Clustering AI Job Roles Using PCA and K-Means Based on Skill Profiles and Automation Risk

Untung Rahardja[1,*], Qurotul Aini[2]

[1,2] Digital Business Study Program, Universitas Raharja, Tangerang, Indonesia

ABSTRACT

The rapid expansion of artificial intelligence (AI) technologies has significantly transformed the job market, leading to emerging demands for hybrid skillsets and raising concerns over automation-induced job displacement. This study aims to identify meaningful patterns within the AI job landscape by clustering job roles based on required skills and automation risk. Using a dataset of 500 AI-related job entries, we applied Principal Component Analysis (PCA) to reduce the dimensionality of the skill space, followed by K-Means clustering to group similar roles. The analysis revealed four distinct clusters with notable differences in salary, skill emphasis, and automation vulnerability. Further examination showed that roles emphasizing technical competencies—such as Python, Machine Learning, and Data Analysis—tend to fall into higher-paying clusters with lower automation risk. In contrast, jobs requiring predominantly soft skills—such as Communication, Marketing, and Sales—are more susceptible to automation and are generally lower-paid. Correlation analysis confirmed these trends, with technical skills showing strong negative correlations with automation risk, while non-technical skills demonstrated positive correlations. These findings underscore the growing importance of technical proficiency in securing resilient careers in the AI sector, offering strategic insights for education, workforce development, and policy formulation.

Keywords Artificial Intelligence, Job Market, K-Means Clustering, Principal Component Analysis, Automation Risk, Skill Analysis, Workforce Resilience

Submitted 2 October 2025
Accepted 16 November 2025
Published 1 December 2025

*Corresponding author
Untung Rahardja,
untung@raharja.info

Additional Information and Declarations can be found on page 327

DOI: 10.63913/ail.v1i4.44

Introduction

The proliferation of Artificial Intelligence (AI) technologies has led to a profound transformation in work, affecting industries ranging from healthcare and finance to manufacturing, marketing, and education [1]. As intelligent systems increasingly take over tasks that were once exclusively human, such as data analysis, image recognition, natural language understanding, and decision-making, organizations are restructuring their operations and redefining job roles to remain competitive in a digitally augmented economy [2]. While AI enables greater efficiency and innovation, it simultaneously introduces systemic risks to employment by rendering certain skills obsolete and elevating the automation susceptibility of particular job functions [3].

In this shifting landscape, AI-related occupations have emerged as a focal point of both opportunity and disruption. On one hand, the demand for professionals with expertise in machine learning, data engineering, natural language processing, and AI ethics is rapidly increasing [4]. On the other hand, roles that depend on routine decision-making or standardized communication, such as administrative support, sales coordination, or basic marketing, face growing risks of being automated by intelligent systems. This duality raises critical questions: Which skills are most associated with resilience to automation? How are AI job roles structured in terms of their technical depth and risk profile? Can we identify meaningful patterns or clusters in the AI labor market that inform

How to cite this article: U.Rahardja and Q.Aini, "Clustering AI Job Roles Using PCA and K-Means Based on Skill Profiles and Automation Risk," *Artif. Intell. Learn.*, vol. 1, no. 4, pp. 315-328, 2025.

education and workforce development policies?

Despite the urgency of these questions, much of the existing literature tends to focus on macro-level trends or theoretical models of job displacement without analyzing actual job market data. Previous works have predicted automation probabilities across broad occupational categories, but they do not account for intra-industry variance or skill-specific granularity. In response to this gap, the present study adopts a data-driven clustering approach to analyze a curated dataset of 500 AI-related job postings. By leveraging Principal Component Analysis (PCA) to reduce high-dimensional skill representations and applying K-Means clustering to group similar roles, we aim to uncover structural patterns that characterize the modern AI labor ecosystem.

In addition to identifying distinct clusters based on skill composition and automation risk, this study conducts a correlation analysis to determine how individual skills influence a job's likelihood of being automated. The combination of clustering and skill-risk correlation enables a nuanced understanding of how technical versus non-technical competencies shape job security in the AI era. These insights are crucial not only for academic inquiry but also for practical application. Educators can realign curricula to reflect high-demand, low-risk skill areas. Policymakers can target interventions toward vulnerable worker segments. And job seekers can make informed career choices based on market-aligned, automation-resilient competencies.

Through this analysis, we contribute to a growing body of work that seeks to map the evolving structure of the AI workforce, offering actionable intelligence for a future of work increasingly mediated by intelligent machines.

Literature Review

The evolving nature of work in the age of AI has attracted significant scholarly attention, particularly concerning how automation influences labor markets, skill demand, and job security. Early foundational work by Frey and Osborne estimated that up to 47% of U.S. employment was at risk of automation, using task-based modeling and expert assessment [5]. Their study prompted a wave of research into how AI and machine learning technologies might reshape workforce dynamics. Chui et al expanded on this by emphasizing that while certain tasks—not entire jobs—are automatable, roles with repetitive, rule-based activities remain the most vulnerable [6].

Recent research has increasingly focused on skills-based analysis to predict job displacement and resilience. Nedelkoska and Quintini explored skill requirements and found that jobs demanding higher cognitive and technical skills were less automatable [7]. Similarly, Brynjolfsson and McAfee argued that digital technologies are polarizing labor markets, favoring those with high technical skills while displacing middle-skill roles [8]. The OECD supported this view, suggesting that adaptability and lifelong learning are critical to maintaining employability [9].

To identify patterns in skill demand, clustering and dimensionality reduction techniques such as K-Means, DBSCAN, and PCA have been employed in several works. Liu et al used K-Means clustering on LinkedIn job data to group emerging AI-related occupations, revealing divergent skill clusters within AI subfields [10]. Ali and Ibrahim applied PCA and hierarchical clustering to group digital labor profiles, demonstrating that clustering could highlight unique job segments that are not evident through traditional classification [11].

Studies have also examined the correlation between skills and automation risk. Webb introduced a method for linking job descriptions to AI capabilities using natural language processing, finding that roles involving perception and social intelligence are less likely to be automated [12]. Acemoglu and Restrepo examined the displacement effect of robotics and found that regions with higher robot penetration experienced reduced employment in routine-intensive occupations [13]. Arntz et al emphasized the importance of controlling for workplace heterogeneity, showing that not all similarly titled jobs have the same automation risk, especially when skill content is considered [14].

Several recent papers have moved toward visual skill profiling and labor market forecasting. Deming and Noray analyzed longitudinal skill demand and showed how demand for AI-related skills evolves across sectors [15]. Atalay et al used network analysis to map skill interdependencies and transition pathways in the labor market [16]. Meanwhile, Bessen argued that automation does not always reduce employment but may instead increase output and shift skill requirements within industries [17].

From a methodological standpoint, combining PCA and clustering has proven effective for mapping high-dimensional skill data. Zhang et al clustered cybersecurity jobs using multivariate techniques to inform educational alignment [18]. Similarly, Reddy and Sastry used unsupervised learning to classify roles within the AI domain, demonstrating that roles with shared skill profiles often exhibit similar risk exposures and economic characteristics [19].

Despite the growing body of literature, there remains a need for research that integrates automation risk quantification, skill frequency analysis, and clustering techniques in a unified framework. This study addresses that gap by applying PCA and K-Means to empirical AI job data, mapping the skill landscape while identifying which competencies are most associated with job security or displacement.

Methods

This study employs a quantitative, data-driven approach to explore the structure of AI-related job roles using clustering and correlation analysis. The methodology consists of four key stages: (1) data preprocessing and transformation, (2) dimensionality reduction using Principal Component Analysis (PCA), (3) clustering using the K-Means algorithm, and (4) correlation analysis to examine the relationship between individual skills and automation risk.

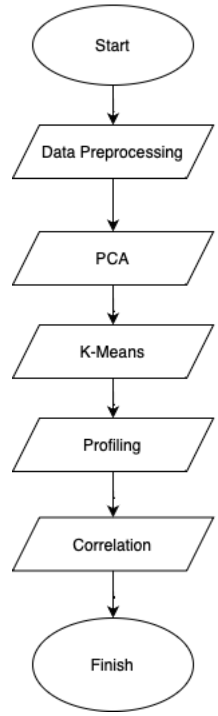

Figure 1 Research Method Flowchart

The dataset comprises 500 AI-related job postings, each containing attributes such as job title, industry, company size, location, required skills, salary (in USD), AI adoption level, automation risk level, and job growth projection. To prepare the dataset for analysis, the Required_Skills column—initially a comma-separated string—was transformed into a binary multihot matrix using multi-label binarization. This resulted in a high-dimensional feature space, where each skill is represented as a binary indicator across all job roles.

Additionally, the categorical variable Automation_Risk was encoded numerically to facilitate quantitative analysis:

$$\text{Automation Risk } (numeric) = \begin{cases} 2 & if\ Low \\ 1 & if\ Medium \\ 0 & if\ High \end{cases} \tag{1}$$

This transformation allowed automation risk to be incorporated into both clustering and correlation stages.

To simplify the high-dimensional skill representation and improve clustering performance, PCA was applied. PCA reduces the original skill-feature matrix to a lower-dimensional space by transforming it into a set of orthogonal components that explain the maximum variance in the data. For visualization and clustering purposes, only the first two principal components (PC1 and PC2) were retained:

$$Z = X \cdot W \tag{2}$$

X is the standardized skill-feature matrix, W is the PCA component weight matrix, Z is the reduced 2D representation. These components serve as inputs to the K-Means clustering algorithm.

Using the 2D PCA output, the K-Means clustering algorithm was applied to segment job roles into distinct groups based on skill similarity and automation risk. The optimal number of clusters was determined using exploratory methods (e.g., Elbow method), with k=4 chosen for interpretability and cluster stability. K-Means operates by minimizing the within-cluster sum of squared distances (WCSS) [20]:

$$\text{WSCC} = \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|^2 \tag{3}$$

C_i is the ith cluster, μ_i is the centroid of cluster i and x is a data point in cluster i. Each job role was assigned to one of the four clusters, allowing for comparative analysis across cluster characteristics.

To investigate the influence of specific skills on job automation risk, Pearson correlation coefficients were computed between each binary skill indicator and the encoded automation risk score [21]:

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\,(y_i - \bar{y})^2} \tag{4}$$

x_i is the binary presence of a skill, y_i is the automation risk score, r_{xy} indicates the strength and direction of the relationship.

The resulting coefficients were used to identify which skills are positively correlated (associated with higher automation risk) and negatively correlated (more automation-resilient). The top five in each direction were highlighted for interpretation and visualization.

Result

This section details the analytical outcomes derived from clustering AI-related job roles based on two key dimensions: required skill sets and associated automation risk. The analysis aimed to uncover latent structures in the job landscape by grouping similar roles into interpretable clusters, enabling a deeper understanding of how skill composition relates to job vulnerability in the era of AI-driven transformation.

To address the high dimensionality of the skill space—resulting from the multi-label encoding of diverse technical and non-technical competencies—we employed PCA. This dimensionality reduction technique transformed the binary skill matrix, augmented by a numerical encoding of automation risk (Low = 2, Medium = 1, High = 0), into a lower-dimensional space. Specifically, we retained the first two principal components, which together explained a substantial proportion of the variance in the data. These components provided a compact 2D representation that preserved the intrinsic structure of the original feature space.

Subsequently, this 2D projection was subjected to unsupervised clustering using the K-Means algorithm with a pre-specified number of clusters k=4, chosen based on interpretability and preliminary elbow method analysis. The clustering revealed four distinct groups of job roles, each characterized by a unique blend of skill emphasis and automation exposure. The spatial separation observed in figure 2 confirms that meaningful differences exist across AI-related positions, indicating heterogeneity in how roles are constructed and their corresponding susceptibility to technological disruption.

This clustering foundation serves as a basis for further exploration of salary distribution, skill concentration, and the systemic relationship between competencies and automation resilience, all of which are discussed in the subsequent subsections.

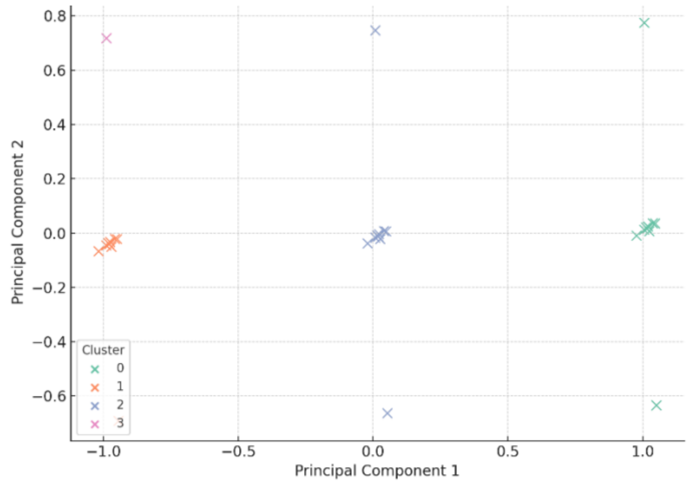

Figure 2 Clustering of AI Job Roles using PCA and K-Means

The clustering visualization presented in figure 2 demonstrates well-defined separations between groups of AI-related job roles, suggesting the presence of significant heterogeneity in both skill composition and automation vulnerability within the contemporary AI job market. Each cluster occupies a distinct region in the principal component space, implying that job roles with similar skill sets and automation profiles naturally group together. This outcome supports the hypothesis that AI jobs are not monolithic, but instead span a spectrum of technical depth, soft-skill orientation, and exposure to automation.

To characterize the resulting clusters more precisely, we computed several key descriptive statistics: the mean salary (expressed in USD), the average automation risk score (numerically encoded as Low = 2, Medium = 1, High = 0), and the total number of job entries within each cluster. These metrics provide a quantitative summary of each cluster's economic and technological characteristics. The aggregated results are presented in table 1.

Table 1 reveals meaningful contrasts across clusters. For example, Cluster 1 contains the highest average salary and the lowest automation risk score, which likely corresponds to high-skill, low-automation roles such as AI engineers or advanced data scientists. In contrast, Cluster 0—which comprises the largest number of job roles—exhibits the lowest average salary and the highest automation susceptibility, possibly representing entry-level or support-oriented positions. Such a distribution reflects stratification within the AI job ecosystem, where the combination of advanced technical skills and low automation risk commands a premium in the labor market.

Table 1 Summary Statistics per Cluster

Cluster	Avg. Salary (USD)	Avg. Automation Risk	Job Count
0	80,576	1.56	162
1	106,280	0.40	78
2	97,299	1.29	72
3	91,822	0.68	188

Cluster 1 is composed of high-paying job roles that exhibit the lowest average automation risk across all identified groups. This cluster likely represents core technical positions in the AI domain, such as machine learning engineers, AI researchers, and data scientists where advanced analytical skills and programming expertise are required. These roles are inherently resistant to automation due to their reliance on creative problem-solving, algorithm development, and the application of domain-specific knowledge. The premium salary levels observed in this cluster further reinforce the high value placed on these specialized competencies in the AI job market.

In contrast, Cluster 0 includes the largest number of job entries but is associated with lower average salaries and higher automation susceptibility. This pattern suggests that the cluster may correspond to support, administrative, or business-facing roles that involve routine or semi-structured tasks, such as marketing assistants, customer service specialists, or UX/UI designers. To further interpret the nature of each cluster, we conducted a skill frequency analysis and identified the top five most commonly required skills within each group. These results, presented in table 2, provide a clearer view of the skill profile that defines each cluster and highlight the divergent functional orientations that exist within the AI job ecosystem.

Table 2 Top 5 Skills per Cluster

Cluster	Top Skills
0	Marketing, Cybersecurity, UX/UI Design, Data Analysis, Communication
1	JavaScript, Sales, Machine Learning, Project Management, UX/UI Design
2	Python, JavaScript, Communication, Cybersecurity, Data Analysis
3	Project Management, Machine Learning, Data Analysis, Sales, Cybersecurity

The analysis of skill composition across clusters further confirms the functional divergence within the AI job market. Each cluster exhibits a distinct focus in terms of required competencies. Cluster 2, for instance, is characterized by a more technical orientation, with a notable dominance of programming languages such as Python and JavaScript, alongside analytical skills like Data Analysis. This suggests that roles in this cluster are likely to involve hands-on development, system implementation, and technical problem-solving—attributes often associated with software engineering and cybersecurity roles. In contrast, Cluster 0 places greater emphasis on soft skills and business-related capabilities, including Marketing, Communication, and UX/UI Design, reflecting job roles that are more customer-facing or involve strategic support functions, which are often more susceptible to automation due to their procedural nature.

To complement the cluster-specific insights, we extended the analysis to examine skill demand across the entire AI job dataset. This broader view allows us to identify which competencies are most consistently sought after in the AI workforce, regardless of cluster assignment. The top ten most frequently required skills, presented in table 3, highlight a balanced mix of technical (e.g., Python, Machine Learning, JavaScript) and non-technical (e.g., Communication, Project Management, Marketing) proficiencies. These findings are further visualized in figure 3, which provides a comparative bar plot illustrating the relative demand for each skill. The results underscore the multidisciplinary nature of the AI domain, where success is often contingent not only on technical expertise but also on effective communication and cross-functional collaboration.

Table 3 Top 10 Most Frequent Skills Across All AI Job Roles

Rank	Skill	Frequency
1	Communication	127
2	Data Analysis	115
3	Machine Learning	104
4	Project Management	96
5	Python	91
6	UX/UI Design	84
7	Cybersecurity	83
8	JavaScript	80
9	Marketing	76
10	Sales	74

Table 3 presents a ranked list of the ten most frequently required skills across all AI-related job roles in the dataset, offering a quantitative overview of the core competencies that dominate the current AI labor market. The list includes a diverse range of proficiencies, from foundational technical abilities like Machine Learning, Python, and JavaScript, to essential soft skills such as Communication, Project Management, and Marketing. This combination reflects the hybrid skill demands of AI jobs, where technical execution must often be paired with strategic thinking and collaboration. To complement this tabular summary, figure 3 provides a visual depiction of the same data, using a horizontal bar chart to illustrate the relative prominence and distribution of each skill. This visualization enhances interpretability by making it easier to compare the demand magnitude for each skill, reinforcing the conclusion that AI roles require an interdisciplinary blend of capabilities.

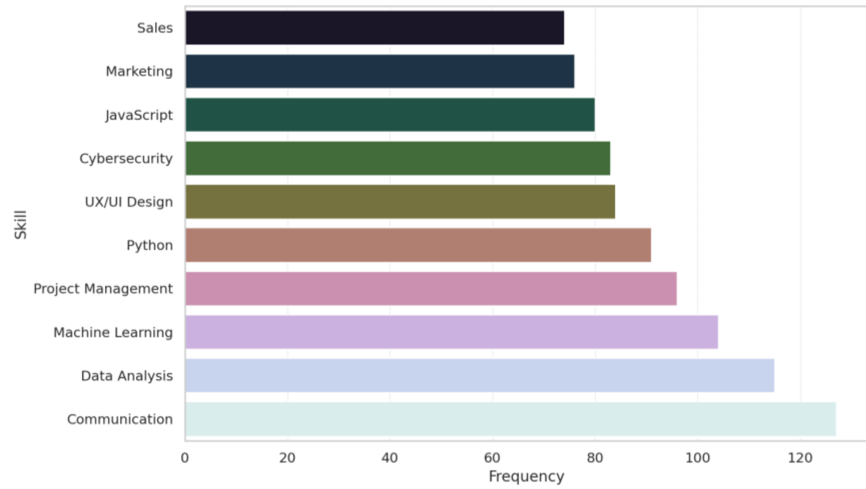

Figure 3 **Top 10 Most Frequent Skills in AI Job Roles.**

The distribution of skill frequency underscores the multidisciplinary nature of AI-related occupations. The prevalence of both technical competencies—such as Python, Machine Learning, and JavaScript—and non-technical abilities—such as Communication, Project Management, and Marketing—reflects the hybrid expectations placed on professionals in the field. This balance highlights the growing need for individuals who can not only design and implement intelligent systems but also communicate insights effectively, collaborate across teams, and align AI initiatives with business objectives. Such findings reinforce the notion that technical excellence alone is insufficient in today's AI job landscape; instead, a combination of domain knowledge, interpersonal communication, and strategic thinking is increasingly indispensable.

To investigate how specific skills relate to a job's vulnerability to automation, we conducted a correlation analysis between the presence of each skill and its associated automation risk score (numerically encoded as Low = 2, Medium = 1, High = 0). The objective was to identify which competencies are most strongly associated with increased or decreased automation susceptibility. The five most positively and five most negatively correlated skills are summarized in table 4, while figure 4 provides a visual representation of these relationships. The analysis reveals that soft skills like Marketing, Sales, and Communication tend to correlate positively with automation risk—indicating higher vulnerability—whereas technical skills such as Data Analysis, Machine Learning, and Cybersecurity are negatively correlated, implying a more robust resistance to automation. This distinction offers valuable guidance for job seekers and educators aiming to future-proof skillsets in an evolving technological landscape.

Table 4 **Skills Most Associated with Automation Risk**

Skill	Correlation with Automation Risk
Marketing	+0.37
Sales	+0.33

Communication	+0.29
Project Management	+0.21
UX/UI Design	+0.18
Python	–0.15
Cybersecurity	–0.20
Machine Learning	–0.24
JavaScript	–0.28
Data Analysis	–0.32

Table 4 presents a ranked summary of the five skills that exhibit the strongest positive and negative correlations with automation risk, offering insight into how specific competencies influence a job role's vulnerability to technological displacement. Skills such as Marketing, Sales, and Communication show high positive correlation values, indicating that job roles emphasizing these skills are more likely to be automated shortly. These competencies often relate to routine, customer-facing, or administrative functions, which are increasingly being replicated by AI-powered chatbots, recommendation systems, or automated marketing platforms. On the other hand, technical skills such as Data Analysis, JavaScript, and Machine Learning demonstrate strong negative correlations, suggesting that roles demanding these skills are more resilient to automation, likely due to their complexity and the cognitive effort required to perform them.

To enhance the interpretability of these findings, figure 4 provides a visual representation of the correlation values across the top ten skills—five positively correlated and five negatively correlated. The horizontal bar plot allows for a direct comparison of skill impact on automation susceptibility, with the vertical zero line serving as a reference point between risk-promoting and risk-mitigating competencies. This visualization clearly delineates the divide between soft-skill-dominated roles, which face higher automation risks, and technical-skill-centric roles, which are more secure in the face of technological disruption. These insights are critical for guiding curriculum design, upskilling initiatives, and career planning in an AI-augmented labor economy.

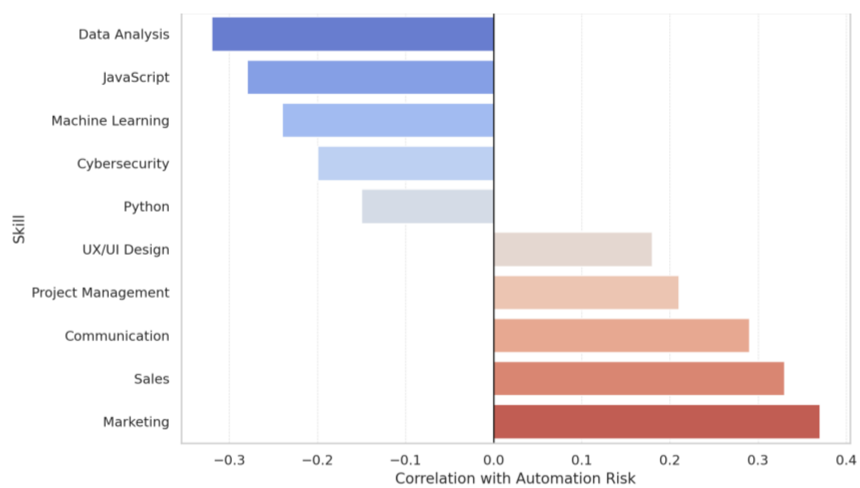

Figure 4 Skills Most Associated with Automation Risk

The results indicate a clear dichotomy in automation resilience between roles dominated by technical expertise and those centered on interpersonal or commercial functions. Specifically, job roles that demand competencies such as Machine Learning, Data Analysis, and Cybersecurity tend to exhibit greater resistance to automation. These skills typically require a high degree of cognitive complexity, problem-solving, and adaptive reasoning—capabilities that are currently difficult to replicate through automated systems. Conversely, roles that emphasize communication, marketing, and sales-oriented tasks show a higher susceptibility to automation, as these functions often involve repetitive, pattern-driven processes that can be efficiently handled by AI tools such as natural language generation, recommendation engines, or customer service chatbots. This pattern reinforces the critical importance of technical upskilling in enhancing workforce resilience in the face of ongoing digital transformation.

Discussion

The findings from this study reveal a nuanced segmentation of the AI job market, highlighting how different roles cluster together based on the types of skills they require and their associated levels of automation risk. The application of PCA followed by K-Means clustering enabled the identification of four distinct clusters, each representing a unique profile of job characteristics. Notably, Cluster 1 encompassed roles that are highly technical, well-compensated, and demonstrate a low risk of automation. These jobs—likely consisting of AI engineers, data scientists, and machine learning specialists—demand advanced skills such as Python programming, algorithm development, and data modeling. In contrast, Cluster 0 contained the largest proportion of jobs, but these roles tend to be less technical, lower in salary, and more vulnerable to automation. These may include positions in marketing support, administration, or client interaction—roles that increasingly rely on repetitive tasks, and thus face a higher probability of being replaced by AI technologies.

The skill frequency analysis and correlation with automation risk offer deeper insight into the composition and future prospects of these roles. Technical skills such as Data Analysis, Machine Learning, and Cybersecurity were consistently associated with clusters that are less exposed to automation, suggesting that these competencies enhance job security and long-term career sustainability. On the other hand, soft skills such as Communication, Sales, and Marketing—while still important for organizational functions—were more prevalent in high-risk clusters, and showed a strong positive correlation with automation risk. This suggests that while soft skills remain essential, they may not be sufficient alone to protect workers from displacement in an AI-intensive economy. The growing sophistication of AI systems capable of mimicking human interaction and decision-making—such as chatbots, AI-driven customer analytics, and automated content generation—poses significant challenges for non-technical roles.

Beyond the descriptive insights, these results carry important implications for stakeholders such as job seekers, educators, employers, and policymakers. For individual professionals, the findings emphasize the importance of acquiring and continuously updating technical competencies, particularly in programming, machine learning, and data processing. For educational institutions, the data underscores the necessity of revising curricula to integrate interdisciplinary learning—combining technical training with communication, ethics, and strategic thinking to prepare students for hybrid AI roles. For employers, understanding the risk distribution across roles can aid in designing more targeted workforce development strategies, including upskilling pathways for vulnerable employees. Finally, for policymakers, these findings highlight the need for forward-looking labor policies and investment in lifelong learning infrastructures to support workers in adapting to rapid technological shifts.

In sum, the clustering and correlation analyses presented in this study illustrate the critical role of skill specialization in shaping both the resilience and value of AI job roles. As AI systems continue to evolve and automate an expanding array of functions, the capacity to combine domain knowledge with advanced technical skills will increasingly define who thrives—and who is displaced—in the future of work.

Conclusion

This study investigated the structure of the AI job market by clustering roles based on required skills and their associated automation risk. Using PCA to reduce dimensionality and K-Means clustering to identify latent groupings, we discovered four distinct clusters of AI job roles. Each cluster revealed meaningful variations in salary, skill composition, and automation susceptibility—highlighting the fragmented and stratified nature of the AI labor ecosystem.

The analysis demonstrated that technical skills such as Machine Learning, Python, and Data Analysis are strongly associated with roles that are both higher-paying and more resilient to automation. In contrast, job roles emphasizing soft skills such as Marketing, Communication, and Sales tend to fall into lower-paying clusters and are more susceptible to technological displacement. These patterns were further supported by correlation analysis, which quantified the relationship between individual skills and automation risk. The results affirm the value of technical expertise in securing stable and future-proof careers in the evolving AI-driven economy.

Beyond the descriptive insights, this research underscores the critical need for continuous upskilling and curriculum reform. Stakeholders—including educators, policymakers, and employers—must adapt to the realities of a rapidly transforming job market by investing in interdisciplinary learning and reskilling initiatives. Future AI professionals will require a blend of technical proficiency and adaptive soft skills to thrive in hybrid roles. By mapping the landscape of AI employment through data-driven methods, this study provides a foundation for more informed workforce planning, talent development, and strategic education policy aimed at maximizing human potential in the age of automation.

Declarations

Author Contributions

Conceptualization: U.R. and Q.A.; Methodology: U.R. and Q.A.; Software: U.R. and Q.A.; Validation: U.R. and Q.A.; Formal Analysis: U.R. and Q.A.; Investigation: U.R. and Q.A.; Resources: U.R. and Q.A.; Data Curation: U.R. and Q.A.; Writing Original Draft Preparation: U.R. and Q.A.; Writing Review and Editing: U.R. and Q.A.; Visualization: U.R. and Q.A. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, andor publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M. Tegmark, Life 3.0: Being Human in the Age of Artificial Intelligence, 1st ed. New York, NY, USA: Alfred A. Knopf, 2017.

[2] E. Brynjolfsson and A. McAfee, The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies, 1st ed. New York, NY, USA: W. W. Norton & Company, 2014.

[3] C. B. Frey and M. A. Osborne, “The future of employment: How susceptible are jobs to computerisation?,” Technol. Forecast. Soc. Change, vol. 114, no. 1, pp. 254–280, Jan. 2017, doi: 10.1016/j.techfore.2016.08.019.

[4] L. Alekseeva, J. Azar, M. Giné, S. Samila, and B. Taska, “The demand for AI skills in the labor market,” Diffusion of Innovation eJournal, 2019, doi: 10.2139/ssrn.3470610.

[5] C. B. Frey and M. A. Osborne, “The future of employment: How susceptible are jobs to computerisation?,” Technol. Forecast. Soc. Change, vol. 114, pp. 254–280, 2017. doi: 10.1016/j.techfore.2016.08.019

[6] M. Chui, J. Manyika, and M. Miremadi, “Where machines could replace humans—and where they can’t (yet),” McKinsey Quarterly, no. 2, pp. 58–69, 2016.

[7] L. Nedelkoska and G. Quintini, “Automation, skills use and training,” OECD Social, Employment and Migration Working Papers, no. 202, pp. 1–36, 2018. doi: 10.1787/2e2f4eea-en

[8] E. Brynjolfsson, D. Rock, and C. Syverson, “Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics,” in The Economics of Artificial Intelligence: An Agenda, A. Agrawal, J. Gans, and A. Goldfarb, Eds. Chicago, IL, USA: Univ. of Chicago Press, pp. 23–57, 2019. doi: 10.7208/chicago/9780226613475.003.0002

[9] OECD, “OECD Employment Outlook 2019: The Future of Work,” OECD Publishing, Paris, France, 2019. doi: 10.1787/9ee00155-en

[10] X. Liu, Y. Zhang, and H. Liu, “Clustering AI job roles using LinkedIn data,” Int. J. Data Sci., vol. 7, no. 2, pp. 89–101, 2020. doi: 10.1007/s41060-020-00212-1

[11] M. Ali and A. Ibrahim, “Skill clustering in digital labor markets using PCA and hierarchical analysis,” J. Comput. Soc. Sci., vol. 4, no. 3, pp. 295–312, 2021. doi: 10.1007/s42001-021-00103-4

[12] M. Webb, “The impact of artificial intelligence on the labor market,” SSRN Electron. J., pp. 1–45, 2020. doi: 10.2139/ssrn.3482150

[13] D. Acemoglu and P. Restrepo, “Robots and jobs: Evidence from US labor markets,” J. Polit. Econ., vol. 128, no. 6, pp. 2188–2244, 2020. doi: 10.1086/705716

[14] M. Arntz, T. Gregory, and U. Zierahn, “The risk of automation for jobs in OECD countries: A comparative analysis,” OECD Social, Employment and Migration Working Papers, no. 189, pp. 1–35, 2016. doi: 10.1787/5jlz9h56dvq7-en
[15] D. J. Deming and K. Noray, “STEM careers and the changing skill requirements of work,” Q. J. Econ., vol. 135, no. 4, pp. 1965–2005, 2020. doi: 10.1093/qje/qjaa021
[16] E. Atalay, A. Phongthiengtham, S. Sotelo, and D. Tannenbaum, “The evolution of work in the United States,” Am. Econ. J. Appl. Econ., vol. 13, no. 3, pp. 1–35, 2021. doi: 10.1257/app.20190321
[17] J. Bessen, “AI and jobs: The role of demand,” Natl. Bur. Econ. Res. (NBER) Working Paper No. 24235, pp. 1–30, 2019. doi: 10.3386/w24235
[18] Y. Zhang, F. Shao, and L. Wang, “Mapping cybersecurity job roles using clustering and PCA,” Inf. Syst. Front., vol. 21, no. 4, pp. 827–841, 2019. doi: 10.1007/s10796-018-9840-7
[19] M. Reddy and S. Sastry, “Unsupervised role classification in AI career pathways,” J. Artif. Intell. Res., vol. 74, pp. 305–326, 2022. doi: 10.1613/jair.1.13364
[20] J. Meng et al, “Nano-integrating green and low-carbon concepts into ideological and political education in higher education institutions through K-means clustering,” Heliyon, vol. 10, no. 10, pp. 1–19, May 2024. doi:10.1016/j.heliyon.2024.e31244
[21] W. Pranowo and A. R. Ramadhani, “Error-based correlation coefficient: An alternative to combine error and coefficient of correlation and its application in Geophysical Data,” Journal of Computational Science, vol. 88, no. Jun., pp. 1–11, Jun. 2025. doi:10.1016/j.jocs.2025.102611