

**EVALUASI DISTRIBUSI HASIL SELEKSI
PPDB SMA NEGERI DI KOTA SURAKARTA
MENGUNAKAN METODE DECISION TREE**

NASKAH PUBLIKASI

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS KOMUNIKASI DAN INFORMATIKA



Diajukan oleh :

Aziz Ridha Utama

Nurgiyatna, M.Sc, Ph.D

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS MUHAMMADIYAH SURAKARTA**

JULI 2014

HALAMAN PENGESAHAN

Publikasi Ilmiah dengan Judul

**EVALUASI DISTRIBUSI HASIL SELEKSI
PPDB SMA NEGERI DI KOTA SURAKARTA
MENGUNAKAN METODE DECISION TREE**

Yang dipersiapkan dan disusun oleh :

Aziz Ridha Utama

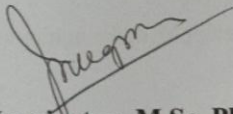
L 200 100 135

Telah disetujui pada :

Hari :

Tanggal :

Pembimbing



Nurgiyatna, M.Sc, Ph.D

NIK : 881

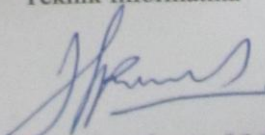
Publikasi ilmiah ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar sarjana

Tanggal

Mengetahui,

Ketua Program Studi

Teknik Informatika



Dr. Heru Supriyono, M.Sc.

NIK : 970



UNIVERSITAS MUHAMMADIYAH SURAKARTA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
PROGRAM STUDI TEKNIK INFORMATIKA

Jl. A Yani Tromol Pos 1 Pabelan Kartasura Telp. (0271)717417, 719483 Fax (0271) 714448
Surakarta 57102 Indonesia. Web: <http://informatika.ums.ac.id> Email: informatika@fki.ums.ac.id

SURAT KETERANGAN LULUS PLAGIASI

/A.3-II.3/INF-FKI/VII/2014

Assalamu'alaikum Wr. Wb

Biro Skripsi Program Studi Teknik Informatika menerangkan bahwa :

Nama : AZIZ RIDHA UTAMA
NIM : L200100135
Judul : EVALUASI DISTRIBUSI HASIL SELEKSI PPDB SMA NEGERI DI
KOTA SURAKARTA MENGGUNAKAN METODE DECISION
TREE
Program Studi : Teknik Informatika
Status : **Lulus**

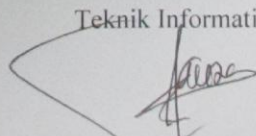
Adalah benar-benar sudah lulus pengecekan plagiasi dari Naskah Publikasi Skripsi, dengan menggunakan aplikasi Turnitin.

Demikian surat keterangan ini dibuat agar dipergunakan sebagaimana mestinya.

Wassalamu'alaikum Wr. Wb

Surakarta, 10 Juli 2014

Biro Skripsi
Teknik Informatika


Fauzan Natsir, S.Kom

Turnitin Originality Report

**EVALUASI DISTRIBUSI HASIL SELEKSI
PPOB SMA NEGERI DI KOTA
SURABAYA MENGGUNAKAN METODE
DECISION TREE** by Aziz Ridha Utama

From September 2014 (publikasi maret 2014)

Similarity Index	Similarity by Source
20%	Internet Sources 18%
	Publications 0%
	Student Papers 0%

Processed on 08-Jul-2014 15:02 WIT sources:
ID: 438475760
Word Count: 2733

SOURCES:

- Word Count: 2733

1 4% match (Internet from 12-Jun-2014)
<http://vi-subang.blogspot.com/2013/05/pohon-keputusan-id3-dan-c45-menggunakan.html>

2 2% match (student papers from 16-Jun-2014)
Class: publikasi maret 2014
Assignment:
Paper ID: 435061278

3 2% match (Internet from 31-May-2012)
http://ppdbsolo.net/docs/Panduan_PPDB_2011_Kota_Surakarta.pdf

4 1% match (Internet from 25-Jun-2013)
<http://www.unisbank.ac.id/ojs/index.php/fil1/article/download/348/225>

5 1% match (Internet from 09-Dec-2012)
http://www.bmkp.go.id/Puslitbang/filePDF/Dokumen_585_Volume_12_Nomor_2_September_2011_Pemertan_Dala_Mingam.pdf

6 1% match (student papers from 02-Jul-2013)
Submitted to iGroup on 2013-07-02

7 1% match (student papers from 04-Feb-2014)
Class: publikasi maret 2014
Assignment:
Paper ID: 393366374

8 1% match (student papers from 12-Jun-2014)
Class: publikasi maret 2014
Assignment:
Paper ID: 434379926

9 1% match (Internet from 16-Jun-2013)
http://digilib.mercubuana.ac.id/e_perpus/artikel.php

10 1% match (Internet from 16-Apr-2010)
<http://warnadunia.com/tag/smk/>

11 1% match (student papers from 08-Jul-2014)
Class: publikasi maret 2014
Assignment:
Paper ID: 438403484

12 1% match (student papers from 04-Feb-2014)
Class: publikasi maret 2014
Assignment:
Paper ID: 393366361

13 < 1% match (Internet from 06-Jul-2014)

EVALUASI DISTRIBUSI HASIL SELEKSI PPDB SMA NEGERI DI KOTA SURAKARTA MENGUNAKAN METODE DECISION TREE

Aziz Ridha Utama, Nurgiyatna
Teknik Informatika, Fakultas Komunikasi dan Informatika,
Universitas Muhammadiyah Surakarta
Email : aziz.ridha.utama@gmail.com

Pesatnya perkembangan teknologi informasi mampu membuat sebuah sistem informasi dapat menyimpan berbagai jenis data dalam jumlah besar. Jika sistem informasi tersebut bersifat penting dalam menunjang suatu kegiatan maka data - data yang tersimpan pada sistem tersebut juga bersifat penting. Apabila data - data yang ada tersedia dalam jumlah besar maka data tersebut dapat dimanfaatkan untuk membentuk suatu informasi yang memiliki nilai guna bagi pemilik sistem informasi tersebut. Salah satu contohnya adalah banyaknya data hasil seleksi Penerimaan Peserta Didik Baru (PPDB) tingkat SMA Negeri di Kota Surakarta yang dimiliki oleh Dinas Pendidikan Pemuda dan Olahraga Kota Surakarta. Cara mengatasi permasalahan pembentukan informasi dari kumpulan data tersebut dengan cara menerapkan proses *data mining*. Adapun metode yang dipakai pada proses *data mining* untuk menganalisa penelitian ini adalah metode *decision tree* dengan *algoritma ID3*. Pada penerapan *algoritma ID3* ini diperlukan perhitungan nilai *entropy* dan *information gain* dari tiap atribut yang ada dalam *data training*. Data hasil seleksi PPDB tingkat SMA Negeri di Kota Surakarta digunakan sebagai *data training*. Berdasarkan analisa dari *data training* yang ada menunjukkan bahwa adanya suatu atribut yang mempunyai dominasi cukup tinggi terhadap hasil seleksi PPDB tersebut, atribut tersebut adalah atribut akumulasi nilai mata pelajaran ujian akhir nasional tingkat SMP. Kemudian ketepatan dalam penerapan proses *data mining* dengan metode *decision tree algoritma ID3* juga berdasarkan seberapa banyak data yang digunakan sebagai *data training*.

Kata Kunci : *Data Mining, Decision tree, Algoritma ID3, Data Training, Entropy, Information Gain*

PENDAHULUAN

Pesatnya perkembangan teknologi informasi pada masa sekarang ini, mampu membuat sebuah sistem informasi dapat menyimpan berbagai jenis data dalam jumlah besar. Jika sistem informasi yang menyimpan data dalam jumlah besar tersebut bersifat penting maka data - data yang ada pada sistem informasi tersebut juga bersifat penting. Apabila data - data tersebut tersedia dalam jumlah besar maka

data tersebut dapat dimanfaatkan untuk membentuk suatu informasi yang memiliki nilai guna bagi pemilik sistem informasi tersebut. Berdasarkan pentingnya *knowledge* atau pengetahuan terhadap suatu pembentukan pola informasi yang memiliki nilai guna sebagai bagian dari proses pemanfaatan data dalam jumlah besar maka mendorong munculnya suatu cabang ilmu baru untuk mengatasi permasalahan tersebut, suatu

cabang ilmu yang biasa disebut ilmu *data mining*.

Salah satu metode yang digunakan dalam *data mining* adalah metode *decision tree*, *Decision tree* adalah suatu metode yang menggunakan representasi model struktur pohon (*tree*) dimana setiap simpulnya (*node*) merupakan representasi dari atribut, cabangnya (*edge*) merupakan representasi dari nilai yang ada pada atribut, dan daunnya (*leaf*) merupakan representasi dari kelas target.

Salah satu contoh penerapan kemajuan teknologi informasi tersebut adalah dengan adanya suatu sistem penerimaan peserta didik baru online (PPDB Online) adalah suatu sistem informasi yang berhubungan dengan penyeleksian penerimaan peserta didik baru tingkat SMP/MTs dan SMA/MA secara transparan dan *real time* berbasis pada perkembangan teknologi informasi. Sistem ini merupakan sistem yang berada pada suatu jaringan sehingga dapat diakses melalui sms maupun internet.

Pada sistem tersebut menyimpan berbagai data yang berhubungan dengan hasil seleksi penerimaan peserta didik baru tingkat SMP/MTs dan SMA/MA. Dengan mengimplementasikan *data mining* pada data hasil seleksi PPDB tingkat SMA Negeri di Kota Surakarta, analisa *data mining* tersebut nantinya dapat digunakan sebagai suatu informasi

yang memiliki nilai guna bagi Dinas Pendidikan Pemuda dan Olahraga Kota Surakarta terkhusus pada Dinas Pendidikan.

Berdasarkan hal tersebut maka peneliti mengangkat judul “Evaluasi Distribusi Hasil Seleksi PPDB SMA Negeri di Kota Surakarta Menggunakan Metode *Decision Tree*”.

LANDASAN TEORI

1. *Data Mining*

Data mining, sering juga disebut dalam suatu istilah lain sebagai *knowledge discovery in database* (KDD). *Knowledge discovery in database* adalah suatu kegiatan yang meliputi aspek pengumpulan data, pemakaian data, historis untuk menemukan keteraturan dari suatu pola atau hubungan dalam set data berukuran besar (Santoso, 2007).

2. Pohon Keputusan (*Decision Tree*)

Decision tree merupakan salah satu metode dalam *data mining* dengan menggunakan representasi dari *tree* atau struktur pohon. Dimana setiap simpulnya (*node*) merupakan representasi dari atribut, cabangnya (*edge*) merupakan representasi dari nilai yang ada pada atribut, dan daunnya (*leaf*) merupakan representasi dari kelas target.

3. Algoritma ID3

Algoritma ID3 (Iterative Dichotomiser Three) adalah salah satu algoritma pembentuk struktur pohon keputusan (*decision tree*). *Algoritma ID3* ini dikembangkan oleh seorang tokoh yang bernama J. Ross Quinlan. *Algoritma ID3* ini membangun pohon keputusan (*decision tree*) secara *top - down* (atas - bawah). *Algoritma ID3* ini menggunakan konsep perhitungan dari nilai *entropy* dan nilai *information gain*.

Rumus untuk perhitungan dari pencarian nilai *entropy* ditunjukkan dalam persamaan matematis sebagai berikut :

$$\text{Entropy}(S) = -P_+ \log_2 P_+ - P_- \log_2 P_-$$

Dimana :

S merupakan ruang (data) sample yang digunakan sebagai *data training*, P_+ merupakan jumlah data yang bersolusi positif (mendukung) pada data sample untuk kriteria tertentu, dan P_- merupakan jumlah data yang bersolusi negatif (tidak mendukung) pada data sample untuk kriteria tertentu.

Rumus untuk perhitungan dari pencarian suatu nilai *information gain* dari suatu atribut A ditunjukkan dalam persamaan matematis sebagai berikut :

$$\text{Gain}(S,A) \equiv \text{Entropy}(S) - \sum_{V \in \text{Value}(A)} \frac{S_V}{S} \text{Entropy}(S_V)$$

Dimana :

A adalah atribut yang ada pada data training, V adalah suatu nilai yang mungkin untuk atribut A, $\text{Value}(A)$ adalah himpunan yang mungkin untuk atribut A, $|S_V|$ adalah jumlah sample untuk nilai V, $|S|$ adalah jumlah seluruh sample data yang digunakan sebagai *data training*, dan $\text{Entropy}(S_V)$ adalah nilai *entropy* untuk sample - sample yang memiliki nilai V.

METODOLOGI PENELITIAN

1. Identifikasi Masalah

Langkah awal dari proses penelitian ini adalah mencari referensi - referensi dan hasil - hasil penelitian yang membahas tentang *data mining* metode *decision tree* dengan *algoritma ID3*. Hasil pencarian tersebut digunakan sebagai bahan pertimbangan pembelajaran, dan pengambilan keputusan dalam melaksanakan penelitian tentang proses evaluasi distribusi hasil seleksi PPDB tingkat SMA Negeri di Kota Surakarta menggunakan metode *decision tree* karena memang belum ada yang melakukan penelitian yang berkaitan dengan permasalahan tersebut. Juga terjadinya penumpukan data - data hasil seleksi PPDB tingkat SMA Negeri yang tiap tahunnya terus bertambah.

2. Penentuan Atribut Dominan

Atribut dominan merupakan atribut dari kriteria calon pendaftar yang sering diutamakan dalam proses analisa. Atribut tersebut berisi data - data yang sangat vital dalam proses pembentukan *decision tree*.

Dalam studi kasus evaluasi distribusi hasil seleksi PPDB tingkat SMA Negeri di Kota Surakarta menggunakan metode *decision tree*, ada 5 atribut dominan yang dipakai

pada penelitian ini, atribut tersebut adalah asal peserta didik baru, *gender* peserta didik baru, prestasi yang dimiliki peserta didik baru, total akumulasi nilai mata pelajaran UAN tingkat SMP, dan sekolah SMA Negeri hasil seleksi PPDB. Gambaran penentuan atribut dominan dapat dilihat pada **Tabel 1**.

Tabel 1. Penentuan Atribut dari Hasil PPDB

Variabel	Atribut	Objek	Kategori
X1	Asal	Dalam kota Luar kota	DK LK
X2	Gender	Laki - laki Perempuan	L P
X3	Prestasi	DL dan prestasi yang memiliki Nilai Prestasi = 0	Ada Tidak
X4	Nilai_Mapel	Nilai mapel > 37,5 35 < Nilai mapel <= 37,5 Nilai mapel <= 35	Range1 Range2 Range3
Y	Hasil_PPDB	SMAN_1 SMAN_3 SMAN_4	SMAN_1 SMAN_3 SMAN_4

3. Penentuan Sample Data

Untuk mendapatkan jumlah sample data yang diperlukan dalam penelitian ini dengan menggunakan rumus persamaan *slovin* seperti yang tertulis dalam rumus persamaan berikut ini.

$$n = \frac{N}{1 + N e^2}$$

Dimana :

n adalah ukuran jumlah sample data yang dibutuhkan, N adalah ukuran atau jumlah populasi keseluruhan data yang ada, e adalah persentase toleransi dari

ketidaktepatan data sample yang dibutuhkan.

4. Penentuan Metode

Metode yang digunakan dalam penelitian ini adalah metode *decision tree* dengan *algoritma ID3* sebagai suatu metode dalam membangun struktur pohon dari hasil seleksi PPDB tingkat SMA Negeri di Kota Surakarta. *Algoritma ID3* ini menggunakan konsep pencarian perhitungan nilai *entropy* dan nilai *information gain*.

HASIL DAN PEMBAHASAN

1. Penentuan Sample Data

Apabila diketahui jumlah peserta didik baru hasil seleksi PPDB tingkat SMA Negeri terkhusus SMA Negeri 1 Surakarta, SMA Negeri 3 Surakarta, dan SMA Negeri 4 Surakarta memiliki jumlah peserta didik baru sebanyak 1218 dan untuk ketoleransian

kesalahan sample data dari data jumlah peserta didik baru tersebut adalah 5%. Maka jumlah sample data yang diperlukan yaitu :

$$n = \frac{N}{1 + N e^2}$$

$$n = 1218 / (1 + (1218 \times (0,05)^2))$$

$$n = 1218 / (1 + (1218 \times 0,0025))$$

$$n = 1218 / (1 + 3,045)$$

$$n = 1218 / 4,045$$

$$n = 301,1125 \text{ peserta didik baru}$$

Jadi jika dibulatkan menjadi 301 peserta didik baru yang digunakan sebagai bahan sample data.

2. Pengolahan Data

Pada tahap pengolahan data dimaksudkan untuk menyusun suatu kumpulan sample data seperti pada **Gambar 1** dibawah menjadi sebuah kumpulan *data training* seperti pada **Gambar 2** dibawah, sesuai dengan aturan penentuan atribut dominan sehingga memudahkan data untuk dianalisa.

No	Nama	Asal	Gender	Prestasi	Nilai_Mapel	Nilai_Total	Hasil_PPDB
1	Data ke - 1	Dalam Kota	Perempuan	1,5	38,05	39,55	SMAN_1
2	Data ke - 2	Luar Kota	Laki-laki	1,75	37,6	39,35	SMAN_1
3	Data ke - 3	Luar Kota	Laki-laki	1,25	37,9	39,15	SMAN_3
4	Data ke - 4	Luar Kota	Laki-laki	1	37,95	38,95	SMAN_1
5	Data ke - 5	Dalam Kota	Perempuan	1,75	36,95	38,7	SMAN_1
6	Data ke - 6	Dalam Kota	Perempuan	1,75	36,85	38,6	SMAN_1
7	Data ke - 7	Dalam Kota	Perempuan	1	37,5	38,5	SMAN_4
8	Data ke - 8	Dalam Kota	Perempuan	1	37,35	38,35	SMAN_1
9	Data ke - 9	Luar Kota	Perempuan	0	38,35	38,35	SMAN_1
10	Data ke - 10	Luar Kota	Laki-laki	2	36,35	38,35	SMAN_4
11	Data ke - 11	Luar Kota	Laki-laki	1	37,35	38,35	SMAN_4
12	Data ke - 12	Luar Kota	Perempuan	0	38,35	38,35	SMAN_4
13	Data ke - 13	Dalam Kota	Laki-laki	0	38,3	38,3	SMAN_3
14	Data ke - 14	Dalam Kota	Perempuan	1	37,3	38,3	SMAN_4

Gambar 1. Sample Data Hasil Seleksi PPDB

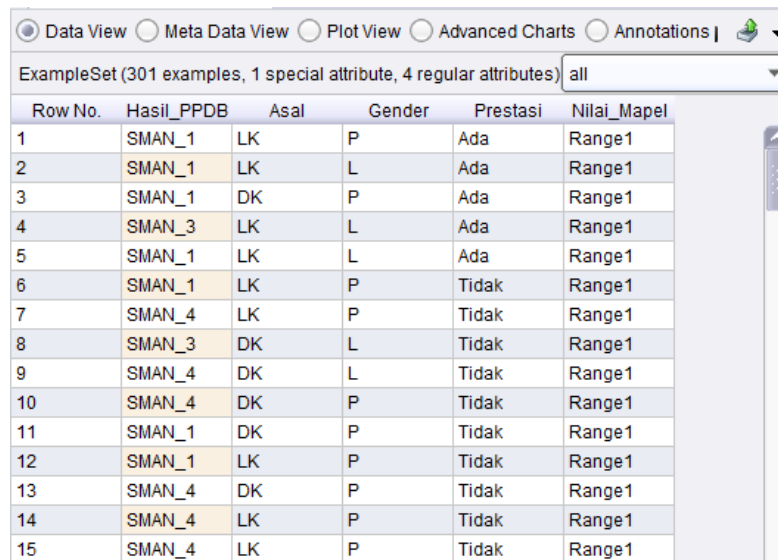
No	Nama	Asal	Gender	Prestasi	Nilai_Mapel	Hasil_PPDB
1	Data ke - 25	LK	P	Ada	Range1	SMAN_1
2	Data ke - 2	LK	L	Ada	Range1	SMAN_1
3	Data ke - 1	DK	P	Ada	Range1	SMAN_1
4	Data ke - 3	LK	L	Ada	Range1	SMAN_3
5	Data ke - 4	LK	L	Ada	Range1	SMAN_1
6	Data ke - 9	LK	P	Tidak	Range1	SMAN_1
7	Data ke - 12	LK	P	Tidak	Range1	SMAN_4
8	Data ke - 13	DK	L	Tidak	Range1	SMAN_3
9	Data ke - 15	DK	L	Tidak	Range1	SMAN_4
10	Data ke - 16	DK	P	Tidak	Range1	SMAN_4
11	Data ke - 18	DK	P	Tidak	Range1	SMAN_1
12	Data ke - 20	LK	P	Tidak	Range1	SMAN_1
13	Data ke - 21	DK	P	Tidak	Range1	SMAN_4
14	Data ke - 22	LK	P	Tidak	Range1	SMAN_4

Gambar 2. Data Training Hasil Seleksi PPDB

3. Implementasi dalam Rapid Miner

Data training hasil seleksi PPDB tingkat SMA Negeri yang telah tersimpan pada media penyimpanan *repository rapid miner* dapat dilihat dalam bentuk tampilan visual *data view* yang disediakan oleh *rapid miner*. *Data view* adalah suatu tampilan visual

yang memaparkan tentang ilustrasi catatan seluruh data dari *data training* yang tersimpan pada media penyimpanan *repository rapid miner*. Tampilan visual *data view* dari *data training* hasil seleksi PPDB tingkat SMA Negeri dapat dilihat pada **Gambar 3.**

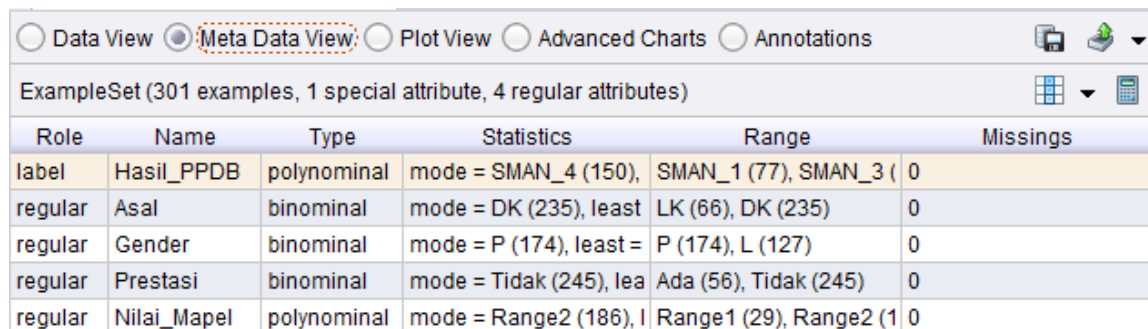


Row No.	Hasil_PPDB	Asal	Gender	Prestasi	Nilai_Mapel
1	SMAN_1	LK	P	Ada	Range1
2	SMAN_1	LK	L	Ada	Range1
3	SMAN_1	DK	P	Ada	Range1
4	SMAN_3	LK	L	Ada	Range1
5	SMAN_1	LK	L	Ada	Range1
6	SMAN_1	LK	P	Tidak	Range1
7	SMAN_4	LK	P	Tidak	Range1
8	SMAN_3	DK	L	Tidak	Range1
9	SMAN_4	DK	L	Tidak	Range1
10	SMAN_4	DK	P	Tidak	Range1
11	SMAN_1	DK	P	Tidak	Range1
12	SMAN_1	LK	P	Tidak	Range1
13	SMAN_4	DK	P	Tidak	Range1
14	SMAN_4	LK	P	Tidak	Range1
15	SMAN_4	LK	P	Tidak	Range1

Gambar 3. Tampilan Visual *Data View* dari *Data Training* Hasil Seleksi PPDB

Hasil penyimpanan dari *data training* hasil seleksi PPDB tingkat SMA Negeri juga dapat dilihat dalam bentuk tampilan visual *meta data view* yang disediakan oleh *rapid miner*. *Meta data view* adalah tampilan visual yang mengilustrasikan ringkasan data

dari seluruh data yang menjadi *data training* yang telah tersimpan pada media penyimpanan *repository rapid miner*. Tampilan visual *meta data view* dari *data training* hasil seleksi PPDB tingkat SMA Negeri dapat dilihat pada **Gambar 4.**



Role	Name	Type	Statistics	Range	Missings
label	Hasil_PPDB	polynomial	mode = SMAN_4 (150),	SMAN_1 (77), SMAN_3 (	0
regular	Asal	binominal	mode = DK (235), least	LK (66), DK (235)	0
regular	Gender	binominal	mode = P (174), least =	P (174), L (127)	0
regular	Prestasi	binominal	mode = Tidak (245), lea	Ada (56), Tidak (245)	0
regular	Nilai_Mapel	polynomial	mode = Range2 (186), l	Range1 (29), Range2 (1	0

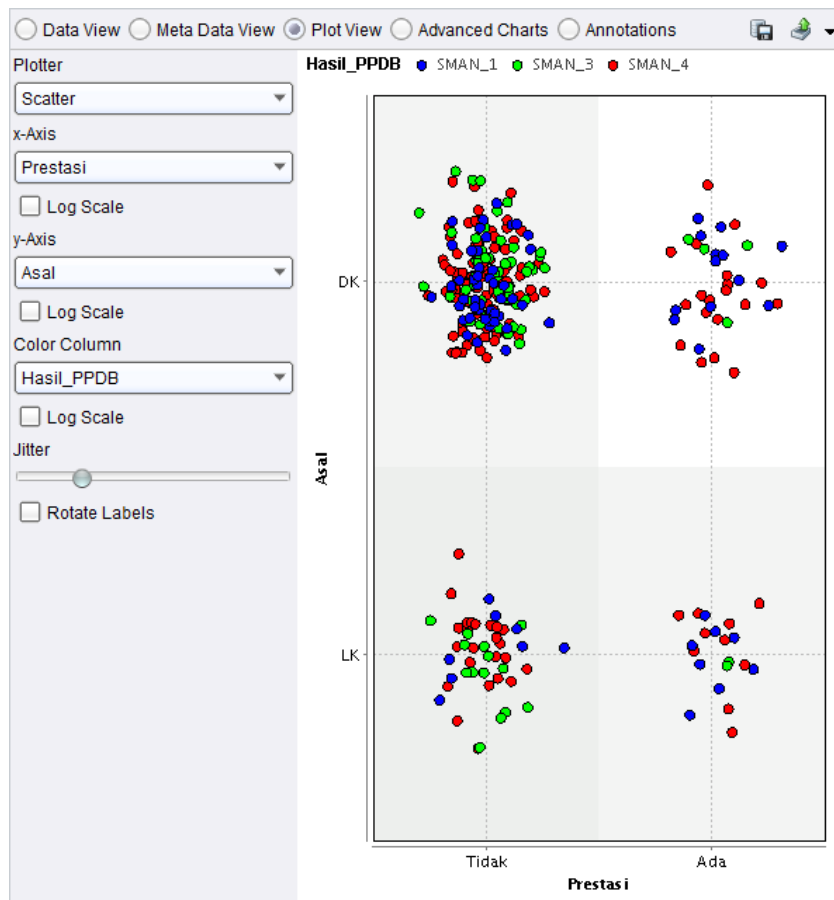
Gambar 4. Tampilan Visual *Meta Data View* dari *Data Training* Hasil Seleksi PPDB

Hasil penyimpanan dari *data training* hasil seleksi PPDB tingkat SMA Negeri juga dapat dilihat dalam bentuk tampilan visual *plot view*. Pada tampilan visual *plot view* disediakan model penggambaran *plotter* data

secara *scatter*, *histogram*, *pie* maupun model penggambaran lainnya untuk dapat dilihat sebaran distribusi data berdasarkan pada setiap atribut yang ada dimiliki oleh *data training* hasil seleksi PPDB tingkat SMA Negeri.

Tampilan visual *plot view* model *scatter* dari *data training* hasil seleksi

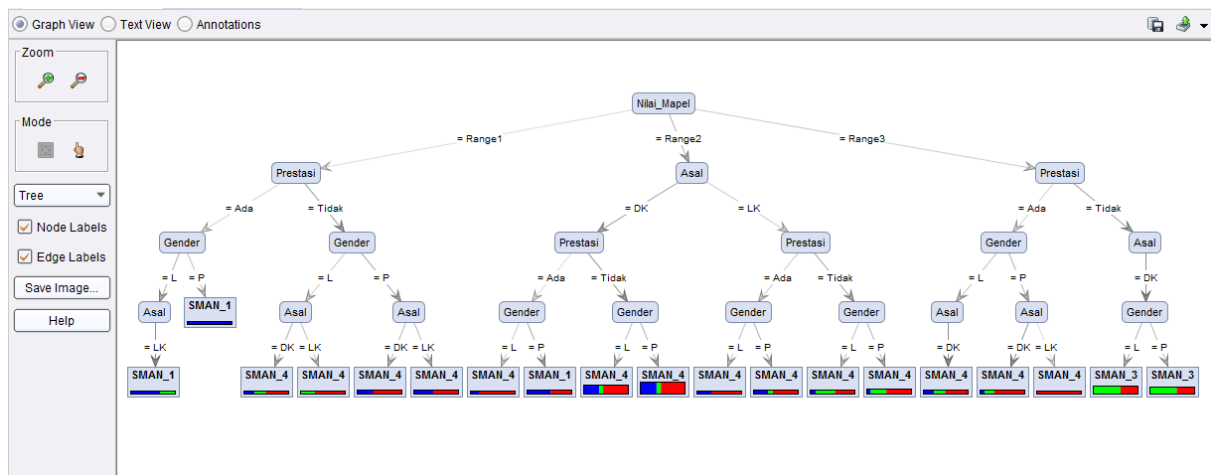
PPDB tingkat SMA Negeri dapat dilihat pada **Gambar 5**.



Gambar 5. Tampilan Visual *Plot View* dari *Data Training* Hasil Seleksi PPDB

Berdasarkan tampilan visual *plot view* model *scatter* pada **Gambar 5** dapat dilihat bahwa asal peserta didik baru yang berasal dari dalam kota baik yang memiliki prestasi maupun tidak memiliki prestasi lebih mendominasi data pada *data training* hasil seleksi PPDB tingkat SMA Negeri dibandingkan dengan asal peserta didik baru yang berasal dari luar kota baik yang memiliki prestasi maupun tidak memiliki prestasi.

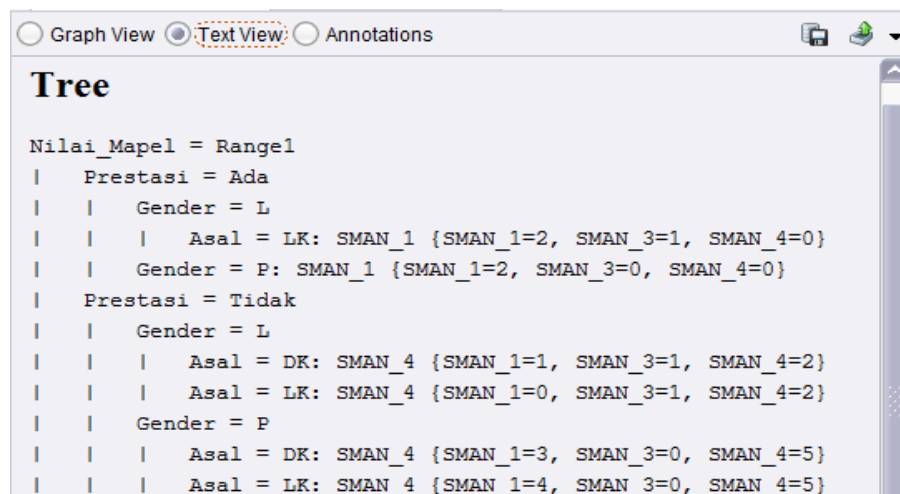
Pohon keputusan (*decision tree*) dari data training hasil seleksi PPDB tingkat SMA Negeri yang telah terinputkan dan telah dianalisa dengan *algoritma ID3* pada *rapid miner* ditampilkan dalam bentuk tampilan visual *graph view*. Tampilan visual *graph view* menampilkan *decision tree* (struktur pohon keputusan) dari data training hasil seleksi PPDB tingkat SMA Negeri dapat dilihat pada **Gambar 6**.



Gambar 6. Graph View Decision Tree dari Data Training Hasil Seleksi PPDB

Pada *rapid miner*, *decision tree* dari data training hasil seleksi PPDB tingkat SMA Negeri juga dapat dilihat dalam bentuk tampilan visual *text view*. Pada tampilan visual *text view*, *decision tree* akan dipecah kedalam

suatu tulisan aturan yang merupakan representasi dari *decision tree* yang ada pada *graph view*. Tampilan visual *text view decision tree* dari data training hasil seleksi PPDB tingkat SMA Negeri bisa dilihat pada **Gambar 7**.



Gambar 7. Text View Decision Tree dari Data Training Hasil Seleksi PPDB

Berdasarkan tampilan visual *graph view* dan *text view* yang telah diperoleh dari *software rapid miner*, *decision tree* hasil seleksi PPDB

tingkat SMA Negeri tersebut apabila ditulis kedalam bentuk suatu tabel aturan, maka diperoleh tabel aturan dari *decision tree* hasil seleksi PPDB

tingkat SMA Negeri yang diterangkan

seperti pada **Tabel 2**.

Tabel 2. Tabel Aturan dari *Decision Tree* Hasil Seleksi PPDB

No	Kondisi 1	Kondisi 2	Kondisi 3	Kondisi 4	Target
1	Range1	Ada	Laki – laki	Luar kota	SMAN 1
2	Range1	Ada	Perempuan	-	SMAN 1
3	Range1	Tidak	Laki – laki	Dalam kota	SMAN 4
4	Range1	Tidak	Laki – laki	Luar kota	SMAN 4
5	Range1	Tidak	Perempuan	Dalam kota	SMAN 4
6	Range1	Tidak	Perempuan	Luar kota	SMAN 4
7	Range2	Dalam kota	Ada	Laki – laki	SMAN 4
8	Range2	Dalam kota	Ada	Perempuan	SMAN 1
9	Range2	Dalam kota	Tidak	Laki – laki	SMAN 4
10	Range2	Dalam kota	Tidak	Perempuan	SMAN 4
11	Range2	Luar kota	Ada	Laki – laki	SMAN 4
12	Range2	Luar kota	Ada	Perempuan	SMAN 4
13	Range2	Luar kota	Tidak	Laki – laki	SMAN 4
14	Range2	Luar kota	Tidak	Perempuan	SMAN 4
15	Range3	Ada	Laki – laki	Dalam kota	SMAN 4
16	Range3	Ada	Perempuan	Dalam kota	SMAN 4
17	Range3	Ada	Perempuan	Luar kota	SMAN 4
18	Range3	Tidak	Dalam kota	Laki – laki	SMAN 3
19	Range3	Tidak	Dalam kota	Perempuan	SMAN 3

Berdasarkan **Tabel 2** apabila salah satu aturan dari tabel aturan tersebut dalam aturan **If - Then** seperti contoh aturan nomor 1 maka aturan tersebut akan ditulis dengan kalimat sebagai berikut “**If** Akumulasi nilai mata pelajaran UAN SMP peserta didik baru berada dikisaran lebih dari 37,5 (Range1) **and** peserta didik baru memiliki prestasi **and** peserta didik baru bergender laki - laki **and** peserta didik baru berasal dari luar kota **Then** memiliki kecenderungan mendaftar pada SMA Negeri 1 Surakarta”.

Dalam melakukan evaluasi distribusi hasil seleksi PPDB tingkat SMA Negeri menggunakan metode *decision tree*, penentuan kelas target dan pembagian objek nilai dari tiap atribut pada *data training* sangat menentukan tingkat akurasi tree yang dibuat. Besar persentase kebenaran dari suatu struktur pohon (*tree*) sangat dipengaruhi oleh seberapa banyak data yang digunakan sebagai bagian dari *data training*.

KESIMPULAN

Berdasarkan penelitian dengan menggunakan data hasil seleksi PPDB tingkat SMA Negeri di Kota Surakarta sebagai *data training* yang kemudian dilakukan proses analisa dalam kaitannya dengan proses analisa pembentukan struktur pohon keputusan (*decision tree*) maka dapat ditarik kesimpulan bahwa :

1. Implementasi *data mining* metode *decision tree* dapat dilakukan, baik dengan analisa perhitungan manual *algoritma ID3* dengan bersumber pada penentuan nilai *entropy* dan *information gain* maupun analisa dengan penerapan *algoritma ID3* yang telah disediakan modulnya pada software *rapid miner*.
2. Implementasi *data mining* metode *decision tree* pada data training hasil seleksi PPDB tingkat SMA Negeri di Kota Surakarta menghasilkan adanya salah satu atribut yang mempunyai peran paling berpengaruh atau dominasi paling tinggi dalam pembentukan pohon keputusan (*decision tree*), atribut tersebut adalah atribut nilai mapel (akumulasi nilai mata pelajaran UAN tingkat SMP)

karena atribut tersebut menjadi *root node* (simpul akar) dari struktur pohon keputusan (*decision tree*) dan atribut tersebut memiliki nilai *information gain* terbesar pada saat penentuan root node.

SARAN

Berdasarkan dari kesimpulan yang telah diperoleh maka dapat diberikan saran sebagai berikut :

1. Data hasil seleksi PPDB sebaiknya dikelola dan disimpan dengan baik karen dapat memberikan manfaat dalam pembentukan *decision tree*.
2. Perlu adanya penelitian yang lebih lanjut dalam rangka mengembangkan pencarian informasi yang lebih banyak dan mendalam mengenai hasil seleksi PPDB berdasarkan atribut lain yang belum dimanfaatkan pada proses *data mining* metode *decision tree algoritma ID3* dalam penelitian ini seperti nilai masing - masing mata pelajaran ujian akhir nasional (UAN) tingkat sekolah menengah pertama (SMP).

DAFTAR PUSTAKA

- Santosa, Budi (2007). *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta : Penerbit Graha Ilmu.
- Kusrini, dan Emha Taufik (2009). *Algoritma Data Mining*. Yogyakarta : Penerbit Andi.
- Sunjana (2010). *Aplikasi Mining Data Mahasiswa Dengan metode Klasifikasi Decision Tree*. Yogyakarta : Seminar Nasional Aplikasi Teknologi Informasi. ISSN : 1907 – 5022
- Seran, Krisantus Jumarto Tey (2013). *Analisis dan Usulan Solusi Sistem Pendukung Keputusan Penentuan Beasiswa Menggunakan Algoritma ID3*. Strata 2 Thesis. Yogyakarta : Universitas Atma Jaya Yogyakarta.
- Dalimunthe, Amirhud (2012). *Implementasi Data Mining Classification dengan Metode Decision Tree (Menggunakan Algoritma ID3 dan C.45)*. Strata 1 Skripsi. Medan : Universitas Negeri Medan.
- Elmande, Widodo (2012). *Pemilihan Criteria Splitting Algoritma Iterative Dichotomiser 3 (ID3) untuk Penentuan Kualitas Beras : Studi Kasus Pada Perum Bulog Divre Kampung*. Lampung : JURNAL TELEMATIKA MKOM Vol. 4 No. 1, Maret 2012
ISSN : 2085 - 725X
- Safii, M (2011). *Penetapan Strategi Pemasaran dengan Pendekatan Model Aturan Pohon Keputusan Menggunakan Algoritma Iterative Dichotomiser 3 (ID3) di Perguruan Tinggi Swasta*. Strata 2 Thesis. Medan : Universitas Sumatera Utara.
- Utami, Winda Pangesti (2012). *Penerapan Algoritma Iterative Dichotomiser Three untuk Pemilihan Dosen Pembimbing (Studi Kasus : FTI UKSW)*. Strata 1 Skripsi. Salatiga : Universitas Kristen Satya Wacana.

BIODATA PENULIS

Nama : Aziz Ridha Utama
Tempat, Tanggal Lahir : Surakarta, 7 November 1991
Jenis Kelamin : Laki - laki
Agama : Islam
Jurusan : Teknik Informatika
Perguruan : Universitas Muhammadiyah Surakarta
Alamat : Jl. A. Yani Tromol Pos 1, Pabelan, Kartasura
Telp / Fax : (0271) 717417, 719483 / (0271) 714448
Alamat Rumah : Sekip Rt 01 Rw 23, Kadipiro, Banjarsari, Surakarta
Kode Pos : 57136
No. HP : +6281802912200
Alamat E - mail : aziz.ridha.utama@gmail.com