JURNAL TEKNOLOGI DAN OPEN SOURCE
Vol. 9, No. 1, June 2026, pp. 242~254
e-ISSN: 2622-1659, accredited Four Grade by Kemenristekdikti, Decree No: 152/E/KPT/2023
DOI: https://doi.org/10.36378/jtos.v9i1.5674

Divorce Pattern Clustering Using The K-Prototypes

Ahsin Ilallah[1], Zaehol Fatah[2], Achmad Baijuri[3]
[1,2,3] Departemen Of Information Systems, Science And Technology, Ibrahimy University, Indonesia.

Article Info

Article history:

Received 05 23, 2026
Revised 06 02,2026
Accepted 06 20, 2026

Keywords:

Clustering
K-Prototypes
Perceraian
Banyuwangi

ABSTRACT

Divorce is a complex social problem that continues to increase every year, including within the jurisdiction of the Banyuwangi Religious Court. Various factors such as early-age marriage, economic problems, prolonged conflict, and domestic violence are the main triggers of divorce. This study aims to classify divorce patterns of the Banyuwangi community using the K-Prototypes clustering method, which is capable of handling mixed numerical and categorical data. The data used are secondary data from the Case Tracking Information System (SIPP) of the Banyuwangi Religious Court in 2025, totaling 5,570 cases. The clustering process was carried out through preprocessing stages, determining the gamma parameter, data encoding, and implementing the K-Prototypes algorithm using Python and the Streamlit framework. The results show the formation of two main clusters. Cluster 0 (59.7%) is dominated by young couples with an average plaintiff age of 29.8 years, short marriage duration of 6.3 years, and 0.7 children. Cluster 1 (40.3%) is dominated by mature couples with an average plaintiff age of 45.5 years, long marriage duration of 17.4 years, and 1.2 children. The developed SKPP application successfully facilitates users in uploading data, preprocessing, clustering, and exporting results. Therefore, the K-Prototypes method is effective for analyzing divorce patterns and can serve as a supporting tool for the Religious Court in formulating more targeted prevention policies.

Corresponding Author:

Ahsin Ilallah
Departemen Of Information Systems, Science And Technology
Ibrahimy University
Situbondo, Indonesia
Email: ahsinilallah15@gmail.com

1. Introduction

The religious court is a court of first instance charged eith receiving, examining, and adjudicating every request or claim at the earliest and most fundamental stage [1]. Over a certain period, hundreds of divorce cases were recorded as having been filed with the religious courts [2]. Divorce may be understood as the dissolution of the bond between a man and a woman as established by law or religion arising from a complete absence of mutual attraction, trust, and compatibility, which consequently leads to disharmony within the the household [3]. Naturally, divorce is driven by several influencing factors. These include husbands' abusive behavior toward their wives, early marriage, economic issues, and prolonged disputes. Consequently, the Banyuwangi Religious Court faces difficulties in decision-making as the number of divorces continues to increase month after month and year after year.

The phenomenon of divorce has become a serious concern at both global and national levels. United Nations (UN) demographic data from 2025 indicates that the highest divorce rates in the world occur in

Russia (73%), the United States (58%), and China (58%) [4]. In the European region, Portugal ranks first with a divorce rate reaching 55%, followed by the Czech Republic (53%) and Hungary (51%) [5]. Countries such as Spain (46%), France (45%), and Germany (42%) also recorded significant divorce rates [6]. This fact demonstrates that marital instability is a universal issue that transcends cultural, religious, and economic boundaries [7].

In the Asian region, divorce trends also show an alarming increase. South Korea recorded a divorce rate of 48%, Japan 35%, and India 22%, despite traditionally having a strong culture of marriage [8]. In Indonesia, the divorce situation is no less concerning. Based on data from the Religious Courts Agency of the Supreme Court of the Republic of Indonesia, as reported by the Central Bureau of Statistics (BPS), there were 394,608 divorce cases in 2024 [9]. This figure, although representing a 14.9% decrease from 463,654 cases in 2023, is still considered very high within the historical context of religious courts in Indonesia [10]. An even more striking fact is that 78% of the total divorces (308,956 cases) were Cerai Gugat (fault divorce/divorce by female petition), which are lawsuits filed by the wives. This phenomenon indicates a dynamic shift in gender relations and domestic conflict resolution patterns in Indonesia. [11]. The Minister of Religious Affairs of the Republic of Indonesia stated in his annual report that approximately 40% of divorces occur among couples married for less than five years, with the primary factors being economic problems, infidelity, and incompatibility [12]. Furthermore, the Ministry of Women's Empowerment and Child Protection noted that 25% of divorce cases are triggered by domestic violence (KDRT), which is predominantly experienced by women of productive age [13].

Certainly, divorce is driven by several contributing factors. These factors include economic problems; financial issues are the most common cause of divorce in Indonesia. BPS data (2025) shows that 45% of divorces are triggered by economic instability, including unemployment, irregular income, or accumulating debt. Couples with incomes below the regional minimum wage have a 2.5 times higher risk of divorce compared to those with stable incomes. [14]. Secondly, education and the age of marriage act as major contributing factors. Early marriage (under the age of 19) contributes to 30% of divorce cases in Indonesia. A low level of education, particularly among women, also correlates positively with divorce rates [15].

Thirdly is the factor of continuous disputes and arguments. Prolonged conflict without proper resolution is the primary cause of divorce in the *Cerai Gugat* (divorce by female petition) category. This is because a lack of effective communication and the inability to manage conflict serve as strong predictors of divorce, regardless of the couple's economic background. [16]. And fourthly is domestic violence (KDRT). Although often hidden, domestic violence serves as a significant factor in filing for *Cerai Gugat* (divorce by female petition). Komnas Perempuan (2025) noted that 65% of domestic violence cases were reported by wives aged 25-35, with psychological violence being the most dominant form. [13]. And there are many other contributing factors to divorce. Consequently, the Banyuwangi Religious Court faces difficulties in decision-making as the number of divorces continues to rise from month to month and year to year.

One approach that can be applied is the utilization of data mining with clustering algorithms to systematically analyze and group divorce patterns [17]. Data mining has been recognized since the 1990s, a period when information utilization became highly crucial in various fields, including business and marketing, engineering and science, as well as entertainment and art. Its primary objective is to identify previously unknown patterns. [18]. In data mining, the clustering method is a reliable approach to categorize data into classes based on specific shared characteristics. [19]. This method is also referred to as data segmentation in several other applications, as clustering divides a large dataset into smaller groups. [20].

The K-Prototypes method is a clustering algorithm used for mixed data type distances. [21]. This algorithm was first introduced by Huang (1998) and has undergone various refinements. [22]. The K-Prototypes algorithm works by calculating the Euclidean distance for numerical data and frequency matching for categorical data, thereby allowing it to cluster data based on varied characteristics. [23]. These two distances are combined using a balancing parameter called gamma (γ) [24]. Therefore, K-Prototypes is highly suitable for clustering divorce patterns.

The primary objective of this study is to cluster divorce patterns using the K-Prototypes clustering method on divorce data of the Banyuwangi community.

2. Research Method

2.1. Research Type

This study is a quantitative research that utilizes a data mining approach with clustering techniques to identify divorce patterns. The method employed is K-Prototypes, which is capable of handling mixed data.

2.2. Data Source

The data used in this study constitutes secondary data. It was obtained from an interview with Nazrul Rachmadi, a member of the IT Planning and Reporting department. The dataset covers the year 2025.

The accessible link for this data is https://sipp.pa-banyuwangi.go.id/. This dataset contains plaintiffs' cases recorded in the Case Tracking Information System (SIPP).

Table 1. 2025 Divorce Data Table

No	Case Type	Plaintiff Age	Defendant Age	Plaintiff Occupation	Defendant Occupation	Number of Children	Marriage Duration ((Years	Plaintiff Education	Defendant Education	Divorce Factor
1	Divorce Lawsuit	32	45	Self-employed	Self-employed	2	15	Elementary school	Upper secondary school	Economic issues
2	Divorce by Husband's Initiation	23	24	Welding workshop employee	Housewife	1	4	Elementary school	Elementary school	Economic issues
3	Divorce by Husband's Initiation	40	36	Daily casual laborer	Housewife	0	14	Elementary school	Elementary school	Continuous disputes and quarrels

2.3. Research Variables

The research variables were determined based on the availability of data in the Case Tracking Information System (SIPP) of the Banyuwangi Religious Court for the 2025 period. The variables used consist of 10 variables, which are divided into 4 numerical variables and 6 categorical variables.

Table 2. Numerical & Categorical Variables Table

No	variable name	data type	scale	description
1	Plaintiff age	Numerical	Ratio	Age of the party filing for divorce
2	Defendant age	Numerical	Ratio	Age of the party being divorced
3	Number of children	Numerical	Ratio	Number of children from the marriage
4	Marriage duration	Numerical	Ratio	Duration of marriage until divorce
5	Case type	Categorical	Nominal	Type of divorce filed
6	Plaintiff occupation	Categorical	Nominal	Occupation of the plaintiff
7	Defendant occupation	Categorical	Nominal	Occupation of the defendant
8	Plaintiff	Categorical	Ordinal	Highest

	education			education level of the plaintiff
9	Defendant education	Categorical	Ordinal	Highest education level of the defendant
10	Divorce factor	Categorical	Nominal	Main cause of divorce

2.4. Data Collection Techniques

Data collection in this study was conducted through two techniques: interviews and documentation.

2.4.1. Interviews

An in-person interview was conducted with Mr. Nazrul Rachmadi, S.Kom., a member of the IT Planning and Reporting staff specializing in server management at the Banyuwangi Religious Court. The objectives of the interview included:

1. Understanding the organizational structure and the division of work units at the Banyuwangi Religious Court.
2. Identifying the applications utilized in managing case data.
3. Discerning the mechanisms for accessing and extracting data from the SIPP (Case Tracking Information System).
4. Obtaining permission and technical guidance for research data collection.

2.4.2. Documentation

Following the directives from the informant, the documentation technique was carried out through the following steps:

1. System Access: Accessing the official SIPP website of the Banyuwangi Religious Court via the link https://sipp.pa-banyuwangi.go.id/.
2. Data Selection: Selecting the specific data required for the research.
3. Data Extraction: Exporting the data that met the criteria into CSV or PDF file formats as needed.
4. Data Table Compilation: Organizing the obtained data into structured tables.

2.5. Population and Sample

2.5.1. Population

All divorce cases recorded in the Case Tracking Information System (SIPP) of the Banyuwangi Religious Court in 2025 constitute the data scope of this study. Based on the collected results, the total population of registered divorce cases is 5,570 cases.

Table 3. Population

Case type	Total cases	Percentage
Divorce Lawsuit	4.435 cases	79,6%
Divorce by Husband's Initiation	1.135 cases	20,4%
Total	5.570 cases	100%

2.5.2. Sample

The sampling mechanism applied is total sampling (saturated sampling), meaning that the entire population is utilized as the research sample. The use of total sampling is based on the following considerations:

1. The population size (5,570 cases) is still computationally feasible to process using the available hardware specifications.
2. Avoiding potential bias that might arise if only a subset of the data were taken.
3. Maximizing the representation of diverse divorce patterns in the Banyuwangi region. All data has been validated and is ready for analysis.

2.6. Data Preprocessing

Prior to clustering using the K-Prototypes method, the obtained raw data must undergo a preprocessing stage to ensure the data is ready for analysis.

3.6.1. Data Cleaning

Data cleaning is the initial preprocessing step that identifies and corrects inaccuracies, incompleteness, and inconsistencies in the raw data. This process includes removing duplicate records, correcting formatting errors, and validating data ranges. Data cleaning is essential to prevent calculation errors, avoid algorithmic bias, and produce accurate clusters. Without proper cleaning, the K-Prototypes algorithm may generate misleading results, as it cannot process corrupted or inconsistent data effectively.

3.6.2. Handling Missing Value

The next step is to handle missing values within the dataset. This treatment is highly critical because the K-Prototypes algorithm cannot process data containing missing values. In this study, rows with missing values were removed entirely (listwise deletion) since only two missing values were detected from 5,570 rows. This approach preserves data integrity without introducing bias. Alternatively, imputation methods such as mean or mode substitution could be used for larger missing data proportions.

3.6.3. Gamma Parameter

The K-Prototypes algorithm possesses a distinct mechanism to balance the contribution between numerical and categorical variables through the gamma parameter (γ).

$$D(x,y) = \underbrace{\sum_{j=1}^{p}(x_j - y_j)^2}_{Euclidean\ Distance(Numerical)} + \gamma \underbrace{\sum_{j=p+1}^{m}\delta(x_j,y_j)}_{Simple\ Matching\ Distance\ (Categorical)} \quad (1)$$

Where :

1. p = number of numerical variables
2. m = total number of variables
3. γ = balancing parameter (weight)
4. $\delta(x_j, y_j)$ = 0 if the values are identical, 1 if they differ.

The γ value in this study was determined using a grid search method within the range of [0.1, 0.2, ..., 1.0] to achieve optimal clustering results. Consequently, numerical data (plaintiff's age, defendant's age, number of children, marriage duration) were used in their original scale, while categorical data were processed using the matching method.

3.6.4. Categorical Data Encoding

Similar to all computational algorithms, the K-Prototypes algorithm cannot process data in text format directly. Therefore, an encoding process is required to convert categorical data into a numerical representation (numbers) without losing its informational meaning.

Based on the measurement scale type, categorical data is divided into two types: ordinal and nominal. For ordinal variables such as the plaintiff's and defendant's education, the Label Encoding method was used with the following mapping: Elementary School (SD)=1, Junior High School (SMP)=2, Senior High School (SMA)=3, Diploma=4, Bachelor's Degree (S1)=5, Master's Degree (S2)=6, and No Schooling (Tidak Sekolah)=0. For nominal variables such as Case Type, plaintiff's and defendant's occupation, and Divorce Cause, the One-Hot Encoding method was applied, which generates separate binary columns for each unique category. Following the encoding process, all data has been transformed into numerical representations and is ready for clustering with the K-Prototypes algorithm.

3.6.5. Summary of Preprocessing Results

After undergoing all preprocessing stages, which included data cleaning, handling missing values, determining the gamma parameter, and encoding categorical data, a summary was obtained indicating that the data is fully prepared for clustering analysis.

3. Result and Discussion

The results of this study explain the application of the K-Prototypes algorithm based on divorce case data obtained from the Banyuwangi Religious Court through the Case Tracking Information System (SIPP). The clustering process was conducted using the Python programming language. This study aims to present the facts uncovered from the data, providing factual information regarding the actual divorce patterns occurring in the Banyuwangi region.

3.1. Dashboard interface

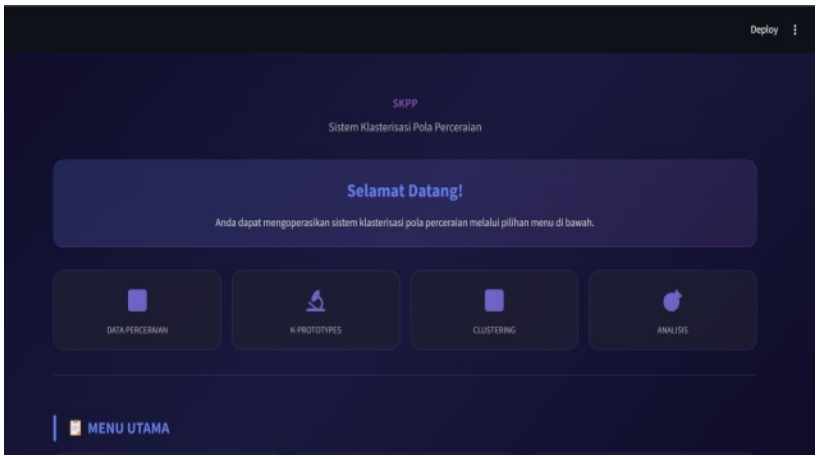


Figure 1. Dashboard interface

The SKPP (Divorce Pattern Clustering System) dashboard, as shown in Figure 1, is designed using Streamlit, a web-based framework specifically developed for interactive Python data applications. This dashboard serves as the main gateway that connects users with all the clustering analysis features available within the SKPP system. Upon accessing the application, users are greeted with a clean and intuitive interface that presents four main menu options: "Data Perceraian" (Divorce Data), "K-Prototypes Clustering," "Analysis," and "Deploy." Each menu is strategically organized to guide users through the entire clustering workflow, starting from data upload, preprocessing, clustering algorithm execution, result visualization, to final data export. The dashboard's responsive design ensures that users, including court administrators without programming backgrounds, can navigate the system easily. Thus, this interface significantly reduces the technical barriers typically associated with implementing machine learning algorithms, making divorce pattern analysis more accessible and practical for the Banyuwangi Religious Court.

3.2. Menu utama interface

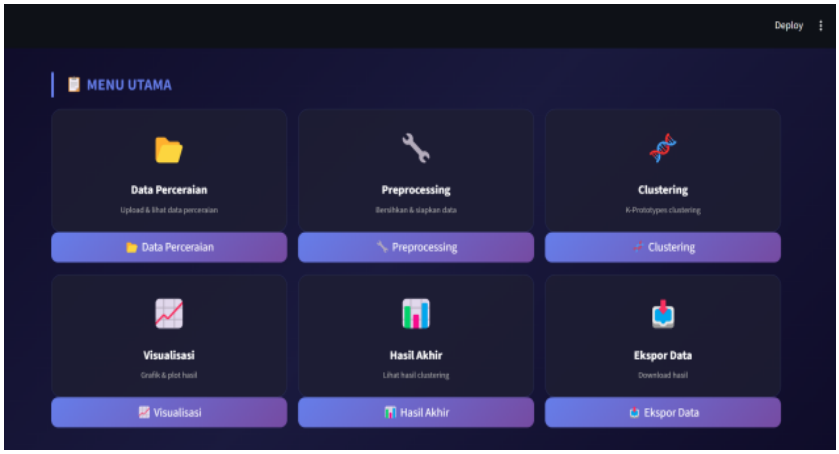


Figure 2. Menu utama interface

Figure 2 displays the main menu interface of the Divorce Pattern Clustering System (SKPP) application. This main menu serves as the navigation hub that assists users in seamlessly accessing the various features available within the system. The menu is organized into six key sections, guiding users through a complete clustering workflow. The first section, "Data Perceraian" (Divorce Data), allows users to upload and edit the divorce dataset. The second section, "Preprocessing," is designed to clean and prepare the raw data before clustering. The third section, "Clustering," executes the K-Prototypes algorithm to group divorce cases based on their characteristics. The fourth section, "Visualisasi" (Visualization), generates graphs and plots to present the clustering results graphically. The fifth section, "Hasil Akhir" (Final Results), displays the complete clustering outcomes with assigned cluster labels. Finally, the sixth section, "Ekspor Data" (Export Data), enables users to download the clustering results in CSV or Excel format for further analysis. This structured menu design ensures that even users without technical backgrounds can follow the clustering process step by step, from data upload to final export, making the SKPP application a practical and user-friendly tool for the Banyuwangi Religious Court.

3.3. Divorce data upload interface

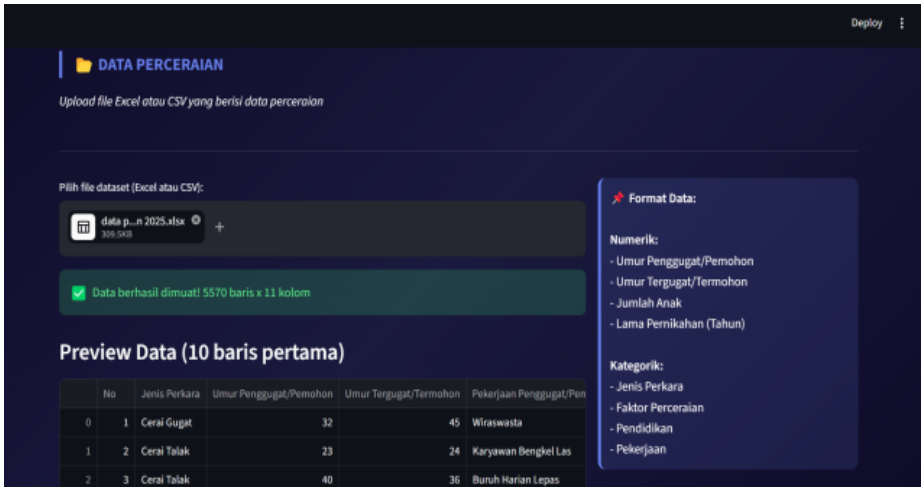


Figure 3. Divorce data upload interface

Figure 3 displays the data upload interface of the SKPP application. In this view, users are provided with the option to upload the divorce dataset file in Excel or CSV format. The interface shows the selected file name, in this case "data p_n 2025.xlsx" with a file size of 205.5 KB. Once the file is uploaded, the system automatically reads and validates the data. A successfully uploaded file is indicated by a green notification with the message "Data Berhasil Dimuat! 5570 baris x 11 kolom" (Data Loaded Successfully! 5570 rows x 11 columns), confirming that the data file has been correctly read. The interface also displays a preview of the first 10 rows and categorizes the variables into numerical (plaintiff age, defendant age, number of children, marriage duration) and categorical (case type, divorce factor, education, occupation), allowing users to verify the data before proceeding to preprocessing.

3.4. data preprocessing interface

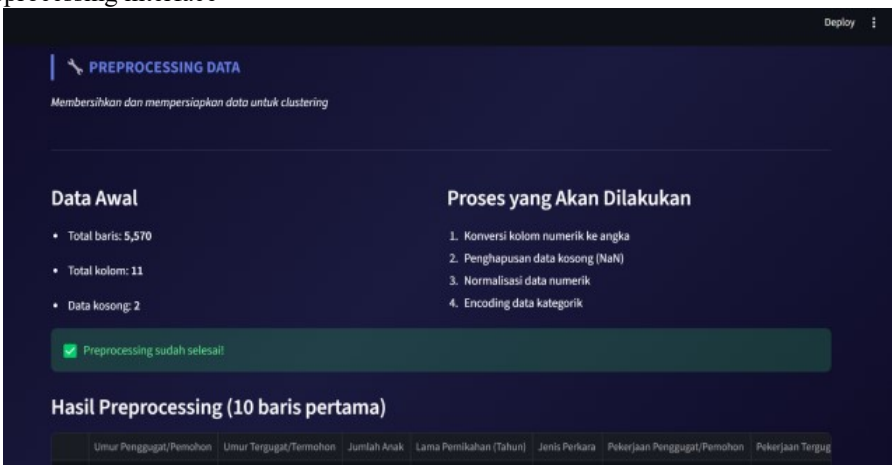


Figure 4. data preprocessing interface

Figure 4 illustrates the data preprocessing interface of the SKPP application. Once the divorce dataset has been successfully uploaded into the system, the next critical step is Data Preprocessing. This process functions to clean and prepare the raw data before the clustering process is executed using the K-Prototypes algorithm. As shown in the interface, the system initially displays key information about the raw dataset, including a total of 5,570 rows, 11 columns, and 2 missing values. The preprocessing workflow consists of four sequential steps: (1) converting numerical columns to numeric data types, (2) removing missing values (NaN), (3) normalizing numerical data to ensure all variables are on a comparable scale, and (4) encoding categorical data into numerical representations using Label Encoding and One-Hot Encoding methods. Once the preprocessing is completed, the system displays a success notification and shows the first 10 rows of the preprocessed dataset. This allows users to verify that the data is now clean, complete, and

ready for clustering analysis. By automating these preprocessing steps, the SKPP application significantly reduces manual effort and minimizes the risk of errors that could affect the accuracy of the clustering results.

	Umur Penggugat/Pemohon	Umur Tergugat/Termohon	Jumlah Anak	Lama Pernikahan (Tahun)	Jenis Perkara	Pekerjaan Penggugat/Pemohon	Pekerjaan Tergu
0	32	45	2	15	Cerai Gugat	Wiraswasta	Wiraswasta
1	23	24	1	4	Cerai Talak	Karyawan Bengkel Las	mengurus ruma
2	40	36	0	14	Cerai Talak	Buruh Harian Lepas	Ibu Rumah Tang
3	22	33	1	2	Cerai Gugat	mengurus rumah tangga	karyawan swast
4	43	36	3	10	Cerai Talak	Karyawan Swasta	Mengurus Rumа
5	30	28	1	8	Cerai Talak	Swasta TKI	Swasta
6	41	46	0	1	Cerai Gugat	mengurus rumah tangga	Pegawai Negeri
7	34	33	2	10	Cerai Talak	Karyawan Swasta	mengurus ruma
8	61	56	0	6	Cerai Talak	pensiunan Guru	Guru PNS
9	45	40	2	21	Cerai Talak	Penjahit	Asisten Rumah

Figure 5. Preprocessed Dataset

Figure 5 displays the first 10 rows of the dataset after successful preprocessing. The table shows that the data is now clean, free from missing values, and uniformly formatted. All numerical columns have been properly converted, and categorical columns are ready for encoding. Therefore, the user can confidently proceed to the clustering stage.

3.5. K-Prototypes Clustering Process Interface

3.6.

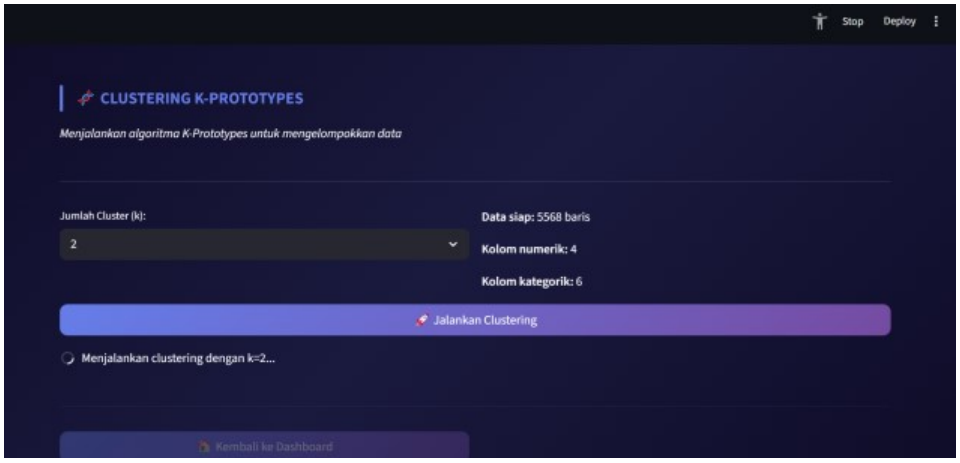


Figure 6. K-Prototypes Clustering Process Interface

Figure 6 presents the K-Prototypes clustering process interface within the SKPP application. This page is the core component of the system, as it functions to execute the K-Prototypes algorithm to group the preprocessed divorce data into meaningful clusters. As shown in the interface, users can specify the number of clusters (k) to be generated. In this study, the number of clusters is set to 2, based on preliminary analysis and evaluation metrics. Before executing the clustering algorithm, the system displays essential information about the prepared dataset, including that 5,568 rows of data are ready for processing, consisting of 4 numerical columns (plaintiff age, defendant age, number of children, and marriage duration) and 6 categorical columns (case type, plaintiff occupation, defendant occupation, plaintiff education, defendant education, and divorce factor). Once the user clicks the "Jalankan Clustering" (Run Clustering) button, the system initiates the K-Prototypes algorithm, which calculates distances using Euclidean distance for numerical attributes and Simple Matching distance for categorical attributes, balanced by the gamma parameter (γ). The interface also provides a "Kembali ke Dashboard" (Back to Dashboard) button, allowing users to easily navigate back to the main menu. This user-friendly design ensures that even non-technical users can execute complex clustering algorithms without writing a single line of code.

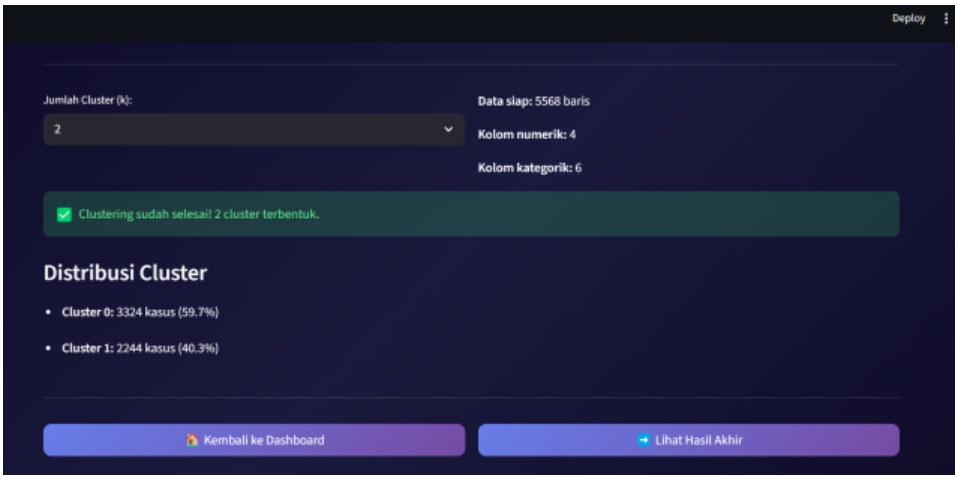


Figure 7. Cluster Distribution

Figure 7 presents the cluster distribution results after executing the K-Prototypes algorithm on the preprocessed divorce dataset. The system successfully formed two distinct clusters from 5,568 rows of data. Cluster 0 consists of 3,324 cases, which represents 59.7% of the total dataset. Cluster 1 consists of 2,244 cases, representing 40.3% of the total dataset. From these results, it is evident that Cluster 0 is the dominant cluster with the largest number of members, containing approximately 60% of all divorce cases. This uneven distribution indicates that the majority of divorce cases share similar characteristics that distinguish them from the minority group.

3.7. Final Cluster Results Interface

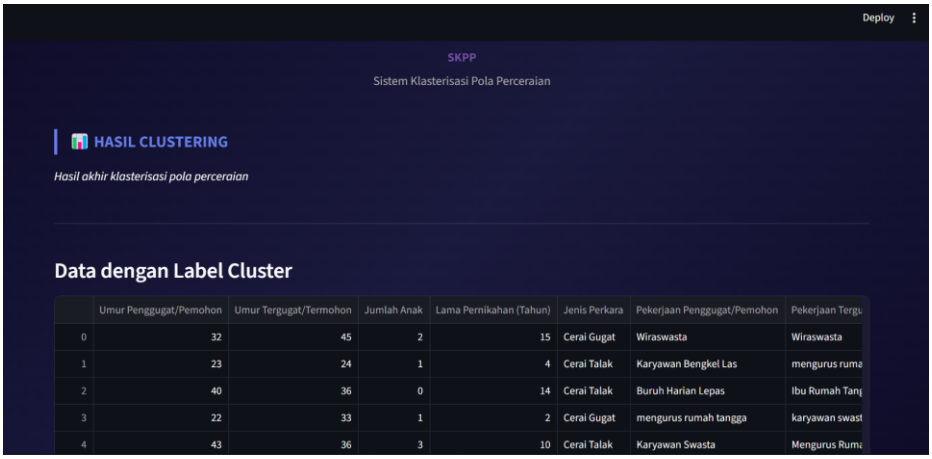


Figure 8. Data with Cluster Labels

Figure 8 displays the divorce dataset after the clustering process is completed. Each divorce case has been assigned a cluster label indicating which group the data belongs to. The table shows the first five rows with their respective attributes, including plaintiff age, defendant age, number of children, marriage duration, case type, occupations, and cluster labels.

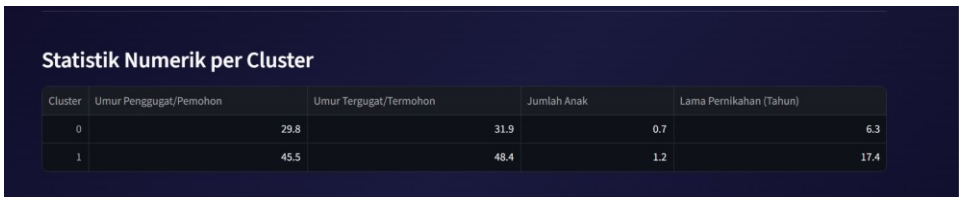


Figure 9. Numerical Statistics per Cluster

Figure 9 presents the average numerical values for each cluster. Cluster 0 has an average plaintiff's age of 29.8 years, an average defendant's age of 31.9 years, an average number of children of 0.7, and an

average marriage duration of 6.3 years. This cluster is characterized by younger couples with relatively short marriages and few children. Meanwhile, Cluster 1 has an average plaintiff's age of 45.5 years, an average defendant's age of 48.4 years, an average number of children of 1.2, and an average marriage duration of 17.4 years. This cluster represents mature couples with longer marriages and more children. The significant age gap between clusters suggests that age is a key distinguishing factor in divorce patterns at the Banyuwangi Religious Court.

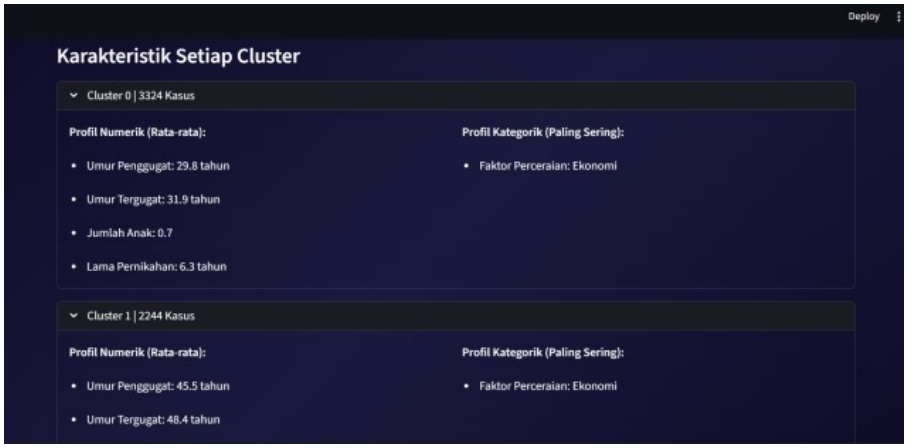


Figure 10. Characteristics of Each Cluster

Figure 10 summarizes the numerical and categorical profiles of both clusters. Cluster 0, consisting of 3,324 cases (59.7%), has an average plaintiff age of 29.8 years, defendant age of 31.9 years, 0.7 children, and marriage duration of 6.3 years. The most frequent divorce factor in this cluster is economic issues. Cluster 1, consisting of 2,244 cases (40.3%), has an average plaintiff age of 45.5 years, defendant age of 48.4 years, 1.2 children, and marriage duration of 17.4 years. The most frequent divorce factor in this cluster is also economic issues. These profiles show that while both clusters share economic factors as the primary cause of divorce, they differ significantly in age, marriage duration, and number of children.

3.8. Visualitation

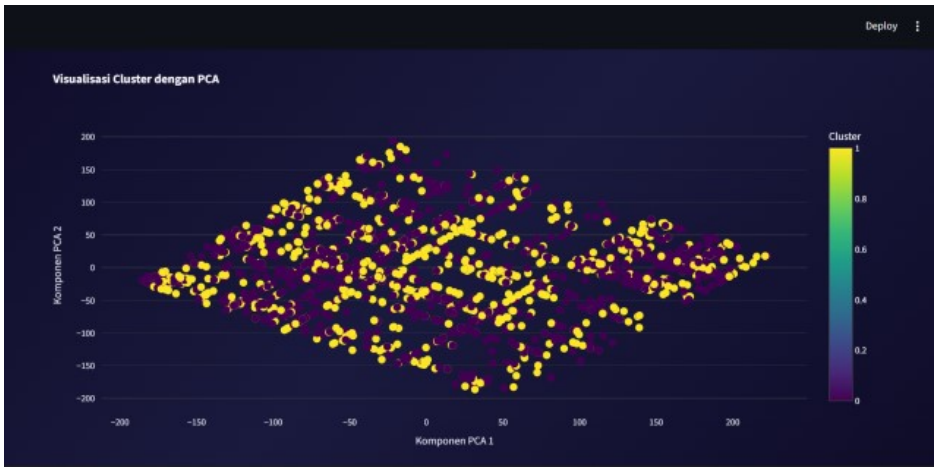


Figure 11. PCA (Principal Component Analysis)

The next page is the Cluster Visualization page, which is designed to present the results of the clustering process in a visual and easily interpretable format. This page helps users understand the grouping patterns formed from the analyzed data and supports more effective data exploration and decision-making. By providing visual representations of cluster distributions, users can identify similarities and differences among data groups more intuitively than through tabular data alone.

One of the primary visualizations available on this page is Principal Component Analysis (PCA), which is used for dimensionality reduction. PCA is a multivariate statistical technique that transforms a large number of correlated variables into a smaller set of independent variables known as principal components. These components retain most of the original information while reducing the complexity of the dataset.

Through this transformation, high-dimensional data can be projected into a two-dimensional or three-dimensional space, making cluster patterns easier to observe and analyze.

The PCA visualization enables users to evaluate the separation and compactness of clusters, identify potential outliers, and assess the overall quality of the clustering results. As a result, the Cluster Visualization page serves as an important analytical tool for interpreting clustering outcomes and communicating data insights more effectively. [10].

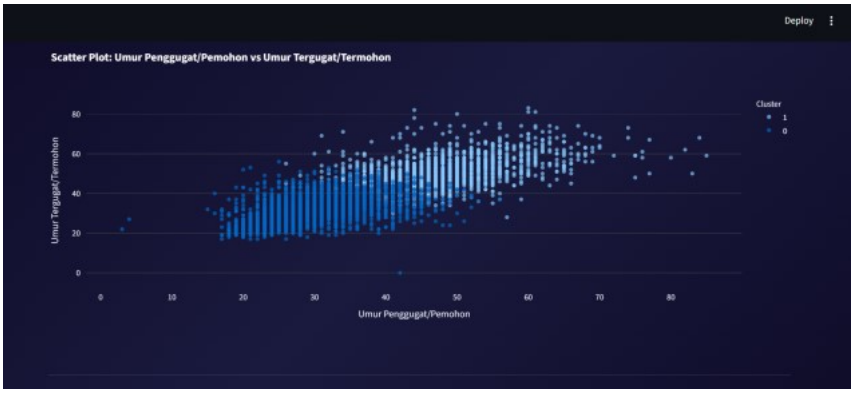
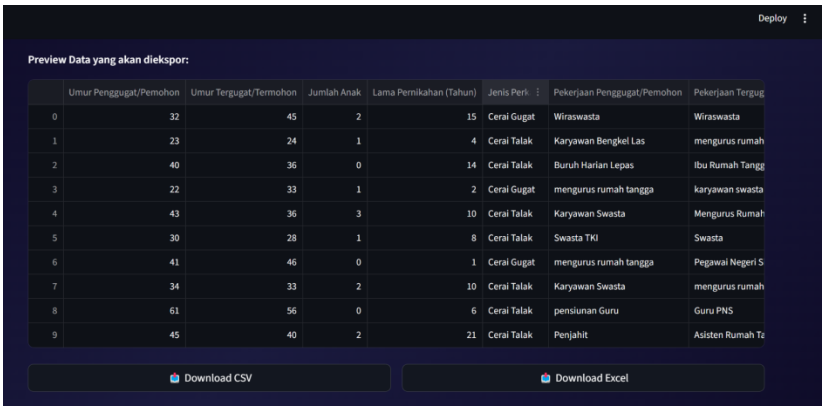


Figure 12. Scatter Plot

Figure 12 presents the scatter plot of plaintiff age versus defendant age, colored by cluster assignment. The plot clearly shows a positive correlation, indicating that older plaintiffs tend to have older defendants. Cluster 0 (blue) is dominated by younger couples aged 20–40 years, while Cluster 1 (red) is dominated by mature-aged couples aged 40–70 years. The separation between the two clusters is quite distinct, with minimal overlap between groups. This clear visual distinction demonstrates that age is a key differentiating factor in divorce patterns. The successful separation confirms that the K-Prototypes algorithm effectively grouped the divorce data based on age characteristics, validating the clustering results.

3.9. Data Export Interface



	Umur Penggugat/Pemohon	Umur Tergugat/Termohon	Jumlah Anak	Lama Pernikahan (Tahun)	Jenis Perk	Pekerjaan Penggugat/Pemohon	Pekerjaan Tergug
0	32	45	2	15	Cerai Gugat	Wiraswasta	Wiraswasta
1	23	24	1	4	Cerai Talak	Karyawan Bengkel Las	mengurus rumah
2	40	36	0	14	Cerai Talak	Buruh Harian Lepas	Ibu Rumah Tangg
3	22	33	1	2	Cerai Gugat	mengurus rumah tangga	karyawan swasta
4	43	36	3	10	Cerai Talak	Karyawan Swasta	Mengurus Rumah
5	30	28	1	8	Cerai Talak	Swasta TKI	Swasta
6	41	46	0	1	Cerai Gugat	mengurus rumah tangga	Pegawai Negeri S
7	34	33	2	10	Cerai Talak	Karyawan Swasta	mengurus rumah
8	61	56	0	6	Cerai Talak	pensiunan Guru	Guru PNS
9	45	40	2	21	Cerai Talak	Penjahit	Asisten Rumah Ta

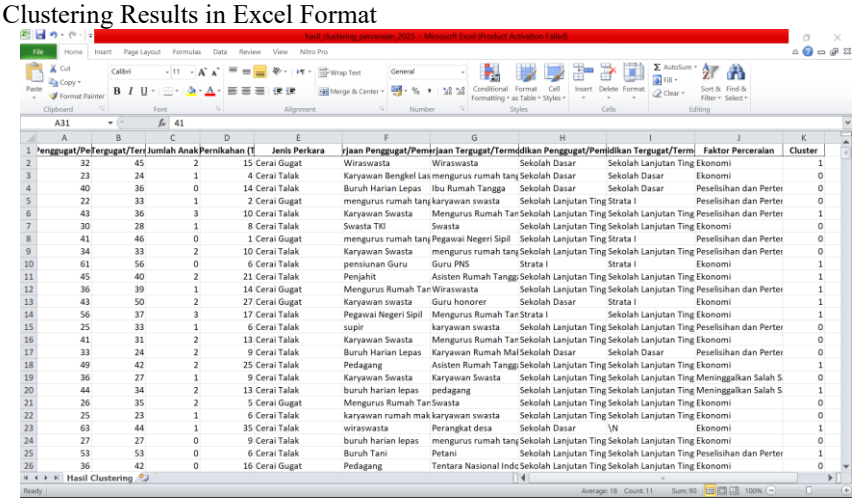
Figure 13. Data Export Interface

Figure 14 presents a preview of the dataset prepared for export, allowing users to review the clustering results before downloading the data. The table displays the first ten records from the divorce dataset and includes several important attributes, such as the age of the plaintiff, age of the defendant, number of children, duration of marriage, type of case, and the occupations of both parties. Each row represents an individual divorce case that has been processed and assigned to a specific cluster based on the clustering analysis.

This preview feature helps users verify the accuracy and completeness of the data before exporting it to external formats. To support flexibility and ease of use, the interface provides two export options: “Download CSV” and “Download Excel.” Users can select the format that best suits their analytical or reporting needs. The exported file contains the complete dataset of 5,570 divorce cases, including all relevant attributes and the cluster labels generated by the system.

The export functionality plays an important role in extending the usability of the SKPP application. By enabling offline access to clustering results, court administrators can conduct further statistical analysis, generate reports, archive data, or share findings with stakeholders and policymakers. This feature enhances data accessibility and supports evidence-based decision-making without requiring continuous access to the main application.

3.10. Exported Clustering Results in Excel Format



	A	B	C	D	E	F	G	H	I	J	K
1	enggugat/Pe	ergugat/Ter	Jumlah Anak	Pernikahan (T	Jenis Perkara	rjaan Penggugat/Pem	rjaan Tergugat/Term	dikan Penggugat/Pem	dikan Tergugat/Term	Faktor Perceraian	Cluster
2	32	45	2	15	Cerai Gugat	Wiraswasta	Wiraswasta	Sekolah Dasar	Sekolah Lanjutan Ting	Ekonomi	1
3	23	24	1	4	Cerai Talak	Karyawan Bengkel Las	mengurus rumah tan	Sekolah Dasar	Sekolah Dasar	Ekonomi	0
4	40	36	0	14	Cerai Talak	Buruh Harian Lepas	Ibu Rumah Tangga	Sekolah Dasar	Sekolah Dasar	Peselisihan dan Perter	0
5	22	33	1	2	Cerai Gugat	mengurus rumah tan	karyawan swasta	Sekolah Lanjutan Ting	Strata I	Peselisihan dan Perter	0
6	43	36	3	10	Cerai Talak	Karyawan Swasta	Mengurus Rumah Tan	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Peselisihan dan Perter	1
7	30	28	1	8	Cerai Talak	Swasta TKI	Swasta	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Ekonomi	0
8	41	46	0	1	Cerai Gugat	mengurus rumah tan	Pegawai Negeri Sipil	Sekolah Lanjutan Ting	Strata I	Peselisihan dan Perter	0
9	34	33	2	10	Cerai Talak	Karyawan Swasta	mengurus rumah tan	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Peselisihan dan Perter	0
10	61	56	0	6	Cerai Talak	pensiunan Guru	Guru PNS	Strata I	Strata I	Ekonomi	1
11	45	40	2	21	Cerai Talak	Penjahit	Asisten Rumah Tangg	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Ekonomi	1
12	36	39	1	14	Cerai Gugat	Mengurus Rumah Tan	Wiraswasta	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Peselisihan dan Perter	1
13	43	50	2	27	Cerai Gugat	Karyawan swasta	Guru honorer	Sekolah Dasar	Strata I	Ekonomi	1
14	56	37	3	17	Cerai Talak	Pegawai Negeri Sipil	Mengurus Rumah Tan	Strata I	Sekolah Lanjutan Ting	Ekonomi	1
15	25	33	1	6	Cerai Talak	supir	karyawan swasta	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Peselisihan dan Perter	0
16	41	31	2	13	Cerai Talak	Karyawan Swasta	Mengurus Rumah Tan	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Ekonomi	0
17	33	24	2	9	Cerai Talak	Buruh Harian Lepas	Karyawan Rumah Mal	Sekolah Dasar	Sekolah Dasar	Peselisihan dan Perter	0
18	49	42	2	25	Cerai Talak	Pedagang	Asisten Rumah Tangg	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Ekonomi	1
19	36	27	1	9	Cerai Talak	Karyawan Swasta	Karyawan Swasta	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Meninggalkan Salah S	0
20	44	34	2	13	Cerai Talak	buruh harian lepas	pedagang	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Meninggalkan Salah S	1
21	26	35	2	5	Cerai Gugat	Mengurus Rumah Tan	Swasta	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Ekonomi	0
22	25	23	1	6	Cerai Talak	karyawan rumah mak	karyawan swasta	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Ekonomi	0
23	63	44	1	35	Cerai Talak	wiraswasta	Perangkat desa	Sekolah Dasar	\N	Ekonomi	1
24	27	27	0	9	Cerai Talak	buruh harian lepas	mengurus rumah tan	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Ekonomi	0
25	53	53	0	6	Cerai Talak	Buruh Tani	Petani	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Peselisihan dan Perter	1
26	36	42	0	16	Cerai Gugat	Pedagang	Tentara Nasional Indo	Sekolah Lanjutan Ting	Sekolah Lanjutan Ting	Ekonomi	0

Figure 14. Exported Clustering Results in Excel Format

Figure 14 displays the exported clustering results in Excel format. The file contains comprehensive information about all divorce cases, including case identity numbers, demographic data of both parties (plaintiff and defendant), number of children, marriage duration, and case type. Additionally, the dataset includes the occupations and education levels of both spouses, the primary divorce factor, and most importantly, the cluster classification results. Each row represents one divorce case with its assigned cluster label (Cluster 0 or Cluster 1). This exported file allows court administrators to perform further analysis, create reports, or share findings with policymakers. The Excel format ensures accessibility and ease of use for non-technical users.

4. Conclusion

This study successfully applied the K-Prototypes clustering method to analyze divorce patterns at the Banyuwangi Religious Court using 5,570 divorce cases recorded in the Case Tracking Information System (SIPP) in 2025. The K-Prototypes algorithm proved to be highly effective for handling mixed data types, combining Euclidean distance for four numerical variables (plaintiff age, defendant age, number of children, and marriage duration) with Simple Matching distance for six categorical variables (case type, plaintiff and defendant occupations, plaintiff and defendant education levels, and divorce factor). The optimal gamma parameter was determined through grid search, and the clustering process was implemented using Python and the Streamlit framework, resulting in the SKPP (Divorce Pattern Clustering System) application. The analysis yielded two distinct clusters. Cluster 0 (59.7%, 3,324 cases) is characterized by young couples with an average plaintiff age of 29.8 years, marriage duration of 6.3 years, and 0.7 children. Cluster 1 (40.3%, 2,244 cases) is dominated by mature couples with an average plaintiff age of 45.5 years, marriage duration of 17.4 years, and 1.2 children. Economic issues emerged as the dominant divorce factor in both clusters. PCA visualization and scatter plots confirmed clear separation between clusters, primarily based on age. The SKPP application successfully facilitates data upload, preprocessing, clustering, and result exportation. Therefore, the K-Prototypes method is highly recommended as an effective analytical tool for divorce pattern analysis, enabling the Religious Court to formulate more targeted and data-driven prevention policies.

Acknowledgement

The author would like to express their sincere gratitude to Mr. Nazrul Rachmadi, S.Kom., a member of the IT Planning and Reporting staff at the Banyuwangi Religious Court, for providing the permission, guidance, and assistance required during the divorce case data collection. Gratitude is also extended to the entire staff of the Banyuwangi Religious Court, the thesis advisors, fellow students of the Information

Systems Department at Ibrahimy University Situbondo, and the author's beloved family for their continuous support, prayers, and motivation, which enabled the successful completion of this research.

References

[1] E. Nurjamal, "Procedural Law in Religious Courts," Edu Publisher, pp. 11, Jun. 2020.
[2] Sofyang, "Law, Divorce and Domestic Violence: Dynamics, Factors and Countermeasures," Dewa Publishing, pp. 5-6, Sep. 2025.
[3] I. B. Sabrani, "Implementation of the K-Prototype Algorithm on Traffic Violation Data Clustering in Malang City," Universitas Islam Negeri Maulana Malik Ibrahim Malang, pp. 12-13, Jan. 2025.
[4] United Nations, Department of Economic and Social Affairs, World Marriage Data 2025. New York: UN Publications, 2025.
[5] Eurostat, Marriage and Divorce Statistics 2024. Luxembourg: Publications Office of the European Union, 2024.
[6] Organisation for Economic Co-operation and Development (OECD), Family Database: Marriage and Divorce Rates. Paris: OECD Publishing, 2025.
[7] S. Bertoli dan V. Marchetta, "Divorce, Remarriage, and Economic Well-Being: A Cross-Country Analysis," Journal of Population Economics, vol. 36, no. 2, pp. 567-598, 2023.
[8] World Health Organization (WHO), Global Status Report on Family Health. Geneva: WHO Press, 2025.
[9] Badan Pusat Statistik (BPS), Indonesia Statistics 2025. Jakarta: BPS RI, 2025.
[10] Supreme Court of the Republic of Indonesia, Religious Courts Agency, Annual Report of Religious Courts 2024. Jakarta: MA RI, 2025.
[11] Ministry of Women Empowerment and Child Protection of the Republic of Indonesia, Indonesian Women's Profile 2025. Jakarta: KemenPPPA, 2025.
[12] Ministry of Religious Affairs of the Republic of Indonesia, National Marriage and Divorce Data 2025. Jakarta: Kemenag RI, 2025.
[13] National Commission on Violence Against Women (Komnas Perempuan), Annual Notes (CATAHU) on Violence Against Women 2025. Jakarta: Komnas Perempuan, 2025.
[14] U. Hasanah, "The Effect of Parental Divorce on Children's Psychology," Jurnal Agenda, vol. 2, no. 1, pp. 20-34, 2024.
[15] D. Suryani and N. Fitriani, "Province Clustering Based on Marriage and Divorce Rates in Indonesia Using Hierarchical Clustering," Journal of Data Analysis and Information Systems, vol. 7, no. 2, pp. 81-95, 2024.
[16] S. Rizkiyah, I. Z. Rizqin, M. A. B. Putri, S. S. M. Wara, and W. S. J. Saputra, "Segmentation of Divorce Factors by Province in Indonesia in 2024 Using K-Means and DBSCAN," Jurnal Ilmiah Ilmu Komputer, vol. 8, no. 1, pp. 45-62, 2025.
[17] S. Anggaraini and A. H. Hasugian, "Identifying Dominant Factors of Divorce in Marbau Selatan Village Using K-Means Clustering," Jurnal COSITTE, vol. 5, no. 2, pp. 1-10, 2025.
[18] E. T. Naldy and A. Andri, "Application of Data Mining for Analysis of Consumer Purchase Lists Using the Apriori Algorithm in MDN Building Store Sales Transactions," Jurnal Nasional Ilmu Komputer, vol. 2, no. 2, pp. 89–101, 2021. doi: 10.47747/jurnalnik.v2i2.525.
[19] A. Ilallah et al., "Application of Data Mining Using the Naïve Bayes Method to Predict the Divorce Rate of Young Couples in Banyuwangi City," Jamastika, vol. 4, no. 1, pp. 124-126, Apr. 2025.
[20] L. Magdalena and R. Fahrudin, "Application of Data Mining for Cooperatives in West Java Using the Clustering Method at the Ministry of Cooperatives and SMEs," 2019. doi: https://doi.org/10.51920/jd.v9i2.120.
[21] R. R. Muhima et al., "Complete Guide to Clustering Algorithms: Concepts, Manual Calculations, and Programs," Penerbit Andi, pp. 2, Apr. 2022.
[22] A. Ahmad dan L. Dey, "A Hybrid Clustering Approach for Mixed Data Using K-Prototypes and Genetic Algorithm," Journal of Big Data, vol. 9, no. 1, pp. 1-20, 2022.
[23] K. N. Reihanah et al., "Clustering of Small and Medium Industry Characteristics in Kendari City Using the K-Prototypes Algorithm," Jurnal Gaussian, vol. 12, no. 3, 2023.
[24] G. M. Putra dan R. S. Perdana, "Optimasi Parameter Gamma pada Algoritma K-Prototype untuk Klasterisasi Data Kesehatan Ibu Hamil," Jurnal Teknologi Informasi dan Ilmu Komputer, vol. 11, no. 3, pp. 455-466, 2024.
[25] Ameliya et al., "Application of Principal Component Analysis to Determine Factors Affecting Poverty in North Sumatra," Jurnal Matematika, Ilmu Pengetahuan Alam, Kebumian dan Angkasa, vol. 4, no. 1, pp. 2, Jan. 2026.