Department of Digital Business
Journal of Artificial Intelligence and Digital Business (RIGGS)
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 4 No. 3 (2025) pp: 2225-2231
P-ISSN: 2963-9298, e-ISSN: 2963-914X

Web Scraping Analysis of Job Platform Adoption in Banten's Manufacturing Sector

Ferdian Bangkit Wijaya, Weksi Budiaji
[1,2]Statistics, Engineering, Universitas Sultan Ageng Tirtayasa
[1]ferdian.bangkit@untirta.ac.id*, [2]budiaji@untirta.ac.id

Abstract

This study addresses the critical disconnect between Indonesia's manufacturing sector and the digitally native Gen Z workforce, focusing on Banten Province, a region with high youth unemployment despite its industrial concentration. We quantitatively assess the digital presence of manufacturing companies on key recruitment platforms, including LinkedIn, Jobstreet, and the government's Karirhub portal, to quantify this gap. Using web scraping techniques with Python, company profile data was systematically collected and analyzed. The findings reveal a limited digital footprint, with overall company presence recorded at 43.24% on LinkedIn, 42.70% on Karirhub, and a notably lower 32.97% on Jobstreet. Significant disparities exist across subsectors; consumer-facing industries like Food, Beverage, and Automotive show high digital engagement, while sectors such as Textiles, Electronics, and Non-metallic Minerals lag considerably. Notably, the Tobacco industry was found exclusively on the government's Karirhub platform. This confirmed digital divide hinders effective talent acquisition and limits job seekers' access to credible information. We conclude that a strategic imperative exists for manufacturers to enhance their digital recruitment strategies. This is crucial not only for attracting Gen Z talent but also for aligning with Indonesia's national digitalization agenda to reduce unemployment.

Keywords: *Digital Recruitment, Manufacturing Industry, Generation Z, Web Scraping, Youth Unemployment*

1. Introduction

Digital transformation in Indonesia's manufacturing sector has become a pivotal driver for enhancing national industrial competitiveness amidst global economic competition. The implementation of digital technologies, such as the Internet of Things (IoT), real-time analytics, and digital asset management, has demonstrably improved energy efficiency and productivity within the Indonesian manufacturing sector [1]. Furthermore, digitalization plays a crucial role in operational efficiency, employee productivity, and customer satisfaction. However, the limitations of technological infrastructure remain a primary challenge to its widespread adoption [2]. The "Making Indonesia 4.0" initiative, launched by the government, serves as a primary catalyst for accelerating this digital transformation, aiming to bolster industrial competitiveness and efficiency through the comprehensive application of digital technologies.

The manufacturing sector is a cornerstone of Indonesia's economy, contributing 18.67% to the Gross Domestic Product (GDP). In 2023, the national manufacturing industry production index grew by 2.41%, which in turn drove the sector's GDP growth to 4.64% [3]. Banten Province, as a major industrial hub, contributes an average of 30.36% annually to the national GDP from the processing industry sector [4]. Despite its significant economic contribution, Banten faces a substantial unemployment challenge, with rates reaching 10% in 2020. A key factor underlying this issue is the mismatch between technological advancements and the adaptive readiness of the local workforce [5].

According to the 2020 Population Census by Statistics Indonesia (BPS), the majority of Indonesia's population comprises Generation Z (born 1997–2012) and Millennials (born 1981–1996), accounting for 27.94% and 25.87% of the total population, respectively [6]. Combined, these two generations constitute over 50% of the population in their productive years, presenting significant potential for economic growth, innovation, and a demographic dividend projected through 2045 [7]. Nevertheless, in 2023, approximately 22.25% of Gen Z in Indonesia, equivalent to 9.9 million people, were recorded as unemployed, positioning them as the demographic with a high national open unemployment rate [8]. In Banten, the open unemployment rate (TPT) for Gen Z reached 9.05%, the second-highest in the country after West Java [9]. This high figure signifies a critical gap between industrial demand and the competencies of the young workforce, particularly in highly urbanized regions.

Generation Z is known for its heavy reliance on digital platforms for job searching. Social media channels like Instagram, Facebook, and LinkedIn, among other digital platforms, serve as their primary means for finding job vacancies and building a personal brand, which has been shown to enhance their competitiveness in the labor market [10]. Consequently, online recruitment (e-recruitment) is the preferred method for Gen Z due to its accessibility and time efficiency. Corporate reputation and branding are also critical factors in their decision-making process when applying for jobs [11]. Members of Gen Z actively investigate the authenticity and digital footprint of companies on platforms including LinkedIn, Instagram, and official websites to verify their existence and professionalism [12]. This digital due diligence is performed to avoid fraudulent job postings and recruitment scams by cross-referencing information across multiple platforms and assessing the consistency and transparency of corporate communications [13]. According to a survey by JakPat, the most widely used job search platform is Jobstreet (51.4%), followed by LinkedIn (38%) [14]. Additionally, the government's Karirhub platform, now integrated into the SIAPkerja system under the Ministry of Manpower, is intended to address labor market challenges through data and service provision [15].

Research conducted by Roso et al. on 150 manufacturing firms in Indonesia indicates that the adoption of digital technologies such as cloud computing and data analytics is positively correlated with a 22% increase in operational efficiency [16]. This finding aligns with previous studies highlighting the importance of developing centralized data infrastructure to support automation in the Industry 4.0 era [17]. The relevance of computational approaches in this study area is also substantiated by the authors' prior research [18], which successfully mapped industrial clusters using a machine learning algorithm to support policy formulation. Building on this premise, this study utilizes web scraping technology (Selenium) implemented in the Python programming language as a relevant solution for analyzing the digital activities of the manufacturing industry, specifically concerning company profiles and available job vacancies. This research aims to address the challenges associated with the digitalization of the manufacturing sector and to contribute to reducing unemployment by enhancing informational transparency for job seekers.

2. Research Methods

This study employs a quantitative methodology centered on web scraping, an automated process of extracting large-scale data from websites. This technique is essential in the context of digitalization, as it allows for the systematic collection of publicly available data that is not otherwise offered in a structured, machine-readable format. The primary data collection was conducted using a custom script written in Python. The process involved programmatically navigating to three target job recruitment platforms (LinkedIn, Jobstreet, and Karirhub) and systematically extracting information related to the online presence of manufacturing companies located in Banten Province. The selection of tools was guided by the need to handle both static and dynamic web content effectively. Python served as the core programming language for developing the web scraping scripts due to its extensive ecosystem of libraries designed for data extraction and analysis.

Selenium to navigate modern, dynamic websites that heavily rely on JavaScript to load content, Selenium was utilized. Selenium automates web browser actions, allowing the script to mimic human interactions such as scrolling, clicking buttons, and waiting for elements to load before extraction [19]. Its compatibility with various web drivers, such as Chrome Driver, makes it a versatile tool for accessing content that is inaccessible to static scrapers [20]. The primary role of Selenium in this research was to render the full HTML of target pages, especially on platforms with infinite scrolling or asynchronous data loading.

BeautifulSoup is once Selenium rendered a web page, the resulting static HTML source was passed to BeautifulSoup for parsing. BeautifulSoup is a Python library renowned for its efficiency and simplicity in navigating, searching, and modifying the parse tree of an HTML or XML document [21]. It excels at extracting specific data points, such as company names, locations, and job posting details, from the complex DOM structure [22]. While BeautifulSoup itself cannot execute JavaScript [23], its synergy with Selenium provides a robust framework: Selenium handles the dynamic content, and BeautifulSoup performs the detailed extraction, a common and effective approach in modern web scraping [24], [25].

2.1. Data

This research utilized a combination of secondary data sources to construct its dataset. The foundational data was sourced from the Processing Industry Company Directory of Banten Province on 2024, published by Statistics Indonesia (BPS) of Banten, which provided an initial list of manufacturing companies with a particular focus on Serang District. To assess the digital presence of these companies, this directory was augmented with data collected

DOI: https://doi.org/ 10.31004/riggs.v4i3.2282

via web scraping from three key online platforms in June 2025: LinkedIn, a professional networking site used for corporate profiles; JobStreet, a major regional job search portal; and Karirhub, the official career platform managed by Indonesia's Ministry of Manpower. From each platform, basic company profile information was extracted. Additional macro-level context on the national industrial landscape was obtained from the Indonesian Manufacturing Industry Directory.

2.2. Data Collection and Analysis

The data collection and analysis process was executed through a custom web scraping script developed in Python. Essential libraries, including requests, BeautifulSoup4, Selenium, and webdriver-manager, were utilized to support this process. Due to the dynamic, JavaScript-rendered nature of the target websites, a Selenium WebDriver with ChromeDriver was configured to automate browser interactions. This setup enabled the script to first perform automated logins on each platform, a necessary prerequisite for accessing company data. Following a successful login, the script was programmed to navigate to the specific company search pages on LinkedIn, JobStreet, and Karirhub. Once the target pages were fully loaded by the WebDriver, BeautifulSoup was employed to parse the HTML source code, allowing for the identification and extraction of relevant company profile information by targeting specific HTML elements within the page structure.

3. Results and Discussions

Based on data from the Indonesian Manufacturing Industry Directory for 2023-2024, there is a significant difference in the number of manufacturing industries across various provinces. West Java ranks first with 8,239 industries, accounting for approximately 25.4% of the total national manufacturing industries. It is followed by East Java with 6,151 industries (19.0%), and Central Java with 5,017 industries (15.5%). Banten Province, the focus of this research, ranks fourth with 3,397 industries, or 10.5% of the national total. This can be seen in Table 1.

Table 1: Number of Manufacturing Industries in Each Province

Year of Publication	Province	Number of Industries	Year of Publication	Province	Number of Industries
2024	Banten	3397	2024	Sulawesi Tengah	122
2024	DKI Jakarta	1359	2024	Sulawesi Tenggara	144
2024	Jawa Barat	8239	2024	Sulawesi Utara	114
2024	Jawa Tengah	5017	2024	Kalimantan Barat	374
2024	Yogyakarta	617	2024	Kalimantan Selatan	194
2024	Jawa Timur	6151	2024	Kalimantan Tengah	167
2024	Aceh	813	2024	Kalimantan Timur	291
2024	Sumatera Utara	1207	2024	Kalimantan Utara	40
2024	Sumatera Barat	260	2023	Maluku	36
2024	Riau	396	2023	Maluku Utara	172
2024	Jambi	194	2023	Bali	456
2024	Sumatera Selatan	331	2024	NTB	213
2023	Bengkulu	79	2024	NTT	76
2023	Kepulauan Riau	708	2024	Papua	17
2024	Kepulauan Bangka Belitung	152	2024	Papua Selatan	8
2024	Lampung	374	2024	Papua Tengah	9
2024	Gorontalo	51	2024	Papua Pegunungan	4
2024	Sulawesi Barat	27	2023	Papua Barat	42
2024	Sulawesi Selatan	536	2024	Papua Barat Daya	25

The concentration of manufacturing industries on the island of Java, particularly in West Java, East Java, Central Java, and Banten, indicates that this region remains the centre of national industrial growth. This phenomenon confirms the existence of industrial polarization centred on Java, especially in the *Jabodetabek* area and its

DOI: https://doi.org/ 10.31004/riggs.v4i3.2282

surroundings. However, the imbalance in industrial distribution between regions also presents its own challenges, particularly concerning the equitable distribution of economic development and employment absorption outside of Java. The disparity in industrial distribution leads to uneven economic development and job opportunities between regions in Indonesia. Islands outside of Java, such as Kalimantan, Sulawesi, Maluku, and Papua, have a very low percentage of industries (each less than 2%), reflecting a structural imbalance in the development of the national industry. This phenomenon aligns with previous research that indicates disparities in development between Western Indonesia (KBI) and Eastern Indonesia (KTI). This situation reinforces the urgency of digital transformation and the strengthening of industrial ecosystems in provinces with a high number of industries, such as Banten, to enhance competitiveness and operational efficiency. On the other hand, provinces with a low number of industries require specific strategies to encourage the growth of the manufacturing sector, whether through investment, infrastructure development, or improving the competency of the local workforce.

Table 2: Manufacturing Industry and Unemployed in Banten 2024

District	Number of Industries	Percent of Unemployed
Pandeglang	15	8.09%
Lebak	43	6.23%
Tangerang	1822	6.06%
Serang	370	9.18%
Tangerang City	852	5.92%
Cilegon City	83	6.08%
Serang City	33	7.12%
Tangerang Selatan City	179	5.09%

Table 2 shows that Serang District, with 370 manufacturing industries, has the highest unemployment rate in Banten Province (9.18%). Meanwhile, South Tangerang City, with fewer industries (179), records the lowest unemployment rate (5.09%). Similarly, Tangerang District, which has the highest concentration of industries in Banten (1,822 industries), shows a relatively moderate unemployment rate (6.06%). These findings indicate that the quantity of industries does not necessarily correlate negatively with the unemployment rate. This result is consistent with previous studies highlighting that the mismatch between technological advancements and the readiness of the local workforce is a major determinant of high unemployment rates, particularly in areas with high urbanization such as Serang District. The high unemployment rate in Serang District (9.18%) suggests a structural gap between industrial demand and labour supply. This phenomenon can be explained from several perspectives: first, the uneven digital transformation in the manufacturing sector in Banten contributes to the low absorption of the workforce. Second, the demographic characteristics of the labour force in Serang District, predominantly composed of Generation Z (27.94%) and Millennials (25.87%), face challenges in adapting to the needs of the manufacturing industry. This indicates a discrepancy between the career expectations of the younger generation and the reality of industrial demands.

Table 3: Availability of Industry Accounts on Job Search Platforms

Platform	Industry Accounts Availability	
	Active Accounts	% Accounts
Linkedin	160	43.24%
Jobstreet	122	32.97%
Karirhub	158	42.70%

The data from web scraping on digital job search platforms, as presented in Table 3, reveals significant variation in the digital presence of manufacturing industries in Banten Province. Of the total 370 manufacturing industries identified in Serang District, only 43.24% (160 out of 370 industries) have a presence on *LinkedIn*, 32.97% on *JobStreet*, and 42.70% on *Karirhub*. These findings indicate a substantial digital gap in the recruitment strategies of manufacturing industries in Serang District, where more than 50% of industries lack adequate digital representation on major job search platforms. The absence of a digital presence for most manufacturing industries

DOI: https://doi.org/ 10.31004/riggs.v4i3.2282

on job search platforms confirms an information asymmetry in the labour market in Serang District. This phenomenon contributes to the paradox previously identified, where there is a coexistence of a high concentration of industries (370 industries) with a similarly high unemployment rate (9.18%). This information asymmetry hinders the effective diffusion of job vacancy information, particularly for Gen Z and Millennial workers who heavily rely on digital platforms for job information.

An analysis of digital engagement based on industry classification, as shown in Table 4, reveals significant disparities in the utilization of digital recruitment platforms across manufacturing subsectors. The Beverage Industry demonstrates the highest digital presence, with 75% on all three platforms (*LinkedIn*, *JobStreet*, and *Karirhub*), indicating the adoption of a comprehensive omni-channel strategy in its recruitment process. The Food Industry also shows substantial digital engagement (67.5% on *LinkedIn*, 57.5% on *JobStreet*, 47.5% on *Karirhub*), followed by the Motor Vehicle Industry (66.7% on *LinkedIn*, 50% on *JobStreet*, 66.7% on *Karirhub*). The high prevalence of digital engagement in these three sectors reflects the consumer-facing nature of these industries, making them more responsive to digitalization trends. In contrast, several manufacturing subsectors show significant limitations in adopting digital platforms. The Tobacco Processing Industry displays an imbalanced digital pattern (0% on *LinkedIn*, 0% on *JobStreet*, 100% on *Karirhub*), indicating exclusive reliance on government platforms. The Garment Industry shows total absence on *Karirhub* (0%) despite moderate presence on commercial platforms (50%). Similarly, the Computer, Electronics, and Optics Industry demonstrates relatively low digital presence (25% on *LinkedIn*, 0% on *JobStreet*, 50% on *Karirhub*), despite operating in a technology-based sector.

Table 4: Manufacturing industry Account Activeness on job search platforms by industry type

Industry Type based on BPS	Active Platform		
	Linkedin	Jobstreet	Karirhub
Industri Alat Angkutan lainnya	66.7%	16.7%	41.7%
Industri Bahan Kimia	42.0%	34.8%	42.0%
Industri Barang Galian bukan logam	24.4%	24.4%	29.3%
Industri Barang Logam Bukan Mesin	42.9%	32.1%	46.4%
Industri Batubara dan minyak bumi	33.3%	33.3%	33.3%
Industri Farmasi, Obat Kimia, Tradisional	50.0%	75.0%	25.0%
Industri Furnitur	50.0%	33.3%	16.7%
Industri Karet dan Plastik	29.3%	22.0%	34.1%
Industri Kayu, Barang dari kayu dan gabus	42.9%	28.6%	28.6%
Industri Kendaraan Bermotor, Trail, dan Semi Trail	66.7%	50.0%	66.7%
Industri Kertas dan barang dari kertas	53.8%	46.2%	61.5%
Industri Komputer, Elektronik, Optik	25.0%	0.0%	50.0%
Industri Kulit, Barang dari kulit, alas kaki	40.0%	26.7%	46.7%
Industri Logam Dasar	36.7%	16.7%	36.7%
Industri Makanan	67.5%	57.5%	47.5%
Industri Mesin dan perlengkapan ytdl	25.0%	25.0%	50.0%
Industri Minuman	75.0%	75.0%	75.0%
Industri Pakaian Jadi	50.0%	50.0%	0.0%
Industri Pengolahan Lainnya	44.4%	44.4%	55.6%
Industri Pengolahan Tembakau	0.0%	0.0%	100.0%
Industri Peralatan Listrik	56.3%	50.0%	62.5%
Industri Tekstil	41.7%	8.3%	58.3%

DOI: https://doi.org/ 10.31004/riggs.v4i3.2282

4. Conclusion

The distribution of national manufacturing industries remains highly concentrated on the island of Java, with West Java, East Java, Central Java, and Banten as the main contributors. In Banten Province, Serang District stands out with a significant number of industries (370 industries); however, it still records the highest open unemployment rate in the region (9.18%), indicating a mismatch between the availability of industries and the absorption of local labor. Further analysis shows that only a portion of industries in Serang District actively utilize job search digital platforms, with a presence of 43.24% on LinkedIn, 32.97% on JobStreet, and 42.70% on Karirhub, meaning that information asymmetry in the labor market remains a major obstacle. A more detailed classification based on industry types reveals that subsectors such as the Beverage Industry, Food Industry, and Motor Vehicle Industry have the highest digital engagement across all three job search platforms, indicating better digital adoption and potentially broader access to labor. In contrast, subsectors such as the Tobacco Processing Industry, Garment Industry, Computer, Electronics, and Optics Industry show very low or uneven digital presence, meaning job opportunities in these sectors are less optimally communicated to job seekers. Optimizing the use of digital platforms by all industry types is key to enhancing the transparency of labor information and improving the effectiveness of labor absorption in regions with high unemployment rates, such as Serang District. This study has limitations in terms of geographic and temporal scope, and further research is needed to analyze the longitudinal dynamics of digital transformation in labor recruitment. The development of an artificial intelligence-based recommendation system that integrates job seeker profile data with specific manufacturing industry needs could be a focus for future research. Additionally, a comparative analysis of the effectiveness of various digital platforms in facilitating labor matching based on demographic characteristics should also be explored. Overall, optimizing digital platforms within the employment ecosystem is a strategic imperative in bridging the gap between industrial demand and labor supply.

Reference

[1] D. Setiawan, B. S. Nugroho, A. Nurrahman, F. A. Hamid, and K. Saputra, "The Effect of IoT Technology, Real-Time Analytics, and Digital Asset Management on Energy Efficiency and Productivity in Indonesia's Manufacturing Industry," *West Science Social and Humanities Studies*, vol. 2, no. 06, pp. 960–971, 2024, doi: 10.58812/wsshs.v2i06.977.

[2] M. Nugroho, F., Fauzi, A., Firdaus, R., & Marna, "The Role of Digital Transformation in Enhancing Organizational Performance in Multisectoral Industries," vol. 5, no. 11, pp. 1263–1272, 2024, doi: https://doi.org/10.46799/jst.v5i11.1022.

[3] BPS Indonesia, "Perkembangan Indeks Produksi Industri Manufaktur 2020," *Jakarta : Badan Pusat Statistik*, pp. 9–15, 2024, [Online]. Available: https://www.bps.go.id/id/publication/2021/08/18/0fbecb97aae729e237657dc6/perkembangan-indeks-produksi-industri-manufaktur-2020.html

[4] BPS Indonesia, "TRIWULAN III-2024 PEREKONOMIAN BANTEN TUMBUH 4,93 PERSEN (y-on-y)." Accessed: May 01, 2025. [Online]. Available: https://banten.bps.go.id/id/news/2024/11/05/55/triwulan-iii-2024-perekonomian-banten-tumbuh-4-93-persen--y-on-y-.html

[5] G. C. Mutaqin and I. Amaliah, "Pengaruh Laju Pertumbuhan Sektor Industri, Investasi Dan Upah Terhadap Penyerapan Tenaga Kerja Sektor Industri Di Provinsi Banten Tahun 2011-2020," *Bandung Conference Series: Economics Studies*, vol. 3, no. 2, pp. 458–466, 2023, doi: 10.29313/bcses.v3i2.8771.

[6] A. Devina, "Consumers Behavior Evaluation and Segmentation of Millennial and Generation Z in Indonesia Instant Noodle Market," *International Journal of Current Science Research and Review*, p., 2022, doi: 10.47191/ijcsrr/v5-i4-38.

[7] R. Pramadian and P. Belgiawan, "Emerging Market Entry: Innovative Marketing Approaches for Tirtagangga Hot Spring Resort Targeting Gen Z and Millennials," *International Journal of Current Science Research and Review*, p., 2023, doi: 10.47191/ijcsrr/v6-i6-57.

[8] S. J. D. Ayu and I. Diksa, "Determinants of Unemployment Among Gen Z in South Sulawesi," *Jurnal Ketenagakerjaan*, 2024, doi: 10.47198/jnaker.v19i3.382.

[9] N. K. S. Muhammad Hanri, "Potret Gen Z yang Menganggur: Mengungkap Realita di Balik Angka," *Labor Market Brief*, vol. 5, no. 6, pp. 1–7, 2024.

[10] N. M. Trang, B. McKenna, W. Cai, and A. M. Morrison, "I do not want to be perfect: investigating generation Z students' personal brands on social media for job seeking," *Information Technology and People*, vol. 37, no. 2, pp. 793–814, 2024, doi: 10.1108/ITP-08-2022-0602.

[11] M. Ariani and Zulhawati, "Digital transformation in the recruitment process: Revealing Gen Z's interest in career opportunities," *International Journal of Research in Business and Social Science (2147- 4478)*, p., 2024, doi: 10.20525/ijrbs.v13i9.3804.

DOI: https://doi.org/ 10.31004/riggs.v4i3.2282

[12] H. Rani, J. Jalih, and L. Widyowati, "Indonesian Generation Z Work Expectation and Intention to Apply for Job: Role of Social Media," *Quantitative Economics and Management Studies*, p., 2022, doi: 10.35877/454ri.qems831.

[13] R. Geofanny and F. Faraz, "Employer Branding and E-recruitment Against Interest in Applying for Generation Z Jobs," *Psikostudia : Jurnal Psikologi*, p., 2023, doi: 10.30872/psikostudia.v12i1.9922.

[14] Katadata, "6 Platform Lowongan Kerja Paling Banyak Digunakan di Indonesia." Accessed: May 01, 2025. [Online]. Available: https://databoks.katadata.co.id/ketenagakerjaan/statistik/dd6adbdcb128ea1/6-platform-lowongan-kerja-paling-banyak-digunakan-di-indonesia

[15] B. Teguh Gunawan, Z. Rajagukguk, F. A. Nasution, Muhyiddin, and S. A. Al-Ayubbi, "Identitas Digital Ketenagakerjaan Pada Sistem Informasi Pasar Kerja di Indonesia: Sebuah Konsep," *Jurnal Ketenagakerjaan*, vol. 17, no. 1, pp. 1–15, 2022, doi: 10.47198/naker.v17i1.123.

[16] E. Roso *et al.*, "Transformasi Digital di Industri Manufaktur : Dampak pada Efisiensi Operasional," *Elektriese: Jurnal Sains dan Teknologi Elektro*, vol. 14, no. 02, pp. 203–211, 2024, doi: https://doi.org/10.47709/elektriese.v14i02.4809.

[17] M. Javaid, A. Haleem, R. Singh, and R. Suman, "Significant Applications of Big Data in Industry 4.0," *Journal of Industrial Integration and Management*, p., 2021, doi: 10.1142/S2424862221500135.

[18] F. B. Wijaya, W. Budiaji, and A. S. Wicaksono, "Applied Machine Learning DBSCAN for Identifying Clusters of Micro and Small Industries," *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 2, pp. 380–386, May 2025, doi: 10.31004/riggs.v4i2.515.

[19] Lakkakula. S. Lakshmi, D. Sumithra, K. Jhansi, and J. Vandana, "Web Scraping And Data Analysis For Online Shopping With Selenium," *IARJSET*, p., 2024, doi: 10.17148/iarjset.2024.11329.

[20] A.-M. Tanasa, S. Oprea, and A. Bâra, "Web Scraping and Review Analytics. Extracting Insights from Commercial Data," *Ovidius University Annals. Economic Sciences Series*, p., 2024, doi: 10.61801/ouaess.2023.2.45.

[21] S. Pant, Er. N. Yadav, Milan, M. Sharma, Y. Bedi, and A. Raturi, "Web Scraping Using Beautiful Soup," *2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS)*, vol. 1, pp. 1–6, 2024, doi: 10.1109/ICKECS61492.2024.10617017.

[22] H. Bhoir, K. Jayamalini, and A. Info, "Web Crawling on News Web Page using Different Frameworks," *Int J Sci Res Sci Technol*, p., 2021, doi: 10.32628/cseit2174120.

[23] Y. Kryvenchuk and M. Burak, "COMPARATIVE ANALYSIS OF SELENIUM AND BEAUTIFULSOUP EFFICIENCY," *Herald of Khmelnytskyi National University*, p., 2022, doi: 10.31891/2307-5732-2022-305-1-50-52.

[24] A. Abodayeh, R. Hejazi, W. Najjar, L. Shihadeh, and R. Latif, "Web Scraping for Data Analytics: A BeautifulSoup Implementation," *2023 Sixth International Conference of Women in Data Science at Prince Sultan University (WiDS PSU)*, pp. 65–69, 2023, doi: 10.1109/WiDS-PSU57071.2023.00025.

[25] F. Chen, "Research on Real-time E-commerce Price Comparison System Using Python Web Scraping Technology," *International Journal of Computer Science and Information Technology*, p., 2024, doi: 10.62051/ijcsit.v4n2.18.

DOI: https://doi.org/ 10.31004/riggs.v4i3.2282