ECONOMIC CONSEQUENCES OF ARTIFICIAL INTELLIGENCE AND ROBOTICS‡

What Can Machines Learn and What Does It Mean for Occupations and the Economy?†

By Erik Brynjolfsson, Tom Mitchell, and Daniel Rock*

Rapid advances in machine learning (ML) are poised to generate significant economic value and transform numerous occupations and industries. Machine learning, as described in Brynjolfsson and Mitchell (2017), is a sub-field of artificial intelligence (AI) that studies the question "How can we build computer programs that automatically improve their performance at some task through experience?" We believe it is also a "general purpose technology" (GPT), a technology that becomes pervasive, improves over time, and generates complementary innovation (Bresnahan and Trajtenberg 1995).

Recent rapid progress in ML has been driven largely by an approach called deep learning, and has made it possible for machines to match or surpass humans in certain types of tasks, especially those involving image and speech recognition, natural language processing, and predictive analytics. So far, the realized economic effects are small relative to the potential offered by this new GPT (Brynjolfsson, Rock, and Syverson 2017). This reflects the time lags of years or even decades before GPTs generate substantial economic value. Entrepreneurs and innovators take time to adopt new technologies, reconfigure existing work, discover new business processes, and co-invent complementary technologies (Bresnahan and Greenstein 1996). Reorganization of economic activity is an important determinant of the returns to innovation.

Concern about the coming wave of automation's impact on employment is growing. For instance, Acemoglu and Restrepo (2017) connect the adoption of robots to reduced employment and wages in local labor markets. A study by the McKinsey Global Institute suggested that about half of the work activities people perform could be automated with current technology (Manyika et al. 2017). While advances in ML are impressive, and automation is already having significant effects on many parts of the workforce, we are far from artificial general intelligence (AGI) which would match humans in all cognitive areas. This raises the question of which tasks will be most affected by ML and which will be relatively unaffected.

In particular, a key insight of Autor, Levy, and Murnane (2003) is that an occupation can be viewed as a bundle of tasks, some of which offer better applications for technology than others. As with studies of routine task automation, the impact of machine learning on employment is a function of the suitability of machine learning for specific work activities.

Furthermore, as noted by Levy (2017), the differential effectiveness of ML in different tasks suggests that the impact of ML diffusion will be uneven across occupations.

We first examine the channels by which ML can affect the workforce. Next, we apply Brynjolfsson and Mitchell's (2017) rubric for evaluating the potential for applying machine learning to tasks to the 2,069 work activities, 18,156 tasks, and 964 occupations in the O*NET database. From this, we build measures of what we call "suitability for machine learning" (SML) for labor inputs in the US economy. We then discuss measures of the potential for reorganization.

In the case of ML, we find that (i) most occupations in most industries have at least some tasks that are SML; (ii) few if any occupations have all tasks that are SML; and (iii) unleashing ML potential will require significant redesign of the task content of jobs, as SML and non-SML tasks within occupations are unbundled and rebundled.

Our findings suggest that a shift is needed in the debate about the effects of AI on work: away from the common focus on full automation of many jobs and pervasive occupational replacement toward the redesign of jobs and reengineering of business processes. Our evidence suggests that ML technologies will indeed be pervasive, but that within jobs, the SML of work tasks varies greatly. We suggest that variability in task-level SML is an indicator for the potential reorganization of a job, as the high and low SML tasks within a job can be separated and re-bundled. The focus of researchers, as well as managers and entrepreneurs, should be not (just) on automation, but on job redesign.

I. Machine Learning and Task Automation

Most of the recent progress in ML performance has been made by a specific class of algorithms called deep neural networks, or more generally, deep learning systems.[1] Although the basic structure of some of these models is decades old, significant new algorithmic advances have also been made. Interest and progress have reignited as computational costs in model training have fallen dramatically with improving hardware and new architectures.[2]

Past automation using explicit rules or manually written computer algorithms to automate tasks has had a significant impact on productivity and the workforce (Acemoglu and Autor 2011; Autor and Dorn 2013; Autor, Levy, and Murnane 2003). However, applications were limited to areas where knowledge was codified, or at least codifiable, because of Polanyi's Paradox (Polanyi 1966)—the fact that we "know more than we can tell." ML models circumvent Polanyi's Paradox by inferring the mapping function between inputs and outputs (in the case of supervised learning) automatically. While not always interpretable or explainable, these ML models open up a new set of possibilities for automation and complementarities to labor (Autor 2014). The types of tasks affected by ML will be quite different from those affected in past waves of automation.

Because of their capacity to learn highly nonlinear functions with near-automatic input space transformations, deep neural nets (DNNs) are currently the algorithms with some of the most obvious economic potential at the automation frontier. DNN software can be extended to new domains formerly closed to digitization by the high cost or impossibility of writing explicit maps of inputs to outputs and policies.

Suboptimal bundling of tasks in jobs can block potential productivity gains from ML. Consider the case of Leontief production, where all task inputs are complements such that production possibilities are constrained by the minimum of inputs. Bundling SML and non-SML tasks prevents specialization and locks up potential productivity gains. If the cost of ML capital (and SML task wage) were zero, workers would prefer to switch to tasks that ML cannot do. If firms only offer labor contracts that have a preset mixture of SML and non-SML tasks, all of the labor effort put toward SML tasks has an output opportunity cost increasing in efficiency units of forgone potential non-SML labor. ML could be doing those tasks, and the firm could increase profit if it were to reorganize job bundles.

One criterion for whether a task is SML is that the set of actions and the corresponding set of

---

outputs for the task can be measured sufficiently well that a machine can learn the mapping between the two sets. If ML substitutes for the tasks which produce the least noisy performance signals, then rebundling residual tasks in new jobs transfers risk from the firm to its workers.[3] Under a model of hidden action as in Holmström and Milgrom (1991), this will affect job design, compensation, and organization of work.[4]

II. What Can Machines Learn?

Successful application of machine learning is contingent on a variety of task characteristics and contextual factors of work activities. We use the O*NET content model for 964 occupations in the US economy joined to 18,156 specific tasks at the occupation level, which are further mapped to 2,069 direct work activities (DWAs) shared across occupations. We score each DWA for its Suitability for ML using a slightly extended version of the task evaluation rubric in Brynjolfsson and Mitchell (2017). The rubric we apply has 23 distinct statements to be evaluated on a 5-point scale varying from "strongly disagree" to "strongly agree."[5]

While we find it daunting to try to imagine all the ways a task could be automated—matching wits with the collective ingenuity of all the world's entrepreneurs—the scope of tasks that are SML, as ML currently exists, is much more constrained and definable. Evaluating worker activities with the rubric has the benefit of focusing on what ML can do and avoiding grouping all forms of automation together. The rubric is applied to each DWA to generate initial SML scores using CrowdFlower, a human

Table 1—Suitability for Machine Learning: Summary Statistics

	Occupations	Tasks	DWAs
Mean SML	3.47	3.47	3.47
SD of SML	0.11	0.31	0.32
Minimum SML	2.78	2.38	2.38
25th percentile SML	3.40	3.25	3.25
75th percentile SML	3.50	3.68	3.70
Max SML	3.90	4.48	4.48
Count	966	19,612	2.069

intelligence task crowdsourcing platform.[6] High values of SML offer an indication of where ML might have the greatest potential to transform a job.

There are a number of important conceptual caveats to this application of the SML rubric. The rubric focuses on technical feasibility. It is silent on the economic, organizational, legal, cultural, and societal factors influencing ML adoption. Additionally, we are focused on relatively near-term opportunities.[7] Matching the evolving state of the art in ML in the future will require updating the rubric accordingly.

Table 1 summarizes the SML measures for occupations, tasks, and activities. Table 2 presents the occupations with the five highest and five lowest values for SML. In addition, readers may be interested to know that occupation "economist" scores close to average (SML of 3.46). The variance of occupation-level SML is considerably lower than that of the tasks. As one would expect, job bundling of tasks provides some diversification with respect to machine learning exposure. Figure 1 shows counts of occupation-level proportions of tasks above the fiftieth, seventy-fifth, and ninetieth percentile for SML. Many occupations have several high SML tasks bundled with low SML tasks.

The within-occupation standard deviation of task SML scores is 0.596 (17.2 percent of the mean SML score of 3.466), revealing a high

---

[3]Performance measurement is directly related to the industrial potential for reinforcement learning algorithms as well. For instance, researchers at Google DeepMind report that they have implemented a neural net system that reduced cooling costs by 40 percent compared to the same data center when it was optimized by their human engineers (see https://deepmind.com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-40/).

[4]For instance, workers may need to be compensated for taking on bundles of tasks with noisier average performance when machines handle measurable tasks. This has the implication that over time worker performance will become harder to evaluate since the most measurable tasks tend to be the most suitable for ML. Brynjolfsson, Mitchell, and Rock (2018) has more detail on this point.

[5]Rubric details are available in the supplementary materials.

[6]The supplementary materials detail how the raw CrowdFlower dataset is built and processed. This dataset is sourced from our companion paper (Brynjolfsson, Mitchell, and Rock 2018). In addition to our measures included here (based on averages of median ratings of activities), we also evaluate more complex boolean combinations of the scores in the companion paper.

[7]For example, we have considered extensive physical activity a challenge for implementation of machine learning.

Table 2—Lowest and Highest 5 SML Score
Occupations

Low SML occupations	SML	High SML occupations	SML
Massage therapists	2.78	Concierges	3.9
Animal scientists	3.09	Mechanical drafters	3.9
Archeologists	3.11	Morticians, undertakers, and funeral directors	3.89
Public address system and other announcers	3.13	Credit authorizers	3.78
Plasterers and stucco masons	3.14	Brokerage clerks	3.78

Panel A. SML score versus occupational log
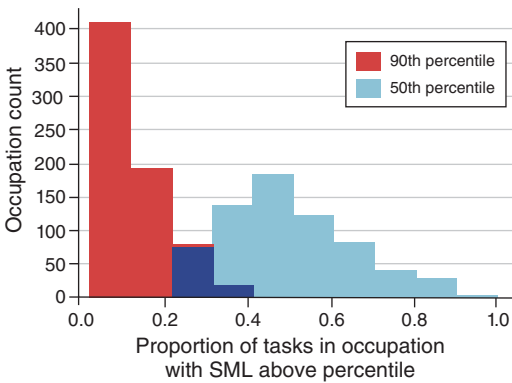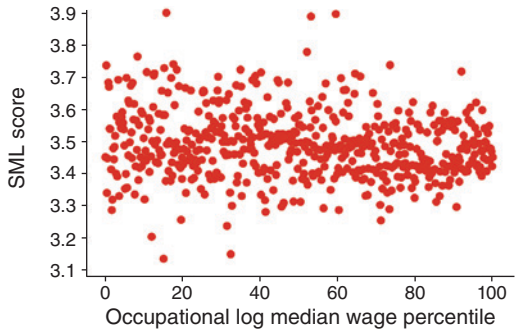median wage percentile

Panel B. SML versus occupational wage bill percentile

Figure 1. Frequency Counts of Occupational Task
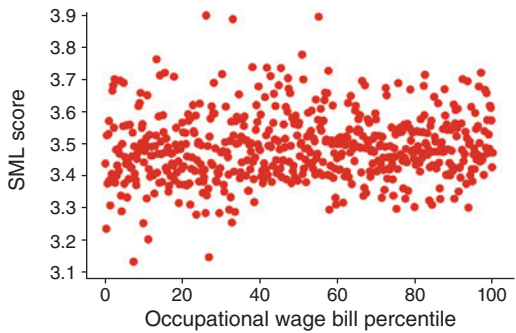Proportions Above Ninetieth, Seventy-fifth, and
Fiftieth Percentiles

Figure 2. Task-Level SML with Occupation versus
Occupational Wage and Wage Bill Percentile
(BLS 2016)

Note: Tasks are weighted by importance from the O*NET
database.

level of variability for the potential of machine learning within jobs. Jobs with higher scores in "sdSML" (within-occupation standard deviation of SML) have higher potential for reorganization.

Machine learning is a very different technology from earlier types of automation and it affects a very different set of tasks. While the last waves of automation lead to increase inequality and wage polarization as routine cognitive tasks were automated (Autor and Dorn 2013) it's not clear that ML will have the same effects. The correlation coefficients of SML with (log median) wage percentile and wage bill (BLS employment times wage) percentiles are very low: -0.14 and 0.10, respectively.

Furthermore, for sdSML, the correlation coefficients with wage and total wage bill percentiles are 0.17 and 0.002. This suggests that the next wave of automation and reengineering may affect a different part of the labor force than the last one. However, it's important to note that the

ex ante potential of ML may differ from its ultimate implementation, as other factors come to bear.

Even though SML correlation with wage and total wage expenditure percentiles is low, the actual implementation of ML technologies by managers and integrators may not follow the SML rankings. If technological change is directed, the implementation of ML by managers and entrepreneurs will be focused on the high wage bill tasks with higher SML.

III. Conclusion

Automation technologies have historically been the key driver of increased industrial productivity. They have also disrupted employment and the wage structure systematically. However, our analysis suggests that ML will affect very

different parts of the workforce than earlier waves of automation. Furthermore, tasks within jobs typically show considerable variability in SML, while few (if any) jobs can be fully automated using ML. Machine learning technology can transform many jobs in the economy, but full automation will be less significant than the reengineering of processes and the reorganization of tasks.[8]

REFERENCES

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks, and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, Vol. 4, edited by Orley Ashenfelter and David Card.

Acemoglu, Daron, and Pascual Restrepo. 2017. "Robots and Jobs: Evidence from US Labor Markets." Unpublished.

Autor, David H. 2014. "Polanyi's Paradox and the Shape of Employment Growth." In *Reevaluating Labor Market Dynamics*, Federal Reserve Bank of St. Louis Economic Policy Proceedings 129–78.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5): 1553–97.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118 (4): 1279–333.

Bresnahan, Timothy F., and Shane Greenstein. 1996. "Technical Progress and Co-invention in Computing and in the Uses of Computers." *Brookings Papers on Economic Activity: Microeconomics:* 1–77.

Bresnahan, Timothy F., and Manuel Trajtenberg. 1995. "General Purpose Technologies 'Engines of Growth'?" *Journal of Econometrics* 65 (1): 83–108.

Brynjolfsson, Erik, and Tom Mitchell. 2017. "What Can Machine Learning Do? Workforce Implications." *Science* 358 (6370): 1530–34.

Brynjolfsson, Erik, Tom Mitchell, and Daniel Rock. 2018. "The Technological Content of Occupational Change." Unpublished.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2017. "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics." National Bureau of Economic Research Working Paper 24001.

Holmström, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7: 24–52.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. "Deep Learning." *Nature* 521 (7553): 436–44.

Levy, Frank. 2017. "Computers and Populism: Artificial Intelligence, Jobs, and Politics in the Near Term." Unpublished.

Manyika, James, Michael Chui, Mehdi Miremadi, Jacques Bughin, Katy George, Paul Willmott, and Martin Dewhurst. 2017. *A Future That Works: Automation, Employment and Productivity*. Chicago: McKinsey Global Institute.

Polanyi, Michael. 1966. "The Logic of Tacit Inference." *Philosophy* 41 (155): 1–18.

---

[8] We might see, for example, large-scale machine learning platform companies contracted to automate aspects of various jobs. The wage and employment effects of these contracts are ambiguous given possible channels of demand elasticity, complementary task efforts, and substitutes.