

Article

A Methodological Framework to Predict Future Market Needs for Sustainable Skills Management Using AI and Big Data Technologies

Naif Radi Aljohani ^{1,*}, Muhammad Ahtisham Aslam ¹, Alaa O. Khadidos ¹ and Saeed-Ul Hassan ²

¹ Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia; maaslam@kau.edu.sa (M.A.A.); aokhadidos@kau.edu.sa (A.O.K.)

² Department of Computing and Mathematics, Manchester Metropolitan University, Manchester M15 6BH, UK; s.ul-hassan@mmu.ac.uk

* Correspondence: nraljohani@kau.edu.sa

Abstract: Analysing big data job posts in Saudi cyberspace to describe the future market need for sustainable skills, this study used the power of artificial intelligence, deep learning, and big data technologies. The study targeted three main stakeholders: students, universities, and job providers. It provides analytical insights to improve student satisfaction, retention, and employability, investigating recent trends in the essential skills pinpointed as enhancing the social effect of learning, and identifying and developing the competencies and talents required for the Kingdom of Saudi Arabia's (KSA's) digital transformation into a regional and global leader in technology-driven innovation. The methodological framework comprises smart data processing, word embedding, and case-based reasoning to identify the skills required for job positions. The study's outcomes may promote the alignment of KSA's business and industry to academia, highlighting where to build competencies and skills. They may facilitate the parameterisation of the learning process, boost universities' ability to promote learning efficiency, and foster the labour market's sustainable evolution towards technology-driven innovation. We believe that this study is crucial to Vision 2030's realisation through a long-term, inclusive approach to KSA's transformation of knowledge and research into new employment, innovation, and capacity.

Keywords: job big data; deep learning; sustainable education; smart decision-making; student satisfaction; employability



Citation: Aljohani, N.R.; Aslam, M.A.; Khadidos, A.O.; Hassan, S.-U. A Methodological Framework to Predict Future Market Needs for Sustainable Skills Management Using AI and Big Data Technologies. *Appl. Sci.* **2022**, *12*, 6898. <https://doi.org/10.3390/app12146898>

Academic Editors: Boris Aberšek and Andrej Flöglie

Received: 13 June 2022

Accepted: 5 July 2022

Published: 7 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

One of the goals of the Vision 2030 programme is to enhance educational sustainability by developing a comprehensive framework for high-quality research, flexible learning, and identification of viable routes for young Saudi professionals [1,2]. As well as their innovation and research potential, Saudi institutions' development of students' skills and capacities [3] has improved dramatically [4], suggesting a shift to digital transformation and sustainable innovation in the Kingdom of Saudi Arabia (KSA). Another shift observed in Saudi institutions has been the high-quality work published through stronger collaboration with the international and regional scientific community [5,6].

Scientific research is considered a reliable indicator of any country's development [7]. In KSA, it is linked to remarkable improvements in technology, economics, and, ultimately, the quality of human life [8]. The world has witnessed significant progress in various fields of science (mathematics, computer science, engineering, etc.), and most countries fully understand that investment in scientific research is vital for the emergence of science and technology, ultimately reflected in a better standard of living [9].

In the past decade, KSA has encouraged scientific research and development in sectors such as technology, commerce, science, agriculture, and industry [10]. However, researchers have noted an absence of strategic research, clear-cut policies, and sufficient research

planning for the capacity-building that will uplift Saudi youth's skills to meet the needs of the job market [11,12]. This study argues that, to align the skills of the young Saudi population to the relevant job market, it is highly desirable to analyse the gap between the market and university curricula [13]. Such analysis can enhance students' awareness of study planning tools, help to build their career pathway [14,15], and boost their engagement and retention [16]. To the best of our knowledge, so far, no study has presented a data-driven analysis between "required skills" in the Saudi job market and university students' "acquired skills" from their training and education. Therefore, we seek to fill this space with the following contributions and guidelines for policymakers:

1. First, we make use of the data about big data jobs available on the web and about jobs in the Saudi information and technology market;
2. Second, we explore the latest trends in required skills that promote the social impact of learning by identifying the required competencies and talents for KSA's digital transformation into a regional and global leader in technology-driven innovation;
3. Third, to facilitate data integration from multiple resources, we investigate the significance of deep learning-supported word embedding methods to map the "required" skills against "acquired" skills;
4. Fourth, to inform individuals' decisions on courses, we employ case-based reasoning to identify the gap between the market's "required" and their own "acquired" skills;
5. Finally, to advance the strategic goals of Vision 2030 through developing KSA's long-term capacities, we aim to deliver innovative research in the domains of data and AI.

This study contributes directly to four integrative dimensions and beyond. First, in the corpus of knowledge of AI, ML, and learning analytics [17], through scientific methodology we employ algorithms and techniques to match KSA's labour market needs to the sophisticated job, skill, and competency profiles associated with Vision 2030. Second, in the Saudi AI market we contribute a cost-effective, efficient, and high-performing approach to developing a platform for the skills and competencies essential to long-term innovation and digital transformation. Third, we contribute to the Saudi innovation ecosystem a realistic, progressive service capable of sustained commercialisation. Finally, this study contributes by linking cutting-edge research and development capabilities to future market and labour demands, orchestrating a scientific, close-to-market approach to merging Saudi academia and Saudi industry.

The rest of the paper is organised as follows: the next section presents a literature review followed by the details of the machine learning and deep learning approaches. Further, we present a case study using data from the Saudi information technology job market. Finally, we present the concluding remarks, along with future research directions.

2. Literature Review

Six key issues, identified as the theoretical background to the Saudi Vision 2030 programme, were used to organise this section.

Meeting the expectations of students at higher education institutions: Higher education is an important source of skilled labour in today's modern, post-industrial economy, ensuring future prosperity. Students consider employability as the capacity to obtain and maintain future employment; however, the smooth transfer to further education and the labour market is a lengthy and complex process. By examining the demands of the target employment markets and adapting programmes to match those needs, universities must put greater emphasis on meeting students' expectations. Semantic-based education [18] and analytical applications [19] are critical to academic courses' knowledge representation, modelling the learning outcomes and categorising the professions and the necessary abilities. All may be used for analysis and prediction [20].

Enhancing education completion rates: According to Burke [21], one in every three students fails to complete their course within six years of enrolment. Non-completion of a course has financial ramifications not just for the student, the student's family, the institution, and the government, but also for the student's personal and professional life.

Furthermore, it has an impact on society and the economy through the loss of critical skills and information [22]. There are also financial and reputational implications for institutions, making student retention and success a major concern for universities worldwide [23].

Promoting student satisfaction: Many theoretical models have been presented to describe the fundamental notion of student happiness and retention in higher education. Hsu et al. [24] investigated alumni happiness and offer a deconstructed alumni satisfaction model, with strategic management maps to identify areas for continuous development. They determined that course design, which should be relevant to real-world career possibilities, has a critical influence on student pleasure. Mello and Wattret [25] emphasised the necessity to incorporate academic and employability skills into the curriculum and present a skillset approach to build employability into the design of a university degree.

Mapping and visualising educational data for decision-making: Thanks to concept mapping and graphic depiction of various ideas and course subjects (CLO), students can easily comprehend a course's learning outcomes. Similarly, skill representation and mapping of course topics can be beneficial both to student satisfaction and learning strategy design. Boyle et al. [26] described the mapping concept in education and semantic-based knowledge [26]. The proposed work has a theoretical model for semantic-based curricular planning and measurement of students' comprehension of ideas and themes. Students may use its outcomes to plan their professional futures better; hence, it may improve student retention and employability.

Using machine learning for classifying and linking learning and jobs: In addition to quickly altering educational learning patterns and objectives, the area of information technology (IT) has enabled online access to a huge volume of job-related data. Classifying and locating relevant information in such a mass of data is a difficult endeavour [27]. In a supervised classification learning algorithm/method, a computer program both trains and learns from an available dataset then applies what it has learned to categorise fresh datasets. Applying ontological categorisation and reasoning to domain data can improve outcomes by adding a further layer to supervised learning [28]. It can also help to classify job adverts, including their required and skill criteria, helping to find any gap between a student's skills and the job's requirements [29].

Integrating academia and job the market: Due to the changing nature of work settings and the market's diverse requirements as a result of the introduction of new technologies, it has become increasingly difficult to integrate university curricula and the job market. At the outset of their job search, graduates are typically hesitant, wasting time and resources in looking at jobs that are not a good fit.

Employers want graduates with subject-specific skills and expertise [30]. Small and medium-sized businesses employ the majority of graduates, and these firms typically provide less training and development than bigger, more conventional companies [31]. Rather than invest money in developing staff's employability skills, these companies hire graduates who already possess the ability to function properly. Therefore, an employer-driven strategy is necessary to bridge the gap between academic courses and the labour market [32]. Moreover, if students have a clear understanding of their academic path from the very first year, through understanding the connection between their studies and their professional ambitions, they make more sensible decisions.

Summary: The literature review provides a theoretical background to the study's strong connection to Vision 2030, which states the following. The qualities and capabilities of our children are among our most valued and beloved belongings. We try to establish a culture that fosters determination, provides chances for all, and aids everyone in learning the skills they need to achieve their own goals to help individuals attain their most significant potential. To that end, we are strengthening the economy's capacity to provide diverse job opportunities and establishing a new paradigm in attracting global talents and credentials to the Kingdom of Saudi Arabia.

3. Methodological Framework

This section discusses the methodology of the study's technique, focusing on its analysis of the future Saudi market's need for sustainable skills management. The study aimed to enhance universities' student satisfaction and retention by improving students' understanding of the skills necessary for their future career options. It adopted a semantic-based approach to developing an effective conceptual and analytical model to support students' study through skills planning and to motivate them through career planning, thus improving their satisfaction and future employability. As shown in Figure 1, our methodology comprises three modules: a data layer, a word-embedding layer, and a mapping layer.

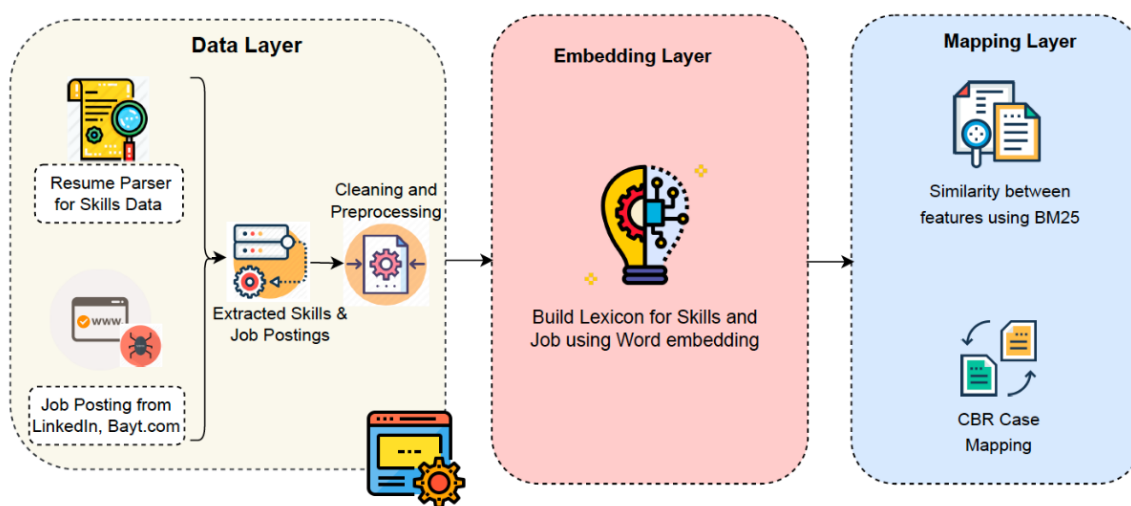


Figure 1. Methodological framework: data layer, word embedding layer, and mapping layer.

3.1. Data Layer

Gathering the data on careers and skills (in any possible available format) was crucial to this phase. The data layer emphasises the various data entities and their interaction through career and skill components such as opportunity awareness, self-awareness, and transition learning. By mapping job-specific skills, the concept of choice learning is covered by the skills component. A job description specifies the abilities required for a position. It is widely known that students are more likely to complete their studies if they have a real desire to achieve and can see the personal or career benefits from acquiring the skill set of their chosen course of study [33–37].

In order to extract jobs-related data, we explored various job portals such as LinkedIn (<https://www.linkedin.com>, accessed on 30 March 2022) and Bayt.com (<https://www.bayt.com>, accessed on 30 March 2022). Bayt.com is the leading job portal connecting job searchers with potential employers in the Middle East region. Thousands of job openings are added daily to this award-winning platform by top firms. Job postings were first gathered from these well-known job websites using the free web scraping services provided by Gresper's (<https://www.grespr.com>, accessed on 30 March 2022).

From users' resumes, we collected candidates' data, and advanced resume parsing techniques were used to extract certain information. Our resume parser combines AI neural networks and data science techniques to extract structured data. Unlike a standard data extraction parser, it is designed to recognise information from CVs. In order to construct a thorough user profile automatically, specific fields were extracted, such as candidates' personal information, technical skills, programming languages, frameworks, certifications, and soft skills. The machine learning algorithm was trained using the education, skills, and work experience subsections of the resumes. Rather than just Natural Language Processing (NLP) and data extraction, it uses deep transfer learning on recent open-source models to segment, identify, and extract the relevant information. We employed image-based

object detection and machine learning algorithms to segment and analyse the resume, determining both the proper reading sequence and the specific sections.

Next, we input the structural information to the sequence tagger to perform Named Entity Recognition (NER) and find the keywords. It is important to note that a separate neural network handles each section and its keywords. We performed data cleaning steps to remove phone numbers, locations, and additional HTML tags that came along with the job data during the parsing were also removed. Additionally, we employed text-matching techniques (see Section 3.3.2) to perform contextual mapping between the skills. To ensure optimal performance our models were trained on our database of thousands of English language resumes. We built a comprehensive dictionary for all five aspects: technical skills, programming languages, frameworks, certification, and soft skills. Once the data about skills and jobs were gathered and were cleansed using the R-Tool, which is effective because it uses both stop-word and free-word techniques.

3.2. Embedding Layer

In this module, for our primary source we used the list of available jobs and skills data, then performed data cleansing. The data layer extracted and analysed job advertisement data from multiple sources, so we first performed data pre-processing, filtering the raw data. These data underwent additional processing (cleaning and deriving indicators and measures). With the help of a domain expert, the skills were then extracted to create a skills map for the IT domain. Each job title might be satisfied by many job adverts; in such situations, reasoning helped to rank them based on the labelled data's weight and frequency, according to the job classification.

3.2.1. Word2Vec

We investigated several weighting schemes, including latent semantic indexing and Word2Vec models such as FastText, to increase the words in jobs and skill, create lexicons with more terms, and extend the terms to be matched. Word embedding is based on the idea that every word in a language can be represented by a set of real numbers (a vector). Word embeddings are N-dimensional vectors whose values attempt to represent words' meaning and context. A legitimate word vector can be any set of integers yet, to be effective, a set for vocabulary should capture organically the meaning of the words, their relationships, and the context in which they are used.

3.2.2. FastText

FastText is a conceptual framework (see Figure 2) in which word embeddings are regarded as the sum of their character-level embeddings rather than just the word itself, and may thus be applied to a variety of text problems. FastText is a Facebook-developed open-source framework that allows experts to learn text representations and classifiers for efficient text classification. This strategy is particularly adapted to distributing power through words made up of common roots encountered in training. Two terms, such as 'counter' and 'revolution', can be combined to explain the less-common word 'counterrevolution', which is otherwise too unusual to be learnt from a dictionary. Character-level embedding also supports learning to misspell jargon, such as 'pls', 'bcz', and so on, which may help to exchange the linkages between languages with a shared origin.

FastText creates a sampling table to eliminate frequent words. The theoretical background behind this work is that frequently repeated words contain less information than uncommon words, and that their representation will not alter significantly after multiple instances of the exact phrase. Equation (1) is the mathematical representation of calculating the probability of a word using a frequency word:

$$P(w) = \sqrt{\frac{t}{f(w)}} + \frac{t}{f(w)}, \quad f(w) = \frac{\text{count}_w}{\text{Total no of Tokens}} \quad (1)$$

where $t = 10 \times 10^{-4}$ and $f(w)$ is the frequency of occurrence of word w . The option is used to change the default threshold manually. The threshold value t has a meaning in fastText different from that in the original word2vec, and should be adjusted accordingly. A word is eliminated if the chance of discard is more significant than a random selection from a uniform distribution between 0 and 1 during the training stage. This mechanism is only applied to unsupervised models; in a supervised model, the words are not discarded.

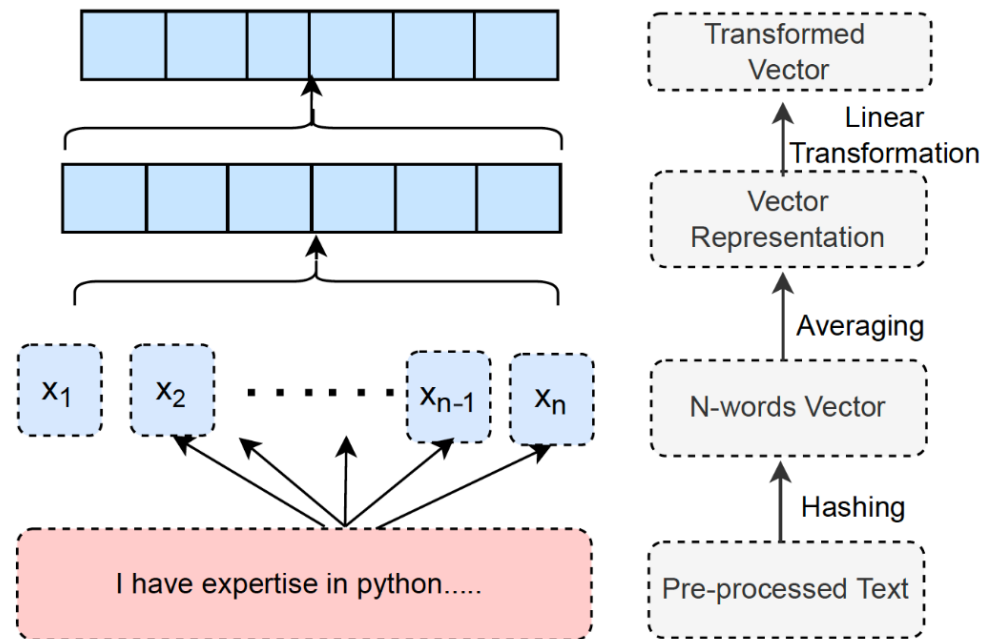


Figure 2. Architecture of fastText embedding.

3.3. Mapping Layer

In order to map jobs to skills, we used case-based reasoning (CBR). After analysing a sample of the weighted scores generated by several systems, the BM25 technique was found to be the most effective at matching skill words from a previous case to skills from the present case. In the following section, we discuss the details of the CBR and BM25 matching method.

3.3.1. Case-Based Reasoning

CBR is any problem-solving strategy that employs previous solutions to a similar problem. It assumes that information can be gained from previous experience, that this helps to avoid paths leading to failure, and that an established solution can be tailored to the current challenge. CBR can be found in many contexts. Google Maps, for example, employs it to estimate the duration of your journey from point A to point B by looking at the patterns of previous users to determine how long it took them. It creates conclusions about how long your journey will take, even if it starts from two slightly different sites. By tackling memory, learning, planning, and problem solving, CBR lays the framework for a new generation of intelligent computer systems that can solve challenges and adapt to new situations. The “intelligent” utilisation of knowledge from previously solved problems (or cases) in CBR is based on the assumption that the more similar the two issues are, the more similar will be the solution. In general, the CBR procedure involves the following steps:

1. *Retrieve*—retrieve from memory the event most like the current situation;
2. *Reuse*—propose a solution based on prior experience, modifying it to the new situation’s requirements;
3. *Revise*—re-evaluate the solution’s relevance and usefulness in the new situation;
4. *Retain*—keep this new approach to problem solving in the memory system.

CBR calculates a similarity score, using weighted averages to measure the likeness between the various cases to map them against the same solution. Equation (2) shows the formula to calculate the weighted average between cases:

$$\text{Similarity (new case, old case)} = \frac{1}{W} \sum_{i,j=1}^n \text{Similarity Score } (f_i, f_j) \times W_i \quad (2)$$

where $W = W_1 + W_2 + W_3 + \dots + W_n$.

3.3.2. Okapi BM25

We employed Okapi BM25 [38], the state-of-the-art matching formula, to calculate the similarity between the various features in different cases: Similarity score (f_i, f_j) . After analysing a sample of the weighted scores generated by several systems, the BM25 technique was found to be the most effective at matching skill words from the previous case to skills in the present case. It is important to note that in CBR every new problem is mapped to an existing case. Therefore, we used BM25 to match the existing case to a new problem. Equation (3) shows the formula for BM25:

$$\text{BM25}(Q, L) = \sum_{t \in q} \left\{ \text{IDF}(q_i) \times \frac{(k_1 + 1) f(q_i, D)}{k_1 \left((1 - b) + b \times \left(\frac{l_i}{l_{av}} \right) \right) + f(q_i, D)} \right\} \quad (3)$$

where Q represents the existing case and L the new case, and k_1, k_3 , and b are the constants.

4. Case Study: Analysing the Future Saudi Market's Needs for Sustainable Skills Management

In this section, we present a Saudi case study on the job market in emerging technological sectors by compiling data from LinkedIn.com and Bayt.com job marketplaces specific to the sectors of data and AI. We included over 400 jobs active on 30 March 2022 for big data engineers, developers and engineers, and database systems. In addition, we collected the CVs of over fifty students at a public sector Saudi university for use as data on the education and training that they had received.

The method described above helped to identify Saudi job market trends in the selected area and the relevant skills required across job sectors. It matched job sectors to individuals through semantic mapping techniques, recommended the jobs and skills required across the selected job markets, and finally displayed the results using Power BI dashboards with drilldown and roll-up functionality to present various analytical insights.

In the following subsections, we first present the Saudi job market dashboard, then three further dashboards to recommend jobs and relevant skills for individual profiles in the areas of data science, NoSQL, and software engineering.

4.1. Job Market Dashboard

Figure 3a shows a Saudi job market dashboard consisting of over 400 jobs active on 30 March 2022 across the areas of big data engineers, developers and engineers, and database systems. Of these jobs, over 55% are associated with big data engineering, followed by developers and engineers at about 30%, and database systems at 11%. Interestingly, about 95% of all these jobs were either entry-level or mid-level, showing that there is an emerging Saudi technical job market in the areas of data and AI. We also found that over 50% of industries are associated with information technology and software technology.

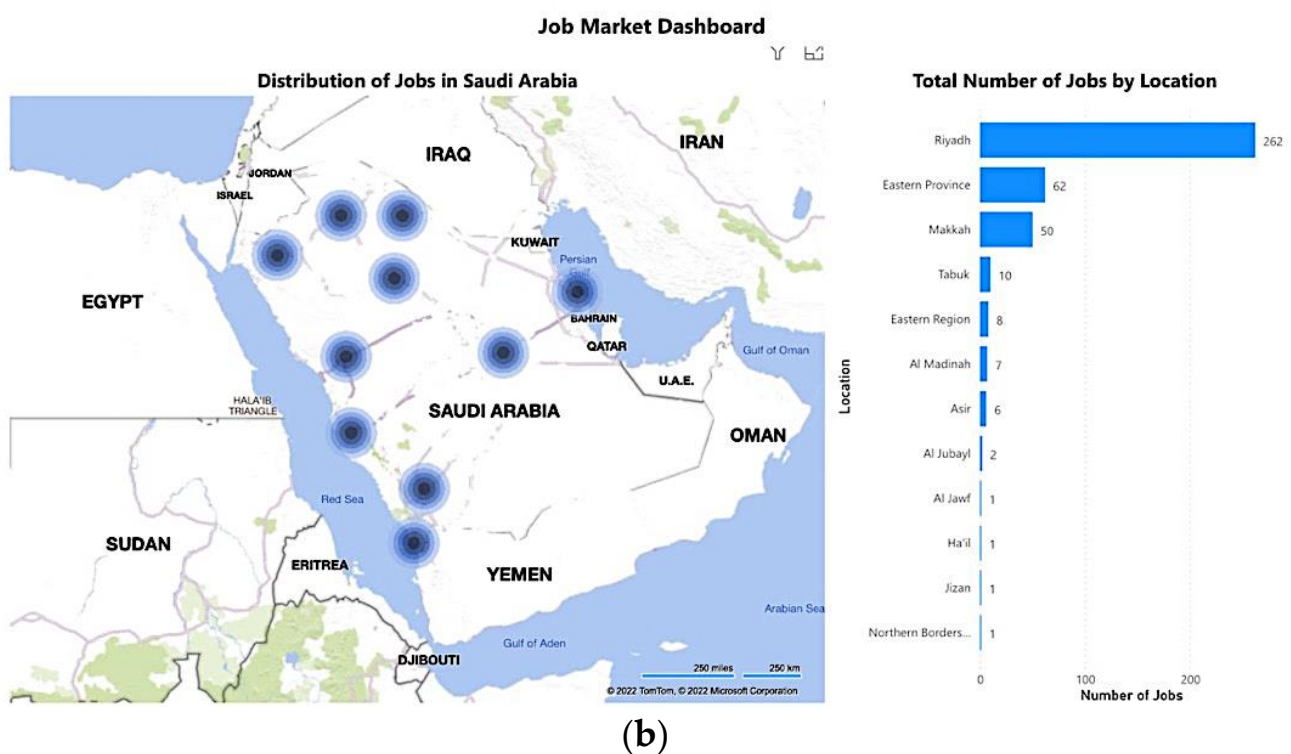
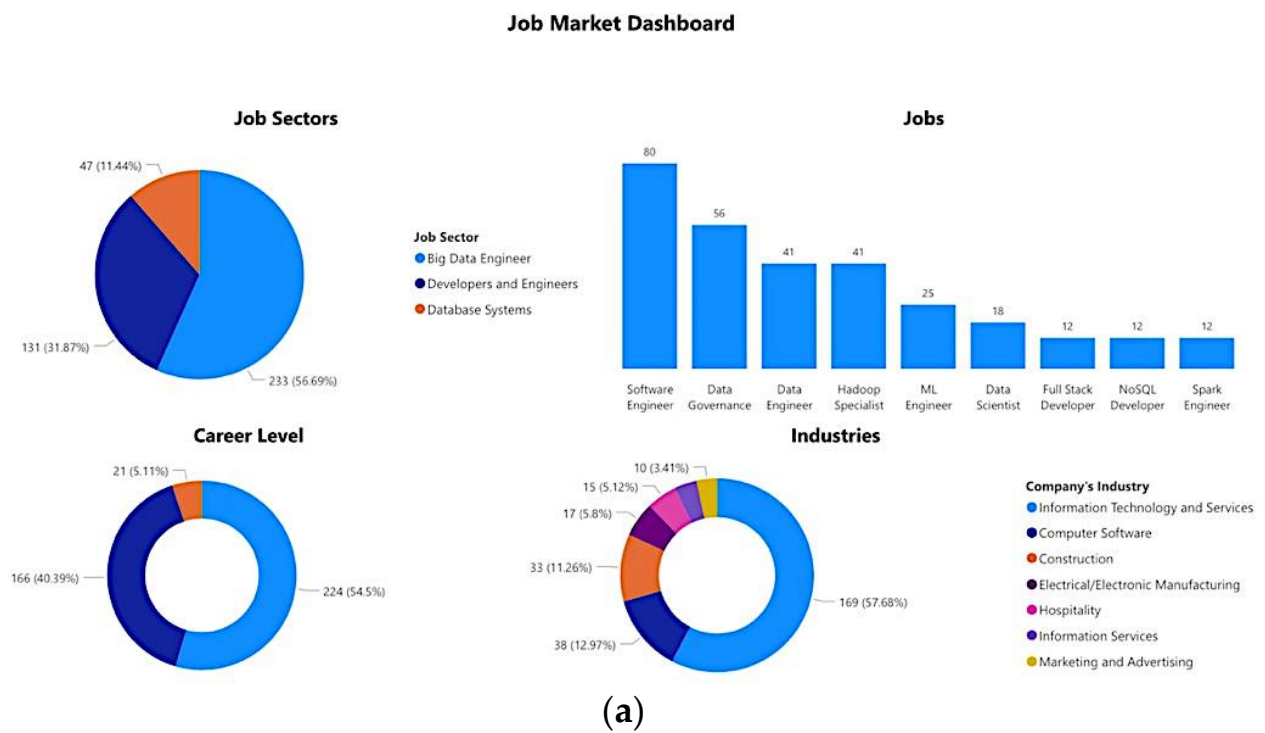
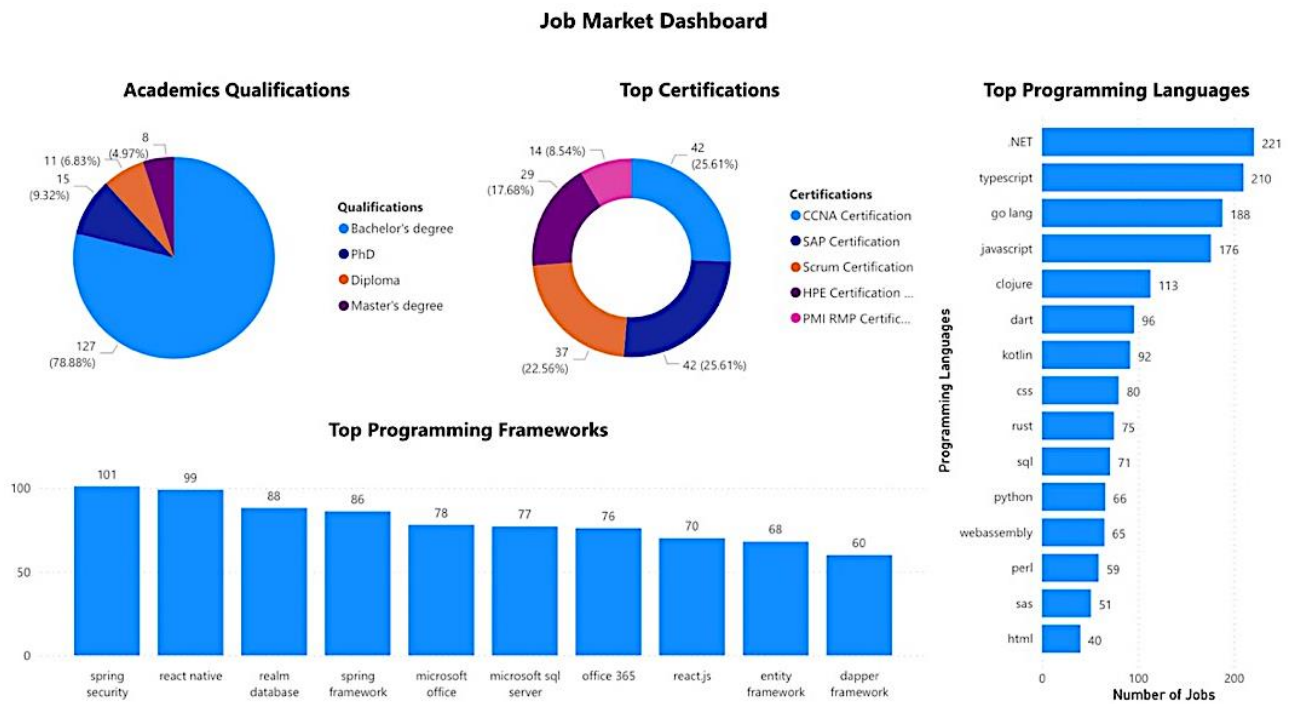
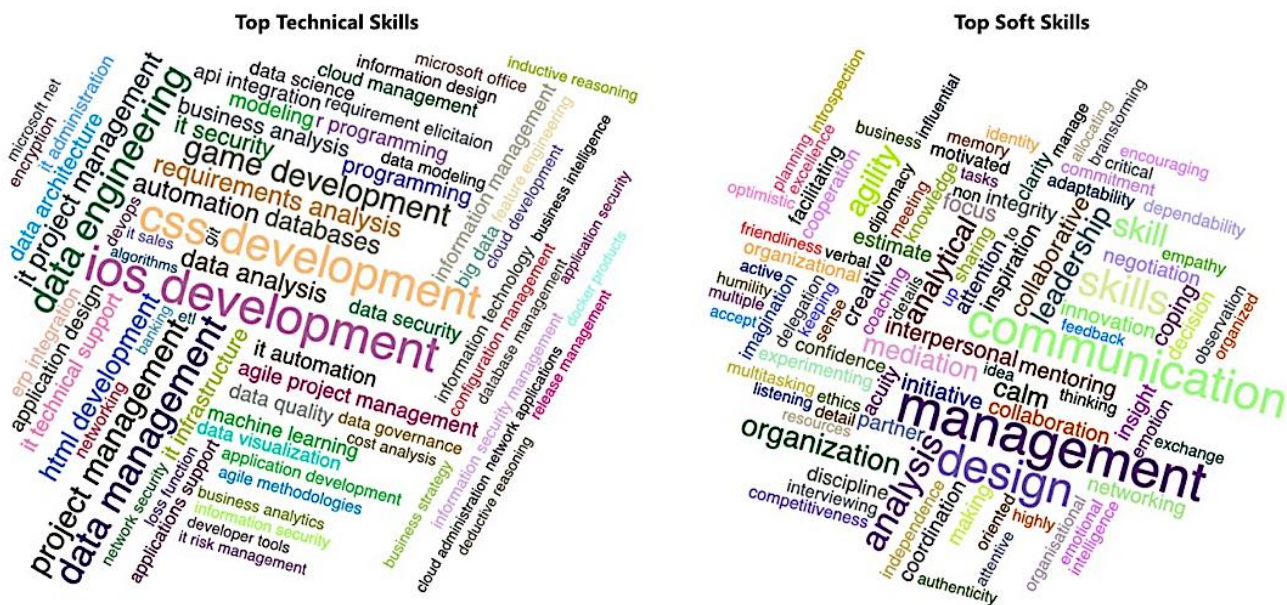


Figure 3. Cont.



(c)

Job Market Dashboard

(d)

Figure 3. (a). Saudi job market dashboard across job sectors, career levels, and industries. (b) Saudi job market dashboard across locations. (c) Saudi job market dashboard across academic qualifications, certifications, programming languages and frameworks. (d) Saudi job market dashboard across technical and soft skills.

Analysis of the distribution of jobs across the regions revealed the interesting insight that those in the selected areas are mostly in only Riyadh or Eastern Province (see Figure 3b).

Furthermore, we show the Saudi job market data points needed across the qualifications, certifications, and programming languages and frameworks in the selected job sectors (see Figure 3c). Here, .NET appears to be in most of the jobs, representing C# and its ecosystem, as .NET is not a programming language but a platform for building software. Surprisingly, a Ph.D. degree is a requirement for around 10% of jobs in the selected sector, indicating that a large number of companies are focused on research and development, which is indeed the backbone of any expanding industry. Furthermore, we show the top required certifications and the top programming languages and frameworks for the Saudi job market in the selected fields.

Lastly, we show the top technical and soft skills required by the job markets in the selected sectors (see Figure 3d). We found that the top three technical skills most required by the market are “ios development”, “css development” and “game development”. Interesting, soft skills such as “management”, “communication”, and “design” appear to be the top soft skill required for the selected jobs in the technology sector.

Figure 3a,b show an analysis of over 400 jobs from the Saudi job market. The top three job sectors are Software Engineer, Data Governance, and Data Engineer, and Figure 3c,d show associated data points across all job sectors. To better understand required programming languages, frameworks, and technical and soft skills associated with respective job sectors, we drill down to analyse data points concerning specific jobs in the following sections. Our proposed method extracts content from the respective CVs of individuals interested in excelling in a given job sector (e.g., Data Scientists, NoSQL or Software Engineer). It then matches their profile data points with the relevant jobs available. Further, the proposed method shows recommended job titles within a selected job sector along with the recommended job certifications and technical and soft skills required by the market.

4.2. Recommendations for Data Science Jobs

In this subsection, we demonstrate the Power BI dashboard to recommend jobs and relevant skills to individual profiles in the areas of data science. Figure 4a shows the recommended job titles and required certification, along with technical and soft skills required by an individual in the study in the job category relevant to data science. The gap analysis between the aspiring student wanting to become a data scientist and the skills required by the job market makes several recommendations across various data points. The analysis showed that the recommended job for the selected individual is data scientist, followed by data analyst and data engineer. Although the certifications data only show IBM and Azure, note that Amazon AWS and Google Cloud are also very desired certifications in international job market.

The dashboard (see Figure 4a) also highlighted the recommended certification, such as IBM Data Science or IBM Data Science Professional, as the most significant in this sector. While machine learning, statistics, and data visualisation are the most required technical skills, the most prominent soft skills in the sector design appear to be responsibility and analysis. Finally, we found the topmost recommended programming frameworks to be realm database and ipython, and the most highly recommended programming languages for this job sector are python, dart, and .net (see Figure 4b).

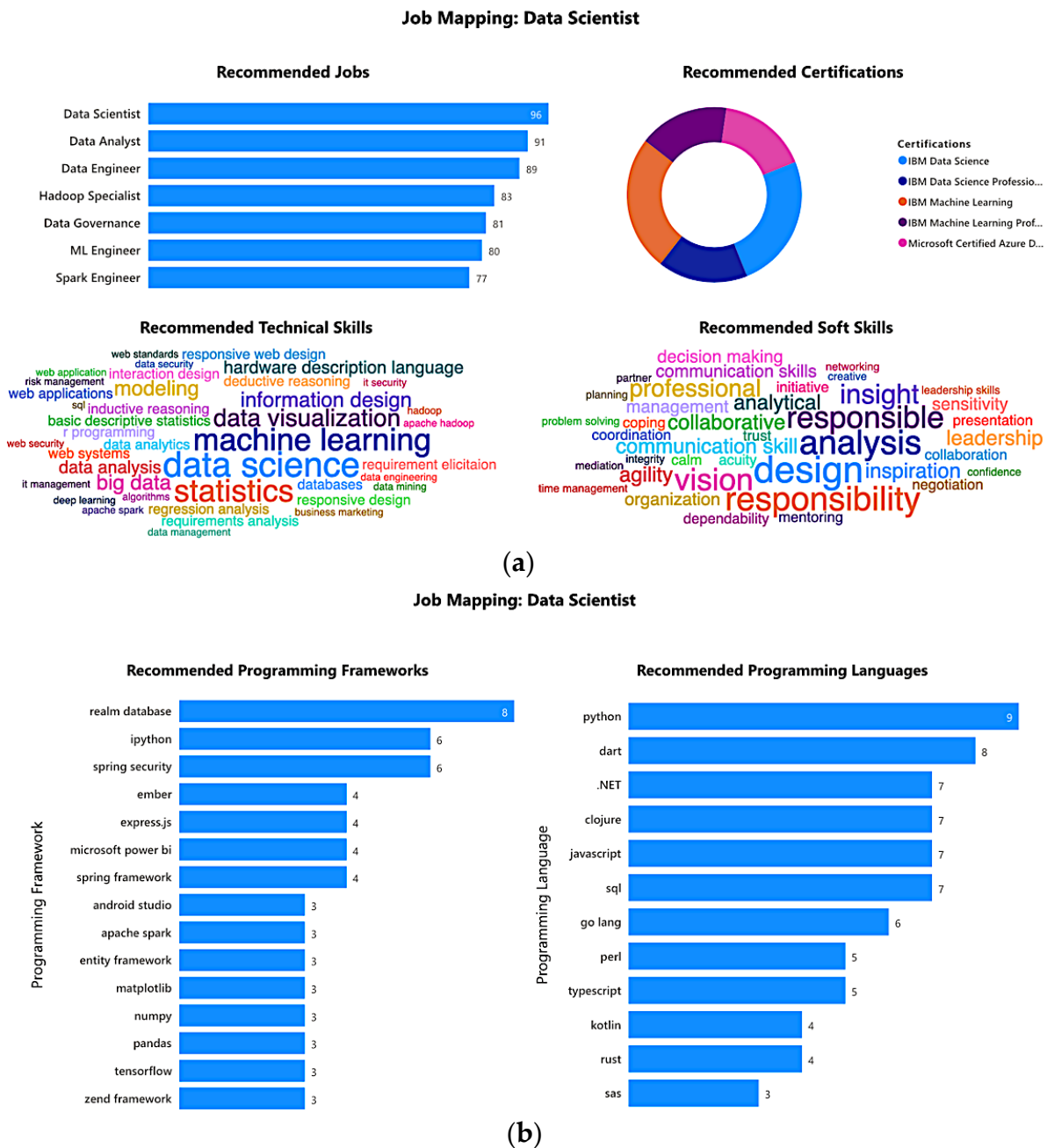


Figure 4. (a) Recommended jobs, certifications, and technical and soft skills for data scientists. (b) Recommended programming frameworks and languages for data scientists.

4.3. Recommendations for NoSQL Jobs

In this subsection, we show individuals' most recommended jobs and relevant skills in the areas of NoSQL. Across the job categories relevant to NoSQL jobs, we show recommended job titles and the certification required by an individual in the study, along with the required technical and soft skills (see Figure 5a). The gap analysis between software engineers and the requirements of the job market skills shows several recommendations across various data points; for the selected individual in NoSQL, MySQL developer is among the most recommended jobs. The dashboard also highlights as the most significant in this sector the recommended certifications of Master in SQL Development and Oracle Cloud Certification. While web development and sql development appear to be the most required technical skills, the most prominent soft skills in this sector appear to be responsibility and communication skills. Finally, we found that the topmost recommended programming

frameworks for this job sector include Microsoft sql server, realm database, and spring, and programming languages such as JavaScript, .net, and go lang (see Figure 5b).

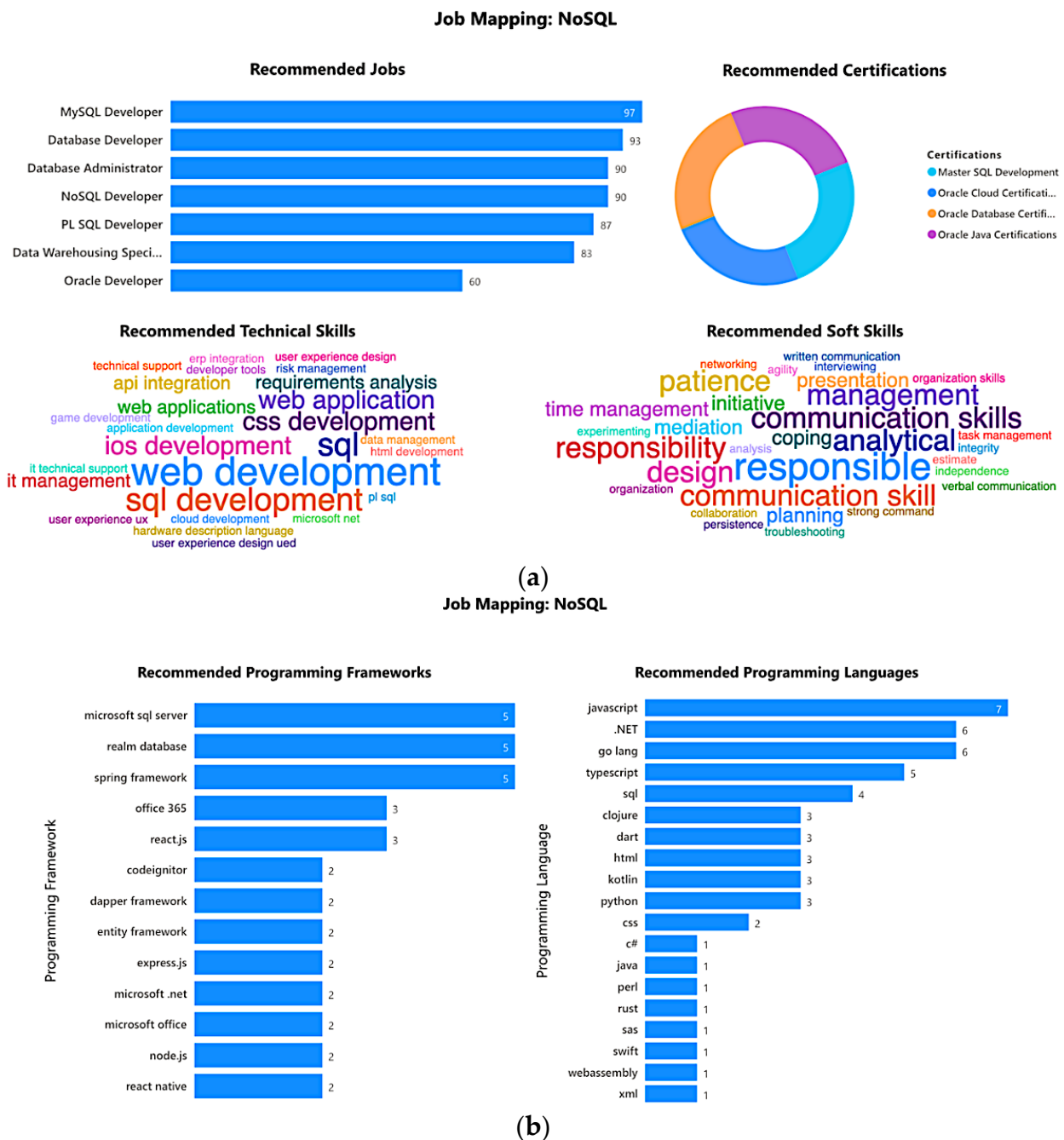


Figure 5. (a) Recommended jobs, certifications, and technical and soft skills for NoSQL. (b) Recommended programming frameworks and languages for NoSQL.

4.4. Recommendations for Software Engineering Jobs

In this subsection, across the job categories relevant to software engineering, we show the recommended job titles and required certifications, along with the required technical and soft skills for an individual in the study (see Figure 6a). The analysis showed that Full Stack is among the most recommended jobs for the selected individual in software engineering. The dashboard also highlighted scrum certification as among the most significant in this sector. While IT management and web development appear to be the most required technical skills, the most prominent soft skills required in this sector appear to

be design, management, and agility. Finally, we found that react native and express.js are among the top recommended programming frameworks, and the programming languages most recommended for this job sector are .net and typescript (see Figure 6b).

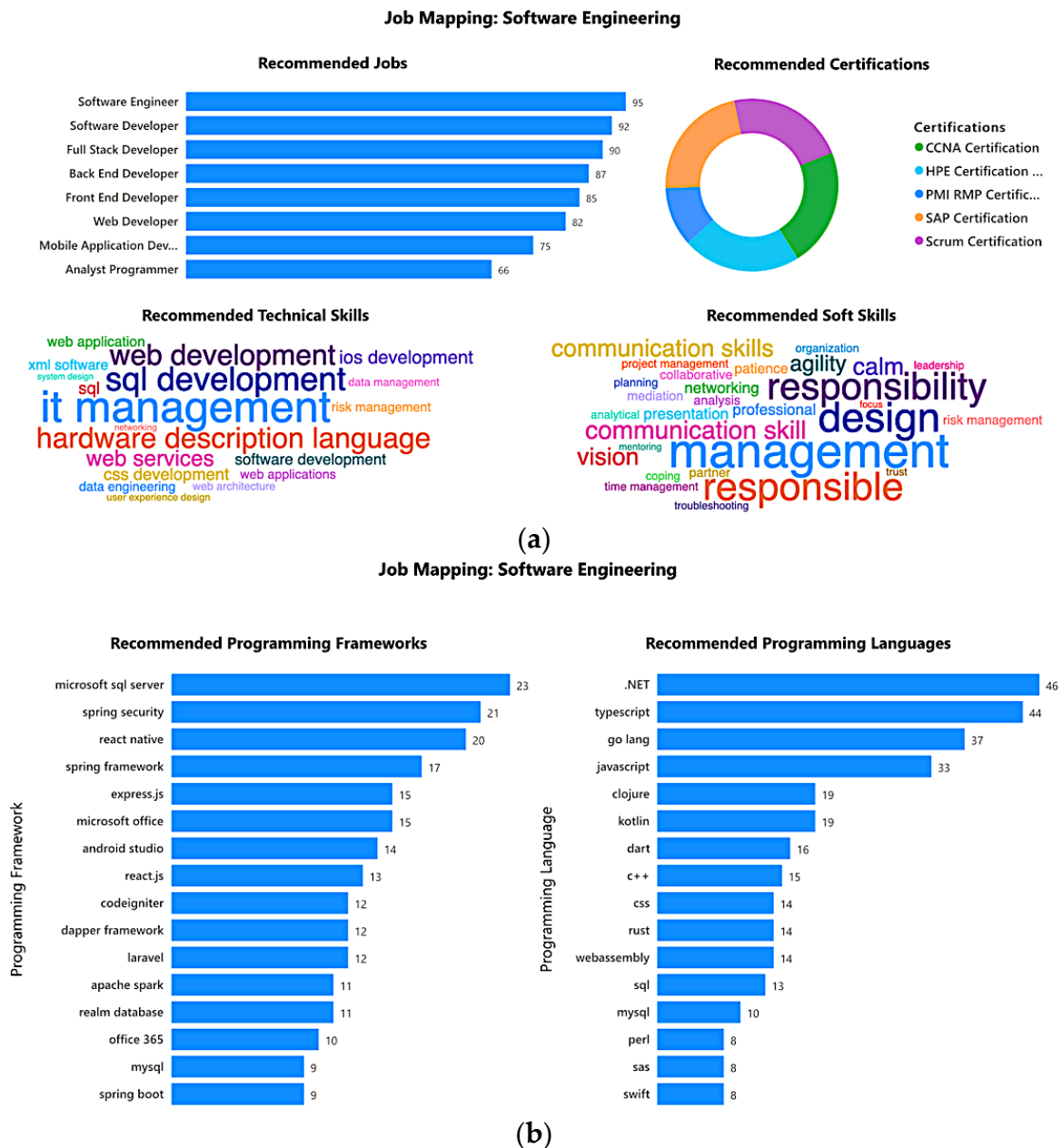


Figure 6. (a) Recommended jobs, certifications, and tech and soft skills for software engineers. (b) Recommended programming frameworks and languages for software engineering.

5. Conclusions

In this study, through the quantitative interpretation of data, we provided analytical insights for students, universities, and employers to assist them in matching skills and jobs. By focusing the technical aspects of a job profile, the work helps students who need extra education or training to develop the abilities necessary for a certain position. This study revealed the latest developments in AI, ML, and learning analytics, and we introduced novel, innovative approaches for a viable and sustainable framework for skills and competencies related to digital transformation in the Saudi economy. We showed that models that can monitor and describe current and future market needs are relevant to sophisticated, technology-driven specialities and expertise—a crucial initiative.

The study is a manifestation of the “Thriving Economy Rewarding Opportunities” initiative. It delivers a sophisticated infrastructure that jointly targets the following. First is the sustainable integration of academia and industry, based on know-how transfer and an innovative, machine-learning infrastructure that permits continuous monitoring of the skills and competencies required by industry for the personalisation and customisation of training content in educational and academic environments. Next is a systematic mapping of job profiles and skills and competencies that are essential to socially inclusive, sustainable economic growth in the field of technology-driven innovation. Then is an industry-oriented ecosystem of advanced learning analytics to describe learning efficiency for working. Finally is a unique Saudi knowledge management system to attract the talent needed for the KSA’s digital transformation.

This paper is a part of KSA’s ongoing investment in education and training to ensure that young Saudi men and women are prepared for future careers. It promotes the idea that Saudi children, no matter where they live, will benefit from a more comprehensive, multifaceted education. In future, this research could support the development of a dependable, efficient dashboard powered by machine learning algorithms to analyse Saudi and global web data, providing a novel and custom system to plan successful industry-oriented careers supporting Vision 2030 goals. Furthermore, this study helps provide guidelines and prepare the young Saudi generation for the digital transformation under the fourth industrial revolution (i.e., industry 4.0). This revolution covers the use of digital systems; sensors in the Internet of Things (IoT); industrial automation, AI and robotics; data and remote operation. That enables improved planning and forecasting, increases sustainable manufacturing of products, and provides competitive advances in the production life cycle.

One of the limitations of this work is that the case study in this paper only represents Saudi job data; thus, the results cannot be generalised to the international job market in the selected fields. Nevertheless, the Saudi job analyses encourage the alignment of business and industrial sectors to academia, identifying the critical areas to develop capabilities and skills.

Author Contributions: Conceptualization, S.-U.H. and N.R.A.; methodology, S.-U.H.; software, M.A.A. and A.O.K.; validation, N.R.A., M.A.A. and A.O.K.; investigation, S.-U.H. and N.R.A.; resources, S.-U.H. and N.R.A.; data curation, S.-U.H. and N.R.A.; writing—review and editing, N.R.A., M.A.A., A.O.K. and S.-U.H.; visualization, S.-U.H. and N.R.A.; supervision, S.-U.H. and N.R.A.; project administration, N.R.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Institutional Fund Projects under grant no. IFPRC-104-611-2020. Therefore, the authors gratefully acknowledge technical and financial support from the Ministry of Education and King Abdulaziz University, Jeddah, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. AlArjani, A.; Modibbo, U.M.; Ali, I.; Sarkar, B. A new framework for the sustainable development goals of Saudi Arabia. *J. King Saud Univ. Sci.* **2021**, *33*, 101477. [\[CrossRef\]](#)
2. Aljohani, N.R.; Aslam, A.; Khadidos, A.O.; Hassan, S.U. Bridging the skill gap between the acquired university curriculum and the requirements of the job market: A data-driven analysis of scientific literature. *J. Innov. Knowl.* **2022**, *7*, 100190. [\[CrossRef\]](#)
3. Alfalih, A.A. The role of sustainable entrepreneurship and corporate social performance on social innovation: The case of the private industrial sector in Saudi Arabia. *J. Knowl. Econ.* **2021**, 1–16. [\[CrossRef\]](#)
4. Alwakid, W.; Aparicio, S.; Urbano, D. The influence of green entrepreneurship on sustainable development in Saudi Arabia: The role of formal institutions. *Int. J. Environ. Res.* **2021**, *18*, 5433. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Hassan, S.U.; Sarwar, R.; Muazzam, A. Tapping into intra-and international collaborations of the Organization of Islamic Cooperation states across science and technology disciplines. *Sci. Public Policy* **2016**, *43*, 690–701. [\[CrossRef\]](#)

6. Sarwar, R.; Soroya, S.H.; Muazzam, A.; Sabah, F.; Iqbal, S.; Hassan, S.U. A bibliometric perspective on technology-driven innovation in the Gulf Cooperation Council (GCC) countries in relation to its transformative impact on international business. In *Technology-Driven Innovation in Gulf Cooperation Council (GCC) Countries: Emerging Research and Opportunities*; IGI Global: Hershey, PA, USA, 2019; pp. 49–66.
7. Abad-Segura, E.; Gonzalez-Zamar, M.D. Sustainable economic development in higher education institutions: A global analysis within the SDGs framework. *J. Clean. Prod.* **2021**, *294*, 126–133. [\[CrossRef\]](#)
8. Alotaibi, H. Knowledge positioning and innovation in strategic alignment for SESCO and Saudi Arabia's economy based on its 2030 vision. *Afr. J. Bus. Manag.* **2021**, *15*, 232–241.
9. MacLilwain, C. Science economics: What science is really worth. *Nature* **2010**, *465*, 682–684. [\[CrossRef\]](#)
10. Taala, W. Impact of Saudi Digital Library (SDL) to Saudi research output: A review. *Open Access Libr.* **2019**, *6*, 1. [\[CrossRef\]](#)
11. Baqadir, A.; Patrick, F.; Burns, G. Addressing the skills gap in Saudi Arabia: Does vocational education address the needs of private sector employers? *J. Vocat. Educ.* **2011**, *63*, 551–561. [\[CrossRef\]](#)
12. Hasanov, F.J.; Mikayilov, J.I.; Javid, M.; Al-Rasasi, M.; Joutz, F.; Alabdullah, M.B. Sectoral employment analysis for Saudi Arabia. *Appl. Econ.* **2021**, *53*, 5267–5280. [\[CrossRef\]](#)
13. Tomy, S.; Pardede, E. Course map: A career-driven course planning tool. In *International Conference on Computational Science and Its Applications*; Springer: Cham, Switzerland, 2018; pp. 185–198.
14. Nawaz, R.; Sun, Q.; Shardlow, M.; Kontonatsios, G.; Aljohani, N.R.; Visvizi, A.; Hassan, S.U. Leveraging AI and Machine Learning for National Student Survey: Actionable Insights from Textual Feedback to Enhance Quality of Teaching and Learning in UK's Higher Education. *Appl. Sci.* **2022**, *12*, 514. [\[CrossRef\]](#)
15. Tomy, S.; Pardede, E. Map my career: Career planning tool to improve student satisfaction. *IEEE Access* **2019**, *7*, 132950–132965. [\[CrossRef\]](#)
16. Tomy, S.; Pardede, E. An entrepreneurial intention model focussing on higher education. *Int. J. Entrep. Behav. Res.* **2020**, *26*, 1423–1447. [\[CrossRef\]](#)
17. Waheed, H.; Hassan, S.-U.; Aljohani, N.R.; Wasif, M. A bibliometric perspective of learning analytics research landscape. *Behav. Inf. Technol.* **2018**, *37*, 941–957. [\[CrossRef\]](#)
18. Safder, I.; Hassan, S.-U.; Visvizi, A.; Noraset, T.; Nawaz, R.; Tuarob, S. Deep learning-based extraction of algorithmic metadata in full-text scholarly documents. *Inf. Process. Manag.* **2020**, *57*, 102269. [\[CrossRef\]](#)
19. Aljohani, N.R.; Fayoumi, A.; Hassan, S.U. A Novel Deep Neural Network-Based Approach to Measure Scholarly Research Dissemination Using Citations Network. *Appl. Sci.* **2022**, *11*, 10970. [\[CrossRef\]](#)
20. Mezhoudi, N.; Alghamdi, R.; Aljunaid, R.; Krichna, G.; Düşteğör, D. Employability prediction: A survey of current approaches, research challenges and applications. *J. Ambient. Intell. Humaniz. Comput.* **2021**, 1–17. [\[CrossRef\]](#)
21. News.com.au, Nation of Dropouts: University Completion Rates Drop to a New Low. 20 January 2017. Available online: <https://www.news.com.au/finance/work/careers/nation-of-dropouts-university-completion-rates-drop-to-a-new-low/news-story/1265f4d9872db263694aaa74f815c432> (accessed on 13 June 2022).
22. Ibrahim, S.A.E.S. The role of social media in shaping the trends of university youth towards the requirements of the knowledge economy. *J. Southwest Jiaotong Univ.* **2021**, *56*. [\[CrossRef\]](#)
23. Nieuwoudt, J.E.; Pedler, M.L. Student retention in higher education: Why students choose to remain at university. *J. Coll. Stud. Retent. Res. Theory Pract.* **2021**, 1521025120985228. [\[CrossRef\]](#)
24. Hsu, S.H.; Wang, Y.C.; Cheng, C.J.; Chen, Y.F. Developing a decomposed alumni satisfaction model for higher education institutions. *Total Qual. Manag. Bus. Excell.* **2016**, *27*, 979–996. [\[CrossRef\]](#)
25. Mello, L.V.; Wattret, G. Developing transferable skills through embedding reflection in the science curriculum. *Biophys. Rev.* **2021**, *13*, 1–7. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Boyle, C.F.; Dykstra, D.I.; Monarch, I.A. Using knowledge representation to study conceptual change in students for teaching Physics. In *12th Annual Conf. CSS Pod*; Psychology Press: New York, NY, USA, 2022; pp. 788–795.
27. Xu, D.; Ruan, C. Modern theoretical tools for understanding and designing next-generation information retrieval system. In *Proceedings of the 15th ACM International Conference on Web Search and Data (WSDM)*, Virtual, 21–25 February 2022; pp. 1635–1637.
28. Chen, H.; Luo, X. An automatic literature knowledge graph and reasoning network modeling framework based on ontology and natural language processing. *Adv. Eng. Inform.* **2019**, *42*, 100959. [\[CrossRef\]](#)
29. Succi, C.; Canovi, M. Soft skills to enhance graduate employability: Comparing students and employers' perceptions. *Stud. High. Educ.* **2020**, *45*, 1834–1847. [\[CrossRef\]](#)
30. Bennett, D. Graduate employability and higher education: Past, present and future. *HERDSA Rev. High. Educ.* **2019**, *5*, 31–61.
31. Cox, S.; King, D. Skill sets: An approach to embed employability in course design. *Educ. Train.* **2006**, *48*, 262–274. [\[CrossRef\]](#)
32. Avram, S. Do workers like employer-driven flexibility? Experimental evidence on work and pay uncertainty and willingness to work. *Eur. J. Work. Organ. Psychol.* **2022**, 1–13. [\[CrossRef\]](#)
33. Crosling, G.; Heagney, M.; Thomas, L. Improving student retention in higher education. *Aust. Univ. Rev.* **2009**, *51*, 9–18.
34. Waheed, H.; Hassan, S.U.; Aljohani, N.R.; Hardman, J.; Alelyani, S.; Nawaz, R. Predicting academic performance of students from VLE big data using deep learning models. *Comput. Hum. Behav.* **2020**, *104*, 106189. [\[CrossRef\]](#)

35. Hassan, S.U.; Waheed, H.; Aljohani, N.R.; Ali, M.; Ventura, S.; Herrera, F. Virtual learning environment to predict withdrawal by leveraging deep learning. *Int. J. Intell. Syst.* **2019**, *34*, 1935–1952. [[CrossRef](#)]
36. Wasif, M.; Waheed, H.; Aljohani, N.R.; Hassan, S.U. Understanding student learning behavior and predicting their performance. In *Cognitive Computing in Technology-Enhanced Learning*; IGI Global: Hershey, PA, USA, 2019; pp. 1–28.
37. Waheed, H.; Anas, M.; Hassan, S.U.; Aljohani, N.R.; Alelyani, S.; Edifor, E.E.; Nawaz, R. Balancing sequential data to predict students at-risk using adversarial networks. *Comput. Electr. Eng.* **2021**, *93*, 107274. [[CrossRef](#)]
38. Robertson, S.E.; Walker, S.; Jones, S.; Hancock-Beaulieu, M.M.; Gatford, M. Okapi at TREC-3. In *NIST Special Publication 500–226: Overview of the Third Text REtrieval Conference (TREC-3)*; Centre for Interactive Systems Research: Gaithersburg, MD, USA, 1995; pp. 109–126.