

**SKRIPSI**  
**PENERAPAN *DATA MINING* UNTUK MENENTUKAN  
POTENSI KEPATUHAN WAJIB PAJAK BERDASARKAN  
DATA PBB MENGGUNAKAN ALGORITMA C4.5  
(STUDI KASUS UPT BBPD ILIR BARAT II)**



Oleh :

**BELLA MOULINA      09031281320020**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS SRIWIJAYA  
2018**

**LEMBAR PENGESAHAN**

**SKRIPSI**

**PENERAPAN *DATA MINING* UNTUK MENENTUKAN POTENSI  
KEPATUHAN WAJIB PAJAK BERDASARKAN DATA PBB  
MENGUNAKAN ALGORITMA C4.5  
(STUDI KASUS UPT BBPD ILIR BARAT II)**

Sebagai salah satu syarat untuk menyelesaikan studi  
Di Program Studi Sistem Informasi Jenjang Strata S1

Oleh :

**Bella Maulina**

**09031281320020**

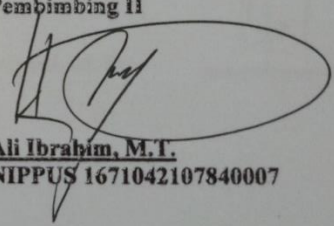
Indralaya,

2018

Pembimbing I

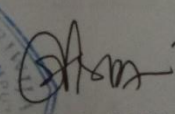
Pembimbing II

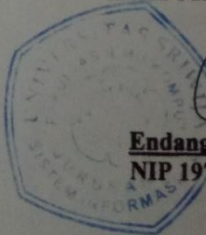
  
**Mgs. Afrivan Firdaus, M.IT**  
NIP 198202122006041003

  
**Ali Ibrahim, M.T.**  
NIPPUS 1671042107840007

Mengetahui,

Ketua Jurusan Sistem Informasi,

  
**Endang Lestari, S.Kom, M.T**  
NIP 197811172006042001



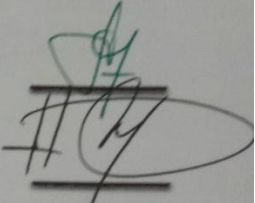
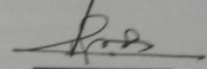
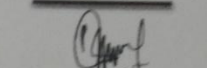
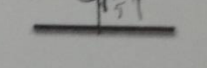
HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

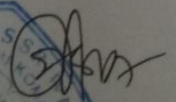
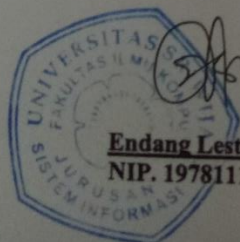
Hari : Senin

Tanggal : 15 Januari 2018

Tim Penguji

- |                               |                              |   |
|-------------------------------|------------------------------|---|
| 1. Ketua (Pembimbing I)       | : Mgs. Afriyan Firdaus, M.IT |    |
| 2. Sekretaris (Pembimbing II) | : Ali Ibrahim, M.T.          |    |
| 3. Anggota I                  | : Dr. Ermatita, M.Kom        |   |
| 4. Anggota II                 | : Putri Eka Sevtiyuni, M.T.  |  |

Mengetahui,  
Ketua Jurusan Sistem Informasi,

  
  
**Endang Lestari, S.Kom, M.T**  
NIP. 197811172006042001

## LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama : Bella Moulina  
NIM : 09031281320020  
Program Studi : Sistem Informasi  
Judul Skripsi : PENERAPAN *DATA MINING* UNTUK MENENTUKAN POTENSI KEPATUHAN WAJIB PAJAK BERDASARKAN DATA PBB MENGGUNAKAN ALGORITMA C4.5 (STUDI KASUS UPT BPPD ILIR BARAT II)  
Hasil Pengecekan *Software iThenticate/Turnitin* : 9 %

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil penjiplakan / plagiat. Apabila ditemukan unsur penjiplakan / plagiat dalam laporan tugas akhir ini, maka saya bersedia menerima saksi akademik dari Universitas Sriwijaya.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapaun.



Indralaya, Januari 2018



**Bella Moulina**  
NIM. 09031281320020

## HALAMAN PERSEMBAHAN

*“ You were given this hard life because you are strong enough  
to face it”*

*Kupersembahkan ini untuk:*

- ✓ *Allah S.W.T*
- ✓ *Kedua Orang Tuaku*
- ✓ *Adik-adikku*
- ✓ *Keluargaku*
- ✓ *Yayasan Karya Salemba  
Empat*
- ✓ *Paguyuban KSE UNSRI*
- ✓ *Octuplet*
- ✓ *Teman-temanku*
- ✓ *Almamater Kebanggaanku*

## KATA PENGANTAR

Segala puji dan syukur Penulis sampaikan kepada Allah SWT, yang Maha. Pengasih dan Maha Penyayang. Karena berkat rahmat dan hidayah, serta inayah-Nya lah sehingga penulis dapat dapat menyelesaikan laporan skripsi ini yang berjudul “**Penerapan *Data Mining* untuk Menentukan Potensi Kepatuhan Wajib Pajak Berdasarkan data PBB Menggunakan Algoritma C4.5 (Studi Kasus UPT BPPD Ilir Barat II)**”

Selama penyusunan laporan tugas akhir ini tidak terlepas dari bimbingan dan bantuan dari berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada :

1. Bapak Jaidan Jauhari, S.Pd., M.T. Selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
2. Ibu Endang Lestari, S.Kom,M.T Selaku Ketua Program Studi Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Bapak Mgs. Afriyan Firdaus, M.IT dan Bapak Ali Ibrahim, M.T.. Selaku Dosen Pembimbing Tugas Akhir.
4. Bapak Sodikin, S.E., M.Si Selaku Kepala BPPD Kota Palembang.
5. Ibu Rahmawati, S.T., M.Si Selaku Kepala UPT BPPD Ilir Barat II Palembang.
6. Bambang Wicaksono Selaku Kepala Badan Kesatuan Bangsa dan Politik Kota Palembang yang telah memberikan izin pengambilan data.

7. Bapak M. Fahrul, Mbak Maretta Primaputri, dan Bapak Yanto Selaku pegawai BPPD yang telah memberikan data-data yang diperlukan untuk penelitian.
8. Kedua Orang tua dan seluruh keluarga besar yang telah memberikan do'a restu, motivasi, bantuan moril dan materil.
9. Eta, S.SI, Mutia, S.SI, Kurnia, S.SI, Alfa, S.SI, Reri, S.SI, Paulus, S.SI, dan Dhana Mawapres Ilkom yang selalu setia, memberikan masukan dan saling mambantu selama perjuangan kuliah.
10. Seluruh Teman Jurusan Sistem Informasi angkatan 2013.
11. Semua pihak yang telah memberi bantuan dan dukungan kepada penulis dan tidak dapat disebutkan satu-persatu.

Semoga ALLAH SWT membalas kebaikan dan ketulusan semua pihak yang telah membantu menyelesaikan laporan tugas akhir ini dengan melimpahkan rahmat dan karunia-Nya. Dalam penulisan tugas akhir ini, Penulis menyadari bahwa laporan ini masih belum sempurna, oleh karena itu penulis sangat mengharapkan kritik dan saran yang bersifat membangun demi menyempurnakan laporan ini. Penulis berharap semoga laporan ini dapat bermanfaat bagi para pembaca dan kita semua.

Indralaya, 2018

Bella Moulina

09031281320020

**PENERAPAN *DATA MINING* UNTUK MENENTUKAN  
POTENSI KEPATUHAN WAJIB PAJAK  
BERDASARKAN DATA PBB MENGGUNAKAN ALGORITMA C4.5  
(STUDI KASUS UPT BPPD ILIR BARAT II)**

**Oleh:**

**Bella Moulina      09031281320020**

**ABSTRAK**

Kegiatan operasional yang terjadi di UPT BPPD Ilir Barat II Palembang dapat menghasilkan dan mengumpulkan banyak data PBB. Data PBB yang ada di UPT BPPD Ilir Barat II saat ini hanya tersimpan di *database* aplikasi Pajak Online. Data yang dihasilkan terkumpul dan berukuran besar dan belum dimanfaatkan untuk dianalisis yang hasilnya berupa informasi berharga. Salah satunya pengolahan data menggunakan teknik *data mining*. Dengan adanya *data mining* diharapkan dapat meningkatkan pengolahan atas data dalam aplikasi Pajak Online sehingga dapat membuat data tersebut dimanfaatkan. Penelitian ini bertujuan untuk mengklasifikasikan potensi kepatuhan wajib pajak pbb menggunakan algoritma C4.5. Metodologi yang digunakan adalah *Nine Methodology* Kimball untuk *data warehouse* dan *decision tree* mengikuti *Cross Industry Standard Process for Data Mining* (CRISP-DM) untuk *data mining*. Sedangkan pengujian sistem menggunakan *Black Box Testing* untuk mengukur apakah sistem berjalan dengan baik sesuai dengan yang diharapkan. Hasil Penelitian ini akan memberikan informasi dalam mengelompokkan wajib pajak pbb sesuai dengan potensi kepatuhannya.

**Kata Kunci :** *data mining*, algoritma C4.5, potensi kepatuhan, wajib pajak pbb.



**DATA MINING IMPLEMENTATION TO DETERMINE  
POTENTIAL ADHERENCE TAXPAYER  
BASED ON PROPERTY TAXES DATA USING C4.5 ALGORITM  
( CASE STUDY IN UPT BPPD ILIR BARAT II)**

**By:**

**Bella Moulina            09031281320020**

**ABSTRACT**

*Operations that occur in the UPT BPPD Ilir Barat II Palembang can produce and collect the number of property taxes data. property taxes data in UPT BPPD Ilir Barat II only stored in Online Tax Application's database. Lots of data has not been used to analyze and produce valuable information. One of it is data processing using data mining techniques. The data mining implementation is expected to improve the processing of data in the application of Online Tax so that it can improve the data utilization. This study aims to classify potential adherence taxpayer of property taxes using C4.5 algoritm. Methodology that used Nine Methodology Kimball for data warehouse and decision tree to follow the stages Cross Industry Standard Process for Data Mining (CRISP-DM) for data mining. While testing the system using the Black Box Testing to measure whether the system is running properly in accordance with the expected. Results of this research will provide information in classify taxpayer of property taxes according their potential adherence.*

**Kata Kunci :** *data mining, algoritma C4.5, potential adherence, taxpayer of property taxes.*

## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>I</b>
<b>LEMBAR PENGESAHAN .....</b>	<b>II</b>
<b>HALAMAN PERSETUJUAN .....</b>	<b>III</b>
<b>LEMBAR PERNYATAAN .....</b>	<b>IV</b>
<b>HALAMAN PERSEMBAHAN .....</b>	<b>V</b>
<b>KATA PENGANTAR.....</b>	<b>VI</b>
<b>ABSTRAK .....</b>	<b>VIII</b>
<b>ABSTRACT .....</b>	<b>IX</b>
<b>DAFTAR ISI.....</b>	<b>X</b>
<b>DAFTAR GAMBAR.....</b>	<b>XIII</b>
<b>DAFTAR TABEL .....</b>	<b>XV</b>
<b>DAFTAR PERSAMAAN.....</b>	<b>XIX</b>
<b>BAB I.....</b>	<b>1</b>
<b>PENDAHULUAN.....</b>	<b>1</b>
1.1 LATAR BELAKANG .....	1
1.2 TUJUAN .....	4
1.3 MANFAAT.....	4
1.4 BATASAN MASALAH .....	4
<b>BAB II .....</b>	<b>6</b>
<b>TINJAUAN PUSTAKA .....</b>	<b>6</b>
2.1 <i>DATA WAREHOUSE</i> (GUDANG DATA) .....	6
2.1.1 Pengertian <i>Data Warehouse</i> (Gudang Data).....	6
2.1.2 Karakteristik <i>Data Warehouse</i> (Gudang Data) .....	6
2.1.3 Metodologi <i>Data Warehouse</i> (Gudang Data) .....	7
2.1.4 Skema <i>Data Warehouse</i> (Gudang Data) .....	9
2.1.4.1 Tabel Fakta.....	9
2.1.4.2 Tabel Dimensi .....	9
2.1.4.3 <i>Star Schema</i> (Skema Bintang) .....	9
2.2 <i>DATA MINING</i> (PENAMBANGAN DATA) .....	11
2.2.1 Definisi <i>Data Mining</i> (Penambangan Data).....	11
2.2.2 Arsitektur <i>Data Mining</i> (Penambangan Data) .....	11
2.2.3 Metodologi <i>Data Mining</i> (Penambangan Data).....	12
2.2.4 Pengelompokkan <i>Data Mining</i> (Penambangan Data).....	14
2.3 KLASIFIKASI.....	15
2.3.1 Pengertian Klasifikasi .....	15
2.3.2 Model klasifikasi.....	15

2.4	<i>DECISION TREE</i> (POHON KEPUTUSAN) .....	16
2.4.1	Pengertian <i>Decision Tree</i> (Pohon Keputusan) .....	16
2.4.2	Konsep <i>Decision Tree</i> (Pohon Keputusan) .....	17
2.5	ALGORITMA C4.5 .....	17
2.6	PAJAK .....	20
2.6.1	Pengertian Pajak .....	20
2.6.2	Pajak Bumi dan Bangunan (PBB) .....	20
2.6.3	Kepatuhan Wajib Pajak .....	21
<b>BAB III</b> .....		<b>24</b>
<b>METODOLOGI PENELITIAN</b> .....		<b>24</b>
3.1	KERANGKA PIKIR PENELITIAN .....	24
3.2	METODE PENGUMPULAN DATA .....	24
3.2.1	Jenis Data .....	24
3.2.2	Sumber Data .....	25
3.2.3	Pengumpulan Data .....	25
3.3	METODE ANALISIS DATA .....	26
3.3.1	Metodologi <i>Data Warehouse</i> (Gudang Data) .....	26
3.3.2	Skema Data Warehouse (Gudang Data) .....	28
3.3.3	Metodologi <i>Data Mining</i> (Penambangan Data) .....	29
3.3.3.1	Fase Pemahaman Bisnis .....	29
3.3.3.2	Fase Pemahaman Data .....	30
3.3.3.3	Fase Persiapan Data .....	30
3.3.3.4	Fase Pemodelan .....	31
3.3.3.5	Fase Evaluasi .....	31
3.3.3.6	Fase Pelaksanaan .....	32
<b>BAB IV</b> .....		<b>33</b>
<b>ANALISIS DATA</b> .....		<b>33</b>
4.1	ANALISIS PENGUMPULAN DATA .....	33
4.1.1	Wawancara .....	33
4.1.2	Observasi .....	33
4.1.3	Pengambilan Data .....	33
4.2	METODOLOGI <i>DATA WAREHOUSE</i> (GUDANG DATA) .....	33
4.2.1	Pemilihan Proses .....	34
4.2.2	Memilih <i>Grain</i> .....	34
4.2.3	Identifikasi dan Membuat Dimensi yang Sesuai .....	34
4.2.4	Memilih Fakta .....	34
4.2.5	Menentukan Data Pre-Kalkulasi dari Tabel Fakta .....	34
4.2.6	Melengkapi Tabel Dimensi .....	35
4.2.7	Memilih Durasi <i>Database</i> .....	38
4.2.8	Melacak Dimensi yang Berubah Secara Perlahan .....	39
4.2.9	Menentukan Prioritas dan Mode dari <i>Query</i> .....	39
4.2.9.1	<i>ETL Design</i> .....	39
4.3	Skema <i>Data Warehouse</i> .....	52
4.3.1	<i>Star Schema</i> (Skema Bintang) .....	52
4.3.2	Metadata .....	53
4.4	METODOLOGI <i>DATA MINING</i> (PENAMBANGAN DATA) .....	55

4.4.1	Fase Pemahaman Bisnis .....	55
4.4.2	Fase Pemahaman Data .....	56
4.4.3	Fase Persiapan Data .....	56
4.4.4	Fase Pemodelan.....	57
4.4.5	Fase Evaluasi.....	127
<b>BAB V</b>	<b>.....</b>	<b>129</b>
<b>HASIL DAN PEMBAHASAN</b>	<b>.....</b>	<b>129</b>
5.1	HASIL .....	129
5.1.1	Hasil Data Warehouse.....	129
5.1.2	Hasil Data Mining .....	129
5.2	PEMBAHASAN.....	130
5.2.1	Halaman <i>Login</i> .....	131
5.2.2	Halaman <i>Home</i> .....	131
5.2.3	Halaman Data PBB .....	132
5.2.4	Halaman OLAP Kepatuhan.....	134
5.2.5	Halaman C4.5.....	137
5.2.6	Halaman Kinerja .....	140
5.2.7	Halaman Penentu Keputusan .....	143
5.3	HASIL UJI COBA .....	144
5.3.1	<i>Test Case Login</i> .....	145
5.3.2	<i>Test Case Data PBB</i> .....	146
5.3.3	<i>Test Case Olap Kepatuhan</i> .....	147
5.3.4	<i>Test Case C4.5</i> .....	148
5.3.5	<i>Test Case Kinerja</i> .....	149
5.3.6	<i>Test Case Penentu Keputusan</i> .....	150
5.4	PENGUJIAN POHON KEPUTUSAN MENGGUNAKAN ALGORITMA C4.5 .....	151
5.4.1	<i>Input</i> data PBB oleh user.....	151
5.4.2	<i>Output</i> Hasil Keputusan Potensi Kepatuhan .....	152
<b>BAB VI</b>	<b>.....</b>	<b>153</b>
<b>KESIMPULAN DAN SARAN</b>	<b>.....</b>	<b>153</b>
6.1	KESIMPULAN .....	153
6.2	SARAN.....	154
<b>DAFTAR PUSTAKA</b>	<b>.....</b>	<b>155</b>
<b>LAMPIRAN</b>		

## DAFTAR GAMBAR

Gambar 2.1 Star Schema.....	10
Gambar 2.2 Siklus CRISP-DM.....	12
Gambar 2.3 <i>Decision Tree</i> .....	16
Gambar 2.4 Flowchart Algoritma C4.5 .....	19
Gambar 3.1 Kerangka Pikir Penelitian.....	24
Gambar 3.2 Siklus CRISP-DM.....	29
Gambar 3.3 Data PBB.....	30
Gambar 4.1 ETL Dimensi Waktu .....	40
Gambar 4.2 Data PBB.....	41
Gambar 4.3 ETL Dimensi OP .....	42
Gambar 4.4 ETL Dimensi WP.....	43
Gambar 4.5 ETL Dimensi Pekerjaan .....	44
Gambar 4.6 ETL Dimensi Status .....	45
Gambar 4.7 ETL Dimensi NJOP Bumi.....	46
Gambar 4.8 ETL Dimensi NJOP Bangunan .....	48
Gambar 4.10 ETL Dimensi Sppt.....	50
Gambar 4.11 ETL Fakta Kepatuhan .....	51
Gambar 4.12 Star Schema.....	52
Gambar 4.13 Metadata Dimensi Waktu.....	53
Gambar 4.14 Metadata Dimensi OP .....	53
Gambar 4.15 Metadata Dimensi WP .....	54
Gambar 4.16 Metadata Dimensi Pekerjaan.....	54
Gambar 4.17 Metadata Dimensi Status.....	54
Gambar 4.18 Metadata Dimensi Njop Bumi.....	54
Gambar 4.19 Metadata Dimensi Njop Bangunan .....	54
Gambar 4.20 Metadata Dimensi Sppt.....	54
Gambar 4.21 Atribut Penentu .....	57
Gambar 4.22 Pohon Keputusan Hasil Perhitungan Node 1 .....	62
Gambar 4.23 Pohon Keputusan Hasil Perhitungan Node 1.1 .....	66
Gambar 4.24 Pohon Keputusan Hasil Perhitungan Node 1.1.1 .....	68
Gambar 4.25 Pohon Keputusan Hasil Perhitungan Node 1.1.1.1 .....	70
Gambar 4.26 Pohon Keputusan Hasil Perhitungan Node 1.1.1.2 .....	72
Gambar 4.27 Pohon Keputusan Hasil Perhitungan Node 1.1.2 .....	73
Gambar 4.28 Pohon Keputusan Hasil Perhitungan Node 1.1.3 .....	75
Gambar 4.29 Pohon Keputusan Hasil Perhitungan Node 1.1.4 .....	77
Gambar 4.30 Pohon Keputusan Hasil Perhitungan Node 1.1.5 .....	79
Gambar 4.31 Pohon Keputusan Hasil Perhitungan Node 1.1.6 .....	81
Gambar 4.32 Pohon Keputusan Hasil Perhitungan Node 1.1.7 .....	83
Gambar 4.33 Pohon Keputusan Hasil Perhitungan Node 1.1.8 .....	85
Gambar 4.34 Pohon Keputusan Hasil Perhitungan Node 1.2.1 .....	86
Gambar 4.35 Pohon Keputusan Hasil Perhitungan Node 1.2.2 .....	87
Gambar 4.36 Pohon Keputusan Hasil Perhitungan Node 1.2.3 .....	88
Gambar 4.37 Pohon Keputusan Hasil Perhitungan Node 1.2.4 .....	89
Gambar 4.38 Pohon Keputusan Hasil Perhitungan Node 1.2.5 .....	90

Gambar 4.39 Pohon Keputusan Hasil Perhitungan Node 1.2.6 .....	91
Gambar 4.40 Pohon Keputusan Hasil Perhitungan Node 1.2.7 .....	92
Gambar 4.41 Pohon Keputusan Hasil Perhitungan Node 1.2.8 .....	93
Gambar 4.42 Pohon Keputusan Hasil Perhitungan Node 1.3.1 .....	94
Gambar 4.43 Pohon Keputusan Hasil Perhitungan Node 1.3.2 .....	96
Gambar 4.44 Pohon Keputusan Hasil Perhitungan Node 1.4.1 dan 1.4.2 .....	97
Gambar 4.45 Pohon Keputusan Hasil Perhitungan Node 1.4.3 .....	98
Gambar 4.46 Pohon Keputusan Hasil Perhitungan Node 1.4.4 .....	99
Gambar 4.47 Pohon Keputusan Hasil Perhitungan Node 1.4.5 .....	100
Gambar 4.48 Pohon Keputusan Hasil Perhitungan Node 1.4.6 .....	101
Gambar 4.49 Pohon Keputusan Hasil Perhitungan Node 1.4.7 .....	102
Gambar 4.50 Pohon Keputusan Hasil Perhitungan Node 1.4.8 .....	103
Gambar 4.51 Pohon Keputusan Hasil Perhitungan Node 1.5.1 .....	104
Gambar 4.52 Pohon Keputusan Hasil Perhitungan Node 1.5.2 .....	106
Gambar 4.53 Pohon Keputusan Hasil Perhitungan Node 1.6.1 .....	107
Gambar 4.54 Pohon Keputusan Hasil Perhitungan Node 1.6.2 .....	109
Gambar 4.55 Pohon Keputusan Hasil Perhitungan Node 1.7.1 .....	111
Gambar 4.56 Pohon Keputusan Hasil Perhitungan Node 1.7.2 .....	112
Gambar 4.57 Pohon Keputusan Hasil Perhitungan Node 1.8.1, 1.8.2 dan 1.8.6	113
Gambar 4.58 Pohon Keputusan Hasil Perhitungan Node 1.8.3 .....	114
Gambar 4.59 Pohon Keputusan Hasil Perhitungan Node 1.8.4 .....	115
Gambar 4.60 Pohon Keputusan Hasil Perhitungan Node 1.8.5 .....	116
Gambar 4.61 Pohon Keputusan Hasil Perhitungan Node 1.8.7 .....	117
Gambar 4.62 Pohon Keputusan Hasil Perhitungan Node 1.8.8 .....	118
Gambar 5.1 Halaman <i>Login</i> .....	132
Gambar 5.2 Halaman <i>Home</i> .....	132
Gambar 5.3 Halaman Menu Data PBB .....	134
Gambar 5.4 Halaman Menu Partisi Data .....	134
Gambar 5.5 Halaman Olap Berdasarkan Dimensi Waktu .....	135
Gambar 5.6 Halaman Olap Berdasarkan Dimensi Bumi .....	136
Gambar 5.7 Halaman Olap Berdasarkan Dimensi Njop Bangunan .....	136
Gambar 5.8 Halaman Olap Berdasarkan Dimensi Sppt.....	137
Gambar 5.9 Halaman Olap Berdasarkan Dimensi Pekerjaan .....	138
Gambar 5.10 Halaman Menu C4.5 Pohon Keputusan .....	139
Gambar 5.11 Halaman Menu Perhitungan.....	141
Gambar 5.0.12 Halaman Menu Kinerja C4.5 .....	142
Gambar 5.0.13 Halaman Menu Proses Kinerja.....	142
Gambar 5.14 Halaman Menu Tabel Penilaian .....	143
Gambar 5.15 Halaman Menu Grafik.....	143
Gambar 5.16 Halaman Menu Grafik Berdasarkan SPPT.....	144
Gambar 5.17 Halaman Menu Penentu Keputusan .....	144
Gambar 5.18 Input Data PBB .....	151
Gambar 5.19 Hasil Keputusan Potensi Kepatuhan .....	152
Gambar 5.20 Hasil Keputusan Pada Database .....	152
Gambar 5.19 Rule untuk Keputusan Potensi Kepatuhan .....	152

## DAFTAR TABEL

Tabel 4.1 Tabel Matrik <i>Grain</i> .....	34
Tabel 4.2 Tabel Dimensi .....	35
Tabel 4.3 Tabel Dimensi Waktu .....	36
Tabel 4.4 Tabel Dimensi OP .....	36
Tabel 4.5 Tabel Dimensi WP .....	37
Tabel 4.6 Tabel Dimensi Pekerjaan .....	37
Tabel 4.7 Tabel Dimensi Status .....	37
Tabel 4.8 Tabel Dimensi NJOP Bumi.....	38
Tabel 4.9 Tabel Dimensi NJOP Bangunan .....	38
Tabel 4.10 Tabel Dimensi SPPT .....	38
Tabel 4.11 Tabel Durasi <i>Database</i> .....	39
Tabel 4.12 Keterangan Atribut Data .....	42
Tabel 4.13 Tabel Kelompok NJOP Bumi .....	47
Tabel 4.14 Perhitungan Node 1.1 .....	65
Tabel 4.15 Perhitungan Node 1.1.1 .....	68
Tabel 4.16 Perhitungan Node 1.1.1.1 .....	69
Tabel 4.17 Perhitungan Node 1.1.1.2.....	71
Tabel 4.18 Perhitungan Node 1.1.2.....	72
Tabel 4.19 Perhitungan Node 1.1.2.1 .....	73
Tabel 4.20 Perhitungan Node 1.1.2.2.....	73
Tabel 4.21 Perhitungan Node 1.1.3.....	74
Tabel 4.22 Perhitungan Node 1.1.3.1 .....	75
Tabel 4.23 Perhitungan Node 1.1.3.2.....	75
Tabel 4.24 Perhitungan Node 1.1.4.....	76
Tabel 4.25 Perhitungan Node 1.1.4.1 .....	76
Tabel 4.26 Perhitungan Node 1.1.4.2.....	77
Tabel 4.27 Perhitungan Node 1.1.5.....	78
Tabel 4.28 Perhitungan Node 1.1.5.1.....	78
Tabel 4.29 Perhitungan Node 1.1.5.2.....	78
Tabel 4.30 Perhitungan Node 1.1.6.....	79
Tabel 4.31 Perhitungan Node 1.1.6.1.....	80
Tabel 4.32 Perhitungan Node 1.1.6.2.....	80
Tabel 4.33 Perhitungan Node 1.1.7.....	81
Tabel 4.34 Perhitungan Node 1.1.7.1.....	82
Tabel 4.35 Perhitungan Node 1.1.7.2.....	82
Tabel 4.36 Perhitungan Node 1.1.8.....	83
Tabel 4.37 Perhitungan Node 1.1.8.1.....	83
Tabel 4.38 Perhitungan Node 1.1.8.2.....	84
Tabel 4.39 Perhitungan Node 1.2.....	85
Tabel 4.40 Perhitungan Node 1.2.1.....	85
Tabel 4.41 Perhitungan Node 1.2.1.1.....	85
Tabel 4.42 Perhitungan Node 1.2.2.....	86
Tabel 4.43 Perhitungan Node 1.2.2.1.....	86
Tabel 4.44 Perhitungan Node 1.2.2.2.....	86

Tabel 4.45 Perhitungan Node 1.2.3.....	87
Tabel 4.46 Perhitungan Node 1.2.3.1.....	87
Tabel 4.47 Perhitungan Node 1.2.4.....	88
Tabel 4.48 Perhitungan Node 1.2.4.1.....	88
Tabel 4.49 Perhitungan Node 1.2.5.....	88
Tabel 4.50 Perhitungan Node 1.2.5.1.....	89
Tabel 4.51 Perhitungan Node 1.2.5.2.....	89
Tabel 4.52 Perhitungan Node 1.2.6.....	89
Tabel 4.53 Perhitungan Node 1.2.6.1.....	90
Tabel 4.54 Perhitungan Node 1.2.6.2.....	90
Tabel 4.55 Perhitungan Node 1.2.7.....	90
Tabel 4.56 Perhitungan Node 1.2.7.1.....	91
Tabel 4.57 Perhitungan Node 1.2.8.....	91
Tabel 4.58 Perhitungan Node 1.2.8.1.....	91
Tabel 4.59 Perhitungan Node 1.2.8.2.....	92
Tabel 4.60 Perhitungan Node 1.3.....	92
Tabel 4.61 Perhitungan Node 1.3.1.....	93
Tabel 4.62 Perhitungan Node 1.3.1.2.....	93
Tabel 4.63 Perhitungan Node 1.3.1.3.....	93
Tabel 4.64 Perhitungan Node 1.3.1.4.....	93
Tabel 4.65 Perhitungan Node 1.3.1.5.....	93
Tabel 4.66 Perhitungan Node 1.3.1.7.....	93
Tabel 4.67 Perhitungan Node 1.3.1.8.....	93
Tabel 4.68 Perhitungan Node 1.3.2.....	94
Tabel 4.69 Perhitungan Node 1.3.2.2.....	94
Tabel 4.70 Perhitungan Node 1.3.2.3.....	94
Tabel 4.71 Perhitungan Node 1.3.2.4.....	95
Tabel 4.72 Perhitungan Node 1.3.2.5.....	95
Tabel 4.73 Perhitungan Node 1.3.2.6.....	95
Tabel 4.74 Perhitungan Node 1.3.2.7.....	95
Tabel 4.75 Perhitungan Node 1.3.2.8.....	95
Tabel 4.76 Perhitungan Node 1.4.....	96
Tabel 4.77 Perhitungan Node 1.4.2.....	96
Tabel 4.78 Perhitungan Node 1.4.2.1.....	96
Tabel 4.79 Perhitungan Node 1.4.2.2.....	96
Tabel 4.80 Perhitungan Node 1.4.3.....	97
Tabel 4.81 Perhitungan Node 1.4.3.1.....	97
Tabel 4.82 Perhitungan Node 1.4.3.2.....	97
Tabel 4.83 Perhitungan Node 1.4.4.....	98
Tabel 4.84 Perhitungan Node 1.4.4.1.....	98
Tabel 4.85 Perhitungan Node 1.4.5.....	99
Tabel 4.86 Perhitungan Node 1.4.5.1.....	99
Tabel 4.87 Perhitungan Node 1.4.6.....	100
Tabel 4.88 Perhitungan Node 1.4.6.1.....	100
Tabel 4.89 Perhitungan Node 1.4.6.2.....	100
Tabel 4.90 Perhitungan Node 1.4.7.....	101
Tabel 4.91 Perhitungan Node 1.4.7.1.....	101
Tabel 4.92 Perhitungan Node 1.4.8.....	102



Tabel 4.93 Perhitungan Node 1.4.8.1.....	102
Tabel 4.94 Perhitungan Node 1.4.8.2.....	102
Tabel 4.95 Perhitungan Node 1.5.....	103
Tabel 4.96 Perhitungan Node 1.5.1.....	103
Tabel 4.97 Perhitungan Node 1.5.1.1.....	103
Tabel 4.98 Perhitungan Node 1.5.1.2.....	103
Tabel 4.99 Perhitungan Node 1.5.2.....	104
Tabel 4.100 Perhitungan Node 1.5.2.2.....	104
Tabel 4.101 Perhitungan Node 1.5.2.3.....	104
Tabel 4.102 Perhitungan Node 1.5.2.4.....	104
Tabel 4.103 Perhitungan Node 1.5.2.5.....	104
Tabel 4.104 Perhitungan Node 1.5.2.7.....	104
Tabel 4.105 Perhitungan Node 1.5.2.8.....	105
Tabel 4.106 Perhitungan Node 1.6.....	105
Tabel 4.107 Perhitungan Node 1.6.1.....	106
Tabel 4.108 Perhitungan Node 1.6.1.2.....	106
Tabel 4.109 Perhitungan Node 1.6.1.3.....	106
Tabel 4.110 Perhitungan Node 1.6.1.4.....	106
Tabel 4.111 Perhitungan Node 1.6.2.....	107
Tabel 4.112 Perhitungan Node 1.6.2.1.....	107
Tabel 4.113 Perhitungan Node 1.6.2.2.....	107
Tabel 4.114 Perhitungan Node 1.6.2.3.....	107
Tabel 4.115 Perhitungan Node 1.6.2.4.....	107
Tabel 4.116 Perhitungan Node 1.6.2.5.....	107
Tabel 4.117 Perhitungan Node 1.6.2.6.....	107
Tabel 4.118 Perhitungan Node 1.6.2.7.....	107
Tabel 4.119 Perhitungan Node 1.6.2.8.....	108
Tabel 4.120 Perhitungan Node 1.7.....	108
Tabel 4.121 Perhitungan Node 1.7.1.....	109
Tabel 4.122 Perhitungan Node 1.7.1.2.....	109
Tabel 4.123 Perhitungan Node 1.7.1.3.....	109
Tabel 4.124 Perhitungan Node 1.7.1.4.....	109
Tabel 4.125 Perhitungan Node 1.7.1.6.....	109
Tabel 4.126 Perhitungan Node 1.7.1.7.....	109
Tabel 4.127 Perhitungan Node 1.7.1.8.....	109
Tabel 4.128 Perhitungan Node 1.7.2.....	110
Tabel 4.129 Perhitungan Node 1.7.2.2.....	110
Tabel 4.130 Perhitungan Node 1.7.2.3.....	110
Tabel 4.131 Perhitungan Node 1.7.2.4.....	111
Tabel 4.132 Perhitungan Node 1.7.2.5.....	111
Tabel 4.133 Perhitungan Node 1.7.2.6.....	111
Tabel 4.134 Perhitungan Node 1.7.2.7.....	111
Tabel 4.135 Perhitungan Node 1.7.2.8.....	111
Tabel 4.136 Perhitungan Node 1.8.....	112
Tabel 4.137 Perhitungan Node 1.8.2.....	112
Tabel 4.138 Perhitungan Node 1.8.2.1.....	112
Tabel 4.139 Perhitungan Node 1.8.2.2.....	112
Tabel 4.140 Perhitungan Node 1.8.3.....	113

Tabel 4.141 Perhitungan Node 1.8.3.1.....	113
Tabel 4.142 Perhitungan Node 1.8.4.....	114
Tabel 4.143 Perhitungan Node 1.8.4.1.....	114
Tabel 4.144 Perhitungan Node 1.8.4.2.....	114
Tabel 4.145 Perhitungan Node 1.8.5.....	115
Tabel 4.146 Perhitungan Node 1.8.5.1.....	115
Tabel 4.147 Perhitungan Node 1.8.5.2.....	115
Tabel 4.148 Perhitungan Node 1.8.7.....	116
Tabel 4.149 Perhitungan Node 1.8.7.1.....	116
Tabel 4.150 Perhitungan Node 1.8.7.2.....	116
Tabel 4.151 Perhitungan Node 1.8.8.....	117
Tabel 4.152 Perhitungan Node 1.8.8.1.....	117
Tabel 4.153 Perhitungan Node 1.8.8.2.....	117
Tabel 4.154 Hasil Pengujian dengan Matriks Konfusi .....	128
Tabel 5.1 Hasil Pengujian .....	144
Tabel 5.2 <i>Test Case Login</i> .....	146
Tabel 5.3 <i>Test Case Data PBB</i> .....	146
Tabel 5.4 <i>Test Case Olap Kepatuhan</i> .....	147
Tabel 5.5 <i>Test Case C4.5</i> .....	148
Tabel 5.6 <i>Test Case Kinerja</i> .....	149
Tabel 5.7 <i>Test Case Penentu Keputusan</i> .....	150

## DAFTAR PERSAMAAN

Persamaan 2.1 Rumus Entrophy .....	17
Persamaan 2.2 Information Gain.....	18
Persamaan 2.3 Gain Ratio .....	18
Persamaan 2.4 Split Info .....	18
Persamaan 3.1 Rumus Akurasi .....	32



# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Menurut Soemarsono dalam (Handayani dkk, 2014) menyatakan bahwa “Pajak Bumi dan Bangunan (PBB) adalah pajak tidak langsung yang dikenakan atas benda berupa harta tak gerak yaitu bumi dan bangunan”. Penyetoran PBB dapat dilakukan melalui Anjungan Tunai Mandiri (ATM), langsung datang ke Bank yang ditunjuk SPPT, atau Unit Pelaksana Teknis (UPT) setiap wilayah kecamatan.

UPT Badan Pengelola Pendapatan Daerah Iilir Barat II merupakan salah satu Unit Pelaksana Teknis Badan Pengelola Pajak Daerah (BPPD) Kota Palembang yang bergerak di bidang pengolahan pajak daerah. UPT BPPD Iilir Barat II memiliki tugas untuk memberikan laporan hasil pembayaran PBB oleh Wajib Pajak khususnya di wilayah kecamatan Iilir Barat II. Data PBB yang diberikan oleh UPT merupakan data yang perlu diolah untuk mengetahui informasi yang terkandung didalamnya.

Laporan data PBB tersebut bisa memberikan informasi bagi BPPD. Laporan dari data PBB tersebut yang saat ini dijadikan oleh BPPD untuk melakukan kebijakan-kebijakan yang bertujuan dalam meningkatkan penerimaan PBB.

Data PBB yang ada di UPT BPPD Iilir Barat II saat ini hanya tersimpan di *database*. Laporan yang dihasilkan dari aplikasi Pajak Online terkumpul dan berukuran besar dan belum dimanfaatkan untuk dianalisis yang hasilnya berupa

informasi berharga. Salah satunya pengolahan data menggunakan teknik *data mining* atau pembangan data. Dengan adanya *data mining* diharapkan dapat meningkatkan pengolahan atas data dalam aplikasi Pajak *Online* sehingga dapat membuat data tersebut dimanfaatkan.

Sebelumnya penelitian mengenai implementasi teknik *data mining* di bidang perpajakan sudah pernah dilakukan. Penelitian yang dilakukan (Johari & Mardiani, 2013), mengenai pemanfaatan *data mining* untuk mengelompokkan jumlah kendaraan menggunakan algoritma *Expectation Maximation* (EM). Penelitian ini menghasilkan 15 *cluster* dari data pembayaran pajak kendaraan bermotor selama 3 tahun. Serta dapat disimpulkan bahwa pendapatan pajak kendaraan bermotor tertinggi berasal dari kendaraan bermerk Mitsubishi dan status pengendara adalah karyawan swasta. Namun pada penelitian ini, penerapan *data mining* berfokus pada pengelompokkan jumlah kendaraan di suatu wilayah dan belum menganalisis potensi kepatuhan wajib pajak dalam membayar pajak kendaraan. Selain itu, teknik *data mining* yang digunakan adalah clustering dan hanya ruang lingkup pajak kendaraan.

Menurut Khamami (Potensi Pemanfaatan *Data Mining* di Kementerian Keuangan, 2014), *data mining* dapat membantu organisasi untuk mengambil informasi paling penting dari *data warehouse*. Contoh potensi pemanfaatannya yaitu *data mining* untuk memprediksi siapa Wajib Pajak (WP) yang tidak patuh dan seberapa besar tingkat kepatuhannya. WP patuh atau tidak dapat diidentifikasi dengan menggunakan metode pohon keputusan. Namun, artikel ini hanya membahas potensi dan belum melakukan penelitian mengenai penerapannya.

Metode *decision tree* atau pohon keputusan adalah salah satu teknik dalam *data mining* yang dapat digunakan untuk klasifikasi data. Pada umumnya ciri-ciri untuk kasus yang cocok untuk diterapkan *decision tree* menurut Santoso dalam (Yolanda, 2015) yaitu:

1. Data dinyatakan dengan pasangan atribut lainnya. Misalnya atribut temperature dan nilainya adalah dingin.
2. Label atau *output* data biasanya bernilai diskrit. Label ini bisa bernilai ‘ya’ atau ‘tidak’, ‘sakit’ atau ‘tidak sakit’, ‘diterima’ atau ‘ditolak’.

Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk membentuk *decision tree* yang dikembangkan oleh Ross Quinlan. Algoritma C4.5 memiliki beberapa kelebihan misalnya dapat mengolah data numerik dan diskret, dapat menangani nilai atribut yang hilang, menghasilkan aturan-aturan yang mudah diinterpretasikan dan tercepat diantara algoritma-algoritma yang lain. Penelitian yang dilakukan (Dhika, 2015), membandingkan algoritma C4.5, *Naïve Bayes*, dan *Neural Network* dalam pemilihan mitra kerja penyedia jasa transportasi. Untuk mengukur ketiga kinerja algoritma tersebut digunakan *Confusion Matrix* dan Kurva ROC. Hasilnya adalah algoritma yang memiliki nilai accuracy dan ROC paling tinggi adalah algoritma C4.5 kemudian *Naïve Bayes* dan terakhir *Neural Network*. Berdasarkan latar belakang di atas, penulis tertarik untuk melakukan penelitian dengan judul **“Penerapan Data Mining untuk Menentukan Potensi Kepatuhan Wajib Pajak Berdasarkan Data PBB Menggunakan Algoritma C4.5 (Studi Kasus UPT BPPD Ilir Barat II)”**.

## 1.2 Tujuan

Tujuan dari tugas akhir ini adalah untuk menerapkan *data mining* dalam menentukan potensi kepatuhan wajib pajak berdasarkan data Pajak Bumi dan Bangunan (PBB) Kecamatan Ilir Barat II menggunakan algoritma C4.5.

## 1.3 Manfaat

Manfaat yang diharapkan dalam tugas akhir ini, antara lain :

1. Meningkatkan pengolahan atas data sehingga dapat membuat data tersebut dimanfaatkan.
2. Mengidentifikasi potensi kepatuhan wajib pajak apakah patuh atau tidak.
3. Memenuhi kebutuhan informasi pihak eksekutif untuk menunjang keputusan yang bertujuan dalam meningkatkan penerimaan PBB.

## 1.4 Batasan Masalah

Agar permasalahan yang dibahas tidak terlalu luas dan dapat lebih fokus dalam melakukan penelitian, maka dilakukan pembatasan masalah, antara lain:

1. Penelitian ini hanya akan menggunakan data PBB Kecamatan Ilir Barat II.
2. Informasi yang akan digali pada *data mining* adalah potensi kepatuhan wajib pajak.
3. Menggunakan metode *Nine Step Methodology* Kimball untuk *data warehouse*.
4. Menggunakan teknik *decision tree* (pohon keputusan) dengan algoritma C4.5 mengikuti tahapan metode *Cross Industry Standard process for Data Mining* (CRISP-DM) untuk *data mining*.



5. Sistem yang akan dikembangkan adalah sistem klasifikasi potensi kepatuhan wajib pajak berbasis web dengan menggunakan bahasa pemrograman PHP.
6. Pada penelitian ini faktor yang mempengaruhi kepatuhan wajib pajak dalam membayar PBB ada 4 yaitu Pekerjaan, NJOP Bumi, NJOP Bangunan dan SPPT.
7. Hasil pada data target yang akan dianalisis pada tugas akhir ini adalah apakah wajib pajak berpotensi patuh atau tidak tanpa seberapa besar tingkat kepatuhannya.

## DAFTAR PUSTAKA

- Abdillah, G., Putra, F. A., & Renaldi, F. (2016). Penerapan Data Mining Pemakaian Air Pelanggan untuk Menentukan Klasifikasi Potensi Pemakaian Air Pelanggan Baru PDAM Tirta Raharja Menggunakan Algoritma K-Means. *SENTIKA 2016*, 1-9.
- Ananda. (2015). *Faktor-faktor yang Mempengaruhi Kepatuhan Wajib Pajak dalam Membayar Pajak Bumi dan Bangunan dengan Pendapatan Masyarakat Sebagai Variabel Moderating*. Medan: Universitas Sumatera Utara.
- Dhika, H. (2015). Kajian Komparasi Penerapan Algoritma C4.5, Naive Bayes, dan Neural Network dalam Pemilihan Mitra Kerja Penyedia Jasa Transportasi : Studi Kasus CV Viradi Global Pratama. *SNIT 2015*, 1-6.
- Handayani, W. T., Santoso, S., & Sohidin. (2014). Analisis Faktor-faktor yang Mempengaruhi Realisasi Penerimaan Pajak Bumi dan Bangunan pada Kecamatan Jebres Kota Surakarta. *Jupe UNS Vol 2 No 3 Hal 224 s/d 237*, 4.
- Hidayah, N. U. (2014). *Analisis Faktor-faktor yang Mempengaruhi Kepatuhan Wajib Pajak dalam Membayar PBB*. Yogyakarta: Universitas Negeri Yogyakarta.
- Johari, I., Soeyapto, D., & Mardiani. (2013). Penerapan Data Mining untuk Data Jumlah Kendaraan Menggunakan Algoritma EM pada Dispenda Kota Palembang. *Jurnal STMIK MDP*, 1-10.
- Kahono, S. (2003). *Pengaruh Sikap Wajib Pajak Terhadap Kepatuhan Wajib Pajak dalam Pembayaran Pajak Bumi dan Bangunan (Studi Empiris di Wilayah KP.PBB Semarang)*. Semarang: Tesis Mahasiswa Magister Sains Akuntansi.
- Khamami, H. (2014, Desember 19). *Potensi Pemanfaatan Data Mining di Kementerian Keuangan*. Retrieved Maret 8, 2017, from Website BPPK Kemenkeu: <http://www.bppk.kemenkeu.go.id/publikasi/artikel/419-artikel-teknologi-informasi/20242-potensi-pemanfaatan-data-mining-di-kementerian-keuangan>
- Noviyanti, S., Effendi, R., & W, C. Y. (n.d.). Pengaruh Kesadaran Wajib Pajak, Pengetahuan dan Pemahaman Peraturan Perpajakan, Kualitas Pelayanan

dan Ketegasan Sanksi Perpajakan Terhadap Kepatuhan WPOP (Studi Kasus KPP Pratama Ilir Barat Palembang). 3-4.

Samudra, H. D. (n.d.). Pengaruh SPPT, Sanksi, Pendapatan Wajib Pajak Terhadap Kepatuhan Wajib Pajak Bumi dan Bangunan. *Jurnal Universitas Dian Nuswantoro*, 17.

*Undang Undang Republik Indonesia Nomor 28 Tahun 2009 tentang Perubahan Keempat atas Undang-undang Nomor 18 Tahun 1997 tentang Pajak Daerah dan Retribusi Daerah.* (n.d.).

Untari, D. (n.d.). Data Mining untuk Menganalisa Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Metode Decision Tree C4.5. *Jurnal Fakultas Ilmu Komputer*, 1-8.

Widodo, P. P., Handayanto, R. T., & Herlawati. (2013). *Penerapan Data Mining dengan Matlab*. Bandung: Rekayasa Sains.

Yolanda, Y. (2015). *Penerapan Data Mining Dalam Klasifikasi Hubungan Antara Kejadian Katarak dengan Diabetes Melitus Menggunakan Algoritma C4.5*. Palembang: Jurusan Sistem Informasi.

Yusnidar, J., Sunarti, & Prasetya, A. (2015). Faktor-faktor yang Mempengaruhi Kepatuhan Wajib Pajak dalam Pembayaran Pajak Bumi dan Bangunan Perdesaan dan Perkotaan. *Jurnal Perpajakan (JEJAK) Vol 1 No.1*, 3.