Article

INDUSTRY & Higher Education

Industry and Higher Education
2022, Vol. 36(1) 63–73

DOI: 10.1177/0950422221990990
journals.sagepub.com/home/ihe
SAGE

An investigation of skill requirements in artificial intelligence and machine learning job advertisements

Amit Verma and Kamal Lamsal
Missouri Western State University, USA

Payal Verma
University of Kansas, USA

Abstract

Due to the advent of big data and efficient computational resources, artificial intelligence (AI) and machine learning (ML) have seen massive growth in recent years. Informatics degree programs are scrambling to meet the ever-increasing market demand of such professions. To explore the skillsets required for AI and ML positions, the authors conducted a content analysis of online job advertisements posted on Indeed.com. They present a ranking of the relevant skills for the two positions. Further, they performed a pairwise comparison between AI and ML positions. Overall, it was observed that technical skills like data mining, programming, statistics and big data are more valued for ML positions than for AI positions. On the contrary, AI positions tend to be more generic, with an emphasis on communication skills. These clearly defined skills can be valuable for the hiring process as well as to revamp existing course curricula to cater to the increasing market demand.

Keywords

Artificial intelligence, content analysis, curriculum design/development, information systems, machine learning, skill requirements

According to the LinkedIn's 2020 Emerging Job Report (Romeo, 2020), artificial intelligence (AI) and machine learning (ML) jobs have been identified among the top emerging jobs of the year. The demand for these positions has grown by 74% annually in the last 4 years, made possible by the rapid advancement in new technologies and big data. These findings are echoed by a Gartner Study (Meulen and Petty, 2017) which forecasts massive disruption of AI/ML in all industries (Wang and Siau, 2019). Both studies argue that the demand for AI/ML jobs is widespread. Despite the high demand, however, there is a shortage of skilled talent, not only in the information technology (IT) industry but also in other traditional fields such as marketing, finance, healthcare and supply chain (Wilson et al., 2017). Thus, hire needs are spread across many industries, with most jobs requiring some knowledge of AI/ML technologies. Due to the interchangeable use of AI and ML for job advertisement and hire, there is a need for better stratification of the job skills between the two to make the hiring process better defined and more efficient. Hence, one of the purposes of this research is to establish the similarities and differences between the two using the developed skill classification framework.

Recruiting for these AI/ML jobs in the industry can be extremely tough without training potential recruits early on in an academic setting. More specifically, new training modules are needed to prepare skilled employees for a competitive marketplace (Romeo, 2020). Skill development and technology adoption typically begin at the college or university level. With this realization, an overall increase in the number of analytics programs offering AI/ML skill training is becoming apparent. Due to ever-changing technology standards, there is a need for universities to continuously

Corresponding author:
Amit Verma, Craig School of Business, Missouri Western State University, 315L Popplewell Hall, 4525 Downs Drive, Saint Joseph, MO 64507, USA.
Email: averma@missouriwestern.edu

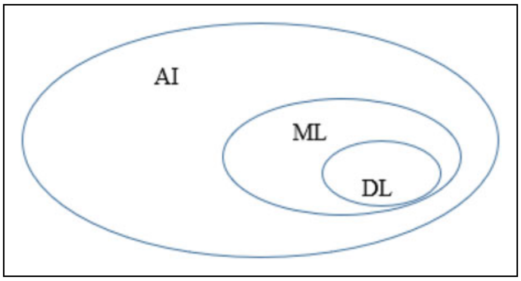

Figure 1. Hierarchy of different domains.

adapt their existing curricula to stay relevant for the skill requirements of these jobs.

There is also a growing need to clarify the definitions of job categories related to AI/ML. The current study attempts to address this by utilizing the associated skill categories. For this purpose, we examine the similarities and differences between AI and ML jobs with respect to different skills. A content analysis technique was used to extract keywords from online job postings to determine the frequencies of specific skills corresponding to each skill category. Using this, we identified the most critical skills as well as the associated skill category in our data sample and were able to assign relative importance to each skill for AI/ML jobs. The findings can also assist in curriculum design such that the most valuable skills are emphasized in the existing analytics curricula. Furthermore, the specific AI/ML courses can be updated to reflect current in-demand software tools and applications.

The job positions that are part of this study represent professional areas of AI and ML. As discussed above, these two domains are closely interrelated. At present, the two terms are ambiguously defined and are often used interchangeably in the literature. According to Goskel and Bozkurt (2019), Woschank et al. (2020) and Wang and Siau (2019), the AI field deals with the development of technology capable of performing tasks related to human intelligence. More specifically, AI is the field of computer science dedicated to the creation of intelligent systems that try to mimic human behavior based on stipulated rules and algorithms (Jakhar and Kaur, 2020). ML and Deep Learning (DL) methods are integral parts of AI. This framework is summarized in Figure 1.

ML, although used interchangeably with AI, is a branch of AI that uses automated learning without any human intervention for a better understanding of data (Agrawal et al., 2019; Woschank et al., 2020). ML models are usually applied to large datasets for pattern recognition, classification and prediction tasks (Mitić, 2019). This is usually achieved by minimizing an error or loss function. Depending on whether the output labels are known or unknown in advance, ML is further classified as supervised or unsupervised ML, respectively. DL is a subset of ML that involves higher-accuracy models requiring computationally intensive resources like graphical processing units (Jakhar and Kaur, 2020). The algorithmic components for DL are typically neural networks, with important application areas like natural language processing, speech recognition, image recognition, recommender systems, etc. We would also like to emphasize that data/business analytics is different from AI. In data analytics projects, the focus is mostly on descriptive statistics based on historical data (see Verma et al. (2019) for more details). Therefore, this stream is also referred to as descriptive analytics. Sometimes, we could also use past data to predict what might happen in the future. These predictions rely on human interventions to understand data, identify trends and create and test hypotheses. This stream is called prescriptive analytics. On the other hand, AI/ML systems are self-reliant and typically involve minimal supervision.

The interest in AI and ML technologies has been growing since the last decade and universities would benefit from an in-depth study of the competency requirements in these fields. In this context, our aim is to answer the following research questions: (1) Which skills are required for AI and ML professionals? (2) What are the similarities and differences between the two job positions?

In this study, we analyzed job advertisements across the USA and provided a skills breakdown for the AI/ML positions. Using this perspective, we offer an alternative point of view of employers' expectations. Using the sorted skills list based on relative frequency, educators can reconfigure existing informatics programs so that the emphasis is on the skillsets required in the job market. However, we do not provide specific recommendations in terms of curriculum design, reserving that topic for future research.

We propose to study only the "demand side" of the US job market by analyzing the job advertisements posted by employers. Future research could consider the gaps in the "supply side" by studying relevant university programs: this could entail a study of the course descriptions or course syllabi or by a survey of administrators and faculty members responsible for the design of such programs. Another research opportunity could explore the localization effects and study the differences in requirements across multiple countries.

Our main contributions are the following:

1. We highlight the key technologies needed for AI and ML positions and summarize the similarities and differences between the two. This is based on our competence framework, which presents a broad classification of required skills sought by the employers in the two areas.
2. Our overview can provide guidance to human resources teams for the creation of job advertisements. It could also lead to the design of specialized in-demand training modules for working professionals in the industry.

3. Educational institutions can incorporate our findings into their curricula. The results could be utilized to create new courses or redesign existing course offerings. Furthermore, students interested in a related career could highlight these in-demand skills in their job applications.

This study supplements the existing literature on curriculum design in informatics programs by identifying key in-demand skills for two widely popular job categories. In addition to assisting academic institutions better align their coursework with the demands of the jobs, the results of our study could be utilized by employers seeking highly skilled workers in AI/ML. The results could also help in the curation of more consistent and better-defined job positions, highlighting the needed skillsets.

The paper is organized as follows. First, we discuss the impact of AI/ML on different disciplines. Next, we examine the stream in the literature dealing with employers' perspectives of job requirements. We also investigate the pedagogical perspective, which provides a link between job requirements and degree programs. The classification framework, research method and dataset are then discussed, followed by a discussion of the results focusing on each skill's relative importance corresponding to a specific job position. Additionally, we examine the similarities and dissimilarities between AI and ML positions. Finally, we outline the conclusions and directions for future research.

Literature review

The role of AI/ML has been paramount in technological development leading to enhanced industry output (see Lu (2019) for a complete survey on AI models and application areas). Vochozka et al. (2018) estimated that the economic impact of AI-based solutions on GDP would be around 24 billion dollars by 2035. Those authors also document the estimated number of AI use cases, job impact, and global merger and acquisition activity. Wamba-Taguimdje et al. (2020) studied the influence of AI on firm performance. Their results highlighted that AI could benefit processes in finance, marketing and administration. They also found that the proper implementation of AI technology requires a reconfiguration of the existing business process.

The impact of AI on education has also been well studied. Researchers have identified both benefits and problems associated with its use. For example, McArthur et al. (2005) summarized current AI applications in education, such as the Intelligent Tutoring System, and predicted future use cases. They emphasized that a widescale implementation of AI in the classroom could be problematic. On the other hand, Roll and Wylie (2016) captured the evolution of AI technology adoption in the classroom through a bibliographic analysis, and argued that integrating AI into the learning environment could facilitate the performance of various tasks and the achievement of process-oriented goals. Along the same lines, Chassignol et al. (2018) identified four different streams of application: educational content, innovative teaching method, enhanced assessment tool and improved communication channel. More recently, Shar (2019) examined students' perspective of AI adoption in education through a survey-based study and found both positive and negative impacts. They established that variables like age and gender influenced individual performance and AI adoption.

Over the last couple of decades, several attempts have been made to integrate AI into the existing academic curriculum. For instance, utilizing computer-based modeling software, Schlanberger (1989), in one of the first studies to explore the integration of AI into the business curriculum, provided a framework for integrating AI applications and methodology into the MBA program. Schlanberger (1989) proposes a framework for information systems planning based on four stages of the growth model, which can be utilized for curriculum design and resource planning. The lifecycle of an information system adoption is divided into four stages: initiation, expansion, formalization and integration. This study also highlights the support needed as the university passes through each stage of this lifecycle, along with providing examples of AI application development software and course syllabi.

Similarly, Aiken (1991) argued for the importance of AI foundations in mathematics, science, engineering and computer science curricula and also emphasized the value of a laboratory component or hands-on experience with current AI tools. Holder and Cook (2001) implemented this framework in the classroom using a web-based multimedia tool to develop and test AI algorithms: this simulation tool drew a lot of interest and elevated the confidence of students in the class content. Baldwin-Morgan (1995) noted that AI topics should be integrated into the accounting curriculum, and students should be exposed to AI tools before encountering them in the workplace. The author argued that students should be motivated with examples using expert systems and neural networks in auditing, taxation, cost accounting, management accounting and financial accounting. Testa et al. (2001) designed a tool called Emergent-Design to integrate AI concepts into the architecture curriculum. This allowed students to learn about diverse roles in a design team while being exposed to complex and dynamic design scenarios. More recently, Chedrawi and Howayeck (2019) argued that the Fourth Industrial Revolution or Industry 4.0 is attributable to the technological advancements of AI. They proposed a model to integrate AI into the accreditation program of the Association to Advance Collegiate Schools of Business (AACSB) through expert systems. This is expected to create a reliable and efficient process, minimizing time, costs and errors.

The impact of Industry 4.0 on jobs is also recognized by Romeo (2020), who noted that rapid advancements in AI had dramatically changed the hiring landscape. The shortage of skilled AI talent has forced businesses to look for innovative solutions like upskilling. In this practice, firms retrain their existing employees to develop expertise in AI. The influence of Industry 4.0 on job profiles and skills was also explored in Fareri et al. (2020) using text mining. They analyzed job descriptions from a multinational company, Whirlpool, with jobs and skills classified according to a dictionary of technology and methods. The authors provided measures to estimate the Industry 4.0 readiness of employees. Pejic Bach et al. (2020) also studied Industry 4.0 job descriptions, using text mining to develop a job profile that highlighted the required skills. The first group of these job profiles focused on the Internet of Things and smart production design and control. The second group was more generic, with application areas like supply chain, customer service and enterprise software.

In addition to Industry 4.0, other related fields of interest are business analytics, data analytics and big data. Numerous studies have explored the skill requirements in these professions using job advertisements. For example, Debortoli et al. (2014) studied the difference between big data and business intelligence job advertisements from an online employment portal. Their approach was to develop a competency framework based on a Latent Semantic Analysis of the text descriptions, and they found that business skills were as important as technical skills for both job positions. Gardiner et al. (2018) analyzed 1,216 job advertisements including “big data” in their title. Their research, built on a framework of skill categories, found that analytical skills and soft skills were in high demand in addition to traditional hard skills. Lovaglio et al. (2018) web scraped information and communication technology (ICT) and statistician Italian job advertisements between June and September 2005, with the aim of defining these positions in terms of the skills required. They also wanted to differentiate the statistician from the ICT positions. Their results indicate that statisticians require more analytical and computing skills than soft skills. Verma et al. (2019) provided a skill breakdown for related professions, such as Business Analyst, Business Intelligence Analyst, Data Analyst and Data Scientist, using content analysis of job advertisements. They concluded that the skills essential for these positions were decision making, organization, communication and data management. More recently, Anton et al. (2020) studied the competencies required for three occupation fields: data science and engineering, software engineering and development, and business development and sales. Their mixed methods approach combined a content analysis of scientific literature with text mining of job advertisements. They presented the critical managerial and technical skills required for these positions.

The current study contributes to the existing literature by categorizing the AI and ML jobs in the analytics domain. AI and ML jobs have received limited attention, and therefore an in-depth analysis of the required skills is warranted. Also, we conduct a pairwise analysis of AI and ML jobs and discuss the similarities and differences between the two. Thus, this study aims to shed light on a significant job category in the Industry 4.0 era. Lastly, the study contributes to the existing literature by introducing a novel set of skill categories using a skill classification framework for these two positions.

Research method

Classification framework

In this section, we describe how we developed a classification framework consisting of skill categories and a related set of skills. These skills were mapped against each job position. We used the comprehensive list of skills for various business and data analytics positions provided by Verma et al. (2019). We added more skills categories with associated skills to reflect the current market requirements for AI/ML positions. For this purpose, we manually investigated 100 job advertisements for each type. We asked three independent raters to validate our classification framework by placing skills into appropriate skill categories. The initial inter-coder reliability based on the alpha coefficient was observed to be 0.88. The skills and associated skill categories, along with sample keywords, are presented in Table 1.

Technique

Our technique of data collection and extraction is summarized in Figure 2. We used data from a popular online job portal, Indeed.com. The collected data, based on US job advertisements, was analyzed using content analysis (Neuendorf, 2016), a qualitative method that enables the tasks of extracting relevant information by counting the number of occurrences of specific keywords. The relevant information was extracted from the job description field. The sentences were split into individual keywords known as unigrams. The consecutive words form bigrams and trigrams based on the distance between neighbors. As part of the preprocessing routine, extraneous stopwords were removed from the analysis. The remaining set of keywords was matched against a keywords list for each skill presented in Table 1. In this way, we were able to determine which skills were required for each job position. Aggregating this knowledge across multiple job positions of the same type (AI or ML), we were able to determine which skills were more important.

Our process can be divided into the following steps:

1. *Data collection.* Web scraping was the primary method of data collection, using Python. We queried Indeed.com for titles, including “Artificial

Table 1. Classification framework.

Skill category	Skills	Keywords
Communication	Written	Copywriting, Editing, Blogging, Content Creation, Story-ideation
	Verbal	Verbal, Oral, Cold calling
	Presentation	Present, Presentation, Report
	Generic	Responsible, Determined, Competitive, Witty, Success-oriented
Employee Attributes	Motivation	Motivated, Ambition, Willingness to learn, Delivering result, Continuous learning
	Time management	Time management, Timely manner, Prioritize time, deadline-driven
	Detail-oriented	Attention to detail, Eye for detail, Accuracy, Precision
	Attitude	Can do, Go-getter, Self-learner, Self-directed, Positive attitude
	Independence	Independence, Without supervision, Autonomous
	Adaptability	Adaptable, Flexible, Multitasking
	Confidence	Confident, Decisive
	Other	Funny, Smiling, High-energy, Reliable, Proactive
Occupational Attributes	Enterprise system	ERP, CRM, SCM, SAP, PeopleSoft, Oracle, Integration, SAAS
	Visualization	Visualization, Tableau, Lumira, Crystal Reports, d3, d3.js
	Programming	Mathematical programming, Scala, Python, C#, C++, VB, Excel Macros, PERL, C, Java, Visual Basic, VB.NET, VBA, COBOL, FORTRAN, S, SPLUS, BASH, Javascript, ASP.NET, JQUERY, JBOSS
	Project management	Project management, PERT, CPM, PERT/CPM, change management, project budget, project documentation, PMP, Microsoft Project, Gantt Chart
	Modeling	Neural networks, linear programming, integer programming, goal programming, queuing, genetic algorithms, expert systems
	Scraping	Scraping, web scraping, crawling, web crawling
	Hardware	Hardware, architecture, devices, printer, storage, desktop, pc, server, workstation, mainframe, legacy, system architecture
	Networks	Internet, LAN, WAN, networking, cloud computing, client server, distributed computing, network security, ubiquitous computing, TCP/IP
	Statistics	Statistics, SPSS, SAS, Excel, Stata, MATLAB, probability, hypothesis testing, regression, pandas, scipy, sps, spotfire, scikits.learn, splunk, h2o, R, STATA, Statistical programming
	Data mining	Classification, text mining, web mining, stream mining, knowledge discovery, anomaly detection, associations, outlier, classify, association, estimation, prediction, forecasting, machine learning, decision trees
	Structured data	SQL, relational database, Oracle, SQL Server, DB2, relational DBMS, Microsoft Access, data model, data management, entity relationship, data warehouse, DBMS, transactional database, sql server, db2, Cassandra, mongo db, mysql, postgresql, oracle db
	Big data	Big data, Unstructured data, Data variety, Data velocity, Data volume, Hadoop, Hive, Pig, Spark, MapReduce, Presto, Mahoot, NoSQL, Spark, shark, oozie, zookeeper, flume
	Decision making	Reporting, analysis, modeling, design, problem-solving, implementation, testing, analytical, strategic thinking
	MS Office	MS Office, MS PowerPoint, presentation, MS Word, communication, documentation
Interpersonal	Interpersonal	Team management, Collaboration, Cooperation, Networking, Client relationship
Problem Solving	Problem solving	Problem solving, Troubleshoot, Conflict resolution, Solve issue, Critical thinker
	Creativity	Creative, Out of box, Storyteller
	Process design	Design Process, Improve process, Continuous improvement, Operations management
Administrative	Administrative	Issue management, Posting schedule, Product launch, social calendar
Analytical	Analytical	Insight, Identify trend, Summarize finding, Analyze trend, Synthesize information, Draw conclusion, Propose solution, Google Analytics, ArcGIS, GIS, QGIS, Data analytics, Business analytics
Research	Research	Data gathering, Data collection, Data reporting, Monitor trend, Monitor performance
Numeracy	Numeracy	Financial Management, Bookkeeping, Accountancy
Foreign Language	Foreign language	Spanish, French, German, Italian, Chinese

Intelligence" and "Machine Learning" through its open-source API, whose input arguments are search phrase, location and time period. We focused on jobs within the USA between July and December 2019. The API output includes the job URL and other fields such as job title, summary and location. The job summary field is a very abstract summary of the job description. Since we are interested in the

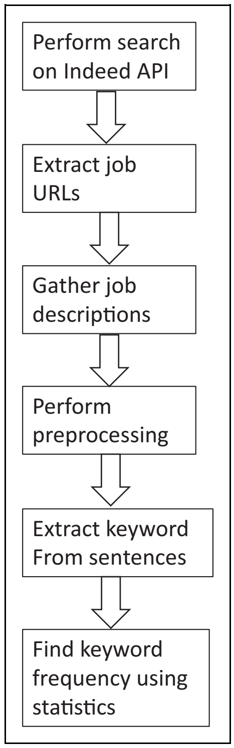

Figure 2. Summary of the data collection routine.

complete job description, we designed a utility in Python which downloaded these job URLs and stored them locally.

2. *Data parsing*. An advantage of relying on data from Indeed.com is that its web pages are standardized. Thus, the relevant information regarding job description can always be found under a specific HTML tag. This job description field contains a complete discussion of the requirements set by the employer for a specific job position. Thus, skill-based information could be located here. Our parser in Python extracted the correct tag corresponding to the job description field. Finally, we stored all job descriptions for each job position corresponding to AI or ML type.
3. *Data preprocessing.* The job description field of each job posting consists of multiple sentences which are further broken down into individual keywords. These keywords contain contextual information through unigrams, bigrams and trigrams. Note that unigram represents each word separately, while bigram represents two consecutive words in a specific sentence. For this purpose, each sentence is broken down into individual words. Thereafter, stopwords like a, an, the, and, in, for, etc. are removed from the analysis. The remaining individual words are identified as unigrams. Neighboring unigrams form bigrams and trigrams, depending on the distance between the neighbors. These n-grams were crucial to our analysis since the keywords belonging to each skill could be easily matched with the n-grams. If a match was found, we declared that a specific skill was required for a job advertisement.
4. *Data statistics*. In this step, aggregate statistics regarding the skill requirements of AI/ML job positions were calculated. The n-grams from the previous step were compared against the keywords for each skill described in Table 1. Thus, we established which skills were demanded for each job advertisement. If a match was found, we concluded that a given job advertisement was linked to a corresponding skill in a specific skill category. This process was repeated for all job advertisements corresponding to AI/ML job type. For a specific job title, we collected the number of jobs associated with each skill. After that, we determined the relative frequency of occurrence of each skill by dividing with the total number of job postings of each type. We converted the relative frequency to a percentage count and present it here in Tables 4 and 5. These data will be utilized in the next section to classify skills concerning their relative importance for different job types.

Table 2. Geographical distribution of AI and ML jobs.

Rank	State (percentage count of AI jobs)	State (percentage count of ML jobs)
1	CA (34.50%)	CA (33.29%)
2	WA (10.83%)	NY (12.24%)
3	MA (7.00%)	MA (7.88%)
4	NY (6.83%)	WA (5.59%)
5	TX (4.92%)	VA (5.18%)

Results

We queried Indeed.com for job titles containing the phrases "artificial intelligence" and "machine learning" and obtained 1,200 and 1,700 entry-level jobs, respectively, between July and December 2019. The sample size is large enough for content analysis studies (Gardiner et al., 2018; Verma et al., 2019). The geographical distribution in terms of the top five states for the respective jobs is shown in Table 2. As is evident from the results, most of these jobs are located in California, Washington, and New York. The IT industry has seen a dramatic rise in the demand for AI/ML skilled workers, especially in the metropolitan areas of San Francisco, Seattle, Boston and New York City.

Next, we provide an in-depth analysis of the skill requirements for the AI and ML job categories. The five most common skill categories required for each job

Table 3. Majors required for AI and ML jobs.

Rank	Major (percentage count of AI jobs)	Major (percentage count of ML jobs)
1	Machine learning (70.00%)	Machine learning (87.76%)
2	Engineering (68.83%)	Engineering (67.88%)
3	Business (61.25%)	Computer science (59.59%)
4	Computer science (59.17%)	Business (48.00%)
5	Artificial intelligence (47.00%)	Mathematics (42.65%)
6	Mathematics (37.92%)	Data science (28.82%)
7	Data science (28.75%)	Statistics (26.94%)

category are listed in Tables 4 and 5. These skill categories are sorted in descending order with respect to the number of job advertisements in which at least one skill associated with the skill category appears. For each skill category, we also report the top five skills. Again, these skills are ranked with respect to the percentage count associated with a specific skill relative to the total number of job postings for a specific job category.

To establish whether there is a significant difference between the two job types, we ran a hypothesis test. The percentage count of a specific skill with respect to AI and ML jobs facilitates a paired two-sample t-test. In this way, we compared the percentage relevance of each skill for AI against ML jobs. The null hypothesis was used to determine whether the mean difference between the two samples of observations is zero. We observed a p-value of 0.043, which helped us conclude that a minor difference exists between the AI and ML job categories with a significance level of 5%. This is explained by the fact that ML is usually considered as a subset of AI. Hence, many skill requirements will be the same for AI and ML jobs.

We also extracted the majors required for the job advertisements using regular expressions, which enables pattern matching in strings. For this purpose, we used the sentences in the job description of each job posting. We searched for keywords in the vicinity of major identifying information like BS, BA, Bachelor's, etc. These keywords were related majors like artificial intelligence, machine learning, data science, computer science, information technology, etc. The results are presented in Table 3.

Note that the percentage columns do not add up to 100% since one job posting could demand multiple majors, like a Bachelor's in data science or engineering. Traditional degree programs in engineering, mathematics, and computer science are still demanded for AI/ML jobs. However, the newer niches like data science, artificial intelligence, and machine learning have started to become more popular. We could observe that business degrees were also required, since many of these positions dealt with domain expertise. This is more pronounced for AI positions since ML positions tend to be slightly more technical.

Table 4. Results for AI positions.

Skill category	Skill	Percentage count (%)
Occupation		
	Decision making	87.42
	Data mining	70.50
	Programming	65.58
	Statistics	37.58
	Big data	32.75
Employee		
	Attitude	75.75
	Time management	64.25
	Motivation	44.33
	Independence	36.08
	Attention to detail	23.00
Communication		
	General	73.83
	Verbal	58.92
	Written	30.75
	Presentation	19.83
Interpersonal		
	Team	47.25
	Personal	13.17
	Networking	12.75
Problem Solving		
	General	32.50
	Creativity	29.17
	Process design	28.75

Next, we focus our attention on the similarities and differences between AI and ML positions. We compare the percentage count columns in Tables 4 and 5 and present an in-depth analysis of the required skills. The occupational skills are ranked at the top for both AI and ML positions. Decision-making skills are critical for both AI and ML positions. These include critical thinking skills while dealing with multiple data sources. The ability to synthesize information and draw conclusions to derive insights is essential in this data-driven digital economy. Knowledge of data mining techniques is more valued for ML positions: these consist of classical techniques like regression, classification, outlier detection and forecasting, with an increased focus on verticals like web mining, text mining and stream mining. Statistical and big data skills also hold more significance for ML positions. The classical constructs of probability and hypothesis testing still hold some value. However, this field is dominated by various niche software packages like Numpy, Scikit-learn, Pandas, etc. along with traditional offerings like R, SPSS, SAS, etc. The advent of big data in terms of volume, variety and velocity has necessitated new modeling paradigms. It has also generated an interest in software like Hadoop, Mapreduce, Nosql, etc.

The focus on hard skills like programming, data modeling, statistics, data mining and structured data is more prominent for the ML positions. Moreover, for the ML job type, more job postings require advanced knowledge of

Table 5. Results for ML positions.

Skill category	Skill	Percentage count (%)
Occupation		
	Data mining	98.00
	Decision making	86.88
	Programming	82.53
	Statistics	58.76
	Big data	39.41
Employee		
	Attitude	98.00
	Time management	52.71
	Motivation	41.71
	Independence	35.24
	Attention to detail	18.47
Communication		
	General	61.24
	Verbal	40.35
	Written	33.47
	Presentation	17.76
Interpersonal		
	Team	44.82
	Personal	10.65
	Networking	5.24
Problem solving		
	Process design	24.88
	General	23.59
	Creativity	17.47

neural networks constructs like Convolution Neural Network, Deep Neural Network, Recurrent Neural Network (modeling skills). The programming skills in demand include Python and R in addition to the traditional languages like C, C++ and Java. These jobs also demand prior experience in design frameworks like Pytorch, Theano, Tensorflow, etc. Thus, the ML positions appear to be more technical. On the other hand, an overall systems design perspective through knowledge of enterprise systems like Oracle, SAP, etc. is more highlighted for AI positions. Hence, AI job positions are more generic in the sense that holistic process understanding is critical for success.

Employee attributes have become increasingly crucial in the workplace. Attitude is the most sought-after skill for both these positions. A can-do or go-getter attitude with a strong desire to learn is vital for success-oriented and ambitious professionals. Time management skills are especially important for millennials. Employees with a disposition toward setting and achieving priorities in a timely manner are valued in a fast-paced and dynamic business environment. The next key soft skills of motivation and independence are interrelated. Independent workers are self-starters who complete tasks with minimal supervision. They are also proactive and take complete ownership of their projects. This requires them to be highly motivated, enthusiastic and passionate. These traits are needed for success in any job. The next skill, attention to detail, is considered as a sign of an organized and attentive individual who can be an asset to any company. These findings, in terms of employee attributes, should be incorporated into the academic discourse, and universities should adapt their curricula accordingly.

The next important skill category is communication. The most common generic subtype includes keywords for responsible and determined employees who are ready to take the initiative and can deal with stress in a dynamic fast-paced environment. Note that this skill category is significantly more valued for AI positions – because most of the AI job requirements are generic and involve processing project requirements for systems-design framework development. This requires gathering various inputs from multiple team members and collating the results of the analysis. Hence verbal and written skills are more sought-after for AI positions than for ML positions, which are somewhat technical. Presentation skills are also more vital for AI professionals since the project design phase typically involves meetings and brainstorming sessions with upper management.

Organization skills, such as interpersonal skills, are also critical for both job categories. They involve collaboration and cooperation in a team setting. Traits like leadership and networking with external clients and internal stakeholders are highly valued. The problem-solving skill category emphasizes individuals who are rational thinkers and can use sound judgment to troubleshoot and solve business issues – these are very useful in designing a business strategy for the long-term growth of a firm. The creativity or innovation skillset acknowledges the entrepreneurial spirit of the employee. Moreover, new out-of-the-box ideas promote an employee's intellectual development and generate new revenue channels for the organization. Lastly, the process design subtype of problem-solving signifies the relevance of the integrated processes at the firm level. Employees are expected to design and improve processes with the objective of minimizing costs, time and errors. Thus, knowledge of product design, operations management, supply chains and marketing channels has become an important skill to acquire in the Industry 4.0. era.

Discussion

Our research identified 35 skills organized into 10 different skill categories for AI and ML professions (presented in Table 1). While the literature has provided some insights related to AI (Anton et al., 2020; Balan, 2019; Bawack et al., 2019), our comprehensive analysis is centered on empirical data. More specifically, we explore the employer's perspective through an analysis of the job advertisements.

Most of the recent interest in AI and ML technologies has emanated from the information technology sector. Hence, the jobs are typically found in the metropolitan hubs

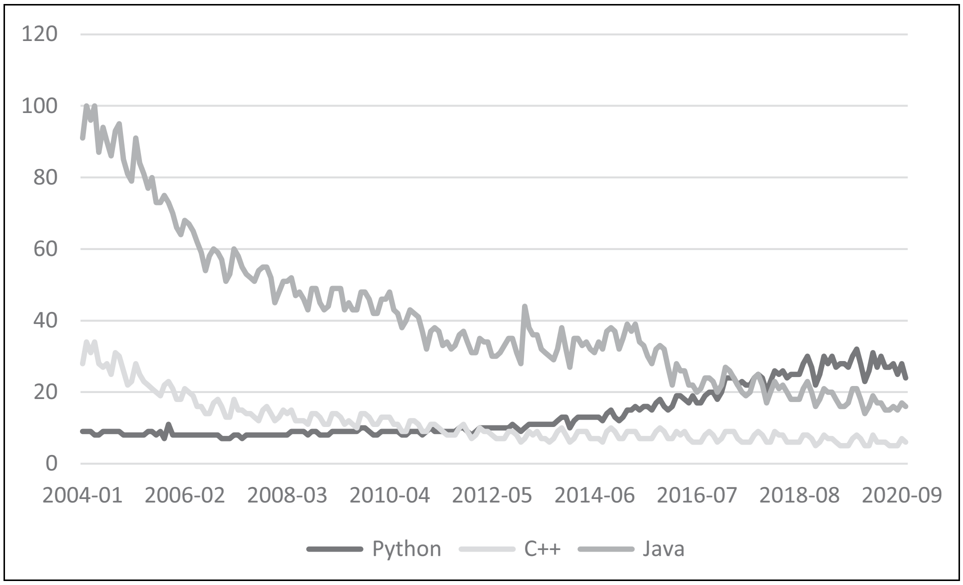

Figure 3. Google search volume in the US for the search queries of the different programming languages (Source: Google Trends).

in New York, Washington and California. From Table 3, it is clear that the business-oriented majors are more in demand for AI professions. The business-related competencies provide domain-level expertise in fields like healthcare and digital marketing. Moreover, we observe that statistics majors are more valued for the ML professions – this emphasizes that the required ML job skills are somewhat more technical, with a special focus on data-driven applications in banking, finance and supply chain management.

We identified a number of similarities between AI and ML job postings. The business skills and generic IT concepts are common to both. It is well established that all strategic decision-makers could benefit from a sound understanding of their business to increase the effectiveness of AI solutions (Brynjolfsson and Mitchell, 2017; Mohanty and Vyas, 2018). Technical expertise in data analysis, visualization and reporting skills are also highly valued. Tasks like classification and prediction are central to the areas of supervised and unsupervised learning. These skills are highly sought after by employers in addition to outlier detection and forecasting techniques. The classical field of data mining concerns itself with pattern extraction on structured data. Due to the abundance of data in terms of volume, variety and speed, the newer domains of web mining and stream mining have started to become more popular. They require expertise in unstructured database management as well as some specific software programs. These jobs increasingly require decision-making by extracting information from web-based and unstructured data sources related to text mining, web mining and social networks. We have seen from the results that these big data skills have started to gain prominence: this is evident from the focus on big data tools like Hadoop, Mapreduce, Hive, etc. Moreover, skills in traditional relational databases like SQL and Oracle have declined in popularity.

We also note that there are some clear differences between AI skills and ML skills. The ML competency requirements place more emphasis on computer programming skills. The newer programming languages like Python are slowly catching up with traditional languages like C++ and Java. These findings match the insight provided by the Google Trends data with respect to the volume of search queries (cf. Figure 3). Almost 60 percent of the ML jobs we analyzed asked for strong software programming skills and statistical knowledge, whereas AI jobs required much less "programming" and statistical knowledge. In addition, ML jobs demand data analysis and machine learning skills like outlier detection, regression, classification, etc. In comparison, there is less mention of such terms in AI job advertisements.

We also want to mention that the current job listings demand more expertise in "open-source" tools. This reduces the barriers of entry for these professions. Even though large vendors like SAS, SPSS, SAP, etc. provide various analytical solutions in these domains, the job market does not place more emphasis on them.

Conclusions and future research

In this research, we studied job descriptions for AI/ML positions in the USA. We utilized content analysis to develop a ranking of key skill categories required for each job type. Moreover, we performed a pair-wise comparison of the top five skill categories for AI and ML. We found that occupational skills were the most sought-after for the two job positions. ML positions are somewhat more technical and place more emphasis on data mining, programming, statistics and big data skills. On the other hand, AI positions assign more importance to communication skills, signaling that they require interaction with other team members and upper management while working on a proof-of-concept for systems design projects.

Our current study identifies and ranks skills that are currently required for AI/ML positions. Our findings could be used in two ways. First, university programs in management information systems, data science, data analytics, and business analytics could revamp their existing offerings related to AI/ML. Thus, our study could be beneficial for designing associated undergraduate and graduate degree programs. Second, the findings could be used by hiring managers looking for skilled AI/ML workers. By using the proposed skills, the hiring team could create well-structured and consistent job descriptions. In this way, the hiring team could utilize their employees efficiently.

Future research could involve a localized study of some specific US states. This would benefit local employers as well as local universities. Another research direction would be to investigate job descriptions from other professional fields – like accounting, finance, marketing and healthcare. In addition, scholars could analyze the existing curricula of degree programs in the AI/ML domain. In this way, the findings of the present study could be used in parallel with the coursework of degree programs in related fields. By identifying the skill gaps in the existing programs, we could make more concrete recommendations regarding curriculum design.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Amit Verma https://orcid.org/0000-0003-0187-5514

References

Agrawal A, Gans JS and Goldfarb A (2019) Artificial intelligence: the ambiguous labor market impact of automating prediction. *Journal of Economic Perspectives* 33(2): 31–50.

Aiken RM (1991) The new hurrah: creating a fundamental role for artificial intelligence in the computing science curriculum. *Education and Computing* 7(1–2): 119–124.

Anton E, Behne A and Teuteberg F (2020) The humans behind artificial intelligence – an operationalization of AI competencies. In: *Proceeding of the 28th European conference on information systems (ECIS)*, Marrakeh, Morocco, 15–17 June.

Balan C (2019) Potential influence of artificial intelligence on the managerial skills of supply chain executives. *Calitatea* 20(S3): 17–24.

Baldwin-Morgan AA (1995) Integrating artificial intelligence into the accounting curriculum. *Accounting Education* 4(3): 217–229.

Bawack RE, Wamba SF and Carillo K (2019) Artificial intelligence in practice: implications for IS research. In: *Twenty-fifth Americas conference on information systems*, Cancun, Mexico, 15–17 August, p. 11.

Brynjolfsson E and Mitchell T (2017) What can machine learning do? Workforce implications. *Science* 6370(358): 8.

Chassignol M, Khoroshavin A, Klimova A, et al. (2018) Artificial intelligence trends in education: a narrative overview. *Procedia Computer Science* 136: 16–24.

Chedrawi C and Howayeck P (2019) Artificial intelligence a disruptive innovation in higher education accreditation programs: expert systems and AACSB. In: Harfouche A and Baghdadi Y (eds) *ICT for a Better Life and a Better World*. Cham: Springer, pp. 115–129.

Debortoli S, Müller O and vom Brocke J (2014) Comparing business intelligence and big data skills. *Business & Information Systems Engineering* 6(5): 289–300.

Fareri S, Fantoni G, Chiarello F, et al. (2020) Estimating Industry 4.0 impact on job profiles and skills using text mining. *Computers in Industry* 118: 103222.

Gardiner A, Aasheim C, Rutner P, et al. (2018) Skill requirements in big data: a content analysis of job advertisements. *Journal of Computer Information Systems* 58(4): 374–384.

Goksel N and Bozkurt A (2019) Artificial intelligence in education: current insights and future perspectives. In: Sisman-Ugur S and Eby G (eds) *Handbook of Research on Learning in the Age of Transhumanism*. Pennsylvania, PA: IGI Global, pp. 224–236.

Holder LB and Cook DJ (2001) A client-server computational tool for integrated artificial intelligence curriculum. *Journal of Computing in Higher Education* 12(2): 52–69.

Jakhar D and Kaur I (2020) Artificial intelligence, machine learning and deep learning: definitions and differences. *Clinical and Experimental Dermatology* 45(1): 131–132.

Lovaglio PG, Cesarini M, Mercorio F, et al. (2018) Skills in demand for ICT and statistical occupations: evidence from web-based job vacancies. *Statistical Analysis and Data Mining: The ASA Data Science Journal* 11(2): 78–91.

Lu Y (2019) Artificial intelligence: a survey on evolution, models, applications and future trends. *Journal of Management Analytics* 6(1): 1–29.

McArthur D, Lewis M and Bishary M (2005) The roles of artificial intelligence in education: current progress and future prospects. *Journal of Educational Technology* 1(4): 42–80.

Meulen R and Petty C (2017, December 13) Gartner says by 2020, artificial intelligence will create more jobs than it eliminates. Available at: https://www.gartner.com/en/newsroom/press-releases/2017-12-13-gartner-says-by-2020-artificial-intelligence-will-create-more-jobs-than-it-eliminates (accessed 15 May 2020).

Mitić V (2019) Benefits of artificial intelligence and machine learning in marketing. In: *Sinteza 2019 – International scientific conference on information technology and data related research*, Siberia, 20 April 2019, pp. 472–477. Singidunum University.

Mohanty S and Vyas S (eds) (2018) Board to CEO: "What's your AI strategy?" In: *How to Compete in the Age of Artificial Intelligence*. Berkeley, CA: Apress, pp. 75–90.

Neuendorf KA (2016) *The Content Analysis Guidebook*. Thousand Oaks, CA: Sage Publications.

Pejic-Bach M, Bertoncel T, Meško M, et al. (2020) Text mining of Industry 4.0 job advertisements. *International Journal of Information Management* 50: 416–431.

Roll I and Wylie R (2016) Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education* 26(2): 582–599.

Romeo J (2020, January 20) AI and machine learning jobs gain growing popularity. Available at: https://hrexecutive.com/in-search-of-artificial-intelligence/ (accessed 1 May 2020).

Schlanberger L (1989) A strategy for integrating artificial intelligence technology into a graduate business curriculum. *Education and Computing* 5(3): 189–196.

Shar TH (2019) *Students' perspectives on artificial intelligence in education*. Doctoral Dissertation, Waikato Institute of Technology, New Zealand.

Testa P, O'Reilly UM, Weiser D, et al. (2001) Emergent design: a crosscutting research program and design curriculum integrating architecture and artificial intelligence. *Environment and Planning B: Planning and Design* 28(4): 481–498.

Verma A, Yurov KM, Lane PL, et al. (2019) An investigation of skill requirements for business and data analytics positions: a content analysis of job advertisements. *Journal of Education for Business* 94(4): 243–250.

Vochozka M, Kliestik T, Kliestikova J, et al. (2018) Participating in a highly automated society: How artificial intelligence disrupts the job market. *Economics, Management, and Financial Markets* 13(4): 57–62.

Wamba-Taguimdje SL, Wamba SF, Kamdjoug JRK, et al. (2020) Influence of artificial intelligence (AI) on firm performance: the business value of AI-based transformation projects. *Business Process Management Journal* 26(7): 1893–1924.

Wang W and Siau K (2019) Artificial intelligence, machine learning, automation, robotics, future of work and future of humanity: a review and research agenda. *Journal of Database Management (JDM)* 30(1): 61–79.

Wilson HJ, Daugherty P and Bianzino N (2017) The jobs that artificial intelligence will create. *MIT Sloan Management Review* 58(4): 14.

Woschank M, Rauch E and Zsifkovits H (2020) A review of further directions for artificial intelligence, machine learning, and deep learning in smart logistics. *Sustainability* 12(9): 3760.