performance of skill demand forecasting at both the Region&L2-O and Company granularity levels is similar, significant differences emerge when forecasting skills experiencing structural breaks. This suggests that skills undergoing structural breaks display more predictable patterns at the Region&L2-O granularity level compared to the Company level, making them relatively easier to forecast.

8

Table 4: Performance comparison on data with structural breaks on RRMSE and SMAPE.

Model	L1-Occupation (%)		L2-Occupation (%)		Region&L1-O (%)		Region&L2-O (%)		Company (%)	
	SMAPE	RRMSE	SMAPE	RRMSE	SMAPE	RRMSE	SMAPE	RRMSE	SMAPE	RRMSE
LSTM	43.78	58.05	48.93	84.46	46.64	78.31	42.03	58.48	68.38	187.30
SegRNN	39.22	37.80	43.09	51.14	45.17	54.31	39.41	41.40	57.45	89.65
CHGH	44.91	66.31	48.90	84.87	45.43	78.32	39.79	58.89	68.36	189.91
PreDyGAE	52.35	47.15	56.56	59.31	52.06	61.22	44.13	42.31	70.26	106.88
Transformer	50.01	64.47	53.10	84.95	46.50	86.56	47.67	61.23	64.92	177.43
Autoformer	63.46	54.08	68.62	64.14	87.93	68.97	88.95	63.85	115.00	100.60
Informer	51.11	58.40	51.89	89.70	47.81	80.86	44.90	57.55	65.11	177.16
Reformer	50.79	61.59	51.51	84.53	46.86	84.15	40.81	58.59	72.36	181.36
FEDformer	62.83	54.37	64.37	60.84	72.24	58.55	80.03	54.29	103.27	100.65
NStransformer	45.36	47.46	47.63	61.85	43.04	57.60	36.72	39.72	170.57	113.87
PatchTST	40.89	51.48	43.26	59.69	41.51	51.85	34.74	43.12	55.26	122.56
DLinear	43.14	53.20	45.25	61.13	45.26	58.80	41.15	41.71	57.65	115.24
TSMixer	54.31	61.31	76.08	99.84	85.12	95.81	117.39	93.66	160.55	102.23
FreTS	42.44	54.59	48.24	81.17	45.39	75.43	39.85	57.83	68.39	187.94
FiLM	38.96	37.82	44.23	44.52	44.95	43.06	40.05	30.80	56.37	80.77
Koopa	46.45	58.59	47.13	66.28	42.60	66.20	36.24	47.48	58.98	131.77

5 Conclusion

In this work, we introduced Job-SDF, a dataset designed for training and benchmarking job-skill demand forecasting models. Compiled from millions of public job advertisements collected from online recruitment platforms, this dataset includes monthly recruitment demand for 2,324 types of skills across 52 occupations, 521 companies, and 7 regions. Using this dataset, we validated a wide range of time-series forecasting approaches, including statistical models, RNN-based models, Transformer-based models, MLP-based models, Graph-based models, and Fourier-based models. Furthermore, we conducted extensive experiments to compare the performance of various methods in predicting skill demand at different granularities. We hope that Job-SDF will facilitate further research in this field.

Acknowledgements

This work was supported in part by the National Key R&D Program of China (Grant No.2023YFF0725001), in part by the National Natural Science Foundation of China (Grant No.92370204), in part by the guangdong Basic and Applied Basic Research Foundation (Grant No.2023B1515120057), in part by Guangzhou-HKUST (GZ) Joint Funding Program (Grant No.2023A03J0008), Education Bureau of Guangzhou Municipality, in part by Nansha Postdoctoral Research Project, and in part by the National Natural Science Foundation of China (Grant No.62176014), the Fundamental Research Funds for the Central Universities.

References

[1] David Autor et al. The polarization of job opportunities in the us labor market: Implications for employment and earnings. *Center for American Progress and The Hamilton Project*, 6:11–19, 2010.

[2] James J Heckman, Jora Stixrud, and Sergio Urzua. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, 24(3):411–482, 2006.

[3] Chuan Qin, Hengshu Zhu, Dazhong Shen, Ying Sun, Kaichun Yao, Peng Wang, and Hui Xiong. Automatic skill-oriented question generation and recommendation for intelligent job interviews. *ACM Transactions on Information Systems*, 42(1):1–32, 2023.

[4] Marios Kokkodis and Panagiotis G Ipeirotis. Demand-aware career path recommendations: A reinforcement learning approach. *Management science*, 67(7):4362–4383, 2021.

[5] Rui Zha, Chuan Qin, Le Zhang, Dazhong Shen, Tong Xu, Hengshu Zhu, and Enhong Chen. Career mobility analysis with uncertainty-aware graph autoencoders: A job title transition perspective. *IEEE Transactions on Computational Social Systems*, 11(1):1205–1215, 2023.

9

[6] Rui Zha, Ying Sun, Chuan Qin, Le Zhang, Tong Xu, Hengshu Zhu, and Enhong Chen. Towards unified representation learning for career mobility analysis with trajectory hypergraph. *ACM Transactions on Information Systems*, 42(4):1–28, 2024.

[7] Xiaoshan Yu, Chuan Qin, Qi Zhang, Chen Zhu, Haiping Ma, Xingyi Zhang, and Hengshu Zhu. Disco: A hierarchical disentangled cognitive diagnosis framework for interpretable job recommendation. *arXiv preprint arXiv:2410.07671*, 2024.

[8] Ying Sun, Hengshu Zhu, Lu Wang, Le Zhang, and Hui Xiong. Large-scale online job search behaviors reveal labor market shifts amid covid-19. *Nature Cities*, 1(2):150–163, 2024.

[9] Le Dai, Yu Yin, Chuan Qin, Tong Xu, Xiangnan He, Enhong Chen, and Hui Xiong. Enterprise cooperation and competition analysis with a sign-oriented preference network. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 774–782, 2020.

[10] Hengshu Zhu, Hui Xiong, Fangshuang Tang, Qi Liu, Yong Ge, Enhong Chen, and Yanjie Fu. Days on market: Measuring liquidity in real estate markets. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 393–402, 2016.

[11] Chuan Qin, Hengshu Zhu, Tong Xu, Chen Zhu, Liang Jiang, Enhong Chen, and Hui Xiong. Enhancing person-job fit for talent recruitment: An ability-aware neural network approach. In *The 41st international ACM SIGIR conference on research & development in information retrieval*, pages 25–34, 2018.

[12] Chuan Qin, Hengshu Zhu, Chen Zhu, Tong Xu, Fuzhen Zhuang, Chao Ma, Jingshuai Zhang, and Hui Xiong. Duerquiz: A personalized question recommender system for intelligent job interview. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2165–2173, 2019.

[13] Chuan Qin, Hengshu Zhu, Tong Xu, Chen Zhu, Chao Ma, Enhong Chen, and Hui Xiong. An enhanced neural network approach to person-job fit in talent recruitment. *ACM Transactions on Information Systems (TOIS)*, 38(2):1–33, 2020.

[14] Kaichun Yao, Jingshuai Zhang, Chuan Qin, Peng Wang, Hengshu Zhu, and Hui Xiong. Knowledge enhanced person-job fit for talent recruitment. In *2022 IEEE 38th International Conference on Data Engineering (ICDE)*, pages 3467–3480. IEEE, 2022.

[15] Likang Wu, Zhaopeng Qiu, Zhi Zheng, Hengshu Zhu, and Enhong Chen. Exploring large language model for graph data understanding in online job recommendations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 9178–9186, 2024.

[16] Shuqing Bian, Xu Chen, Wayne Xin Zhao, Kun Zhou, Yupeng Hou, Yang Song, Tao Zhang, and Ji-Rong Wen. Learning to match jobs with resumes from sparse interaction data using multi-view co-teaching network. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pages 65–74, 2020.

[17] Yong Luo, Huaizheng Zhang, Yonggang Wen, and Xinwen Zhang. Resumegan: an optimized deep representation learning framework for talent-job fit via adversarial learning. In *Proceedings of the 28th ACM international conference on information and knowledge management*, pages 1101–1110, 2019.

[18] Yoosof Mashayekhi, Nan Li, Bo Kang, Jefrey Lijffijt, and Tijl De Bie. A challenge-based survey of e-recruitment recommendation systems. *ACM Computing Surveys*, 56(10):1–33, 2024.

[19] Yong Luo, Huaizheng Zhang, Yongjie Wang, Yonggang Wen, and Xinwen Zhang. Resumenet: A learning-based framework for automatic resume quality assessment. In *2018 IEEE International Conference on Data Mining (ICDM)*, pages 307–316. IEEE, 2018.

[20] Dazhong Shen, Chuan Qin, Hengshu Zhu, Tong Xu, Enhong Chen, and Hui Xiong. Joint representation learning with relation-enhanced topic models for intelligent job interview assessment. *ACM Transactions on Information Systems (TOIS)*, 40(1):1–36, 2021.

10

[21] Feihu Jiang, Chuan Qin, Kaichun Yao, Chuyu Fang, Fuzhen Zhuang, Hengshu Zhu, and Hui Xiong. Enhancing question answering for enterprise knowledge bases using large language models. In *International Conference on Database Systems for Advanced Applications*, pages 273–290. Springer, 2024.

[22] David J Deming. The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, 132(4):1593–1640, 2017.

[23] Liyi Chen, Chuan Qin, Ying Sun, Xin Song, Tong Xu, Hengshu Zhu, and Hui Xiong. Collaboration-aware hybrid learning for knowledge development prediction. In *Proceedings of the ACM on Web Conference 2024*, pages 3976–3985, 2024.

[24] Yunfei Zhang, Chuan Qin, Dazhong Shen, Haiping Ma, Le Zhang, Xingyi Zhang, and Hengshu Zhu. Relicd: a reliable cognitive diagnosis framework with confidence awareness. In *2023 IEEE International Conference on Data Mining (ICDM)*, pages 858–867. IEEE, 2023.

[25] William Rasdorf, Joseph E Hummer, and Stephanie C Vereen. Data collection opportunities and challenges for skilled construction labor demand forecast modeling. *Public Works Management & Policy*, 21(1):28–52, 2016.

[26] Joshua Healy, Kostas Mavromaras, and Peter J Sloane. Adjusting to skill shortages in australian smes. *Applied Economics*, 47(24):2470–2487, 2015.

[27] Lutz Bellmann and Olaf Hübler. The skill shortage in german establishments before, during and after the great recession. *Jahrbücher für Nationalökonomie und Statistik*, 234(6):800–828, 2014.

[28] Subhro Das, Sebastian Steffen, Wyatt Clarke, Prabhat Reddy, Erik Brynjolfsson, and Martin Fleming. Learning occupational task-shares dynamics for the future of work. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, pages 36–42, 2020.

[29] Maysa Malfiza Garcia de Macedo, Wyatt Clarke, Eli Lucherini, Tyler Baldwin, Dilermando Queiroz Neto, Rogerio Abreu de Paula, and Subhro Das. Practical skills demand forecasting via representation learning of temporal dynamics. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, pages 285–294, 2022.

[30] Wenshuo Chao, Zhaopeng Qiu, Likang Wu, Zhuoning Guo, Zhi Zheng, Hengshu Zhu, and Hao Liu. A cross-view hierarchical graph learning hypernetwork for skill demand-supply joint prediction. In *AAAI*, 2024.

[31] Adebiyi A Ariyo, Adewumi O Adewumi, and Charles K Ayo. Stock price prediction using the arima model. In *2014 UKSim-AMSS 16th international conference on computer modelling and simulation*, pages 106–112. IEEE, 2014.

[32] Shengsheng Lin, Weiwei Lin, Wentai Wu, Feiyu Zhao, Ruichao Mo, and Haotong Zhang. Segrnn: Segment recurrent neural network for long-term time series forecasting. *arXiv preprint arXiv:2308.11200*, 2023.

[33] Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, pages 11106–11115, 2021.

[34] Haixu Wu, Jiehui Xu, Jianmin Wang, and Mingsheng Long. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. *Advances in neural information processing systems*, 34:22419–22430, 2021.

[35] Yong Liu, Haixu Wu, Jianmin Wang, and Mingsheng Long. Non-stationary transformers: Exploring the stationarity in time series forecasting. *Advances in Neural Information Processing Systems*, 35:9881–9893, 2022.

[36] Tian Zhou, Ziqing Ma, Qingsong Wen, Xue Wang, Liang Sun, and Rong Jin. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting. In *International conference on machine learning*, pages 27268–27286. PMLR, 2022.

11

[37] Nikita Kitaev, Łukasz Kaiser, and Anselm Levskaya. Reformer: The efficient transformer. *arXiv preprint arXiv:2001.04451*, 2020.

[38] Ailing Zeng, Muxi Chen, Lei Zhang, and Qiang Xu. Are transformers effective for time series forecasting? In *Proceedings of the AAAI conference on artificial intelligence*, pages 11121–11128, 2023.

[39] Shiyu Wang, Haixu Wu, Xiaoming Shi, Tengge Hu, Huakun Luo, Lintao Ma, James Y Zhang, and Jun Zhou. Timemixer: Decomposable multiscale mixing for time series forecasting. *arXiv preprint arXiv:2405.14616*, 2024.

[40] Tian Zhou, Ziqing Ma, Qingsong Wen, Liang Sun, Tao Yao, Wotao Yin, Rong Jin, et al. Film: Frequency improved legendre memory model for long-term time series forecasting. *Advances in Neural Information Processing Systems*, 35:12677–12690, 2022.

[41] Yong Liu, Chenyu Li, Jianmin Wang, and Mingsheng Long. Koopa: Learning non-stationary time series dynamics with koopman predictors. *Advances in Neural Information Processing Systems*, 36, 2024.

[42] Saba Sareminia. A support vector based hybrid forecasting model for chaotic time series: Spare part consumption prediction. *Neural Processing Letters*, 55(3):2825–2841, 2023.

[43] Chao Chen, Jamie Twycross, and Jonathan M Garibaldi. A new accuracy measure based on bounded relative error for time series forecasting. *PloS one*, 12(3):e0174202, 2017.

[44] Alan Manning. We can work it out: the impact of technological change on the demand for low-skill workers. *Scottish Journal of Political Economy*, 51(5):581–608, 2004.

[45] John M Abowd, John C Haltiwanger, Julia Lane, Kevin L McKinney, and Kristin Sandusky. Technology and the demand for skill: an analysis of within and between firm differences, 2007.

[46] Chuan Qin, Le Zhang, Yihang Cheng, Rui Zha, Dazhong Shen, Qi Zhang, Xi Chen, Ying Sun, Chen Zhu, Hengshu Zhu, et al. A comprehensive survey of artificial intelligence techniques for talent analytics. *arXiv preprint arXiv:2307.03195*, 2023.

[47] Chinhui Juhn. Wage inequality and demand for skill: evidence from five decades. *ILR Review*, 52(3):424–443, 1999.

[48] Jong-Wha Lee and Dainn Wie. Technological change, skill demand, and wage inequality: Evidence from indonesia. *World Development*, 67:238–250, 2015.

[49] Ying Sun, Fuzhen Zhuang, Hengshu Zhu, Qi Zhang, Qing He, and Hui Xiong. Market-oriented job skill valuation with cooperative composition neural network. *Nature communications*, 12(1):1992, 2021.

[50] George S Benson and Edward E Lawler III. Raising skill demand: Generating good jobs. *Transforming the US Workforce Development System. Urbana: Labor and Employment Relations Association*, 2011.

[51] Margaret Hilton. *Research on future skill demands: A workshop summary*. National Academies Press, 2008.

[52] Xi Chen, Chuan Qin, Zhigaoyuan Wang, Yihang Cheng, Chao Wang, Hengshu Zhu, and Hui Xiong. Pre-dygae: Pre-training enhanced dynamic graph autoencoder for occupational skill demand forecasting. In *Proceedings of the 33th International Joint Conference on Artificial Intelligence*, 2024.

[53] Manuel Cifuentes, Jon Boyer, David A Lombardi, and Laura Punnett. Use of o* net as a job exposure matrix: a literature review. *American journal of industrial medicine*, 53(9):898–914, 2010.

[54] Filippo Chiarello, Gualtiero Fantoni, Terence Hogarth, Vito Giordano, Liga Baltina, and Irene Spada. Towards esco 4.0–is the european classification of skills in line with industry 4.0? a text mining approach. *Technological Forecasting and Social Change*, 173:121177, 2021.

12

[55] Chuan Qin, Kaichun Yao, Hengshu Zhu, Tong Xu, Dazhong Shen, Enhong Chen, and Hui Xiong. Towards automatic job description generation with capability-aware neural networks. *IEEE Transactions on Knowledge and Data Engineering*, 35(5):5341–5355, 2022.

[56] Chuyu Fang, Chuan Qin, Qi Zhang, Kaichun Yao, Jingshuai Zhang, Hengshu Zhu, Fuzhen Zhuang, and Hui Xiong. Recruitpro: A pretrained language model with skill-aware prompt learning for intelligent recruitment. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 3991–4002, 2023.

[57] Kaichun Yao, Chuan Qin, Hengshu Zhu, Chao Ma, Jingshuai Zhang, Yi Du, and Hui Xiong. An interactive neural network approach to keyphrase extraction in talent recruitment. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, pages 2383–2393, 2021.

[58] Feihu Jiang, Chuan Qin, Jingshuai Zhang, Kaichun Yao, Xi Chen, Dazhong Shen, Chen Zhu, Hengshu Zhu, and Hui Xiong. Towards efficient resume understanding: A multi-granularity multi-modal pre-training approach. *arXiv preprint arXiv:2404.13067*, 2024.

[59] Zhiyuan Wang, Xovee Xu, Weifeng Zhang, Goce Trajcevski, Ting Zhong, and Fan Zhou. Learning latent seasonal-trend representations for time series forecasting. *Advances in Neural Information Processing Systems*, 35:38775–38787, 2022.

[60] Gregory C Chow. Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society*, pages 591–605, 1960.

[61] Defu Cao, Yujing Wang, Juanyong Duan, Ce Zhang, Xia Zhu, Congrui Huang, Yunhai Tong, Bixiong Xu, Jing Bai, Jie Tong, et al. Spectral temporal graph neural network for multivariate time-series forecasting. *Advances in neural information processing systems*, 33:17766–17778, 2020.

[62] Ting Guo, Feng Hou, Yan Pang, Xiaoyun Jia, Zhongwei Wang, and Ruili Wang. Learning and integration of adaptive hybrid graph structures for multivariate time series forecasting. *Information Sciences*, 648:119560, 2023.

[63] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[64] Sean J Taylor and Benjamin Letham. Forecasting at scale. *The American Statistician*, 72(1):37–45, 2018.

[65] Chen Zhu, Hengshu Zhu, Hui Xiong, Pengliang Ding, and Fang Xie. Recruitment market trend analysis with sequential latent variable models. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 383–392, 2016.

[66] Qi Zhang, Tong Xu, Hengshu Zhu, Lifu Zhang, Hui Xiong, Enhong Chen, and Qi Liu. After-shock detection with multi-scale description based neural network. In *2019 IEEE International Conference on Data Mining (ICDM)*, pages 886–895. IEEE, 2019.

[67] Qi Zhang, Hengshu Zhu, Ying Sun, Hao Liu, Fuzhen Zhuang, and Hui Xiong. Talent demand forecasting with attentive neural sequential model. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 3906–3916, 2021.

[68] Qi Zhang, Hengshu Zhu, Qi Liu, Enhong Chen, and Hui Xiong. Exploiting real-time search engine queries for earthquake detection: A summary of results. *ACM Transactions on Information Systems (TOIS)*, 39(3):1–32, 2021.

[69] Dazhong Shen, Qi Zhang, Tong Xu, Hengshu Zhu, Wenjia Zhao, Zikai Yin, Peilun Zhou, Lihua Fang, Enhong Chen, and Hui Xiong. A machine learning-enhanced robust p-phase picker for real-time seismic monitoring. *arXiv preprint arXiv:1911.09275*, 2019.

[70] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

13

[71] Yuqi Nie, Nam H Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. *arXiv preprint arXiv:2211.14730*, 2022.

[72] Kun Yi, Qi Zhang, Wei Fan, Shoujin Wang, Pengyang Wang, Hui He, Ning An, Defu Lian, Longbing Cao, and Zhendong Niu. Frequency-domain mlps are more effective learners in time series forecasting. *Advances in Neural Information Processing Systems*, 36, 2024.

[73] Liyi Chen, Zhi Li, Tong Xu, Han Wu, Zhefeng Wang, Nicholas Jing Yuan, and Enhong Chen. Multi-modal siamese network for entity alignment. In *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*, pages 118–126, 2022.

[74] Liyi Chen, Zhi Li, Weidong He, Gong Cheng, Tong Xu, Nicholas Jing Yuan, and Enhong Chen. Entity summarization via exploiting description complementarity and salience. *IEEE Transactions on Neural Networks and Learning Systems*, 34(11):8297–8309, 2023.

[75] Dazhong Shen, Chuan Qin, Chao Wang, Hengshu Zhu, Enhong Chen, and Hui Xiong. Regularizing variational autoencoder with diversity and uncertainty awareness. *arXiv preprint arXiv:2110.12381*, 2021.

[76] Miao Chen, Chao Wang, Chuan Qin, Tong Xu, Jianhui Ma, Enhong Chen, and Hui Xiong. A trend-aware investment target recommendation system with heterogeneous graph. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2021.

[77] Keerti Rawal and Aijaz Ahmad. Mining latent patterns with multi-scale decomposition for electricity demand and price forecasting using modified deep graph convolutional neural networks. *Sustainable Energy, Grids and Networks*, page 101436, 2024.

[78] Benedek Rozemberczki, Paul Scherer, Yixuan He, George Panagopoulos, Alexander Riedel, Maria Astefanoaei, Oliver Kiss, Ferenc Beres, Guzman Lopez, Nicolas Collignon, and Rik Sarkar. PyTorch Geometric Temporal: Spatiotemporal Signal Processing with Neural Machine Learning Models. In *Proceedings of the 30th ACM International Conference on Information and Knowledge Management*, page 4564–4573, 2021.

[79] Haixu Wu, Tengge Hu, Yong Liu, Hang Zhou, Jianmin Wang, and Mingsheng Long. Timesnet: Temporal 2d-variation modeling for general time series analysis. In *International Conference on Learning Representations*, 2023.

[80] Aldo Pareja, Giacomo Domeniconi, Jie Chen, Tengfei Ma, Toyotaro Suzumura, Hiroki Kanezashi, Tim Kaler, Tao Schardl, and Charles Leiserson. Evolvegcn: Evolving graph convolutional networks for dynamic graphs. In *Proceedings of the AAAI conference on artificial intelligence*, pages 5363–5370, 2020.

[81] Youngjoo Seo, Michaël Defferrard, Pierre Vandergheynst, and Xavier Bresson. Structured sequence modeling with graph convolutional recurrent networks. In *Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December 13-16, 2018, Proceedings, Part I 25*, pages 362–373. Springer, 2018.

[82] Ling Zhao, Yujiao Song, Chao Zhang, Yu Liu, Pu Wang, Tao Lin, Min Deng, and Haifeng Li. T-GCN: A Temporal Graph ConvolutionalNetwork for Traffic Prediction. *IEEE Transactions on Intelligent Transportation Systems*, pages 3848–3858, 2020.

[83] Jinyin Chen, Xueke Wang, and Xuanheng Xu. Gc-lstm: Graph convolution embedded lstm for dynamic network link prediction. *Applied Intelligence*, pages 1–16, 2022.

[84] Aynaz Taheri, Kevin Gimpel, and Tanya Berger-Wolf. Learning to represent the evolution of dynamic graphs with recurrent models. In *Companion Proceedings of The 2019 World Wide Web Conference*, WWW '19, New York, NY, USA, 2019. Association for Computing Machinery.

14

A   Ethics Statement

In accordance with the Code of Ethics as outlined by NeurIPS [1], we present the following discussion on the ethical considerations and potential societal impact of our research.

Privacy.   Our data is collected from online platforms, encompassing job advertisements, which is public by various companies. This collection method ensures that all information utilized is already available to the public, having been disclosed by companies for the purpose of recruitment. Consequently, our analysis inherently respects privacy norms, as it involves no proprietary or confidential data, relying solely on publicly shared statistics and general information.

Consent.   We have extensively collected job advertisement data from online platforms where the information is openly and transparently available. Our analysis is based on the statistical data derived from these job advertisements, rather than the original textual information, significantly minimizing the risk of privacy breaches. Therefore, the use of this data is permissible.

Deprecated Datasets.   Not applicable.

Copyright and Fair Use.   The library used in our research is publicly available and distributed under the MIT License. We list the used assets along with their licenses as follows:

- PyTorch Geometric Temporal [78] consists of various dynamic and temporal geometric deep learning, embedding, and spatio-temporal regression methods from a variety of published research papers. This library is open-source and available at `https://github.com/benedekrozemberczki/pytorch_geometric_temporal`, which is publicly available and distributed under the MIT License.
- Time Series Library [79] is an open-source library for deep learning researchers, especially for deep time series analysis. This library is available at `https://github.com/thuml/Time-Series-Library`, which is publicly available and distributed under the MIT License.

Representative Evaluation Practice.   Job-SDF has extensively covered skill demand data at multiple granular levels, including regions, occupations, and companies. The detailed nature of the data proves invaluable in advancing tasks related to skill demand prediction. It significantly assists job seekers in identifying employment opportunities, employers in recruiting suitable talent, and educational institutions in tailoring their curricula to meet market demands. Moreover, this comprehensive dataset not only facilitates the analysis of skill trends but also serves as a platform for verifying time-series algorithms within a specialized application domain.

Safety.   Our research does not employ technologies that directly inflict harm on individuals.

Security.   Since the data sources are all publicly available, there is no risk of security incidents.

Discrimination.   Our data is strictly limited to market analysis and does not judge the superiority or inferiority of skills themselves, thus it does not involve any issues of discrimination.

Deception & Harassment.   We firmly believe that our data cannot be used for deceptive practices or harassment activities.

Human Rights.   Not applicable.

Bias and Fairness.   Since our data is collected from the public job postings of a subset of companies, there may be statistical biases in skill demands across different industries or regions. While it may not represent the entirety, it effectively captures the main trends in skill demands.

---

[1]https://neurips.cc/public/EthicsGuidelines

15

B Code and Data Availability

The dataset is CC BY-NC-SA 4.0 licensed, accessible via the URL `https://github.com/Job-SDF/benchmark`. We also provide a website for the project: `https://job-sdf.github.io/` with information on how to use the data and code.

C Computational Resource

Due to inherent design and size constraints of the models combined with varying data sizes at different granularities, the deployment environments for each model are distinct. The CHGH model, which requires over 80GB of memory, is exclusively deployed on CPU platforms to accommodate its substantial resource demands. In contrast, the PreDyGAE model operates solely on GPU infrastructure, leveraging the computational efficiencies of the NVIDIA A800 GPUs. For other models, deployment strategies are tailored according to the granularity of the data. Experiments at the labor market, regions, L1 occupations, L2 occupations, and Region & L1 occupations granularities are conducted on GPUs, capitalizing on the enhanced processing capabilities of these units for handling moderate data volumes. However, at the granularities of Region & L2 and company, where data volumes are significantly larger, deployment shifts to CPUs. Overall, the training time of different models are shown in Table 5.

Table 5: Training time (minute) of different models for job skill demand forecsting.

Model	Market	Region	L1-O	L2-O	R&L1-O	R&L2-O	Company
LSTM	0-0.5	0-0.5	0-0.5	0-0.5	0-0.5	37.7	39.0
SegRNN	0-0.5	0-0.5	0-0.5	0-0.5	0-0.5	342.8	458.2
CHGH	17.7	132.8	170.2	258.3	1300.3	490.6	6604.2
PreDyGAE	1-10	16.5	30.0	48.1	48.1	88.2	126.2
Transformer	0-0.5	0-0.5	0-0.5	0.5-1	0.5-1	128.2	166.5
Autoformer	0-0.5	0-0.5	0-0.5	0-0.5	0.5-1	304.3	325.0
Informer	0-0.5	0-0.5	0-0.5	0-0.5	0-0.5	133.8	171.7
Reformer	0-0.5	0-0.5	0-0.5	0-0.5	0-0.5	36.0	52.5
FEDformer	0-0.5	0-0.5	0-0.5	0-0.5	0-0.5	193.8	198.5
NStransformer	0-0.5	0.5-1	0.5-1	0-0.5	0.5-1	128.5	195.7
PatchTST	0-0.5	0-0.5	0-0.5	0-0.5	0.5-1	1202.8	2558.0
DLinear	0-0.5	0-0.5	0.5-1	0-0.5	0-0.5	20.0	39.1
TSMixer	0-0.5	0-0.5	0-0.5	0-0.5	0-0.5	24.0	97.0
FreTS	0-0.5	0-0.5	0-0.5	0-0.5	0-0.5	85.0	200.0
FiLM	0-0.5	0.5-1	0.5-1	1-10	1-10	598.0	1464.7
Koopa	0-0.5	0-0.5	0-0.5	0-0.5	0-0.5	38.3	68.5

D Additional Experimental Results

Due to page limitations in the main text, we have included additional experimental content in the appendix. First, we present the results of repeated trials of the benchmark models discussed in the main text in the first subsection. Subsequently, we focus on the performance of existing benchmark models in predicting demand for low-frequency skills. Further, we have constructed a co-occurrence relationship between skills as prior knowledge based on the training set and employed various Graph Neural Network (GNN)-based multivariate time series forecasting methods in the task of job skill demand forecasting, demonstrating promising results. Finally, considering that the skill demand proportion may be more meaningful than the skill demand volume in certain contexts, we have constructed a dataset for skill demand proportion and showcased the performance of benchmark models on this task.

D.1 Repeated Experiments on Job Skill Demand Forecasting

To demonstrate the robustness and reliability of our experimental results, we first repeated the experiments multiple times as described in the main text. Additionally, we extended our analysis to include experiments across the entire labor market and at various regional granularities.

16

Implementation Details. We utilized the Time-series-Library [2] to implement some of the models. The hyperparameters were uniformly set as follows: a learning rate of 0.0001, 20 epochs of training, a hidden layer dimension of 2048, the GELU activation function, and MSE loss as the loss function. Early stopping was employed to prevent overfitting by terminating training early when necessary. Data from the first 24 months were used for pre-training, and the model was fine-tuned using the next 6 months to capture trend changes. Finally, the model was used to infer skill demands for the last 6 months. All other hyperparameters were kept consistent with those in the original paper. To ensure the reliability of our findings, we repeated these experiments four times, using random seeds set to 0, 1, 2, and 3, respectively.

Overall Performance Table 6 displays the mean and standard deviation results of repeated experiments on the benchmark models for the job skill demand forecasting task as presented in the main text. Initially, we supplement the experimental results at the overall labor market level and regional granularity, where the RMSE averages over 1000. In cases of coarser granularity, due to the larger base of demand values, the prediction deviations are significant.

Performance on Skill Demand Series with Structural Breaks Table 7 presents the results of repeated experiments on forecasting skill demand sequences that have undergone structural breaks. Initially, the overall errors are quite pronounced, underscoring the challenge of accurately predicting these skills. Moreover, FiLM performs well on most metrics, which further verifies its robustness.

D.2 Job Skill Demand Forecasting for Low-Frequency Skills

In multigranular skill demand sequences, a significant number of skills remain inactive or in low frequency over extended periods. These skills might continue to have low demand in the future (indicating low importance), or they might suddenly gain interest from certain professions or companies, leading to rapid growth. In this study, we define low-frequency skills as those that appear fewer than twice in the time slices of the training set. Predicting the demand for these skills is challenging because their data points are predominantly zero during training, resulting in a lack of effective observational data. Therefore, we specifically present the demand prediction results of the existing benchmark models for these low-frequency skills.

Results. We continued to test the demand prediction effect on low-frequency skills using the benchmark models described in the main text, and the results are shown in Table 8. From this, we can draw the following conclusions: Firstly, there is a significant increase in the error on the RRMSE metric, indicating that low-demand skills are difficult to predict accurately. Secondly, Koopa has the best predictive performance in this scenario. We also found that the performance of SegRNN significantly decreases, suggesting that SegRNN's segment learning approach is not suitable for predicting low-frequency skill demands due to a lack of effective observational data, rendering the learning segments meaningless.

D.3 Skill Co-occurrence Graph Enhanced Job Skill Demand Forecasting

In the task of job skill demand forecasting, fully leveraging the inter-relationships among different skills is beneficial for downstream tasks. Therefore, we construct a prior graph with co-occurrence frequency from the training data to include as a dataset component. Given a set of granularities $i, j, \ldots, k$, we constructed the skill co-occurrence graph as $\mathcal{G}^{i,j,\ldots,k} = (\mathcal{V}^{i,j,\ldots,k}, \mathcal{E}^{i,j,\ldots,k})$, where $\mathcal{V}^{i,j,\ldots,k}$ is the extended skill set under the multiple granularities. The edge weight $e_{v,v'} \in \mathcal{E}^{i,j,\ldots,k}$ between nodes v and v' is determined by the co-occurrence frequency of the node pair v, v' in the job advertisement data for training $\mathcal{P}^{train}$. Specifically, given $v = (a^i, a^j, \ldots, a^k, s)$ and $v' = (a^{i\prime}, a^{j\prime}, ..., a^{k\prime}, s')$, $e_{v,v'}$ is calculated as:

$$e_{v,v'} = \sum_{p \in \mathcal{P}^{train}} \prod_{x \in \{a^i, a^j, \ldots, a^k, a^{i\prime}, a^{j\prime}, ..., a^{k\prime}, s, s'\}} \mathbf{1}_p(x \in p). \tag{4}$$

This information will serve as prior knowledge, reflecting global inter-skill dependency patterns.

---

[2]https://github.com/thuml/Time-Series-Library

17

Table 6: Overall performance comparisons on repeated experiments.

	Model	Market	Region	L1-O	L2-O	R&L1-O	R&L2-O	Company
MAE	LSTM	$314.54_{\pm 0.57}$	$49.92_{\pm 0.0}$	$24.43_{\pm 4.91}$	$8.04_{\pm 0.87}$	$4.44_{\pm 0.47}$	$1.45_{\pm 0.15}$	$1.49_{\pm 0.13}$
	SegRNN	$190.05_{\pm 0.37}$	$35.92_{\pm 0.47}$	$16.37_{\pm 3.73}$	$5.81_{\pm 0.73}$	$3.68_{\pm 0.5}$	$1.32_{\pm 0.25}$	$1.23_{\pm 0.2}$
	CHGH	$315.47_{\pm 0.04}$	$50.03_{\pm 0.02}$	$25.62_{\pm 3.23}$	$8.04_{\pm 0.87}$	$4.43_{\pm 0.48}$	$1.51_{\pm 0.7}$	$1.54_{\pm 0.73}$
	PreDyGAE	$189.95_{\pm 0.01}$	$35.84_{\pm 0.09}$	$21.72_{\pm 1.15}$	$6.81_{\pm 0.21}$	$4.08_{\pm 0.15}$	$1.85_{\pm 0.43}$	$1.85_{\pm 0.56}$
	Transformer	$340.72_{\pm 0.79}$	$54.98_{\pm 0.13}$	$27.43_{\pm 4.9}$	$8.93_{\pm 1.24}$	$4.92_{\pm 0.83}$	$1.75_{\pm 0.36}$	$1.69_{\pm 0.39}$
	Autoformer	$465.86_{\pm 2.78}$	$60.9_{\pm 0.64}$	$31.97_{\pm 8.13}$	$9.97_{\pm 1.6}$	$6.68_{\pm 0.21}$	$2.6_{\pm 0.17}$	$2.85_{\pm 0.42}$
	Informer	$340.76_{\pm 1.6}$	$55.08_{\pm 0.35}$	$27.54_{\pm 4.87}$	$8.81_{\pm 1.26}$	$4.87_{\pm 0.91}$	$1.73_{\pm 0.39}$	$1.69_{\pm 0.39}$
	Reformer	$344.9_{\pm 2.62}$	$54.83_{\pm 0.2}$	$27.35_{\pm 4.79}$	$8.88_{\pm 1.3}$	$4.9_{\pm 0.9}$	$1.71_{\pm 0.42}$	$1.8_{\pm 0.24}$
	FEDformer	$469.46_{\pm 4.52}$	$60.37_{\pm 0.19}$	$31.87_{\pm 8.22}$	$9.54_{\pm 1.9}$	$5.83_{\pm 1.1}$	$2.41_{\pm 0.39}$	$2.52_{\pm 0.08}$
	NStransformer	$208.23_{\pm 3.43}$	$37.39_{\pm 0.46}$	$19.13_{\pm 1.61}$	$6.29_{\pm 0.5}$	$3.79_{\pm 0.31}$	$1.26_{\pm 0.1}$	$1.64_{\pm 0.44}$
	PatchTST	$204.94_{\pm 12.41}$	$36.33_{\pm 2.23}$	$18.69_{\pm 3.45}$	$6.12_{\pm 0.89}$	$3.69_{\pm 0.54}$	$1.23_{\pm 0.18}$	$1.21_{\pm 0.18}$
	DLinear	$201.05_{\pm 1.95}$	$35.45_{\pm 0.56}$	$18.23_{\pm 1.48}$	$5.89_{\pm 0.41}$	$3.55_{\pm 0.28}$	$1.21_{\pm 0.12}$	$1.18_{\pm 0.11}$
	TSMixer	$517.95_{\pm 1.49}$	$67.56_{\pm 11.06}$	$30.83_{\pm 8.66}$	$10.54_{\pm 2.19}$	$5.95_{\pm 0.12}$	$3.58_{\pm 2.16}$	$6.75_{\pm 6.58}$
	FreTS	$310.62_{\pm 4.57}$	$49.52_{\pm 0.73}$	$23.1_{\pm 6.05}$	$7.74_{\pm 1.11}$	$4.3_{\pm 0.59}$	$1.41_{\pm 0.17}$	$1.45_{\pm 0.17}$
	FiLM	$201.7_{\pm 24.86}$	$36.18_{\pm 2.29}$	$16.79_{\pm 3.51}$	$5.9_{\pm 0.75}$	$3.7_{\pm 0.42}$	$1.28_{\pm 0.13}$	$1.31_{\pm 0.12}$
	Koopa	$205.12_{\pm 4.95}$	$35.95_{\pm 0.5}$	$19.9_{\pm 0.01}$	$6.19_{\pm 0.13}$	$3.68_{\pm 0.13}$	$1.2_{\pm 0.05}$	$1.15_{\pm 0.06}$
RMSE	LSTM	$1799.7_{\pm 0.49}$	$341.97_{\pm 0.07}$	$308.58_{\pm 103.99}$	$148.32_{\pm 29.18}$	$66.87_{\pm 13.95}$	$30.01_{\pm 6.11}$	$35.35_{\pm 8.1}$
	SegRNN	$941.59_{\pm 1.08}$	$194.73_{\pm 1.59}$	$154.02_{\pm 41.75}$	$78.84_{\pm 9.14}$	$38.6_{\pm 3.97}$	$18.21_{\pm 2.05}$	$21.05_{\pm 4.58}$
	CHGH	$1803.56_{\pm 0.73}$	$344.15_{\pm 0.11}$	$335.97_{\pm 67.99}$	$148.94_{\pm 29.54}$	$67.12_{\pm 14.29}$	$19.3_{\pm 12.73}$	$20.61_{\pm 14.53}$
	PreDyGAE	$925.27_{\pm 0.01}$	$193.94_{\pm 0.04}$	$186.36_{\pm 1.4}$	$84.55_{\pm 1.44}$	$40.37_{\pm 1.6}$	$18.79_{\pm 1.27}$	$22.05_{\pm 3.48}$
	Transformer	$1829.87_{\pm 2.96}$	$360.87_{\pm 6.13}$	$318.17_{\pm 94.1}$	$151.16_{\pm 30.08}$	$67.66_{\pm 14.26}$	$30.42_{\pm 6.37}$	$35.12_{\pm 9.25}$
	Autoformer	$2177.05_{\pm 87.58}$	$313.84_{\pm 8.55}$	$270.26_{\pm 76.23}$	$122.29_{\pm 20.33}$	$62.56_{\pm 4.38}$	$26.57_{\pm 2.26}$	$35.47_{\pm 2.81}$
	Informer	$1854.81_{\pm 3.0}$	$359.97_{\pm 7.72}$	$314.97_{\pm 100.17}$	$148.73_{\pm 28.62}$	$67.01_{\pm 15.41}$	$30.14_{\pm 6.45}$	$35.1_{\pm 9.3}$
	Reformer	$1890.32_{\pm 13.72}$	$357.61_{\pm 1.23}$	$314.03_{\pm 100.12}$	$152.43_{\pm 32.71}$	$67.58_{\pm 15.18}$	$30.21_{\pm 6.76}$	$36.14_{\pm 8.01}$
	FEDformer	$2137.44_{\pm 58.69}$	$321.53_{\pm 2.51}$	$268.73_{\pm 79.24}$	$113.47_{\pm 22.37}$	$56.36_{\pm 12.0}$	$25.64_{\pm 3.57}$	$30.9_{\pm 3.64}$
	NStransformer	$1121.07_{\pm 25.29}$	$236.56_{\pm 12.38}$	$198.03_{\pm 44.33}$	$103.2_{\pm 15.48}$	$45.99_{\pm 8.13}$	$21.42_{\pm 3.62}$	$32.8_{\pm 1.86}$
	PatchTST	$1098.82_{\pm 43.54}$	$220.21_{\pm 7.5}$	$204.56_{\pm 57.97}$	$97.9_{\pm 17.38}$	$45.32_{\pm 9.07}$	$20.88_{\pm 3.94}$	$25.48_{\pm 5.83}$
	DLinear	$1107.0_{\pm 28.41}$	$222.14_{\pm 4.31}$	$211.02_{\pm 51.25}$	$99.81_{\pm 16.61}$	$46.04_{\pm 8.55}$	$21.12_{\pm 3.95}$	$25.7_{\pm 6.25}$
	TSMixer	$2578.82_{\pm 27.27}$	$438.06_{\pm 67.68}$	$350.22_{\pm 143.66}$	$172.63_{\pm 60.23}$	$78.88_{\pm 15.28}$	$48.76_{\pm 17.8}$	$82.91_{\pm 56.64}$
	FreTS	$1763.69_{\pm 25.03}$	$336.65_{\pm 4.04}$	$289.29_{\pm 111.07}$	$140.54_{\pm 31.17}$	$63.72_{\pm 14.52}$	$28.68_{\pm 6.17}$	$34.0_{\pm 7.86}$
	FiLM	$1071.14_{\pm 223.85}$	$216.77_{\pm 38.28}$	$173.2_{\pm 51.16}$	$83.31_{\pm 16.12}$	$39.26_{\pm 8.54}$	$18.23_{\pm 3.85}$	$22.53_{\pm 6.08}$
	Koopa	$1147.88_{\pm 31.89}$	$228.36_{\pm 7.6}$	$227.46_{\pm 43.96}$	$106.03_{\pm 12.93}$	$49.37_{\pm 7.89}$	$22.53_{\pm 3.49}$	$27.06_{\pm 6.28}$
SMAPE(%)	LSTM	$42.33_{\pm 0.04}$	$49.78_{\pm 0.03}$	$42.31_{\pm 0.85}$	$33.14_{\pm 0.27}$	$31.79_{\pm 0.19}$	$22.67_{\pm 0.24}$	$30.42_{\pm 0.14}$
	SegRNN	$39.67_{\pm 0.26}$	$51.14_{\pm 0.53}$	$42.16_{\pm 2.14}$	$34.62_{\pm 1.16}$	$36.02_{\pm 0.65}$	$30.62_{\pm 6.19}$	$32.77_{\pm 0.27}$
	CHGH	$42.21_{\pm 0.01}$	$49.03_{\pm 0.03}$	$40.83_{\pm 0.51}$	$29.98_{\pm 0.35}$	$28.42_{\pm 0.28}$	$46.97_{\pm 9.54}$	$30.69_{\pm 14.64}$
	PreDyGAE	$42.1_{\pm 0.14}$	$64.37_{\pm 1.24}$	$75.3_{\pm 23.21}$	$59.69_{\pm 0.77}$	$46.51_{\pm 11.69}$	$76.6_{\pm 3.58}$	$60.36_{\pm 31.18}$
	Transformer	$49.86_{\pm 0.46}$	$60.59_{\pm 0.02}$	$58.06_{\pm 2.26}$	$52.7_{\pm 7.73}$	$47.37_{\pm 14.81}$	$47.03_{\pm 12.77}$	$45.69_{\pm 16.5}$
	Autoformer	$73.47_{\pm 0.36}$	$79.08_{\pm 0.19}$	$79.44_{\pm 8.36}$	$79.16_{\pm 4.38}$	$86.16_{\pm 3.64}$	$86.32_{\pm 4.73}$	$91.12_{\pm 14.54}$
	Informer	$49.2_{\pm 0.2}$	$60.45_{\pm 0.03}$	$58.59_{\pm 1.59}$	$52.61_{\pm 7.82}$	$48.84_{\pm 12.86}$	$45.52_{\pm 14.82}$	$47.61_{\pm 13.88}$
	Reformer	$48.69_{\pm 0.17}$	$60.48_{\pm 0.17}$	$58.49_{\pm 1.74}$	$51.26_{\pm 9.75}$	$47.8_{\pm 14.24}$	$42.15_{\pm 19.43}$	$52.76_{\pm 6.85}$
	FEDformer	$73.57_{\pm 0.53}$	$78.11_{\pm 0.55}$	$79.26_{\pm 9.09}$	$76.98_{\pm 7.02}$	$78.54_{\pm 4.9}$	$82.05_{\pm 0.3}$	$86.29_{\pm 7.21}$
	NStransformer	$38.88_{\pm 0.09}$	$47.74_{\pm 0.14}$	$38.94_{\pm 0.76}$	$27.49_{\pm 1.08}$	$26.43_{\pm 1.33}$	$15.47_{\pm 0.84}$	$23.75_{\pm 0.41}$
	PatchTST	$37.56_{\pm 2.15}$	$46.52_{\pm 2.49}$	$37.82_{\pm 2.85}$	$26.98_{\pm 2.25}$	$26.7_{\pm 1.42}$	$15.2_{\pm 1.55}$	$22.29_{\pm 2.19}$
	DLinear	$37.7_{\pm 1.82}$	$49.27_{\pm 3.85}$	$41.79_{\pm 0.05}$	$34.65_{\pm 0.28}$	$34.07_{\pm 0.54}$	$26.6_{\pm 0.75}$	$31.42_{\pm 0.65}$
	TSMixer	$73.63_{\pm 1.83}$	$72.27_{\pm 9.41}$	$61.01_{\pm 4.04}$	$68.37_{\pm 3.58}$	$65.19_{\pm 15.78}$	$79.28_{\pm 37.95}$	$95.62_{\pm 54.39}$
	FreTS	$42.81_{\pm 0.56}$	$50.46_{\pm 0.62}$	$41.53_{\pm 1.62}$	$31.26_{\pm 0.98}$	$29.76_{\pm 1.08}$	$18.89_{\pm 1.16}$	$28.37_{\pm 1.03}$
	FiLM	$37.79_{\pm 1.75}$	$46.85_{\pm 4.01}$	$41.03_{\pm 1.29}$	$30.26_{\pm 0.56}$	$29.49_{\pm 0.64}$	$17.65_{\pm 0.37}$	$26.19_{\pm 0.42}$
	Koopa	$36.26_{\pm 0.39}$	$44.7_{\pm 0.95}$	$36.69_{\pm 1.05}$	$25.59_{\pm 0.12}$	$24.51_{\pm 0.09}$	$14.13_{\pm 0.14}$	$20.62_{\pm 0.18}$
RRMSE(%)	LSTM	$52.7_{\pm 0.09}$	$52.87_{\pm 0.03}$	$73.45_{\pm 14.2}$	$97.42_{\pm 12.52}$	$80.24_{\pm 10.81}$	$102.19_{\pm 13.54}$	$210.9_{\pm 33.32}$
	SegRNN	$20.55_{\pm 0.05}$	$22.49_{\pm 0.39}$	$32.2_{\pm 4.91}$	$41.44_{\pm 8.3}$	$39.72_{\pm 8.04}$	$49.27_{\pm 11.53}$	$71.42_{\pm 13.56}$
	CHGH	$52.92_{\pm 0.03}$	$53.46_{\pm 0.02}$	$77.26_{\pm 10.23}$	$98.13_{\pm 12.8}$	$80.86_{\pm 11.35}$	$54.98_{\pm 47.9}$	$100.68_{\pm 96.78}$
	PreDyGAE	$20.07_{\pm 0.0}$	$22.65_{\pm 0.01}$	$51.16_{\pm 26.22}$	$54.87_{\pm 26.22}$	$47.52_{\pm 17.38}$	$64.98_{\pm 30.23}$	$96.4_{\pm 45.03}$
	Transformer	$52.49_{\pm 0.18}$	$55.57_{\pm 1.14}$	$75.98_{\pm 10.71}$	$98.09_{\pm 12.61}$	$82.63_{\pm 5.9}$	$101.34_{\pm 13.21}$	$200.07_{\pm 32.6}$
	Autoformer	$49.81_{\pm 1.73}$	$38.47_{\pm 1.72}$	$54.53_{\pm 0.71}$	$64.59_{\pm 1.09}$	$61.08_{\pm 4.1}$	$71.01_{\pm 3.15}$	$103.89_{\pm 3.92}$
	Informer	$54.11_{\pm 0.2}$	$55.59_{\pm 2.17}$	$71.42_{\pm 12.08}$	$97.76_{\pm 8.25}$	$81.04_{\pm 10.45}$	$102.15_{\pm 10.95}$	$198.51_{\pm 31.17}$
	Reformer	$55.75_{\pm 0.54}$	$54.83_{\pm 0.62}$	$76.12_{\pm 13.48}$	$98.47_{\pm 13.48}$	$81.65_{\pm 8.01}$	$103.29_{\pm 11.35}$	$206.04_{\pm 33.02}$
	FEDformer	$48.25_{\pm 2.89}$	$39.48_{\pm 0.85}$	$56.52_{\pm 2.27}$	$60.81_{\pm 0.74}$	$55.93_{\pm 2.95}$	$69.07_{\pm 0.9}$	$102.11_{\pm 3.78}$
	NStransformer	$26.75_{\pm 0.76}$	$30.42_{\pm 1.94}$	$43.64_{\pm 3.24}$	$59.15_{\pm 1.44}$	$49.08_{\pm 0.17}$	$62.42_{\pm 0.79}$	$107.45_{\pm 6.09}$
	PatchTST	$26.38_{\pm 1.12}$	$28.31_{\pm 1.01}$	$47.51_{\pm 3.34}$	$56.73_{\pm 1.69}$	$47.96_{\pm 2.74}$	$64.14_{\pm 3.05}$	$108.92_{\pm 5.86}$
	DLinear	$26.6_{\pm 1.35}$	$28.58_{\pm 1.32}$	$49.04_{\pm 3.52}$	$58.69_{\pm 1.4}$	$51.47_{\pm 0.38}$	$64.72_{\pm 0.06}$	$110.8_{\pm 1.95}$
	TSMixer	$74.69_{\pm 3.7}$	$69.2_{\pm 15.69}$	$86.56_{\pm 23.18}$	$118.56_{\pm 17.54}$	$99.17_{\pm 10.85}$	$100.88_{\pm 4.01}$	$136.0_{\pm 30.91}$
	FreTS	$51.01_{\pm 0.94}$	$51.53_{\pm 0.82}$	$69.56_{\pm 13.82}$	$92.39_{\pm 10.91}$	$76.55_{\pm 9.53}$	$97.37_{\pm 11.25}$	$200.81_{\pm 23.97}$
	FiLM	$25.72_{\pm 7.38}$	$27.83_{\pm 7.19}$	$36.13_{\pm 1.29}$	$43.32_{\pm 0.5}$	$38.93_{\pm 1.16}$	$48.51_{\pm 0.7}$	$79.66_{\pm 2.5}$
	Koopa	$27.81_{\pm 1.3}$	$29.66_{\pm 1.55}$	$52.74_{\pm 5.07}$	$62.29_{\pm 2.78}$	$56.41_{\pm 1.28}$	$70.91_{\pm 2.82}$	$122.3_{\pm 1.52}$

Benchmark Models   To fully utilize the prior information from the co-occurency graph, we introduce several GNN-based methods for multivariate time series prediction. These methods leverage GNNs to extract the influences between different variables, effectively capturing the relationships among various time series. The specific models are as follows:

- **EvolveGCN** [80]: EvolveGCN introduces a recurrent mechanism to update the network parameters, as GCN parameters, for capturing the dynamism of the graphs. Two methods are introduced: EvolveGCNH, which learns the weight matrix of the graph at each time

18

Table 7: Performance comparisons on skill demand series with structural breaks.

	Model	Market	Region	L1-O	L2-O	R&L1-O	R&L2-O	Company
MAE	LSTM	$423.99_{\pm0.56}$	$109.63_{\pm0.04}$	$101.68_{\pm13.13}$	$63.39_{\pm4.97}$	$21.02_{\pm1.85}$	$8.55_{\pm0.59}$	$26.23_{\pm1.67}$
	SegRNN	$256.67_{\pm0.54}$	$76.09_{\pm1.01}$	$68.08_{\pm5.62}$	$44.93_{\pm0.88}$	$16.12_{\pm0.24}$	$7.07_{\pm0.47}$	$18.91_{\pm0.85}$
	CHGH	$425.17_{\pm0.04}$	$109.82_{\pm0.06}$	$104.41_{\pm9.23}$	$63.54_{\pm5.0}$	$18.11_{\pm1.93}$	$15.16_{\pm4.33}$	$19.75_{\pm13.35}$
	PreDyGAE	$296.27_{\pm0.01}$	$85.05_{\pm5.06}$	$84.11_{\pm8.87}$	$56.28_{\pm7.2}$	$16.57_{\pm0.84}$	$9.51_{\pm0.19}$	$22.9_{\pm0.13}$
	Transformer	$460.45_{\pm2.31}$	$120.62_{\pm0.4}$	$113.11_{\pm13.19}$	$69.45_{\pm7.05}$	$22.4_{\pm2.76}$	$9.42_{\pm0.88}$	$26.93_{\pm4.12}$
	Autoformer	$632.29_{\pm2.76}$	$131.92_{\pm1.52}$	$131.01_{\pm21.72}$	$75.4_{\pm7.07}$	$25.82_{\pm0.93}$	$11.76_{\pm0.39}$	$36.65_{\pm6.8}$
	Informer	$462.88_{\pm1.99}$	$120.89_{\pm0.68}$	$113.36_{\pm13.21}$	$68.13_{\pm7.46}$	$22.22_{\pm2.91}$	$9.1_{\pm1.26}$	$26.82_{\pm4.06}$
	Reformer	$465.72_{\pm3.72}$	$120.58_{\pm0.34}$	$112.51_{\pm13.12}$	$69.02_{\pm7.59}$	$22.4_{\pm2.88}$	$9.09_{\pm1.44}$	$28.16_{\pm2.29}$
	FEDformer	$645.02_{\pm6.45}$	$131.05_{\pm0.27}$	$130.18_{\pm22.59}$	$72.12_{\pm9.14}$	$23.06_{\pm2.34}$	$11.01_{\pm0.59}$	$32.89_{\pm1.1}$
	NStransformer	$285.18_{\pm6.1}$	$81.54_{\pm1.2}$	$80.2_{\pm2.03}$	$49.66_{\pm0.33}$	$16.88_{\pm0.26}$	$7.07_{\pm0.2}$	$28.34_{\pm10.69}$
	PatchTST	$282.34_{\pm17.19}$	$79.1_{\pm4.75}$	$80.46_{\pm2.76}$	$48.45_{\pm3.13}$	$16.29_{\pm1.29}$	$7.01_{\pm0.41}$	$19.65_{\pm1.48}$
	DLinear	$277.33_{\pm2.32}$	$77.47_{\pm0.99}$	$77.64_{\pm3.23}$	$46.5_{\pm0.16}$	$16.03_{\pm0.08}$	$6.64_{\pm0.12}$	$18.73_{\pm0.51}$
	TSMixer	$698.27_{\pm6.79}$	$146.22_{\pm23.62}$	$130.91_{\pm21.4}$	$88.39_{\pm4.38}$	$28.49_{\pm1.37}$	$16.46_{\pm8.55}$	$79.02_{\pm69.45}$
	FreTS	$419.07_{\pm6.85}$	$108.63_{\pm1.75}$	$98.46_{\pm14.62}$	$62.17_{\pm5.14}$	$20.54_{\pm1.82}$	$8.47_{\pm0.54}$	$25.89_{\pm1.53}$
	FiLM	$277.66_{\pm34.84}$	$78.58_{\pm6.24}$	$69.56_{\pm6.12}$	$45.37_{\pm2.5}$	$15.26_{\pm0.87}$	$6.79_{\pm0.39}$	$19.86_{\pm0.98}$
	Koopa	$282.85_{\pm6.89}$	$78.75_{\pm1.42}$	$83.25_{\pm7.31}$	$48.75_{\pm1.55}$	$16.92_{\pm0.47}$	$6.93_{\pm0.13}$	$19.16_{\pm0.11}$
RMSE	LSTM	$2148.33_{\pm0.79}$	$534.61_{\pm0.67}$	$704.41_{\pm136.89}$	$466.21_{\pm60.24}$	$178.29_{\pm26.25}$	$58.54_{\pm5.62}$	$192.72_{\pm30.76}$
	SegRNN	$1130.19_{\pm0.81}$	$304.2_{\pm2.76}$	$387.1_{\pm3.14}$	$264.53_{\pm10.99}$	$110.27_{\pm3.44}$	$36.38_{\pm1.34}$	$120.84_{\pm7.95}$
	CHGH	$2153.77_{\pm0.87}$	$537.97_{\pm0.18}$	$735.74_{\pm97.15}$	$468.27_{\pm61.0}$	$159.45_{\pm15.71}$	$41.0_{\pm28.76}$	$203.78_{\pm87.33}$
	PreDyGAE	$1510.41_{\pm0.01}$	$403.02_{\pm0.06}$	$529.39_{\pm58.83}$	$388.54_{\pm43.46}$	$159.45_{\pm15.71}$	$36.48_{\pm2.19}$	$162.0_{\pm20.48}$
	Transformer	$2154.53_{\pm30.61}$	$562.07_{\pm12.46}$	$717.62_{\pm125.1}$	$474.06_{\pm63.8}$	$179.2_{\pm25.64}$	$61.55_{\pm5.56}$	$191.91_{\pm36.18}$
	Autoformer	$2566.07_{\pm102.19}$	$489.69_{\pm10.56}$	$621.06_{\pm80.33}$	$387.57_{\pm33.41}$	$155.94_{\pm0.51}$	$58.1_{\pm4.51}$	$187.41_{\pm19.68}$
	Informer	$2212.83_{\pm2.36}$	$559.86_{\pm11.51}$	$715.12_{\pm132.16}$	$467.05_{\pm58.69}$	$178.41_{\pm28.75}$	$57.53_{\pm7.65}$	$191.85_{\pm36.5}$
	Reformer	$2248.94_{\pm15.68}$	$558.75_{\pm0.8}$	$713.91_{\pm131.53}$	$478.6_{\pm70.65}$	$179.85_{\pm27.33}$	$59.48_{\pm9.48}$	$195.6_{\pm31.87}$
	FEDformer	$2551.16_{\pm77.11}$	$501.84_{\pm4.46}$	$621.78_{\pm81.74}$	$363.84_{\pm35.37}$	$149.67_{\pm19.37}$	$57.89_{\pm2.21}$	$166.63_{\pm10.36}$
	NStransformer	$1337.16_{\pm47.86}$	$373.7_{\pm20.75}$	$473.97_{\pm10.71}$	$334.17_{\pm14.36}$	$128.79_{\pm8.11}$	$39.55_{\pm1.82}$	$176.86_{\pm17.5}$
	PatchTST	$1328.54_{\pm49.25}$	$347.07_{\pm11.24}$	$499.82_{\pm22.78}$	$321.16_{\pm15.89}$	$125.31_{\pm13.05}$	$39.74_{\pm1.04}$	$144.13_{\pm14.98}$
	DLinear	$1341.55_{\pm32.44}$	$350.8_{\pm6.91}$	$506.06_{\pm18.99}$	$325.15_{\pm16.25}$	$130.09_{\pm10.18}$	$38.99_{\pm1.16}$	$144.53_{\pm18.45}$
	TSMixer	$3040.93_{\pm86.91}$	$678.15_{\pm105.71}$	$816.57_{\pm184.07}$	$582.2_{\pm93.85}$	$211.98_{\pm22.73}$	$115.88_{\pm67.93}$	$436.64_{\pm301.19}$
	FreTS	$2105.11_{\pm31.62}$	$526.0_{\pm6.74}$	$679.7_{\pm130.16}$	$451.97_{\pm53.49}$	$173.34_{\pm22.83}$	$57.76_{\pm4.47}$	$188.29_{\pm25.82}$
	FiLM	$1298.06_{\pm265.59}$	$341.89_{\pm60.42}$	$426.14_{\pm19.46}$	$276.0_{\pm13.7}$	$111.46_{\pm9.34}$	$34.32_{\pm1.86}$	$128.54_{\pm16.33}$
	Koopa	$1386.85_{\pm35.21}$	$360.64_{\pm11.92}$	$532.64_{\pm14.5}$	$338.26_{\pm12.88}$	$138.63_{\pm9.35}$	$41.39_{\pm0.09}$	$152.48_{\pm17.55}$
SMAPE(%)	LSTM	$37.23_{\pm0.05}$	$45.67_{\pm0.02}$	$45.95_{\pm1.99}$	$50.34_{\pm1.28}$	$47.36_{\pm0.66}$	$42.47_{\pm0.4}$	$70.15_{\pm1.62}$
	SegRNN	$32.88_{\pm0.2}$	$43.35_{\pm0.39}$	$41.54_{\pm2.12}$	$46.46_{\pm3.08}$	$47.11_{\pm1.77}$	$44.29_{\pm4.45}$	$59.48_{\pm1.85}$
	CHGH	$37.21_{\pm0.01}$	$45.55_{\pm0.03}$	$46.29_{\pm1.26}$	$50.3_{\pm1.28}$	$46.22_{\pm0.72}$	$45.92_{\pm21.79}$	$77.34_{\pm37.44}$
	PreDyGAE	$32.0_{\pm0.02}$	$43.25_{\pm0.18}$	$49.33_{\pm2.76}$	$51.02_{\pm5.06}$	$50.59_{\pm1.35}$	$63.2_{\pm17.41}$	$70.72_{\pm0.42}$
	Transformer	$44.18_{\pm0.71}$	$53.31_{\pm0.19}$	$53.13_{\pm2.85}$	$56.88_{\pm3.45}$	$55.61_{\pm8.32}$	$55.47_{\pm7.12}$	$73.21_{\pm7.57}$
	Autoformer	$67.46_{\pm0.27}$	$68.4_{\pm0.03}$	$70.67_{\pm6.58}$	$71.92_{\pm3.02}$	$83.21_{\pm4.31}$	$83.71_{\pm4.79}$	$99.62_{\pm14.04}$
	Informer	$43.4_{\pm0.31}$	$53.28_{\pm0.16}$	$53.54_{\pm2.22}$	$56.42_{\pm4.14}$	$56.23_{\pm7.69}$	$54.33_{\pm8.61}$	$73.34_{\pm7.52}$
	Reformer	$43.1_{\pm0.01}$	$53.09_{\pm0.1}$	$53.42_{\pm2.4}$	$56.23_{\pm4.31}$	$55.84_{\pm8.2}$	$52.75_{\pm10.9}$	$76.27_{\pm3.57}$
	FEDformer	$68.1_{\pm0.29}$	$67.4_{\pm0.48}$	$70.57_{\pm7.06}$	$70.35_{\pm5.46}$	$76.14_{\pm3.56}$	$79.91_{\pm0.11}$	$95.19_{\pm7.37}$
	NStransformer	$32.85_{\pm0.14}$	$42.35_{\pm0.05}$	$44.17_{\pm1.08}$	$47.32_{\pm0.29}$	$44.23_{\pm1.08}$	$38.15_{\pm1.31}$	$103.89_{\pm60.87}$
	PatchTST	$31.96_{\pm2.0}$	$41.14_{\pm2.29}$	$42.63_{\pm1.59}$	$45.79_{\pm2.31}$	$43.83_{\pm2.11}$	$37.55_{\pm2.57}$	$58.04_{\pm2.54}$
	DLinear	$31.38_{\pm1.08}$	$40.56_{\pm1.37}$	$41.8_{\pm1.23}$	$44.83_{\pm0.38}$	$45.42_{\pm0.14}$	$41.61_{\pm0.42}$	$58.22_{\pm0.52}$
	TSMixer	$68.55_{\pm1.22}$	$67.06_{\pm11.47}$	$61.4_{\pm6.47}$	$75.25_{\pm0.75}$	$72.63_{\pm11.4}$	$83.29_{\pm31.13}$	$109.31_{\pm46.77}$
	FreTS	$37.47_{\pm0.7}$	$46.12_{\pm0.91}$	$45.36_{\pm2.66}$	$50.03_{\pm1.63}$	$46.82_{\pm1.3}$	$41.25_{\pm1.28}$	$70.0_{\pm1.47}$
	FiLM	$32.01_{\pm0.16}$	$41.26_{\pm0.46}$	$41.12_{\pm1.97}$	$45.56_{\pm1.21}$	$45.99_{\pm0.95}$	$40.91_{\pm0.79}$	$58.02_{\pm1.51}$
	Koopa	$31.01_{\pm0.02}$	$39.92_{\pm0.13}$	$42.95_{\pm3.2}$	$45.56_{\pm1.21}$	$41.88_{\pm0.65}$	$35.91_{\pm0.3}$	$57.19_{\pm1.63}$
RRMSE(%)	LSTM	$53.94_{\pm0.16}$	$53.9_{\pm0.12}$	$73.77_{\pm14.35}$	$98.43_{\pm12.75}$	$93.46_{\pm13.83}$	$65.39_{\pm6.3}$	$227.07_{\pm36.31}$
	SegRNN	$21.23_{\pm0.03}$	$22.98_{\pm0.43}$	$32.33_{\pm4.99}$	$41.84_{\pm8.49}$	$44.18_{\pm9.25}$	$34.49_{\pm6.31}$	$73.51_{\pm14.73}$
	CHGH	$54.21_{\pm0.03}$	$54.5_{\pm0.02}$	$77.63_{\pm10.33}$	$99.15_{\pm13.03}$	$93.96_{\pm14.27}$	$53.56_{\pm32.26}$	$105.96_{\pm104.02}$
	PreDyGAE	$50.72_{\pm0.01}$	$33.15_{\pm0.01}$	$46.59_{\pm9.64}$	$55.32_{\pm12.77}$	$57.56_{\pm12.47}$	$45.53_{\pm6.19}$	$91.47_{\pm23.2}$
	Transformer	$52.69_{\pm0.67}$	$56.3_{\pm1.62}$	$76.3_{\pm10.8}$	$99.08_{\pm12.9}$	$94.57_{\pm7.31}$	$67.56_{\pm5.78}$	$215.96_{\pm35.18}$
	Autoformer	$51.07_{\pm0.19}$	$39.44_{\pm1.34}$	$54.75_{\pm0.61}$	$65.4_{\pm1.15}$	$64.75_{\pm3.85}$	$54.46_{\pm8.57}$	$106.51_{\pm5.39}$
	Informer	$55.42_{\pm0.04}$	$56.22_{\pm2.17}$	$71.7_{\pm12.14}$	$98.81_{\pm8.31}$	$93.29_{\pm11.35}$	$64.2_{\pm6.07}$	$213.07_{\pm32.78}$
	Reformer	$56.91_{\pm0.64}$	$55.93_{\pm0.41}$	$76.53_{\pm13.64}$	$99.47_{\pm13.64}$	$95.12_{\pm10.02}$	$67.05_{\pm7.72}$	$221.45_{\pm36.6}$
	FEDformer	$49.35_{\pm2.5}$	$40.99_{\pm1.28}$	$56.8_{\pm2.22}$	$61.57_{\pm0.66}$	$63.27_{\pm4.31}$	$54.5_{\pm0.19}$	$105.93_{\pm4.82}$
	NStransformer	$27.47_{\pm1.23}$	$31.55_{\pm2.22}$	$43.9_{\pm3.25}$	$60.16_{\pm1.54}$	$56.5_{\pm1.0}$	$39.75_{\pm0.02}$	$116.26_{\pm2.18}$
	PatchTST	$27.53_{\pm1.1}$	$29.31_{\pm1.0}$	$47.82_{\pm3.34}$	$57.57_{\pm1.19}$	$55.03_{\pm2.9}$	$40.96_{\pm1.97}$	$114.76_{\pm7.12}$
	DLinear	$27.83_{\pm1.38}$	$29.66_{\pm1.38}$	$49.37_{\pm3.5}$	$59.59_{\pm1.4}$	$59.48_{\pm0.62}$	$40.74_{\pm0.89}$	$116.83_{\pm1.45}$
	TSMixer	$76.26_{\pm6.24}$	$69.83_{\pm15.63}$	$86.8_{\pm23.27}$	$119.17_{\pm17.64}$	$113.43_{\pm16.08}$	$77.61_{\pm14.65}$	$141.79_{\pm36.12}$
	FreTS	$52.23_{\pm1.04}$	$52.51_{\pm0.92}$	$69.87_{\pm13.95}$	$93.32_{\pm11.09}$	$88.52_{\pm11.95}$	$63.64_{\pm5.3}$	$215.64_{\pm25.29}$
	FiLM	$26.91_{\pm7.65}$	$28.84_{\pm7.48}$	$36.38_{\pm1.32}$	$43.93_{\pm0.54}$	$44.44_{\pm1.26}$	$31.45_{\pm0.6}$	$83.15_{\pm2.17}$
	Koopa	$28.98_{\pm1.27}$	$30.77_{\pm1.6}$	$53.08_{\pm5.03}$	$63.24_{\pm2.77}$	$65.24_{\pm0.87}$	$44.14_{\pm3.04}$	$129.36_{\pm2.2}$

step as a hidden state, and EvolveGCNO, which directly employs the weight evolution as a hidden state output, decoupled from node embedding.

- **GConvGRU** [81]: This model integrates convolutional neural networks (CNNs) on graphs to identify spatial structures and recurrent neural networks (RNNs) to detect dynamic patterns. Two architectures, GConvGRU and GConvLSTM, are explored for the Graph Convolutional Recurrent Network (GCRN).

- **TGCN** [82]: The temporal graph convolutional network (T-GCN) model, which is in combination with the graph convolutional network (GCN) and gated recurrent unit (GRU). Specifically, the GCN is used to learn complex topological structures to capture spatial

19

Table 8: Performance comparisons on skill demand series with low-frequency.

Metric	Model	Market	Region	L1-O	L2-O	R&L1-O	R&L2-O	Company
MAE	LSTM	$33.69_{\pm0.44}$	$26.88_{\pm0.05}$	$16.49_{\pm0.01}$	$12.11_{\pm0.04}$	$12.28_{\pm0.02}$	$8.92_{\pm0.1}$	$12.6_{\pm0.02}$
	SegRNN	$51.5_{\pm0.57}$	$46.12_{\pm0.28}$	$25.65_{\pm0.39}$	$18.03_{\pm0.03}$	$22.08_{\pm1.92}$	$19.1_{\pm3.22}$	$19.37_{\pm0.88}$
	CHGH	$32.37_{\pm0.0}$	$24.99_{\pm0.02}$	$14.2_{\pm0.04}$	$9.32_{\pm0.01}$	$9.45_{\pm0.0}$	$67.89_{\pm10.03}$	$77.01_{\pm5.58}$
	PreDyGAE	$113.67_{\pm1.73}$	$157.17_{\pm18.68}$	$343.32_{\pm0.15}$	$64.46_{\pm13.49}$	$24.16_{\pm0.13}$	$60.41_{\pm55.15}$	$68.89_{\pm62.89}$
	Transformer	$54.84_{\pm1.99}$	$52.95_{\pm0.09}$	$46.1_{\pm0.13}$	$45.47_{\pm0.07}$	$45.64_{\pm0.02}$	$45.1_{\pm0.01}$	$45.43_{\pm0.0}$
	Autoformer	$132.98_{\pm11.42}$	$106.13_{\pm1.65}$	$110.06_{\pm0.19}$	$97.66_{\pm1.23}$	$98.28_{\pm0.46}$	$95.26_{\pm1.1}$	$90.23_{\pm0.54}$
	Informer	$54.42_{\pm1.83}$	$52.76_{\pm0.13}$	$46.1_{\pm0.0}$	$45.52_{\pm0.01}$	$45.69_{\pm0.02}$	$45.07_{\pm0.01}$	$45.47_{\pm0.0}$
	Reformer	$52.72_{\pm0.3}$	$53.25_{\pm0.15}$	$46.19_{\pm0.07}$	$45.58_{\pm0.02}$	$45.65_{\pm0.01}$	$45.08_{\pm0.03}$	$45.49_{\pm0.03}$
	FEDformer	$133.07_{\pm6.09}$	$101.47_{\pm1.47}$	$109.51_{\pm1.3}$	$96.43_{\pm1.0}$	$95.83_{\pm1.48}$	$94.18_{\pm0.48}$	$90.71_{\pm0.47}$
	NStransformer	$57.96_{\pm1.83}$	$34.24_{\pm0.3}$	$17.99_{\pm0.37}$	$10.84_{\pm0.04}$	$11.46_{\pm0.02}$	$6.45_{\pm0.02}$	$11.9_{\pm0.0}$
	PatchTST	$47.25_{\pm6.07}$	$33.21_{\pm3.3}$	$17.31_{\pm1.85}$	$10.39_{\pm1.03}$	$10.91_{\pm1.1}$	$6.1_{\pm0.67}$	$11.4_{\pm1.23}$
	DLinear	$51.7_{\pm7.31}$	$42.08_{\pm8.83}$	$29.64_{\pm9.87}$	$25.18_{\pm10.9}$	$25.6_{\pm10.9}$	$22.53_{\pm11.72}$	$25.79_{\pm10.77}$
	TSMixer	$102.23_{\pm14.26}$	$64.77_{\pm10.8}$	$51.89_{\pm9.94}$	$52.61_{\pm1.14}$	$40.06_{\pm7.0}$	$34.02_{\pm2.2}$	$41.32_{\pm2.03}$
	FreTS	$38.55_{\pm0.83}$	$28.89_{\pm0.96}$	$16.6_{\pm0.58}$	$11.19_{\pm0.47}$	$11.46_{\pm0.51}$	$7.44_{\pm0.47}$	$11.94_{\pm0.48}$
	FiLM	$51.42_{\pm20.47}$	$34.97_{\pm10.99}$	$18.23_{\pm5.69}$	$10.9_{\pm3.11}$	$11.47_{\pm3.38}$	$6.43_{\pm1.96}$	$11.95_{\pm3.58}$
	Koopa	$43.3_{\pm4.56}$	$29.66_{\pm2.37}$	$15.52_{\pm1.38}$	$9.49_{\pm0.73}$	$9.95_{\pm0.82}$	$5.53_{\pm0.48}$	$10.35_{\pm0.86}$
RMSE	LSTM	$99.83_{\pm0.77}$	$306.02_{\pm0.33}$	$168.19_{\pm0.01}$	$140.57_{\pm0.0}$	$247.24_{\pm0.02}$	$125.6_{\pm0.01}$	$272.77_{\pm0.04}$
	SegRNN	$141.21_{\pm0.31}$	$452.49_{\pm0.09}$	$233.37_{\pm0.06}$	$146.34_{\pm0.22}$	$216.77_{\pm0.01}$	$117.81_{\pm0.63}$	$227.59_{\pm0.0}$
	CHGH	$100.44_{\pm0.02}$	$306.28_{\pm0.01}$	$168.23_{\pm0.01}$	$140.54_{\pm0.0}$	$247.21_{\pm0.01}$	$134.20_{\pm0.0}$	$208.01_{\pm17.09}$
	PreDyGAE	$169.76_{\pm1.46}$	$542.42_{\pm24.84}$	$771.72_{\pm50.15}$	$195.02_{\pm16.23}$	$218.94_{\pm0.65}$	$119.76_{\pm109.33}$	$188.17_{\pm171.77}$
	Transformer	$112.3_{\pm5.16}$	$311.19_{\pm1.43}$	$176.74_{\pm0.52}$	$150.63_{\pm0.53}$	$252.9_{\pm0.1}$	$137.39_{\pm0.13}$	$278.45_{\pm0.13}$
	Autoformer	$249.03_{\pm36.15}$	$364.32_{\pm9.52}$	$231.67_{\pm3.95}$	$184.29_{\pm1.11}$	$266.1_{\pm0.64}$	$172.38_{\pm1.84}$	$281.53_{\pm0.17}$
	Informer	$111.59_{\pm6.84}$	$312.15_{\pm0.35}$	$176.18_{\pm0.19}$	$150.46_{\pm0.16}$	$252.84_{\pm0.01}$	$137.07_{\pm0.12}$	$279.56_{\pm0.36}$
	Reformer	$107.11_{\pm3.02}$	$314.04_{\pm1.77}$	$176.73_{\pm0.57}$	$150.62_{\pm0.23}$	$252.85_{\pm0.14}$	$137.02_{\pm0.12}$	$279.79_{\pm0.61}$
	FEDformer	$233.86_{\pm32.3}$	$346.25_{\pm1.54}$	$227.64_{\pm2.23}$	$182.54_{\pm0.06}$	$264.12_{\pm0.64}$	$171.43_{\pm0.76}$	$280.57_{\pm1.11}$
	NStransformer	$159.4_{\pm10.66}$	$329.86_{\pm4.31}$	$182.83_{\pm8.95}$	$136.01_{\pm2.69}$	$249.63_{\pm0.44}$	$129.74_{\pm5.67}$	$255.48_{\pm4.76}$
	PatchTST	$126.71_{\pm17.24}$	$339.92_{\pm16.66}$	$182.49_{\pm8.84}$	$134.57_{\pm2.89}$	$233.14_{\pm1.45}$	$120.94_{\pm1.7}$	$253.85_{\pm3.35}$
	DLinear	$109.06_{\pm5.5}$	$320.09_{\pm2.51}$	$173.51_{\pm2.67}$	$133.82_{\pm2.37}$	$235.23_{\pm1.35}$	$123.01_{\pm2.84}$	$254.7_{\pm1.91}$
	TSMixer	$186.01_{\pm18.76}$	$359.57_{\pm28.14}$	$199.07_{\pm10.44}$	$160.52_{\pm5.4}$	$261.65_{\pm6.67}$	$139.7_{\pm3.15}$	$296.15_{\pm1.83}$
	FreTS	$104.43_{\pm1.38}$	$310.94_{\pm2.43}$	$170.36_{\pm1.47}$	$139.96_{\pm0.74}$	$239.77_{\pm4.49}$	$122.58_{\pm2.07}$	$266.82_{\pm3.88}$
	FiLM	$129.11_{\pm32.28}$	$343.63_{\pm34.15}$	$185.3_{\pm16.68}$	$135.8_{\pm1.72}$	$235.13_{\pm10.73}$	$122.84_{\pm2.06}$	$256.54_{\pm7.91}$
	Koopa	$111.96_{\pm8.52}$	$317.8_{\pm3.82}$	$171.28_{\pm3.23}$	$134.11_{\pm0.67}$	$242.79_{\pm2.91}$	$124.64_{\pm0.36}$	$260.87_{\pm0.46}$
SMAPE(%)	LSTM	$26.1_{\pm0.39}$	$19.43_{\pm0.17}$	$15.18_{\pm0.01}$	$12.88_{\pm0.07}$	$13.01_{\pm0.02}$	$11.46_{\pm0.17}$	$13.01_{\pm0.03}$
	SegRNN	$33.71_{\pm0.71}$	$29.55_{\pm0.4}$	$22.04_{\pm0.56}$	$18.44_{\pm0.11}$	$24.16_{\pm2.97}$	$24.8_{\pm1.63}$	$19.83_{\pm1.32}$
	CHGH	$23.61_{\pm0.0}$	$15.99_{\pm0.04}$	$11.08_{\pm0.09}$	$7.8_{\pm0.03}$	$7.86_{\pm0.01}$	$9.61_{\pm1.08}$	$8.07_{\pm2.03}$
	PreDyGAE	$83.59_{\pm0.72}$	$82.33_{\pm4.38}$	$114.25_{\pm1.42}$	$48.0_{\pm7.33}$	$23.97_{\pm0.1}$	$47.34_{\pm43.22}$	$48.31_{\pm44.1}$
	Transformer	$49.05_{\pm1.26}$	$48.23_{\pm0.0}$	$48.68_{\pm0.14}$	$50.78_{\pm0.08}$	$50.92_{\pm0.01}$	$52.6_{\pm0.01}$	$50.28_{\pm0.0}$
	Autoformer	$80.95_{\pm1.5}$	$76.73_{\pm0.44}$	$81.69_{\pm0.03}$	$80.14_{\pm0.55}$	$80.43_{\pm0.21}$	$81.25_{\pm0.49}$	$76.73_{\pm0.25}$
	Informer	$48.87_{\pm0.78}$	$48.07_{\pm0.25}$	$48.69_{\pm0.03}$	$50.83_{\pm0.01}$	$50.96_{\pm0.02}$	$52.57_{\pm0.01}$	$50.29_{\pm0.0}$
	Reformer	$48.32_{\pm1.06}$	$48.33_{\pm0.02}$	$48.76_{\pm0.08}$	$50.87_{\pm0.03}$	$50.93_{\pm0.0}$	$52.59_{\pm0.02}$	$50.28_{\pm0.02}$
	FEDformer	$82.56_{\pm1.23}$	$75.32_{\pm0.52}$	$81.88_{\pm0.57}$	$79.7_{\pm0.39}$	$79.55_{\pm0.62}$	$80.83_{\pm0.21}$	$76.96_{\pm0.21}$
	NStransformer	$32.62_{\pm0.69}$	$18.84_{\pm0.31}$	$11.59_{\pm0.02}$	$7.3_{\pm0.01}$	$7.81_{\pm0.01}$	$4.76_{\pm0.0}$	$7.76_{\pm0.0}$
	PatchTST	$29.48_{\pm1.53}$	$18.23_{\pm1.46}$	$11.15_{\pm1.04}$	$7.01_{\pm0.71}$	$7.51_{\pm0.76}$	$4.56_{\pm0.52}$	$7.46_{\pm0.8}$
	DLinear	$42.19_{\pm8.16}$	$35.1_{\pm10.8}$	$31.64_{\pm12.55}$	$30.55_{\pm14.04}$	$30.93_{\pm14.03}$	$30.35_{\pm15.14}$	$30.52_{\pm13.82}$
	TSMixer	$68.79_{\pm5.48}$	$51.84_{\pm6.45}$	$48.47_{\pm6.43}$	$54.61_{\pm0.92}$	$45.87_{\pm6.93}$	$43.34_{\pm2.84}$	$45.68_{\pm1.95}$
	FreTS	$29.63_{\pm0.39}$	$19.71_{\pm0.62}$	$14.01_{\pm0.56}$	$10.6_{\pm0.64}$	$10.86_{\pm0.67}$	$8.45_{\pm0.75}$	$11.09_{\pm0.61}$
	FiLM	$30.74_{\pm10.37}$	$18.56_{\pm6.37}$	$11.42_{\pm4.06}$	$7.19_{\pm2.55}$	$7.7_{\pm2.79}$	$4.69_{\pm1.69}$	$7.64_{\pm2.84}$
	Koopa	$28.27_{\pm2.34}$	$16.67_{\pm1.38}$	$10.15_{\pm0.9}$	$6.35_{\pm0.57}$	$6.79_{\pm0.63}$	$4.09_{\pm0.39}$	$6.73_{\pm0.63}$
RRMSE(%)	LSTM	$432.63_{\pm21.64}$	$1364.7_{\pm202.93}$	$1057.49_{\pm28.76}$	$1468.93_{\pm15.86}$	$2581.13_{\pm33.65}$	$1485.07_{\pm28.13}$	$1828.92_{\pm33.42}$
	SegRNN	$114.7_{\pm0.21}$	$125.97_{\pm0.16}$	$120.24_{\pm0.18}$	$146.81_{\pm0.73}$	$130.23_{\pm0.45}$	$130.24_{\pm0.87}$	$133.03_{\pm0.54}$
	CHGH	$452.99_{\pm0.65}$	$1662.7_{\pm2.11}$	$1394.35_{\pm4.28}$	$1824.0_{\pm1.1}$	$3436.23_{\pm14.46}$	$289.09_{\pm14.07}$	$406.07_{\pm8.08}$
	PreDyGAE	$104.17_{\pm0.99}$	$216.23_{\pm2.0}$	$201.73_{\pm0.3}$	$220.44_{\pm4.62}$	$229.42_{\pm0.22}$	$165.21_{\pm59.52}$	$272.3_{\pm66.0}$
	Transformer	$172.07_{\pm30.78}$	$513.13_{\pm29.31}$	$320.11_{\pm12.98}$	$269.97_{\pm3.45}$	$468.03_{\pm0.41}$	$242.99_{\pm4.1}$	$466.5_{\pm1.48}$
	Autoformer	$89.47_{\pm0.4}$	$161.16_{\pm10.32}$	$136.15_{\pm6.04}$	$134.91_{\pm1.17}$	$195.47_{\pm1.88}$	$134.43_{\pm1.31}$	$191.17_{\pm2.0}$
	Informer	$193.88_{\pm30.22}$	$538.02_{\pm31.16}$	$316.88_{\pm4.27}$	$275.64_{\pm0.03}$	$470.15_{\pm2.27}$	$246.67_{\pm0.09}$	$462.58_{\pm1.78}$
	Reformer	$200.52_{\pm32.93}$	$536.21_{\pm13.15}$	$324.86_{\pm11.04}$	$276.66_{\pm0.14}$	$470.18_{\pm0.3}$	$247.32_{\pm0.89}$	$460.86_{\pm4.08}$
	FEDformer	$96.72_{\pm6.76}$	$176.64_{\pm0.39}$	$136.62_{\pm1.69}$	$136.93_{\pm0.31}$	$199.25_{\pm1.04}$	$134.8_{\pm0.73}$	$189.29_{\pm1.49}$
	NStransformer	$90.14_{\pm7.16}$	$219.18_{\pm39.19}$	$193.48_{\pm36.91}$	$218.73_{\pm0.68}$	$406.29_{\pm7.04}$	$272.29_{\pm31.63}$	$259.15_{\pm8.34}$
	PatchTST	$95.85_{\pm5.0}$	$210.2_{\pm36.63}$	$199.12_{\pm34.35}$	$266.55_{\pm38.3}$	$391.36_{\pm46.33}$	$298.48_{\pm41.01}$	$278.87_{\pm33.7}$
	DLinear	$90.72_{\pm2.96}$	$241.41_{\pm14.46}$	$219.68_{\pm17.61}$	$281.05_{\pm37.87}$	$431.86_{\pm48.87}$	$296.1_{\pm49.89}$	$300.39_{\pm25.41}$
	TSMixer	$125.99_{\pm9.23}$	$200.39_{\pm22.34}$	$211.03_{\pm36.38}$	$192.85_{\pm35.49}$	$418.03_{\pm16.7}$	$275.29_{\pm1.9}$	$293.38_{\pm24.47}$
	FreTS	$222.55_{\pm30.4}$	$479.45_{\pm20.87}$	$462.19_{\pm33.86}$	$720.43_{\pm51.2}$	$962.73_{\pm47.52}$	$758.24_{\pm41.67}$	$809.76_{\pm76.22}$
	FiLM	$108.05_{\pm34.05}$	$444.72_{\pm401.85}$	$360.89_{\pm298.05}$	$402.84_{\pm297.56}$	$652.7_{\pm546.99}$	$417.27_{\pm297.85}$	$364.47_{\pm235.25}$
	Koopa	$93.83_{\pm0.11}$	$287.42_{\pm57.85}$	$264.45_{\pm56.56}$	$345.86_{\pm79.42}$	$582.52_{\pm154.25}$	$368.66_{\pm64.94}$	$345.87_{\pm66.81}$

dependence and the gated recurrent unit is used to learn dynamic changes of traffic data to capture temporal dependence.

- **GCLSTM** [83]: GCLSTM is an end-to-end model integrating a Graph Convolution Network (GCN) embedded Long Short-Term Memory network (LSTM) for dynamic network link prediction. The GCN captures local structural properties, while the LSTM learns temporal features across snapshots of a dynamic network.

- **DyGrEncoder** [84]: This approach combines a sequence-to-sequence encoder-decoder model with gated graph neural networks (GGNNs) and long short-term memory networks

20

Table 9: Performance comparisons on skill demand series with GNN-based methods.

	Model	Market	Region	L1-O	L2-O	R&L1-O	R&L2-O	Company
MAE	EvolveGCNH	$1053.18_{\pm804.38}$	$30.07_{\pm22.97}$	$16.51_{\pm12.61}$	$5.43_{\pm4.15}$	$2.92_{\pm2.23}$	$1.11_{\pm0.85}$	$0.96_{\pm0.73}$
	EvolveGCNO	$151.13_{\pm115.43}$	$27.01_{\pm20.63}$	$15.4_{\pm11.76}$	$5.04_{\pm3.85}$	$2.95_{\pm2.25}$	$1.21_{\pm0.92}$	$0.89_{\pm0.68}$
	GConvGRU	$570.31_{\pm435.58}$	$92.96_{\pm71.0}$	$46.43_{\pm35.46}$	$11.45_{\pm8.75}$	$5.47_{\pm4.18}$	$1.4_{\pm1.07}$	$1.04_{\pm0.79}$
	TGCN	$741.05_{\pm565.99}$	$96.43_{\pm73.65}$	$55.41_{\pm42.32}$	$13.21_{\pm10.09}$	$6.28_{\pm4.4}$	$1.63_{\pm1.24}$	$1.11_{\pm0.85}$
	GCLSTM	$729.48_{\pm557.15}$	$57.13_{\pm43.63}$	$25.84_{\pm19.74}$	$11.49_{\pm8.78}$	$5.67_{\pm4.33}$	$1.61_{\pm1.23}$	$1.07_{\pm0.82}$
	GConvLSTM	$741.45_{\pm566.29}$	$93.61_{\pm71.5}$	$46.57_{\pm35.57}$	$11.19_{\pm8.55}$	$5.57_{\pm4.25}$	$1.49_{\pm1.14}$	$1.09_{\pm0.83}$
	DyGrEncoder	$732.53_{\pm559.48}$	$92.21_{\pm70.43}$	$46.96_{\pm35.87}$	$12.26_{\pm9.36}$	$5.57_{\pm4.26}$	$1.47_{\pm1.13}$	$1.07_{\pm0.82}$
RMSE	EvolveGCNH	$4998.7_{\pm3817.82}$	$166.34_{\pm127.04}$	$178.55_{\pm136.37}$	$76.14_{\pm58.15}$	$35.12_{\pm26.82}$	$16.97_{\pm12.96}$	$19.33_{\pm14.77}$
	EvolveGCNO	$709.19_{\pm541.65}$	$143.72_{\pm109.77}$	$164.16_{\pm125.38}$	$71.57_{\pm54.67}$	$33.21_{\pm25.37}$	$15.04_{\pm11.49}$	$17.93_{\pm13.7}$
	GConvGRU	$2968.75_{\pm2267.42}$	$579.66_{\pm442.72}$	$442.53_{\pm337.99}$	$170.09_{\pm129.91}$	$81.95_{\pm62.59}$	$31.75_{\pm24.25}$	$30.65_{\pm23.41}$
	TGCN	$3163.46_{\pm2416.13}$	$581.49_{\pm444.12}$	$450.79_{\pm344.29}$	$172.13_{\pm131.47}$	$83.33_{\pm63.64}$	$32.19_{\pm24.56}$	$30.77_{\pm23.5}$
	GCLSTM	$3159.61_{\pm2413.19}$	$497.97_{\pm380.33}$	$371.13_{\pm283.45}$	$170.23_{\pm130.01}$	$82.66_{\pm63.13}$	$32.25_{\pm24.63}$	$30.81_{\pm23.53}$
	GConvLSTM	$3166.41_{\pm2418.38}$	$581.26_{\pm443.94}$	$442.72_{\pm338.13}$	$169.43_{\pm129.4}$	$82.38_{\pm62.92}$	$32.29_{\pm24.66}$	$30.88_{\pm23.58}$
	DyGrEncoder	$3161.61_{\pm2414.72}$	$579.78_{\pm442.81}$	$443.49_{\pm338.72}$	$171.91_{\pm131.3}$	$82.27_{\pm62.83}$	$32.02_{\pm24.46}$	$30.76_{\pm23.49}$
SMAPE(%)	EvolveGCNH	$76.69_{\pm58.58}$	$41.67_{\pm31.83}$	$39.77_{\pm30.38}$	$33.11_{\pm25.29}$	$28.99_{\pm22.14}$	$18.79_{\pm14.35}$	$28.36_{\pm21.66}$
	EvolveGCNO	$45.32_{\pm34.61}$	$56.75_{\pm43.35}$	$46.34_{\pm35.39}$	$35.09_{\pm26.8}$	$29.33_{\pm22.4}$	$33.19_{\pm25.35}$	$25.99_{\pm19.85}$
	GConvGRU	$63.93_{\pm48.82}$	$62.41_{\pm47.67}$	$38.65_{\pm29.52}$	$31.09_{\pm23.74}$	$23.67_{\pm18.08}$	$12.93_{\pm9.88}$	$20.5_{\pm15.66}$
	TGCN	$71.13_{\pm54.32}$	$51.37_{\pm39.23}$	$56.23_{\pm42.95}$	$37.21_{\pm28.42}$	$25.83_{\pm19.73}$	$21.74_{\pm16.6}$	$23.35_{\pm17.83}$
	GCLSTM	$60.49_{\pm46.2}$	$38.68_{\pm29.54}$	$43.6_{\pm33.3}$	$28.15_{\pm21.5}$	$24.34_{\pm18.59}$	$18.57_{\pm14.18}$	$21.74_{\pm16.6}$
	GConvLSTM	$64.79_{\pm49.49}$	$44.77_{\pm34.2}$	$35.9_{\pm27.42}$	$26.21_{\pm20.02}$	$21.67_{\pm16.55}$	$13.95_{\pm10.66}$	$23.05_{\pm17.61}$
	DyGrEncoder	$61.55_{\pm47.01}$	$44.74_{\pm34.17}$	$38.12_{\pm29.11}$	$26.94_{\pm20.58}$	$26.49_{\pm20.23}$	$18.42_{\pm14.07}$	$23.75_{\pm18.14}$
RRMSE(%)	EvolveGCNH	$68.68_{\pm52.46}$	$19.57_{\pm14.95}$	$32.99_{\pm25.2}$	$36.53_{\pm27.9}$	$32.1_{\pm24.52}$	$41.01_{\pm31.32}$	$65.93_{\pm50.35}$
	EvolveGCNO	$16.21_{\pm12.38}$	$17.85_{\pm13.63}$	$28.93_{\pm22.09}$	$34.6_{\pm26.43}$	$30.4_{\pm23.22}$	$32.27_{\pm24.64}$	$58.07_{\pm44.35}$
	GConvGRU	$418.84_{\pm319.89}$	$989.03_{\pm755.38}$	$1111.69_{\pm849.06}$	$701.09_{\pm535.46}$	$414.17_{\pm316.33}$	$307.07_{\pm234.53}$	$385.69_{\pm294.58}$
	TGCN	$3478.62_{\pm2656.84}$	$1037.13_{\pm792.12}$	$112164.83_{\pm85667.3}$	$948.04_{\pm724.08}$	$501.07_{\pm382.7}$	$338.75_{\pm258.73}$	$404.5_{\pm308.94}$
	GCLSTM	$3083.62_{\pm2355.15}$	$176.53_{\pm134.83}$	$159.96_{\pm122.17}$	$721.9_{\pm551.36}$	$467.49_{\pm357.05}$	$334.26_{\pm255.3}$	$425.46_{\pm324.95}$
	GConvLSTM	$4479.9_{\pm3421.58}$	$1123.32_{\pm857.95}$	$1101.0_{\pm840.9}$	$640.31_{\pm489.05}$	$449.12_{\pm343.02}$	$382.23_{\pm291.94}$	$451.71_{\pm345.0}$
	DyGrEncoder	$3386.49_{\pm2586.47}$	$1010.93_{\pm772.11}$	$1247.51_{\pm952.8}$	$1004.62_{\pm767.29}$	$450.57_{\pm344.13}$	$331.31_{\pm253.04}$	$418.67_{\pm319.77}$

Table 10: Performance comparisons on skill demand series with structural breaks with GNN-based methods.

	Model	Market	Region	L1-O	L2-O	R&L1-O	R&L2-O	Company
MAE	EvolveGCNH	$1385.2_{\pm1057.96}$	$61.79_{\pm47.19}$	$58.05_{\pm44.33}$	$36.04_{\pm27.53}$	$10.69_{\pm8.17}$	$5.56_{\pm4.25}$	$13.25_{\pm10.12}$
	EvolveGCNO	$195.87_{\pm149.6}$	$53.86_{\pm41.14}$	$52.46_{\pm40.07}$	$32.76_{\pm25.02}$	$10.85_{\pm8.28}$	$5.21_{\pm3.98}$	$12.22_{\pm9.33}$
	GConvGRU	$749.83_{\pm572.69}$	$204.83_{\pm156.44}$	$189.48_{\pm144.72}$	$92.29_{\pm70.49}$	$26.94_{\pm20.58}$	$9.23_{\pm7.05}$	$17.26_{\pm13.18}$
	TGCN	$990.21_{\pm756.28}$	$210.7_{\pm160.92}$	$222.48_{\pm169.92}$	$100.93_{\pm77.09}$	$29.27_{\pm22.36}$	$9.93_{\pm7.58}$	$17.87_{\pm13.65}$
	GCLSTM	$977.98_{\pm746.94}$	$128.33_{\pm98.02}$	$102.44_{\pm78.24}$	$92.71_{\pm70.81}$	$27.94_{\pm21.34}$	$10.25_{\pm7.83}$	$17.81_{\pm13.61}$
	GConvLSTM	$993.55_{\pm758.84}$	$207.86_{\pm158.76}$	$190.31_{\pm145.35}$	$90.33_{\pm68.99}$	$27.51_{\pm21.01}$	$9.93_{\pm7.58}$	$18.0_{\pm13.75}$
	DyGrEncoder	$982.05_{\pm750.06}$	$204.89_{\pm156.49}$	$191.91_{\pm146.57}$	$98.78_{\pm75.44}$	$27.41_{\pm20.93}$	$9.6_{\pm7.33}$	$17.51_{\pm13.38}$
RMSE	EvolveGCNH	$5799.81_{\pm4429.68}$	$261.83_{\pm199.97}$	$371.99_{\pm284.11}$	$227.88_{\pm174.05}$	$87.98_{\pm67.2}$	$29.87_{\pm22.82}$	$100.77_{\pm76.96}$
	EvolveGCNO	$863.11_{\pm659.21}$	$226.98_{\pm173.36}$	$341.34_{\pm260.7}$	$213.99_{\pm163.44}$	$82.25_{\pm62.82}$	$25.03_{\pm19.11}$	$93.36_{\pm71.3}$
	GConvGRU	$3463.23_{\pm2645.08}$	$890.59_{\pm680.2}$	$922.56_{\pm704.62}$	$508.61_{\pm388.46}$	$198.01_{\pm151.24}$	$69.96_{\pm53.43}$	$158.49_{\pm121.05}$
	TGCN	$3688.76_{\pm2817.34}$	$893.17_{\pm682.17}$	$938.93_{\pm717.12}$	$514.25_{\pm392.77}$	$201.03_{\pm153.54}$	$71.44_{\pm54.56}$	$159.07_{\pm121.49}$
	GCLSTM	$3684.35_{\pm2813.97}$	$768.16_{\pm586.69}$	$776.42_{\pm593.0}$	$508.99_{\pm388.74}$	$199.63_{\pm152.47}$	$71.71_{\pm54.77}$	$159.26_{\pm121.64}$
	GConvLSTM	$3692.17_{\pm2819.94}$	$892.93_{\pm681.98}$	$922.93_{\pm704.9}$	$506.71_{\pm387.0}$	$198.99_{\pm151.98}$	$71.93_{\pm54.94}$	$159.59_{\pm121.89}$
	DyGrEncoder	$3686.65_{\pm2815.72}$	$890.76_{\pm680.33}$	$924.49_{\pm706.09}$	$513.77_{\pm392.4}$	$198.73_{\pm151.79}$	$70.95_{\pm54.19}$	$159.01_{\pm121.44}$
SMAPE(%)	EvolveGCNH	$77.85_{\pm59.46}$	$31.5_{\pm24.06}$	$30.67_{\pm23.43}$	$33.05_{\pm25.24}$	$33.43_{\pm25.54}$	$29.61_{\pm22.61}$	$38.66_{\pm29.53}$
	EvolveGCNO	$31.65_{\pm24.17}$	$33.88_{\pm25.88}$	$30.29_{\pm23.14}$	$31.57_{\pm24.11}$	$34.17_{\pm26.1}$	$37.43_{\pm28.59}$	$37.69_{\pm28.78}$
	GConvGRU	$49.4_{\pm37.73}$	$44.31_{\pm33.84}$	$44.6_{\pm34.06}$	$38.91_{\pm29.72}$	$34.25_{\pm26.16}$	$27.57_{\pm21.06}$	$37.43_{\pm28.59}$
	TGCN	$71.56_{\pm54.65}$	$49.2_{\pm37.58}$	$120.31_{\pm91.89}$	$48.37_{\pm36.94}$	$36.76_{\pm28.08}$	$32.09_{\pm24.51}$	$39.02_{\pm29.8}$
	GCLSTM	$63.9_{\pm48.8}$	$31.28_{\pm23.89}$	$28.84_{\pm22.03}$	$39.16_{\pm29.91}$	$35.03_{\pm26.76}$	$32.18_{\pm24.58}$	$37.87_{\pm28.93}$
	GConvLSTM	$68.73_{\pm52.49}$	$45.39_{\pm34.66}$	$44.43_{\pm33.94}$	$38.47_{\pm29.38}$	$33.75_{\pm25.77}$	$28.62_{\pm21.86}$	$37.96_{\pm28.99}$
	DyGrEncoder	$64.22_{\pm49.05}$	$43.72_{\pm33.39}$	$44.73_{\pm34.16}$	$43.73_{\pm34.1}$	$35.07_{\pm27.29}$	$30.42_{\pm23.23}$	$37.41_{\pm28.57}$
RRMSE(%)	EvolveGCNH	$68.88_{\pm52.61}$	$20.26_{\pm15.47}$	$33.11_{\pm25.29}$	$36.88_{\pm28.17}$	$35.03_{\pm26.76}$	$26.91_{\pm20.55}$	$68.99_{\pm52.69}$
	EvolveGCNO	$17.01_{\pm12.99}$	$18.47_{\pm14.11}$	$28.97_{\pm22.12}$	$34.88_{\pm26.64}$	$32.23_{\pm24.61}$	$20.17_{\pm15.41}$	$59.82_{\pm45.69}$
	GConvGRU	$426.16_{\pm325.49}$	$1072.11_{\pm818.84}$	$1215.75_{\pm928.55}$	$763.66_{\pm583.25}$	$485.64_{\pm370.91}$	$255.02_{\pm194.77}$	$427.8_{\pm326.74}$
	TGCN	$3708.55_{\pm2832.45}$	$1159.31_{\pm885.44}$	$156814.93_{\pm119769.38}$	$1119.75_{\pm855.22}$	$621.0_{\pm474.3}$	$290.69_{\pm222.02}$	$451.93_{\pm345.17}$
	GCLSTM	$3210.27_{\pm2451.88}$	$184.79_{\pm141.14}$	$164.56_{\pm125.68}$	$786.29_{\pm600.54}$	$553.67_{\pm422.87}$	$291.57_{\pm222.69}$	$475.27_{\pm362.99}$
	GConvLSTM	$4672.71_{\pm3568.84}$	$1217.46_{\pm929.85}$	$1206.73_{\pm921.66}$	$695.01_{\pm530.83}$	$527.58_{\pm402.95}$	$326.17_{\pm249.12}$	$505.82_{\pm386.33}$
	DyGrEncoder	$3526.47_{\pm2693.38}$	$1093.03_{\pm834.81}$	$1365.47_{\pm1042.9}$	$1110.66_{\pm848.28}$	$525.05_{\pm401.02}$	$279.39_{\pm213.39}$	$465.82_{\pm355.78}$

(LSTMs). The encoder captures temporal dynamics in an evolving graph, and the decoder reconstructs the dynamics using the encoded representation.

We implement these benchmark models using the PyG library [3] and demonstrate the effectiveness of these GNN-based methods in skill demand forecasting.

Results We have implemented a series of graph-based multivariate time series forecasting methods based on the co-occurrence graph and verified their experimental effects under the three scenarios discussed above. Firstly, Table 9 presents the overall performance of the methods based on the co-occurrence graph for skill demand forecasting. It is observed that the prediction accuracy significantly

---

[3] https://github.com/benedekrozemberczki/pytorch_geometric_temporal

21

Table 11: Performance comparisons on low-frequency skill demand series with GNN-based methods.

	Model	Market	Region	L1-O	L2-O	R&L1-O	R&L2-O	Company
MAE	EvolveGCNH	$35.53_{\pm27.13}$	$2.28_{\pm1.74}$	$1.47_{\pm1.13}$	$0.55_{\pm0.42}$	$0.31_{\pm0.23}$	$0.14_{\pm0.11}$	$0.19_{\pm0.15}$
	EvolveGCNO	$27.87_{\pm21.29}$	$3.35_{\pm2.56}$	$1.69_{\pm1.29}$	$0.57_{\pm0.43}$	$0.31_{\pm0.24}$	$0.27_{\pm0.21}$	$0.17_{\pm0.13}$
	GConvGRU	$63.05_{\pm48.16}$	$3.06_{\pm2.34}$	$0.25_{\pm0.19}$	$0.18_{\pm0.14}$	$0.11_{\pm0.09}$	$0.05_{\pm0.04}$	$0.11_{\pm0.09}$
	TGCN	$7.37_{\pm5.63}$	$2.16_{\pm1.65}$	$0.13_{\pm0.1}$	$0.61_{\pm0.46}$	$0.26_{\pm0.2}$	$0.15_{\pm0.12}$	$0.15_{\pm0.11}$
	GCLSTM	$0.62_{\pm0.47}$	$0.41_{\pm0.32}$	$1.37_{\pm1.04}$	$0.15_{\pm0.11}$	$0.12_{\pm0.09}$	$0.09_{\pm0.07}$	$0.12_{\pm0.09}$
	GConvLSTM	$0.79_{\pm0.6}$	$0.43_{\pm0.33}$	$0.2_{\pm0.15}$	$0.13_{\pm0.1}$	$0.1_{\pm0.08}$	$0.06_{\pm0.05}$	$0.13_{\pm0.1}$
	DyGrEncoder	$1.04_{\pm0.79}$	$0.47_{\pm0.36}$	$0.24_{\pm0.18}$	$0.13_{\pm0.1}$	$0.15_{\pm0.11}$	$0.09_{\pm0.07}$	$0.14_{\pm0.11}$
RMSE	EvolveGCNH	$144.53_{\pm110.39}$	$11.08_{\pm8.46}$	$11.65_{\pm8.9}$	$5.52_{\pm4.22}$	$3.48_{\pm2.66}$	$1.7_{\pm1.3}$	$1.81_{\pm1.38}$
	EvolveGCNO	$83.5_{\pm63.77}$	$10.36_{\pm7.91}$	$12.82_{\pm9.79}$	$5.74_{\pm4.38}$	$3.8_{\pm2.9}$	$1.95_{\pm1.49}$	$1.65_{\pm1.26}$
	GConvGRU	$63.06_{\pm48.16}$	$3.65_{\pm2.79}$	$1.19_{\pm0.91}$	$0.87_{\pm0.66}$	$1.59_{\pm1.22}$	$0.81_{\pm0.62}$	$1.72_{\pm1.31}$
	TGCN	$21.13_{\pm16.14}$	$8.89_{\pm6.79}$	$1.13_{\pm0.86}$	$3.12_{\pm2.38}$	$2.32_{\pm1.77}$	$1.11_{\pm0.85}$	$1.78_{\pm1.36}$
	GCLSTM	$1.63_{\pm1.24}$	$2.48_{\pm1.89}$	$2.8_{\pm2.14}$	$0.86_{\pm0.66}$	$1.61_{\pm1.23}$	$0.83_{\pm0.63}$	$1.73_{\pm1.32}$
	GConvLSTM	$1.23_{\pm0.94}$	$2.1_{\pm1.6}$	$1.14_{\pm0.87}$	$0.85_{\pm0.65}$	$1.6_{\pm1.22}$	$0.81_{\pm0.62}$	$1.74_{\pm1.33}$
	DyGrEncoder	$1.49_{\pm1.14}$	$2.12_{\pm1.62}$	$1.15_{\pm0.88}$	$0.86_{\pm0.66}$	$1.59_{\pm1.22}$	$0.81_{\pm0.62}$	$1.73_{\pm1.32}$
SMAPE(%)	EvolveGCNH	$47.39_{\pm36.19}$	$42.25_{\pm32.27}$	$34.5_{\pm26.35}$	$25.49_{\pm19.47}$	$20.27_{\pm15.48}$	$12.66_{\pm9.67}$	$21.3_{\pm16.27}$
	EvolveGCNO	$115.04_{\pm87.86}$	$89.67_{\pm68.48}$	$48.24_{\pm36.84}$	$28.83_{\pm22.02}$	$20.7_{\pm15.81}$	$29.63_{\pm22.63}$	$19.09_{\pm14.58}$
	GConvGRU	$131.13_{\pm100.15}$	$105.97_{\pm80.93}$	$26.95_{\pm20.58}$	$23.04_{\pm17.6}$	$13.04_{\pm9.96}$	$6.02_{\pm4.6}$	$11.92_{\pm9.1}$
	TGCN	$83.42_{\pm63.71}$	$48.87_{\pm37.32}$	$11.74_{\pm8.97}$	$28.63_{\pm21.87}$	$14.45_{\pm11.04}$	$16.13_{\pm12.32}$	$15.27_{\pm11.67}$
	GCLSTM	$42.8_{\pm32.69}$	$36.75_{\pm28.07}$	$43.27_{\pm33.05}$	$17.57_{\pm13.42}$	$14.03_{\pm10.72}$	$12.17_{\pm9.3}$	$13.44_{\pm10.26}$
	GConvLSTM	$53.85_{\pm41.13}$	$36.93_{\pm28.2}$	$21.67_{\pm16.55}$	$14.71_{\pm11.24}$	$10.07_{\pm7.69}$	$7.15_{\pm5.46}$	$14.91_{\pm11.39}$
	DyGrEncoder	$62.53_{\pm47.76}$	$39.84_{\pm30.43}$	$26.26_{\pm20.06}$	$15.61_{\pm11.92}$	$17.33_{\pm13.23}$	$12.71_{\pm9.7}$	$16.23_{\pm12.39}$
RRMSE(%)	EvolveGCNH	$66.67_{\pm50.92}$	$67.56_{\pm51.6}$	$66.87_{\pm51.08}$	$67.29_{\pm51.39}$	$73.11_{\pm55.84}$	$73.29_{\pm55.97}$	$138.14_{\pm105.51}$
	EvolveGCNO	$66.67_{\pm50.92}$	$67.59_{\pm51.63}$	$66.81_{\pm51.03}$	$67.22_{\pm51.34}$	$69.95_{\pm53.43}$	$71.57_{\pm54.66}$	$119.39_{\pm91.18}$
	GConvGRU	$66.46_{\pm50.76}$	$75.73_{\pm57.84}$	$122.73_{\pm93.74}$	$153.14_{\pm116.96}$	$297.91_{\pm227.54}$	$242.77_{\pm185.42}$	$241.43_{\pm184.39}$
	TGCN	$66.51_{\pm50.8}$	$68.02_{\pm51.95}$	$676.51_{\pm516.7}$	$68.73_{\pm52.5}$	$89.88_{\pm68.65}$	$92.87_{\pm70.93}$	$190.19_{\pm145.26}$
	GCLSTM	$60.79_{\pm46.43}$	$103.27_{\pm78.88}$	$70.15_{\pm53.58}$	$169.24_{\pm129.26}$	$359.25_{\pm274.38}$	$214.34_{\pm163.7}$	$272.3_{\pm207.97}$
	GConvLSTM	$65.93_{\pm50.36}$	$189.03_{\pm144.38}$	$156.39_{\pm119.44}$	$166.51_{\pm127.18}$	$345.76_{\pm264.08}$	$287.94_{\pm219.92}$	$285.4_{\pm217.98}$
	DyGrEncoder	$60.35_{\pm46.09}$	$171.99_{\pm131.36}$	$167.79_{\pm128.15}$	$171.6_{\pm131.06}$	$325.47_{\pm248.58}$	$279.07_{\pm213.14}$	$269.12_{\pm205.54}$

declines across the overall labor market. However, as the granularity of the forecast becomes finer, the model performance improves, and at a finer granularity, the EvolveGCN method outperforms the state-of-the-art (SOTA) methods mentioned in the main text considerably. We analyze that the finer the granularity, the more accurately the co-occurrence graph reflects the associations between skills, while coarser granularity might introduce excessive noise leading to decreased model performance. The fine-grained co-occurrence graph accurately reflects the interrelationships between skills at different granularities, which aids in enhancing the model's prediction accuracy. Secondly, we find significant differences in the RRMSE metric among these methods, with EvolveGCN showing superior performance because it can learn the evolution of GCN parameter weights over time, thus capturing the evolving dependencies among edges. Therefore, based on the provided co-occurrence graph, it can effectively learn the evolution of skill relationships, which is beneficial for dynamic prediction of skill demand.

For skill demand forecasting in scenarios involving structural breaks, as shown in Table 10, the improvements in the methods based on the co-occurrence graph are greater than those in the overall skill demand forecasting task. This suggests that skills experiencing structural breaks have strong interconnections, and the co-occurrence graph helps the model to identify the patterns of skill demand sequences that are likely to undergo structural breaks, thus further enhancing the prediction effectiveness for this category of skills.

In the task of predicting low-frequency skills, as shown in Table 11, methods like GConvLSTM significantly outperform EvolveGCN. This is due to the sparse observable data for these skills, which leads to sparse connectivity edges on the co-occurrence graph.

D.4 Job Skill Demand Proportion Forecasting

In the main text, we discussed the issue of skill demand prediction. However, consider a scenario where the number of skill postings for a particular occupation is very low, leading to a low demand for that occupational skills. Nevertheless, these skills might constitute a significant portion of the profession's core competencies. Therefore, using skill demand alone may not adequately measure the importance of these skills within the occupation. To address this, we introduce an extended dataset that includes the skill demand propotion. We define the skill demand proportion as:

$$R_{s,t}^i = [R_{s,t,a^i}^i]_{a^i \in \mathcal{A}^i}, \quad R_{s,t,a^i}^i = \frac{\sum_{p \in \mathcal{P}_t} \mathbf{1}(s \in p) \cdot \mathbf{1}(a^i \in p)}{\sum_{p \in \mathcal{P}_t} \mathbf{1}(a^i \in p)}, \tag{5}$$

22

Table 12: Performance comparisons on skill demand proportion forecasting.

	Model	Market	Region	L1-O	L2-O	R&L1-O	R&L2-O	Company
MAE(%)	LSTM	$0.14_{\pm0.0}$	$0.49_{\pm0.01}$	$1.21_{\pm0.03}$	$2.26_{\pm0.05}$	$2.43_{\pm0.03}$	$3.51_{\pm0.06}$	$2.69_{\pm0.03}$
	SegRNN	$0.15_{\pm0.02}$	$1.25_{\pm0.02}$	$2.63_{\pm0.11}$	$4.2_{\pm0.04}$	$6.6_{\pm1.12}$	$10.16_{\pm2.44}$	$5.52_{\pm0.53}$
	CHGH	$0.1_{\pm0.01}$	$0.19_{\pm0.0}$	$0.35_{\pm0.0}$	$0.64_{\pm0.01}$	$0.7_{\pm0.0}$	$0.91_{\pm0.07}$	$1.03_{\pm0.01}$
	PreDyGAE	$0.08_{\pm0.01}$	$0.09_{\pm0.0}$	$0.12_{\pm0.02}$	$0.23_{\pm0.02}$	$0.18_{\pm0.0}$	$0.27_{\pm0.15}$	$0.29_{\pm0.16}$
	Transformer	$0.62_{\pm0.02}$	$3.89_{\pm0.02}$	$10.87_{\pm0.05}$	$20.58_{\pm0.01}$	$22.02_{\pm0.01}$	$31.26_{\pm0.02}$	$23.47_{\pm0.01}$
	Autoformer	$1.29_{\pm0.14}$	$8.84_{\pm0.15}$	$24.5_{\pm0.22}$	$43.99_{\pm0.94}$	$47.53_{\pm0.67}$	$66.32_{\pm1.02}$	$50.02_{\pm0.66}$
	Informer	$0.55_{\pm0.02}$	$3.83_{\pm0.03}$	$10.86_{\pm0.03}$	$20.6_{\pm0.01}$	$22.04_{\pm0.01}$	$31.23_{\pm0.01}$	$23.48_{\pm0.0}$
	Reformer	$0.56_{\pm0.01}$	$3.85_{\pm0.01}$	$10.9_{\pm0.03}$	$20.65_{\pm0.02}$	$22.01_{\pm0.01}$	$31.25_{\pm0.01}$	$23.48_{\pm0.0}$
	FEDformer	$1.42_{\pm0.04}$	$8.58_{\pm0.18}$	$25.11_{\pm0.59}$	$44.15_{\pm0.6}$	$47.07_{\pm0.66}$	$65.95_{\pm0.23}$	$50.0_{\pm0.07}$
	NStransformer	$0.06_{\pm0.0}$	$0.09_{\pm0.0}$	$0.11_{\pm0.0}$	$0.19_{\pm0.0}$	$0.24_{\pm0.0}$	$0.34_{\pm0.0}$	$0.34_{\pm0.0}$
	PatchTST	$0.06_{\pm0.0}$	$0.09_{\pm0.01}$	$0.11_{\pm0.02}$	$0.18_{\pm0.04}$	$0.23_{\pm0.05}$	$0.33_{\pm0.07}$	$0.33_{\pm0.06}$
	DLinear	$0.27_{\pm0.15}$	$1.65_{\pm1.07}$	$4.63_{\pm3.09}$	$8.76_{\pm5.86}$	$9.4_{\pm6.26}$	$13.33_{\pm8.89}$	$10.05_{\pm6.65}$
	TSMixer	$0.85_{\pm0.03}$	$3.15_{\pm0.69}$	$7.57_{\pm1.43}$	$18.88_{\pm0.49}$	$15.12_{\pm3.85}$	$18.76_{\pm2.85}$	$17.77_{\pm1.58}$
	FreTS	$0.12_{\pm0.0}$	$0.31_{\pm0.04}$	$0.69_{\pm0.11}$	$1.26_{\pm0.21}$	$1.37_{\pm0.22}$	$1.94_{\pm0.32}$	$1.54_{\pm0.24}$
	FiLM	$0.06_{\pm0.01}$	$0.09_{\pm0.01}$	$0.12_{\pm0.01}$	$0.19_{\pm0.02}$	$0.25_{\pm0.03}$	$0.35_{\pm0.05}$	$0.34_{\pm0.05}$
	Koopa	$0.06_{\pm0.0}$	$0.08_{\pm0.0}$	$0.09_{\pm0.0}$	$0.15_{\pm0.01}$	$0.2_{\pm0.01}$	$0.27_{\pm0.02}$	$0.28_{\pm0.01}$
RMSE(%)	LSTM	$0.61_{\pm0.01}$	$1.74_{\pm0.04}$	$2.89_{\pm0.08}$	$4.06_{\pm0.09}$	$4.16_{\pm0.05}$	$5.21_{\pm0.1}$	$4.43_{\pm0.05}$
	SegRNN	$1.32_{\pm0.17}$	$4.98_{\pm0.11}$	$6.49_{\pm0.22}$	$7.7_{\pm0.09}$	$11.23_{\pm1.78}$	$14.58_{\pm3.23}$	$9.18_{\pm0.83}$
	CHGH	$0.36_{\pm0.01}$	$0.49_{\pm0.0}$	$0.7_{\pm0.0}$	$1.12_{\pm0.02}$	$1.14_{\pm0.0}$	$2.11_{\pm0.32}$	$2.31_{\pm0.07}$
	PreDyGAE	$0.23_{\pm0.0}$	$0.29_{\pm0.0}$	$0.36_{\pm0.0}$	$0.85_{\pm0.0}$	$0.91_{\pm0.0}$	$1.15_{\pm1.05}$	$1.29_{\pm0.72}$
	Transformer	$6.22_{\pm0.12}$	$16.65_{\pm0.08}$	$27.9_{\pm0.14}$	$38.45_{\pm0.03}$	$39.77_{\pm0.01}$	$47.4_{\pm0.02}$	$40.97_{\pm0.01}$
	Autoformer	$15.11_{\pm2.51}$	$38.94_{\pm0.61}$	$63.72_{\pm0.49}$	$82.56_{\pm1.83}$	$86.21_{\pm1.22}$	$100.99_{\pm1.58}$	$87.83_{\pm1.13}$
	Informer	$5.63_{\pm0.15}$	$16.41_{\pm0.13}$	$27.89_{\pm0.05}$	$38.5_{\pm0.03}$	$39.79_{\pm0.01}$	$47.36_{\pm0.01}$	$40.98_{\pm0.0}$
	Reformer	$5.62_{\pm0.05}$	$16.45_{\pm0.05}$	$27.96_{\pm0.11}$	$38.58_{\pm0.02}$	$39.76_{\pm0.03}$	$47.38_{\pm0.02}$	$40.98_{\pm0.02}$
	FEDformer	$17.24_{\pm0.68}$	$37.52_{\pm0.77}$	$64.93_{\pm1.42}$	$82.85_{\pm1.16}$	$85.38_{\pm1.21}$	$100.35_{\pm0.34}$	$87.74_{\pm0.1}$
	NStransformer	$0.21_{\pm0.0}$	$0.28_{\pm0.0}$	$0.35_{\pm0.0}$	$0.77_{\pm0.01}$	$0.86_{\pm0.0}$	$1.69_{\pm0.0}$	$1.24_{\pm0.0}$
	PatchTST	$0.21_{\pm0.01}$	$0.28_{\pm0.02}$	$0.35_{\pm0.03}$	$0.75_{\pm0.04}$	$0.84_{\pm0.06}$	$1.65_{\pm0.1}$	$1.22_{\pm0.07}$
	DLinear	$2.21_{\pm1.49}$	$6.05_{\pm4.1}$	$10.27_{\pm6.97}$	$14.16_{\pm9.59}$	$14.64_{\pm9.91}$	$17.51_{\pm11.77}$	$15.11_{\pm10.19}$
	TSMixer	$10.5_{\pm0.04}$	$14.32_{\pm2.43}$	$21.09_{\pm3.64}$	$34.92_{\pm1.37}$	$25.27_{\pm5.66}$	$26.48_{\pm3.26}$	$31.07_{\pm2.82}$
	FreTS	$0.47_{\pm0.02}$	$0.96_{\pm0.07}$	$1.56_{\pm0.12}$	$2.22_{\pm0.17}$	$2.29_{\pm0.17}$	$2.99_{\pm0.2}$	$2.54_{\pm0.17}$
	FiLM	$0.21_{\pm0.0}$	$0.28_{\pm0.0}$	$0.35_{\pm0.02}$	$0.76_{\pm0.04}$	$0.85_{\pm0.06}$	$1.67_{\pm0.11}$	$1.22_{\pm0.07}$
	Koopa	$0.21_{\pm0.0}$	$0.27_{\pm0.0}$	$0.33_{\pm0.01}$	$0.72_{\pm0.01}$	$0.79_{\pm0.02}$	$1.57_{\pm0.04}$	$1.17_{\pm0.01}$
SMAPE(%)	LSTM	$0.26_{\pm0.0}$	$0.92_{\pm0.02}$	$2.25_{\pm0.05}$	$4.21_{\pm0.09}$	$4.52_{\pm0.05}$	$6.5_{\pm0.11}$	$5.0_{\pm0.05}$
	SegRNN	$0.27_{\pm0.03}$	$2.09_{\pm0.03}$	$4.55_{\pm0.17}$	$7.38_{\pm0.06}$	$11.1_{\pm1.63}$	$16.79_{\pm3.54}$	$9.58_{\pm0.83}$
	CHGH	$0.19_{\pm0.01}$	$0.36_{\pm0.0}$	$0.69_{\pm0.0}$	$1.24_{\pm0.02}$	$1.36_{\pm0.0}$	$1.78_{\pm0.01}$	$1.77_{\pm0.06}$
	PreDyGAE	$0.09_{\pm0.0}$	$0.19_{\pm0.0}$	$0.22_{\pm0.0}$	$0.28_{\pm0.0}$	$0.39_{\pm0.0}$	$0.64_{\pm0.27}$	$0.57_{\pm0.3}$
	Transformer	$0.81_{\pm0.02}$	$4.78_{\pm0.01}$	$13.28_{\pm0.04}$	$25.1_{\pm0.0}$	$26.89_{\pm0.01}$	$38.12_{\pm0.01}$	$28.71_{\pm0.01}$
	Autoformer	$1.13_{\pm0.02}$	$7.42_{\pm0.05}$	$20.74_{\pm0.1}$	$38.35_{\pm0.38}$	$41.28_{\pm0.29}$	$58.08_{\pm0.43}$	$43.84_{\pm0.29}$
	Informer	$0.75_{\pm0.02}$	$4.74_{\pm0.02}$	$13.27_{\pm0.03}$	$25.12_{\pm0.0}$	$26.9_{\pm0.01}$	$38.1_{\pm0.01}$	$28.72_{\pm0.0}$
	Reformer	$0.75_{\pm0.01}$	$4.75_{\pm0.01}$	$13.31_{\pm0.02}$	$25.16_{\pm0.02}$	$26.87_{\pm0.01}$	$38.12_{\pm0.01}$	$28.72_{\pm0.0}$
	FEDformer	$1.17_{\pm0.01}$	$7.35_{\pm0.08}$	$21.02_{\pm0.24}$	$38.42_{\pm0.24}$	$41.09_{\pm0.28}$	$57.94_{\pm0.1}$	$43.84_{\pm0.03}$
	NStransformer	$0.11_{\pm0.0}$	$0.17_{\pm0.0}$	$0.22_{\pm0.0}$	$0.35_{\pm0.0}$	$0.45_{\pm0.0}$	$0.61_{\pm0.0}$	$0.63_{\pm0.0}$
	PatchTST	$0.11_{\pm0.01}$	$0.17_{\pm0.02}$	$0.21_{\pm0.05}$	$0.34_{\pm0.09}$	$0.44_{\pm0.1}$	$0.59_{\pm0.14}$	$0.61_{\pm0.11}$
	DLinear	$0.43_{\pm0.19}$	$2.58_{\pm1.39}$	$7.19_{\pm4.0}$	$13.59_{\pm7.59}$	$14.59_{\pm8.11}$	$20.67_{\pm11.52}$	$15.62_{\pm8.62}$
	TSMixer	$0.92_{\pm0.03}$	$4.05_{\pm0.67}$	$10.03_{\pm1.36}$	$23.97_{\pm1.21}$	$21.46_{\pm4.26}$	$27.6_{\pm3.5}$	$23.83_{\pm1.45}$
	FreTS	$0.22_{\pm0.01}$	$0.59_{\pm0.07}$	$1.31_{\pm0.21}$	$2.4_{\pm0.41}$	$2.63_{\pm0.43}$	$3.69_{\pm0.62}$	$2.95_{\pm0.46}$
	FiLM	$0.11_{\pm0.02}$	$0.17_{\pm0.02}$	$0.22_{\pm0.02}$	$0.36_{\pm0.03}$	$0.46_{\pm0.06}$	$0.63_{\pm0.09}$	$0.64_{\pm0.09}$
	Koopa	$0.1_{\pm0.0}$	$0.15_{\pm0.01}$	$0.18_{\pm0.01}$	$0.28_{\pm0.01}$	$0.37_{\pm0.01}$	$0.49_{\pm0.03}$	$0.52_{\pm0.03}$
RRMSE(%)	LSTM	$43.76_{\pm0.35}$	$81.48_{\pm0.6}$	$85.53_{\pm0.64}$	$91.34_{\pm0.35}$	$93.23_{\pm0.15}$	$97.48_{\pm0.09}$	$96.6_{\pm0.07}$
	SegRNN	$72.41_{\pm4.5}$	$96.92_{\pm0.17}$	$96.4_{\pm0.21}$	$96.99_{\pm0.11}$	$98.79_{\pm0.46}$	$99.24_{\pm0.25}$	$98.15_{\pm0.28}$
	CHGH	$27.4_{\pm0.7}$	$37.66_{\pm0.07}$	$37.45_{\pm0.03}$	$53.25_{\pm0.58}$	$58.32_{\pm0.03}$	$80.91_{\pm5.07}$	$70.89_{\pm2.0}$
	PreDyGAE	$20.64_{\pm0.0}$	$28.83_{\pm0.0}$	$40.02_{\pm0.01}$	$40.12_{\pm0.0}$	$47.5_{\pm0.0}$	$76.9_{\pm42.81}$	$66.68_{\pm33.49}$
	Transformer	$97.99_{\pm0.07}$	$99.73_{\pm0.01}$	$99.8_{\pm0.0}$	$99.89_{\pm0.0}$	$99.92_{\pm0.0}$	$99.97_{\pm0.0}$	$99.96_{\pm0.0}$
	Autoformer	$99.63_{\pm0.12}$	$99.95_{\pm0.0}$	$99.96_{\pm0.0}$	$99.97_{\pm0.0}$	$99.98_{\pm0.0}$	$99.99_{\pm0.0}$	$99.98_{\pm0.0}$
	Informer	$97.59_{\pm0.12}$	$99.72_{\pm0.0}$	$99.8_{\pm0.0}$	$99.89_{\pm0.0}$	$99.92_{\pm0.0}$	$99.97_{\pm0.0}$	$99.96_{\pm0.0}$
	Reformer	$97.57_{\pm0.03}$	$99.72_{\pm0.01}$	$99.81_{\pm0.0}$	$99.89_{\pm0.0}$	$99.92_{\pm0.0}$	$99.97_{\pm0.0}$	$99.96_{\pm0.0}$
	FEDformer	$99.75_{\pm0.02}$	$99.95_{\pm0.0}$	$99.96_{\pm0.0}$	$99.97_{\pm0.0}$	$99.98_{\pm0.0}$	$99.99_{\pm0.0}$	$99.98_{\pm0.0}$
	NStransformer	$16.44_{\pm0.44}$	$22.17_{\pm0.31}$	$19.66_{\pm0.16}$	$37.34_{\pm0.26}$	$45.81_{\pm0.26}$	$75.98_{\pm0.13}$	$61.84_{\pm0.13}$
	PatchTST	$16.32_{\pm0.88}$	$21.85_{\pm1.31}$	$19.5_{\pm1.44}$	$36.66_{\pm1.59}$	$44.97_{\pm1.95}$	$75.59_{\pm0.49}$	$61.55_{\pm1.34}$
	DLinear	$78.33_{\pm15.19}$	$95.41_{\pm3.54}$	$96.69_{\pm2.57}$	$97.94_{\pm1.58}$	$98.55_{\pm1.11}$	$99.37_{\pm0.44}$	$98.87_{\pm0.84}$
	TSMixer	$99.29_{\pm0.02}$	$99.61_{\pm0.13}$	$99.64_{\pm0.11}$	$99.86_{\pm0.02}$	$99.8_{\pm0.11}$	$99.85_{\pm0.0}$	$99.91_{\pm0.05}$
	FreTS	$34.81_{\pm1.43}$	$61.32_{\pm3.03}$	$66.35_{\pm3.08}$	$77.25_{\pm2.53}$	$81.41_{\pm2.37}$	$92.32_{\pm1.5}$	$90.05_{\pm1.64}$
	FiLM	$16.49_{\pm0.39}$	$21.81_{\pm0.76}$	$19.48_{\pm0.05}$	$36.44_{\pm0.97}$	$44.79_{\pm2.15}$	$75.91_{\pm10.46}$	$61.27_{\pm8.05}$
	Koopa	$16.47_{\pm0.04}$	$21.47_{\pm0.18}$	$18.71_{\pm0.07}$	$35.58_{\pm0.07}$	$43.65_{\pm0.38}$	$75.85_{\pm2.58}$	$61.26_{\pm1.98}$

where $a^i \in p$ represents a job advertisement p containing the attribute a^i under granularity i. Similarly, we can further define skill demand proportions $R_{s,t}^{i,j,...,k}$ across multiple granularities $\{i, j, ..., k\}$ by

23

calculating:

$$R_{s,t,\overline{a}}^{i,j,...,k} = \frac{\sum_{p \in \mathcal{P}_t} \mathbf{1}(s \in p) \cdot \mathbf{1}(a^i \in p \wedge a^j \in p \wedge ... \wedge a^k \in p)}{\sum_{p \in \mathcal{P}_t} \mathbf{1}(a^i \in p \wedge a^j \in p \wedge ... \wedge a^k \in p)}, \qquad (6)$$

where $\overline{a} = \{a^i, a^j, ..., a^k\}$, $a^i \in \mathcal{A}^i$, $a^j \in \mathcal{A}^j$, ..., $a^k \in \mathcal{A}^k$, and $R_{s,t}^{i,j,...,k} \in \mathbb{R}^{|\mathcal{A}^i||\mathcal{A}^j|...|\mathcal{A}^k|}$.

Results We continue to utilize the benchmark models described in the main text for this task, and the results, as shown in Table 12, lead to the following conclusions: Firstly, the best-performing model on the task of forecasting the proportion of skill demand is Koopa. This model, integrating time series decomposition and Fourier transformations, effectively captures the distribution changes in demand proportions. Secondly, there is a significant variation in performance across models in this task. For example, models like DLinear perform poorly on this task, though they are reasonably effective in skill demand forecasting. We analyze that predicting percentages is distinct from forecasting skill demand, as percentage predictions are also influenced by the demand for other skills at the same granularity. Therefore, simple linear models are not advantageous for capturing the complex interrelations and influences among multiple pieces of information.

E Data Structure and Components

Our dataset comprises five components for each granularity level: job skill demand sequences, job skill demand proportion sequences, ID mapping index, the indexes of skills with structural breaks, and skill co-occurrence graph. Each component is structured as follows: (1) *Job Skill Demand Sequences:* These are presented in tabular files, where each row represents a specific skill, and each column corresponds to a different time slice (month). Each cell within the table contains a numerical value that reflects the demand for the respective skill during that month. (ii) *Job Skill Demand Proportion Sequences:* This component is also formatted in tabular files similar to the skill demand sequences. However, each cell in these tables displays a value between 0 and 1, representing the proportion of demand defined in Eq 6. This provides a normalized view of skill demand across different granularities. (iii) *ID mapping index:* In the dataset, various elements such as regions, occupations, companies, and skills are represented using unique identifiers (IDs) for the convenience of experimentation and analysis. An index table is provided that maps each ID to the actual names of regions, occupations, and skills, facilitating clear and effective data interpretation. The names of companies, however, are withheld due to potential privacy concerns. This selective anonymization ensures that while company-related data remains confidential, researchers can still access necessary details about other variables. During the review period, we open part of id-to-entity mappings to illustrate its use and structure. (iv) *Indexes of skills with structural breaks:* In the provided dataset, data concerning skills that have experienced structural breaks are organized in JSON format. Each granularity level is represented by a separate JSON file, which contains a list of indexes. These indexes correspond to the skills that have undergone structural breaks and can be directly mapped to the skill indexes in the skill demand sequences. The purpose of supplying this data is to facilitate research on the demand trends of skills that have exhibited structural breaks, enabling a detailed analysis of their demand dynamics over time. (v) *Skill Co-occurrence Graph:* This data is provided as a set of triples (skill ID_1, skill ID_2, frequency of co-occurrence), forming a collection that outlines the co-occurrence relationships between skills. Each triple indicates how frequently two skills are mentioned or required together within the job advertisements in the training data, serving as a prior knowledge graph to enhance predictive modeling by capturing relationships between skills.

F Datasheet for Datasets

F.1 Motivation

- **For what purpose was the dataset created?** The dataset in this paper was created specifically for the task of skill demand forecasting. The creation of this dataset aims to address two primary objectives: first, to fill the existing gap in publicly available datasets for skill demand forecasting; second, to offer a dataset that encompasses multiple levels of granularity. This diversity in granularity enables the possibility of predicting skill trends at various levels of detail. By providing such a comprehensive dataset, researchers can gain a more thorough understanding of and ability to predict changes in skill demand, thereby

24

facilitating job seekers in finding suitable positions and aids recruiters in hiring the right talent. Furthermore, this dataset can support the development of education, training, and employment policies.

- **Who created the dataset (e.g., which team, research group) and on behalf of which entity (e.g., company, institution, organization)?** Our dataset is jointly developed by a collaborative effort from the following affiliations: University of Science and Technology of China, Computer Network Information Center, Chinese Academy of Sciences, Institute of Artificial Intelligence, Beihang University, and Artificial Intelligence Thrust, The Hong Kong University of Science and Technology (Guangzhou).

- **Who funded the creation of the dataset?** This work was supported by University of Science and Technology of China, Computer Network Information Center, Chinese Academy of Sciences, and The Hong Kong University of Science and Technology (Guangzhou).

F.2 Composition

- **What do the instances that comprise the dataset represent (e.g., documents, photos, people, countries)?** See the Section E.

- **How many instances are there in total (of each type, if appropriate)?** We collected millions of public job advertisements, covering 521 companies, 52 occupations, and 7 regions. We further extracted 2335 skills and tracked the monthly demand for these skills, thus ultimately forming sequences of demand for 2335 skills over 36 months, broken down by company, occupation, and region.

- **Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set?** Yes, the dataset encompasses an exhaustive collection of job recruitment demands posted by various companies on online platforms. In the subsequent data processing stages, we employed rule-based matching and Named Entity Recognition (NER) models to extract as many relevant skill demands as possible from each job advertisement. Additionally, we filtered out certain implausible skill terms based on skill frequency analysis and manual review. Therefore, the dataset includes all possible instances except for those terms deemed unreasonable and filtered out during the review process.

- **What data does each instance consist of?** See the Subsection 3.1 and Section E.

- **Is there a label or target associated with each instance?** Yes, given that our dataset is a time series collection, it is structured for unsupervised learning. This means there are no explicit labels or target variables associated with each instance. Instead, the data is used to observe and analyze patterns over time, particularly in skill demand and relationships between skills, without predefined outcomes or classifications.

- **Is any information missing from individual instances?** N/A

- **Are relationships between individual instances made explicit (e.g., user movie ratings, social network links)?** We supplemented our analysis with a co-occurrence graph to explicitly represent the relationships between skills based on their co-occurrence in various contexts.

- **Are there recommended data splits (e.g., training, development/validation, testing)?** Yes, we recommended the data split in Subsection 3.2.

- **Are there any errors, sources of noise, or redundancies in the dataset?** Yes. See the Subsection 3.1

- **Is the dataset self-contained, or does it link to or otherwise rely on external resources (e.g., websites, tweets, other datasets)?** The dataset is self-contained for the tasks described in the paper, which is collected from online websites. a) We can not ensure the linked resources exist for a long time and consistent. b) There is no official archive related to the linked resource. c) There are no restrictions.

- **Does the dataset contain data that might be considered confidential (e.g., data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals' non-public communications)?** No.

- **Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety?** No.

25

F.3 Collection Process

- **How was the data associated with each instance acquired?** See the Section 3.1.
- **What mechanisms or procedures were used to collect the data?** See Section 3.1.
- **If the dataset is a sample from a larger set, what was the sampling strategy (e.g., deterministic, probabilistic with specific sampling probabilities)?** No.
- **Who was involved in the data collection process (e.g., students, crowdworkers, contractors) and how were they compensated (e.g., how much were crowdworkers paid)?** We have collaborated with industry experts who have relevant experience to assist in data annotation for skill dictionary construction. Their participation is voluntary and unpaid.
- **Over what timeframe was the data collected?** The raw data, consisting of job advertisements, was collected over a period in the past three years.
- **Were any ethical review processes conducted (e.g., by an institutional review board)?** No.

F.4 Preprocessing/cleaning/labeling

- **Was any preprocessing/cleaning/labeling of the data done (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)?** Yes. See the Subsection 3.1.
- **Was the "raw" data saved in addition to the preprocessed/cleaned/labeled data (e.g., to support unanticipated future uses)?** No.
- **Is the software that was used to preprocess/clean/label the data available?** No.

F.5 Uses

- **Has the dataset been used for any tasks already?** No.
- **Is there a repository that links to any or all papers or systems that use the dataset?** No.
- **What (other) tasks could the dataset be used for?** N/A
- **Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses?** N/A.
- **Are there tasks for which the dataset should not be used?** N/A.

F.6 Distribution

- **Will the dataset be distributed to third parties outside of the entity (e.g., company, institution, organization) on behalf of which the dataset was created?** No. Our dataset will be managed and maintained by our research group.
- **How will the dataset will be distributed (e.g., tarball on website, API, GitHub)?** The evaluation dataset is released to the public and hosted on GitHub.
- **When will the dataset be distributed?** It has been released now.
- **Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)?** Our dataset will be distributed under the CC BY-SA 4.0 license.
- **Have any third parties imposed IP-based or other restrictions on the data associated with the instances?** No.
- **Do any export controls or other regulatory restrictions apply to the dataset or to individual instances?** No.

F.7 Maintenance

- **Who will be supporting/hosting/maintaining the dataset?** The authors.
- **How can the owner/curator/manager of the dataset be contacted (e.g., email address)?** By the email address.

26

- **Is there an erratum?** N/A.

- **Will the dataset be updated (e.g., to correct labeling errors, add new instances, delete instances)?** Yes, the dataset is expected to be updated with additional job advertisements collected over extended periods, potentially on a biannual basis. These updates will focus on increasing our data volume over time. All updates will be made available and communicated to dataset consumers via GitHub.

- **If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (e.g., were the individuals in question told that their data would be retained for a fixed period of time and then deleted)?** N/A.

- **Will older versions of the dataset continue to be supported/hosted/maintained?** We plan to maintain the newest version only.

- **If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so?** Yes. Contact the authors of the paper.

Checklist

The checklist follows the references. Please read the checklist guidelines carefully for information on how to answer these questions. For each question, change the default **[TODO]** to [Yes] , [No] , or [N/A] . You are strongly encouraged to include a **justification to your answer**, either by referencing the appropriate section of your paper or providing a brief inline description. For example:

- Did you include the license to the code and datasets? [Yes] See Section X.
- Did you include the license to the code and datasets? [N/A]

Please do not modify the questions and only use the provided macros for your answers. Note that the Checklist section does not count towards the page limit. In your paper, please delete this instructions block and only keep the Checklist section heading above along with the questions/answers below.

1. For all authors...

    (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes] Please refer to the Abstract and Introduction sections.

    (b) Did you describe the limitations of your work? [Yes] Please refer to dataset limitation in Section 3.3.

    (c) Did you discuss any potential negative societal impacts of your work? [Yes] Pleas refer to the Appendix A.

    (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes] This work adheres to the NeurIPS Code of Ethics in each respect.

2. If you are including theoretical results...

    (a) Did you state the full set of assumptions of all theoretical results? [N/A]

    (b) Did you include complete proofs of all theoretical results? [N/A]

3. If you ran experiments (e.g. for benchmarks)...

    (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] All the code, data, and instructions needed to reproduce the main experimental results can be found at `https://github.com/Job-SDF/benchmark`.

    (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] We have provided the detailed implementation descriptions in the Appendix.

    (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] Due to the page limit, we report error bars in the Appendix.

    (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes] The experimental environment is reported in the Appendix.

27

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...

    (a) If your work uses existing assets, did you cite the creators? [Yes] We have cited the original paper that produced the code package.

    (b) Did you mention the license of the assets? [Yes] We have stated the detailed version and license of each asset. See the Appendix for the details.

    (c) Did you include any new assets either in the supplemental material or as a URL? [Yes] We have released a new dataset and the codes for the benchmark models.

    (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [Yes] Please refer to the Appendix.

    (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [Yes] The proposed dataset does not contain personally identifiable information or offensive content. See the Appendix for the details.

5. If you used crowdsourcing or conducted research with human subjects...

    (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]

    (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]

    (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

28