Vol. 8, No. 4, August 2026

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

Association Rule Mining for Uncovering Co-occurring Skill Sets in Multi-Portal ICT Job Advertisements

Erick Samwi[1], George Matto[2], Mercy Mbise[3]

[1,2]Department of Information and Communication Technology, Moshi Co-operative University, Moshi, Tanzania

[3] Department of Computer Science and Engineering, University of Dar es Salaam, Dar es Salaam, Tanzania.

Received:
February 17, 2026
Revised:
May 27, 2026
Accepted:
June 24, 2026
Published:
August 22, 2026

Corresponding Author:
Author Name*:
Erick Samwi
Email*:
erick.samwi@mocu.ac.tz

DOI:
10.63158/journalisi.v8i4.1671

Abstract. Rapid changes in the ICT labour market require stakeholders to understand not only which skills are demanded but also how these skills occur together. However, existing studies provide limited insights into co-occurring skill demands in ICT jobs. This study addresses this gap by uncovering skill sets that frequently appear together in ICT job descriptions. A total of 15,653 unique ICT job advertisements published between 2020 and 2025 were collected from LinkedIn, AjiraYako and Mabumbe to capture regional and international contexts. Guided by the Cross-Industry Standard Process for Data Mining (CRISP-DM), association rule mining using the Apriori algorithm identified frequent itemsets and generated rules evaluated using support, confidence and lift. The analysis revealed strong co-occurrences among programming skills, such as C# and C++ (confidence = 1.00; lift = 2.41), although this combination occurred in a small proportion of advertisements (support = 0.0096). Java and Python showed the highest support (0.1818), with confidence values of 0.51 and 0.42 and a lift of 1.19, which indicate a positive association despite belonging to different application domains. By uncovering skill associations beyond individual skill frequencies, the study provides richer labour market intelligence for workforce development and policy formulation.

Keywords: Association Rule Mining, ICT Skills, Job Advertisements, Labour Market Analytics, Web Scraping, Text Mining, Apriori Algorithm

1. INTRODUCTION

The labour market is characterized by dynamic skill requirements in response to technological and organizational changes [1]. As one of the fastest growing industries, the ICT sector is undergoing rapid technological changes, which in turn are transforming the work environment and skill demands [2]. The changing nature of ICT skill demands can be explained by Skill-Biased Technological Change (SBTC) theory, which suggests that technological advancements increase productivity and demand for workers with relevant skills that complement emerging technologies. As a result, employers seek combinations of complementary skills rather than isolated competencies. These demands for combinations of skills make systematic labour market analysis essential for understanding not only which skills are in demand, but also how they are co-occurring within specific job roles [3].

Co-occurring skills refer to combinations of competencies which are jointly required within the same job role. The values of these skills are derived from their simultaneous application rather than as a standalone skill. These skill interdependencies are observed across fields such as engineering where technical design skills co-exist with project management and regulatory compliance skills [4]. Presence of co-occurring skills in job roles shows that modern occupations demand employees with integrated skill sets. In ICT sector, combinations of skills required together are frequently changing because of rapid changes in technological demands [5], [6].

Understanding co-occurring skill sets is important because many ICT roles require combination of complementary skills rather than isolated competencies [7]. While analysis of individual skill frequencies provides useful insights into general demand trends, it may not capture how skills are structured and integrated within specific ICT job roles [8]. This gap restricts insights into how skills are structured and applied within specific job roles and thus limits the ability of educators and policymakers to respond effectively to evolving ICT labour market skill demands.

Traditional market analyses have focused on identifying and ranking individual skills through descriptive statistical analysis which treat skills as independent requirements. While these approaches provide valuable insights into skill requirements, they do not capture how skills are jointly required within job roles [9]. As a result, the interdependence between skills required by employers in modern ICT job roles remains underexplored.

Current studies have analysed job descriptions using text mining and natural language processing techniques to understand skills demand in the labour market. Studies such as that conducted by Fareri *et al.* [10] applied text mining technique to extract skills data from a structured database for jobs and skills of a multinational company, while Hassani *et al.* [11] employed similar techniques to facilitate skill matching using data from recruitment portal. In addition, Li *et al.* [12] and Johnson *et al.* [13] used text mining approaches to extract and categorize skills from job advertisements. Although these studies successfully identified skills demand, their analytical focus was limited to extraction of individual skills. Methodologically, they did not apply techniques capable of uncovering co-occurrence patterns among skills within the same job role.

Classification and forecasting of ICT job skills is another area explored in existing studies. Studies by Sakurada *et al.* [14] and Ternikov [15] classified ICT skills based on job roles, while Pichler and Stehrer [16] focused on forecasting future ICT skill demands based on the trend in the industry. In addition, Ng *et al.* [17] employed predictive approaches to align skills development with technological change. Similarly, Taherdoost [18] provided general insights into skill demand trends but did not examine skills co-occurrence in job descriptions. While these studies advanced understanding of skill evolution and changes in industry demands, they considered skills as independent units of analysis. The techniques applied were not designed to examine how multiple skills are required together within individual job advertisements.

Therefore, this study aimed to uncover co-occurring skill sets within ICT job roles by applying association rule mining to multi-portal ICT job advertisement data. To achieve this objective, the study addressed the following research question: What co-occurring skill sets and cross-domain skill associations emerge from ICT job roles in multi-portal job advertisements? Online ICT job advertisements were specifically used because they are frequently updated and reflect current needs of employers [19], [20]. Unlike printed advertisements and periodic surveys, job portals provide detailed descriptions of roles and skill requirements [21]. To provide a broader representation, ICT job advertisements were collected from both regional and international online portals. The resulting dataset therefore reflects a mixed labour market context comprising Tanzanian, regional and international ICT job advertisements. This enabled the analysis to capture co-occurring skill demands across different ICT labour market contexts.

Existing studies largely identify the most frequently demanded ICT skills but provide limited understanding of how these skills are required together within job roles. The contribution of this study lies in shifting labour market analysis from individual skill frequency analysis to skill interdependency analysis by uncovering co-occurring skill sets in ICT job advertisements. This reveals complementary competencies sought by employers that may be overlooked in frequency-based approaches. Such insights can guide curricula design by informing the integration of related competencies into bundled learning pathways. In addition, the identified skill interdependencies can support workforce planning by guiding targeted reskilling and upskilling initiatives aligned with evolving labour market demands.

2. METHODS

This study adopted Cross-Industry Standard Process for Data Mining (CRISP-DM) to identify co-occurring skill sets within ICT job roles. This methodology provides iterative framework for uncovering hidden patterns in large datasets with varied formats. The methodology fits this study to analyze job advertisements where skills appear as unstructured text and insights are based on association among skills rather than individual occurrences. The study was conducted in line with CRISP-DM six stages namely: (i) business understanding, (ii) data understanding, (iii) data preparation, (iv) modeling, (v) evaluation and (vi) deployment. Each stage and its implementation is described in the subsequent subsections.

2.1 Problem Framing and Analytical Scope

Descriptive statistical analysis was conducted to identify leading ICT job roles and define the analytical scope prior to modelling. Treating all job advertisements as homogenous professional functions would lead to aggregation bias as skill requirements vary across ICT job roles. Therefore, the analysis was conducted at two levels. The first level involved leading job roles to enable identification of role specific skills associations rather than aggregated patterns. The second level analysed the complete ICT dataset to identify broader co-occurrence patterns across domains. This approach enabled both role-specific and sector-wide insights into skill co-occurrences.

2.2 Job Data Sources and Characteristics

2.2.1 Job Data Sources

Job data were collected from three job portals named LinkedIn (https://www.linkedin.com), ajirayako (https://ajirayako.co.tz) and mabumbe (https://mabumbe.com/jobs). The use of multiple portals intended to improve representativeness of skill co-occurrence patterns. This is because job portals differ in structure and job coverage where a single portal captures only a fraction of the overall ICT labour market [22]. In addition, relying on data from a single job portal results in limited scalability and adaptability across different platforms [23].

The selection of job portals was based on their reliability as trusted sources of job advertisements, the ability to keep history of expired job advertisements, and availability of posts from multiple companies as opposed to company specific job sites that only feature advertisements from a single organization. LinkedIn represents global professional networks setting while AjiraYako and Mabumbe cover Tanzanian and regional ICT labour market characteristics. No additional geographic filters were applied beyond those inherent to the selected portals.

2.2.2 Job Data Characteristics

Comparative assessment of Hypertext Markup Language (HTML) fields was done to examine variations in structure and content of job advertisements across three job portals. The assessment involved comparison of fields containing job titles, date posted, employer and the job description. Structural differences analysed informed selection of data acquisition techniques across three job portals.

2.3 Data Acquisition and Preparation

Data collection was conducted in September 2025 and comprised job advertisements published between January 2020 and September 2025. Web scraping techniques were applied and advertisements collected were prepared into a dataset suitable for association rule mining. The study collected only publicly accessible job advertisements and excluded personal information. Job data collected were used solely for academic research purposes and analysed in aggregate form without individual identities. The web scraping was conducted responsibly by limiting requests to avoid disrupting platform operations. The data collection procedures were designed to respect terms of service and access policies of the respective job portals. Portal-specific processing procedures were implemented to accommodate differences in website structure and data organization. Specifically, Selenium was used to simulate user authentication and triggers requirements of LinkedIn portal, while BeautifulSoup (version 4.12.3) library was employed to read HTML tags from Mabumbe and Ajirayako sites. Data preprocessing involved deduplication, aggregation and cleaning to make the dataset ready for association rule modelling.

2.3.1 Scraping of Job Postings from LinkedIn

On LinkedIn portal, jobs are identified using unique identification numbers and are not organized into distinct categories. Therefore, the scraping process involved collecting job titles and descriptions from all available advertisements and storing them in a single csv file. A supervised binary text classification approach was then employed to retain only ICT-related postings. Training labels were automatically derived from job categories available in AjiraYako and Mabumbe, where vacancies were explicitly classified as ICT and non-ICT. These portal-based categories served as reference labels for constructing the labelled dataset. The dataset was randomly divided into training and testing subsets using an 80:20 ratio (random seed = 42). Job descriptions were transformed into numerical feature representations using the Term Frequency-Inverse Document Frequency (TF-IDF) vectorizer implemented in scikit-learn with default settings. A logistic regression classifier was trained on the TF-IDF features to distinguish ICT from non-ICT job advertisements. The classifier was implemented using the scikit-learn library with default parameter settings (solver = 'lbfgs', L2 regularization, and max_iter = 100). Model performance was evaluated using precision, recall and F1-score on the independent test set. No additional manual validation was conducted because classification performance was assessed using the held-out test dataset.

2.3.2 Scraping of Job Postings from AjiraYako and Mabumbe

Ajirayako and Mabumbe web pages are arranged in series for each job category defined in the system. This arrangement makes navigation specific to ICT related jobs easier. The arrangement of pages for ICT jobs is as follows:

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

First page: *'base url'*/job-category/*'category'*/

Second page: *'base url'*/job-category/*'category'*/page/2

nth page: 'base url'/job-category/'category'/page/n

In these sites, jobs are not uniformly structured and are termed differently across advertisements in the same portal. To enable reading all available important job information, the HTML tags for job description containers were located from page source files as indicated in the snippet shown in Figure 1.

```
    <div class="job-desc" itemprop="description">
            </p><div class='code-block code-block-4' style='margin: 8px auto; text-alig
<script async src="https://pagead2.googlesyndication.com/pagead/js/adsbygoogle.js?client=ca

<ins class="adsbygoogle"
     style="display:block"
     data-ad-client="ca-pub-7553421991920529"
     data-ad-slot="4158801095"
     data-ad-format="auto"
     data-full-width-responsive="true"></ins>
<script type="pmdelayedscript" data-cfasync="false" data-no-optimize="1" data-no-defer="1"
     (adsbygoogle = window.adsbygoogle || []).push({});
</script></div>
<p><span style="font-size: 12pt;"><strong>Job Opportunity</strong></span><br />
<span style="font-size: 12pt;"><strong>Position:</strong> IT Support Specialist</span></p>

………………………………………………

<ol start="1">
<li><span style="font-size: 12pt;"><strong>Educational Background:</strong></span><br />
<span style="font-size: 12pt;">Bachelor's degree or Diploma in Computer Science, Informat
<li><span style="font-size: 12pt;"><strong>Technical Skills:</strong></span>
<ul>
<li><span style="font-size: 12pt;"><strong>Programming Languages:</strong> Proficiency in
<li><span style="font-size: 12pt;"><strong>API Support:</strong> Proficiency in troublesh
<li><span style="font-size: 12pt;"><strong>Database Management:</strong> Knowledge of SQL
<li><span style="font-size: 12pt;"><strong>Operating System:</strong> Knowledge and exper

…………………………………………………

</script></div>
```

Figure 1. The job description container located from the job site's page source

Based on the HTML tag of the class that contains the description of the job, the function

job_desc_container = soup.find("div", class_="job-desc", itemprop="description"):

was applied where job requirement information, including education background, technical skills, job summary, duties and responsibilities, qualifications and experience were read from the job site.

2.3.3 Data Preprocessing

Scraping across multiple sources introduces duplicate postings because employers may post the same job on different platforms. Duplicates across portals were identified using a deterministic hashing approach implemented in Python version 3.12.7. The unique identifiers for each job advertisement were generated using *hashlib* library of Python from concatenation of job title, company name, location and date posted. Job advertisements with identical hash values were treated as duplicates. Unique advertisements were merged into a single data-frame using *pandas* (version 2.3.2) library in Python. The final aggregated job data was stored in CSV file. Although this approach effectively identified exact duplicates, advertisements describing the same vacancy with minor textual variations or inconsistencies in metadata may not have been detected.

Aggregated job data were cleaned and standardized for further analysis. This involved creating a uniform representation of text as well as removal of noise including HTML tags and special characters. Records with missing job descriptions were excluded because they could not contribute to skill extraction. In addition, inconsistent date formats were standardized into a common date representation. Advertisements containing incomplete descriptions were retained only when sufficient information to identify skills was available.

Job descriptions were also filtered by identifying keywords such as "requirements", "skills", "qualifications", "technical competencies" and "experience with" to retain only segments with texts relating to technical skill requirement. A domain-based skill dictionary comprising 51 predefined skills was developed to identify ICT skills from the job description segments. The dictionary covered major ICT categories including programming, databases, data science, networking, cloud computing and systems administration. The use of predefined skill taxonomies and dictionaries has been widely adopted in job advertisement analysis to support systematic extraction and classification of skills from unstructured vacancy text [24]. Skill extraction relied on exact matching between dictionary entries and job description texts. The dictionary primarily consisted of predefined single-word skills; therefore, generalized phrase-detection techniques for multi-word skill expressions were not implemented during preprocessing. Although the dictionary mainly consisted of single-word skills, a limited number of predefined multi-word expressions, such as "SQL Server", "Deep Learning", and "Data Analysis", were retained

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

as exact phrases during matching. To minimize false positives, frequently extracted terms were manually inspected to verify their technical relevance.

To facilitate the analysis, the processed dataset was transformed into Boolean transactional format, where each job posting was represented as a transaction and the presence or absence of specific skill as binary (1 or 0) values. The output of preprocessing was a job dataset where each job entry contained clean and standardized skills information. Figure 2 is a screenshot of a subset for job dataset transformed into Boolean transactional format. Table 1 summarizes the preprocessing pipeline applied in this study.

	C#	C++	Dart	Java	JavaScript	Kotlin	PHP	Perl	Python	R	Ruby	Rust	Scala	Swift	TypeScript
.NET Full Stack Developer/Technical Lead	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
DevOps Engineer II	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Embedded Sensor/EW Real-Time Software Engineer	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
GIS Analyst/Programmer (DC local only - Onsite role)	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Infrastructure Specialist Cloud Platforms (IICS) -	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
PowerBI/GIS Programmer Analyst	FALSE	TRUE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Senior Software Engineer - Kubernetes	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Software Engineer - Proprietary languages	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

Figure 2. Subset of job dataset in Boolean transactional format

Table 1. Data preprocessing steps applied to multi-portal ICT job advertisements prior to association rule mining

Step	Operation	Output
1	Aggregate data from multiple portals	Unified dataset
2	Generate hash identifiers and remove duplicates	Unique job advertisements
3	Remove records with missing descriptions	Cleaned dataset
4	Standardize dates and text format	Standardized dataset
5	Extract technical requirement sections	Skill-related text
6	Apply ICT skill dictionary	Extracted skills
7	Verify extracted skills	Validated skill list
8	Convert skills to Boolean transactions	Transactional dataset
9	Export dataset for Apriori analysis	Modelling-ready dataset

2.4 Skills Co-occurrence Modelling

Several analytical techniques can be applied to labour market skill analysis. These techniques include descriptive frequency analysis, clustering methods, classification

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

models and association rule mining. Descriptive frequency approach is effective for identifying the most demanded individual skills, while clustering and classification techniques are applicable in grouping skills in categories. However, these techniques are not designed to uncover interdependency among skills within the same job role.

Job descriptions contain multiple technical skill requirements. The analytical focus of this study was to uncover co-occurring skill sets from these job descriptions rather than individual skill or categorization of skills. Association rule mining was therefore adopted because it is designed to uncover dependencies and co-occurrence relationships within large datasets [25]. This capability makes this technique suitable for modelling interdependence of ICT skill requirements. The algorithm was implemented using *mlxtend.frequent_patterns* module of the Machine Learning Extensions (MLxtend version 0.23.4) Python library. The *apriori()* function of MLxtend was used to compute frequent itemset while, *association_rules()* function was used to generate association rules. Apriori algorithm was selected due to its strength in handling Boolean transactional datasets to discover frequent itemset.

2.5 Model Evaluation

Co-occurrences among ICT skills were modelled using association rules of the form $X \Rightarrow Y$, where (X) (antecedent) and (Y) (consequent) represent sets of skills occurring within the same job advertisement. The strength and usefulness of the rules were evaluated using support, confidence and lift measures as shown in Equation 1.

$$\text{Support}(X \Rightarrow Y) = P(X \cup Y) = \frac{\sigma(X \cup Y)}{N} \qquad (1)$$

where:

$\sigma(X \cup Y)$ is the number of job advertisements containing both $X$and Y,

N is the total number of job advertisements.

Support, defined in Equation (1), measures the proportion of job advertisements in which both the antecedent and consequent occur together as shown in Equation 2.

$$\text{Confidence}(X \Rightarrow Y) = P(Y \mid X) = \frac{\sigma(X \cup Y)}{\sigma(X)} \qquad (2)$$

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

where:

$\sigma(X)$ is the number of advertisements containing the antecedent.

Confidence, shown in Equation (2), estimates the conditional probability that advertisements containing X also contain Y.

$$\text{Lift}(X \Rightarrow Y) = \frac{\text{Confidence}(X \Rightarrow Y)}{\text{Support}(Y)} = \frac{P(X \cup Y)}{P(X)P(Y)} \quad (3)$$

Lift, given in Equation (3), measures the extent to which the co-occurrence of two skills exceeds what would be expected if they occurred independently. A lift with value greater than one (>1) indicates positive associations between skills beyond random co-occurrence, while a value less than one (<1) indicates that the skills co-occur less frequently than expected by chance.

A minimum support threshold of 0.0005 and a minimum confidence threshold of 0.2 were used in the Apriori algorithm. The low support threshold was selected to capture less frequent but meaningful skill combinations across different ICT domains. Consequently, rules with low support despite high confidence and lift values should be interpreted as representing specialized skill combinations observed in a relatively small proportion of job advertisements rather than broadly prevalent ICT skill requirements. The maximum itemset length was restricted to three skills to maintain interpretability and reduce computational complexity.

2.6 Skills Co-occurrence Model Deployment

This study employed visualization techniques to interpret extracted association rules into understandable representations. This deployment mechanism was adopted to support effective communication of complex association patterns to non-technical stakeholders. Network visualizations were used to communicate strength and frequency of skill co-occurrence patterns through size of nodes representing skills and thickness of arrows representing associations. On the other hand, two-dimensional rule visualization was employed to facilitate comparison of multiple rules and identify dominant skill co-occurrence patterns. Figure 3 illustrates the workflow of the research methodology adopted in this study.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

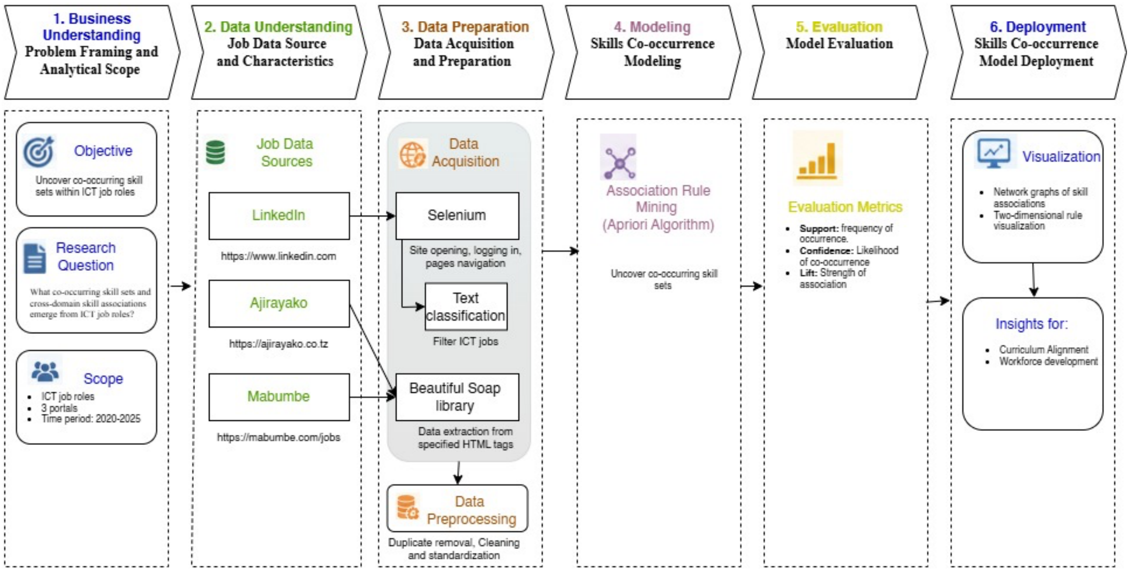

Figure 3. CRISP-DM-based research workflow for uncovering co-occurring skill sets in ICT job advertisements

3. RESULTS AND DISCUSSION

3.1. ICT Job Dataset

A total of 16,455 ICT related jobs were identified from LinkedIn. Classification of ICT jobs from LinkedIn achieved weighted average f1-score of 0.98 which indicates high reliability in distinguishing ICT jobs from non-ICT jobs. For ICT jobs, the model attained a precision of 0.98, recall of 0.86, and F1-score of 0.91. The high precision suggests that most advertisements classified as ICT jobs were correctly identified, while the recall value indicates that a small proportion of ICT job advertisements may not have been captured and retained for analysis. This may result in slight underrepresentation of some skill requirements in the final dataset. The results of classification of ICT and non-ICT jobs scraped from LinkedIn are indicated in Table 2.

Job advertisements amounting to 1300 were scraped from Mabumbe and Ajirazyako job portals. Preprocessing of job advertisements from the three sources resulted in a dataset with 15,653 unique entries. Table 3 summarizes the number of job advertisements obtained from each portal and the resulting dataset used for association rule analysis.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

The final Boolean transaction matrix comprised 15,653 job advertisements and 51 distinct ICT skills extracted from job descriptions. Each row represented a job advertisement, while each column represented a skill encoded as a binary attribute (1 = present, 0 = absent). This matrix served as the input for association rule mining.

Table 2. ICT and non-ICT jobs classification results

Number of training samples: 99962				
Number of testing samples: 24991				
Category	precision	recall	f1-score	support
Non-ICT jobs (0)	0.98	1.00	0.99	21264
ICT jobs (1)	0.98	0.86	0.91	3727
accuracy			0.98	24991
macro avg	0.98	0.93	0.95	24991
weighted avg	0.98	0.98	0.98	24991

Table 3. Summary of job advertisements collected and processed from three different job portals

Processing Stage	**LinkedIn**	**Mabumbe**	**AjiraYako**	**Total**
Raw job advertisements collected	124,953	904	396	126,253
ICT job identification method	Supervised classifier	Portal category	Portal category	–
ICT job advertisements retained	16,455	904	396	17,755
Aggregated ICT job advertisements	–	–	–	17,755
Cross-portal duplicate records removed	–	–	–	2,102
Final unique ICT job advertisements used for analysis	–	–	–	15,653

3.2. Frequent ICT Job Roles and their Linked Skills

1) Frequent ICT Job Roles

Results show that roles related to software engineering dominate the observed dataset. Software Engineer (n = 906; 5.79%) and Senior Software Engineer (n = 892; 5.7%) recorded the highest number of advertisements among the top ICT job roles. Data Science related

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

jobs were also highly represented in the dataset with different titles including Data analyst (n = 380; 2.43%), Senior Data engineer (n = 339; 2.17%), Data Engineer (n = 265; 1.70%) and Data Scientist (n = 226; 1.45%). These findings reflect the increasing demand for data-driven systems and analytics capabilities within ICT sector. Figure 4 presents the distribution of the top 25 ICT job roles. Collectively, these roles account for 5,691 job advertisements, representing 36.4% of the 15,653 unique ICT job advertisements in the final dataset. The lengths of the bars indicate the percentage share of each role relative to the entire dataset, while the embedded labels show the corresponding absolute number of advertisements.

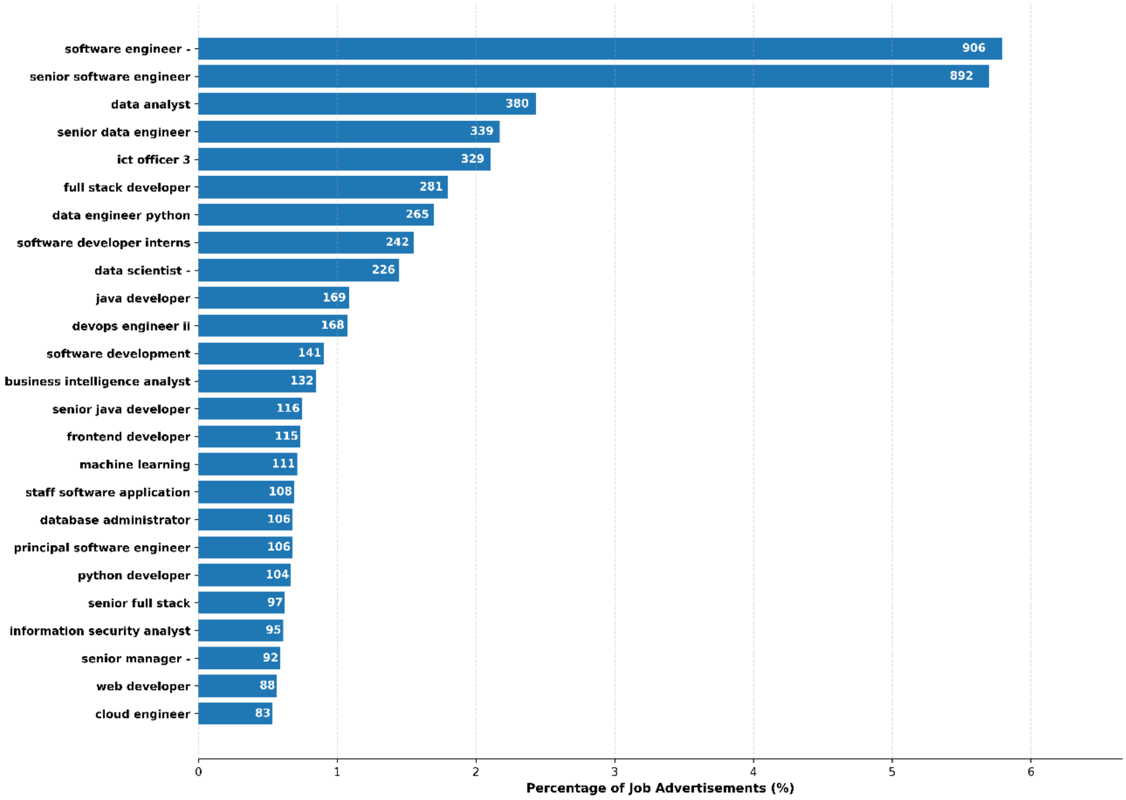

Figure 4. Distribution of 25 most frequent ICT job roles by frequency and percentage of advertisements published between January 2020 and September 2025 on the selected portals

The dominance of programming and data science related jobs reflects the growing reliance of organizations on digital systems, data driven decision making and electronic services. This aligns with Skill-Biased Technological Change (SBTC) theory which suggest that technological advancements increase productivity and demand for workers with relevant skills. Similar demands have been reported in prior studies which identified

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

software development and data related roles as core of ICT employment demand ([10], [12].

2) Skills Linked to Frequent ICT Job Roles

Results show that Python, C++ and Java were the most frequently mentioned individual skills across job descriptions associated with the 25 most frequent ICT job roles. Each of these skills appeared in more than 30% of the 5,691 job advertisements analyzed, as shown in Figure 5. These findings indicate strong demand for competencies related to software development, data processing and enterprise systems.

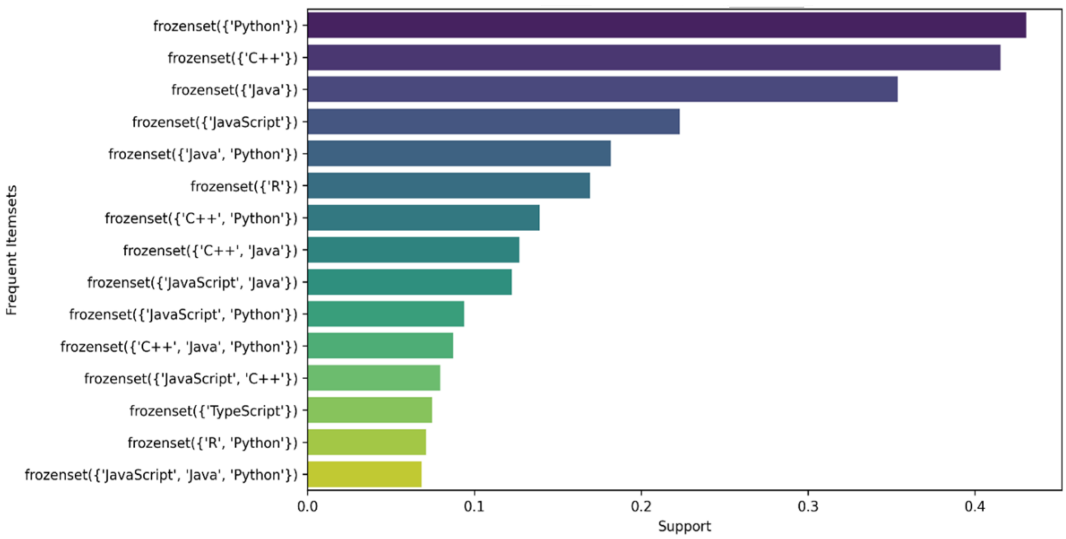

Figure 5. Most frequent skills in job descriptions

Figure 5 further demonstrates that employers frequently seek combinations of skills rather than isolated competencies. For example, the combination of Java and Python emerged as the most frequent co-occurring skill set, followed by combinations of C++ and Python and C++ and Java. While individual skill frequencies identify the technologies most requested by employers, they do not reveal how these skills are bundled within job roles. The frequent itemsets therefore provide additional insights into complementary competencies that may be overlooked in frequency-based analyses. This highlights the added value of association rule mining in uncovering skill interdependencies.

3.2.1. Co-occurring Skill Sets

1) Skill Sets in Leading Job Roles

The top 25 ICT job roles were predominantly programming oriented. Association rule analysis of the 5,691 job advertisements representing these leading roles revealed a strong relationship between C# and C++, with all advertisements mentioning C# also requiring C++ (confidence = 1.00; lift = 2.41). However, the relatively low support value (0.0096) corresponded to approximately 54 job advertisements suggests that the combination occurred in a relatively small proportion of the dataset. In addition, the association between Rust and C++ was observed in approximately 81 advertisements (support = 0.0142) with high confidence (0.82) and strong positive lift (1.97), which indicates a less frequent but meaningful co-occurrence. Kotlin and Java also demonstrated a strong association with confidence of 0.76, lift of 2.15 and co-occurred in 12 advertisements (support = 0.0021), which suggests that Kotlin related jobs frequently require Java skills despite the relatively small number of occurrences. The results also revealed that skills used in different domains can exist together as job requirements. For example, Java and Python demonstrated the highest support among the reported rules (0.1818), corresponding to approximately 1,035 job advertisements in which both skills were specified. This indicates that the combination occurred in a substantial proportion of the analysed job advertisements relative to other reported associations and may therefore reflect broad-demand competencies required across a wider range of ICT roles. Their confidence values (0.51 and 0.42) and lift of 1.19 further suggest a positive association beyond random co-occurrence. The Apriori algorithm generated 263 frequent itemsets and 355 association rules from extracted ICT skills. Table 4 presents the subset of rules selected for interpretation based on the predefined filtering criteria.

Table 4. Association rules showing co-occurrence of skills in ICT job descriptions

	Antecedents	Consequents	Support	Support Count	Confidence	Lift
0	(C#)	(C++)	0.009568	54	1.000000	2.408520
1	(C#)	(Java)	0.005219	30	0.545455	1.542027
2	(C#)	(JavaScript)	0.005219	30	0.545455	2.443121
3	(C#)	(Python)	0.004059	23	0.424242	0.985328
4	(C#)	(TypeScript)	0.003189	18	0.333333	4.456078
5	(C++)	(Java)	0.126993	723	0.305866	0.864698

	Antecedents	Consequents	Support	Support Count	Confidence	Lift
6	(Java)	(C++)	0.126993	723	0.359016	0.864698
7	(JavaScript)	(C++)	0.079443	452	0.355844	0.857058
8	(Kotlin)	(C++)	0.011308	64	0.423913	1.021003
9	(PHP)	(C++)	0.027254	155	0.435185	1.048152
10	(Perl)	(C++)	0.010728	61	0.432033	1.036224
11	(C++)	(Python)	0.139171	792	0.335196	0.778511
12	(Python)	(C++)	0.139171	792	0.323232	0.778511
13	(Ruby)	(C++)	0.013337	76	0.351145	0.845742
14	(Rust)	(C++)	0.014207	81	0.816667	1.969589
15	(Swift)	(C++)	0.015077	86	0.448276	1.079681
16	(TypeScript)	(C++)	0.023485	134	0.313953	0.756163
17	(Dart)	(Java)	0.002320	13	0.666667	1.884699
18	(JavaScript)	(Java)	0.122354	696	0.548052	1.549376
19	(Java)	(JavaScript)	0.122354	696	0.345902	1.549376
20	(Kotlin)	(Java)	0.002096	12	0.760870	2.151016
21	(PHP)	(Java)	0.042621	243	0.680556	1.392498
22	(Perl)	(Java)	0.019158	109	0.465116	1.193643
23	(Java)	(Python)	0.181792	1,035	0.513934	1.193643
24	(Python)	(Java)	0.181792	1,035	0.422222	1.193643

The results provide evidence that ICT jobs increasingly require multi-skilled workers. Several of the identified associations demonstrated high confidence values and lift values greater than one (>1), which indicate that the skills co-occurred more frequently than expected by chance. For example, the associations between C# and C++, Rust and C++, as well as Kotlin and Java suggest complementary technical competencies that are commonly required together in specific software development roles. These findings are consistent with literature which argues that in modern roles, ICT personnel are expected to integrate multiple technologies within the same project [5], [6]. The observed associations among web, software and mobile development skills highlight the complementary nature of backend, frontend, and mobile technologies. For example, the strong association between Kotlin and Java reflects the close relationship between Android application development and enterprise software technologies, while the Rust

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

and C++ association suggests demand for performance-oriented programming skills in systems development. These combinations indicate that employers advertising through the selected portals increasingly value professionals who can contribute across different stages of software development, including implementation, integration, maintenance and optimization. The findings align with current job advertisements which emphasize cross-platform development capabilities rather than narrow technical roles. However, rules with lift values below one (<1), such as C++ with Java, Java with C++, and C++ with Python, indicate weak associations despite their observed co-occurrence. These rules suggest that the corresponding skills may occur together in some job advertisements but are not particularly dependent on each other across the dataset.

The results also revealed associations between skills with different application domains. In particular, the co-occurrence between Java and Python suggests that employers frequently seek both competencies within the same vacancies. Although the lift value was only moderately above one, the relatively high support indicates that the combination appeared in a substantial proportion of job advertisements. This finding aligns with previous job advertisement mining studies that report growing convergence between software engineering, data analytics, cloud computing and artificial intelligence capabilities within modern ICT occupations. The results therefore extend existing labour market analyses beyond identification of individual high-demand skills by revealing how skills are required together in practice.

2) Skill Sets Across ICT Domains

While programming language associations were dominating within the top 25 job roles, the broader analysis of all 15,653 ICT job advertisements revealed co-occurring skills spanning multiple ICT domains. The identified associations involved combinations of skills from networking, web development, databases, data science, programming and DevOps. Although the support values were low (0.001), representing approximately 16 job advertisements for each reported association, the high confidence (0.57–1.00) and lift values (86.37–151.15) indicate strong relationships beyond random co-occurrence whenever they occurred. However, because these rules were derived from a relatively small number of advertisements, the confidence and lift values may be influenced by rare-rule inflation. These patterns should therefore be interpreted as niche skill requirements associated with specialized or emerging ICT roles rather than broadly

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

prevalent labour market demands. The resulting cross-domain skill associations identified from ICT job dataset are presented in Table 5.

Table 5. Co-occurring skills across ICT domains

Sno.	Co-occurring Skill Set	Domain Identification	support	confidence	lift
1	CSS + Firewall	Networking | Web	0.001	1	151.15
2	Kotlin + Firewall	Programming | Networking	0.001	1	151.15
3	Rust + Firewall	Programming | Networking	0.001	1	151.15
4	TypeScript + Firewall	Programming | Networking	0.001	1	151.15
5	Node.js + TCP/IP	Networking | Web	0.001	1	151.15
6	Firewall + SQL Server	Networking | Databases	0.001	0.83	125.96
7	Firewall + HTML	Networking | Web	0.001	0.8	120.92
8	Angular + TCP/IP	Networking | Web	0.001	0.8	120.92
9	TensorFlow + Ruby	Data Science | Programming	0.001	0.75	113.36
10	TensorFlow + TCP/IP	Data Science | Networking	0.001	0.75	113.36
11	TCP/IP + Deep Learning	Data Science | Networking	0.001	0.75	113.36
12	Node.js + Firewall	Networking | Web	0.001	0.71	107.97
13	Angular + Firewall	Networking | Web	0.001	0.71	107.97
14	Routing + Node.js	Networking | Web	0.001	0.67	100.77
15	Firewall + React	Networking | Web	0.001	0.67	100.77
16	Firewall + PostgreSQL	Networking | Databases	0.001	0.67	100.77
17	Perl + Node.js	Programming | Web	0.001	0.6	90.69
18	Rust + Ansible	DevOps | Programming	0.001	0.6	90.69
19	TensorFlow + PHP	Data Science | Programming	0.001	1	90.69
20	C++ + Deep Learning	Programming | Data Science	0.001	0.6	90.69
21	C# + Deep Learning	Programming | Data Science	0.001	1	90.69
22	Swift + TensorFlow	Programming | Data Science	0.001	1	90.69
23	TCP/IP + MongoDB	Networking | Databases	0.001	0.6	90.69
24	Routing + Kotlin	Programming | Networking	0.001	0.6	90.69
25	Kotlin + Data Analysis	Programming | Data Science	0.001	0.6	90.69
26	Rust + TCP/IP	Programming | Networking	0.001	0.6	90.69
27	Node.js + Deep Learning	Data Science | Web	0.001	0.6	90.69
28	Rust + MongoDB	Programming | Databases	0.001	0.57	86.37
29	Rust + SQL Server	Programming | Databases	0.001	0.57	86.37

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

The broader analysis identified associations involving networking, programming, data science, web, databases and DevOps skills. This indicates that complementary skill requirements are evident across diverse ICT domains available in the dataset. These findings suggest that employers advertising through the selected portals increasingly seek combinations of complementary skills across traditional ICT specializations.

3.2.2. Visualization of Skill Associations

1) Visualization of Python Skill Associations

The results show that Python has strong associations with Java, C++, JavaScript, PHP and Scala, as they frequently appear together in job descriptions. This association implies that Python is commonly used alongside these languages in practical work environments. The associations with JavaScript and PHP in data science projects involving web development, as well as with Java and C++ in enterprise system projects, highlight the need for professionals to possess skills across multiple technologies. Presence of associations with R, Kotlin, and Swift show the relevance of Python in data science and mobile development for both android and iOS operating systems. Figure 6 details the result of association of Python with other skills in job advertisements.

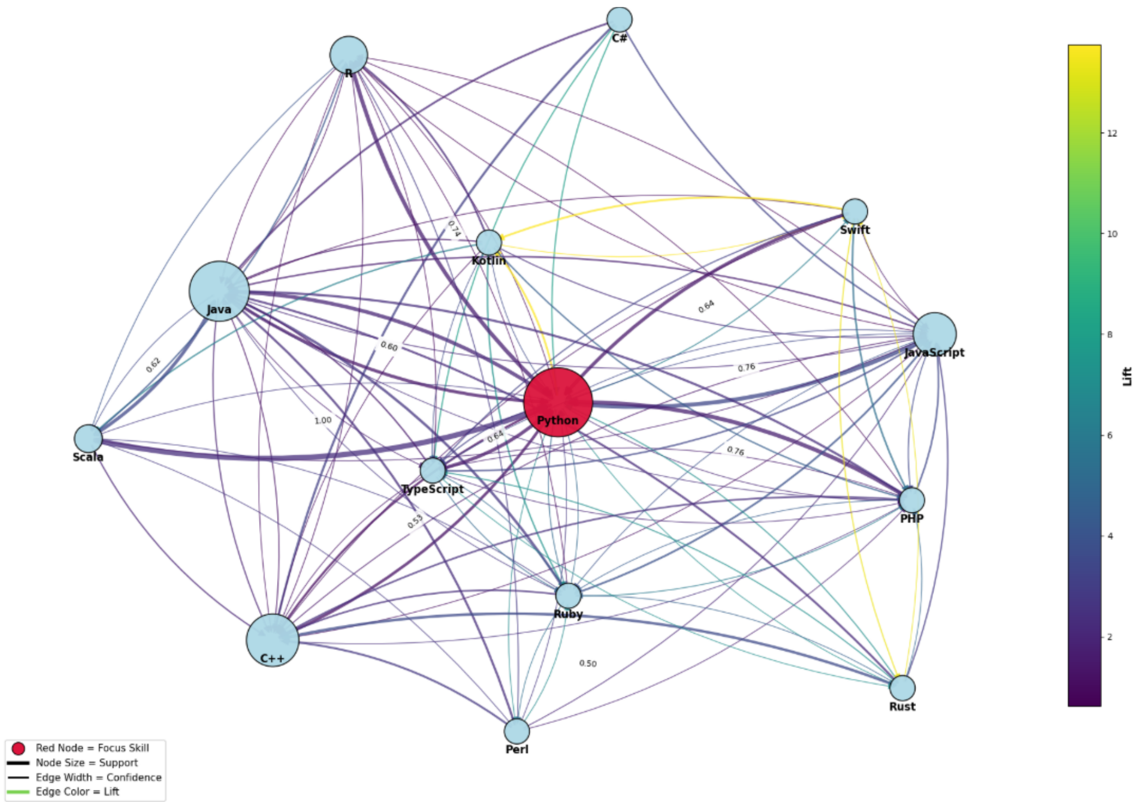

Figure 6. Visualization of association rules involving Python

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

2) Visualization of C++ Skill Associations

The visualization results show that C++ frequently appear together in job descriptions with Java, JavaScript, Python, PHP and Scala. Among them, JavaScript and PHP show stronger associations with C++ through thickness of their connecting arrows as indicated in figure 7.

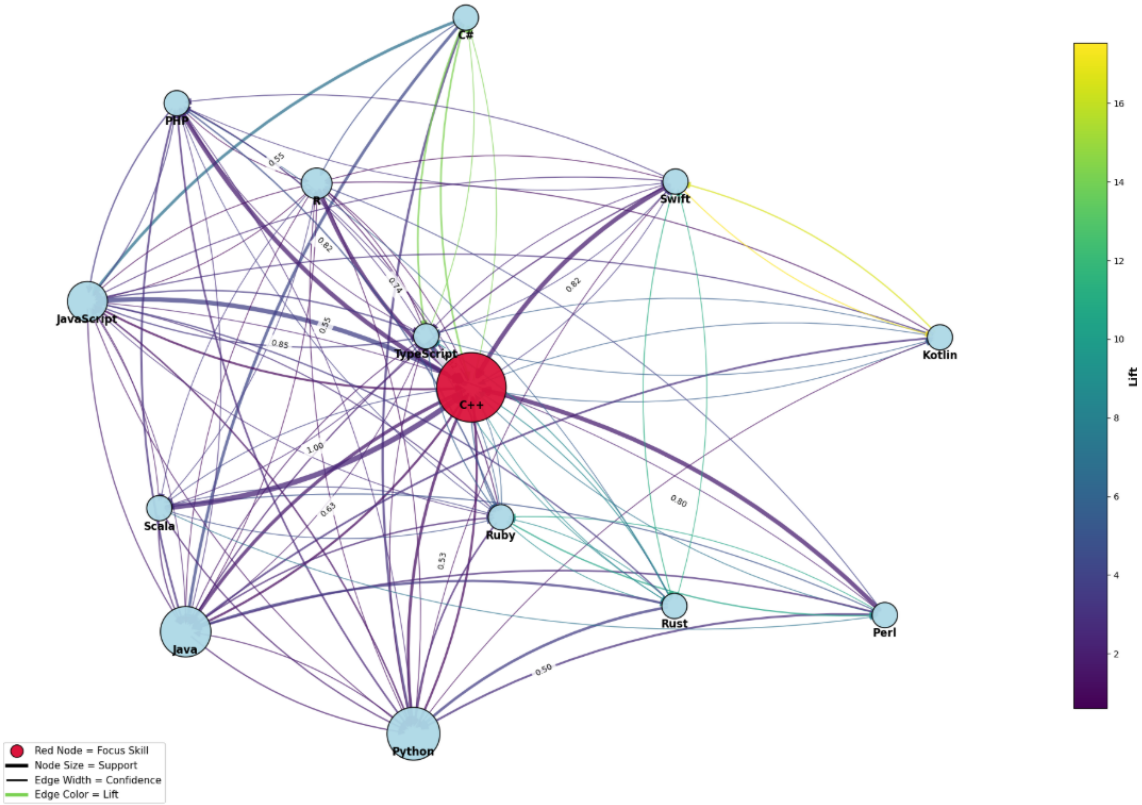

Figure 7. Visualization of association rules involving C++

3) Visualization of Java Skill Associations

The results as indicated in figure 8 reveals that Java has strong association with JavaScript and Python. This shows that there is high probability that these skills appear together as requirements in a particular job advertisement. The connection between Java and Dart suggest requirement for flexibility of skills possessed to work in different types of mobile applications development projects. The results further reveal Java's wide range of applications including web development, enterprise systems and big data systems through its association with PHP, TypeScript and Scala.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

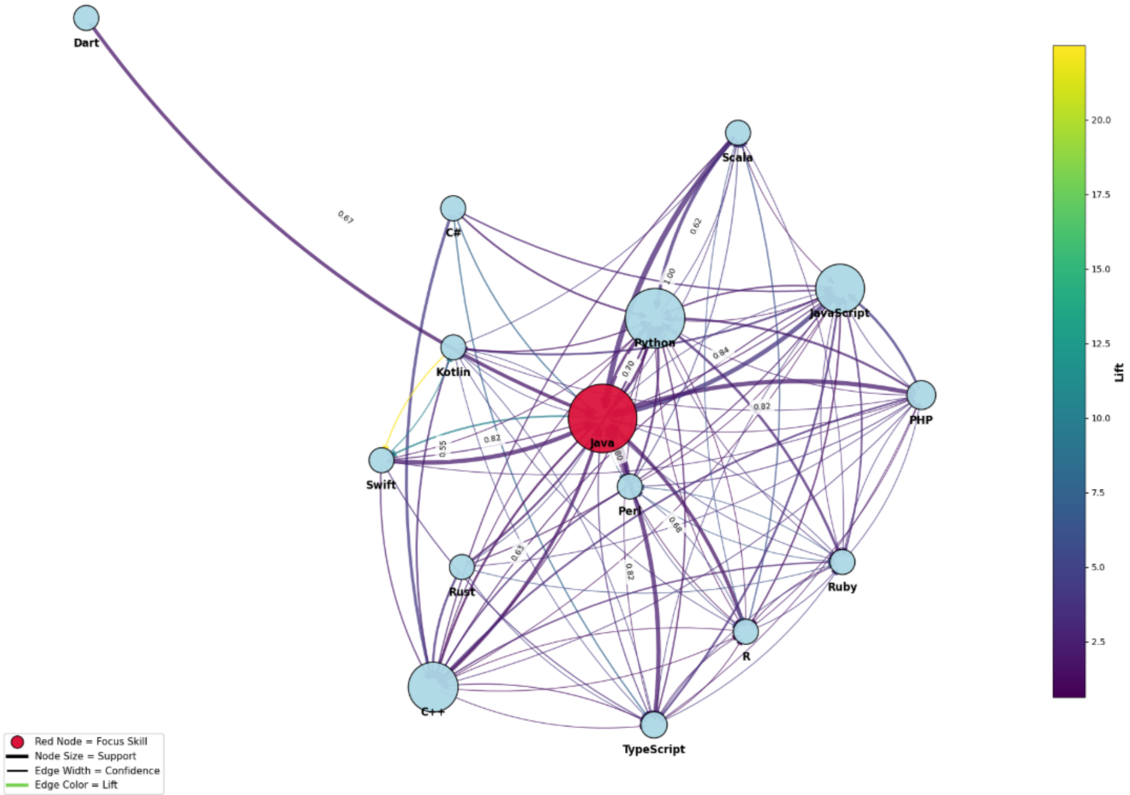

Figure 8. Visualization of association rules involving Java

4) Combined Visualization of Association Rules

The visualization of association rules shows that TypeScript, Ruby, and Kotlin appeared among the most frequent consequent skills across multiple rule combinations. This suggests that these technologies are commonly required alongside other programming skills in ICT job advertisements. However, the strength of these associations should be interpreted using the corresponding support, confidence and lift values rather than the visualization alone. The results generally indicate increasing demand for combinations of complementary programming skills across web, software, and mobile development roles. In addition, the observed association between Scala and R suggests a relationship between technologies commonly used in data-oriented applications. The combined visualization for association rules among skills is presented in figure 9.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

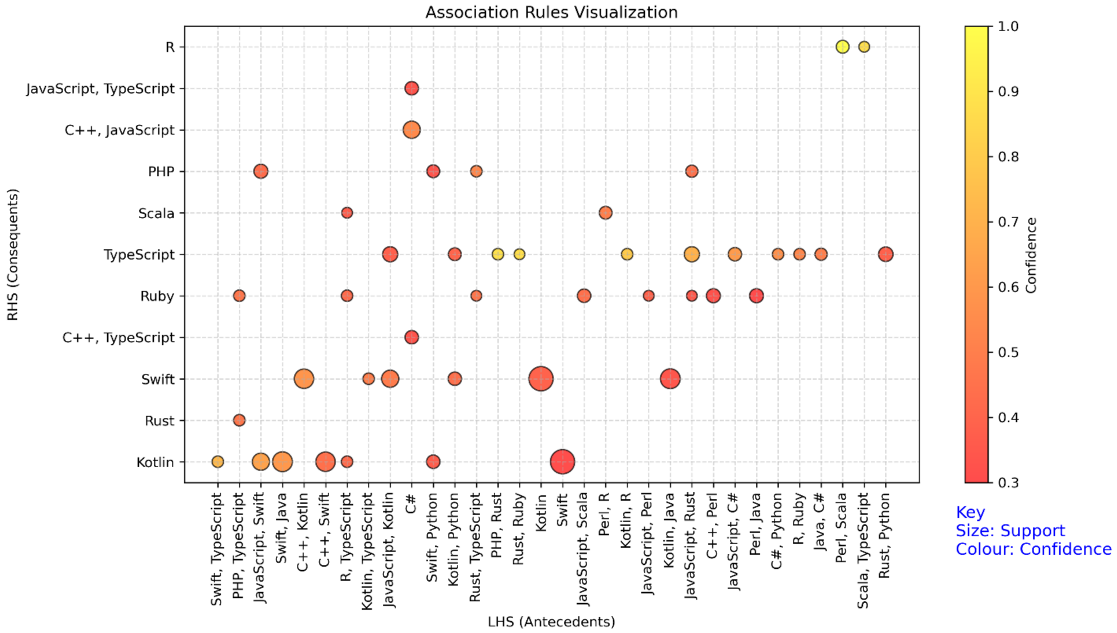

Figure 9. Combined visualization of association rules

3.3. Discussion

The co-occurrence analysis provided insights beyond those obtained from individual skill frequency analysis. While frequency analysis identifies the most demanded skills in the labour market, it does not reveal how those skills are combined within specific job roles. For example, a skill such as Java may appear frequently across job advertisements, which indicates high demand, but association rule mining further reveals that Java is commonly required together with skills such as Python, JavaScript, and Kotlin. In addition, the broader analysis provided insights on complementing skills across diverse ICT domains in the dataset. The identified associations should be interpreted as co-occurrence patterns within job advertisements rather than direct evidence of project-level technology integration or causal relationships between skills.

The identified co-occurring skill sets have important implications for curricula design. The observed associations suggest that employers advertising through the selected portals require combinations of complementary skills within the same job roles. By uncovering relationships among skills, association rule mining provides insights that may be used to design curricula that emphasize complementary skill sets frequently demanded together in the ICT labour market. Curricula developers may consider designing bundled learning modules, integrated projects and interdisciplinary training experiences that reflect these combinations of skills. High-support associations may inform core curriculum modules because they represent broadly demanded competency combinations across ICT occupations, whereas low-support but high-lift associations may inform specialist electives, professional certification programmes and targeted upskilling initiatives aimed at addressing specialized or emerging workforce needs. This approach can better prepare graduates for workplace requirements by exposing them to the skills set that are commonly demanded together in the labour market.

The findings also provide valuable labour market intelligence for policymakers and employers involved in workforce development initiatives. Co-occurring skill patterns offer insights into the combinations of competencies required in emerging ICT occupations and can therefore support evidence-based planning of reskilling and upskilling programmes. Rather than focusing on individual technologies in isolation, training initiatives can prioritize skill bundles that reflect actual employer requirements. Employers may also use these insights to design targeted professional development programmes, while policymakers can use the identified skill combinations to guide investments in digital skills development, workforce readiness programmes and ICT capacity building strategies. These interventions can contribute to narrowing skills gaps and improving alignment between education outcomes and labour market demands.

However, the findings are limited to the selected job portals and publication period between January 2020 and September 2025. Differences in the types of employers and vacancies represented across LinkedIn, AjiraYako, and Mabumbe may have influenced the observed skill frequencies and associations. As reported in previous studies, differences in portal coverage, employer participation and vacancy posting practices can introduce biases into labour market analyses based on online job advertisements [26]. In addition, the ICT classifier may have excluded a small number of relevant ICT vacancies, leading to slight underrepresentation of certain skills and co-occurrence patterns in the final dataset, while the exact-hash deduplication approach may not have detected duplicate advertisements with minor variations in wording or metadata. Skill extraction relied on a predefined dictionary and was limited in handling multi-word skill expressions and variations in equivalent job role titles. Furthermore, the identified skill associations were derived solely from job advertisements and were not validated directly with employers

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

or curriculum experts. Addressing these limitations in future studies would strengthen the reliability and applicability of findings derived from online job advertisements.

4. CONCLUSION

This study applied association rule mining to ICT job advertisements collected from multiple online portals to uncover co-occurring skill requirements within ICT job roles. Moving beyond conventional labour market analyses that focus on individual skill frequencies, the study identified skill interdependencies that reveal how employers combine competencies within the same vacancies. The findings highlighted recurring combinations of skills in the dataset, such as C# with C++, Rust with C++, Kotlin with Java, and Java with Python. These findings highlight growing demand for complementary and cross-domain competencies in the ICT jobs represented in the dataset. By uncovering relationships among skills rather than only identifying frequently demanded skills, the study contributes richer labour market intelligence that can support curricula alignment and workforce development initiatives aimed at preparing professionals for evolving industry requirements. The findings should, however, be interpreted considering limitations related to portal coverage, automated classifier filtering, dictionary-based skill extraction and the limited recognition of multi-word skill expressions. The identified associations should be interpreted as exploratory patterns derived from ICT job advertisements published on the selected portals between January 2020 and September 2025. To address these limitations, future research should extend to more diverse job portals, examine changes in co-occurring skill demands over time, explore advanced NLP techniques for multi-word skills recognition and semantic normalization of job roles, and validate the identified skill associations with industry and curriculum experts.

REFERENCES

[1] M. Tavakoli, A. Faraji, J. Vrolijk, M. Molavi, S. T. Mol, and G. Kismihók, "An AI-based open recommender system for personalized labor market driven education," *Adv. Eng. Inform.*, vol. 52, Art. no. 101508, 2022, doi: 10.1016/j.aei.2021.101508.

[2] T. Ndofirepi and R. Steyn, "Technological entrepreneurship readiness: An analysis across BRICS countries," *J. BRICS Stud.*, vol. 2, no. 2, pp. 19–37, 2024.

[3] L. Li, "Reskilling and upskilling the future-ready workforce for Industry 4.0 and beyond," *Inf. Syst. Front.*, vol. 26, no. 5, pp. 1697–1712, 2022, doi: 10.1007/s10796-022-10308-y.

[4] M. A. Hazrat, N. M. S. Hassan, A. A. Chowdhury, M. G. Rasul, and B. A. Taylor, "Developing a skilled workforce for future industry demand: The potential of digital twin-based teaching and learning practices in engineering education," *Sustainability*, vol. 15, no. 23, Art. no. 16433, 2023, doi: 10.3390/su152316433.

[5] T. Ciarli, M. Kenney, S. Massini, and L. Piscitello, "Digital technologies, innovation, and skills: Emerging trajectories and challenges," *Res. Policy*, vol. 50, no. 7, Art. no. 104289, 2021, doi: 10.1016/j.respol.2021.104289.

[6] A. Jaiswal, C. J. Arun, and A. Varma, "Rebooting employees: Upskilling for artificial intelligence in multinational corporations," in *Artificial Intelligence and International HRM*. London, U.K.: Routledge, 2023, pp. 114–143, doi: 10.4324/9781003377085-5.

[7] T. Bolli and F. Pusterla, "Is technological change really skills-biased? Firm-level evidence of the complementarities between ICT and workers' education," *Econ. Innov. New Technol.*, vol. 32, no. 1, pp. 69–91, 2023, doi: 10.1080/10438599.2020.1871270.

[8] D. Nyangoma, E. M. Adaga, N. J. Sam-Bulya, and G. O. Achumie, "Market trend analysis as a strategic tool for workforce development programs: A data-driven conceptual model," *Int. J. Adv. Multidiscip. Res. Stud.*, vol. 5, no. 2, pp. 1239–1246, 2025, doi: 10.62225/2583049x.2025.5.2.3972.

[9] A. Verma, K. Lamsal, and P. Verma, "An investigation of skill requirements in artificial intelligence and machine learning job advertisements," *Ind. High. Educ.*, vol. 36, no. 1, pp. 63–73, 2022, doi: 10.1177/0950422221990990.

[10] S. Fareri, G. Fantoni, F. Chiarello, E. Coli, and A. Binda, "Estimating Industry 4.0 impact on job profiles and skills using text mining," *Comput. Ind.*, vol. 118, Art. no. 103222, 2020, doi: 10.1016/j.compind.2020.103222.

[11] H. Hassani, C. Beneki, S. Unger, M. T. Mazinani, and M. R. Yeganegi, "Text mining in big data analytics," *Big Data Cogn. Comput.*, vol. 4, no. 1, pp. 1–34, 2020, doi: 10.3390/bdcc4010001.

[12] G. Li, C. Yuan, S. Kamarthi, M. Moghaddam, and X. Jin, "Data science skills and domain knowledge requirements in the manufacturing industry: A gap analysis," *J. Manuf. Syst.*, vol. 60, pp. 692–706, 2021, doi: 10.1016/j.jmsy.2021.07.007.

[13] M. Johnson *et al.*, "Impact of big data and artificial intelligence on industry: Developing a workforce roadmap for a data-driven economy," *Glob. J. Flex. Syst. Manag.*, vol. 22, no. 3, pp. 197–217, 2021, doi: 10.1007/s40171-021-00272-y.

[14] L. Sakurada, C. A. S. Geraldes, F. P. Fernandes, J. Pontes, and P. Leitão, "Analysis of new job profiles for the factory of the future," in *Service Oriented, Holonic and Multi-Agent Manufacturing Systems for Industry of the Future*, T. Borangiu, D. Trentesaux, P. Leitão, O. Cardin, and S. Lamouri, Eds., *Studies in Computational Intelligence*, vol. 952. Cham, Switzerland: Springer, 2021, pp. 262–273, doi: 10.1007/978-3-030-69373-2_18.

[15] A. Ternikov, "Soft and hard skills identification: Insights from IT job advertisements in the CIS region," *PeerJ Comput. Sci.*, vol. 8, Art. no. e946, 2022, doi: 10.7717/peerj-cs.946.

[16] D. Pichler and R. Stehrer, "Breaking through the digital ceiling: ICT skills and labour market opportunities," *WIIW Working Paper*, no. 193, 2021.

[17] D. T. K. Ng, M. Lee, R. J. Y. Tan, X. Hu, J. S. Downie, and S. K. W. Chu, "A review of AI teaching and learning from 2000 to 2020," *Educ. Inf. Technol.*, vol. 28, no. 7, pp. 8445–8501, 2023, doi: 10.1007/s10639-022-11491-w.

[18] H. Taherdoost, "An overview of trends in information systems: Emerging technologies that transform the information technology industry," *Cloud Comput. Data Sci.*, vol. 4, no. 1, pp. 1–16, 2023, doi: 10.37256/ccds.4120231653.

[19] A. Verma, P. Purohit, T. Thornton, and K. Lamsal, "An examination of skill requirements for augmented reality and virtual reality job advertisements," *Ind. High. Educ.*, vol. 37, no. 1, pp. 46–57, 2023, doi: 10.1177/09504222221109104.

[20] D. Acemoglu, D. Autor, J. Hazell, and P. Restrepo, "Artificial intelligence and jobs: Evidence from online vacancies," *J. Labor Econ.*, vol. 40, no. S1, pp. S293–S340, 2022, doi: 10.1086/718327.

[21] I. Khaouja, I. Kassou, and M. Ghogho, "A survey on skill identification from online job ads," *IEEE Access*, vol. 9, pp. 118134–118153, 2021, doi: 10.1109/access.2021.3106120.

[22] N. Kumar, M. Gupta, D. Sharma, and I. Ofori, "Technical job recommendation system using APIs and web crawling," *Comput. Intell. Neurosci.*, vol. 2022, Art. no. 7797548, 2022, doi: 10.1155/2022/7797548.

[23] A. Kumar, K. Chauhan, and J. K. Grewal, "Web scraping job portals," in *Advancements in Communication and Systems*, A. K. Tripathi and V. Shrivastava, Eds. India: SCRS, 2024, pp. 291–303, doi: 10.56155/978-81-955020-7-3-25.

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

[24] J. Napierała, “Enhancing taxonomy-based extraction: Leveraging information from online community platforms for digital skills demand identification in job ads” *Stat. J. IAOS*, vol. 40, no. 3, pp. 591–602, 2024, doi: 10.3233/SJI-230110.

[25] J. A. Diaz-Garcia, M. D. Ruiz, and M. J. Martin-Bautista, "A survey on the use of association rules mining techniques in textual social media," *Artif. Intell. Rev.*, vol. 56, no. 2, pp. 1175–1200, 2023, doi: 10.1007/s10462-022-10196-3.

[26] M. Sostero and E. Fernández-Macías, "The professional lens: What online job advertisements can say about occupational task profiles," European Commission Joint Research Centre, Seville, Spain, Tech. Rep. JRC125917, JRC Working Papers Series on Labour, Education and Technology, no. 2021/13, 2021.