Better Skill-based Job Representations, Assessed via Job Transition Data

Conference Paper · December 2022

DOI: 10.1109/BigData55660.2022.10021087

Better Skill-based Job Representations, Assessed via Job Transition Data

Tyler Baldwin
IBM Research
Almaden, USA
tbaldwin@us.ibm.com

Wyatt Clarke
IBM Research
Yorktown Heights, USA
wyatt.clarke@ibm.com

Maysa M. Garcia de Macedo
IBM Research
Sao Paulo, Brazil
mmacedo@br.ibm.com

Rogerio de Paula
IBM Research
Sao Paulo, Brazil
ropaula@br.ibm.com

Subhro Das
MIT-IBM Watson AI Lab, IBM Research
Cambridge, USA
subhro.das@ibm.com

2022 IEEE International Conference on Big Data (Big Data) | 978-1-6654-8045-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/BigData55660.2022.10021087

***Abstract*—Learning never stops for successful workers, who must grow their careers while coping with the changing expectations of employers. Robust job-skill representations can empower workers by helping them to better decipher viable job changes given their current skill set and guide them toward skills they can learn to meet career goals. In this work we combine threads of research in economics and AI to improve upon existing job-skill representation methodology and performance. We build a benchmark dataset of between-job transitions from US Census data and show that a representation trained on a large set of online job postings via a transformer-based architecture outperforms existing baselines. Further analysis demonstrates that this model is better able to transfer across taxonomies than existing models.**

***Index Terms*—NLP, Big Data, Talent Management, Skill Similarity**

I. Introduction

Workers' skills grow and evolve over their lifetime, at the same time as employers' demands for skills are also changing. Navigating a career in this dynamic environment is hard; it threatens to become harder if innovations in AI and robotics quicken the rate of change in skill demands.

While understanding and aiding in the alleviation of this skill pressure has been of interest to researchers in both AI and economics, the approaches to solving these issues have thus far been divergent. In this work we attempt to bring these threads together both in terms of methodology and evaluation. To do so, we establish a challenge to create skill-based job representations that can faithfully reproduce true job transition patterns—and provide a preliminary answer.

From economics, we draw on a body of work on skill similarity and skill space analysis, used to understand the impact of skill trends on the labor market [1]–[4]. From AI, we draw on methods from job recommendation system research and natural language processing to improve performance over the embedding space methods currently prevalent in the economics research. We show that a transformer-based architecture trained over a large set of online job postings is able to build a job-skill representation that is more predictive of true job transitions than existing baseline methods.

In this work, we:

- Advocate for the adoption of an evaluation setting based on job transitions from US Census data, allowing for consistent comparison of job-skill representations.
- Introduce the same-job prediction task as a means of modeling job-skill relationships.
- Train a transformer-based model via this prediction task and show that its resultant job-skill representations better predict real job transition data than existing baselines.
- Show that the proposed model more robustly handles unseen skills and transfers across taxonomies.

II. Related Work

In economics, jobs have long been seen as a cluster of tasks which a worker must perform or a cluster of skills which the worker provides [4], [5]. As such, skill-based representations are often used to get a more granular picture of the labor market than is found by simply examining jobs. Common skill analysis methods include the examination of similarity within an embedded space [6] or graph-based network analysis [7]. Common use cases include examining the impact of particular changes in the labor market, such as the rise of information technology [8]–[10], automation [5], [11], online labor markets [12], or the gig economy [13]. Other work has explored how skill demand changes over time, in terms of seasonal trends [14], secular changes [15], and across occupation and wage groups [16]–[18].

The most common use case of skill representations for practical task completion is in job recommendation systems. The goal of job recommendation work is to match job seekers, often represented by résumé data, with employment opportunities, often represented by job postings. In contrast to the approach presented here, while skills are one important set of features utilized by many of these methods, they are typically not the sole data source. Most methods attempt to

Statisticians	
Job Sample 1	Job Sample 2
Statistical Reporting	Physics
Linear Algebra	Oral Communication
Calculus	Teamwork/Collaboration
Statistics	MATLAB
Data Collection	Systems Engineering
Algebra	Statistics
Differential Equations	Communication Skills
Surveys	

TABLE I: Skill lists for 2 jobs with the same CPS job title.

	NDCG@k		
k	5	10	20
Word2vec	0.368	0.432	0.471
O*NET	0.360	0.429	0.486
Transformer - Job title jobs	0.411	0.469	0.501
Transformer - CPS jobs	**0.435**	**0.496**	**0.532**

TABLE II: Job transition prediction results.

build user models by employing multiple data sources such as employment history, education history, and social media links; for a review of the field see [19] or [20]. Even the most skill-centric methods rely on additional data sources to improve their models [21], [22].

III. Methodology

A. Data

For model development, we utilize a dataset of millions of online job postings from Lightcast[1], collected over the period 2010 to 2020. Job postings in the Lightcast data have been supplemented by a variety of extracted or inferred meta information. For the purposes of this work we are concerned with three pieces of extracted information: the skills each job posting requires, the title of each job posting, and the job code which links each posting to a normalized job from an ontology. The Lightcast data links to O*NET-SOC codes[2], a variation of the US Bureau of Labor Statistics Standard Occupation Classification coding schema which is further augmented by the O*NET ontology. The Lightcast skill taxonomy consists of 15,305 separate skills[3].

For evaluation, we follow previous work [23], [24] in constructing historical counts of transitions between pairs of jobs documented by the US Census' Current Population Survey (CPS) over the past decade. In the CPS, survey respondents are asked their occupation over a sequence of multiple months. When their response changes, it is counted as an occupation change. CPS occupations are coded using the US Census Occupational Classification System (OCC), a set of 516 job codes. For this work, January 2010 to March 2021 CPS microdata files are accessed using IPUMS-CPS [25], which provides some additional data harmonization and cleaning.

B. Task

To discover relationships between skill sets and associated jobs from job posting data, we model the relationship as a same-job prediction task. That is, given two sets of skills, each associated with a single job posting, we train a model to predict whether they refer to the same normalized job. There are many skills in the taxonomy, and an individual posting will have a comparatively small number of them that it requires; jobs in the dataset required slightly over 9 skills on average. As shown in the example *Statisticians* job in Table I, different postings linked to the same normalized job will often require sets of skills with little to no overlap between them. Thus, this task setting allows for the mining of skill relationships that may be non-apparent from strictly overlap-based methods.

What constitutes "same job" is variable depending on the job classification system employed. As job classifications in the wild are often not linked with an existing job taxonomy, we consider two task settings. In one, we consider two job postings to be of the same job if they share the same assigned CPS job ID, converted from the O*NET-SOC code provided within the dataset using a crosswalk provided by the Census Bureau[4]. To examine the fidelity of our task design when job taxonomy linking is not present, we also examine a narrower definition in which two postings are considered to be the same job only when they have the exact same job title.

C. Model

We adopt a transformer-based deep learning architecture [26] to perform the same-job prediction task. Transfer learning from popular transformer-based models such as BERT [27] has led to advances in the state of the art on many NLP tasks. Skill representations share some similarities to natural language, and as such have the potential to benefit from fine-tuning on top of pretrained language models in the same way.

We employ the standard BERT model with a task setting similar to the next sentence prediction task performed in the original BERT paper, but modified for the prediction of skill representations. Skill representations of two job postings are fed concurrently into the model, separated by a separator token. At the end of the architecture, a classifier performs the prediction over the class (CLS) token, which encapsulates an embedding of the two skill representations together. Due to space constraints, we refer the reader to the original transformer [26] and BERT [27] papers for further detail.

The following settings were used in all experiments in this paper. The transformer architecture consisted of 12 total layers with 12 total attention heads. The hidden layer size was 512, dropout probability was 0.1, and the Adam optimizer was used with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. Fine-tuning was performed on 5,341,226 sampled job posting pairs from 2020, balanced to have an even number of positive and negative examples. Experiments were performed on a Unix architecture system with 2 Tesla V100 GPUs with 32 GB of memory apiece.

[1]https://www.lightcast.io/

[2]https://www.onetcenter.org/taxonomy.html

[3]Our data use a legacy skill taxonomy from Burning Glass Technologies, which merged with Emsi in 2021 to form Lightcast.

[4]https://www2.census.gov/programs-surveys/demo/guidance/industry-occupation/2010-occ-codes-with-crosswalk-from-2002-2011.xls

D. Baselines

To evaluate the performance of our model we compare it to two existing baseline methods: (1) a word2vec-based model [28] and (2) a replication of the O*NET Career Changers Matrix [29]. The word2vec method uses a single layer network and a masked language model style objective to embed each skill in a vector space. The word2vec method is commonly used in existing labor market analysis work for extrapolation of skill data to labor trends [8], [30] and as a skill embedding in job recommendation systems [31], [32]. The word2vec model is trained over the entire corpus of job postings from 2010 to 2020, using the skipgram model setting and an embedding size of 200. We perform mean pooling over the individual tokens which make up a job to produce a job-level embedding, which is then compared against other jobs via cosine similarity.

The Career Changers Matrix is a component of O*NET, intended to help job seekers with work experience find related work requiring minimal retraining [29]. It powers My Skills My Future[5], a U.S. Department of Labor-sponsored recommendation engine used in public employment offices. For each occupation in O*NET, expert voters have scored the importance of around 250 aspects of work (e.g., the importance of "Repairing and Maintaining Mechanical Equipment"). The Career Changers Matrix is calculated using the Euclidean distance between a subset of the 250 dimensions, after which human scorers remove some comparison occupations using expert judgement. The rank-ordered 10 nearest neighbors are published for each occupation. As the experiments performed here require more than just the 10 nearest occupations, we replicate the algorithmic portion of the process as a baseline for this exercise.

IV. Results

To assess the performance of the proposed job-skill representation model, we examine its ability to predict job transition data from the CPS survey. Job transitions were discovered by extracting month-over-month job changes reported by individuals in the survey and aggregated over the entire dataset. To eliminate skews from frequently occurring jobs, job transitions were normalized by dividing by their frequency within the dataset. To minimize noise from infrequently occurring data, jobs that appear less than 100 times as the origin and job-to-job transitions that occurred less than 10 times are excluded from the dataset. After these data cleaning steps are performed, a ranked list of the most likely job transitions from a given job (and associated transition probabilities) can be extracted.

To compare predicted job transitions to the true transitions, we employ the normalized discounted cumulative gain at k (NDCG@k) metric commonly used in job recommendation system evaluation. In this context, NDCG@k measures the proportion of predicted transitions in the top k predictions that are also in the top k true transitions. However, unlike similar metrics such as Precision@k, NDCG@k weighs predictions based on their relative importance in the ranking. Discounted cumulative gain (DCG@k) is defined as:

$$DCG = \sum_{i=1}^{k} \frac{2^{rel_i} - 1}{log_2(i+1)} \tag{1}$$

where rel_i is the relevance of a transition, defined here as the normalized transition probability from the CPS ranking. Once DCG is computed, NDCG can be calculated by dividing the computed DCG by the ideal discounted cumulative gain, the highest possible gain at the current value of k.

For both the proposed model and the word2vec baseline, predicted relevance for a job pair is calculated over representations comprising the top 10 most frequent skills associated with that job in the Lightcast data. Word2vec transitions are ranked via the cosine similarity of their embedding vectors, while the prediction confidence is used in the proposed model. The O*NET baseline uses Euclidean distance between expert-assigned ratings on a vector of job characteristics to rank predictions.

Table II shows the results of the transformer model relative to the baselines. Results are given at $k = 5, 10, 20$. As shown, all versions of the proposed model outperform both baselines on all values of k. Training with same job defined using the CPS taxonomy was more effective than using job titles, but the job title based version still outperformed the baseline models.

To further examine the utility of each model, we conduct experiments on taxonomy transfer. Because there is no agreed upon standard skill taxonomy, and because job posting and other free text data may not use predefined skill terms, a model that is better able to handle unseen skill representations will be more robust. To examine this, we compare job representations comprised of skills from the Lightcast taxonomy to those comprised of skills and tasks from the O*NET ontology. Both skills and tasks were used due to the representation method used in O*NET, which uses a small set of pure skills and supplements this with other categories. O*NET provides a representation for each job; Lightcast skill representations were generated in the same manner as the previous experiment.

Model	Average Rank
Word2vec	81.8
Transformer - BERT	**17.3**

TABLE III: Taxonomy transfer results. Average rank of correct representation when comparing same job represented by 2 different taxonomies.

All CPS jobs were represented by skill clusters from both representations. The models were then asked to predict which O*NET based representation corresponded to each Lightcast based representation. Answers were returned as a ranked list and models were scored by the ranking position of the correct representation (lower rank is preferred). Results for the word2vec baseline and the proposed model are shown in Table III. Because the Career Changer's matrix is an O*NET-centric representation, a Lightcast based representation could not be produced and it was excluded from this experiment. As shown,

[5]https://www.myskillsmyfuture.org/

the BERT-based transformer substantially outperformed the baseline.

V. Conclusion

Robust job-skill representations can reveal similarities in the skills needed for different jobs, empowering workers to navigate their careers and learning. In this work we built upon advances in natural language processing to produce a new job-skill representation model. To do so, we introduced the same-job prediction task, which can be run over associated skill pairs found in a large corpus of job posting data without the need for further data annotation. To test the performance of the introduced model and to encourage further research on job-skill representations, we provide a dataset of job transitions in the labor market mined from US Census Bureau data. We show that such a dataset can be used for assessment of skill-based job predictions regardless of the method, dataset, or taxonomy employed to create them. Results suggest that the transformer-based model introduced here is better able to predict job transitions than existing baseline models. Additionally, using a pretrained language model as a basis for the skill representation was shown to allow for more robust transfer of the model between skill taxonomies.

References

[1] Daron Acemoglu and Pascual Restrepo, "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, 2019.

[2] David H. Autor and David Dorn, "The growth of low-skill service jobs and the polarization of the US Labor Market," *American Economic Review*, 2013.

[3] Erik Brynjolfsson and Andrew McAfee, *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*, WW Norton & Company, 2014.

[4] David Deming and Lisa B Kahn, "Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals," *Journal of Labor Economics*, 2018.

[5] Daron Acemoglu and Pascual Restrepo, "The race between man and machine: Implications of technology for growth, factor shares, and employment," *American Economic Review*, 2018.

[6] B. Johnston, B. Zweig, M. Peran, C. Wang, and R. Rosenfeld, "Estimating skill fungibility and forecasting services labor demand," in *2017 IEEE International Conference on Big Data (Big Data)*, Dec. 2017, pp. 3583–3585.

[7] Ahmad Alabdulkareem, Morgan R. Frank, Lijun Sun, Bedoor AlShebli, César Hidalgo, and Iyad Rahwan, "Unpacking the polarization of workplace skills," *Science Advances*, vol. 4, no. 7, pp. eaao6030, July 2018, Publisher: American Association for the Advancement of Science Section: Research Article.

[8] Enghin Atalay, Phai Phongthiengtham, Sebastian Sotelo, and Daniel Tannenbaum, "New technologies and the labor market," *Journal of Monetary Economics*, vol. 97, pp. 48–67, Aug. 2018.

[9] M. Papoutsoglou, A. Ampatzoglou, N. Mittas, and L. Angelis, "Extracting Knowledge From On-Line Sources for Software Engineering Labor Market: A Mapping Study," *IEEE Access*, vol. 7, pp. 157595–157613, 2019, Conference Name: IEEE Access.

[10] Ann Bartel, Casey Ichniowski, and Kathryn Shaw, "How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills," *The Quarterly Journal of Economics*, 2007.

[11] Erik Brynjolfsson, Tom Mitchell, and Daniel Rock, "What Can Machines Learn and What Does It Mean for Occupations and the Economy?," *AEA Papers and Proceedings*, 2018.

[12] Marios Kokkodis and Panos Ipeirotis, "The Utility of Skills in Online Labor Markets," . *Decision Analytics*, p. 18, 2014.

[13] Keman Huang, Jinhui Yao, and Ming Yin, "Understanding the Skill Provision in Gig Economy from A Network Perspective: A Case Study of Fiverr," *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, no. CSCW, pp. 132:1–132:23, Nov. 2019.

[14] M. Verma and N. Francis, "On Utility of Temporal Embeddings in Skill Matching," in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, Nov. 2017, pp. 237–241, ISSN: 2375-9259.

[15] Panos Alexopoulos and Spyretta Leivaditi, "Modeling, Measuring and Exploiting Concept Drift in the Labour Market Domain," in *SEMANTICS Workshops*, 2017, p. 4.

[16] Martin Fleming, Wyatt Clarke, Subhro Das, Phai Phongthiengtham, and Prabhat Reddy, "The future of work: How new technologies are transforming tasks," Tech. Rep., MIT-IBM Watson AI Lab, 2019.

[17] 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*, 2020.

[18] 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*, 2022, pp. 285–294.

[19] Mauricio Noris Freire and Leandro Nunes de Castro, "e-Recruitment recommender systems: a systematic review," *Knowledge and Information Systems*, vol. 63, no. 1, pp. 1–20, Jan. 2021.

[20] Mohammed Maree, Aseel B. Kmail, and Mohammed Belkhatir, "Analysis and shortcomings of e-recruitment systems: Towards a semantics-based approach addressing knowledge incompleteness and limited domain coverage," *Journal of Information Science*, vol. 45, pp. 713 – 735, 2019.

[21] Aritra Ghosh, Beverly Woolf, Shlomo Zilberstein, and Andrew Lan, "Skill-based Career Path Modeling and Recommendation," in *2020 IEEE International Conference on Big Data (Big Data)*, Dec. 2020, pp. 1156–1165.

[22] Vachik S. Dave, Baichuan Zhang, Mohammad Al Hasan, Khalifeh AlJadda, and Mohammed Korayem, "A Combined Representation Learning Approach for Better Job and Skill Recommendation," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, Torino Italy, Oct. 2018, pp. 1997–2005, ACM.

[23] Marcela Escobari and Natalie Geismar, "Mobility pathways," 11 2020.

[24] Giuseppe Moscarini and Kaj Thomsson, "Occupational and job mobility in the us," *Scandinavian Journal of Economics*, vol. 109, pp. 807–836, 12 2007.

[25] Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren, *Integrated Public Use Microdata Series, Current Population Survey: Version 8.0 [dataset]*, IPUMS, 2020.

[26] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin, "Attention is all you need," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Red Hook, NY, USA, Dec. 2017, NIPS'17, pp. 6000–6010, Curran Associates Inc.

[27] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv:1810.04805 [cs]*, May 2019, arXiv: 1810.04805.

[28] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean, "Efficient Estimation of Word Representations in Vector Space," *arXiv:1301.3781 [cs]*, Sept. 2013, arXiv: 1301.3781.

[29] Matthew T Allen, Gordon Waugh, Megan Shaw, Suzanne Tsacoumis, David Rivkin, Phil Lewis, Mark Brendle, Denise Craven, Christy Gregory, and Dan Connell, "The development and evaluation of a new o*net related occupations matrix," Tech. Rep., National Center for O*NET Development, 9 2012.

[30] Di Tong, Lingfei Wu, and James Allen Evans, "Low-skilled Occupations Face the Highest Re-skilling Pressure," *arXiv:2101.11505 [cs]*, Jan. 2021, arXiv: 2101.11505.

[31] Akshay Gugnani, Vinay Kumar Reddy Kasireddy, and Karthikeyan Ponnalagu, "Generating Unified Candidate Skill Graph for Career Path Recommendation," in *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, Nov. 2018, pp. 328–333, ISSN: 2375-9259.

[32] Jing Zhao, Jingya Wang, Madhav Sigdel, Bopeng Zhang, Phuong Hoang, Mengshu Liu, and Mohammed Korayem, "Embedding-based Recommender System for Job to Candidate Matching on Scale," *arXiv:2107.00221 [cs]*, July 2021, arXiv: 2107.00221.