

## Exploring BSIT Graduates Employability Utilizing K-Means Clustering Technique

Renalyn G. Tecson

College of Computing Studies, Information and Communication Technology, Isabela State University, Echague Main Campus

**ABSTRACT:** As higher education institutions strive to ensure the relevance of academic programs, graduate employability has become a critical metric. This study investigates the employment outcomes of Bachelor of Science in Information Technology (BSIT) graduates from Isabela State University–Echague Campus from 2018 to 2022. Using a descriptive-exploratory research design, the study utilized the K-means clustering algorithm, a data mining technique, to analyze responses from 155 out of 230 graduates with 67.39% response rate. Data were collected through a CHED-aligned graduate tracer questionnaire via Google Forms. The dataset underwent preprocessing, descriptive statistical analysis, and unsupervised machine learning using Google Colab with Python. Findings revealed a high employment rate of 94.19%, with most graduates securing jobs within one year. Dominant job roles included Admin Officer, Office Clerk, and Programmer, reflecting both administrative and IT-focused employment trends. Key factors aiding job placement included educational qualifications, online applications, and personal recommendations. The clustering analysis generated three distinct clusters, which offer insight into graduate readiness and support mechanisms. The outcomes emphasize the importance of academic quality, industry-aligned competencies, and career support in enhancing employability. Results can inform curriculum development, graduate services, and institutional policies.

**KEYWORDS:** BSIT graduates, employability, K-means clustering, data mining, tracer study

### I. INTRODUCTION

Higher Education Institutions (HEIs) produce an increasing number of graduates each year, making graduate employability a critical concern for academic institutions. Graduate tracer studies play a key role in evaluating and improving the effectiveness of academic programs in contemporary higher education (Nudzor & Ansah, 2020). These studies serve as vital indicators of the relevance and impact of educational programs and institutions. Moreover, the quality of outcomes achieved by graduates is a fundamental component of higher education quality (Djan & Adawiyah, 2021). Tracer studies offer valuable insights into graduate outcomes, supporting efforts to enhance instructional delivery and overall institutional quality (Andari, Setiawan, & Rifqi, 2021).

In the case of the Isabela State University, particularly in the College of Computing Studies Information and Communication Technology (CCSICT) in Echague Campus, graduate tracer studies were conducted to determine the employability status of the BSIT graduates from 2011-2016. The result reveals that the employability of the BSIT graduates of the ISU-Echague Campus proved that the high status of employment reflects the quality of performance of graduates and the quality of education the University is providing them (Plata, 2020). Because of educational changes and the rise of disruptive technology, the demand in the workplace is fast-changing, emphasizing the importance of employability skills and literacy that will assure career success and degree program relevancy (Caingcoy, 2021). Moreover, to contribute to the development of high-level human resources capable of advancing national development goals (Cagasa, Belonias, & Cuadra, 2019), and in alignment with the Sustainable Development Goal on Quality Education—which aims, by 2030, to ensure that all learners acquire the knowledge and skills necessary to promote sustainable development (UN.org)—the College recognizes the need for ongoing graduate tracing. This initiative enables the evaluation of the relevance of graduates' acquired skills, the quality of instruction, and how it relates to their employability. Furthermore, since tracer study methodologies provide simple and utilizable results that can consume appropriately at an individual and institutional level (Dumlao & Manago, 2020), it is a potent tool that documents the profile of graduates, which gives implications on the performance of graduates (Deblois, 2021). Therefore, it is vital to apply data mining techniques to determine the intervention suitable for particular graduate students. Data mining techniques are used to determine the factors affecting graduates' employability (Othman, Shan & Yusoff, 2018). In this study, the employment status of BSIT graduates were analyzed, including the relevance of the skills and competencies acquired by the students in the University with their length of time in landing their job. In order for these to realized, the researcher opted to use descriptive and exploratory analysis using K-

## Exploring BSIT Graduates Employability Utilizing K-Means Clustering Technique

means clustering technique to determine the graduates' employability profiles. This study specifically aimed to summarize the employment status and types of the BSIT graduates from 2018-2022; explore the length of time for the graduates in landing their first job and the competencies they acquired in the University that helped them; and, utilize the K-means clustering algorithm to group BSIT graduates based on common attributes related to employability. This is a type of educational data mining technique to determine the graduate competencies and its alignment to their employability status. The result of this study can be used by the University to further improve the quality of education being provided, and most importantly on the revision and development of curriculum relevant to the needs of the industry.

### II. CONCEPTUAL FRAMEWORK

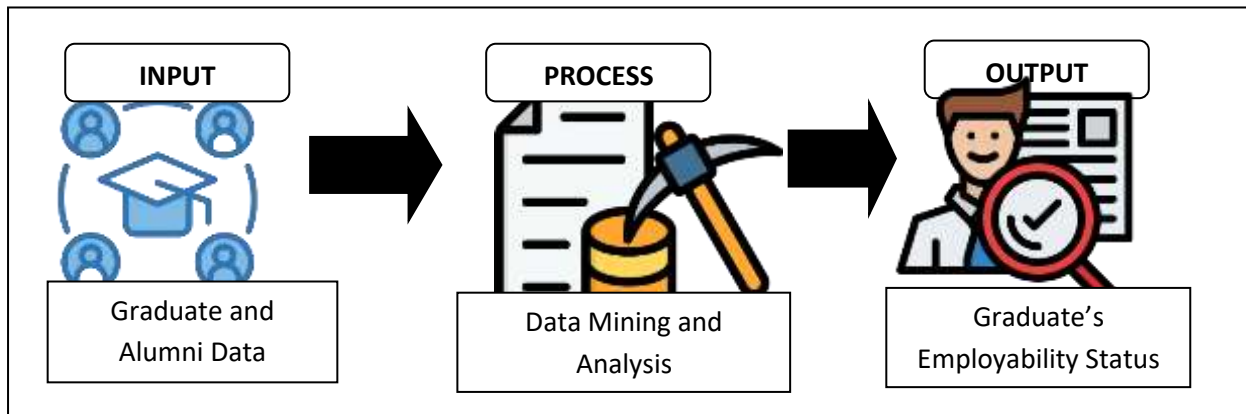


Figure 1: Conceptual Model of the Study

Figure 1 shows the conceptual framework which illustrates the flow of this study, which aims to explore BSIT graduates' employability using the K-means clustering technique. As shown in the figure, the model is structured around three interrelated components: Input, Process, and Output, representing the systematic flow of data collection, analysis, and interpretation in the context of alumni employability.

The **input** involves collecting relevant data from graduates and alumni which was tailored through the graduate tracer questionnaire by the Commission on Higher Education (CHED), including information on year graduated, certifications, acquired skills, job placement timelines, and employment-related details such as the time it took to secure a job post-graduation and the relevance of their current role to their field of study. This dataset serves as the foundation for the analysis. In the **Process** component, the study employs data mining and machine learning, specifically the K-means clustering algorithm, to explore hidden patterns and groupings within the dataset. Before clustering, necessary preprocessing and exploratory analysis steps are applied to ensure analytical integrity. The K-means clustering technique is particularly suited for this analysis as it allows the segmentation of graduates into distinct clusters based on similarities across multiple employability indicators. This clustering process is supported by internal validation metrics such as the silhouette score, which assesses the cohesion and separation of the generated clusters. Through this approach, the study moves beyond traditional descriptive analysis and applies an exploratory method to uncover meaningful profiles among graduates that may not be apparent through simple aggregation of statistics. Lastly, the output phase of the framework is a set of data-driven insights into the employability status of BSIT graduates. These insights can help academic institutions better understand the factors influencing graduate outcomes and inform strategies for curriculum improvement and graduate support.

### III. METHODOLOGY

This study employed a descriptive-exploratory quantitative research design utilizing data mining techniques, particularly the K-means clustering algorithm, to analyze the employability status of Bachelor of Science in Information Technology (BSIT) graduates. The descriptive aspect was used to summarize and describe graduate data, while the exploratory component focused on uncovering hidden patterns or groupings among graduates based on selected employability-related indicators.

## Exploring BSIT Graduates Employability Utilizing K-Means Clustering Technique

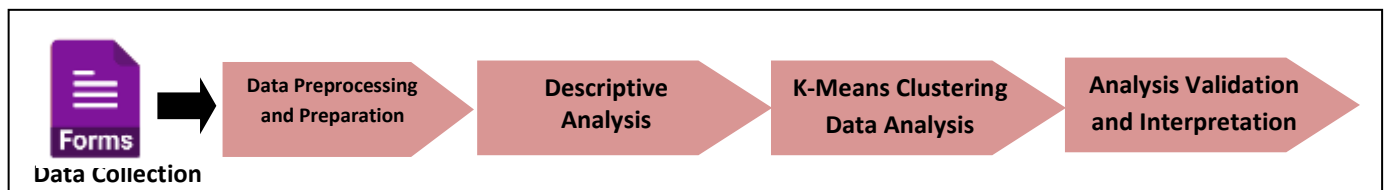


Figure 2: Graduates Employability Analysis Process Flow

Figure 2 illustrates a structured and systematic methodology adopted for analyzing graduates' employability through the application of K-Means Clustering, a machine learning technique widely used in unsupervised data analysis. The process is divided into five main stages: Data Collection, Data Preprocessing and Preparation, Descriptive Analysis, K-Means Clustering Data Analysis, and Analysis Validation and Interpretation.

### A. Data Collection

The initial stage involves the acquisition of data through google forms, which was designed anchored to the prescribed graduate tracer questionnaire of the Commission on Higher Education. The link was distributed to the graduates of 2018 to 2022, through posting in the official facebook page of the college, and facebook messenger group for each batch. The prescribed graduate tracer questionnaire includes the basic demographic profile of the graduates, such as, name, sex, marital status and address. This was followed by another section of the questionnaire which is the graduates employment information, such as, employment status, position or current job, length of time to land their first job, reasons for the delay (if appropriate), factors that facilitated them in getting their first/present job, eligibility, competencies acquired that helped them land their job. The data collected using the google forms were extracted in a microsoft excel format for data analysis. The total graduates from 2018 to 2022 is 230 graduates, and there were 155 respondents or 67.39% of the total graduates participated in the survey.

### B. Data Preprocessing and Preparation

Raw data often contains inconsistencies, missing values, and outliers that can negatively impact the analysis. Therefore, this stage encompasses a series of data cleansing, transformation, and normalization techniques aimed at preparing the dataset for analysis. The extracted results from the google form was prepared, data appropriate for analysis were selected to prepare a dataset. During the data preparation, data validation and removal of duplicate and irrelevant data were conducted. Once the dataset was prepared, activities such as encoding categorical variables, imputing missing values, and scaling numerical features are performed using the google colab with python to ensure that the data is suitable for clustering algorithms.

### C. Descriptive Analysis

Before applying clustering techniques, descriptive statistical analysis is conducted to summarize and explore the dataset. This includes calculating measures particularly frequency distributions, and visual representations through graphs were also performed to understand the underlying patterns and trends in the data. This stage serves to guide the selection of relevant variables for clustering and provides context for interpreting cluster outcomes.

### D. K-Means Clustering Data Analysis

At the core of the research lies the application of K-Means Clustering, an unsupervised learning algorithm that partitions the dataset into a predefined number of clusters (k) based on feature similarity. To perform the K-Means clustering analysis for exploring graduates' employability, the study employed the Python programming language in the Google Colaboratory (Colab) environment. Google Colab provides a cloud-based platform for running Python code, enabling efficient data analysis, visualization, and machine learning implementation without the need for local installations. The steps that were carried out during this phase was shown in Figure 3 below.

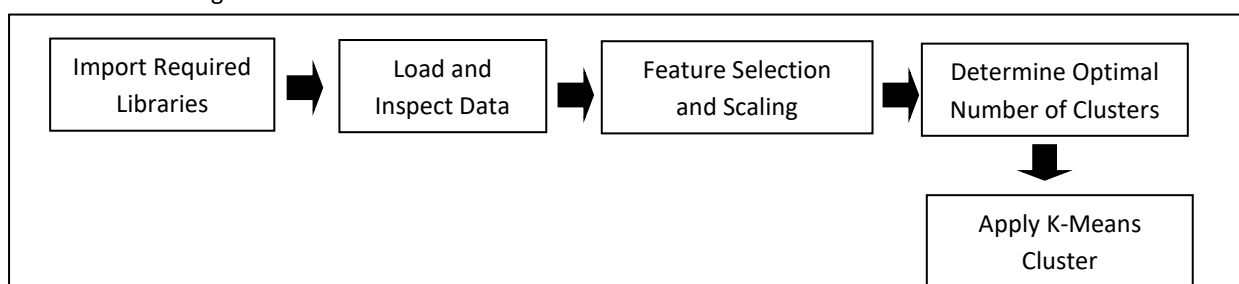


Figure 3: Graduates Data Analysis process using K-Means Clustering Technique

## Exploring BSIT Graduates Employability Utilizing K-Means Clustering Technique

### E. Analysis Validation and Interpretation

The final stage focuses on evaluating the quality and validity of the clusters generated by the K-Means algorithm. Techniques such as the Silhouette Score, Elbow Method, or Cluster Purity was employed to assess the coherence and separation of clusters. Additionally, interpretation of the clustering results is carried out to draw meaningful conclusions about employability factors and their distribution across different graduate profiles. These insights can inform educational institutions, policymakers, and employers about key trends affecting graduate employability.

## IV. RESULTS

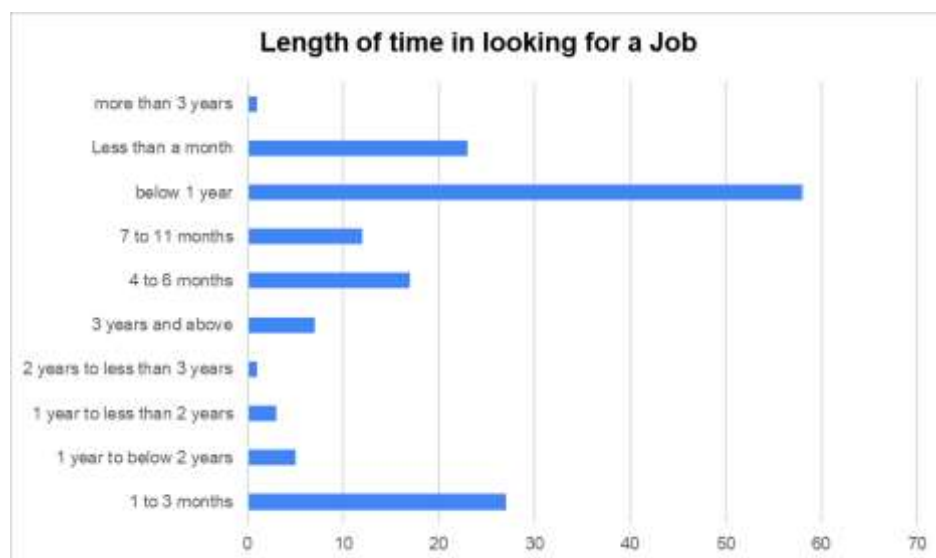
Based on the data gathered from 155 graduates out of 230 graduates from year 2018 to 2022, the following are the results of the analysis conducted and interpretation of the graduate tracer study.

### A. Graduates Employment Status

**Table 1: Distribution of Graduates per Employment Status**

| Employment Status | Frequency | Percentage rate |
|-------------------|-----------|-----------------|
| Employed          | 146       | 94.19%          |
| Self-employed     | 8         | 5.16%           |
| Unemployed        | 1         | 0.65%           |
| TOTAL             | 155       |                 |

Table 1 shows the frequency distribution of the 155 graduates from 2018 to 2022 based on employment status. As shown on the table, there were 146 or 94.19% from the respondent graduates are employed, eight(8) were self-employed, while there is one (1) who declared Unemployed at the time of data collection. The high employment rate is a strong indicator of the effectiveness of the academic program in preparing students for professional integration. It implies that the curriculum, instructional strategies, and experiential learning components such as internships and laboratory hands-on activities are responsive to industry standards and employer expectations. The presence of self-employed graduates further proved that the program's role in cultivating entrepreneurial skills, critical thinking, and innovation attributes essential for those who choose to create their own career paths. These results provide valuable feedback on the performance and relevance of the institution's academic offerings, specifically, the data suggest that the institution is successful in equipping students with practical skills, domain-specific knowledge, and professional values that translate into workforce readiness. The low unemployment rate also implies that graduates are competitive in the job market, which may reflect well on the institution's reputation and the quality of its graduates.

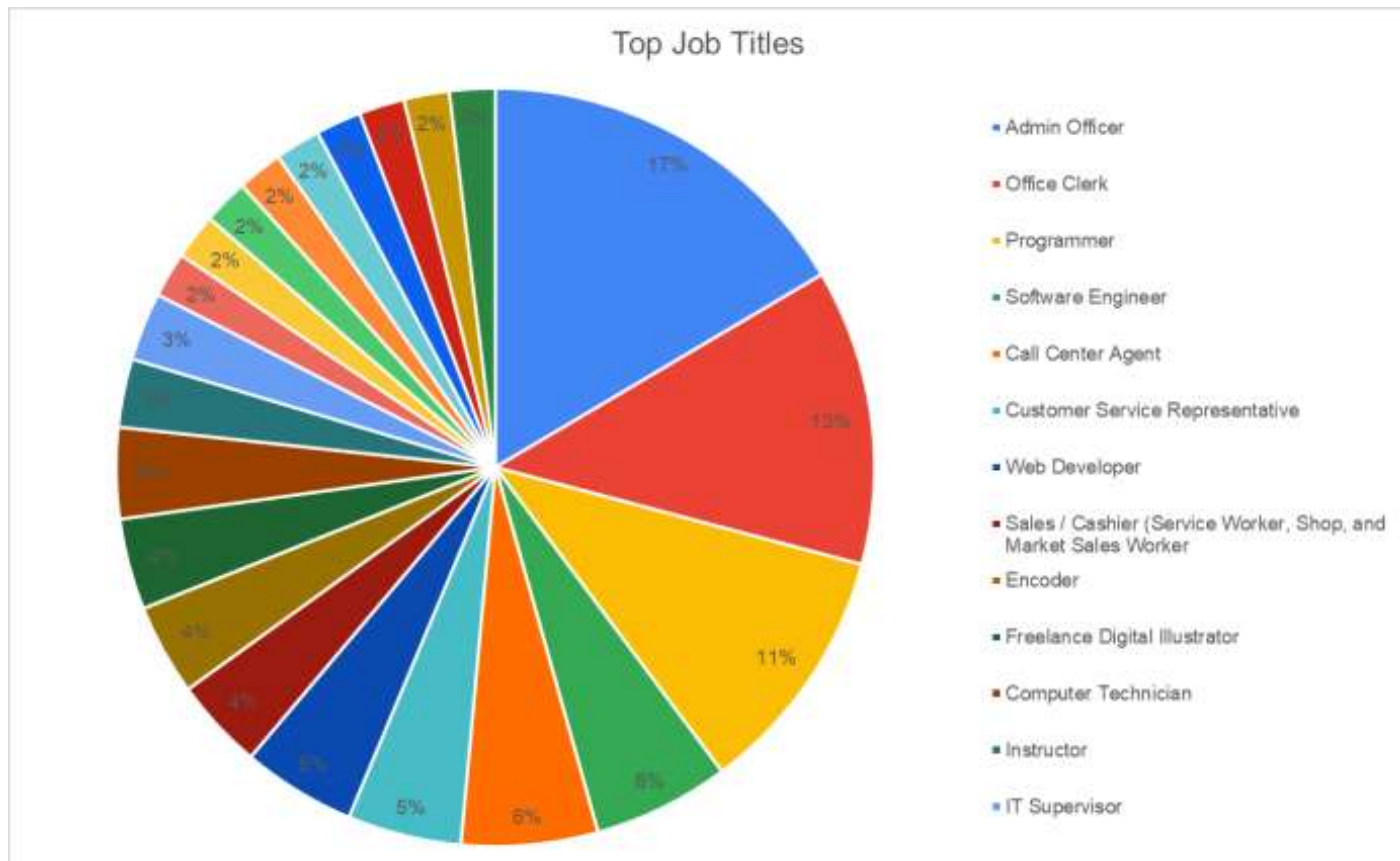


**Figure 4: Result of the Length of time in looking for a Job**

Tracking how long it takes graduates to find work helps the institution assess the quality of its programs and identify areas for improvement. Figure 4 presents the data visualization on how long it took graduates to find a job after completing their studies. The results show that the majority of graduates were able to find employment within one year. Specifically, the largest group, around 65 graduates were reported being employed within less than a year. Among them, about 25 graduates found jobs

## Exploring BSIT Graduates Employability Utilizing K-Means Clustering Technique

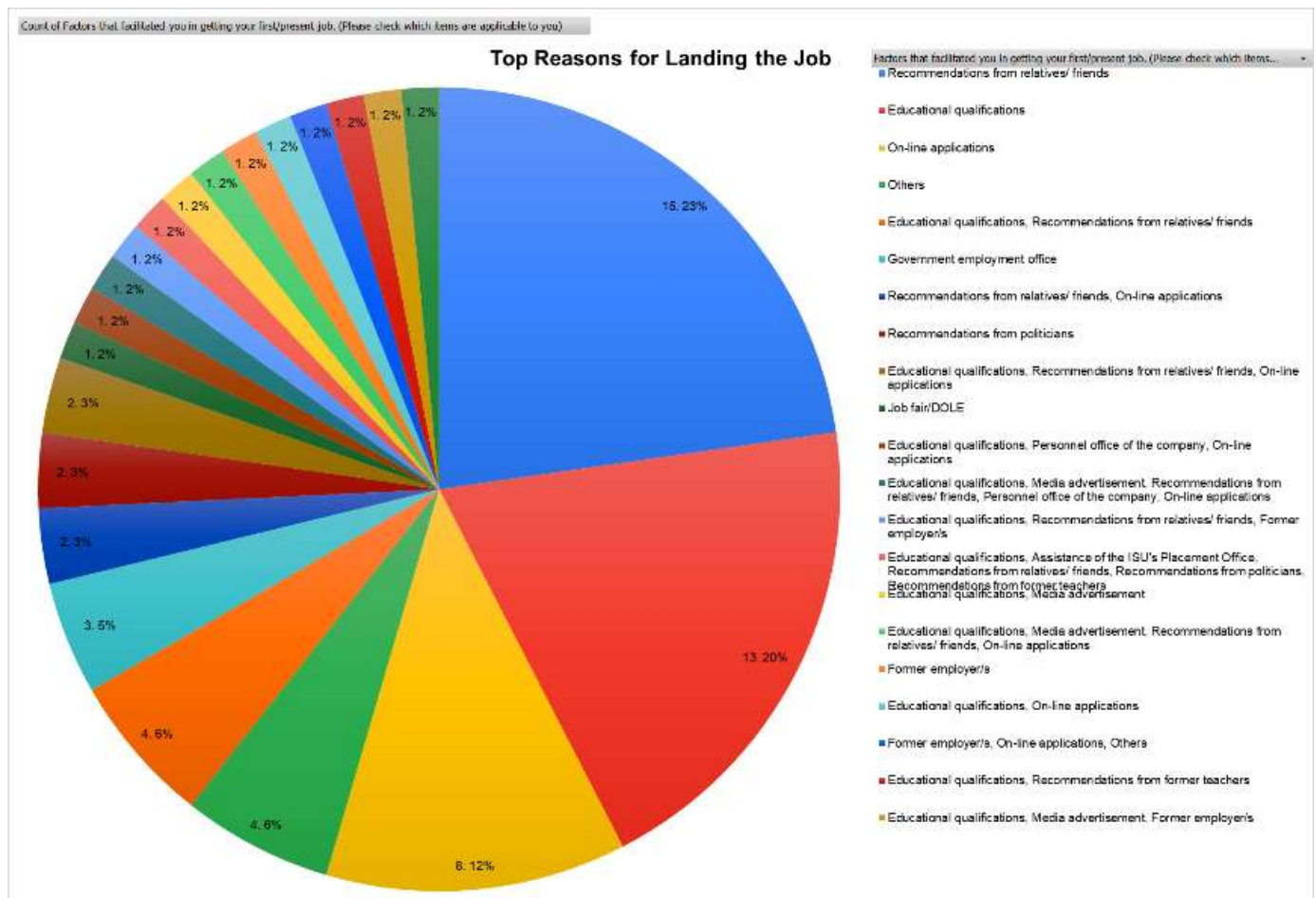
in 1 to 3 months, and 20 graduates secured jobs in less than a month. This indicates that many graduates transitioned quickly into the workforce. Smaller numbers of graduates found jobs between 4 to 6 months and 7 to 11 months after graduation, showing that while some took longer, most were still employed within a reasonable time frame. Only a few graduates experienced longer job searches, ranging from 1 year to over 3 years, which reflects positively on the employability of the group as a whole. These findings imply that the academic program effectively prepares students for employment. The short job-search period for most graduates indicates that they have the skills and qualifications needed by employers. This also implies that the curriculum is aligned with labor market demands.



**Figure 5: Data Visualization of the Top Job Titles of the Graduates**

Figure 5 displays the top job titles held by graduates, highlighting the diversity of career paths pursued. The most common job is Admin Officer, accounting for 17% of the total. This is followed by Office Clerk at 13%, and Programmer at 11%, showing that administrative and IT-related positions are among the most prevalent occupations for graduates. Other notable roles include Software Engineer (6%), Call Center Agent (6%), and Customer Service Representative (5%), suggesting that a considerable number of graduates are employed in both technical and service-oriented roles. Job titles such as Web Developer, Encoder, and Computer Technician, each with 4%–5%, indicate additional demand for digital and clerical skills. Less frequent but still relevant roles include Freelance Digital Illustrator, Instructor, and IT Supervisor, along with a range of other positions each comprising 2%–3% of the distribution. This wide variety of job titles reflects the versatility of graduate skills and their ability to adapt to different sectors. The distribution of job titles indicates that graduates are entering a broad reach of industries, with a strong presence in both administrative and technology-driven roles. This suggests that the academic program equips students with transferable skills applicable across various fields. The importance of IT-related roles (e.g., Programmer, Software Engineer, Web Developer, IT Supervisor) confirms that technical competencies are in demand and that the institution is successfully preparing graduates for careers in digital and technological environments. Additionally, the presence of freelance and instructional roles shows that some graduates are pursuing non-traditional or entrepreneurial career paths, which implies flexibility and initiative.

## Exploring BSIT Graduates Employability Utilizing K-Means Clustering Technique



**Figure 6: Top Reasons for the Graduates in Landing their Job**

Figure 6 illustrates the top reasons graduates attribute to successfully landing their jobs. The most frequently cited factor is educational qualifications, accounting for 15.27% of responses. This suggests that the knowledge and credentials acquired during their academic programs are seen as critical to employability. Online applications with 13.20% from the response rate influences graduates for job acquisition. Additionally, recommendations from relatives or friends got 8.12% and government employment offices, there were 6.63% from the respondents also played significant roles, highlighting the importance of both social networks and institutional support in facilitating employment.

Based on the analysis the following are notable combination of reasons:

- Educational qualifications and recommendations from relatives/friends
- Educational qualifications and online applications
- Educational qualifications and referral by former teachers or employees

These combinations reveal that graduates often benefit from multiple sources of support or advantages when seeking employment. The result also indicates that educational qualifications is a key factor confirms the continuing relevance of academic preparation in improving graduate employability.

### **B. K-Means Clustering Analysis Result**

To explore further the graduates data, the factors that facilitated the graduates in getting their job, and what competencies learned in college did they find useful in their job were clustered as to the graduates employment status and length of time in looking for a job were analyzed using the K-Means Clustering technique. Based on the analysis conducted through Google Colab along with python, the following is the generated clustering report of the data.

## BSIT Graduate Employability Clustering Report

### Cluster 0

Most Common Employment Status: Employed

Typical Job Search Duration: below 1 year

Top Job Facilitators: Recommendations from relatives/ friends

Useful Competencies: Communication Skills, Communication Skills, Information Technology Skills, Hardware and Software Installation Skills, Information Technology Skills, Problem-solving Skills, Critical Thinking Skills

### Cluster 1

Most Common Employment Status: Employed

Typical Job Search Duration: 4 to 6 months

Top Job Facilitators: None

Useful Competencies: None

### Cluster 2

Most Common Employment Status: Employed

Typical Job Search Duration: below 1 year

Top Job Facilitators: Educational qualifications

Useful Competencies: Communication Skills, Human Relation/Interpersonal Skills, Leadership/Managerial Skills, Entrepreneurial Skills, Information Technology Skills, Problem-solving Skills, Critical Thinking Skills, Research and Extension Skills, Programming Skills, System Analysis Skills, Hardware and Software Installation Skills, Network Administration Skills

Figure 7: Clustering Report generated from applying K-Means Clustering Technique for graduates tracer data

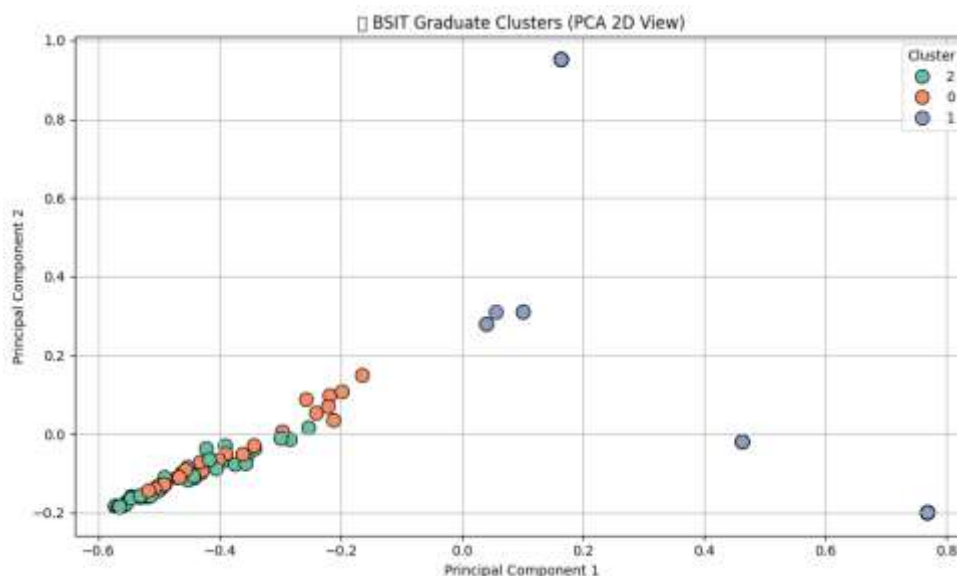


Figure 8: Visual Representation of the Clustering Report

Figures 7 and 8 presents the result of the graduates data analysis utilizing the k-means clustering technique. The clustering analysis presented in the PCA 2D (Figure 8) view offers valuable insights into the employability characteristics of BSIT (Bachelor of Science in Information Technology) graduates. Through the application of unsupervised machine learning techniques, specifically clustering in combination with Principal Component Analysis (PCA) for dimensionality reduction, the dataset reveals three distinct groups of graduates, each characterized by varying levels of employability readiness, academic performance, and career support

## Exploring BSIT Graduates Employability Utilizing K-Means Clustering Technique

access. As shown in Figure 7, Cluster 0 represents the “Network-Driven & Technically Equipped” graduates, occupies a central position within the data distribution, these are graduates who landed jobs relatively quickly due to social connections or referrals. Their employability appears to be supported by practical IT skills and soft skills like communication and problem-solving. These are likely graduates with moderate but well-aligned technical preparation and strong informal job-search networks. In contrast, Cluster 1, referred to as “Unassisted & Skill-Underreported” graduates, is spatially distant from the other clusters, indicating a distinct divergence in employability features. Graduates in this cluster may suffer from a lack of institutional support, career guidance, or self-marketing strategies, which could result in underreporting of skills on résumés, weak portfolio development, or limited job search engagement. The outlier nature of this cluster raises concerns about potential employability gaps that may hinder labor market integration. Cluster 2, identified as the “Well-Prepared & Academically Strong” graduates, is tightly grouped, indicating homogeneity in their employability profiles. Their proximity within the PCA space reflects consistency in high academic performance, clear career planning, and proactive engagement in professional development. This group represents the ideal graduate type, embodying both theoretical knowledge and practical competencies that align with industry expectations. The clustering outcome supports the assertion that academic excellence, when coupled with relevant training and institutional support, significantly enhances employability outcomes. This is the most comprehensively equipped cluster. These graduates clearly value their educational background, and their job placement seems heavily supported by formal qualifications. They demonstrate both technical expertise and soft/hybrid skills, making them highly employable in professional IT roles.

## V. CONCLUSIONS

The findings of this graduate tracer and clustering analysis provide strong empirical evidence of the effectiveness of the BSIT academic program in fostering employability among its graduates. With a high employment rate of 94.19%, and the majority securing jobs within one year after graduation, the program demonstrates responsiveness to labor market demands. The diversity of job titles ranging from administrative to technical roles, further affirms the versatility and interchangeability of the competencies acquired through the curriculum.

The clustering analysis using the K-Means technique, combined with PCA for visualization, revealed three distinct employability profiles among graduates: a) Cluster 0: "Network-Driven & Technically Equipped", these graduates benefit significantly from social referrals and possess practical IT and soft skills, enabling them to transition swiftly into employment; b) Cluster 1: "Unassisted & Skill-Underreported", the outlier group shows limited reporting of employability-enabling factors, suggesting gaps in career guidance, job search readiness, or self-marketing practices; and c) Cluster 2: "Well-Prepared & Academically Strong" group that represents the ideal graduate profile, characterized by strong academic foundations, comprehensive skill sets, and reliance on educational qualifications to secure employment. These differentiated clusters highlight the varied nature of graduates employability. While most graduates thrive through a combination of academic preparation and support systems, a small subset may require targeted interventions to fully realize their employment potential. The study affirms that educational qualifications, technical competencies, and job search facilitation mechanisms (e.g., referrals, online applications) are critical enablers of successful labor market integration.

Hence, the results highlights the importance of sustaining high academic standards, enhancing career support services, and promoting both formal and informal pathways to employment. These insights can inform continuous curriculum improvement, graduate support programs, and institutional policies aimed at strengthening graduate outcomes and employability trajectories in the ever-evolving ICT sector.

## ACKNOWLEDGMENT

The researcher wishes to express sincere gratitude to all the graduates who participated in the study. Their willingness to share their experiences and insights was instrumental in completing the research and enriching its findings. Deep appreciation is also extended to the University for the support and valuable insights from the research department.

## REFERENCES

- 1) Andari, S., Setiawan, A. C., & Rifqi, A. (2021). Educational management graduates: A tracer study from universitas negeri surabaya, indonesia. *IJORER: International Journal of Recent Educational Research*, 2(6), 671-681.
- 2) Cagasan, E. G., Belonias, B. S., & Cuadra, M. E. J. (2019). Graduate Students' Perceived Contribution of Scholarship Grants to Academic Success. *Sci. Humanit. J*, 13(1), 99-113.
- 3) Caingcoy, M. (2021). Scoping Review on Employability Skills of Teacher Education Graduates in the Philippines: A Framework for Curriculum Enhancement. *International Journal of Education & Literacy Studies*, 9(4), 182-188.

## Exploring BSIT Graduates Employability Utilizing K-Means Clustering Technique

- 4) Deblois, E. C. (2021). The Employment Profile of Graduates in a State University in Bicol Region, Philippines. *Journal of Education, Management and Development Studies*, 1(1), 33-41.
- 5) Djan, I., & Adawiyyah, S. R. (2021). A new decade for social changes. *Tech. Soc. Sci. J*, 17, 235-243.
- 6) Dumlao, T. R., & Manago, J. G. (2020). Tracing CCI education graduates: How well are we in developing professional teachers. *CC The Journal: A Multidisciplinary Research Review*, 15, 1-8.
- 7) Nudzor, H. P., & Ansah, F. (2020). Enhancing post-graduate programme effectiveness through tracer studies: the reflective accounts of a Ghanaian nation-wide graduate tracer study research team. *Quality in Higher Education*, 26(2), 192-208.
- 8) Othman, Z., Shan, S. W., Yusoff, I., & Kee, C. P. (2018). Classification techniques for predicting graduate employability. *International Journal on Advanced Science, Engineering and Information Technology*, 8(4-2), 1712-1720.
- 9) Plata, I.T. (2020). Connecting With The Graduates: An Online Employability Graduate Tracer Study (O-EGTS) Of Isabela State University.



There is an Open Access article, distributed under the term of the Creative Commons Attribution – Non Commercial 4.0 International (CC BY-NC 4.0) (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits remixing, adapting and building upon the work for non-commercial use, provided the original work is properly cited.