Measuring Innovations by their Labor Impact

Wajeeha Ahmad [1]

[1]Department of Management Science and Engineering, Stanford University

Introduction

Technological advances, particularly in artificial intelligence (AI), have led to widespread anxiety about the future of human professions. While innovations can automate tasks previously performed by labour, they can also augment human labor by creating new tasks and activities in which humans can be productively employed. Measuring such effects is challenging since the boundary between labor-augmenting and labor-automating innovations is not well-defined.

This project measures the occupational impact of innovations using patent text data. Such measures can help us better understand the impact of innovation on societal outcomes such as employment, wage inequality and the incentives of firms in producing different kinds of innovations.

Background: Using patent text to measure occupational impact

The extent to which a given patent automates and augments human labor is measured by comparing the textual similarity of patents with descriptions of tasks performed in various occupations and occupational titles. This project compares different techniques for obtaining such similarity measures using both static word embeddings and contextual word embeddings.

Approach: Generating automation and augmentation scores

Methodology

The approach for generating automation and augmentation scores for patents, and linking patents to occupations as done by Autor et al. (2021) is shown in Figure 1. I use a similar approach for generating patent-level scores using pretrained BERT-based and MPNET-based sentence transformers in addition to the gloVe embeddings used in prior work.

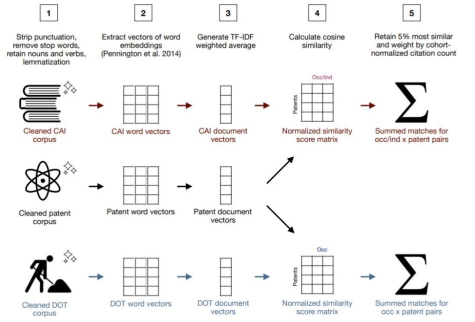

Figure 1. Generating automation and augmentation scores for patents

Linking patents to occupational tasks and titles: Key formulae

Patent-level and occupation-level automation and augmentation scores are calculated for 94,368 patents issued to US firms in 2010 as follows:

1. **Measuring similarity of patent text to occupational text**: For automation and augmentation scores, O_j represents the document vector of occupational tasks and titles, respectively.
$$\text{cosine similarity} = S_c(P_i, O_j) = \frac{\vec{p_i}\vec{o_j}}{|\vec{p_i}||\vec{o_j}|}$$
2. We get **automation and augmentation scores for each patent** according to the below formula:
$$\text{patent score}_i = \max(S_c(P_i, O_j)) - \text{median}(S_c(P_i, O_j)) \;\; \forall j \in J.$$
3. A patent is assigned to be an **automation patent** and/or an **augmentation patent** if its corresponding score is $> \sigma$, where σ is the 95^{th} percentile of the similarity score distribution:
$$I_{p,j} = 1 \;\text{ if }\; S_c(P_i, O_j) > \sigma,$$
4. An **occupation's exposure** to automation and augmentation patents is calculated as follows:
$$\text{occupational exposure}_j = arcsinh(\sum_p I_{p,j}).$$

Results: industry-level scores and matching patent text to occupations

Table 1 shows the automation scores by industry. Since GloVe and MPNet-based embeddings produce scores that show consistent patterns with the baseline Naive Bayes classification and with each other, I further evaluate scores based on these embeddings.

Industry	Naive Bayes	GloVe	BERT-base	MPNet-base
Chemical	0.16	0.87	0.72	0.52
Computers and Communications	0.98	0.92	0.73	0.55
Drugs and Medical	0.30	0.87	0.75	0.53
Electrical and Electronic	0.57	0.90	0.70	0.52
Mechanical	0.48	0.92	0.72	0.54
Others	0.37	0.92	0.73	0.56

Table 1. Mean (non-normalized) automation scores by industry.

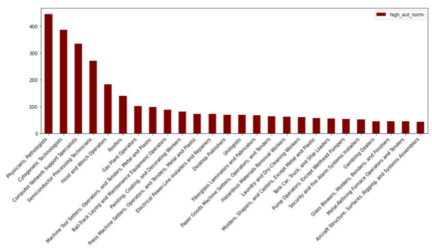

Figure 2. Number of automation patents associated with specific occupations.

Evaluation

To evaluate these measures, I analyze how they relate to the nature of jobs. To do so, I compile measures of the degree to which an occupation's tasks consist of routine, abstract and manual tasks based on work by Autor et al. (2003) and US Census code crosswalks from 1990 to 2018.

This analysis is based on the assumption that automated machines are good at carrying out repetitive tasks and fail at complex abstract tasks. Tables 2 and 3 show that automation measures based on GloVe embeddings better fit this prediction about routine and abstract tasks.

	Routine	Abstract	Manual
(Intercept)	3.76^{***}	3.18^{***}	1.06^{***}
	(0.25)	(0.22)	(0.15)
IHS_aut_patents	0.36^{**}	-0.38^{***}	0.24^{***}
	(0.12)	(0.11)	(0.07)
IHS_aug_patents	0.08	−0.13	−0.10
	(0.11)	(0.10)	(0.06)
Adj. R^2	0.08	0.13	0.08
Num. obs.	118	118	118
RMSE	1.99	1.77	1.17

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 2. Relationship between occupational scores and task content based on normalized GloVe scores.

	Routine	Abstract	Manual
(Intercept)	4.34^{***}	2.24^{***}	1.14^{***}
	(0.23)	(0.19)	(0.13)
IHS_aut_patents	0.05	0.23^{*}	0.05
	(0.12)	(0.10)	(0.07)
IHS_aug_patents	0.23	−0.04	0.03
	(0.12)	(0.10)	(0.07)
Adj. R^2	0.02	0.02	-0.01
Num. obs.	136	136	136
RMSE	2.10	1.78	1.25

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3. Relationship between occupational scores and task content based on normalized MPNet scores.

Table 4 shows that measures based GloVe embeddings better predict routine task content than both MPNet embeddings and the Naive Bayes classification from Mann and Puttman (2021).[1]

Model	Correlation
Naive Bayes	0.28
MPNet	0.05
GloVe	0.31

Table 4. Correlation of an occupation's routine task measure and its associated number of automation patents.

[1] The Naive Bayes correlation from prior work is based on industry-level measures whereas the GloVe and MPNet correlations are based on occupation-level measures.

References

[1] F. Levy Autor, D. H. and R. J. Murnane. The skill content of recent technological change: An empirical exploration. In *The Quarterly Journal of Economics 118(4), 1279–1333*, 2003.

[2] Anna Salomons David Autor, Bryan Seegmiller. New Frontiers: The Origins, and 1940–2018. Content of New Work. In *Working paper*, 2021.

[3] Katja Mann and Lukas Püttmann. Benign Effects of Automation: New Evidence from Patent Texts. In *The Review of Economics and Statistics.*, 2021.

wajeehaa@stanford.edu