

## Applications of Data Mining techniques in multi-employment service facility

**Isagani Mirador Tano**

Quezon City University (QCU)

Quezon City, Philippines

Email: isaganitano062781@yahoo.com

### ABSTRACT

The employment facilitation service is generally perceived as being ‘information rich’ yet ‘knowledge poor’. This agency generates a wealth of data related to employment facilitation, however, there is a lack of research effective analysis tools to discover hidden relationships and trends in data. Knowledge discovery and data mining have found numerous applications in business and scientific domain. Valuable knowledge can be discovered from application of data mining techniques in employment facilitation services. In this study, it is briefly examined the potential use of classification-based data mining techniques such as Rule based, decision tree and Artificial Neural Network to massive volume of employment facilitation service data. In particular a study using classification techniques on the data available at the Quezon City Public Employment Service Office of the Department of Labor and Employment in the Philippines was conducted.

Keywords: employment facilitation service, Data Mining, DOLE, PESO, employment data

Date Submitted: August 28, 2025

Date Accepted: September 29, 2025

Date Published: October 28, 2025

### INTRODUCTION

Republic Act No. 8759, also known as the Public Employment Service Office (PESO) Act of 1999, established the PESO as a mechanism to achieve full employment and equality of opportunities for all Filipinos (Bureau of Local Employment, Department of Labor and Employment [DOLE]). As an institutionalized program of DOLE, PESO serves as an avenue for individuals to explore diverse employment options and seek the type of assistance suited to their needs. Its creation addresses the growing demand for accessible and localized labor market information, serving as a referral and information center for DOLE and other government agencies. Moreover, PESO fosters collaboration among regional and interregional employment networks,

---

DOI: <http://doi.org/10.69651/PIJHSS0404548>

Recommended citation:

Tano, I. M. (2025). Applications of Data Mining techniques in multi-employment service facility. *Pantao (The International Journal of the Humanities and Social Sciences)* 4 (4), 5956-5963.

<http://doi.org/10.69651/PIJHSS0404548>

facilitating the exchange of employment opportunities and ensuring that job seekers and employers can connect efficiently.

In the context of an information-driven society, knowledge has become one of the most valuable assets of any organization. Information Technology (IT) plays a vital role in enhancing the delivery and management of employment facilitation services such as those provided by PESO. The integration of knowledge management systems in these services presents challenges in creating, disseminating, and preserving employment-related knowledge through advanced technological platforms. The pragmatic use of database systems, data warehousing, and knowledge management technologies contributes significantly to the development of decision-support systems in employment facilitation, helping administrators make informed and data-based decisions.

Knowledge discovery in databases (KDD) is a systematic process comprising several critical stages, with data mining serving as its core. As defined by Frawley and Piatetsky-Shapiro (1996), data mining refers to the non-trivial extraction of implicit, previously unknown, and potentially valuable information from large datasets. This technology provides a user-oriented approach to identifying hidden patterns and insights within data. In the context of employment facilitation, these insights can be utilized by PESO administrators to enhance service quality, improve the efficiency of job matching, and optimize the dissemination of information about DOLE's various programs and initiatives.

Through the application of data mining, PESO can significantly improve operations across multiple service areas, such as creating a pool of qualified applicants for referral, minimizing job mismatch, disseminating job vacancies through job fairs, monitoring livelihood and self-employment programs, managing special credit assistance for overseas workers, implementing the Special Program for Employment of Students (SPES), and coordinating the Workers Hiring for Infrastructure Projects (WHIP). Decision-making in employment facilitation traditionally relies on historical data, experiences, and resource constraints; however, data mining and knowledge management technologies offer a data-driven approach that transforms these processes into knowledge-rich environments capable of more strategic and proactive planning.

In key urban centers such as Quezon City, PESO can adopt Knowledge Discovery in Databases (KDD) with the support of skilled personnel who possess expertise in both data analysis and the employment industry. KDD's ability to process large volumes of complex employment data allows for the identification of meaningful trends and the development of evidence-based strategies. Employment analysts and policymakers can draw valuable insights from KDD applications in other sectors, such as human resources and manpower management, to enhance the efficiency of employment programs and policy interventions.

Employment-related data encompass vast and intricate information, including employer and employee profiles, job applicants' work histories, educational attainment, skills, and certifications. To effectively utilize these datasets, PESO must strengthen its capacity for data analysis and digital record management. By computerizing employment records and employing data mining techniques, the agency can address critical challenges in employment management—particularly job mismatch—by uncovering patterns that inform targeted interventions. Ultimately, the strategic application of data mining within PESO can generate actionable knowledge, enabling the agency to deliver more efficient, equitable, and responsive employment facilitation services.

## Statement of the problem

Despite the vast amount of employment data collected by the Public Employment Service Office (PESO), much of it remains underutilized due to the lack of effective analytical tools for extracting meaningful insights. This limitation hinders efficient decision-making, contributes to job mismatch, and reduces the overall effectiveness of employment facilitation programs. Therefore, this study seeks to explore the application of data mining techniques—such as rule induction, decision trees, and artificial neural networks—to analyze and interpret employment data in order to enhance the efficiency and accuracy of services provided by multi-employment service facilities like the Quezon City PESO. Specifically, this study seeks to answer the following questions:

1. What types of employment-related data are available in the Quezon City Public Employment Service Office (QCPESO) that can be utilized for analysis through data mining techniques?
2. How can classification-based data mining techniques such as rule induction, decision tree, and artificial neural network be applied to the employment data of QCPESO?
3. What patterns or relationships can be discovered from the employment data using these data mining techniques?
4. How can the discovered knowledge be used to improve decision-making, reduce job mismatch, and enhance the efficiency of employment facilitation services?
5. What data mining model provides the most accurate and practical results for employment data analysis in a multi-employment service facility?

## METHODOLOGY

This study employed a quantitative-descriptive research design utilizing data mining techniques to explore patterns and relationships within the employment data gathered from the Quezon City Public Employment Service Office (QCPESO). The research focused on identifying how data mining methods such as rule induction, decision tree analysis, and artificial neural networks (ANN) could be applied to enhance the efficiency and accuracy of employment facilitation services. Data mining, as an essential step in knowledge discovery, involves an iterative process of data cleaning, integration, selection, pattern recognition, and knowledge presentation. It differs from traditional data analysis by being discovery-driven and interdisciplinary, drawing upon principles from statistics, machine learning, neural networks, and inductive logic programming (Glymour, Madigan, Predigon, & Smith, 1996). Through these processes, the study aimed to extract useful and previously unknown information from massive datasets that would be difficult to analyze using conventional statistical tools.

The research was conducted at the Quezon City Public Employment Service Office, a major multi-employment service facility operating under the Department of Labor and Employment (DOLE). The study population consisted of employment applicants whose records were stored in the QCPESO database. These records included relevant attributes such as age, gender, educational attainment, district location, work experience, and job classification. A purposive sampling technique was employed to select datasets that best represented the

demographic and employment diversity of the city's labor force. The selection of this locale was strategic, as Quezon City is among the most populous and economically active areas in the Philippines, thereby generating extensive and diverse employment-related data suitable for data mining analysis.

The primary data used in this study were secondary in nature, obtained directly from the QCPESO employment database. The dataset underwent preprocessing and transformation prior to analysis to ensure accuracy, consistency, and completeness. The data mining process followed an iterative cycle, including data cleaning to remove inconsistencies, data integration to consolidate multiple data sources, and data selection to identify relevant variables for analysis. Online Analytical Processing (OLAP) and data warehousing technologies were utilized to organize and store data efficiently, allowing multi-dimensional analysis and rapid retrieval of information (Shams & Frashita, 2001). This enabled the research to convert raw data into structured information, forming the foundation for the development of predictive and classification models.

Three classification-based data mining techniques were applied in the analysis: rule induction, decision tree, and artificial neural network. Rule induction involved extracting significant "if-then" rules from data based on statistical associations to predict applicant-job compatibility. The decision tree model, constructed through algorithms such as ID3 and C4.5 (Han & Kamber, 2001), was used to classify and predict employment outcomes by analyzing variables such as district, education level, and gender. The artificial neural network (ANN) approach simulated human cognitive processes to identify complex, non-linear patterns within the employment data (Lu, Setiono, & Liu, 1996). These models collectively aimed to identify key predictors that influence job placement, reduce mismatch, and enhance data-driven decision-making in employment facilitation services.

The data gathering process adhered to ethical research standards. Prior permission was secured from the Quezon City PESO and the Department of Labor and Employment to access and analyze the employment records. All data were treated with strict confidentiality and anonymity, ensuring that no personally identifiable information was disclosed or used beyond the scope of this research. Data were processed solely for academic and analytical purposes, and findings were reported in aggregate form to maintain ethical integrity.

For data analysis, the study utilized data mining software capable of implementing classification algorithms and generating visual decision models. The data were evaluated using cross-tabulation, classification accuracy measures, and predictive validation to ensure reliability of the models developed. Patterns and insights derived from the models were then interpreted to determine their implications for policy and operational decision-making within QCPESO. Through this methodological approach, the study sought to demonstrate how data mining and knowledge discovery techniques can transform employment data into actionable insights, ultimately improving the delivery of public employment services and supporting data-informed governance.

## RESULTS AND DISCUSSION

### Data mining of QC PESO data

A total of 2,423 applicant records from the Quezon City Public Employment Service Office (QCPESO) were analyzed using data mining techniques, specifically the decision tree method, to uncover hidden relationships and patterns within employment data. These records contained

information on various attributes such as district location, age, gender, educational attainment, and graduation status. The inclusion of these variables allowed the researcher to examine how different demographic and educational factors influenced employment prospects and job matching in the city. The methodology previously discussed ensured that the data underwent cleaning, integration, and transformation through Online Analytical Processing (OLAP) tools, providing a reliable dataset for classification and pattern recognition.

The analysis began by examining the relationship between district location and educational attainment. The results revealed that across the five districts of Quezon City, the majority of applicants had completed college, followed by elementary and high school graduates. Specifically, District 4 recorded the highest percentage of college graduates at 57.8%, followed by District 3 with 51.9% and District 2 with 49.3%. District 1 showed a moderate proportion of college graduates at 46.0%, while District 5 recorded the lowest at 38.7%. These findings indicate that Districts 3 and 4 have relatively higher concentrations of college-educated applicants, suggesting a potentially stronger pool of candidates for professional or specialized employment opportunities. Meanwhile, the significant number of elementary and high school graduates in Districts 1 and 5 reflects the need for more vocational or skills-based employment programs to address the disparity in educational attainment and improve employability across districts.

When examining the age distribution of applicants per district, the data showed that most applicants fell within the 25–34 age group, comprising 62.5% of the total population. This pattern was consistent across all districts, with District 1 recording the highest proportion in this age range at 66.8%, followed closely by Districts 4 and 5 at 67.7% and 64.4%, respectively. The 15–24 age group represented 24.8% of applicants, indicating a substantial presence of young job seekers likely entering the workforce for the first time. Meanwhile, applicants aged 35–44 constituted only 2.2% of the total, while those aged 45 and above made up less than 11%. This distribution suggests that employment facilitation programs should focus on the 25–34 demographic, which represents the prime working age group, while also creating inclusive employment opportunities for older individuals who may face barriers in re-entering the job market.

Gender distribution analysis revealed that female applicants slightly outnumbered males, accounting for 53.6% of the total population compared to 46.4% males. This pattern was consistent across most districts. District 4 had the highest percentage of female applicants at 57.1%, followed by District 2 with 55.9%, and District 1 with 53.8%. District 3 was the only area where male applicants (53.9%) slightly outnumbered females (46.1%). The predominance of female applicants across districts may indicate a higher level of job-seeking activity among women, possibly reflecting their growing participation in the labor market. This insight is vital for employment service offices as it underscores the importance of developing gender-sensitive employment programs and policies that address specific challenges faced by women, particularly in balancing employment and family responsibilities.

Another significant factor analyzed was graduation status, which provided insights into the educational standing of applicants across districts. The findings revealed that the majority of applicants were graduates, comprising 65.7% of the total population, while undergraduates represented 34.3%. District 5 recorded the highest proportion of graduates at 68.4%, followed closely by District 1 at 67.8%, District 2 at 67.5%, and District 3 at 64.6%. District 4 had the lowest percentage of graduates at 57.8%, indicating a higher concentration of undergraduates or

incomplete degree holders in that area. These results highlight the importance of implementing employment and training programs that cater to both graduates seeking professional roles and undergraduates who may require additional skills development or certification to enhance their employability.

Based on the cross-tabulated data from all variables, a decision tree model was constructed to predict job qualification outcomes based on demographic and educational attributes. The model was developed using the Chi-squared Automatic Interaction Detection (CHAID) growing method, with “graduation status” as the dependent variable and “grade completed,” “district ID,” “gender,” and “age” as the independent variables. Cross-validation was employed to ensure model reliability, with parameters set to a maximum tree depth of three, a minimum of 100 cases per parent node, and 50 cases per child node. The final model produced eight nodes, five of which were terminal nodes, representing the different pathways leading to the classification of applicants based on their likelihood of being qualified for specific job categories.

The results of the decision tree analysis confirmed that age, district location, and educational attainment were the most significant predictors of job qualification within the QCPESO dataset. Younger applicants with higher educational levels, particularly college graduates from Districts 3 and 4, demonstrated stronger employability prospects, while older applicants or those with incomplete education tended to fall into categories requiring further training or upskilling. These findings underscore the value of data mining as a tool for transforming raw employment data into actionable knowledge, enabling employment service offices to make informed, data-driven decisions. By identifying these patterns, QCPESO can better align its employment facilitation programs with the specific needs of its constituents, reduce job mismatch, and ensure more equitable access to employment opportunities across all districts of Quezon City.

## CONCLUSION

The study demonstrated the potential of applying data mining techniques to enhance employment facilitation services within the Quezon City Public Employment Service Office (QCPESO). By systematically analyzing 2,423 applicant records, the research revealed how advanced analytical methods such as rule induction, decision tree analysis, and artificial neural networks can transform large volumes of employment data into meaningful insights for better decision-making. The integration of data mining into the employment service process allows for the identification of patterns, correlations, and predictors that traditional manual analysis cannot easily detect. As a result, data-driven knowledge can now inform strategic interventions, improve program delivery, and reduce job mismatch across multiple employment service facilities.

The findings indicated that educational attainment, age, and district location were the most significant factors influencing employability and job qualification among Quezon City applicants. Districts with higher proportions of college graduates, particularly Districts 3 and 4, showed stronger employability trends, while areas with more elementary and high school graduates, such as Districts 1 and 5, revealed the need for targeted skills enhancement and vocational programs. The data also showed that most applicants fell within the 25–34 age bracket, representing the city’s most active labor segment, and that female applicants slightly outnumbered males, underscoring the growing participation of women in the workforce. These insights confirm the importance of localized and data-informed employment strategies that consider demographic and educational disparities among districts.

Through the application of the Chi-squared Automatic Interaction Detection (CHAID) decision tree model, the study confirmed that the relationships among district, age, gender, and educational background can be used to predict employment readiness and qualification outcomes with a high degree of accuracy. This predictive capacity highlights the value of integrating knowledge discovery and data mining in public sector decision-making, particularly in managing labor market information systems.

Ultimately, the study concludes that data mining offers a practical and effective approach for improving employment facilitation services. By harnessing the power of knowledge discovery, QCPESO and other multi-employment service facilities can move beyond data collection toward evidence-based management and policy formulation. This approach supports the broader objectives of the Department of Labor and Employment (DOLE) under Republic Act No. 8759—to ensure full and equal employment opportunities for all. Strengthening data analytics capacity within PESO offices nationwide will not only enhance operational efficiency but also enable the government to respond more effectively to the evolving dynamics of the labor market in the digital age.

## REFERENCES

- Glymour, C., Madigan, D., Pregibon, D., & Smyth, P. (1996). Statistical inference and data mining. *Communications of the ACM*, 39(11), 35–41. <https://doi.org/10.1145/240455.240479>
- Han, J., & Kamber, M. (2001). *Data mining: Concepts and techniques*. Morgan Kaufmann Publishers.
- Lu, H., Setiono, R., & Liu, H. (1996). Effective data mining using neural networks. *IEEE Transactions on Knowledge and Data Engineering*, 8(6), 957–961. <https://doi.org/10.1109/69.553160>
- Miller, A. (1993). *The application of neural networks to imaging and signal processing in astronomy and medicine (Doctoral dissertation)*. University of Southampton.
- Miller, A., Blott, B., & Hames, T. (1992). Review of neural network applications in medical imaging and signal processing. *Medical & Biological Engineering & Computing*, 30(5), 449–464. <https://doi.org/10.1007/BF02457864>
- Shams, K., & Frashita, M. (2001). Data warehousing toward knowledge management. *Topics in Health Information Management*, 21(3), 25–35.
- Weinstein, J., Kohn, K., Grever, M., Viswanadhan, V. N., Rubinstein, L. V., Monks, A., Vitaro, J., & Shoemaker, R. H. (1992). Neural computing in cancer drug development: Predicting mechanism of action. *Science*, 258(5081), 447–451. <https://doi.org/10.1126/science.1411541>

Frawley, W. J., & Piatetsky-Shapiro, G. (1996). Knowledge discovery in databases: An overview. In G. Piatetsky-Shapiro & W. J. Frawley (Eds.), *Knowledge discovery in databases* (pp. 1–30). AAAI Press.

Bureau of Local Employment, Department of Labor and Employment. (1999). Republic Act No. 8759: Public Employment Service Office (PESO) Act of 1999. Government of the Philippines. <https://ble.dole.gov.ph>

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.

Shmueli, G., Bruce, P. C., Gedeck, P., & Patel, N. R. (2020). *Data mining for business analytics: Concepts, techniques, and applications* in R. Wiley.

Turban, E., Sharda, R., Delen, D., & King, D. (2011). *Business intelligence: A managerial approach* (3rd ed.). Prentice Hall.

Han, J., & Cercone, N. (2000). Rule-based knowledge discovery in databases. *IEEE Transactions on Knowledge and Data Engineering*, 12(6), 959–972. <https://doi.org/10.1109/69.877497>

Liao, S. H., Chu, P. H., & Hsiao, P. Y. (2012). Data mining techniques and applications: A decade review from 2000 to 2011. *Expert Systems with Applications*, 39(12), 11303–11311. <https://doi.org/10.1016/j.eswa.2012.02.063>

Department of Labor and Employment (DOLE). (2024). PESO employment information system: Labor market data and analytics report. DOLE Bureau of Local Employment. <https://ble.dole.gov.ph>

Quezon City Public Employment Service Office. (2024). Employment facilitation data report 2023. Quezon City Government.

Kantardzic, M. (2020). *Data mining: Concepts, models, methods, and algorithms* (3rd ed.). Wiley.

Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill Education.

Han, J., Pei, J., Yin, Y., & Mao, R. (2004). Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data Mining and Knowledge Discovery*, 8(1), 53–87. <https://doi.org/10.1023/B:DAMI.0000005258.31418.83>

Witten, I. H., Frank, E., & Hall, M. A. (2017). *Data mining: Practical machine learning tools and techniques* (4th ed.). Morgan Kaufmann.