Journal of Information Technology, computer science and Electrical Engineering (JITCSE)
Vol. 1, No. 2, May 2024, pp. 158~161
ISSN: 3046-4900, DOI: 10.61306/jitcse.v1i2

Analysis of Nursing Home Residents' Identity Completeness Classification Using the Decision Tree Algorithm

Muhammad Indra[1], Darmeli Nasution[2]
[1,2]Master of Information Technology, Universitas Pembangunan Panca Budi, Indonesia

ABSTRACT

This study evaluates the effectiveness of the Decision Tree algorithm in classifying the completeness of nursing home residents' identities based on age. The data used includes identity information from 333 residents, encompassing both Family Cards and Identity Cards (KTP). By applying the Decision Tree C4.5 algorithm, the data is classified into the categories of Incomplete, Sufficiently Complete, and Complete. The analysis results indicate that older residents tend to have less complete identities compared to younger residents. These findings highlight the effectiveness of the Decision Tree algorithm in identifying patterns within identity data, facilitating service planning and administrative management in the nursing home, and ensuring regulatory compliance. This research provides a foundation for improving identity management systems and can be used to optimize administration and protection in nursing homes.

Keyword : Decision Tree, Nursing Homes, Data Mining.

Corresponding Author:
Muhammad Indra,
Master of Information Technology
Universitas Pembangunan Panca Budi
Jl. Jend. Gatot Subroto Km. 4,5 Sei Sikambing 20122, Medan, Indonesia.
Email : or6kp9@gmail.com

Article history:
Received May 17, 2024
Revised May 20, 2024
Accepted May 25, 2024

1. INTRODUCTION

Data Mining is the process of extracting or filtering data using a sufficiently large dataset through a series of procedures to obtain valuable information from the data (Asyuti & Setyawan, 2023). One commonly used Data Mining algorithm for classification is the Decision Tree Algorithm, known for its clarity and simplicity in predicting data and its ease of understanding based on IF-THEN rules (Fatmawati et al., 2023). RapidMiner, as a data processing engine, supports deep data integration and analysis (Apriyani et al., 2023).

A family card is a civil document that regulates family relationships and includes important data such as identification numbers, gender, religion, marital status, and occupation (Pratidina & Assayuti, 2023). Meanwhile, the Identity Card (KTP) is an official identification for Indonesian citizens, essential for various administrative and legal affairs (Supawanhar et al., 2023). In a nursing home, having complete identification such as KTP and family cards is crucial for ensuring proper security and protection for residents, as well as facilitating administrative management and regulatory compliance.

In the context of a nursing home, Data Mining and Decision Tree algorithms can be used to classify the completeness of residents' identification into categories such as Incomplete, Sufficiently Complete, or Complete, and also based on age. This technique allows for automated analysis of identity data, helping to ensure that each resident receives appropriate protection and care according to their age. The use of these methods also facilitates service planning and enhances management effectiveness in the nursing home.

2. RESEARCH METHOD

a. Data Classification, is performed based on age and the completeness of identification, categorized as follows: Incomplete (if neither the Family Card nor the Identity Card (KTP) is available), Sufficiently Complete (if one of the identifications is available), and Complete (if both identifications, the Family Card and the KTP, are available).

b. Data Collection, is a crucial activity because, without data, the research objectives cannot be achieved (Firdaus et al., 2023) the collected data includes the List of Social Welfare Residents of UPTD Elderly

Social Service Kisaran, Jl. Perintis Km. 8 Simpang Tiga Lemang, Kec. Simpang Empat, Kab. Asahan, from January to April 2024.

c. Data preprocessing, is a crucial step in preparing data before performing the classification process. (Nugroho et al., 2023).
d. Model and Analysis, the classification model used is the Decision Tree C4.5, and this model is then analyzed.

3. RESULTS AND DISCUSSION

A. Data Collection

	Name WBS	Wife/Husband's Name	Number of Cl	Age	Gender	Date of Entry	Room	No. KTP	No. KK	Origin Location	Identity
313	Toga Sinambela	-	-	60	Male	08/04/2024	Jeruk	'402171058(	(402021012(	Kisaran	Complete
314	Misnan	-	-	63	Male	08/04/2024	Jeruk	-	-	Tanjungbalai	Incomplete
315	Sudirman	-	-	72	Male	08/04/2024	Jeruk	-	-	Tanjungbalai	Incomplete
316	Irianto	-	-	85	Male	08/04/2024	Jeruk	-	-	Pinangsori	Incomplete
317	Albiner Simamora	-	-	60	Male	08/04/2024	Langsat	-	-	Asahan	Incomplete
318	Rahmawati	-	-	75	Male	15/04/2024	Langsat	-	-	Kisaran Barat	Incomplete
319	Roliyah	-	2	80	Male	15/04/2024	Langsat	'402711241(	(402121107(	Kisaran	Complete
320	Naffiah	-	-	74	Female	15/04/2024	Langsat		-	Tanjung Balai	Incomplete
321	Gustina Wati Sitorus	-	-	64	Female	15/04/2024	Langsat	'830521159(	-	Tanjung Balai	Quite Complete
322	Tomoh Sihombing	-	-	64	Female	17/04/2024	Langsat	-	-	Medan	Incomplete
323	Ayub Silalahi	-	-	60	Female	17/04/2024	Mangga	-	-	Tanjung Balai	Incomplete
324	Saman Ilyas	Darmawati	1	74	Male	17/04/2024	Mangga	'4012811480002		Rantau Prapat	Quite Complete
325	Baharuddin	-	-	68	Male	17/04/2024	Mangga	'4051412070004		Pinangsori	Quite Complete
326	Siti Fatimah	Alm.Amin Panjaitan	-	61	Male	25/04/2024	Nenas	'920450859(	(920051107(	Tanjungbalai	Complete
327	Rubiah	Alm.Sukarno	3	66	Male	25/04/2024	Nenas	-	-	Tanjungbalai	Incomplete
328	Mariati Purba	-	-	60	Female	25/04/2024	Nenas	-	-	Kisaran	Incomplete
329	Ade Rospita Nasution	-	-	60	Female	25/04/2024	Nenas	-	-	Tanjungbalai	Incomplete
330	Mahfuri Siregar	-	-	87	Female	25/04/2024	Nenas	-	-	Siantar	Incomplete
331	Suwarni	-	-	64	Female	25/04/2024	Nenas	'922711258(	(922020720(	Medan	Complete
332	Meri Sianturi	-	-	65	Female	25/04/2024	Nenas	'107500358(	-	Tanjungbalai	Quite Complete
333	Ernawati Lingga	-	-	60	Female	25/04/2024	Pepaya	-	-	Siantar	Incomplete

Fig 1. Data Collection

In Figure 1, the data collection consists of 333 rows, including the header, and includes the following columns: Name WBS, Wife/Husband's Name, Number of Children, Age, Gender, Date of Entry, Room, ID Number, Family Card Number, Origin Location, and Identity.

B. Data Processing

	Age	Identity
313	60	Complete
314	63	Incomplete
315	72	Incomplete
316	85	Incomplete
317	60	Incomplete
318	75	Incomplete
319	80	Complete
320	74	Incomplete
321	64	Quite Complete
322	64	Incomplete
323	60	Incomplete
324	74	Quite Complete
325	68	Quite Complete
326	61	Complete
327	66	Incomplete
328	60	Incomplete
329	60	Incomplete
330	87	Incomplete
331	64	Complete
332	65	Quite Complete
333	60	Incomplete

Fig 2. Data Preprocessing

In Figure 2, after preprocessing, the dataset remains at 333 rows, including the header, but the number of columns is reduced to only Age and Room.

B. Model and Analysis

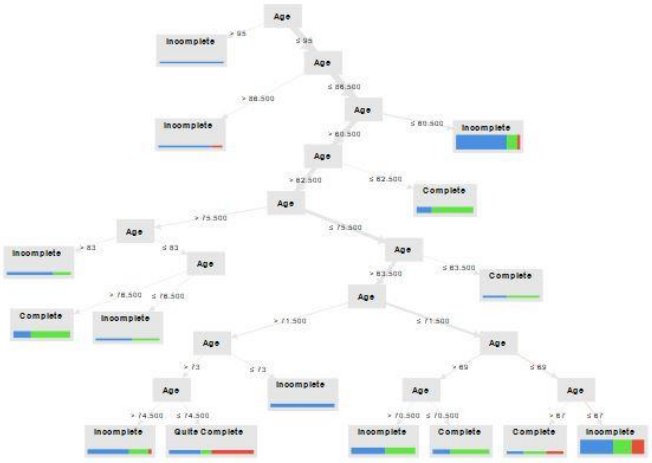

Fig 3. Model Decision Tree

The Decision Tree model shown in Figure 3 classifies data based on age, starting with the root node, which is Age > 95. For ages above 95, the classification is Incomplete. If the age is less than or equal to 95, the tree splits at Age > 86.500. For ages greater than 86.500, the classification is Incomplete, while for ages less than or equal to 86.500, the next split is at Age > 60.500. If the age is greater than 60.500, the subsequent split is at Age > 62.500, which then further splits at Age > 75.500. For ages greater than 75.500, the classification depends on Age > 83. Ages above 83 are classified as Incomplete, while ages less than or equal to 83 continue to the split at Age > 76.500. If the age is greater than 76.500, the classification is Complete; if the age is less than or equal to 76.500, the classification is Incomplete. For ages less than or equal to 75.500, the next split is at Age > 63.500, which further splits at Age > 71.500. If the age is greater than 71.500, further splitting at Age > 73 results in Incomplete if the age is greater than 74.500, and Quite Complete if the age is less than or equal to 74.500. Ages less than or equal to 73 are classified as Incomplete. For ages less than or equal to 71.500, the next split is at Age > 69, with further splitting at Age > 70.500, resulting in Incomplete if the age is greater than 70.500, and Complete if the age is less than or equal to 70.500. Ages less than or equal to 69 continue to the split at Age > 67, resulting in Complete if the age is greater than 67, and Incomplete if the age is less than or equal to 67. Ages less than or equal to 63.500 are classified as Complete. Ages less than or equal to 62.500 are classified as Complete, and ages less than or equal to 60.500 are classified as Incomplete.

Tree

```
Age > 95: Incomplete {Incomplete=5, Complete=0, Quite Complete=0}
Age ≤ 95
|   Age > 86.500: Incomplete {Incomplete=5, Complete=0, Quite Complete=1}
|   Age ≤ 86.500
|   |   Age > 60.500
|   |   |   Age > 62.500
|   |   |   |   Age > 75.500
|   |   |   |   |   Age > 83: Incomplete {Incomplete=8, Complete=3, Quite Complete=0}
|   |   |   |   |   Age ≤ 83
|   |   |   |   |   |   Age > 76.500: Complete {Incomplete=9, Complete=20, Quite Complete=0}
|   |   |   |   |   |   Age ≤ 76.500: Incomplete {Incomplete=4, Complete=3, Quite Complete=0}
|   |   |   |   Age ≤ 75.500
|   |   |   |   |   Age > 63.500
|   |   |   |   |   |   Age > 71.500
|   |   |   |   |   |   |   Age > 73
|   |   |   |   |   |   |   |   Age > 74.500: Incomplete {Incomplete=13, Complete=6, Quite Complete=1}
|   |   |   |   |   |   |   |   Age ≤ 74.500: Quite Complete {Incomplete=6, Complete=2, Quite Complete=8}
|   |   |   |   |   |   |   Age ≤ 73: Incomplete {Incomplete=14, Complete=0, Quite Complete=0}
|   |   |   |   |   |   Age ≤ 71.500
|   |   |   |   |   |   |   Age > 69
|   |   |   |   |   |   |   |   Age > 70.500: Incomplete {Incomplete=15, Complete=13, Quite Complete=0}
|   |   |   |   |   |   |   |   Age ≤ 70.500: Complete {Incomplete=6, Complete=13, Quite Complete=0}
|   |   |   |   |   |   |   Age ≤ 69
|   |   |   |   |   |   |   |   Age > 67: Complete {Incomplete=3, Complete=4, Quite Complete=3}
|   |   |   |   |   |   |   |   Age ≤ 67: Incomplete {Incomplete=34, Complete=19, Quite Complete=13}
|   |   |   |   |   Age ≤ 63.500: Complete {Incomplete=3, Complete=4, Quite Complete=0}
|   |   |   Age ≤ 62.500: Complete {Incomplete=7, Complete=19, Quite Complete=0}
|   |   Age ≤ 60.500: Incomplete {Incomplete=54, Complete=11, Quite Complete=3}
```

Fig 4. Description Decision Tree

In the decision tree shown in Figure 4, the process of classifying data based on age starts with the root node, which is Age > 95. For ages above 95, the classification is Incomplete, with {Incomplete=5, Complete=0, Quite Complete=0}. For ages less than or equal to 95, the decision tree further splits at Age

> 86.500. For ages greater than 86.500, the classification is Incomplete, with {Incomplete=5, Complete=0, Quite Complete=1}. For ages less than or equal to 86.500, the next evaluation is Age > 60.500. If the age is greater than 60.500, the next split is at Age > 62.500, which further splits at Age > 75.500. For ages greater than 75.500, the classification depends on Age > 83. For ages above 83, the classification is Incomplete, with {Incomplete=8, Complete=3, Quite Complete=0}. For ages less than or equal to 83, the split continues at Age > 76.500, which results in Complete if Age > 76.500, with {Incomplete=9, Complete=20, Quite Complete=0}, and Incomplete if Age ≤ 76.500, with {Incomplete=4, Complete=3, Quite Complete=0}. For ages less than or equal to 75.500, the next split is at Age > 63.500, which then splits at Age > 71.500. For ages greater than 71.500, further splitting at Age > 73 results in Incomplete if Age > 74.500, with {Incomplete=13, Complete=6, Quite Complete=1}, and Quite Complete if Age ≤ 74.500, with {Incomplete=6, Complete=2, Quite Complete=8}. Ages less than or equal to 73 are classified as Incomplete with {Incomplete=14, Complete=0, Quite Complete=0}. For ages less than or equal to 71.500, the next split is at Age > 69, with further splitting at Age > 70.500, resulting in Incomplete if Age > 70.500, with {Incomplete=15, Complete=13, Quite Complete=0}, and Complete if Age ≤ 70.500, with {Incomplete=6, Complete=13, Quite Complete=0}. Ages less than or equal to 69 continue to the split at Age > 67, resulting in Complete if Age > 67, with {Incomplete=3, Complete=4, Quite Complete=3}, and Incomplete if Age ≤ 67, with {Incomplete=34, Complete=19, Quite Complete=13}. For ages less than or equal to 63.500, the classification is Complete, with {Incomplete=3, Complete=4, Quite Complete=0}. Ages less than or equal to 62.500 are classified as Complete with {Incomplete=7, Complete=19, Quite Complete=0}, and ages less than or equal to 60.500 are classified as Incomplete with {Incomplete=54, Complete=11, Quite Complete=3}.

4. CONCLUSION

The Decision Tree algorithm is effective in analyzing and classifying the completeness of nursing home residents' identities based on age, categorizing data into Incomplete, Sufficiently Complete, and Complete. The analysis reveals that older residents tend to have incomplete identities, while younger residents are more likely to have complete identities. The application of this algorithm facilitates service planning and administrative management in the nursing home, ensuring appropriate protection and care as well as compliance with regulations.

REFERENCES

Apriyani, P., Dikananda, A. R., & Ali, I. (2023). Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi. Hello World Jurnal Ilmu Komputer, 2(1), 20–33. https://doi.org/10.56211/helloworld.v2i1.230

Asyuti, S., & Setyawan, A. A. (2023). DATA MINING DALAM PENGGUNAAN PRESENSI KARYAWAN DENGA CLUSTER MEANS. In Jurnal Ilmiah Sains Teknologi Dan Informasi (Vol. 1, Issue 1).

Fatmawati, D., Trisnawati, W., Jumaryadi, Y., & Triyono, G. (2023). KLIK: Kajian Ilmiah Informatika dan Komputer Klasifikasi Tingkat Kepuasan Penggunaan Layanan Teknologi Informasi Menggunakan Decision Tree. Media Online), 3(6), 1056–1062. https://doi.org/10.30865/klik.v3i6.803

Firdaus, I., Hidayati, R., Hamidah, R. S., Rianti, R., Cahyuni, R., & Khotimah, K. (2023). Model-Model Pengumpulan Data dalam Penelitian Tindakan Kelas. Jurnal Kreativitas Mahasiswa, 1(2), 105–113.

Nugroho, K., Tjahjaningsih, E., Liana, L., Mohamad, R., & Bhakti, H. (2023). Prediksi Ujaran Kebencian Berbasis Text Pada Sosial Media Menggunakan Metode Neural Network Prediction of Text Based Hate Speech on Socials Media Using the Naural Network. Jurnal Ilmiah Intech: Information Technology Journal of UMUS, 5(01), 60–68.

Pratidina, G., & Assayuti, M. J. (2023). EFEKTIVITAS PELAYANAN PEMBUATAN KARTU KELUARGA (KK). In Karimah Tauhid (Vol. 2, Issue 5).

Supawanhar, Ditasman, Amrullah, Hartono, R., & Silvana, R. J. (2023). Analisis Indeks Kepuasan Masyarakat Dalam Pembuatan Kartu Tanda Penduduk (E-KTP). Jurnal Ekombis Review – Jurnal Ilmiah Ekonomi Dan Bisnis, 11(2), 1539–1539. https://doi.org/10.37676/ekombis.v11i12