Work unit evolution: Using large language models to determine temporal task evolution

Bryan Adams

George Mason University

November 28, 2023

Table of Contents

Motivation

- Identifying relevant knowledge, skills, and abilities (KSAs) has been extensively researched in economics, management and sociology. [2, 10, 13, 16]
- Classifying KSAs is also extensively researched in many fields. [7, 15, 17]
- The majority of previous research focused on unifying KSAs into one taxonomy as "tasks" and relied on survey data which forces participants to bin their occupational requirements. [11, 9]
- Previous research focused on industry or country level aggregation of occupational requirements.
- With the advancement of large language models (LLMs), could we identify occupational requirements for a specific position and how they evolve over time?

USAJobs job ads

Analyzing job ads for the Army Acquisitions Workforce (AAW)

In September 2023, USAJobs[1] contained **2,379,956** of historical job ads with position opening dates from 2012-09-12 to 2027-05-26. The following filtering criteria were applied to determine the relevant AAW job ads. The filtering resulted in a possible **75,914** job ads related to AAW with position opening dates from 2015-02-01 to 2023-10-22.

1. Filter on job ads with a hiring department of *Department of the Army* or *Department of the Defense* reducing the number of job ads to **772,297**.

2. Search job add for the term *acquisition* reducing the number of job ads to **79,071**.

[1] Job ads were pulled from https://developer.USAJobs.gov/API-Reference/

USAJobs job ads

(a) Historical job posting counts by position opening year

Year	Count
2012	1
2013	6
2014	23
2015	143
2016	3893
2017	237451
2018	329440
2019	348783
2020	328452
2021	369111
2022	441709
2023	320874
2024	69
2027	1

(b) Possible AAW historical job posting counts by position opening year

Year	Count
2015	5
2016	39
2017	7537
2018	12033
2019	13087
2020	10320
2021	11619
2022	12567
2023	8707

USAJobs job ad example

Raw job ad

To qualify for a Contract Specialist your resume and supporting documentation must support:
\n
\nA. . Basic DoD 1102 Requirement: Public Law 106-398, Section 808: A.) A baccalaureate degree from an accredited educational institution authorized to grant baccalaureate degrees AND B.) at least 24 semester hours (or equivalent) of study from an ... Creditable specialized experience includes:
\n
\nGS-13:\n
\n \t
Developing contractual strategies.
\n \t
Ensuring acquisition plans are in full compliance with contracting regulations and related Department of Defense (DoD) standards.
\n \t
Planning and conducting negotiations on price, technical requirements, terms, and conditions of the contract.
\n \t
Developing acquisitions strategies and/or determining methods of procurement and ensuring proper performance.
\n
\nGS-12:\n\n
\n \t
Verifies for accuracy purchase requests, prepares acquisition plans, conducts market research and recommends contract type and pricing strategies.
\n \t
Performs price and cost analysis, analyzing unit costs and pricing data and contractors projected costs.
\n \t
Performs full range of contract administration functions.
\n \t
Assists with extensive negotiations which address acquisition related issues with offerors.
\n \t
Assists to prepare and issue contract modifications under long term contracts.
\n \t
Assists in developing contractual strategies.
\n \t
Develops cost and price analysis spreadsheets for long-term contracts.
\n
\nE...

USAJobs job ad example

Removed html tags, common phrases and unnecessary characters.

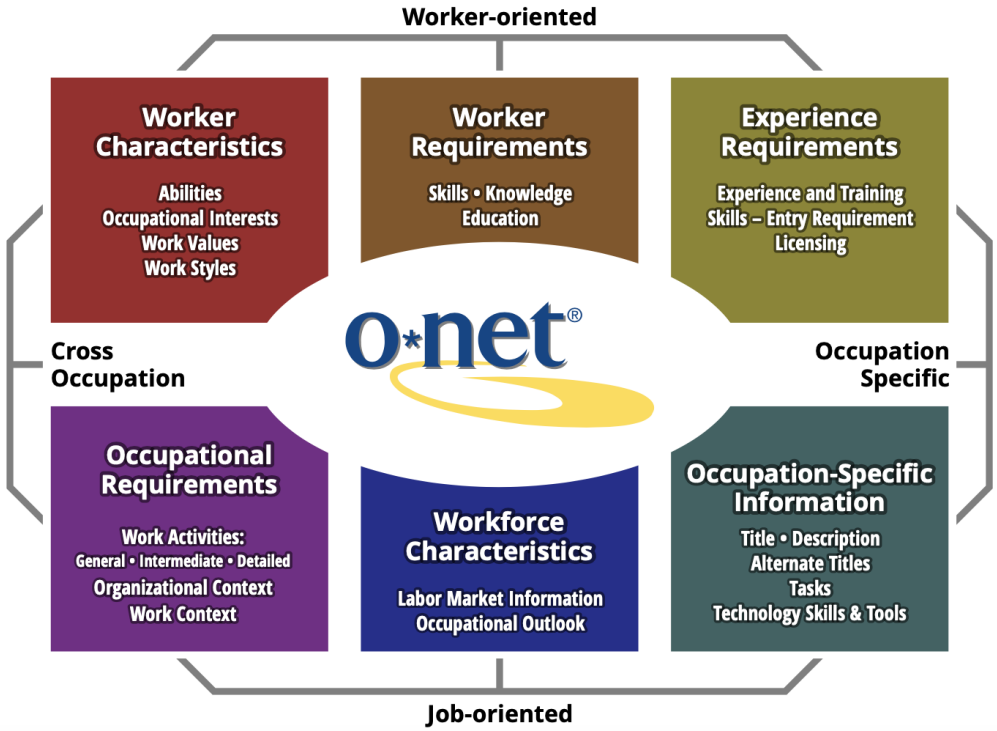
Cleaned job ad

To qualify for a Contract Specialist your resume and supporting documentation must support:A. Basic DoD 1102 Requirement: Public Law 106-398, Section 808: A.) A baccalaureate degree from an accredited educational institution authorized to grant baccalaureate degrees AND B.) at least 24 semester hours (or equivalent) of study from an ... Creditable specialized experience includes:GS-13:Developing contractual strategies.Ensuring acquisition plans are in full compliance with contracting regulations and related Department of Defense (DoD) standards.Planning and conducting negotiations on price, technical requirements, terms, and conditions of the contract.Developing acquisitions strategies and/or determining methods of procurement and ensuring proper performance.GS-12:Verifies for accuracy purchase requests, prepares acquisition plans, conducts market research and recommends contract type and pricing strategies.Performs price and cost analysis, analyzing unit costs and pricing data and contractors projected costs.Performs full range of contract administration functions.Assists with extensive negotiations which address acquisition related issues with offerors.Assists to prepare and issue contract modifications under long term contracts.Assists in developing contractual strategies.Develops cost and price analysis spreadsheets for long-term contracts.

Occupational Information Network (O*NET)

O*NET created and currently maintains occupation specific descriptors. These descriptors created the O*NET content model, which is used to define occupations within Standard Occupational Classification (SOC) system.[2]

- **33 knowledges** [4]
- **46 skills** [14]
- **52 abilities** [8]
- 42 generalized work activities
- 46 aspects of work context
- 21 occupational values
- 17 work style requirements



[2] Data files available at https://www.onetcenter.org/content.html

Example O*NET knowledge definitions[3]

- Administration and Management - Knowledge of business and management principles involved in strategic planning, resource allocation, human resources modeling, leadership technique, production methods, and coordination of people and resources.
- Administrative - Knowledge of administrative and office procedures and systems such as word processing, managing files and records, stenography and transcription, designing forms, and workplace terminology.
- Mathematics - Knowledge of arithmetic, algebra, geometry, calculus, statistics, and their applications.
- Engineering and Technology - Knowledge of the practical application of engineering science and technology. This includes applying principles, techniques, procedures, and equipment to the design and production of various goods and services.

[3] Data files available at https://www.onetcenter.org/database

PaLM - Pathways Language Model [3]

- PaLM uses a standard Transformer model architecture similar to the architecture described in *Attention is all you need*.
- The significant changes are:
  - Uses SwiGLU activation function
  - Parallel Layers in each Transformer block
  - Transformer formulation uses k attention heads
  - RoPE embeddings instead of absolute or relative position embeddings
  - Use SentencePiece for vocabulary
- 540 billion parameters, 118 layers, and 48 attention heads
- PaLM2 is an update on PaLM but trained on a larger more diverse data set and redesigned for more compute efficiency. [1]

SwiGLU activation function [20]

SwiGLU activation function combines the Swish activation function [18] and gated linear units (GLU) [5].

- Swish activation function
  - $Swish(x) = x\sigma(\beta x), \sigma(z) = (1 + e^{-z})^{-1}$
  - Improvement over ReLU $= max(0, x)$
- GLU
  - $h_l(\mathbf{X}) = (\mathbf{X} * \mathbf{W} + b) \otimes (\mathbf{X} * \mathbf{V} + c)$
    - h_l = hidden layer
    - $\mathbf{X}$ = input layer
    - $\mathbf{V}$ = words in model vocabulary
    - $\mathbf{W}$ = learned parameters
- SwiGLU
  - $Swish_\beta(x\mathbf{W} + b) \otimes (x\mathbf{V} + c)$
  - PaLM uses $Swish_\beta(x\mathbf{W}) \otimes (x\mathbf{V})$

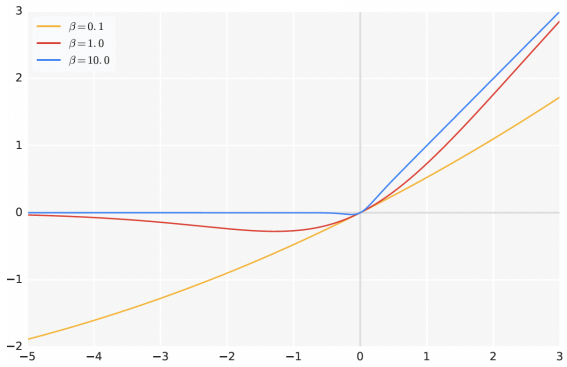


Figure: Swish activation function [18]

Identifing KSAs in job ads

The following is the prompt structure provide to PaLM2 to extract KSAs from a specific job ad.

Leading text provided to PaLM2

<knowledge name> is the <knowledge description>. Look at the job ad and decide if <knowledge name> is required to perform this job. If <knowledge name> is required please provide the specific tasks in a bulletized format from the job ad that relate to <knowledge name>, if not simply say no.

Query provided to PaLM2

Does this job ad require <knowledge name> knowledge to perform the job?

Example PaLM2 response for Administration and Management

Candidate 1: Yes.

The job ad requires the following skills:

- Knowledge of Federal and agency contracting laws and regulations applicable to centralized acquisition of agency-wide commodity and service requirements
- Knowledge of commercial business and industry practices and trends, cost factors, and market conditions along with understanding the military support requirements to be procured to identify potential contractors in limited supplier source situations.
- Developing of pre-negotiation plans, negotiation techniques, and cost and price analysis of procurement proposals to independently decide on proper strategies to procure complex requirements.
- Applying negotiation, persuasion and compromise techniques is required to negotiate price, terms and conditions, delivery schedule, contract modifications, (quantity options and extension of contract performance periods) and settlement of contractor claims in the best interest of the Government.

Identifing KSAs in job ads

- Temperature [0,1] = 0.5 - Controls the randomness of the output, higher values produce more varied responses
- Retrieve **3** candidate answers and if all are **yes** the KSA is added to the job KSA vector

Example KSA vector

[Physics, Mathematics and Science, Economics and Accounting, Business and Management, Education and Training, Mathematics, Chemistry, Engineering and Technology]

How close are two job requirements?

We need to define a vector space to measure how job requirements evolve overtime. In this figure, the job KSAs changed between t and $t + m$ as well as $t + m$ and $t + n$.

Example job KSA vectors

- t = [Engineering and Technology, Mathematics]
- $t + m$ = [Engineering and Technology, Physics]
- $t + n$ = [Business and Management, Economics and Accounting]

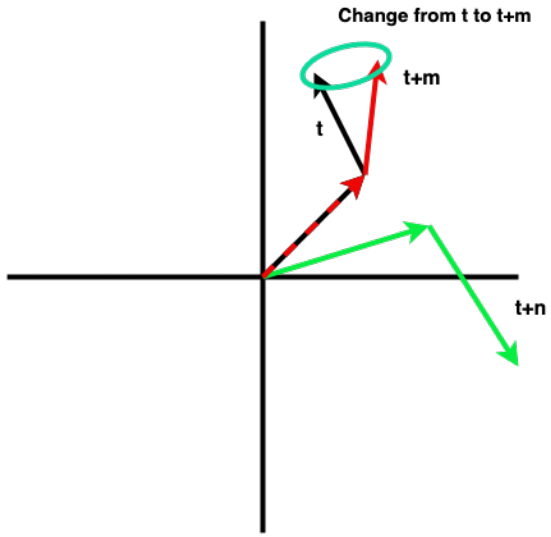
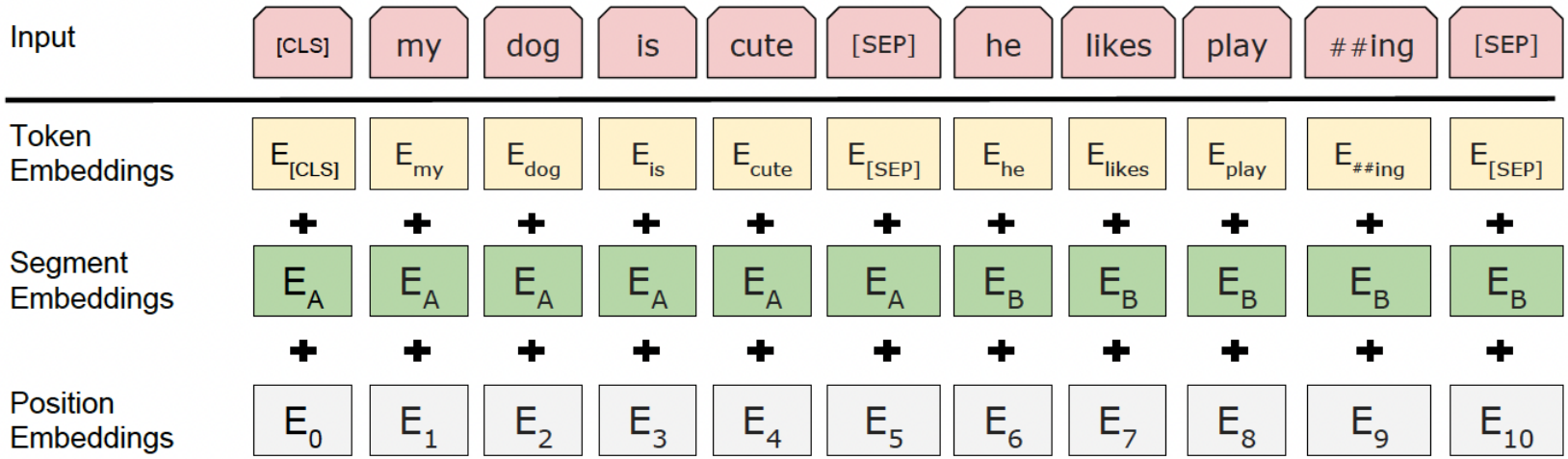


Figure: Representation of job KSA vector changes

Creating vector space for KSAs using BERT

BERT - Bidirectional Encoder Representations from Transformers [6]

- Multi-layer bidirectional Transformer encoder based on *Attention Is All You Need*. [21]
- $\text{BERT}_{\text{BASE}}$ - 12 layers, hidden size is 768, 12 attention heads, 110 million parameters
- Minimized cross entropy loss



Tuning BERT model

Used batch size of **16**, **2** epochs and masked **15%** of the tokens

Tokenized knowledge description

['[CLS]','principles','and','facts','related','to','business','administration',
'and','accounting',',','human','and','material','resource','management','in',
'organizations',',','sales','and','marketing',',','economics',',','and',
'office','information','and','organizing','systems','[SEP]','[PAD]',...]

Whole word masking

['[CLS]','[MASK]','and','[MASK]','related','to','business','[MASK]',
'and','accounting',',','human','and','[MASK]','resource','management',
'in','organizations',',','sales','and','marketing','[MASK]','economics',
'[MASK]','and','office','information','and','organizing','systems', '[SEP]','[PAD]'...]

Wordpiece [22] tokenizing

[['[CLS]','principle','##s','and','facts','related','to','business','administration',
'and','accounting',',','human','and','material','resource','management','in',
'organizations',',','sales','and','marketing',',','economics',',','and',
'office','information','and','organizing','systems','[SEP]']]

Tuning BERT model

Used **Adam** instead of **stochastic gradient descent**. **Adam** is a method for efficient stochastic optimization that only requires first-order gradients with little memory requirement. [12]

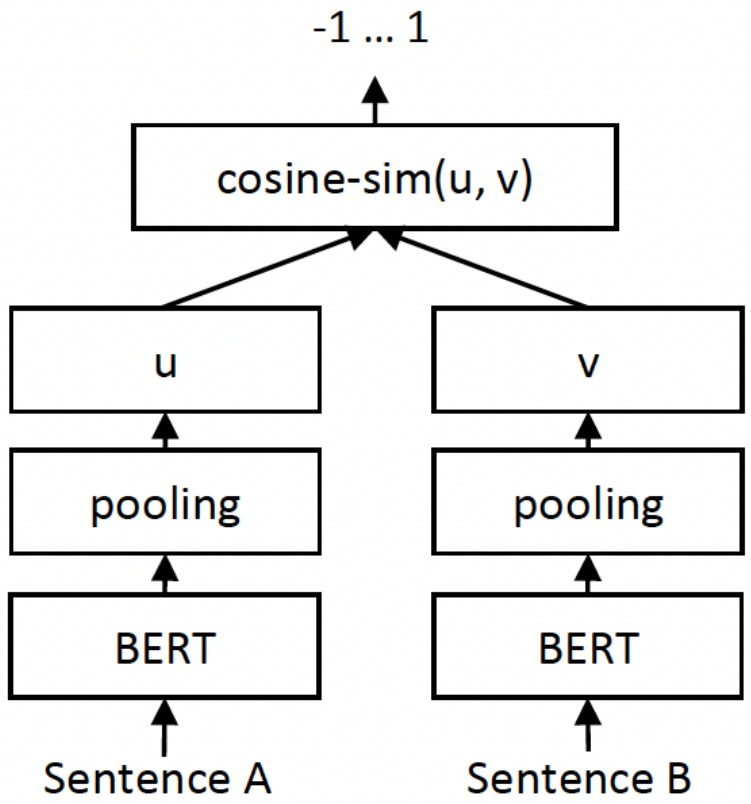
- Learning rate (step size): $\alpha = 0.00005$
- Hyper-parameters for decay rates of moving averages: $\beta_1 = 0.9$, $\beta_2 = 0.98$
- Correct initialization bias

Require: α: Stepsize
Require: $\beta_1, \beta_2 \in [0, 1)$: Exponential decay rates for the moment estimates
Require: $f(\theta)$: Stochastic objective function with parameters θ
Require: θ_0: Initial parameter vector
  $m_0 \leftarrow 0$ (Initialize 1^{st} moment vector)
  $v_0 \leftarrow 0$ (Initialize 2^{nd} moment vector)
  $t \leftarrow 0$ (Initialize timestep)
  while θ_t not converged **do**
    $t \leftarrow t + 1$
    $g_t \leftarrow \nabla_\theta f_t(\theta_{t-1})$ (Get gradients w.r.t. stochastic objective at timestep t)
    $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ (Update biased first moment estimate)
    $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$ (Update biased second raw moment estimate)
    $\widehat{m}_t \leftarrow m_t / (1 - \beta_1^t)$ (Compute bias-corrected first moment estimate)
    $\widehat{v}_t \leftarrow v_t / (1 - \beta_2^t)$ (Compute bias-corrected second raw moment estimate)
    $\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \widehat{m}_t / (\sqrt{\widehat{v}_t} + \epsilon)$ (Update parameters)
  end while
  return θ_t (Resulting parameters)

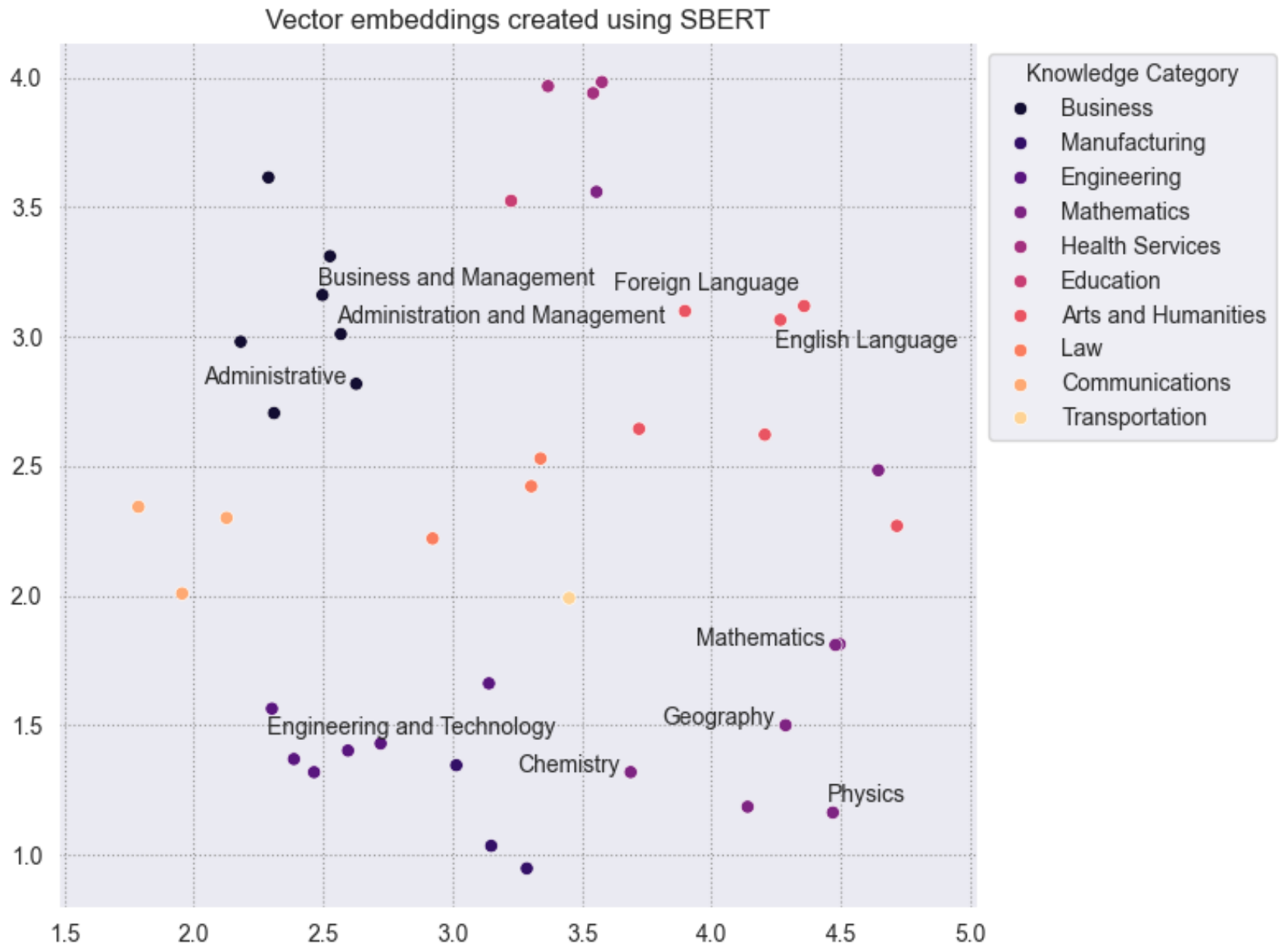
Sentence-BERT (SBERT) [19]

The objective is to determine which sentence pairs are similar (should be close in vector space) and which pairs are dissimilar (should be far away in vector space)

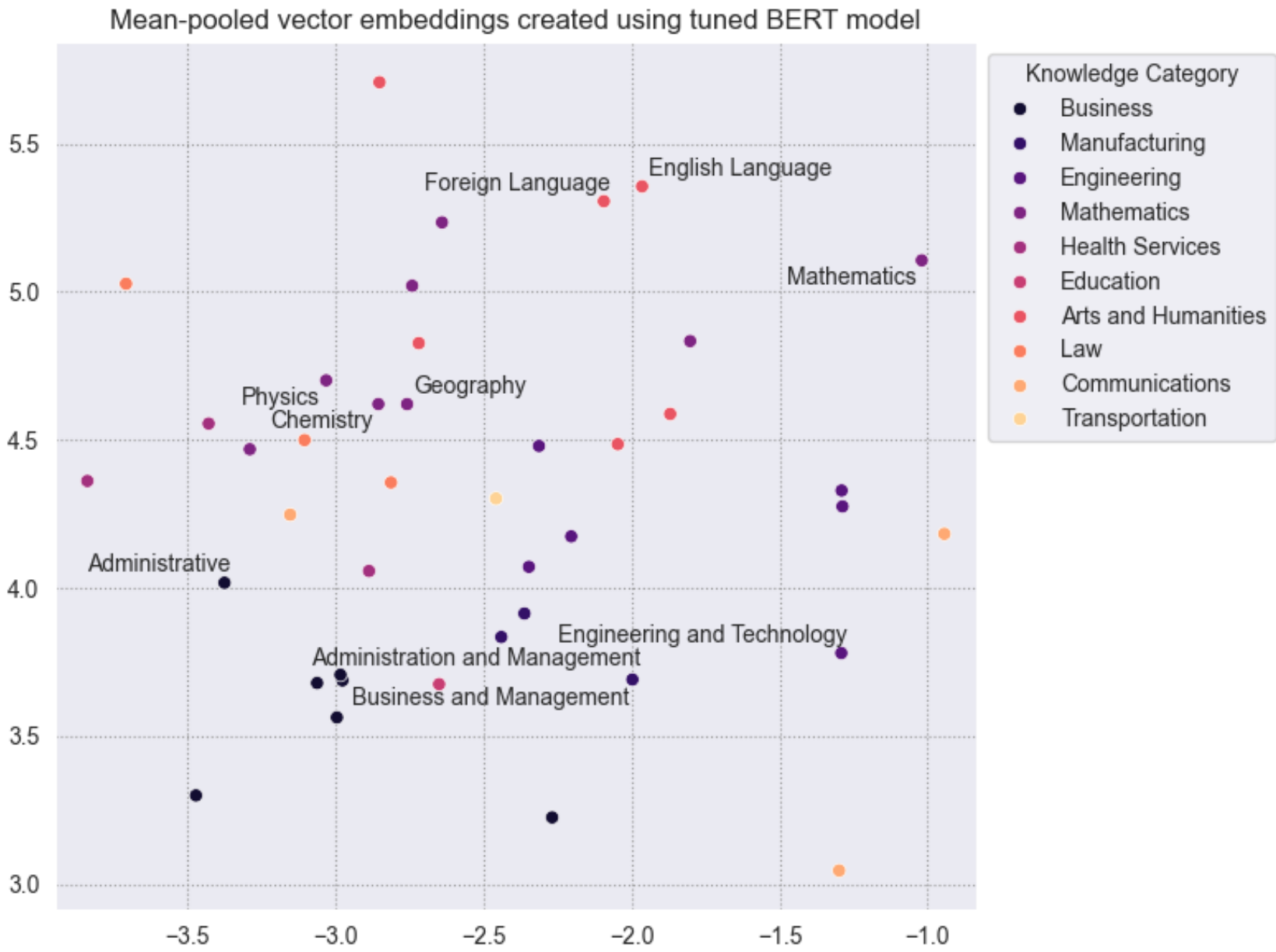
- SBERT adds a pooling operation to the output of BERT
- Mean pools the BERT hidden layers
- Computes the cosine similarity between the two embeddings: $\frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \|\vec{v}\|}$



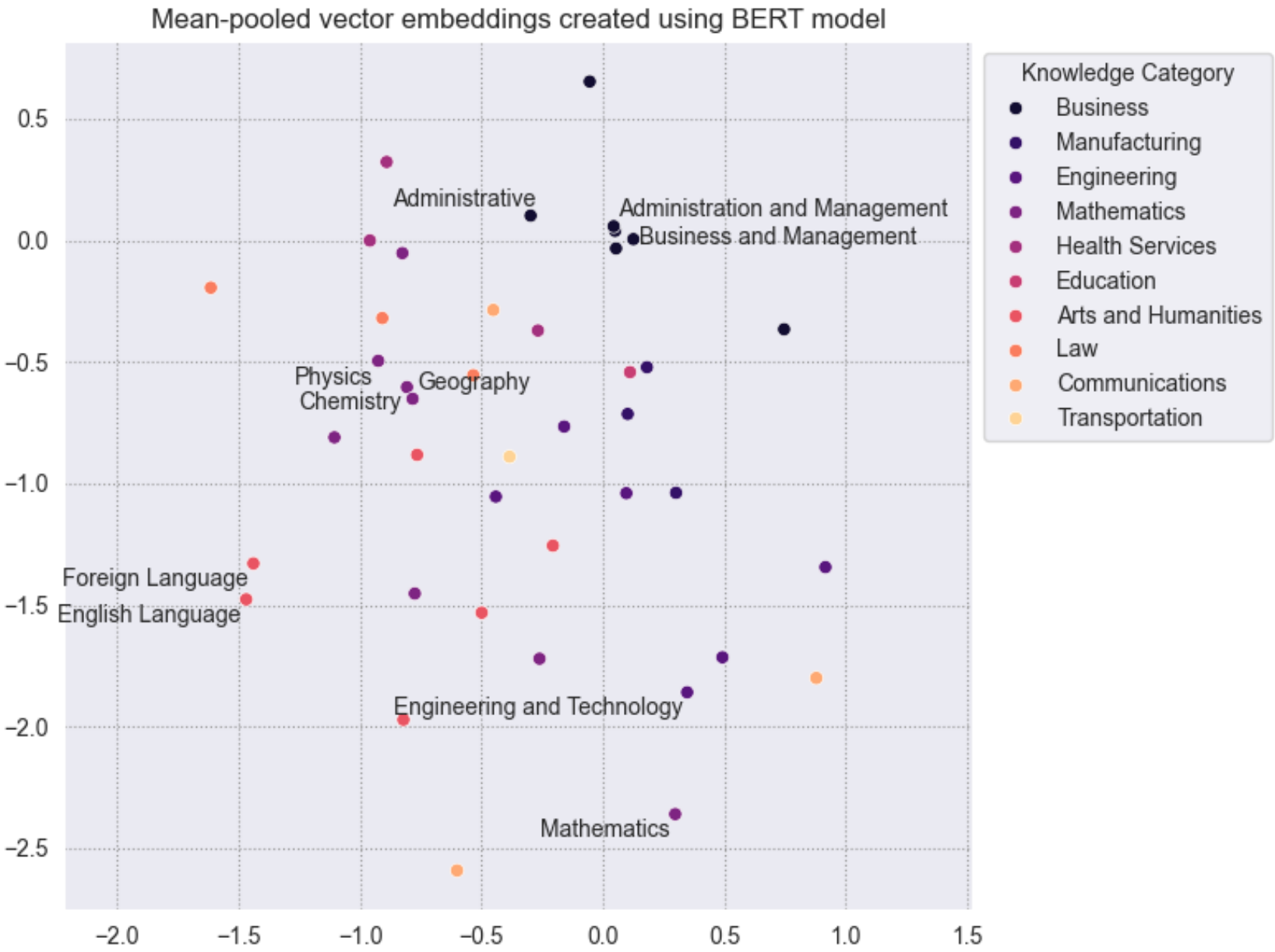
SBERT vector embeddings projection



Tuned BERT vector embeddings projection



BERT vector embeddings projection



Comparison of model results

Table: Cosine similarity between knowledge descriptions

Element 1	Element 2	SBERT	Tuned BERT	BERT base
Business and Management	Administration and Management	0.731	0.859	0.865
Administration and Management	Administrative	0.336	0.759	0.782
Mathematics	Physics	0.264	0.719	0.742
Engineering and Technology	Mathematics	0.277	0.727	0.732
Administrative	Mathematics	0.203	0.702	0.722
English Language	Foreign Language	0.831	0.941	0.965
Mathematics	English Language	0.170	0.706	0.724

Future work

- PaLM2 prompt generation - providing more specificity in prompt will increase validity of response, e.g. add examples of misclassification to the prompt
- Determine the optimal model to use to define the vector space of KSAs

References I

[1] R. Anil, A. M. Dai, and O. Firat, *Palm 2 technical report*, 2023.

[2] G. S. Becker, *Investment in human capital: A theoretical analysis*, Journal of Political Economy, 70 (1962), pp. 9–49.

[3] A. Chowdhery, S. Narang, J. Devlin, M. Bosma, G. Mishra, A. Roberts, P. Barham, H. W. Chung, C. Sutton, S. Gehrmann, P. Schuh, K. Shi, S. Tsvyashchenko, J. Maynez, A. Rao, P. Barnes, Y. Tay, N. Shazeer, V. Prabhakaran, E. Reif, N. Du, B. Hutchinson, R. Pope, J. Bradbury, J. Austin, M. Isard, G. Gur-Ari, P. Yin, T. Duke, A. Levskaya, S. Ghemawat, S. Dev, H. Michalewski, X. Garcia, V. Misra, K. Robinson, L. Fedus, D. Zhou, D. Ippolito, D. Luan, H. Lim, B. Zoph, A. Spiridonov, R. Sepassi, D. Dohan, S. Agrawal, M. Omernick, A. M. Dai, T. S. Pillai, M. Pellat, A. Lewkowycz, E. Moreira, R. Child, O. Polozov, K. Lee, Z. Zhou, X. Wang, B. Saeta, M. Diaz, O. Firat, M. Catasta, J. Wei, K. Meier-Hellstern, D. Eck, J. Dean, S. Petrov, and N. Fiedel, *Palm: Scaling language modeling with pathways*, 2022.

[4] D. P. Costanza, E. A. Fleishman, and J. Marshall-Mies, *Knowledges.*, An occupational information system for the 21st century: The development of O*NET., American Psychological Association, Washington, DC, US, 1999, pp. 71–90.

[5] Y. N. Dauphin, A. Fan, M. Auli, and D. Grangier, *Language modeling with gated convolutional networks*, 2017.

[6] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, *Bert: Pre-training of deep bidirectional transformers for language understanding*, 2019.

References II

[7] J.-P. Ferguson and S. Hasan, *Specialization and career dynamics evidence from the indian administrative service*, Administrative Science Quarterly, 58 (2013), pp. 233–256.

[8] E. A. Fleishman, D. P. Costanza, and J. Marshall-Mies, *Abilities.*, An occupational information system for the 21st century: The development of O*NET., American Psychological Association, Washington, DC, US, 1999, pp. 175–195.

[9] C. Gathmann and U. Schönberg, *How general is human capital? a task-based approach*, Journal of Labor Economics, 28 (2010), pp. 1–49.

[10] R. Gibbons and M. Waldman, *Task-specific human capital*, American Economic Review, 94 (2004), pp. 203–207.

[11] M. Hosseinioun, F. Neffke, Letian, Zhang, and H. Youn, *Nested skills in labor ecosystems: A hidden dimension of human capital*, 2023.

[12] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, 2017.

[13] R. E. Lucas, *On the mechanics of economic development*, 1988.

[14] M. D. Mumford, N. G. Peterson, and R. A. Childs, *Basic and cross-functional skills.*, An occupational information system for the 21st century: The development of O*NET., American Psychological Association, Washington, DC, US, 1999, pp. 49–69.

References III

[15] D. Neal, *Industry-specific human capital: Evidence from displaced workers*, Journal of Labor Economics, 13 (1995), pp. 653–677.

[16] F. Neffke and M. Henning, *Skill relatedness and firm diversification*, Strategic Management Journal, 34 (2013), pp. 297–316.

[17] M. Poletaev and C. Robinson, *Human capital specificity: Evidence from the dictionary of occupational titles and displaced worker surveys, 1984–2000*, Journal of Labor Economics, 26 (2008), pp. 387–420.

[18] P. Ramachandran, B. Zoph, and Q. V. Le, *Searching for activation functions*, 2017.

[19] N. Reimers and I. Gurevych, *Sentence-bert: Sentence embeddings using siamese bert-networks*, 2019.

[20] N. Shazeer, *Glu variants improve transformer*, 2020.

[21] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, *Attention is all you need*, 2017.

[22] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey, J. Klingner, A. Shah, M. Johnson, X. Liu, Łukasz Kaiser, S. Gouws, Y. Kato, T. Kudo, H. Kazawa, K. Stevens, G. Kurian, N. Patil, W. Wang, C. Young, J. Smith, J. Riesa, A. Rudnick, O. Vinyals, G. Corrado, M. Hughes, and J. Dean, *Google's neural machine translation system: Bridging the gap between human and machine translation*, 2016.