

Analysis of Property Tax Payment Compliance Classification in Tebing Tinggi City Using the C4.5 Decision Tree Algorithm


Andysah Putera Utama Siahaan¹, Sulis Sutiono², Sugeng Pranoto³, Sarifudin⁴, Risca Sri Mentari⁵

^{1,2,3,4,5}Master of Information Technology, Universitas Pembangunan Panca Budi, Indonesia

ABSTRACT

This research analyzes property tax payment compliance in Tebing Tinggi City using the C4.5 Decision Tree algorithm. With the rapid advancement of data mining technology, this analysis utilizes classification techniques to identify compliance patterns based on property tax payment data. The research methodology involves data collection, preprocessing, and building the Decision Tree model using RapidMiner. The results indicate that the Decision Tree model can effectively predict compliance levels based on attributes such as Total_Payment and Total_Bill. Individuals with higher payment and bill values tend to be compliant, while those with lower values show less compliance. These findings provide insights for authorities to design more effective strategies to improve tax compliance and identify areas that require special attention in Tebing Tinggi City.

Keyword : Property Tax Payment Compliance, C4.5 Decision Tree, Data Mining.

 This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Corresponding Author:

Andysah Putera Utama Siahaan,
Master of Information Technology
Universitas Panca Budi
Jl. Jend. Gatot Subroto Km. 4,5 Sei Sikambang 20122, Medan, Indonesia.
Email : andiesiahaan@gmail.com

Article history:

Received May 9, 2024
Revised May 20, 2024
Accepted May 11, 2024

1. INTRODUCTION

A government or private institution will always keep up with technological advancements (Damayanti & Rusliyawati, 2023). The rapid advancement of technology today cannot be avoided in its influence (Samsugi et al., 2020). This advancement has created a modern revolution commonly referred to as the information revolution (Arifin & Veza Okta, 2019). One of the technological advancements is Data Mining, which involves a collection of techniques aimed at discovering previously unknown patterns in collected data (Zai, 2022). With the application of data mining, it is possible to perform several tasks including classification, prediction, estimation, and obtaining other useful information from large data sets (Adhitya et al., 2023).

Classification is the process of identifying and grouping an object into the same group or category (Adhitya et al., 2023) Classification in data mining can be performed using the C4.5 algorithm (Fajrillah & Lusiah, 2023). The C4.5 algorithm is commonly used to construct decision trees (Pangaribuan et al., 2019). A decision tree is a flowchart-like structure where each node (internal node) represents an attribute and branches depict the results of tests or attribute values, while leaves represent classes or class distributions (Septhya et al., 2023).

Taxes are a major source of revenue for the country and also a crucial source of funding for national financing (Oktavia & Simbolon, 2023). There are two types of taxes in Indonesia: central taxes and regional taxes. Central taxes are obligations imposed on taxpayers, accumulated and processed by the central government, while regional taxes are obligations managed by local governments (regents or mayors) and are directly collected by local authorities within districts or villages, which are also involved in assisting and carrying out tax collection. One form of regional tax is the Property Tax (PBB) (Agustin & Mustoffa, 2023).

Property Tax (PBB) is a tax managed by the government imposed on individuals as taxpayers for the ownership of land and buildings (Engkus Kurniawan et al., 2023). Property Tax is the revenue that must be paid by an individual or entity that legally owns, controls, and benefits from the land and buildings they possess (Ningrum & Siregar, 2023). Public compliance with tax payments is greatly influenced by the level of public awareness (Mataram, 2023). However, the level of public compliance with Property Tax payments in Tebing Tinggi City is often suboptimal, which can affect regional revenue

and hinder the government's ability to fund various programs and development projects. Applying the C4.5 Decision Tree algorithm in data mining can be a solution to the problem of Property Tax payment compliance in Tebing Tinggi City. This algorithm can be used to analyze Property Tax payment data and predict compliance levels based on various relevant attributes. With insights gained from this predictive model, authorities in Tebing Tinggi City can design more effective strategies to improve tax compliance.

2. RESEARCH METHOD

- a. Compliance Classification, this is the process of grouping Property Tax payment data in Tebing Tinggi City into the categories of 'OBEDIENT' or 'NOT OBEY'.
- b. Data Collection, this is the stage required for analysis. The data is taken from the Property Tax (PBB-P2) records of Tebing Tinggi City for the tax year 2022-2023.
- c. Preprocessing, this is the process of preparing raw data before further processing. Typically, data is processed by removing irrelevant data or transforming data within the system to facilitate processing (Adhitya et al., 2023)
- d. Modeling and Analysis, this is the final stage of the research, where a model is created using RapidMiner to produce a Decision Tree model. RapidMiner can be used for data mining thanks to its advanced computational algorithm technology and computer-based data analysis (Kaamilah, 2023). After modeling is completed, the next step is to conduct analysis based on the model. Analysis is crucial in all fields, events, and conditions (Saputri et al., 2023)

3. RESULTS AND DISCUSSION

A. Data Collection

	A	B	C	D	E	F	G	H
1	NOP	NJOP BUMI	NJOP_BNG	REGULATION	BOOK	TOTAL_BILL	TOTAL_PAYMENT	STATUS
56314	127420320200700900	16000000	6804000	16005 I		32010		0 OBEDIENT
56315	127420120101405920	8446000	0	10558 I		21116		0 OBEDIENT
56316	127420420300501880	4544000	21420000	19955 I		19955		0 OBEDIENT
56317	127420420600500380	41800000	9720000	51900 I		103800		0 OBEDIENT
56318	127420120300200700	133988000	0	167485 II		334970		0 OBEDIENT
56319	127420520500800180	33408000	13752000	46450 I		92900		0 OBEDIENT
56320	127420320500203350	10332000	21420000	27190 I		54380	27190	OBEDIENT
56321	127420320100900900	43740000	0	54675 I		109350	109350	OBEDIENT
56322	127420120300404520	19260000	18600000	47325 I		94650	94650	OBEDIENT
56323	127420420400304870	9536000	0	11920 I		23840	23840	OBEDIENT
56324	127420220100201160	51200000	0	64000 I		128000	131840	NOT OBEY
56325	127420120700400380	247200000	20440000	322050 II		644100	644100	OBEDIENT
56326	127420520300404360	269040000	266000000	656300 III		1312600	1417608	NOT OBEY
56327	127420520400102690	71604000	63000000	155755 II		155755	155755	OBEDIENT
56328	127420420100400220	508320000	0	635400 III		635400	648108	NOT OBEY
56329	127420320500105450	2240000	0	10000 I		20000		0 OBEDIENT
56330	127420420400101750	90585000	0	113231 II		226462		0 OBEDIENT
56331	127420420100404870	17928000	0	22410 I		44820		0 OBEDIENT
56332	127420120200503810	6400000	11880000	10350 I		10350		0 OBEDIENT
56333	127420320500404960	77353000	0	96691 I		193382		0 OBEDIENT

Fig 1. Property Tax Payment Data for Tebing Tinggi City

In Figure 1, it can be seen that the Property Tax payment data for Tebing Tinggi City consists of 56,333 records, including the headers. It includes columns (NOP, NJOP BUMI, NJOP_BNG, REGULATION, BOOK, TOTAL_BILL, TOTAL_PAYMENT, STATUS).

B. Preprocessing

	A	B	C
1	TOTAL_BILL	TOTAL_PAYMENT	STATUS
56314	32010	0	OBEDIENT
56315	21116	0	OBEDIENT
56316	19955	0	OBEDIENT
56317	103800	0	OBEDIENT
56318	334970	0	OBEDIENT
56319	92900	0	OBEDIENT
56320	54380	27190	OBEDIENT
56321	109350	109350	OBEDIENT
56322	94650	94650	OBEDIENT
56323	23840	23840	OBEDIENT
56324	128000	131840	NOT OBEY
56325	644100	644100	OBEDIENT
56326	1312600	1417608	NOT OBEY
56327	155755	155755	OBEDIENT
56328	635400	648108	NOT OBEY
56329	20000	0	OBEDIENT
56330	226462	0	OBEDIENT
56331	44820	0	OBEDIENT
56332	10350	0	OBEDIENT
56333	193382	0	OBEDIENT

Fig 2. Preprocessing of Property Tax Payment Data for Tebing Tinggi City

In Figure 2, after preprocessing, the Property Tax payment data for Tebing Tinggi City still consists of 56,333 records, including the headers. However, the remaining data includes only the columns (TOTAL_BILL, TOTAL_PAYMENT, STATUS)

C. Modeling dan Analysis

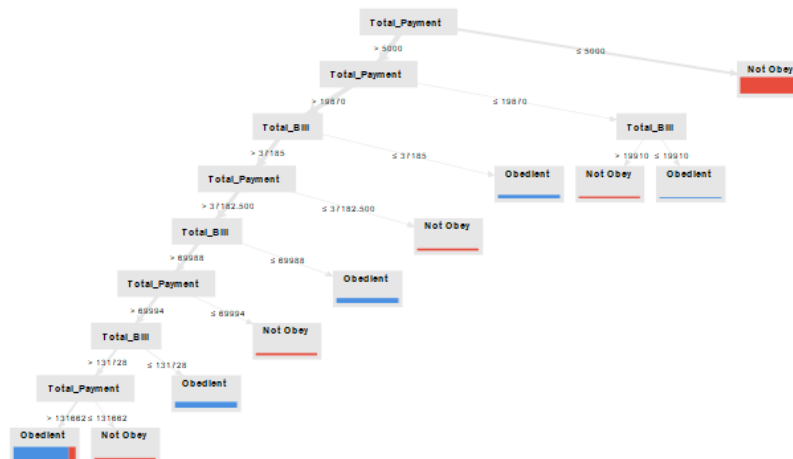


Fig 3. Model Decision Tree

In Figure 3, this Decision Tree model begins with the root node where `Total_Payment > 5000`, dividing the data into two groups: greater than 5000 and less than or equal to 5000. For the group with `Total_Payment` greater than 5000, the analysis continues by comparing `Total_Payment` and `Total_Bill` using various thresholds. If `Total_Payment` is greater than 19,870, the data is analyzed based on `Total_Bill` with a threshold of 37,185, followed by comparisons of `Total_Payment` and `Total_Bill` at the next thresholds. If `Total_Bill` is greater than 69,988 and `Total_Payment` is greater than 69,994, the data is classified as Obedient if `Total_Payment` is greater than 131,662, or Not Obey if it is not. If `Total_Bill` is less than or equal to 69,988, the data is classified as Obedient if `Total_Payment` is less than or equal to 37,182.500, or Not Obey if it is not. For the group with `Total_Payment` less than or equal to 19,870, if `Total_Bill` is greater than 19,910, the data is classified as Not Obey; otherwise, if it is less than or equal to 19,910, it is classified as Obedient. The group with `Total_Payment` less than or equal to 5000 is always classified as Not Obey.

Tree

```
Total_Payment > 5000
|   Total_Payment > 19870
|   |   Total_Bill > 37185
|   |   |   Total_Payment > 37182.500
|   |   |   |   Total_Bill > 69988
|   |   |   |   |   Total_Payment > 69994
|   |   |   |   |   |   Total_Bill > 131728
|   |   |   |   |   |   |   Total_Payment > 131662: Obedient {Obedient=13594, Not Obey=1655}
|   |   |   |   |   |   |   Total_Payment ≤ 131662: Not Obey {Obedient=0, Not Obey=1732}
|   |   |   |   |   |   |   Total_Bill ≤ 131728: Obedient {Obedient=6332, Not Obey=6}
|   |   |   |   |   |   |   Total_Payment ≤ 69994: Not Obey {Obedient=0, Not Obey=1968}
|   |   |   |   |   |   |   Total_Bill ≤ 69988: Obedient {Obedient=5112, Not Obey=5}
|   |   |   |   |   |   |   Total_Payment ≤ 37182.500: Not Obey {Obedient=0, Not Obey=1749}
|   |   |   |   |   |   |   Total_Bill ≤ 37185: Obedient {Obedient=3265, Not Obey=2}
|   |   |   |   |   |   |   Total_Payment ≤ 19870
|   |   |   |   |   |   |   |   Total_Bill > 19910: Not Obey {Obedient=0, Not Obey=1271}
|   |   |   |   |   |   |   |   Total_Bill ≤ 19910: Obedient {Obedient=409, Not Obey=0}
|   |   |   |   |   |   |   Total_Payment ≤ 5000: Not Obey {Obedient=0, Not Obey=19232}
```

Fig 4. Description Tree

In Figure 4, the Description Tree shows the total number of compliant (Obedient) and non-compliant (Not Obey) individuals distributed across various branches based on `Total_Payment` and `Total_Bill` values. For the first branch, where `Total_Payment > 5000`, compliant and non-compliant individuals are further split: in the sub-branch `Total_Payment > 19870`, several conditions result in different numbers of compliant and non-compliant individuals. For instance, under the condition `Total_Payment > 37182.500` and `Total_Bill > 69988`, further sub-branches show that `Total_Payment > 69994` and `Total_Bill > 131728` result in 13,594 compliant individuals and 1,655 non-compliant individuals if `Total_Payment > 131662`, or 6,332 compliant individuals and 6 non-compliant individuals if `Total_Payment ≤ 131662`. In cases where `Total_Bill ≤ 69988`, there are 5,112 compliant individuals and 5 non-compliant individuals, and with `Total_Payment ≤ 69994`, there are 1,968 non-compliant individuals. For `Total_Bill ≤ 37185`, there are 3,265 compliant individuals and 2 non-compliant individuals. In the sub-branch `Total_Payment ≤ 19870`, when `Total_Bill > 19910`, there are 1,271 non-compliant individuals, whereas `Total_Bill ≤ 19910` includes 409 compliant individuals and no non-compliant individuals. Outside these branches, for `Total_Payment ≤ 5000`, there are 19,232 non-compliant individuals.

4. CONCLUSION

This study demonstrates that applying the C4.5 Decision Tree algorithm to classify Property Tax (PBB) payment compliance in Tebing Tinggi City yields significant findings. The Decision Tree model can identify patterns in PBB payment data and predict compliance levels based on attributes such as `Total_Payment` and `Total_Bill`. The analysis results indicate that individuals with higher total payments

and bills are more likely to be compliant, whereas those with lower payments and bills are more likely to be non-compliant. These findings provide valuable insights for authorities to design more effective strategies for improving tax compliance by considering factors that influence compliance levels. By using this model, the government can identify areas requiring special attention and implement targeted policies to enhance PBB payment compliance in Tebing Tinggi City.

REFERENCES

- Adhitya, R. R., Wina Witanti, & Rezki Yuniarti. (2023). PERBANDINGAN METODE CART DAN NAÏVE BAYES UNTUK KLASIFIKASI CUSTOMER CHURN. *INFOTECH Journal*, 9(2), 307–318. <https://doi.org/10.31949/infotech.v9i2.5641>
- Agustin, M. W., & Mustoffa, A. F. (2023). Analisis Partisipasi Masyarakat dalam Membayar Pajak Bumi dan Bangunan (Studi Kasus Desa Puhpelem). *Owner*, 7(3), 1919–1929. <https://doi.org/10.33395/owner.v7i3.1574>
- Arifin, N. Y., & Veza Okta. (2019). DASHBOARD SISTEM APLIKASI PENGELOLAAN OBAT. *Engineering And Technology International Journal*, 1(2), 59–65.
- Damayanti, S., & Rusliyawati. (2023). Analisis Pemanfaatan Google Spreetsheed Untuk Pengendalian Surat (Study Kasus: Dinas Perindustrian Bandar Lampung). *JURNAL TEKNOLOGI DAN SISTEM INFORMASI*, 4(2), 179–184. <https://doi.org/10.33365/jtsi.v4i2.2571>
- Engkus Kurniawan, Oyon Saryono, & Irma Darmawati. (2023). Pengaruh Dimensi Komunikasi Dan Motivasi Kerja Terhadap Efektivitas Pencapaian Target Pajak Bumi Dan Bangunan (PBB). *Jurnal of Management and Social Sciences*, 1(2), 87–99. <https://doi.org/10.59031/jmsc.v1i2.144>
- Fajrillah, & Lusiah. (2023). Sistem Informasi Geografis Fasilitas Kesehatan di Tuntungan Berbasis Android. *Bulletin of Computer Science Research*, 3(3), 257–262. <https://doi.org/10.47065/bulletincsr.v3i3.244>
- Kaamilah, L. L. (2023). Analisis Kelompok Lansia Berdasarkan Kategori Usia Dengan Metode K-Means Clustering. *Jurnal Riset Ilmu Akuntansi*, 2(2).
- Mataram, U. (2023). Tingkat kesadaran Pembayaran Pajak (sTudi di kecamaTan LabuaPi kabuPaTen Lombok baraT) LeveL of Tax PaymentT awareness (a sTudy in LabuaPi subdisTricT, wesT Lombok regency) rusnan johannes johny koynja. *Jurnal Kompilasi Hukum*, 8(2). <https://doi.org/10.29303/jkh.v8i2.156>
- Ningrum, D. F., & Siregar, R. A. (2023). Analisis Penerimaan Pajak Bumi dan Bangunan dalam Meningkatkan Pendapatan Daerah Kota Medan. *JURNAL SYNTAX IMPERATIF : Jurnal Ilmu Sosial Dan Pendidikan*, 4(5), 700–708. <https://doi.org/10.36418/syntax-imperatif.v4i5.301>
- Oktavia, N., & Simbolon, L. (2023). FAKTOR-FAKTOR PENENTU DALAM KEPATUHAN MEMBAYAR PAJAK-PAJAK BUMI DAN BANGUNAN Anang Subardjo Sekolah Tinggi Ilmu Ekonomi Indonesia (STIESIA) Surabaya. *Jurnal Ilmu Dan Riset Akuntansi*, 12(1), 1–17.
- Pangaribuan, J. J., Tedja, C., & Wibowo, S. (2019). PERBANDINGAN METODE ALGORITMA C4.5 DAN EXTREME LEARNING MACHINE UNTUK MENDIAGNOSIS PENYAKIT JANTUNG KORONER. In *PSDKU Medan Jurusan Teknik Informatika INFORMATICS ENGINEERING RESEARCH AND TECHNOLOGY*.
- Samsugi, S., Nurkholis, A., Permatasari, B., Candra Nugroho, A., & Bagus Prasetyo, A. (2020). INTERNET OF THINGS UNTUK PENINGKATAN PENGETAHUAN TEKNOLOGI BAGI SISWA. *Journal of Technology and Social for Community Service (JTSCS)*, 2(2), 173–177. <https://ejurnal.teknokrat.ac.id/index.php/teknobdimas>
- Saputri, S. R., Liyani, Y. A., & Astutik, E. P. (2023). ANALISIS FAKTOR YANG MELATARBELAKANGI KONSOLIDASI DI PT. BANK MANDIRI (PERSERO) TBK. *Bisnis Dan Keuangan TRANSEKONOMIKA | VOLUME*, 3(3). <https://transpublika.co.id/ojs/index.php/Transekonomika>
- Septhya, D., Rahayu, K., Rabbani, S., Fitriana, V., Irawan, Y., & Hayami, R. (2023). Implementasi Algoritma Decision Tree dan Support Vector Machine untuk Klasifikasi Penyakit Kanker Paru. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(1), 15–19.
- Zai, C. (2022). IMPLEMENTASI DATA MINING SEBAGAI PENGOLAHAN DATA. *Portaldata.Org*, 2(3), 1–12.