MaRS

Automation Affecting Canadians

Understanding the impact of machine learning on the Canadian labour force

JUNE 2020

The Team

Writing, Analysis and Visualization

Upeksha Amarasinghe
Data Analyst, MaRS Data Catalyst

Amy Rowe
Data Analyst Intern, MaRS Data Catalyst

Analysis and Visualization

FeiFei Han
Data Analyst Intern, MaRS Data Catalyst

Madhusoodhanan Rajendran
Data Analyst Intern, MaRS Data Catalyst

Concept and Review

Matthias Oschinski
Lecturer, University of Toronto, Munk School of Global Affairs and Public Policy
Former Lead Executive, MaRS Data Catalyst

Report Design

Steve Deery
Associate, Product Management, MaRS Data Catalyst

We would like to thank Asa Motha-Pollock and Melissa Felder for their review and providing quality assurance, as well as Daria Lapteva and Melissa Smith for their guidance in the development of interactive data visualization.

This is an independent study by MaRS Data Catalyst.

Table of Contents

Executive Summary

The Fourth Industrial Revolution is fundamentally changing the way we live, work and relate to one another (World Economic Forum, 2020). A key driver in this revolution is Artificial Intelligence (AI), which is intelligence demonstrated by machines. Machine learning (ML), a sub-discipline of AI, is contributing to an increased scope and scale of the deployment of AI across various aspects of life, and is expected to transform numerous occupations and industries.

In 2020, MaRS Data Catalyst undertook an independent analysis to understand the potential for how machine learning could impact Canadian occupations and industries. More specifically, we look to understand the potential for Canadian occupations to be substituted by machine learning. We based the analysis on research by Brynjolfsson and Mitchell (2017)[1], who developed a rubric to evaluate the potential for machine learning to replace specific occupational tasks in the U.S. Brynjolfsson et al. (2018)[2] apply this rubric to build measures of "suitability for machine learning" (SML) for occupational tasks found in O*NET, an American occupation information database[3].

To understand the potential for ML to impact Canadian occupations and industries, MaRS Data Catalyst mapped SML scores from American occupations in O*NET to Canadian occupations in the National Occupation Classification (NOC) using an O*NET-NOC crosswalk[4]. Our main findings suggest that:

1. There is no correlation between income and the potential impacts of ML;
2. There is a weak negative correlation between the employment population of a Census Metropolitan Area (CMA) and its employment-weighted SML; and,
3. Finance and insurance, management of companies and enterprises, as well as retail trade sectors are relatively more exposed to the impacts of machine learning than agriculture, forestry, fishing and hunting and construction sectors.[5]

This indicates that ML will impact occupations relatively equally regardless of income and that the relative impact across Canada's metropolitan areas is inversely related to population, however, this relationship is weak. And certain industries are at risk of greater exposure based on their composition of occupations and the nature of the tasks required by those occupations.

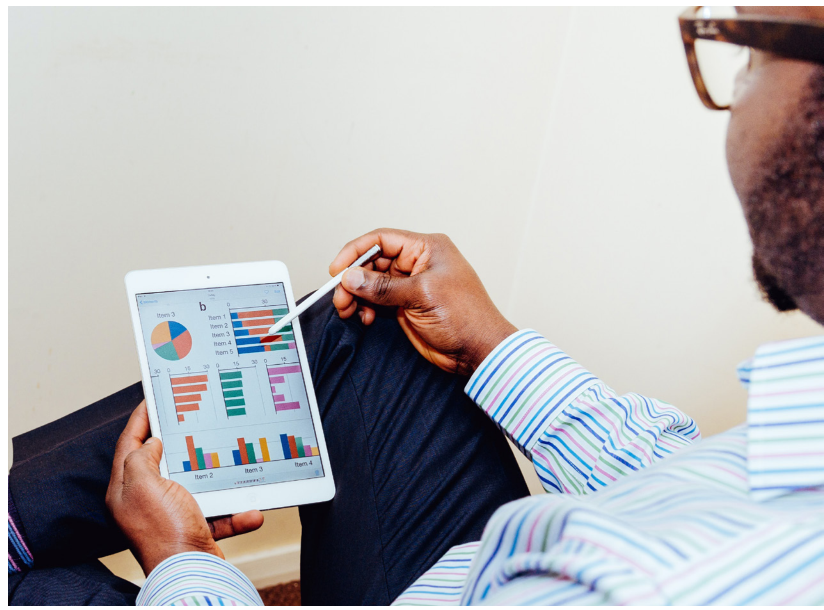
While our O*NET-NOC crosswalk is a one-to-one mapping, a many-to-many mapping is more appropriate in many cases but carries limitations. A many-to-many mapping requires understanding of the proportional composition of the O*NET occupations mapped to a given NOC, without which we are

1. Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. Science Magazine, Vol. 358, Issue 6370, pp. 1530-1534. Retrieved from https://www.cs.cmu.edu/~tom/pubs/Science_WorkforceDec2017.pdf
2. Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy?. American Economic Association Papers and Proceedings, Vol. 108, pp. 43-47. Retrieved from https://www.aeaweb.org/articles?id=10.1257/pandp.20181019
3. Their main findings suggest that (i) ML will affect different occupations than those affected in prior automation waves, (ii) most occupations involve some tasks that are suitable for ML, (iii) few occupations are fully automatable using ML and (iv) realizing the potential for ML requires a redesign of the job task content.

4. This crosswalk is a table that maps the relationship between the U.S. occupations in the O*NET taxonomy and Canadian occupations in the National Occupation Classification.
5. We note that these findings differ from Brynjolfsson et al. (2019) whose research suggests that lower wage occupations will be disproportionately affected as well as those employed in the retailing and transportation industries.

unable to map SML scores between the two taxonomies. Further, as SML scores were produced for occupational tasks, the use of a crosswalk assumes that tasks overlap between matching occupations in both taxonomies.

Further research should look to apply Brynjolfsson and Mitchell's rubric to evaluate the suitability for machine learning of the "main duties" found in each NOC in order to more accurately evaluate the impact of ML on Canadian occupations. The expansion of this rubric to include an assessment of the legal, social and organizational aspects of ML applicability and adoption in addition to other forms of automation, such as robotics, will allow for a more comprehensive understanding of impact of the Fourth Industrial Revolution on Canadian occupations, industries and the economy.



Introduction

The Fourth Industrial Revolution, characterized by the development and utilization of emerging technologies, is currently underway. This fundamental shift in the nature of work has sparked an unprecedented workforce transition. Across many occupations and industries, tasks that were traditionally performed by humans have been automated, causing concern about the impact of artificial intelligence (AI) on the labour force. Specifically, there has been increasing concern about the impact of automation on worker wages and the demand for labour. In their book The Second Machine Age, Brynjolfsson and McAfee (2014, p.11), present a grim prediction of the effect automation will have on employment:

Rapid and accelerating digitization is likely to bring economic rather than environmental disruption, stemming from the fact that as computers get more powerful, companies have less need for some kinds of workers. Technological progress is going to leave behind some people, perhaps even a lot of people, as it races ahead. As we'll demonstrate, there's never been a better time to be a worker with special skills or the right education, because these people can use technology to create and capture value. However, there's never been a worse time to be a worker with only 'ordinary' skills and abilities to offer, because computers, robots, and other digital technologies are acquiring these skills and abilities at an extraordinary rate.

AI is the academic discipline that attempts to mimic human cognitive functions or act in a way that is considered "smart." Kaplin and Haenlin define AI as "a system's ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation" (Kaplin and Haenlein, 2018). Applications of AI can be found all around us, such as Apple's virtual assistant, Siri, and the Tesla Autopilot driver assistance system. AI has enabled intelligent machines to do increasingly more and replace work that was once done by humans.

Artificial intelligence

Intelligence demonstrated by machines akin to human intelligence

Machine learning

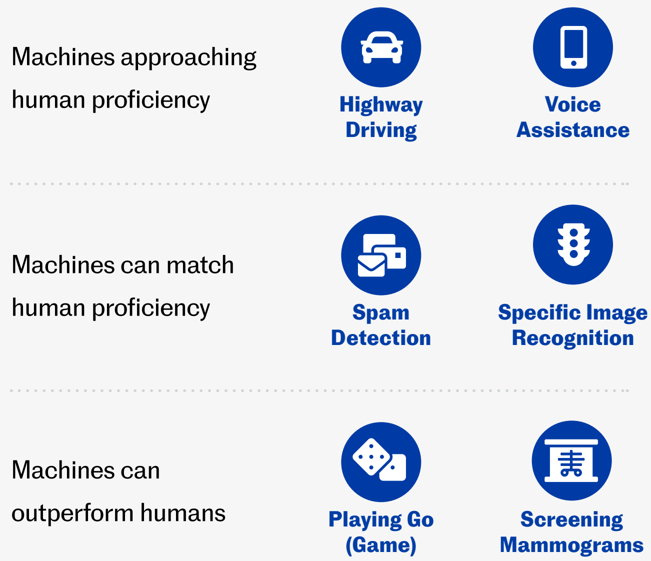
The science of getting machines to learn without being explicitly programmed

Deep learning

Subset of machine learning based on artificial neural networks inspired by neurons in the brain

FIGURE 0
AI, ML and DL Definitions

FIGURE 1
Progress of ML automation on human tasks



The three main drivers of advancements within the AI discipline are (1) exponential growth in computing power — it is estimated that computing power will surpass brain power by 2030, (2) increased data collection — particularly of unstructured data (data that is not organized in a systematic way, such as text data), and (3) breakthroughs in research (McAfee and Brynjolfsson, 2017).

Many recent advancements in AI are due to breakthroughs in subfields of AI known as machine learning (ML) and deep learning (DL). Andrew Ng, a computer scientist and pioneer in online education, known for his efforts to democratize artificial intelligence through online courses, defines ML as the science of getting machines to act without being explicitly programmed. It can also be thought of as a prediction technology, as ML algorithms rely on patterns or inferences from vast amounts of data (Agrawal et al., 2019). Deep learning is a subset of machine learning based on artificial neural networks that are inspired by the structure and function of the brain (see Figure 0 for definitions).

This paper assesses the impact of AI on Canada's labour market, specifically as it relates to ML by applying a methodology developed by Brynjolfsson et al. (2018). Recently, there have been compelling advancements in ML that will generate significant economic value and transform many occupations and industries. Work activities historically performed by humans are now being automated and new tasks are being created to complement these emerging technologies. The AI Index 2018 Annual Report identified the ML applications that have experienced some of the greatest breakthroughs: recommendation systems, natural language processing (text analytics), machine translation and voice recognition (Shoham et al., 2018). Further, machines have now surpassed humans in tasks involving image and speech recognition, predictive analytics and natural language processing (Brynjolfsson et al., 2017). For example, this year, Google's DeepMind AI system outperformed six radiologists in detecting breast cancer from mammogram scans (McKinney, 2020) (see Figure 1 for more examples). A recent study suggests that the impact of AI on the labour market is likely to be different from the impact industrial robots generate. More specifically, it indicates that the proportion of high-skilled workers is greater in firms applying AI and ML technologies compared to those deploying industrial robots (Morikawa, 2020).

ML has seen significant advances in the past two decades, and it is already integrated into the workforce. To better predict how ML will continue to impact the labour market, it is essential to understand which occupational tasks will be most affected by ML. The goal of this paper is to understand the potential effect of ML on the Canadian labour force. We build off the work done by Brynjolfsson et al. (2017), applying their methodology to the Canadian labour force.

Literature Review

AI is being deployed in a variety of industries and we can reasonably expect this trend to continue as further advancements in AI are made. Despite this, there is currently no consensus among economists on whether the digital disruption of the labour force will lead to net job gain or net job loss. The University of the Chicago Booth School of Business polled a panel of economists on the following question:

“Holding labour market institutions and job training fixed, rising use of robots and artificial intelligence is likely to increase substantially the number of workers in advanced countries who are unemployed for long periods (IGM, 2017).”

Of the responses, 28 percent agreed or strongly agreed, 20 percent disagreed or strongly disagreed, and 24 percent were uncertain (IGM, 2017). The relatively even distribution of responses across all answers, reiterates the uncertainty around the impact of AI, and signals a need for further research to help mitigate the rising concerns. Increased awareness and understanding about the impact of AI can assist firms and policy makers in better preparing for the future of work.

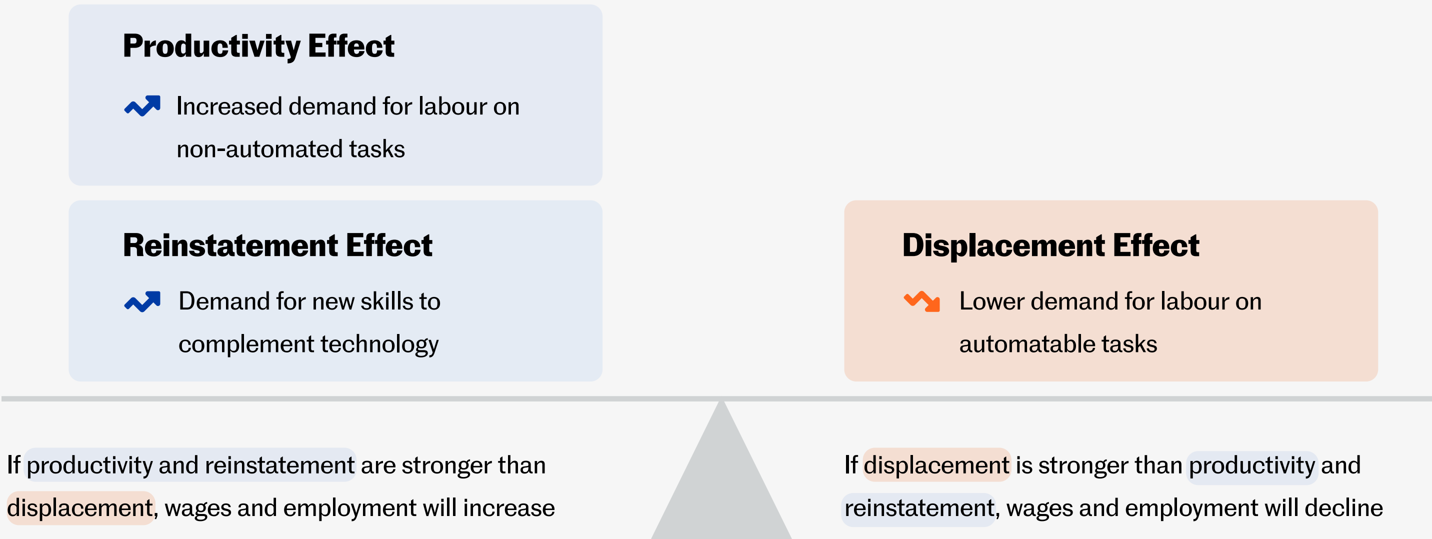
The uncertainty around the impact of AI on employment stems from the fact that there are two opposing forces at work. On the one hand, AI and machines are the drivers behind the automation of tasks historically performed by humans, potentially reducing the demand for labour and wages. On the other hand, automation and digitization increases productivity and generates cost savings which can create new jobs to complement emerging technologies and non-automated tasks (Acemoglu and Restrepo, 2018). The net effect on employment depends on which effect dominates.

Recent research suggests that, for now, the former effect might be more dominant. Using a model in which robots compete against workers, Acemoglu and Restrepo (2018) find that the effects of increased industrial robot usage between 1990 and 2007 on the U.S. labour markets is negative. Industrial robots are typically used in manufacturing to automate certain parts of a production line. To study the impact of automation, they compute for each region in the U.S., a measure of exposure to industrial robots, capturing the spread of robots in the industries in that region. They then compare the evolution of employment and wages in the most affected area to that in the least affected areas and find large negative impacts on both employment and wages. Specifically, according to estimates, one additional robot per thousand workers reduces the employment to population ratio by 0.2 percentage points and depresses wages by 0.42 percent. Overall, high exposure to industrial robots lowers the demand for labour and worker wages.

While these findings paint a rather bleak picture, researchers argue that there is more to the story (Bughin, 2018). Instead of the outright automation and elimination of jobs, companies can harness the power of AI to increase efficiency. Preliminary evidence indicates that early AI adoption can stimulate value creation for a firm (Bughin et al., 2017). By making early investments in AI and positioning themselves for growth, companies can stimulate employment. This can potentially

FIGURE 2

Displacement, productivity & reinstatement effects



offset some of the unemployment caused by automation. However, for this to occur, it requires firms to use AI for innovation. Although it seems like the workforce is becoming more technologically advanced, many firms are hesitant to adopt AI-based architectures. As Bughin (2018) points out, in a sample of 3,000 companies, half are AI resistors and will either not invest in AI, or will do so on a very small scale. Irrespective of the specific reasons, the lack of adoption among these companies can potentially have a negative impact on employment. As the study points out, job loss will not only occur as a result of automation. Job loss will also occur due to the unwillingness to use AI for innovation purposes, leading to lower firm profits and a lower absolute need for labour. This finding, then, suggests a need for enhanced investment in AI in order to boost positive labour market outcomes.

As Morikawa (2020) shows, however, any boost in labour market outcomes might necessitate a greater need for upskilling within the labour force. His study suggests that AI has a great complementarity with cognitive skills, thus requiring a higher skilled workforce.

Acemolgu and Restrepo (2018) describe the two effects of innovation on the labour market as the displacement and productivity effects. A key point behind their argument is that automation replaces specific tasks formerly performed by humans. The displacement effect measures the substitution of machines for this type of human labour. This substitution subsequently causes a decrease in the demand for labour, causing wages and employment to decline. However the displacement effect is offset by the productivity effect. The productivity effect implies that the demand for labour could increase for those tasks that are not being automated. Finally, in addition to the productivity effect, innovation will also lead to the creation of new tasks to complement new technology, which will directly counteract the displacement effect. Acemoglu and Restrepo (2018) refer to this as the reinstatement effect. As such, the impact of automation and digitization on employment and wages ultimately depends on the size of these effects. If the sum of the reinstatement and productivity effect is larger than the displacement effect, we can expect an overall positive impact on employment and wages (see Figure 2).

Evidence of the reinstatement effect can, for example, be found in higher education. Top universities across the country are responding to the changing nature of the workforce, by creating new degree programs aimed at providing students the required skills for jobs that complement emerging technologies. In 2018, the Rotman School of Management at the University of Toronto established the Master of Management Analytics (MMA) program. The goal of the program is to provide students with the skills needed for data-driven decision making with a focus on AI and other technologies. Entering its third year, the program has seen a 121 percent increase in applications since 2018, as students seek the skills needed to remain competitive in the changing labour market. Other universities, such as Queen's University and the University of British Columbia, have also recently begun offering similar degrees. For instance, Queen's Smith School of Business offers the Master of Management in Artificial Intelligence.

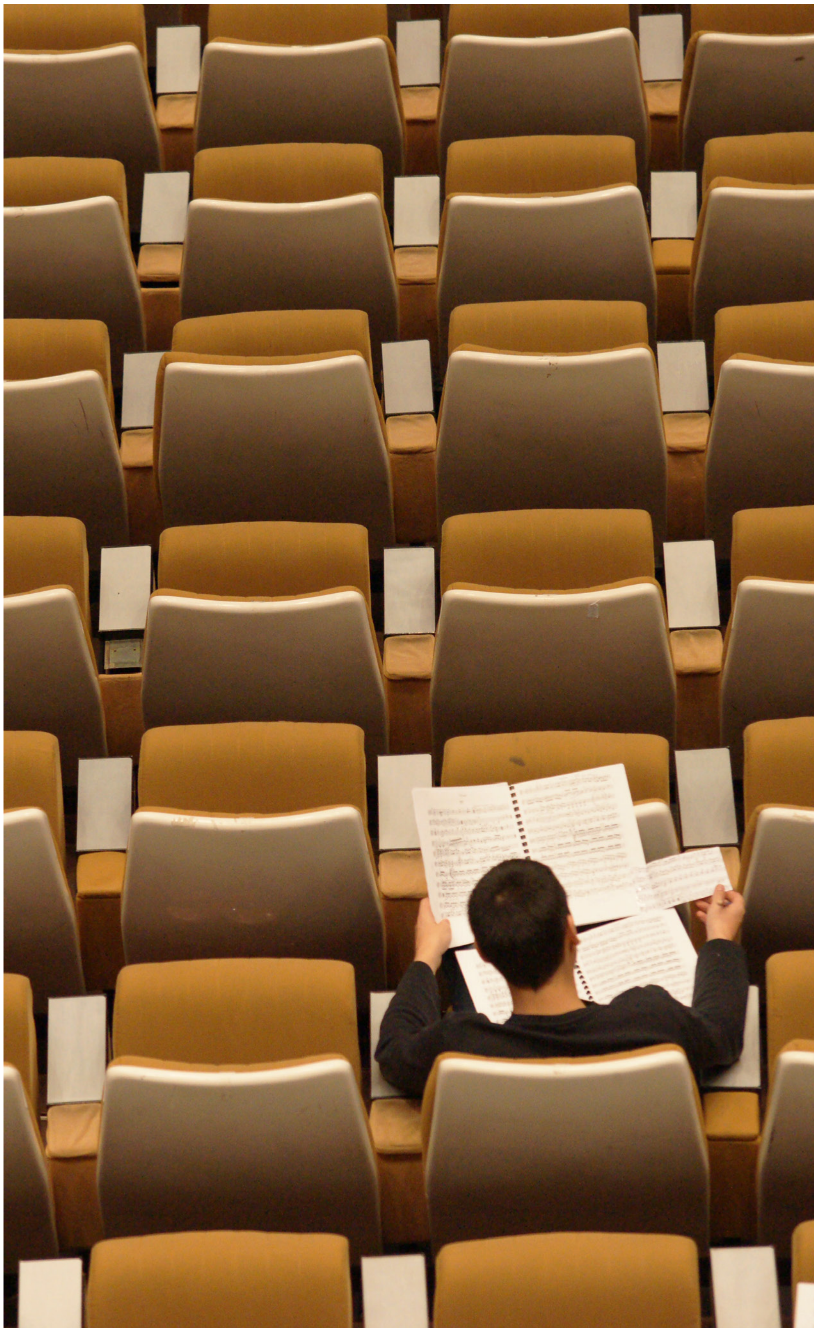
In recent years, advancements in ML have generated significant economic value. Between 2015 and 2018, active AI startups in the U.S. increased by 113 percent, and venture capital funding for these startups increased by 250 percent from 2013 to 2015 (Shoham et al., 2018). To echo the changing nature of the labour economy, such institutions as the Vector Institute and Creative Destruction Lab have been established to champion AI-based innovation.

Despite new advances in the field of AI and widespread integration of ML into the workforce, however, we are still far from the ability of machines to match humans in all cognitive areas (Brynjolfsson et al., 2018). Consequently, some occupational tasks and occupations are more suitable for AI applications than others. In general, any occupation can be viewed as a bundle of tasks. Taking this approach allows for a deeper look at the impacts of AI on jobs. Bryjnolfsson et al. assess the impacts of ML, a subset of AI, on specific work activities by creating a measure called the "suitability for machine learning" (SML). Using O*NET, an American occupational information database, they apply this rubric to evaluate the suitability for ML of 2069 direct work activities (DWA). They then aggregate scores for the DWA to the occupation-level to produce occupation-level SML scores. The authors find that: (1) most occupations across most industries have some tasks that could be automated; (2) virtually no occupations are composed entirely of tasks that can be automated; and (3) in order to achieve the full potential of ML a re-engineering of the task content within occupations is required.

In order to determine the likely impact of AI on the labour market, recent research aims to identify the tasks ML excels at, and to then identify occupational tasks that could be automated. In this context, a study by Agrawal et al. (2019) finds that ML particularly excels at prediction tasks, which can be defined as using data to fill in missing information. Based on their findings, the authors argue that the ML will replace prediction tasks that were traditionally performed by humans. Based on this observation, it can be argued that the impact of AI on employment can be boiled down to the competitive advantage in each occupation. In other words, jobs in which the core tasks involve prediction are more likely to be negatively affected by ML. However, this depends on the availability of the data required to make those predictions.

As described above, the integration of AI in the economy has a productivity effect that can counteract the displacement of human labour caused by machines. As such, machines are most likely to replace humans when they are a substitute for human labour and raise the productivity of those employees for whom they are a complement. This raises the question of how automation and employment will be distributed across types of workers in the economy. At the current state of technology, high-skilled labour is more challenging to automate (Prettner and Strulik, 2019). Therefore, it is likely that widespread automation can increase income inequality (Autor et. al, 2019). Evidence suggests that automated tasks help eliminate primarily blue-collar and clerical work (Levy, 2008). Further, widespread adoption of emerging technologies has reduced the relative bargaining power of low-skilled workers (Madgavkar et al., 2018).

Due to the unequal consequences of automation on high- and low-skilled workers, people may be incentivized to upgrade their skills. Applying an endogenous education model, Prettner and Strulik (2019) show that moving upward to a more future-proof job due to threats of automation motivates a greater number of people to obtain higher education. In the context of accelerated automation and innovation, existing barriers to higher education and limited possibilities for retraining and upskilling for some segments of the labour force can, however, contribute to increasing income and wealth inequality.



Methodology

An occupation can be thought of as a bundle of tasks. Brynjolfsson and Mitchell (2017) developed a rubric to evaluate the potential for ML to substitute human labour for specific work activities. Brynjolfsson et al. (2018) applied this rubric to evaluate the “suitability for machine learning” (SML) of 2,069 direct work activities (DWA) found in O*NET. O*NET is an American occupation information database describing nearly 1,000 occupations. O*NET captures key attributes and characteristics of workers and occupations via standardized job-oriented descriptors and worker-oriented descriptors.

The rubric established by Brynjolfsson and Mitchell (2017) evaluates the criteria required for ML to substitute an occupational task. It consists of 23 statements evaluated on a five-point scale varying from “strongly agree” to “strongly disagree.” For example, the first question in the rubric is:

1. Information needed to complete the task (inputs) and outputs can be explicitly specified in machine-readable format

1: It is very difficult or impossible to identify particular inputs and outputs (e.g. emotions, ideas, impressions)

3: It is possible to create rankings or partial representability of inputs and outputs

5: It is easy to quantify results on a machine/computer (e.g. calculations, concrete inputs and outputs)

FIGURE 3
Averaging task-level SML to occupation level

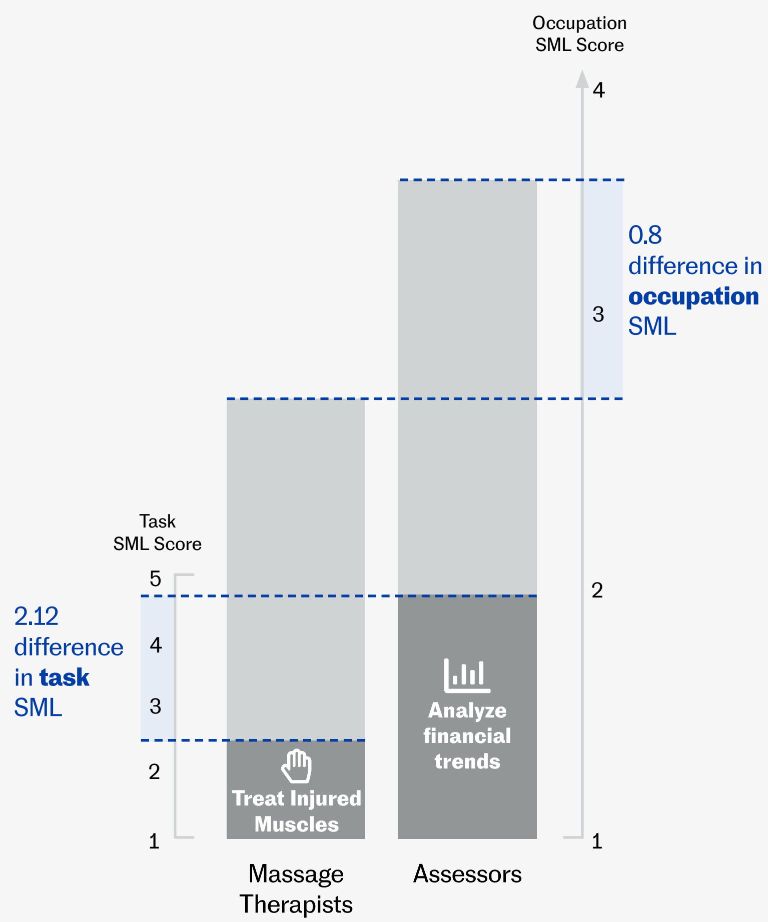
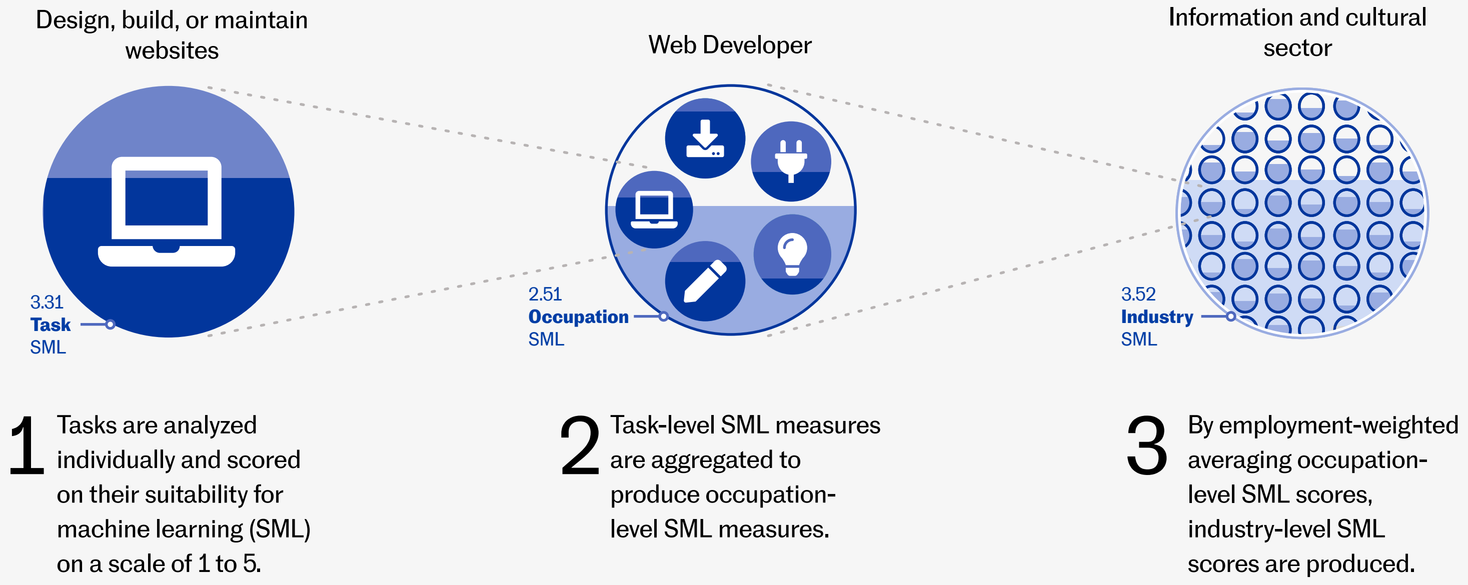


FIGURE 4

Task-based framework for analyzing suitability for machine learning

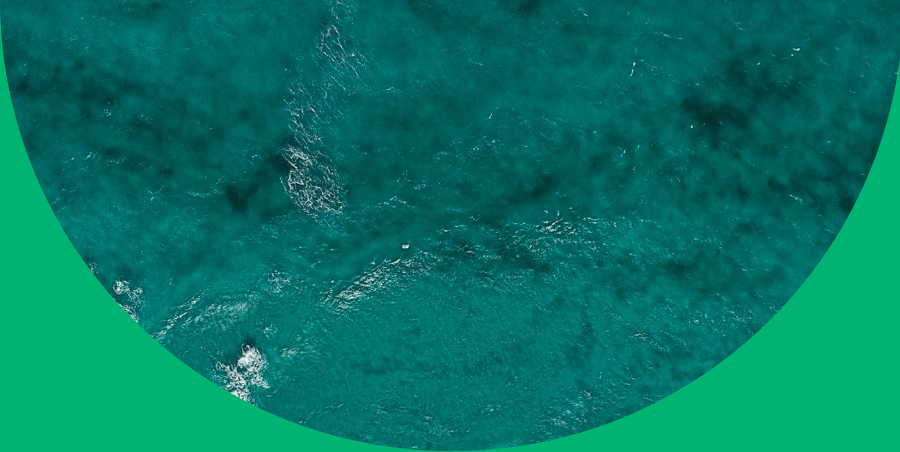


This rubric was applied to each direct work activity (DWA) using CrowdFlower, a human intelligence task crowdsourcing platform. Each DWA was scored by at least seven respondents. In total, 2,069 DWAs were mapped to 18,156 tasks. These tasks were mapped to 964 occupations in O*NET. Task-level SML measures were produced by averaging the SML of the DWAs associated with each task. Occupation-level SML scores were then produced by taking the weighted average, by importance, of the tasks mapped to an occupation (Brynjolfsson et al., 2018).

It is important to note that this rubric focuses solely on the technical feasibility of ML as it currently exists. It is not an evaluation of all forms of automation. As a consequence, as ML evolves, this rubric will need to be updated. Further, “it is silent on the economic, organizational, legal, cultural and societal factors influencing ML adoption” (Brynjolfsson et al., 2018).

The findings presented by Brynjolfsson et al. suggest that nearly all occupations have some tasks that are suitable for ML, while few, if any, occupations are completely suitable to ML. There is significant variation of SML scores between tasks, but little variation of SML score between occupations. In other words, the range of SML scores between tasks is large, as some tasks are very well suited for ML, while others are not appropriate for ML. However, when all the task-wise ratings are aggregated to the occupation level, the difference in SML score between occupations is minimal (see Figure 3). Although the effects of AI will be widespread, it is unlikely that occupations will be completely automated. This signals a need to think about the redesign of jobs and their task composition, instead of the complete automation of jobs.

For the purpose of our analysis, we mapped SML scores from O*NET occupations to Canadian occupations found in the National Occupation Classification (NOC) using a crosswalk we developed. To assess the potential impact of ML across occupations, regional economies, incomes and industries (see Figure 4) in Canada, we used a custom Statistics Canada table of employment statistics from the 2016 Canadian Census and 2019 median hourly wage data from the Government of Canada Job Bank’s Wage report.



Results

We began by mapping SML scores from O*NET occupations to occupations in the Canadian National Occupation Classification (NOC). As described above, an occupation can be interpreted as a bundle of tasks. In this context, a high SML occupation is one where many of its tasks are good candidates for ML, or, in other words, where many of its tasks can be replaced by machine learning. These would be occupations where many tasks meet the following criteria: well-defined inputs, outputs and goals, large data sets either exist or can be created mapping inputs to outputs, is error-tolerant, the task does not change rapidly over time and more. Conversely, a low SML occupation is predominantly composed of tasks where ML would be largely ineffective. For example, tasks involving specialized dexterity, social skills or where tolerance for error is low are just a few of the qualities that would make the current state of ML a poor candidate for.

Occupations with greatest resilience to machine learning

Our findings indicate that the seven occupations in the NOC most resilient to ML are:

1. Massage therapists (NOC 3236)
2. Logging machinery operators (NOC 8241)
3. Plasterers, drywall installers, and finishers and lathers (NOC 7284)
4. Dancers (NOC 5134)
5. Bricklayers (NOC 7281)
6. Athletes (NOC 5251)
7. Land survey technologists and technicians (NOC 2254)

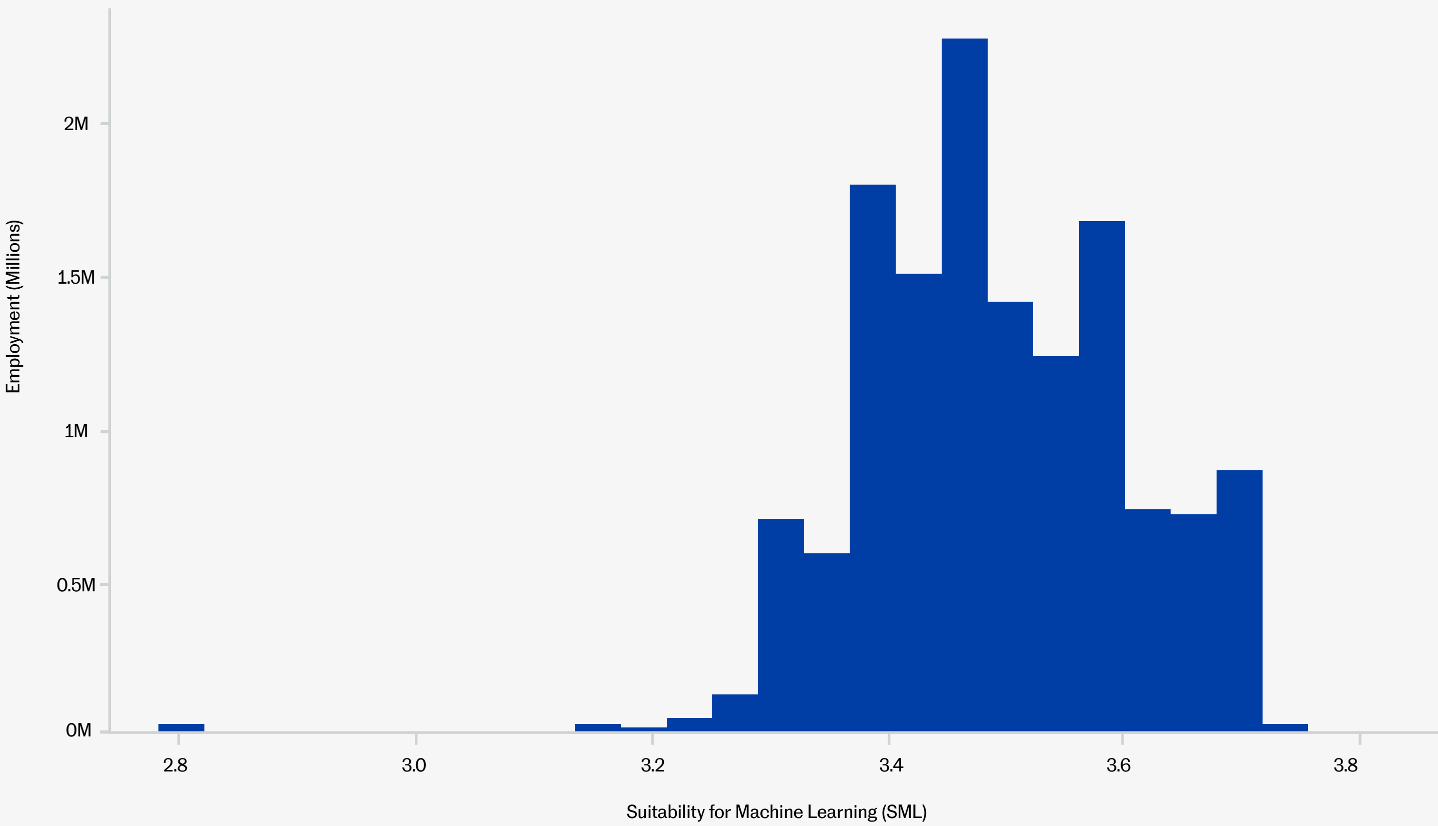
Overall, it is notable that these seven occupations are fairly labour intensive. In addition, among these jobs are many blue-collar jobs (bricklayers, logging machinery operators, plasterers, drywall installers, and finishers and lathers, as well as land survey technologists and technicians). This contradicts Levy (2008), who argued that blue-collar jobs would be most susceptible to automation. This disparity is a result of the nature of SML scores being purely a measure of the technical feasibility for ML to replace an occupational task. Blue-collar jobs may face greater job risk from other advances in technology, such as robotics. Despite potential future risk, this tells us that higher education does not necessarily guarantee a career with greater resiliency to the changing nature of work. At a minimum, it appears that focusing purely on cognitive skills may be misleading when aiming to build a resilient labour force.

Of these seven occupations, massage therapist is the occupation with the least potential for ML to substitute its tasks. This is an understandable finding given that the main duties[6] of a massage therapist are physically intensive and require specialized dexterity in addition to social and interpersonal skills. These tasks are largely unstructured, as is the working environment of a massage therapist. The physical capabilities of ML and its applications to robotics is still fairly clumsy compared to humans.

It also is worth noting that massage therapist is the only occupation in the NOC with a SML score below three. While all

6. The main duties listed in the NOC under massage therapists are as follows: (1) assess clients via movement tests and propose treatment plans (2) explain procedures risks and benefits (3) administer massage techniques (4) suggest home care instructions (5) maintain records of treatment given (6) consult other healthcare professionals.

FIGURE 5 Distribution of Employment by Suitability for Machine Learning



other occupations in the NOC have a SML score greater than three, there is little variance between occupation-level SML scores. This indicates that no occupation is entirely susceptible or unsusceptible to being replaced by ML technologies. As per Brynjolfsson et al. (2018), we must instead begin to think about the redesign of occupational task content rather than the full automation of entire occupations. Which tasks are well-suited to ML and which tasks aren't, and what will the future demand of these tasks look like?

Occupations with greatest suitability for machine learning

Below are the seven occupations with the highest SML scores.

1. Casino occupations (NOC 6533)
2. Graphic arts technicians (NOC 5223)
3. Plateless printing equipment operators (NOC 9471)
4. Boat assemblers and inspectors (NOC 9531)
5. Records management technicians (NOC 1253)
6. Accounting and related clerks (NOC 1431)
7. Accounting technicians and bookkeepers (NOC 1311)

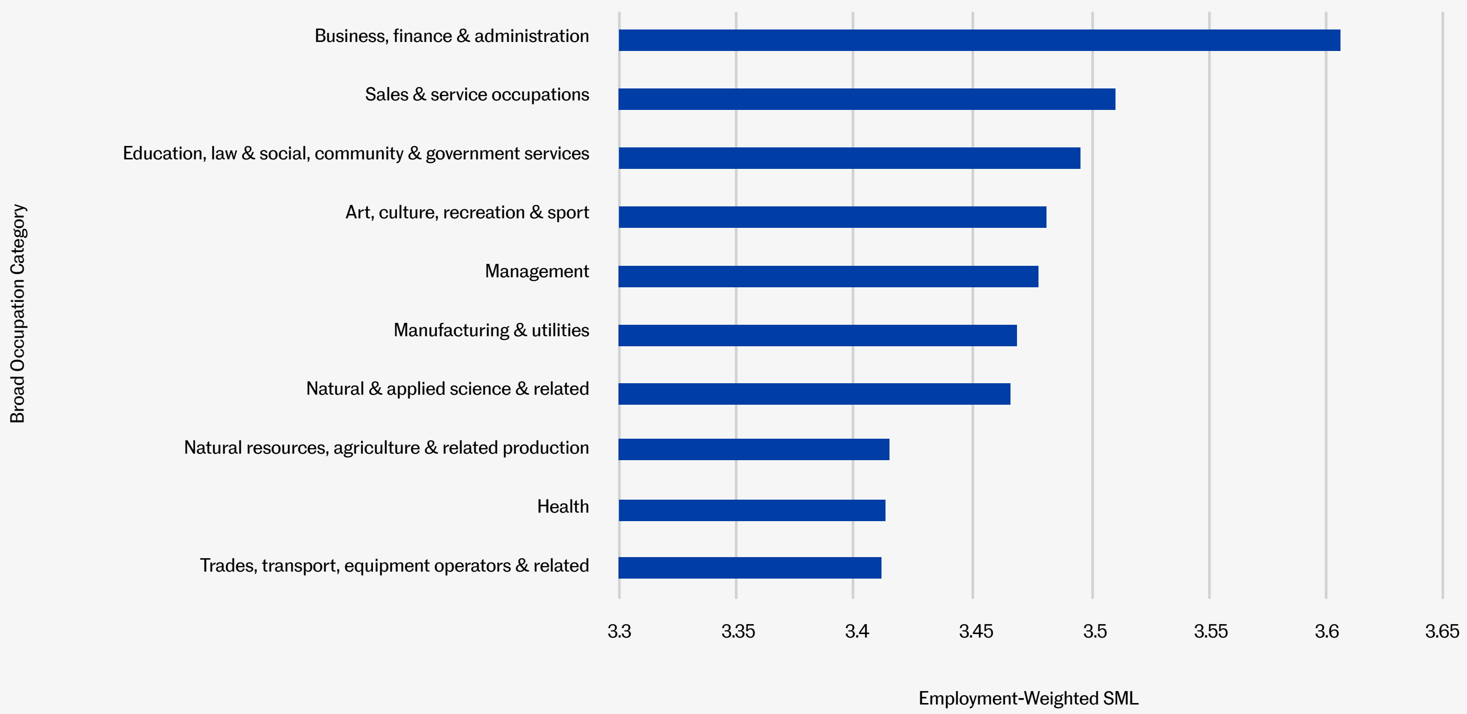
These are occupations with a high volume of tasks well-suited for ML. The top scorer, casino occupations, involve the following duties: (1) operate game tables (2) monitor and assist patrons using slot machines (3) explain rules to patrons and ensure they are followed (4) accept keno wagers (5) determine winners and announce winning numbers (6) calculate and payout winning bets (7) replenish and reset slot machines (8) perform minor repairs to slot machines. It is not difficult to see why these tasks are good candidates for ML — task five for example, determines the winning numbers of a game based on a set of predetermined rules. At its core, ML models are learning functions and therefore would excel at learning how to play numerous games with well-defined rules. Recall, Agrawal, Gans and Goldfarb (2019) recognized that ML excels at prediction, indicating that occupations with the greatest potential to be impacted by ML are those that contain a high proportion of prediction-based tasks. In economics parlance, ML diminishes the comparative advantage humans have over machines (or algorithms) in occupations with a high volume of tasks suitable for ML.

Employment distribution by SML

Despite presenting a ranking of occupations by their SML scores, it is noteworthy that the occupation-level SML range for all occupations in the NOC is from 2.78 to 3.76, inclusive. Hence, while SML is measured on a five-point scale, there is less than one-point difference between the highest and lowest-scoring occupations in the NOC. This means that occupations do not differ significantly in terms of their suitability for ML. Further,

FIGURE 6

Employment-weighted SML by Broad Occupation Category



the occupation-level SML scores are all relatively centered in the middle of the distribution (see Figure 5). There are no occupations lying at either extreme of the SML score range. This indicates that no occupation is completely safe nor at extreme risk of being replaced by ML.

In interpreting these findings, it is important to highlight that this is a result of how occupation-level SML scores are produced, i.e. by aggregating task-level SML scores for the tasks present in a given occupation. A task that is highly suitable for ML could be balanced by one less suited for ML. This requires a need to change how we think about the future of work. As ML can only automate a subset of tasks within a given occupation, it signals a need for the rebundling of tasks within an occupation to complement ML adoption. In addition, we should anticipate the emergence of new tasks as ML is applied.

SML by broad occupation category

The interactive visualization on the landing page of this report allows users to explore and understand the relative impact of ML across occupations as well as provide a sense of the number of Canadians expected to be impacted based on employment by each occupation. Figure 6 shows the employment-weighted SML by broad occupation category. The employment-weighted SML does not vary greatly across occupation categories, with the greatest difference in SML being between business, finance and administration occupations and sales and service occupations. Business, finance and administration occupations are expected to be most impacted among all occupation categories while trades, transport and equipment operators and related occupations are expected to be the least impacted. This is in line with our findings at a sector-level analysis, below.

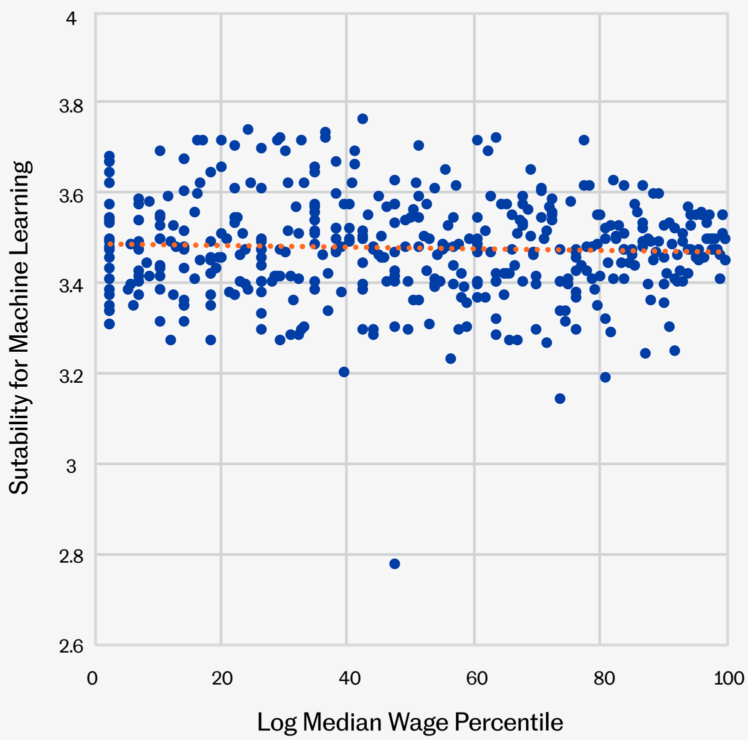
Wage and suitability for machine learning

To determine the impact of automation by income, we plot SML against median hourly wages. Our findings suggest that there is no correlation between hourly wages and SML (see Figure 7). This differs from Brynjolfsson et al. (2019) who found that ML will disproportionately and adversely affect individuals in lower income jobs.

Prettner and Strulik (2019) argue that high-skilled labour is more difficult to automate with the current state of technology. If we assume skill is positively correlated with wage, ie. that high-skilled workers receive greater compensation than low-skilled workers, this analysis contradicts Prettner and Strulik's finding. This could be driven by the fact that our analysis focuses purely on the technical feasibility of the current state of ML, not the current state of technology overall, which Prettner and Strulik consider in evaluating the relationship between skill level and automation. We find there is no correlation between median hourly wage and SML[7], and that the impacts of ML are fairly evenly distributed across wages.

7. Five occupations were omitted from this analysis as they heavily skewed the results due to their extremely high median hourly wages. Four of the five omitted occupations reported median hourly wages higher than $44,000.

FIGURE 7 SML vs. Log Median Wage Percentile



It is important to note, however, that wages differ between cities and industries. This analysis was limited to median hourly wages based on the Toronto region. Further analysis should include wages specific to different industries and regions to more accurately understand the relationship between income and the impacts of ML as it differs by industry and region.

While there is no correlation between SML and wage when looking across all NOCs, this observation does not hold when analyzing this relationship within broad occupation categories. Lower wage occupations in business, finance and administration, manufacturing and utilities, education, law, and social, community and government services are expected to be more impacted by ML while conversely, higher wage occupations in management, natural and applied sciences, arts, culture and recreation, natural resources and agriculture, and sales and service are expected to be more impacted (see Figure 8).

FIGURE 8 SML vs. Median Hourly Wage, by Broad Occupation Category

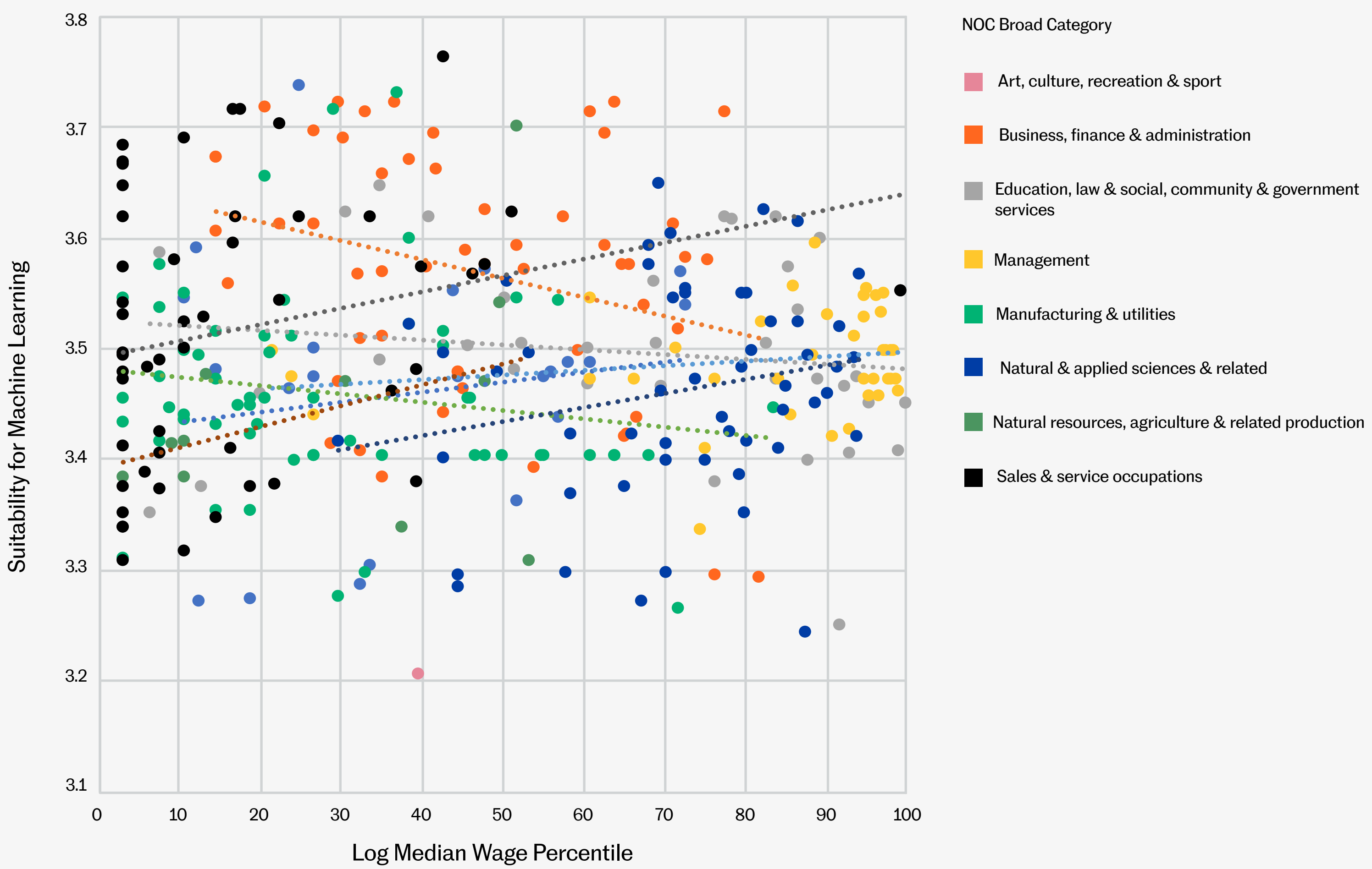
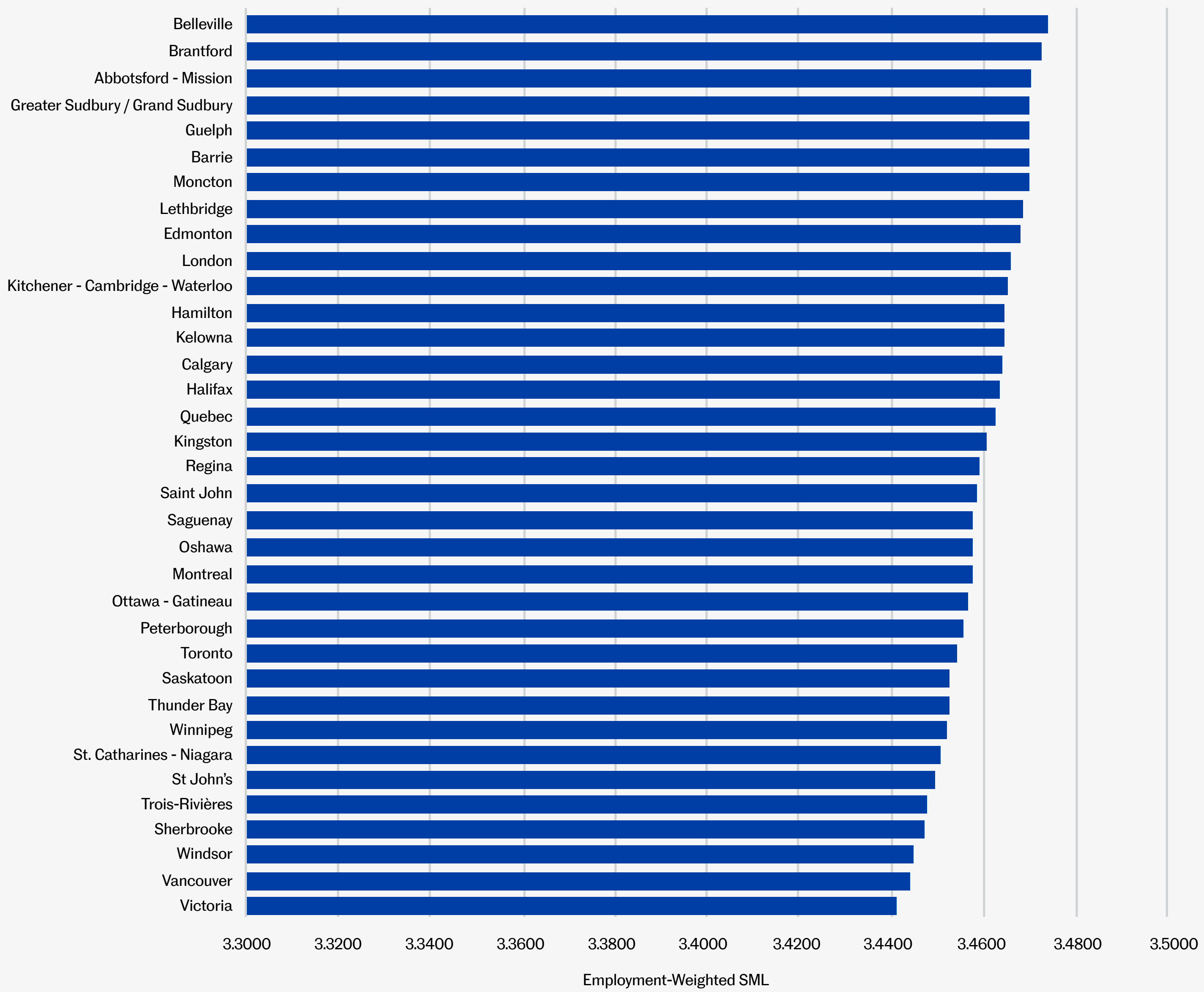


FIGURE 9

Employment-Weighted SML by CMA



Regional exposure to machine learning

To determine the impact of ML across Canada's census metropolitan areas (CMA), we calculate employment-weighted SML scores for each metropolitan area[8]. Brynjolfsson et al. (2019) find that people in smaller cities are more likely to be affected by ML than those in larger ones. Our analysis reveals a weak negative correlation (-0.18) between the employed population of a CMA and its SML score. While the negative correlation indicates that CMAs with a smaller labour force are more likely to be affected by ML, the strength of this relationship is weak. Further, there is little variance in CMA-level SML scores with only a 0.03 difference between the CMA with the highest employment-weighted SML and the lowest employment-weighted SML. Figure 9 shows employment-weighted SML by census metropolitan areas.

Further analysis of the labour force should be performed at a census subdivision or census agglomeration (CA) level (of areas with populations of 10,000 or more). Such geographical granularity would be required to understand the impact of ML within various regions in Canada as well as to understand the relationship between a region's employment size and composition and their exposure to ML.

Despite the low variance in CMA level scores, it is worthwhile to understand the relative impact ML could have on various regions in Canada, however, it is difficult to identify specific drivers due to the diversity of occupations in the labour force composition in these metropolitan areas. Cities are made up of many different types of occupations and the proportion of employment in

8. Statistics Canada defines a CMA as an area consisting of one or more neighbouring municipalities situated around a core with a total population of at least 100,000 of which 50,000 or more live in the core.

each occupation is relatively small. For example, in the Toronto metropolitan area, retail salesperson is the occupation employing the highest number of individuals. However, it only accounts for 4 percent of the employed population. The top 30 occupations by employment account for 30 percent of the employed population in the Toronto CMA.

Sectoral exposure to machine learning

We calculate employment weighted SML scores by two-digit NAICS sectors (see Figure 10). We find that finance and insurance, management of companies and enterprises, and retail trade are relatively more exposed to the impacts of ML compared to agriculture, forestry, fishing and hunting, and construction. This is in line with our findings at a broad occupation level, where we saw that ML had relatively little impact on trades, transport, equipment operators and related occupations.

As such, our findings at the sector level are a result of the relatively higher representation of these occupations in these sectors. Brnyjolfsson, Rock and Tambe's findings suggest that accomodation and food services, transportation, warehousing, and retail trade will be the most impacted sectors in the United States, while education services will be the least impacted, followed by healthcare and social assistance. This is driven by

FIGURE 10

Employment-Weighted SML by Sector

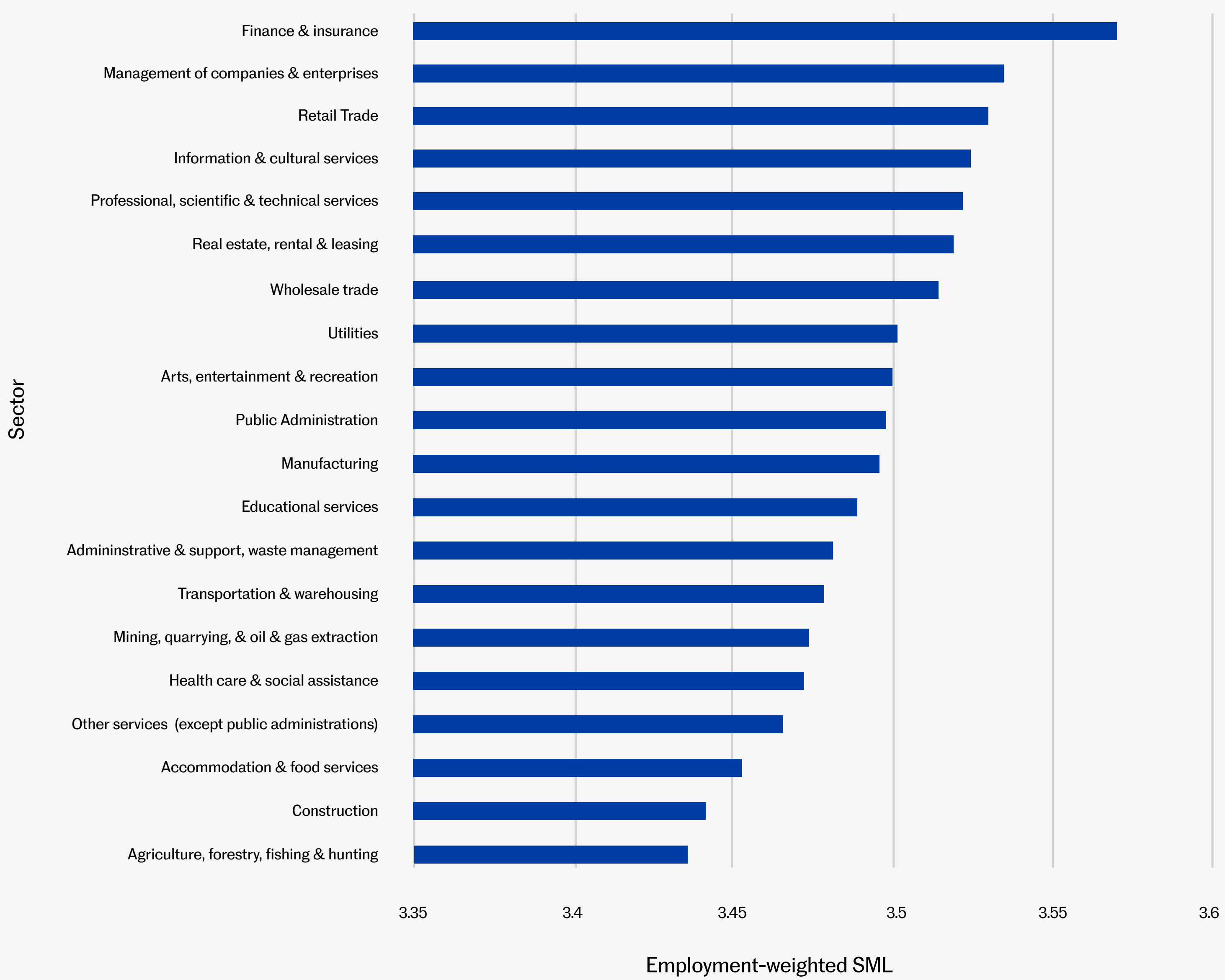
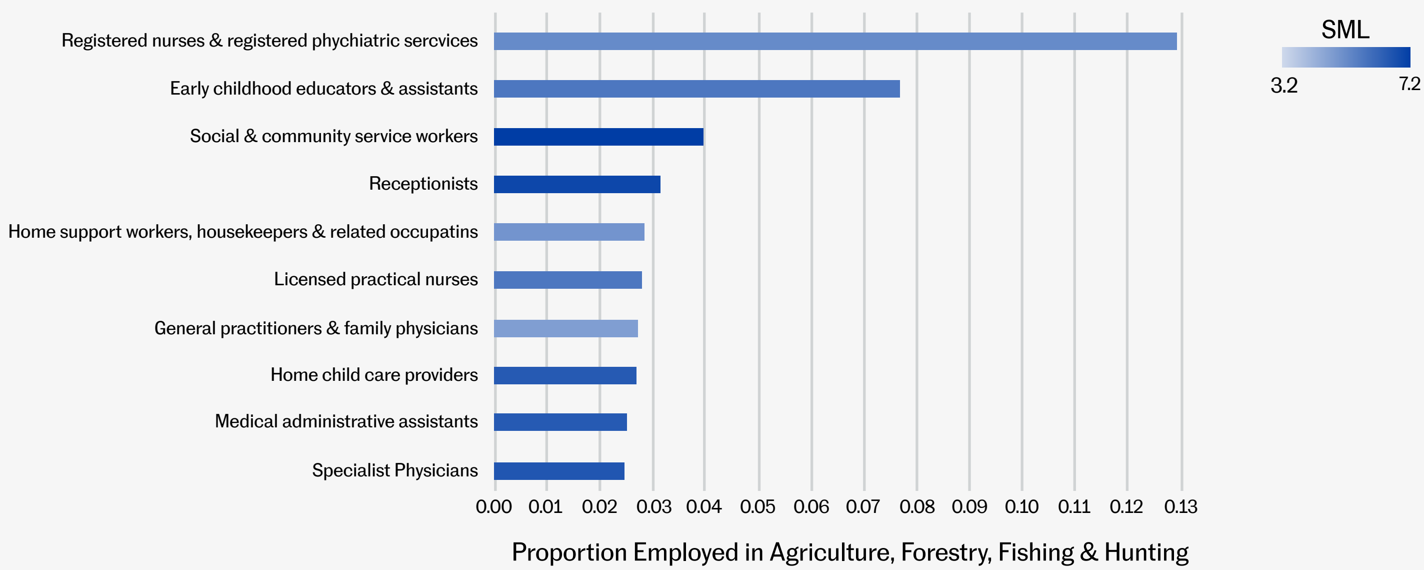


FIGURE 11

Top 10 Occupations by Employment in Agriculture, Forestry, Fishing & Hunting



the employment distribution by occupation in each industry. As this analysis is only for Census Metropolitan Areas in Canada, these sector-level SML measures do not account for employment within smaller regional areas in Canada, which could impact the overall sector-level scores.

To understand the drivers behind these sector-level scores, we look at the top ten occupations by employment and SML in the agriculture, forestry, fishing and hunting, finance and insurance, and retail trade sectors.

Figure 11 shows the ten occupations with the highest employment across all CMAs in the agriculture, forestry, fishing and hunting sector. Together, these occupations account for two-thirds (67 percent) of employment in this sector. Managers and general farm workers account for 45 percent of employment in the sector. All 10 of these occupations have relatively low SML when compared to other NOCs. The employment-weighted SML for these ten occupations in agriculture, forestry, fishing and hunting is 3.41 while for finance and insurance, the SML is 3.58.

Figure 12 shows the top ten occupations in finance and insurance, the sector with the highest employment-weighted SML score. These ten occupations account for more than half (56 percent) of all of employment in this sector, across all the CMAs. As can be observed in the bar plot in Figure 12, the

FIGURE 12

Top 10 Occupations by Employment in Finance & Insurance

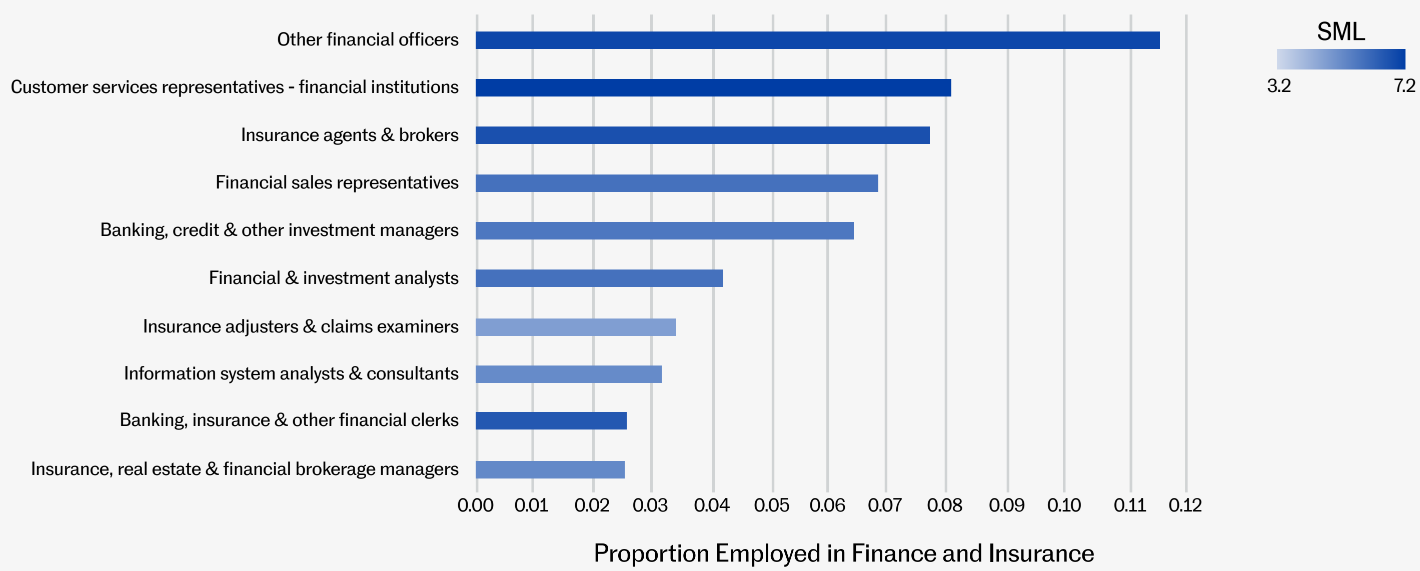
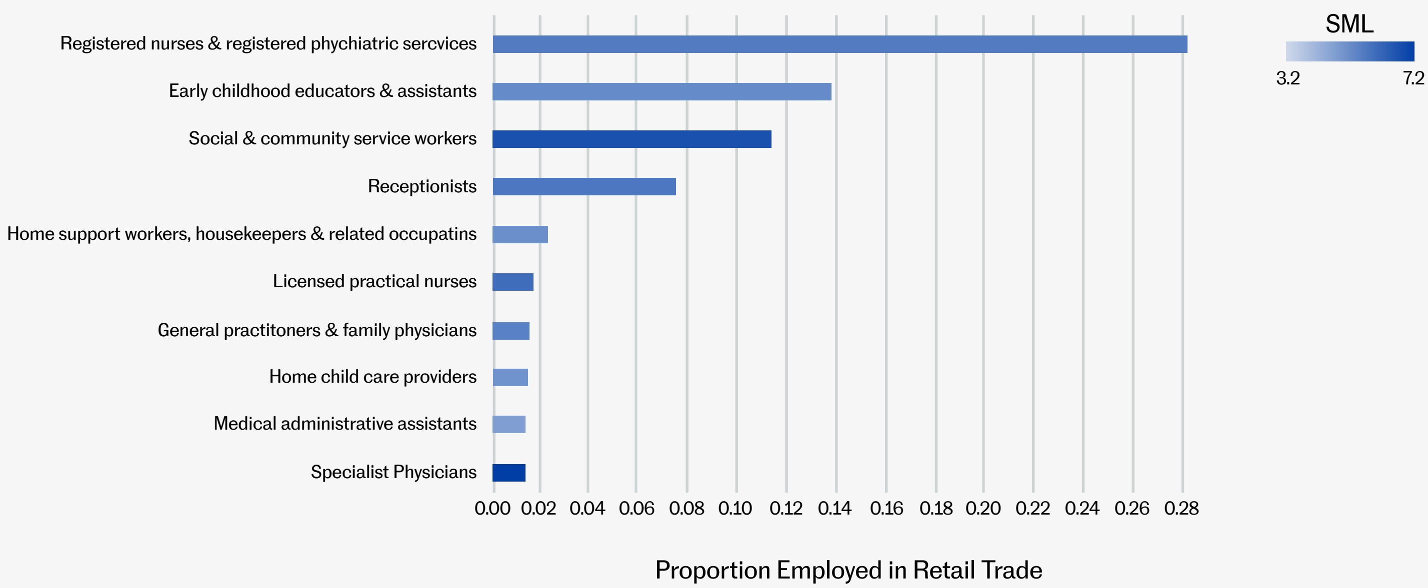


FIGURE 13

Top 10 Occupations by Employment in Retail Trade



majority of these occupations have relatively higher SML than other NOCs, especially when compared to occupations in the agriculture, forestry, fishing and hunting sector. This illustrates the relative exposure between these two sectors, with finance and insurance facing greater potential for redesign of work than agriculture, forestry, fishing and hunting. Relatively higher exposure to ML, however, does not necessarily imply a reduced demand for labour, as ML could be used to complement labour, increase demand by lowering costs or change demand by changing overall income, or change information flows and drive a reorganization of work (Brynjolfsson et al., 2019).

Retail trade is the sector with the highest employment across Canada. Employment in this sector is heavily concentrated in a few occupations, with more than half of all employment in the following three occupations: retail salespersons, retail and wholesale trade managers, and cashiers. The top ten occupations account for 72 percent of employment. Figure 13 displays the relative impact of ML across the top ten occupations in retail trade by employment. We can observe greater variance in SML in the these ten occupations in this sector when compared to agriculture, forestry, fishing and hunting, and finance and insurance.

Limitations and Further Research

While this paper provides some initial insights into the potential for ML to impact Canadian occupations and industries and highlights some of the differences between Canada and the U.S., there is opportunity to strengthen this analysis.

The use of a O*NET-NOC crosswalk with one-to-one mapping limited our ability to understand the potential impact of ML on Canadian occupations. This further contributes to limitations in understanding the impacts of ML at a regional and industry level. Identified below are the specific concerns we have with using a crosswalk.

1. Loss in fidelity: As there are roughly half as many NOCs as there are O*NET occupations, a one-to-one mapping results in a loss of information.

    a. Inconsistent mapping: A many-to-many mapping is more appropriate in most cases. In cases where a single NOC maps to many O*NET occupations, we do not have information on the proportional representations of these O*NET occupations within the NOC. Further, this composition will change by industry and region. This complicates the mapping of SML scores between taxonomies.

2. Task variance: The use of a crosswalk makes an implicit assumption that occupational tasks overlap between matching occupations in both taxonomies. In this case, SML measures were produced for occupational tasks, then aggregated up to an occupation level. A more ideal mapping would be between tasks of the two taxonomies.

To address the above limitations, we recommend that further research apply Brynjolfsson and Mitchell's rubric to evaluate the suitability for ML of the "main duties" found in each NOC. Further, the expansion of this rubric to assess the legal, social and organizational aspects of ML applicability and adoption in addition to other forms of automation, such as robotics, will allow for a more comprehensive understanding of impact of the Fourth Industrial Revolution on Canadian occupations, industries and the economy.

Once this rubric is applied to evaluate the SML for main duties within NOCs, the application of those results on more recent labour force data, including more granular geographic information, would allow for a more current and representative view on the potential impacts of ML on the Canadian labour force and economy. This research can be further augmented with educational attainment data to help individuals identify learning opportunities to upskill and retrain in order to build resilience in the changing nature of work. Including data from the Canadian Occupational Projection System will also allow for a deeper understanding of future labour market conditions on an industrial and occupational basis.

We believe there is potential to increase the understanding of the impacts of ML among job seekers, workers and employers. The translation of this information to tasks found in job descriptions would bring value to these individuals and organizations. This would involve modelling the relationship between the text describing occupational tasks in O*NET and its SML in order to predict the SML for tasks not found in O*NET. For employers, this information could highlight opportunities for innovation within their organizations as well as identify opportunities to upskill and retrain their own workforce to complement ML adoption. Further, this could help inform career decisions of job seekers and workers by providing information around the SML for tasks within their current and prospective jobs. While this is technically very difficult to achieve, we believe there is opportunity to communicate this information beyond the research sphere and provide the value of this information to the broader public.

Conclusion

Artificial intelligence (AI), which is intelligence demonstrated by machines, is widely recognized as being the catalyst of the Fourth Industrial Revolution. Machine learning (ML), a sub-discipline of AI, has enabled the widespread deployment of AI and is expected to transform numerous occupations, industries and economies. This paper assesses the potential impacts of ML on Canadian occupations and industries by applying a methodology developed by Byrnjolfsson et al. (2017 & 2018) to measure the "suitability for machine learning" (SML) of occupations.

Our main findings suggest that ML will impact the labour force relatively equally across occupations regardless of income; however the relative impact across Canada's metropolitan areas is correlated with population and different industries are at greater exposure based on their composition of occupations and the nature of the tasks required by those occupations. These findings differ from Brynjolfsson et al. (2019) whose research suggests that lower wage occupations will be disproportionately affected as well as those employed in the retailing and transportation industries.

We created a O*NET-NOC crosswalk to map SML scores between American and Canadian occupations. This crosswalk is a one-to-one mapping between the O*NET and NOC taxonomies where a many-to-many mapping is more appropriate. However, a many-to-many mapping requires understanding of the proportional composition of the O*NET occupations mapped to a given NOC, without which we are unable to map SML scores between the two taxonomies. Further, with the use of a crosswalk, there is an unvalidated assumption that tasks overlap between matching occupations in both taxonomies.

Further research should look to apply Brynjolfsson and Mitchell's rubric to evaluate the suitability for machine learning of the "main duties" found in each NOC in order to more accurately evaluate the impact of ML on Canadian occupations. The expansion of this rubric to include an assessment of the legal, social and organizational aspects of ML applicability and adoption in addition to other forms of automation, such as robotics, will allow for a more comprehensive understanding of impact of the Fourth Industrial Revolution on Canadian occupations, industries and economy.

References

Acemoglu, D., & Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. Journal of Political Economy 128, no. 6: 2188-2244. Retrieved from https://www.nber.org/papers/w23285

Acemoglu, D., & Restrepo, P. (2018). Artificial Intelligence, Automation and Work. NBER Working Paper Series. Retrieved from https://www.nber.org/papers/w24196

Agrawal, A., Gans, J., & Goldfarb, A. (2019). Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction. The Journal of Economic Perspectives, 33(2), 31-50. Retrieved from www.jstor.org/stable/26621238

Autor, D., Dorn, D., Katz, L., Patterson, C., & Van Reenan, J., (2019). The Fall of the Labor Share and the Rise of Superstar Firms. Quarterly Journal of Economics. Retrieved from https://economics.mit.edu/files/12979

Brynjolfsson, E., Rock, D., & Tambe, P. (2019). How Will Machine Learning Transform the Labour Market? Governance in An Emerging New World. Spring Series, Issue 619. Retrieved from https://www.hoover.org/research/how-will-machine-learning-transform-labor-market

Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy?. American Economic Association Papers and Proceedings, Vol. 108, pp. 43-47. Retrieved from https://www.aeaweb.org/articles?id=10.1257/pandp.20181019

Brynjolfsson, E., & McAfee, A. (2017). What's Driving the Machine Learning Explosion? Harvard Business Review. Retrieved from: https://hbr.org/2017/07/whats-driving-the-machine-learning-explosion

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. Science Magazine, Vol. 358, Issue 6370, pp. 1530-1534. Retrieved from https://www.cs.cmu.edu/~tom/pubs/Science_WorkforceDec2017.pdf

Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. New York, NY: WW Norton.

Bughin, J. (2018). Why AI isn't the death of jobs. MIT Sloan Management Review, 59(4), 42-46. Retrieved from http://myaccess.library.utoronto.ca/login?url=https://search-proquest-com.myaccess.library.utoronto.ca/docview/2057246267?accountid=14771

References Continued

Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Allas, T., Dahlström, P. and Trench, M. (2017). Artificial Intelligence: The Next Digital Frontier?. McKinsey Global Institute. Retrieved from https://www.mckinsey.com/~/media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx

IGM Forum (2017). Robots and Artificial Intelligence. Chicago Booth. Retrieved from http://www.igmchicago.org/surveys/robots-and-artificial-intelligence/

Job Bank. (2019). Wage report. Government of Canada. Retrieved from https://www.jobbank.gc.ca/wagereport/location/geo9219

Kaplan, A., & Haenlein, M., (2018). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations and implications of artificial intelligence. Kelley School of Business, Indiana University. Elsevier Inc. Retrieved from https://www.sciencedirect.com/science/article/pii/S0007681318301393?via%3Dihub

Levy, F. (2018). Computers and populism: artificial intelligence, jobs, and politics in the near term. Oxford Review of Economic Policy, 34(3), 393–417. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3091867

Madgavkar, A., Ramaswamy, S., & Gumbel, P. (2018). What will automation mean for wages and income inequality? New York: McKinsey & Company, Inc. Retrieved from https://www.mckinsey.com/featured-insights/future-of-work/what-will-automation-mean-for-wages-and-income-inequality

McKinney, S.M., Sieniek, M., Godbole, V. et al. (2020). International evaluation of an AI system for breast cancer screening. Nature 577, 89–94. Retrieved from: https://deepmind.com/research/publications/International-evaluation-of-an-artificial-intelligence-system-to-identify-breast-cancer-in-screening-mammography

Morikawa, M. (2020). "Heterogeneous Relationships between Automation Technologies and Skilled Labor: Evidence from a Firm Survey", RIETI Discussion Paper, 20-E-004. Retrieved from https://voxeu.org/article/relationship-between-automation-technologies-and-skilled-labour

References Continued

Prettner, K., & Strulik, H. (2019). Innovation, automation, and inequality: Policy challenges in the race against the machine. Journal of Monetary Economics. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304393219301965

Shoham, Y., Perrault, R., Brynjolfsson, E., Clark, J., Manyika, J., Niebles, J. C., Lyons, T., Etchemendy, J., Grosz, B., & Bauer, Z., (2018). The AI Index 2018 Annual Report. AI Index Steering Committee, Human-Centered AI Initiative, Stanford University, Stanford, CA, December 2018.

Statistics Canada. (2016). 2016 census of Canada employment and median income by occupation, industry, age, sex, and census metropolitan area [Data file]. Ottawa

World Economic Forum. (2020). Fourth Industrial Revolution. Retrieved from https://www.weforum.org/focus/fourth-industrial-revolution#