



## Implementation of Association Rule Mining Using the FP-Growth Algorithm on Non-Procedural Indonesian Migrant Worker (PMI) Data in South Sulawesi

<sup>a1</sup>Dayini Syahirah

<sup>1</sup>Program Studi Manajemen Teknologi Keimigrasian Politeknik Pengayoman  
Indonesia

<sup>a</sup>Kementerian Hukum Republik Indonesia, Tangerang, Indonesia

[dayinisafar@gmail.com](mailto:dayinisafar@gmail.com)

### ABSTRACT

This study examines patterns in non-procedural Indonesian Migrant Workers (PMI) data in South Sulawesi and utilizes them to strengthen data-driven prevention. The method used is association rule mining using the FP-Growth algorithm within the CRISP-DM framework implemented through Altair AI Studio software. Modeling is run based on a minimum support value of 30% with a minimum confidence of 80%) for the Makassar, Pare-Pare, and Palopo Immigration Office datasets. Patterns are retained if the lift value is  $> 1$  and selecting the top 10 patterns for each dataset. The results show consistent frequent itemsets and association rules indicating a general pattern of non-procedural PMI dominated by adult males with destinations in Malaysia with illegal/undocumented issues. The findings can be used as a preventive measure in strengthening interviews and document verification in the passport issuance process and the Immigration fostered village program. The study confirms that the application of FP-Growth with support, confidence, and lift evaluations provides evidence-based insights relevant to a more targeted and effective non-procedural PMI prevention policy by the Immigration Office.

**Keywords:** *Non-Procedural PMI, Association Rule Mining, FP-Growth, Support, Confidence, Lift, CRISP-DM, Immigration*

### INTRODUCTION

The rapid development of information technology in recent decades has changed the way people work and access information in various sectors, including government. Digital transformation requires government agencies to manage data optimally as a basis for evidence-based policy making to improve public services.

Data is now a strategic asset that can be processed to find patterns and trends through data mining techniques. This approach allows the government to identify relationships between attributes in large datasets, thus enabling them to formulate more targeted policies. One relevant technique is association rule mining, which focuses on discovering relationships between attributes in data and has proven effective in analyzing social patterns and public services.

One social issue requiring a data-driven approach is the case of non-procedural Indonesian Migrant Workers (PMI). According to BP2MI data, 110,640 non-procedural PMI were deported from various countries between 2020 and 2024.[5] This

issue not only concerns legal aspects and citizen protection but also poses challenges for immigration agencies in preventing illegal departures. South Sulawesi Province is one of the regions with the highest number of non-procedural repatriations of Indonesian migrant workers (PMI), with 2,681 people repatriated from 2022 to 2024. Most of them come from Bulukumba, Pinrang, and North Toraja districts, with Malaysia as the dominant destination. Immigration offices in Makassar, Pare-Pare, and Palopo play a crucial role in ensuring that every prospective PMI departs through legal channels.

In supporting prevention, the Directorate General of Immigration has implemented the Immigration Assisted Village Program as an early warning system.[6]. This program focuses on community development in high-risk areas. The program's effectiveness depends on the accuracy of determining priority villages for development. Therefore, the application of the association rule mining method using the FP-Growth algorithm is relevant. This algorithm is capable of identifying significant relationships between attributes without repeated candidate generation, making it more efficient in processing large data (Lathifah, 2023). Through non-procedural PMI pattern analysis, the results are expected to help Immigration identify risk factors, determine priority development areas, and strengthen the passport interview process.

## LITERATURE REVIEW

### Non-Procedural Indonesian Migrant Workers

Non-procedural PMI are individuals who go to work abroad without going through the official permit process from the Indonesian government or authorized agencies. In an effort to prevent cases of non-procedural migrant workers (PMI) that have the potential to lead to transnational crime, the presence of the state is crucial as a form of responsibility to protect citizens abroad. This situation makes them vulnerable to human trafficking and exploitation.

### *Association Rule Mining*

*Association rule* is a data mining technique for finding associations and relationships between large sets of data items.[9]. In association rule mining, evaluation of relationships between attributes generally uses three main metrics, namely support, confidence, and lift ratio [10]. Support measures the frequency of occurrence of items, confidence indicates the strength of the association, and lift ratio assesses the significance of the relationship between attributes.

### FP-Growth Algorithm

The FP-Growth algorithm is a development of the Apriori algorithm, a pioneer in association analysis. This algorithm is suitable for large datasets because it processes quickly and produces more comprehensive results. By building a Frequent Pattern Tree (FP-Tree) structure, FP-Growth is able to find patterns without generating repeated candidates.

### CRISP-DM

CRISP-DM (Cross Industry Standard Process for Data Mining) is a standard framework for conducting data mining-based research. This model consists of six main stages: business understanding, data understanding, data preparation,

modeling, evaluation, and deployment. This approach ensures that each analysis process runs systematically and is in line with the research objectives.

## RESEARCH METHODS

This study employed a quantitative approach with association rule mining using the FP-Growth algorithm. The data used were non-procedural migrant worker repatriation data from the South Sulawesi Migrant Workers Agency (BP3MI) for the 2022–2024 period. The analysis was conducted to identify patterns that could support immigration policies in preventing non-procedural migrant worker departures.

The research procedure consists of four main stages, namely introduction, literature study, application of the CRISP-DM model, and interpretation of results. The preliminary stage includes determining the topic, problem formulation, objectives, and benefits of the research. The literature study stage is carried out to strengthen the theoretical basis through a review of relevant literature. In the main stage, the research follows the CRISP-DM framework which includes: (1) business understanding to determine the objectives of preventing non-procedural migrant workers, (2) data understanding to recognize the structure and characteristics of data, (3) data preparation through data cleaning and transformation, (4) modeling with the FP-Growth algorithm, (5) evaluation to assess results using lift ratio values, and (6) deployment to interpret results into policy recommendations. The final stage is the interpretation of results and drawing conclusions, where the patterns formed are used as the basis for compiling policy recommendations for immigration agencies, especially in preventing the departure of non-procedural migrant workers.

The attributes used included gender, age, region of origin, destination country, and problem. The analysis parameters were set at a minimum support value of 30% and a confidence level of 80%. Patterns with a lift ratio greater than 1 were considered significant and retained for further analysis.

## RESULTS AND DISCUSSION

The FP-Growth analysis yielded a number of consistent patterns of inter-attribute associations across three Immigration Office work areas in South Sulawesi: the Makassar Class I Special TPI Immigration Office, the Pare-Pare Class II TPI Immigration Office, and the Palopo Class II Non-TPI Immigration Office. The minimum support parameter was set at 30% and the confidence at 80%, with the lift ratio used to assess the strength of the relationship between attributes. From the processing results using Altair AI Studio, several association patterns were obtained with support, confidence, and lift ratio values that met the threshold. The following are the main results of the FP-Growth analysis:

Table 1. Summary of Pattern Results for Each Scenario

| No. | Number of Records | Number of Attributes | Data Subset                     | Min Support | Min Confidence | Number of Patterns |
|-----|-------------------|----------------------|---------------------------------|-------------|----------------|--------------------|
| 1   | 1,066             | 23                   | Districts/cities in the working | 95%<br>30%  | 80%<br>80%     | 0<br>49            |

|   |     |    | area of the<br>Makassar<br>Immigration<br>Office                                         |     |     |    |
|---|-----|----|------------------------------------------------------------------------------------------|-----|-----|----|
| 2 | 497 | 20 | Districts/cities<br>in the working<br>area of the Pare-<br>Pare<br>Immigration<br>Office | 95% | 80% | 0  |
|   |     |    |                                                                                          | 30% | 80% | 61 |
| 3 | 207 | 16 | Districts/cities<br>in the working<br>area of the<br>Palopo<br>Immigration<br>Office     | 95% | 80% | 0  |
|   |     |    |                                                                                          | 30% | 80% | 69 |

The lift ratio value is used as the main parameter to assess the strength of the relationship between items. A lift value greater than 1 indicates a significant relationship, so the rule with the highest value is prioritized for interpreting the results.[15].

Table II. Results of association rules with the highest lift ratio values in the work area dataset of the Makassar Class I Special Immigration Office TPI

| No. | Premises                                                       | Conclusion                                                                  | Support | Confidence | Elev<br>ator |
|-----|----------------------------------------------------------------|-----------------------------------------------------------------------------|---------|------------|--------------|
| 1   | Regency/City =<br>Bulukumba                                    | Destination Country<br>= Malaysia, Problem<br>=<br>Illegal/Undocument<br>ed | 31.6%   | 94.7%      | 1,052        |
| 2   | Regency/City =<br>Bulukumba                                    | Problem =<br>Illegal/Undocument<br>ed                                       | 31.6%   | 94.7%      | 1,046        |
| 3   | Destination Country =<br>Malaysia, Regency/City<br>= Bulukumba | Problem =<br>Illegal/Undocument<br>ed                                       | 31.6%   | 94.7%      | 1,046        |
| 4   | Gender = M                                                     | Destination Country<br>= Malaysia, Problem<br>=<br>Illegal/Undocument<br>ed | 70.2%   | 91.2%      | 1,014        |
| 5   | Regency/City =<br>Bulukumba                                    | Destination Country<br>= Malaysia, Age =<br>Adult                           | 30.9%   | 92.4%      | 1,014        |

Table III. Results of association rules with the highest lift ratio values in the Class II TPI Pare-Pare Immigration Office work area dataset

| No. | Premises                                                                    | Conclusion                                                                  | Support | Confidence | Elev<br>ator |
|-----|-----------------------------------------------------------------------------|-----------------------------------------------------------------------------|---------|------------|--------------|
| 1   | Destination Country = Malaysia, Age = Adult, Problem = Illegal/Undocumented | Gender = M                                                                  | 66.4%   | 80.7%      | 1,028        |
| 2   | Gender = M                                                                  | Destination Country = Malaysia, Age = Adult, Problem = Illegal/Undocumented | 66.4%   | 84.6%      | 1,028        |
| 3   | Destination Country = Malaysia, Problem = Illegal/Undocumented              | Gender = M                                                                  | 72.2%   | 80.1%      | 1,021        |
| 4   | Gender = M                                                                  | Destination Country = Malaysia, Problem = Illegal/Undocumented              | 72.2%   | 92.1%      | 1,021        |
| 5   | Gender = M, Regency/City = Pinrang                                          | Destination Country = Malaysia, Problem = Illegal/Undocumented              | 32.4%   | 92%        | 1,021        |

Table IV. Results of association rules with the highest lift ratio values in the Class II Non-TPI Palopo Immigration Office work area dataset

| No. | Premises                                | Conclusion                                                                  | Support | Confidence | Elev<br>ator |
|-----|-----------------------------------------|-----------------------------------------------------------------------------|---------|------------|--------------|
| 1   | Gender = M, Regency/City = North Toraja | Destination Country = Malaysia, Problem = Illegal/Undocumented, Age = Adult | 38.2%   | 84.9%      | 1,072        |
| 2   | Regency/City = North Toraja             | Destination Country = Malaysia, Problem = Illegal/Undocumented, Age = Adult | 56%     | 84.1%      | 1,061        |
| 3   | Gender = M, Regency/City = North Toraja | Problem = Illegal/Undocumented, Age = Adult                                 | 38.2%   | 84.9%      | 1,059        |

| No. | Premises                                                                | Conclusion                                  | Support | Confidence  | Elev<br>ator |
|-----|-------------------------------------------------------------------------|---------------------------------------------|---------|-------------|--------------|
| 4   | Destination Country = Malaysia, Gender = M, Regency/City = North Toraja | Problem = Illegal/Undocumented, Age = Adult |         | 38.2% 84.9% | 1,059        |
| 5   | Problem = Illegal/Undocumented, Gender = M, Regency/City = North Toraja | Destination Country = Malaysia, Age = Adult |         | 38.2% 91.9% | 1,051        |

Overall, the analysis shows that the dominant pattern of non-procedural migrant workers across the region is productive-age males, with Malaysia as the destination, and the primary problem being illegal departures. These findings can serve as a basis for Immigration to strengthen data-driven prevention measures and determine priority surveillance areas.

## CONCLUSION

This study shows that non-procedural migrant workers in South Sulawesi are dominated by adult males with destinations in Malaysia and illegal or undocumented migrants. Lift ratios above 1 in each region indicate a significant relationship between attributes. The application of the FP-Growth algorithm has proven effective in identifying these patterns and can be used as a basis for immigration agencies' policies to prevent non-procedural migrant workers.

## REFERENCE

- A. Lathifah, "Cross-Industry Standard Process For Data Mining (CRISP-DM) untuk Menemukan Pola Asosiasi pada Data Tracer Study Lulusan Perguruan Tinggi Keagamaan Islam (Studi Kasus: UIN Syarif Hidayatullah Jakarta)," UIN Syarif Hidayatullah Jakarta, Jakarta, 2023.
- A. R. Muttaqin, A. Wibawa, and K. Nabila, "Inovasi Digital untuk Masyarakat yang Lebih Cerdas 5.0: Analisis Tren Teknologi Informasi dan Prospek Masa Depan," *Jurnal Inovasi Teknologi dan Edukasi Teknik*, vol. 1, no. 12, pp. 880–886, Dec. 2021, doi: 10.17977/um068v1i122021p880-886.
- Amna *et al.*, *DATA MINING*. Padang: PT Global Eksekutif Teknologi, 2023. [Online]. Available: [www.globaleksekutifteknologi.co.id](http://www.globaleksekutifteknologi.co.id)
- BP2MI, "Siaran Pers Komnas Perempuan tentang Peringatan Hari Migran Sedunia 2024," BP2MI. Accessed: Apr. 18, 2025. [Online]. Available: <https://komnasperempuan.go.id/siaran-pers-detail/siaran-pers-komnas-perempuan-tentang-peringatan-hari-migran-sedunia-2024>
- D. Winarti *et al.*, "Penerapan Data Mining untuk Analisa Tingkat Kriminalitas Dengan Algoritma Association Rule Metode FP-Growth," *Jurnal SIMTIKA*, vol. 4, no. 3, 2021.
- E. Buulolo, *Data Mining untuk Perguruan Tinggi*. Yogyakarta: Deepublish, 2020.

- F. Ferdin Bakker and T. Mirwanto, "Kontribusi Peranan Imigrasi Indonesia Dalam Mencegah Dan Memberikan Perlindungan Ham Terhadap Adanya Pekerja Migran Non-Prosedural (Pmi-Np) Dari Kejahatan Transnasional (Contribution Of The Role Of Indonesian Immigration In Preventing And Protecting Human Rights Against Non-Procedural Migrant Workers (Pmi-Np) From Transnational CRIMES)," 2021. [Online]. Available: <https://imigrasi.go.id/berita/detail/humas->
- F. Sulianta, *Basic Data Mining From A to Z*. Bandung, 2023. [Online]. Available: <https://www.researchgate.net/publication/382274667>
- F. Sulianta, *Dasar dan Konsep Algoritma & Pemrograman*. Bandung: Feri Sulianta, 2025.
- J. Magakwe, "Advancing Governance: Role of Data Analytics in Driving Evidence-Based Decision-Making in Public Administration," in *Recent Advances in Public Sector Management*, IntechOpen, 2025. doi: 10.5772/intechopen.114901.
- K. Ananda Mustari, P. Assiroj, B. Hartati, and F. Samuel, "Implementasi Data Mining Pada Instansi Pemerintahan (Systematic Literature Review)," 2024.
- Kantor Imigrasi Kelas I Khusus TPI Batam, "Kantor Imigrasi Batam Gelar Penyuluhan Desa Binaan Imigrasi Tahun 2024 Terkait Pencegahan Tindak Pidana Perdagangan Orang," Kantor Imigrasi Kelas I Khusus TPI Batam. Accessed: Apr. 19, 2025. [Online]. Available: <https://batam.imigrasi.go.id/berita/2024/06/kantor-imigrasi-batam-gelar-penyuluhan-desa-binaan-imigrasi-tahun-2024-terkait-pencegahan-tindak-pidana-perdagangan-orang>
- M. N. P. Ma'ady, B. R. Lidiawaty, A. Dzulkarnain, and A. Ramadan, *Data Mining: Algoritma Dan Contoh Perhitungan Matematis*. Yogyakarta: Deepublish, 2024.
- P. W. Rahayu et al., *BUKU AJAR DATA MINING*. PT Sonpedia Publishing Indonesia, 2024. [Online]. Available: <https://www.researchgate.net/publication/377415198>
- S. Santoso, F. W. Perdana, I. Irwan, B. Setiawan, and P. Purboyo, "Penguatan Fungsi Keimigrasian dalam Rangka Pencegahan Tindak Pidana Perdagangan Orang (TPPO) dalam Pengiriman Buruh Migran Non Prosedural di Wilayah Perbatasan," *Jurnal Indonesia Sosial Teknologi*, vol. 3, no. 02, pp. 333–341, Feb. 2022, doi: 10.36418/jist.v3i2.375.