

**ANALISIS PENERIMAAN CALON MAHASISWA PMDK
POLITEKNIK NEGERI BATAM MENGGUNAKAN
ALGORITMA *DECISION TREE* C4.5**

TUGAS AKHIR

Oleh :

Novitasari 3311201081

Disusun untuk memenuhi syarat kelulusan Program Diploma III



**PROGRAM STUDI TEKNIK INFORMATIKA
POLITEKNIK NEGERI BATAM**

BATAM

2015

HALAMAN PENGESAHAN

**ANALISIS PENERIMAAN CALON MAHASISWA PMDK
POLITEKNIK NEGERI BATAM MENGGUNAKAN ALGORITMA
*DECISION TREE C4.5***

Oleh :

Novitasari (3311201081)

Tugas Akhir ini telah diterima dan disahkan
sebagai persyaratan untuk memperoleh gelar

Ahli Madya

di

**PROGRAM STUDI DIPLOMA 3 TEKNIK INFORMATIKA
POLITEKNIK NEGERI BATAM**

Batam, 13 Juli 2015

Disetujui oleh;

Pembimbing I,

Hilda Widyastuti, M.T.

NIK. 102020

HALAMAN PENGESAHAN

**ANALISIS PENERIMAAN CALON MAHASISWA PMDK
POLITEKNIK NEGERI BATAM MENGGUNAKAN ALGORITMA
*DECISION TREE C4.5***

Oleh :

Novitasari (3311201081)

Tugas Akhir ini telah diterima dan disahkan
sebagai persyaratan untuk memperoleh gelar

Ahli Madya

di

**PROGRAM STUDI DIPLOMA 3 TEKNIK INFORMATIKA
POLITEKNIK NEGERI BATAM**

Batam, 13 Juli 2015

Disetujui oleh;

Pembimbing II,

Selly Artaty Zega, S.ST.

NIK. 113104

HALAMAN PERNYATAAN

Dengan ini, saya:

NIM : 3311201081

Nama : Novitasari

adalah mahasiswa Teknik Informatika Politeknik Batam yang menyatakan bahwa tugas akhir dengan judul:

Analisis Penerimaan Calon Mahasiswa PMDK Politeknik Negeri Batam Menggunakan
Algoritma Decision Tree C4.5

disusun dengan:

1. tidak melakukan plagiat terhadap naskah karya orang lain
2. tidak melakukan pemalsuan data
3. tidak menggunakan karya orang lain tanpa menyebut sumber asli atau tanpa ijin pemilik

Jika kemudian terbukti terjadi pelanggaran terhadap pernyataan di atas, maka saya bersedia menerima sanksi apapun termasuk pencabutan gelar akademik.

Lembar pernyataan ini juga memberikan hak kepada Politeknik Batam untuk mempergunakan, mendistribusikan ataupun memproduksi ulang seluruh hasil Tugas Akhir ini.

Batam, 13 Juli 2015

Novitasari
3311201081

KATA PENGANTAR

Dengan rahmat Tuhan Yang Maha Esa sehingga penulis dapat menyelesaikan Laporan Tugas Akhir yang berjudul “Analisis Penerimaan Calon Mahasiswa PMDK Politeknik Negeri Batam Menggunakan Algoritma *Decision Tree* C4.5 ” ini dengan tepat pada waktunya. Laporan Tugas Akhir ini dibuat untuk memenuhi persyaratan memperoleh gelar Ahli Madya di Program Studi Diploma 3 Teknik Informatika Politeknik Negeri Batam dan memenuhi kriteria penilaian dari mata kuliah wajib yang harus diselesaikan dan dipertanggungjawabkan sebagaimana mestinya.

Dalam kesempatan ini penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada pihak-pihak terkait, terkhusus kepada :

1. Tuhan Yang Maha Esa yang telah memberikan kemudahan dalam mengerjakan Tugas Akhir ini.
2. Kedua orang tua kami tercinta yang telah memberikan dukungan moril maupun materil selama ini.
3. Ibu Meyti Eka Apriyani, MT, selaku Kaprodi Teknik Informatika.
4. Ibu Hilda Widyastuti, M.T, selaku pembimbing I pada Tugas Akhir ini.
5. Ibu Selly Artaty Zega, selaku pembimbing II pada Tugas Akhir ini.
6. Teman-teman sekalian yang tidak mungkin kami sebut satu persatu namanya yang telah memberikan semangat, dorongan serta dukungan dalam pembuatan Tugas Akhir ini.

Banyak pelajaran yang dapat dipetik selama proses pembuatan Tugas Akhir ini, salah satunya ialah menuntut adanya tanggungjawab yang sangat besar dan kesabaran. Tak ada gading yang tak retak, demikian juga Tugas Akhir ini tidak luput dari sedikit banyaknya ketidaksesuaian baik dari kata-kata maupun strukturnya.

Besar harapan penulis adanya masukan ide-ide baru baik berupa kritik dan saran dimana bertujuan demi penyempurnaan Tugas Akhir ini.

Akhir kata dari segenap penulis mengucapkan terima kasih yang sebesar-besarnya, semoga dengan adanya Tugas Akhir yang luar biasa ini dapat menambah dan memperkaya sedikit ilmu dan wawasan, terutama dalam bidang ilmu data mining beserta pemrograman.

Batam, 13 Juli 2015

Penulis

ABSTRAK

ANALISIS PENERIMAAN CALON MAHASISWA PMDK POLITEKNIK NEGERI BATAM MENGGUNAKAN ALGORITMA *DECISION TREE C4.5*

PMDK adalah sistem penerimaan mahasiswa baru yang diselenggarakan oleh suatu Universitas maupun Perguruan Tinggi secara mandiri. Sesuai dengan namanya PMDK merupakan penelusuran minat calon mahasiswa, bakatnya, juga kemampuannya. Jalur PMDK tidak perlu mengikuti tes ujian masuk, hanya dinilai dari beberapa nilai rapot atau beberapa mata pelajaran yang dijadikan syarat untuk mengikuti PMDK.

Dari data yang ada berdasarkan studi kasus di Perguruan Tinggi Politeknik Negeri Batam, dilakukan suatu penerapan untuk menemukan hasil dalam menentukan calon mahasiswa PMDK. Penerapan dilakukan dengan menggunakan teknik *data mining* yang mengekstraksi data untuk memperoleh informasi, serta menggunakan metode *classification*, yang mengklasifikasi data hingga menemukan suatu *rule* yang dapat mengklasifikasi berupa data baru yaitu penerimaan calon mahasiswa PMDK.

Rule yang diperoleh akan diimplementasikan dalam aplikasi yang dapat mengklasifikasikan peserta PMDK diterima atau tidak diterima berdasarkan data yang dimasukkan.

Kata Kunci : *data mining*, *classification*, *rule*, dan PMDK → Penelusuran Minat dan Bakat

ABSTRACT

ANALYSIS OF ACCEPTANCE OF CANDIDATES PMDK Batam Polytechnic USING DECISION TREE ALGORITHM C4.5

PMDK is a new admissions system organized by a University or College independently. As the name implies PMDK is tracking student interest, talent, also his ability. PMDK not need to take the test entrance examination, only assessed from some value rapot or several subjects were used as a requirement to follow PMDK. From the existing data based on case studies in Batam Polytechnic University, performed an application to find results in determining prospective students PMDK. The application is done by using data mining techniques to extract the data to obtain information, as well as using methods of classification, which classifies the data to find a rule to classify the form of new data that admissions PMDK. Rule obtained will be implemented in an application that can classify participants PMDK accepted or not accepted based on the data entered.

Keywords: data mining, classificasion, rule, and PMDK → Interest Search and Talent

DAFTAR ISI

BAB I PENDAHULUAN	1
I.1 Latar Belakang	1
I.2 Rumusan Masalah	2
I.3 Batasan Masalah	2
I.4 Tujuan Penelitian	2
I.5 Sistematika Penulisan	3
BAB II LANDASAN TEORI.....	5
II.1 Pengertian Analisis	5
II.2 Pengertian PMDK.....	5
II.3 Penelitian Sebelumnya	6
II.4 Jalur Masuk Politeknik Negeri Batam	9
II.5 Data Operasional.....	9
II.6 Data Mining	10
II.6.1 Metode Klasifikasi	12
II.6.2 Teknik Decision Tree dalam Metode Klasifikasi	14
II.7 Perangkat-Perangkat Lunak yang Mendukung dalam Analisis.....	16
II.7.1 Microsoft Excel.....	16
II.7.2 Microsoft Visio	17
II.7.3 Netbeans.....	17
II.7.4 Pengertian Java	17
II.7.5 Basis Data (<i>Database</i>)	18
II.7.6 XAMPP	19
II.7.7 UML.....	19
II.8 Distribusi Frekuensi dan Interval Kelas	20
BAB III ANALISIS DAN PERANCANGAN.....	23
III.1 Pengumpulan Data	23
III.2 Pemilihan Atribut	24
III.3 Tahapan Preprocessing Data.....	24
III.3.1 Transformasi Data.....	24
III.3.2 Data Cleaning.....	25
III.3.3 Data <i>Dizcretization</i>	25
III.4 Pemilihan Data Training dan Data Testing	27

III.5 Deskripsi Sistem.....	28
III.6 Diagram ER.....	28
III.7 Spesifikasi Database pmdk_polibatam	29
III.8 Lingkungan Operasional	30
III.9 Spesifikasi Kebutuhan Fungsional dan Non Fungsional.....	30
III.9.1 Spesifikasi Kebutuhan Fungsional	30
III.9.2 Spesifikasi Kebutuhan Fungsional	30
III.10 Use Case.....	30
III.10.1 Diagram <i>Use Case</i>	31
III.11 Sequence Diagram	32
III.11.1 Sequence Diagram Menyeleksi Peserta PMDK	32
III.11.2 Sequence Diagram Mencari Informasi Peserta	33
III.12 Class Diagram	34
III.13 Perancangan Antarmuka	35
III.13.1 Perancangan Antarmuka Halaman Utama.....	36
III.13.2 Perancangan Antarmuka Penyeleksian PMDK	37
III.13.3 Perancangan Antarmuka Informasi Peserta PMDK	38
III.13.4 Perancangan Antarmuka Tentang Pembuat Aplikasi	39
III.13.5 Perancangan Antarmuka Petunjuk.....	39
BAB IV PROSES KLASIFIKASI	40
IV.1 Perhitungan Entropy dan Gain Data Training.....	40
IV.1.1 Penentuan <i>Root Node</i>	40
IV.1.2 Penentuan <i>Node 1.1</i>	41
IV.1.3 Penentuan <i>Node 1.2</i>	41
IV.1.4 Penentuan <i>Node 1.3</i>	42
IV.1.5 Penentuan <i>Node 1.3.1</i>	43
IV.1.6 Penentuan <i>Node 1.3.1.1</i>	44
IV.1.7 Penentuan <i>Node 1.3.1.1.1</i>	45
IV.1.8 Penentuan <i>Node 1.3.1.1.1.1</i>	46
IV.1.9 Penentuan <i>Node 1.3.1.2</i>	47
IV.1.10 Penentuan <i>Node 1.3.2</i>	48
IV.1.11 Penentuan <i>Node 1.3.3</i>	49
IV.1.12 Penentuan <i>Node 1.3.3.1</i>	50
IV.1.13 Penentuan <i>Node 1.3.4</i>	51

IV.1.14 Penentuan <i>Node</i> 1.3.4.1	52
IV.1.15 Penentuan <i>Node</i> 1.3.5	53
IV.1.16 Penentuan <i>Node</i> 1.3.5.1	54
IV.1.17 Penentuan <i>Node</i> 1.3.5.1.1	55
IV.1.18 Penentuan <i>Node</i> 1.3.5.1.1.1	56
IV.1.19 Penentuan <i>Node</i> 1.3.5.1.1.1.1	57
IV.1.20 Penentuan <i>Node</i> 1.3.5.1.1.1.2	58
IV.1.21 Penentuan <i>Node</i> 1.3.6	59
IV.1.22 Penentuan <i>Node</i> 1.4	60
IV.1.23 Penentuan <i>Node</i> 1.4.1	61
IV.1.24 Penentuan <i>Node</i> 1.4.2	62
IV.1.25 Penentuan <i>Node</i> 1.4.2.1	63
IV.1.26 Penentuan <i>Node</i> 1.4.2.2	64
IV.1.27 Penentuan <i>Node</i> 1.4.2.2.1	65
IV.1.28 Penentuan <i>Node</i> 1.4.2.2.1.1	66
IV.1.29 Penentuan <i>Node</i> 1.4.2.2.2	67
IV.1.30 Penentuan <i>Node</i> 1.4.3	68
IV.1.31 Penentuan <i>Node</i> 1.4.3.1	69
IV.1.32 Penentuan <i>Node</i> 1.4.3.2	70
IV.1.33 Penentuan <i>Node</i> 1.4.4	71
IV.1.34 Penentuan <i>Node</i> 1.4.4.1	72
IV.1.35 Penentuan <i>Node</i> 1.4.4.1.1	73
IV.1.36 Penentuan <i>Node</i> 1.4.4.2	74
IV.1.37 Penentuan <i>Node</i> 1.4.5	75
IV.1.38 Penentuan <i>Node</i> 1.4.6	76
IV.1.39 Penentuan <i>Node</i> 1.4.7	77
IV.1.40 Penentuan <i>Node</i> 1.4.7.1	78
IV.1.41 Penentuan <i>Node</i> 1.4.8	79
IV.1.42 Penentuan <i>Node</i> 1.4.9	80
IV.1.43 Penentuan <i>Node</i> 1.4.10	81
IV.1.44 Penentuan <i>Node</i> 1.5	82
IV.1.45 Penentuan <i>Node</i> 1.5.1	83
IV.1.46 Penentuan <i>Node</i> 1.5.2	84
IV.1.47 Penentuan <i>Node</i> 1.5.3	84

IV.1.48 Penentuan <i>Node</i> 1.5.3.1	85
IV.1.49 Penentuan <i>Node</i> 1.5.3.1.1	86
IV.1.50 Penentuan <i>Node</i> 1.5.3.2	87
IV.1.51 Penentuan <i>Node</i> 1.5.4	88
IV.1.52 Penentuan <i>Node</i> 1.5.4.1	89
IV.1.53 Penentuan <i>Node</i> 1.5.5	90
IV.1.54 Penentuan <i>Node</i> 1.5.5.1	91
IV.1.55 Penentuan <i>Node</i> 1.5.6	92
IV.1.56 Penentuan <i>Node</i> 1.5.7	93
IV.1.57 Penentuan <i>Node</i> 1.6	94
IV.1.58 Penentuan <i>Node</i> 1.6.1	94
IV.1.59 Penentuan <i>Node</i> 1.6.1.1	95
IV.1.60 Penentuan <i>Node</i> 1.6.1.1.1	96
IV.1.61 Penentuan <i>Node</i> 1.6.1.1.1.1	97
IV.1.62 Penentuan <i>Node</i> 1.6.2	98
IV.1.63 Penentuan <i>Node</i> 1.6.2.1	99
IV.1.64 Penentuan <i>Node</i> 1.7	100
IV.1.65 Penentuan <i>Node</i> 1.7.1	101
IV.1.66 Penentuan <i>Node</i> 1.7.1.1	102
IV.1.67 Penentuan <i>Node</i> 1.7.1.1.1	102
IV.1.68 Penentuan <i>Node</i> 1.7.1.2	103
IV.1.69 Penentuan <i>Node</i> 1.7.1.3	104
IV.2 Tingkat Error Rate Data	105
BAB V IMPLEMENTASI DAN PENGUJIAN	106
V.1 Algoritma	106
V.1.1 Algoritma Menyeleksi Peserta PMDK	106
V.1.2 Algoritma Mencari Informasi Peserta PMDK	108
V.2 Implementasi Perancangan AntarMuka	108
V.2.1 Hasil Implementasi Perancangan Halaman Utama	108
V.2.2 Hasil Implementasi Perancangan Penyeleksian PMDK	110
V.2.3 Hasil Implementasi Perancangan Informasi Peserta PMDK	111
V.2.4 Hasil Implementasi Perancangan Tentang Pembuat Aplikasi ...	112
V.2.5 Hasil Implementasi Perancangan Tentang Pembuat Aplikasi ...	113
V.3 Pengujian	113

V.3.1	Strategi Pengujian	113
V.3.2	Skenario Pengujian	113
V.3.3	Hasil Pengujian	114
BAB VI	KESIMPULAN DAN SARAN.....	115
VI.1	Kesimpulan	115
VI.2	Saran.....	115
DAFTAR PUSTAKA.....		xviii
LAMPIRAN		129
Lampiran A	(Data Mentah)	A-1
Lampiran B	(Hasil Cleaning)	B-1
Lampiran C	(Data Training)	C-1
Lampiran D	(Data Testing)	D-1
Lampiran E	(Entropy dan GAIN)	E-1
Lampiran F	(Data Error Rate)	F-1
Lampiran G	(Pembentukan Model Aturan Klasifikasi)	G-1

DAFTAR GAMBAR

Gambar 1. Langkah <i>Learning</i>	13
Gambar 2. Langkah <i>Classification</i>	13
Gambar 3. Flowchart Proses Analisis dan Klasifikasi	23
Gambar 4. Deskripsi Sistem	28
Gambar 5. Diagram ER	29
Gambar 6. Diagram <i>Use Case</i>	31
Gambar 7. Sequence Menyeleksi Peserta PMDK	33
Gambar 8. Sequence Diagram Mencari Informasi Peserta.....	34
Gambar 9. Class Diagram.....	35
Gambar 10. Perancangan Antarmuka Halaman Utama.....	36
Gambar 11. Perancangan Antarmuka Penyeleksian PMDK	37
Gambar 12. Perancangan Antarmuka Informasi Peserta PMDK	38
Gambar 13. Perancangan Antarmuka Tentang Pembuat Aplikasi	39
Gambar 14. Perancangan Antarmuka Petunjuk.....	39
Gambar 15. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.....	40
Gambar 16. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.1.....	41
Gambar 17. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.2.....	42
Gambar 18. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.....	43
Gambar 19. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.1.....	44
Gambar 20. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.1.1.....	45
Gambar 21. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.1.1.1.....	46
Gambar 22. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.1.1.1.1.....	47
Gambar 23. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.1.2.....	48
Gambar 24. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.2.....	49
Gambar 25. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.3.....	50
Gambar 26. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.3.1.....	51
Gambar 27. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.4.....	52
Gambar 28. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.4.1.....	53
Gambar 29. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.5.....	54
Gambar 30. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.5.1.....	55
Gambar 31. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.5.1.1.....	56
Gambar 32. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.5.1.1.1.....	57

Gambar 33. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.5.1.1.1.1.....	58
Gambar 34. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.5.1.1.1.2.....	59
Gambar 35. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.3.6.....	60
Gambar 36. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.....	61
Gambar 37. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.1.....	62
Gambar 38. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.2.....	63
Gambar 39. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.2.1.....	64
Gambar 40. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.2.2.....	65
Gambar 41. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.2.2.1.....	66
Gambar 42. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.2.2.1.1.....	67
Gambar 43. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.2.2.2.....	68
Gambar 44. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.2.2.2.....	69
Gambar 45. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.3.1.....	70
Gambar 46. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.3.2.....	71
Gambar 47. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.4.....	72
Gambar 48. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.4.1.....	73
Gambar 49. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.4.1.1.....	74
Gambar 50. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.4.2.....	75
Gambar 51. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.5.....	76
Gambar 52. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.6.....	77
Gambar 53. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.7.....	78
Gambar 54. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.7.1.....	79
Gambar 55. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.8.....	80
Gambar 56. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.9.....	81
Gambar 57. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.4.10.....	82
Gambar 58. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.....	83
Gambar 59. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.1.....	83
Gambar 60. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.2.....	84
Gambar 61. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.3.....	85
Gambar 62. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.3.1.....	86
Gambar 63. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.3.1.1.....	87
Gambar 64. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.3.2.....	88
Gambar 65. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.4.....	89
Gambar 66. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.4.1.....	90

Gambar 67. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.5.....	91
Gambar 68. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.5.1.....	92
Gambar 69. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.6.....	93
Gambar 70. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.5.7.....	93
Gambar 71. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.6.....	94
Gambar 72. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.6.1.....	95
Gambar 73. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.6.1.1.....	96
Gambar 74. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.6.1.1.1.....	97
Gambar 75. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.6.1.1.1.1.....	98
Gambar 76. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.6.2.....	99
Gambar 77. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.6.2.1.....	100
Gambar 78. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.7.....	101
Gambar 79. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.7.1.....	101
Gambar 80. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.7.1.1.....	102
Gambar 81. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.7.1.1.1.....	103
Gambar 82. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.7.1.2.....	104
Gambar 83. <i>Decision Tree</i> Hasil Perhitungan <i>Node</i> 1.7.1.3.....	104
Gambar 84. Hasil Implementasi Perancangan Halaman Utama.....	108
Gambar 85. Hasil Implementasi Halaman Utama Submenu Menu.....	109
Gambar 86. Hasil Implementasi Halaman Utama Submenu Tentang.....	110
Gambar 87. Hasil Implementasi Perancangan Penyeleksian PMDK	110
Gambar 88. Hasil Implementasi Perancangan Informasi Peserta PMDK	111
Gambar 89. Hasil Implementasi Perancangan Tentang Pembuat Aplikasi	112
Gambar 90. Perancangan Tentang Pembuat Aplikasi	113
Gambar 91. Skenario Pengujian	113

DAFTAR TABEL

Tabel 1. Perbandingan dengan Penelitian Sebelumnya.....	6
Table 2. Contoh Kasus Data Pengujian	15
Tabel 3. Data Nilai Bahasa Indonesia Mahasiswa Universitas Y	21
Tabel 4. Data Nilai Bahasa Indonesia Mahasiswa Universitas Y dengan Range..	22
Tabel 5. Tranformasi Data.....	25
Tabel 6. Range Rata-rata Sem 1-5.....	27
Tabel 9. Spesifikasi Tabel PMDK.....	29
Tabel 7. Kebutuhan Fungsional.....	30
Tabel 8. Kebutuhan Non Fungsional.....	30
Tabel 10. Hasil Pengujian Sistem.....	114

BAB I

PENDAHULUAN

I.1 Latar Belakang

PMDK (Penelusuran Minat dan Bakat) adalah sistem penerimaan mahasiswa baru yang diselenggarakan oleh suatu Universitas maupun Perguruan Tinggi secara mandiri. Sesuai dengan namanya PMDK merupakan penelusuran minat calon mahasiswa, bakatnya, juga kemampuannya. Jalur PMDK tidak perlu mengikuti tes ujian masuk, hanya dinilai dari beberapa nilai rapot atau beberapa mata pelajaran yang dijadikan syarat untuk mengikuti PMDK.

Forum Direktur Politeknik Negeri se-Indonesia (FDPNI) pada tahun 2014 menetapkan pola Penerimaan Mahasiswa Baru (PMB) jalur Penelusuran Minat dan Kemampuan Politeknik Negeri (PMDK-PN) akan dilakukan bersama dan diikuti seluruh Politeknik Negeri se-Indonesia (sebanyak 38 Politeknik Negeri) secara *online*. Hal ini dilaksanakan dengan berdasarkan UU No. 12 Tahun 2012, PP No. 66 dan PP No. 34 Tahun 2010 (www.polibatam.ac.id).

Di Kampus Politeknik Negeri Batam, banyak calon mahasiswa meminati jalur PMDK tersebut dikarenakan calon mahasiswa yang mendaftar melalui jalur ini tidak perlu mengikuti ujian tes masuk, hanya dinilai dari beberapa mata pelajaran yang dijadikan syarat untuk mengikuti PMDK. Namun juga tidak semua calon mahasiswa bisa lulus dalam seleksi ini, karena sistem PMDK tersebut memiliki ketentuan nilai untuk bisa lulus. Semakin banyak calon mahasiswa yang mengikuti jalur PMDK tersebut dapat menyulitkan pihak universitas dalam menyeleksi calon mahasiswa yang mendaftar melalui jalur PMDK. Proses seleksi yang dilakukan pun masih secara manual sehingga membutuhkan waktu yang cukup lama untuk menyeleksi data calon mahasiswa.

Dari data yang ada, dapat dilakukan pengkajian untuk menyeleksi calon mahasiswa PMDK. Namun belum ditemukan data mana yang paling tepat untuk menyeleksi calon mahasiswa PMDK tersebut. Untuk itu, analisis terhadap data-data calon mahasiswa perlu dilakukan untuk menemukan susunan atau pola yang dapat digunakan dalam mengklasifikasikan data calon mahasiswa PMDK.

Berdasarkan studi kasus di Perguruan Tinggi Politeknik Negeri Batam, dilakukan suatu penerapan untuk menemukan hasil dalam menentukan calon mahasiswa PMDK. Penerapan dilakukan dengan menggunakan teknik *data mining* yang mengekstraksi data untuk memperoleh informasi, serta menggunakan metode *classification*, yang mengklasifikasi data hingga menemukan suatu *rule* yang dapat mengklasifikasi berupa data baru yaitu penerimaan calon mahasiswa PMDK.

Pada Tugas Akhir ini, akan dibuat “Analisis Penerimaan Calon Mahasiswa PMDK Politeknik Negeri Batam Menggunakan Algoritma *Decision Tree C4.5*” sehingga menghasilkan rule yang akan dipakai dalam mengklasifikasikan data baru yang belum diketahui kelas labelnya (penerimaan calon mahasiswa PMDK). Rule tersebut diimplementasikan ke dalam suatu aplikasi yang bertujuan untuk membuat suatu *interface* dalam mengklasifikasikan penerimaan peserta PMDK baru.

I.2 Rumusan Masalah

Adapun beberapa rumusan masalah dalam penulisan laporan Tugas Akhir adalah sebagai berikut :

1. Apa saja atribut-atribut yang digunakan dalam menentukan penerimaan calon mahasiswa PMDK.
2. Bagaimana langkah-langkah menganalisis penerimaan calon mahasiswa PMDK
3. Proses analisis akan menghasilkan hasil *rule* klasifikasi, bagaimana mengimplementasikannya kedalam suatu *interface* aplikasi.

I.3 Batasan Masalah

Batasan masalah yang terdapat dalam penulisan laporan Tugas Akhir sebagai berikut :

1. Data mahasiswa yang dianalisis adalah data mahasiswa angkatan 2012-2013 Politeknik Negeri Batam.
2. Tidak menangani data PMDK UN dalam menganalisis penerimaan calon mahasiswa.

I.4 Tujuan Penelitian

Tugas akhir ini ditulis dengan tujuan sebagai berikut :

1. Menentukan atribut-atribut yang akan digunakan sebelum menganalisis data penerimaan calon mahasiswa.

2. Menganalisis data calon mahasiswa PMDK dengan menggunakan metode *decision tree* C4.5 sehingga menghasilkan hasil *rule* klasifikasi yang menentukan calon mahasiswa PMDK tersebut diterima atau tidak.
3. Mengimplementasikan hasil *rule* klasifikasi ke dalam *interface* aplikasi penerimaan calon mahasiswa PMDK.

I.5 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam Tugas Akhir ini sebagai berikut:

BAB I

Bab ini menjelaskan latar belakang, rumusan masalah, batasan masalah, tujuan, dan sistematika penulisan dalam penerapan metode klasifikasi peserta PMDK.

Bab II

Bab ini berisi teori-teori yang berhubungan dengan penerimaan peserta PMDK dan ulasan penelitian-penelitian yang pernah dilakukan.

Bab III

Bab ini memuat uraian langkah-langkah dalam menganalisis peserta PMDK

Bab IV

Bab ini berisi proses perhitungan *entropy* dan *gain* terhadap data *training* untuk menentukan *node* suatu *decision tree*. Perhitungan *gain* dan *entropy* dilakukan dengan menggunakan *Microsoft Excel* yang telah mengandung rumus perhitungan *gain* dan *entropy*.

Bab V

Proses analisis telah dilewati dan menghasilkan *rule* yang akan dipakai dalam mengklasifikasikan data baru yang belum diketahui kelas labelnya (penerimaan calon mahasiswa PMDK). *Rule* tersebut diimplementasikan ke dalam suatu aplikasi yang bertujuan untuk membuat suatu *interface* dalam mengklasifikasikan penerimaan peserta PMDK suatu data baru.

Bab VI

Kesimpulan dan saran yang berisi tentang kesimpulan dari hasil klasifikasi hingga diimplementasi kedalam aplikasi yang dibuat pada Tugas Akhir serta saran pengembangan aplikasi mengenai penyempurnaan ide yang dapat dilakukan terhadap aplikasi yang dibuat.

BAB II

LANDASAN TEORI

Berikut adalah berbagai sumber tinjauan, metode yang akan digunakan, dan juga perangkat aplikasi apa saja yang akan digunakan dalam membangun tugas akhir ini.

II.1 Pengertian Analisis

Menurut (Jogiyanto, 2001) analisis berasal dari bahasa Yunani yaitu “analisis”. Analisis dalam bahasa Latin berasal dari kata “ana” dengan arti kembali dan “luein” dengan arti melepas. Berdasarkan asal kata tersebut maka analisis adalah suatu proses yang menguraikan atau memisahkan informasi yang utuh ke dalam bagian-bagian komponen dengan maksud untuk mengidentifikasi dan mengevaluasi permasalahan-permasalahan yang ada serta hasil yang diperoleh dapat dikaji lebih lanjut.

Dalam menganalisis ada langkah-langkah dasar yang harus dilakukan supaya memudahkan untuk melakukan koordinasi dan pengawasan. Adapun langkah-langkah menganalisis sebagai berikut:

1. *Identify*, yaitu mengidentifikasi masalah.
2. *Understand*, yaitu memahami kerja yang ada.
3. *Collecting*, yaitu mengumpulkan data yang diperlukan.
4. *Analyze*, yaitu menganalisis.
5. *Report*, yaitu membuat laporan hasil analisis.

II.2 Pengertian PMDK

PMDK adalah sistem penerimaan mahasiswa baru yang diselenggarakan oleh suatu Universitas maupun Perguruan Tinggi secara mandiri. Dalam sistem ini setiap peserta hanya dapat memilih program studi yang ada di Universitas atau Perguruan Tinggi tersebut. Sesuai dengan namanya PMDK merupakan penelusuran minat calon mahasiswa, bakatnya, juga kemampuannya. Pendaftaran calon peserta atau mahasiswa melalui jalur PMDK tersebut tidak dikenakan biaya apapun karena semua biaya telah dibebankan pada anggaran pemerintah. Penerimaan Mahasiswa Baru melalui Jalur Penelusuran Minat dan Kemampuan (PMDK) ini diperuntukkan bagi calon mahasiswa dari lulusan SLTA yang memiliki prestasi akademik (termasuk nilai UN murni yang tinggi) maupun non-akademik. Peserta yang dinyatakan mendapat beasiswa PMDK akan dibebaskan biaya pendidikan semester 1.

Semakin tinggi tingkat prestasi siswa yang ditunjukkan dari nilai raport atau nilai ujian nasional murni atau penghargaan non akademik yang diperoleh, maka semakin besar kesempatan untuk memperoleh beasiswa PMDK.

Calon mahasiswa yang lulus pada program PMDK mendapat kesempatan untuk memperoleh beasiswa yaitu beasiswa PMDK dengan mengajukan pendaftaran beasiswa terlebih dahulu. Kesempatan mendapatkan PMDK diperuntukkan bagi peserta PMDK yang memiliki prestasi akademik dan non akademik atau nilai ujian nasional murninya tertinggi. Peserta yang dinyatakan mendapat beasiswa PMDK berupa pembebasan biaya hidup selama 1 semester (www.polibatam.ac.id).

II.3 Penelitian Sebelumnya

Pada bagian ini peneliti akan membahas perbedaan penelitian dengan empat penelitian sebelumnya terkait Tugas Akhir ini dapat dilihat seperti tabel 1.

Tabel 1. Perbandingan dengan Penelitian Sebelumnya

Perbandingan	Penelitian 1	Penelitian 2	Penelitian 3	Penelitian 4	Penelitian Sekarang
Objek yang diteliti	Analisis kualitas mahasiswa	Memprediksi kinerja akademik mahasiswa	Memprediksi kinerja mahasiswa	Membandingkan tingkat akurasi dari masing-masing algoritma	Analisis penerimaan calon mahasiswa PMDK
Studi kasus	Politeknik Negeri Batam	-	-	-	Politeknik Negeri Batam
Algoritma	<i>Decision tree C4.5</i>	ID3 dan C4.5	ID3, C4.5, dan REPTree	<i>logistic regression, decision tree c4.5, naïve bayes dan neural network</i>	<i>Decision tree C4.5</i>
<i>Software pengembang</i>	<i>Visual basic</i>	-	-	-	Java netbeans

Penjelasan lebih lanjut mengenai penelitian analisis data sebelumnya yang terdapat pada tabel 1 yaitu:

1. Analisis Kualitas Mahasiswa Politeknik Negeri Batam Program Studi Teknik Informatika Berdasarkan Jalur Masuk Kuliah” (Zega, 2011)

Pada penelitian ini dilakukan analisis dengan menggunakan teknik *data mining* yang mengekstraksi data untuk memperoleh informasi, serta menggunakan metode *classification*, yang mengklasifikasikan data hingga menemukan suatu *rule* yang dapat mengklasifikasikan data baru yang belum diketahui kategori kualitasnya, yaitu tingkat kualitas mahasiswa Politeknik Negeri Batam Program Studi Teknik Informatika. Selain itu, hasil analisis akan memberikan informasi tentang pengaruh jalur masuk kuliah dalam menentukan tingkat kualitas mahasiswa. *Rule* yang diperoleh akan diimplementasikan dalam aplikasi yang dapat mengklasifikasikan tingkat kualitas suatu mahasiswa berdasarkan data yang dimasukkan.

Kualitas mahasiswa menjadi tolak ukur dalam menghasilkan sumber manusia yang kompeten. Kualitas itu melekat pada setiap mahasiswa sesuai dengan apa yang sudah mahasiswa gali selama proses perkuliahan dijalani. Kualitas tersebut dapat dilihat dari Indeks Prestasi Kumulatif (IPK), kemampuan untuk menyelesaikan perkuliahan tepat waktu, Surat Peringatan (SP) dan faktor lainnya. Semua data tersebut disimpan ke dalam basis data mahasiswa, hingga mahasiswa masuk di kampus Politeknik Negeri Batam hingga dinyatakan lulus. Data yang banyak itu menjadi dokumentasi yang jika dibutuhkan, dapat diambil kembali.

Dari data yang ada, dapat dilakukan pengkajian untuk menentukan kualitas mahasiswa. Namun, belum ditemukan data mana yang paling tepat untuk menilai kualitas tersebut. Untuk itu, analisis terhadap data-data mahasiswa perlu dilakukan untuk menemukan *pattern* atau pola yang dapat digunakan dalam mengklasifikasikan data mahasiswa berdasarkan tingkat kualitasnya. Selain itu, perlu ditemukan persentase tingkat kualitas mahasiswa untuk masing-masing jalur masuk kuliah. Politeknik Negeri Batam memiliki dua jalur masuk kuliah, yaitu melalui Ujian Masuk Politeknik Negeri Batam (UMPB) dan melalui Penelusuran Minat dan Keterampilan (PMDK).

Dari analisis tersebut, diharapkan diperoleh informasi mengenai pola yang dipakai dalam menentukan tingkat kualitas mahasiswa Politeknik Negeri Batam Program Studi Teknik Informatika, seberapa besar pengaruh jalur masuk kuliah dalam menentukan tingkat

kualitas mahasiswa, serta persentase tingkat kualitas mahasiswa untuk kedua jalur masuk kuliah tersebut.

2. “Predicting Students Performance Using ID3 and C4.5 Classification Algorithms (Adhartrao, 2013)”.

Pada penelitian ini dilakukan metode klasifikasi untuk memprediksi bagaimana kinerja akademik mahasiswa yang baru terdaftar pada waktu yang mendatang. Sehingga sistem ini nantinya mampu mengidentifikasi mahasiswa mana yang menjanjikan dan juga memberikan mereka kesempatan untuk memperhatikan dan meningkatkan mahasiswa-mahasiswa yang mungkin akan mendapatkan nilai yang lebih rendah. Atribut atau variabel yang digunakan pada penelitian ini yaitu jenis kelamin, hasil nilai ujian dari 10 dan kelas 12, nilai dan peringkat dari ujian masuk universitas, serta nilai dari tahun pertama mahasiswa tersebut masuk universitas. Pada penelitian ini algoritma klasifikasi yang digunakan yaitu ID3 dan C4.5 dan didapat dari hasil bahwa algoritma C4.5 memiliki waktu eksekusi algoritma yang lebih cepat dibandingkan dengan algoritma ID3.

3. Using Data Mining Technique to Analyze Student’s Performance (Ahmed, 2013)

Pada penelitian ini dilakukan metode klasifikasi untuk memprediksi kinerja dari mahasiswa yang dibagi menjadi 5 *class* yang berbeda, yaitu apakah mahasiswa tersebut memiliki kinerja *excellent*, *very good*, *good*, *acceptable* dan *fail*. Pada penelitian ini juga melakukan perbandingan 3 buah algoritma klasifikasi yaitu ID3, C4.5 dan REPTree. Dari hasil penelitian didapat hasil bahwa algoritma C4.5 memiliki akurasi dan kecepatan dalam membuat *decision tree* yang lebih baik daripada kedua algoritma lainnya, serta algoritma C4.5 memiliki tingkat error yang lebih kecil dibandingkan dengan kedua algoritma lainnya.

4. Analisis Komparasi Algoritma Klasifikasi Data Mining untuk Prediksi Mahasiswa Non Aktif (Hastusti, 2012)

Pada penelitian ini dilakukan metode klasifikasi untuk membandingkan tingkat akurasi dari masing-masing algoritma. Algoritma yang digunakan adalah *logistic regression*, *decision tree c4.5*, *naïve bayes* dan *neural network*. Selain itu penelitian ini juga dilakukan untuk memprediksi mahasiswa tersebut tergolong aktif atau non aktif. Hasil dari proses klasifikasi dievaluasi dengan menggunakan *cross validation*, *confusion matrix*, *ROC Curve* dan *T-Test* untuk mengetahui algoritma klasifikasi data mining yang paling akurat untuk prediksi mahasiswa non aktif. Dari hasil penelitian didapat bahwa algoritma *decision tree C4.5* memiliki akurasi yang paling tinggi.

II.4 Jalur Masuk Politeknik Negeri Batam

Politeknik Negeri Batam menyediakan jalur masuk kuliah, yaitu melalui Penelusuran Minat dan Keterampilan (PMDK). Untuk jalur PMDK pihak Politeknik Negeri Batam melakukan penyeleksian terhadap data calon mahasiswa, yaitu :

1. Asal sekolah, pihak Politeknik Negeri Batam akan memilih sekolah yang unggul dan berkualitas, sehingga memberikan nilai yang lebih bagus untuk calon dari sekolah tersebut.
2. Prestasi Akademik, yang meliputi nilai rata-rata rapor tiap semester dan peringkat di kelas.
3. Prestasi Non Akademik, meliputi penghargaan yang pernah diraih, tahun dan tingkatnya.
4. Kesesuaian minat bakat terhadap jurusan yang dipilih di Politeknik Negeri Batam.
5. Kesesuaian jurusan di Sekolah Menengah Atas (SMA) atau Sekolah Menengah Kejuruan (SMK) terhadap jurusan yang dipilih di Politeknik Negeri Batam. Misalnya, untuk jurusan Teknik Elektronika, calon mahasiswa yang diprioritas berasal dari SMA jurusan IPA atau SMK jurusan Teknik Elektronika.

Setelah melalui tahap penyeleksian, maka akan dilakukan wawancara. Wawancara tersebut berguna untuk memastikan data-data dari calon mahasiswa benar dan akurat. Politeknik Negeri Batam menargetkan mahasiswa yang diterima melalui jalur masuk PMDK berkisar 15-20% dari jumlah calon mahasiswa yang mendaftar melalui PMDK.

II.5 Data Operasional

Data operasional yang diambil adalah mahasiswa PMDK Polteknik Negeri Batam angkatan 2012-2013.

Data yang diambil untuk angkatan 2012-2013 mencakup :

1. Data calon mahasiswa PMDK meliputi Pilihan Prodi, Jurusan SLTA, Status Sekolah, dan rata-rata rapor dari semester 1 s/d 5.
2. Prestasi Non Akademik meliputi penghargaan.
3. Wawancara meliputi Nilai Motivasi, Nilai Prestasi Akademik, Nilai Prestasi Non Akademik, dan Nilai Pemecahan Masalah.

II.6 Data Mining

Data mining adalah proses mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan terkait, serta menemukan hubungan, *pattern* dan kecenderungan dengan memeriksa sekumpulan data dari basis data yang besar[5]. Proses *Data mining* berguna untuk menggali nilai tambah dari suatu kumpulan data, bisa berupa pengetahuan selama hal tersebut tidak ditemukan jika dianalisis secara manual. Sehingga, proses *Data mining* dapat menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk menemukan nilai tambah tersebut.

Ada beberapa faktor yang mendorong kemajuan luar biasa dalam bidang *Data mining*, antara lain (Larose, 2005):

1. Pertumbuhan yang cepat dalam pengumpulan data.
2. Penyimpanan data dalam *data warehouse*, sehingga seluruh perusahaan memiliki akses ke dalam *database* yang handal.
3. Adanya peningkatan akses data melalui *web* dan intranet.
4. Perkembangan teknologi perangkat lunak untuk *data mining*.
5. Perkembangan yang hebat terhadap kemampuan komputasi dan kapasitas penyimpanan data.

Dari penjelasan di atas, dapat ditarik tiga kesimpulan tentang *data mining*, yaitu :

1. *Data mining* merupakan proses otomatis terhadap data yang sudah ada.
2. Data yang akan diproses berupa data yang sangat besar.
3. Tujuan *data mining* adalah mendapatkan pola atau hubungan yang mungkin memberikan indikasi yang bermanfaat.

Hubungan yang dicari dalam *data mining* dapat berupa hubungan antara dua atau lebih dalam satu dimensi. Misalnya dalam dimensi penerimaan calon mahasiswa PMDK, kita dapat melihat keterhubungan antara mata pelajaran dengan kemampuan calon mahasiswa PMDK atau prestasi non akademik, maupun dengan prestasi akademiknya. Selanjutnya, penemuan pola merupakan keluaran dari data mining. Misalkan sebuah bank ingin meningkatkan jumlah customer terhadap promo layanan kartu kredit, maka bank tersebut akan mencari pola pelanggan yang ada untuk mengetahui minat calon pelanggan potensial dan yang tidak potensial.

Walaupun data mining merupakan proses otomatis, namun hal ini tidak menggantikan posisi tangan manusia. Manusia harus ikut aktif dalam setiap fase pada proses data mining. Campuran manusia disini berupa eksplorasi dan analisis terhadap data. Sehingga ada keilmuan yang harus dikuasai. Metode dan algoritma memang disediakan dalam data mining untuk menganalisis data, namun tidak menutup kemungkinan terjadinya kesalahan yang fatal jika pengguna tidak menerapkan atau memilih metode dan algoritma yang tepat.

Data mining merupakan tahapan inti (core) dalam proses KDD (Knowledge Discovery in Database). Langkah-langkah proses KDD dijelaskan sebagai berikut:

1. Memahami ruang lingkup aplikasi, berupa pengetahuan sebelumnya yang relevan, dan tujuan dari aplikasi ini.
2. Menciptakan sebuah data target, dan data selection. Pemilihan atau seleksi data dari data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi inilah yang akan digunakan pada tahapan data mining.
3. Data Cleaning/ Data Preprocessing. Sebelum tahap data mining, perlu dilakukan tahap cleaning terhadap data yang telah diseleksi. Proses ini mencakup membuang duplikasi data, mengisi data yang kosong, memperbaiki data yang tidak konsisten.
4. Data Reduction/ Data Transformation. Data reduction membuat representasi data yang mewakili data aslinya untuk memperkecil volume data, namun tetap menghasilkan analisis data yang sama. Coding merupakan proses data yang telah dipilih dan dibersihkan, sehingga data tersebut sesuai untuk proses data mining. Proses coding merupakan proses kreatif yang sangat tergantung pada jenis atau pola informasi yang akan dicari dalam database.
5. Menentukan teknik atau metode pada data mining. Metode tersebut berupa summarization, classification, regrestion, association, clustering.
6. Menentukan algoritma dari metode yang dipilih. Metode dan algoritma dalam data mining sangat bervariasi, pemilihan yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.
7. Data mining. Mencari pola yang menarik dan terkait.
8. Interpretation/ Evaluation. Pola yang dihasilkan dari data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini disebut tahap interpretation. Juga melakukan evaluasi terhadap informasi dan pola yang dihasilkan terhadap fakta atau hipotesis sebelumnya.

Ada beberapa teknik atau metode pengelompokan data sesuai dengan tugasnya masing-masing, yaitu (Larose, 2005):

1. Deskripsi
2. Estimasi
3. Prediksi
4. Klasifikasi
5. Pengklusteran
6. Asosiasi

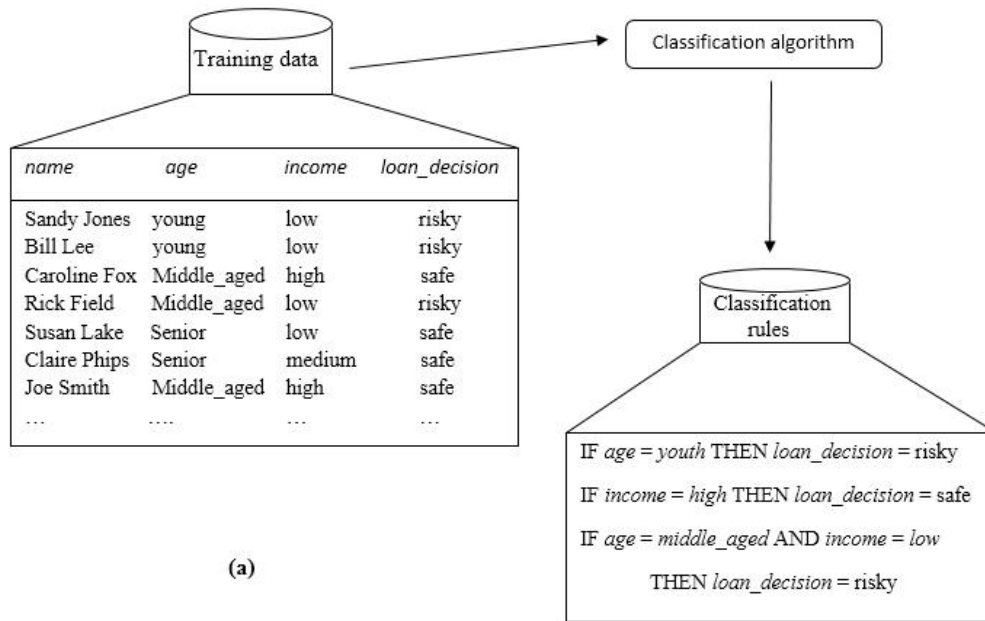
II.6.1 Metode Klasifikasi

Klasifikasi adalah suatu proses menemukan sebuah model berdasarkan data training dan nilai kelas label dari atribut target serta menggunakannya untuk mengklasifikasi suatu data baru (Han, 2006). Data training adalah data yang siap untuk di-mining yang telah melewati data preprocessing. Sedangkan data testing adalah data yang digunakan untuk menguji rule klasifikasi yang diperoleh dari data training. Misalnya, petugas bank yang mengurus pinjaman harus menganalisis data yang ada untuk mempelajari calon kreditur yang tergolong “safe” dan “risky” terhadap bank. Contoh lain, seorang manager marketing AllElectronics membutuhkan data analisis untuk membantu perkiraan jenis komputer yang akan dibeli oleh seorang customer dengan profile yang ada.

Ada dua langkah dalam proses data classification (Han dan Kamber, 2006), yaitu:

1. *Learning*

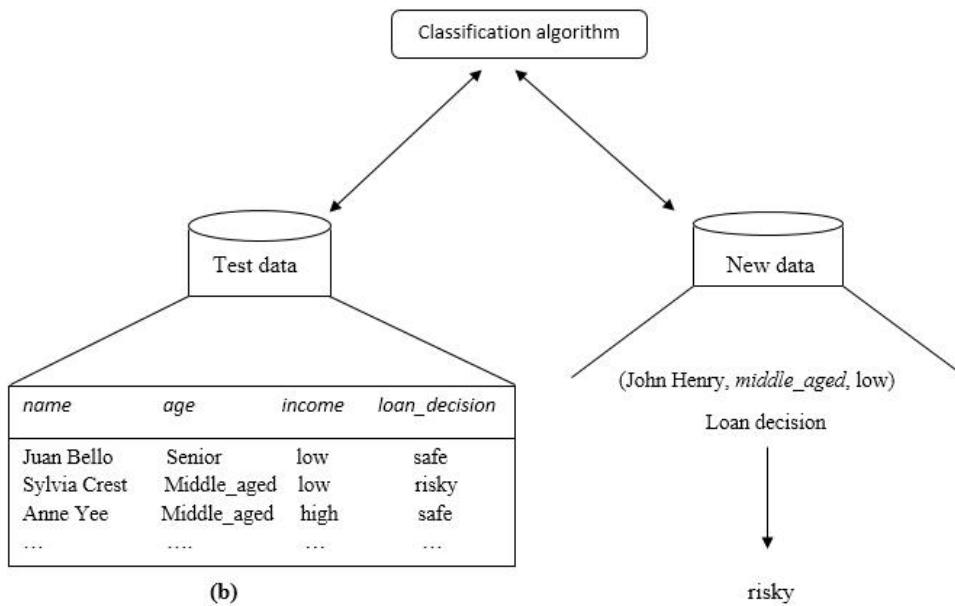
Membangun algoritma klasifikasi dengan menganalisis atau “belajar dari” data training. Karena kelas label untuk setiap data training telah tersedia/ diketahui, maka metode klasifikasi tergolong supervised learning. Dari analisis data training tersebut, terbentuklah *classification rules*. Ilustrasi pada Gambar 1.



Gambar 1. Langkah Learning

2. *Classification*

Dari classification rules, dilakukan pengujian terhadap data testing untuk memperkirakan/ mengestimasi akurasi rule yang diperoleh. Jika rule tepat, maka dapat diaplikasikan untuk data yang baru. Ilustrasi pada Gambar 2.



Gambar 2. Langkah Classification

Untuk menguji data, dapat menggunakan cara *accuracy* yaitu:

1. Pengujian dengan cara membandingkan kelas label yang diketahui dengan hasil tes sampel labelnya mirip atau tidak.

2. Pengujian dengan ketepatan.
3. Pengujian dengan melihat sedikitnya nilai yang tidak sesuai dengan pola yang ditentukan (*over-fitting*) untuk *data training*.

II.6.2 Teknik Decision Tree dalam Metode Klasifikasi

Decision tree atau pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Decision tree mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan-aturan. Aturan tersebut lebih mudah dipahami dengan bahasa alami. Pohon keputusan juga berguna untuk mengeksplorasi data dan menemukan model yang tersembunyi dari data dengan sebuah target variabel.

Pohon keputusan adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan record yang lebih kecil dengan memperhatikan variabel tujuannya (Turban, 2005). Pohon keputusan juga dapat digunakan untuk mengestimasi atau memperkirakan nilai dari variabel kontinyu.

Data dalam pohon keputusan biasanya dinyatakan dalam bentuk tabel dengan atribut dan record. Atribut menyatakan suatu parameter yang dibuat sebagai kriteria dalam pembentukan pohon. Misalkan untuk menentukan pergi bermain tenis, kriteria yang diperhatikan adalah cuaca, angin, dan temperatur. Salah satu atribut merupakan atribut yang menyatakan data solusi per item data yang disebut target atribut. Atribut memiliki nilai-nilai yang dinamakan dengan instance. Misalkan atribut cuaca mempunyai instance berupa cerah, berawan, dan hujan.

Salah satu algoritma untuk membentuk pohon keputusan adalah C4.5. Secara umum langkah untuk membangun algoritma C4.5 adalah (Kusrini, 2009):

1. Pilih atribut yang menjadi root.
2. Buat cabang untuk setiap nilai.
3. Bagi kasus dalam cabang.
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut yang menjadi root, diperoleh dari nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain, digunakan rumus pada persamaan 1.

$$GAIN(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad \text{Persamaan 1}$$

dengan :

- S : himpunan kasus
- A : atribut
- n : jumlah partisi atribut A
- $|S_i|$: jumlah kasus pada partisi ke i
- $|S|$: jumlah kasus dalam S

Untuk menghitung nilai *entropy* dapat digunakan rumus pada persamaan 2.

$$Entropy(S) = - \sum_{i=1}^n p_i * \log_2 p_i \quad \text{Persamaan 2}$$

dengan :

- S : himpunan Kasus
- A : fitur
- n : jumlah partisi S
- p_i : proporsi dari S_i terhadap S

Setelah menghitung entropy dan gain maka akan menghasilkan *decision tree*. Decision tree kemudian diuji untuk mengetahui estimasi keakuratannya. Semakin sedikit error rate (kesalahan) yang dihasilkan dari decision tree maka semakin akurat decision tree yang dihasilkan. Untuk menghitung error rate digunakan persamaan 3.

$$E_{r_i} (E) = \frac{b}{j_i} \frac{d}{ah d} \frac{e}{ah d} * 100\% \quad \text{Persamaan 3}$$

Contoh penggunaan:

1. Dari decision tree yang dihasilkan maka dilakukan pengujian. Hasil pengujian disediakan pada tabel 2.

Table 2. Contoh Kasus Data Pengujian

Nama	Usia	Berat	Kelamin	Hipertensi (Class Label)	Berdasarkan Bobot
Abi	muda	overweight	pria	ya	ya
Rio	muda	underweight	pria	tidak	tidak
Ina	muda	average	wanita	tidak	tidak
Riko	tua	overweight	pria	tidak	tidak
Iman	tua	overweight	pria	ya	tidak
Dodo	muda	underweight	pria	tidak	tidak
Rini	tua	overweight	wanita	ya	ya
Giring	tua	average	pria	tidak	tidak

Penyelesaian:

Jumlah data = 8

Jumlah data error = 1

$$E r (E) = \frac{b}{jv} \frac{d}{hd} \frac{e}{e} * 100\%$$

$$Error\ rate\ (ER) = \frac{1}{8} * 100\%$$

$$Error\ rate\ (ER) = 12,5\ \%$$

II.7 Perangkat-Perangkat Lunak yang Mendukung dalam Analisis

Aplikasi yang nantinya akan dibangun tentunya memiliki perangkat-perangkat untuk menunjang terbentuknya aplikasi, diantaranya adalah sebagai berikut:

II.7.1 Microsoft Excel

Menurut Ananda (2011), Microsoft Excel merupakan perangkat lunak (*software*) yang mengolah data, meliputi perhitungan dasar, penggunaan fungsi-fungsi, pembuatan grafik, dan manajemen data. Selain itu, Excel juga dapat digunakan untuk menyelesaikan berbagai urusan administratif dari yang sederhana sampai kompleks. Proses menganalisis data peminjaman buku di perpustakaan Otorita Batam menggunakan Microsoft Excel 2013. Proses menganalisis data yang dimaksud salah satunya adalah proses *generalisasi* data.

Alasan menggunakan Microsoft Excel dalam pengolahan data peminjaman buku karena beberapa kelebihanannya, antara lain:

1. Mempunyai kemampuan menampung data yang cukup besar dengan 1 juta baris dan 16.000 kolom dalam 1 sheet. Jadi dalam 1 sheet bisa menampung jawaban 1 juta responden dan 16 ribu jawaban atau pertanyaan.
2. Microsoft Excel mempunyai format yang paling populer dan fleksibel jadi sebagian besar software data entry ada fasilitas konversi ke format Excel atau format lain yang bisa dibaca Excel. Jika dibutuhkan pengguna Excel bisa konversi balik dari Excel ke software statistik lainnya.
3. Microsoft Excel mempunyai program penggunaan rumus yang sangat lengkap sehingga mempermudah pengolahan angka untuk menghasilkan dokumen yang lebih canggih.
4. Menggunakan Pivot Tables, pengguna bisa kerja lebih efektif karena semua tabel summary yang pengguna rencanakan dapat dibuat dahulu walaupun data belum masuk

semua. Setiap ada data masuk otomatis pivot table akan me-refresh sehingga tabel akan terupdate sendiri.

II.7.2 Microsoft Visio

Menurut Pascal (2007), Microsoft Visio adalah software pembuat diagram dan flowchart berbasis vektor yang awalnya dibuat oleh Visio Corporation. Visio mulai terintegrasi di dalam Microsoft Office sejak Office 2007. Sejumlah fitur baru ditambahkan, terutama di integrasi antara data dan diagram, serta otomatisasi dalam pembuatan bagan dan diagram.

Microsoft Visio yang digunakan adalah Microsoft Visio 2013 untuk membuat FP-tree dari algoritma FP-growth. Kelebihan dari Microsoft Visio 2013 yaitu:

1. Tool-toolnya sangat mudah dipahami.
2. Cara penggunaannya sangat mudah.
3. Lebih banyak pilihan gambarnya.

II.7.3 Netbeans

Menurut Whitten (2002), Netbeans adalah Integrated Development Environment (IDE) berbasis java dari Sun Microsystem yang berjalan diatas Swing. Netbeans merupakan software development yang open source dengan kata lain software ini dibawah pengembangan bersama dan bebas biaya. Alasan digunakannya Netbeans dalam pembuatan sistem penerapan data mining ini karena sesuai dengan kemampuan laptop yang digunakan untuk pembuatan sistem hingga running sistem.

Netbeans yang akan digunakan untuk sistem penerapan data mining ini adalah Netbeans 7. Netbeans 7 memiliki beberapa fitur baru yang lebih sempurna dibandingkan Netbeans versi sebelumnya. Beberapa fitur baru Netbeans 7 yaitu JDK 7, weblogic server, oracle database, glassfish, java, Java EE, bahasa pemrograman web, PHP, dan C/C++. Aplikasi java memiliki komponen GUI yang disimpan pada kontainer yang biasa disebut form. Bahasa pemrograman java mengumpulkan komponen antarmuka dari form GUI yang telah terbentuk. Pada Netbeans komponen GUI ini telah tersedia. Hal ini mempermudah pengguna dalam membangun dan merancang form java.

II.7.4 Pengertian Java

Java menurut definisi Sun dalam M. Shalahudin (2009 :17) adalah nama untuk sekumpulan teknologi untuk membuat dan menjalankan perangkat lunak pada komputer *standalone*

ataupun pada lingkungan jaringan. Meskipun pada awal saat dirilis sekitar tahun 90-an. Java dirancang untuk digunakan pada sistem-sistem kecil seperti TV kabel atau *home theater*, sekarang sudah merambah keseluruhan aplikasi pada komputer bahkan beberapa institusi pendidikan beralih dari pemrograman dengan Pascal dan C++ ke pemrograman Java. Java adalah suatu bahasa pemrograman di dunia software komputer yang di kenal sebagai bahasa pemrograman tingkat tinggi yang mudah di pelajari terutama bagi seorang programmer yang sebelumnya telah mengenal bahasa/pemrograman C/C++.

II.7.5 Basis Data (*Database*)

Database atau Basis Data adalah sekumpulan data yang saling terhubung satu dengan yang lainnya atau sekumpulan *table* yang saling terhubung satu dengan yang lainnya. dan fungsi dari *database* adalah menyimpan suatu data pada tabel-tabel dan dikumpulkan menjadi satu dengan *database*. *Database* juga bisa di umpamakan sebagai sebuah rumah dengan beberapa kamar-kamar dan sebuah property seperti almari meja belajar tempat tidur itu bisa di sebut dengan data *querynya* (Arbie, 2004).

Ada beberapa bagian bagian dari *database* yaitu:

A. Komponen *Database*

Komponen yang terdapat pada suatu *database* antara lain :

i) *Table*

Sebuah komponen yang digunakan untuk menyimpan suatu data yang telah di akses dan dimasukkan kedalamnya.

ii) *Record*

Isi atau data dari table tersebut yang telah dikelola. *Record* dapat mempunyai beberapa macam data. Dan data berfariasi tersebut di simpan kedalam table dan itulah yang disebut *record*.

iii) *Field*

Pemberian identitas suatu data dimana data tersebut akan di letakkan. Sesuai dengan pengelompokan datanya.

B. Struktur *Database*

Strukture *database* adalah suatu pengaturan *field-field* pada suatu tabel pada *database*.

Beberapa struktur *database* sebagai berikut:

i) *Nama Field*

Digunakan sebagai suatu pemberian identitas atau member keterangan pada *field*.

ii) *Type Data*

Pemberian suatu tipe pada *field* sesuai dengan identitas yang telah diberikan.

iii) *Ukuran Data*

Pemberian suatu panjang atau banyak data yang telah di inputkan.

iv) *Keterangan*

Memberikan suatu keterangan atau deskripsi pada sebuah *field*.

II.7.6 XAMPP

Menurut Madcoms (2009), XAMPP adalah salah satu paket software web server yang terdiri dari Apache, MySQL, PHP, dan phpMyAdmin. Perangkat lunak komputer ini memiliki kelebihan untuk bisa berperan sebagai server web Apache untuk simulasi pengembangan website. Tool pengembangan web ini mendukung teknologi web populer seperti PHP, MySQL, dan Perl. XAMPP juga dilengkapi fitur manajemen database PHPMyAdmin seperti pada server hosting sungguhan, sehingga pengembang web dapat mengembangkan aplikasi web berbasis database secara mudah.

XAMPP akan digunakan sebagai tempat pembuatan dan penyimpanan database dari aplikasi penerapan data mining ini. XAMPP yang digunakan untuk pembuatan database adalah XAMPP versi 1.8.3.5. Alasan lebih memilih XAMPP dibandingkan tempat pembuatan database lainnya karena XAMPP memiliki beberapa kelebihan, yaitu:

1. Dapat berperan sebagai server web Apache untuk simulasi pengembangan website.
2. Melalui XAMPP, programmer dapat menguji aplikasi web yang dikembangkan dan merepresentasikannya ke pihak lain secara langsung dari komputer, tanpa perlu terkoneksi ke internet.
3. XAMPP dilengkapi fitur manajemen database PHPMyAdmin seperti pada server hosting sungguhan, sehingga pengembang web dapat mengembangkan aplikasi web berbasis database secara mudah.
4. XAMPP dapat dijalankan di sistem operasi Windows 2000/XP/Vista/7 dan sistem operasi lainnya.

II.7.7 UML

Menurut Triandini (2012), *StarUML* adalah *platform* pemodelan perangkat lunak yang mendukung UML (*Unified Modelling Language*). *StarUML* secara aktif mendukung pendekatan MDA (*Model Driven Architecture*) dengan mendukung konsep UML *profile*. *StarUML* memiliki 9 fitur diagram yang mendukung, antara lain:

1. *Use Case Diagram*
2. *Class Diagram*
3. *Sequence Diagram*
4. *State Machine Diagram*
5. *Activity Diagram*
6. *Component Diagram*
7. *Deployment Diagram*
8. *UML Profile*
9. *UML Stereotype*

3. *Sequence Diagram*

8. *Deployment Diagram*

4. *Collaboration Diagram*

9. *Composite Structure Diagram* (UML versi 2.0)

5. *Statechart Diagram*

StarUML digunakan untuk membuat beberapa pemodelan sistem yang menerapkan data mining ini. Diagram yang digunakan untuk pemodelan sistem hanya 3 dari 9 diagram, yaitu use case diagram, class diagram, dan sequence diagram. StarUML memiliki beberapa keunggulan, yaitu: ringan, open source, dan simbol-simbolnya mudah di mengerti. StarUML unggul dalam hal kustomisasi lingkungan kerja pengguna, dan memiliki ekstensibilitas tinggi pada fungsionalitasnya. StarUML yang digunakan untuk pemodelan sistem ini yaitu versi 5.0.2.

II.8 Distribusi Frekuensi dan Interval Kelas

Distribusi Frekuensi adalah pengelompokkan data dalam beberapa kelas atau kelompok, kemudian dihitung banyaknya data yang masuk ke dalam tiap kelas. Dalam menentukan kelas, digunakan rumus distribusi frekuensi pada persamaan 4.

$$k = 1 + 3.322 \log n$$

Persamaan 4

Keterangan:

k = banyaknya kelas

n = banyaknya nilai observasi atau data

Prinsip-prinsip dalam pembentukan Distribusi Frekuensi yaitu:

1. Tentukan banyaknya kelas dengan rumus 6.

2. Tentukan interval/selang kelas

Satu data hanya dapat dimasukkan ke dalam satu kelas (jadi tidak terjadi overlapping).

3. Sorting data (Ascending atau Descending)

Agar range data diketahui dan mempermudah penghitungan frekuensi tiap kelas.

Untuk menentukan interval kelas dapat digunakan rumus persamaan 7.

$$c = \frac{X_n - X_1}{k}$$

Persamaan 5

Keterangan:

c = perkiraan besarnya (jumlah interval kelas)

X_n = data tertinggi

X_1 = data terendah

k = jumlah kelas

Contoh:

1. Data dibawah ini merupakan data nilai mahasiswa Universitas X mata kuliah Matematika:

Tabel 3. Data Nilai Bahasa Indonesia Mahasiswa Universitas Y

Nama Mahasiswa	Nilai
Cika	9,5
Doni	4,0
Gino	7,8
Raka	9,0
Lidya	8,8
Yunita	7,8
Yudi	9,0
Rina	8,9

Buatlah range dari nilai matematika tersebut!

Penyelesaian:

$$n = 8$$

$$X_n = 9,5$$

$$X_1 = 4,0$$

Tentukan jumlah kelas tersebut dengan persamaan 6.

$$k = 1 + 3,322 \log n$$

$$k = 1 + 3,322 \log (8)$$

$$k = 4,322 * 0,90308998$$

$$k = 3,90315492377$$

$$k = 3 \text{ (diambil angka depan)}$$

Kemudian menentukan interval kelas dengan menggunakan persamaan 7.

$$c = \frac{X - X}{k}$$

$$c = \frac{9,5 - 4,0}{3}$$

$$c = 1,83$$

Jadi interval kelas yang akan dibentuk adalah 1,83.

$$\text{Nilai A : } 7,68 \leq x \leq 9,5$$

$$\text{Nilai A : } 5,84 \leq x \leq 7,67$$

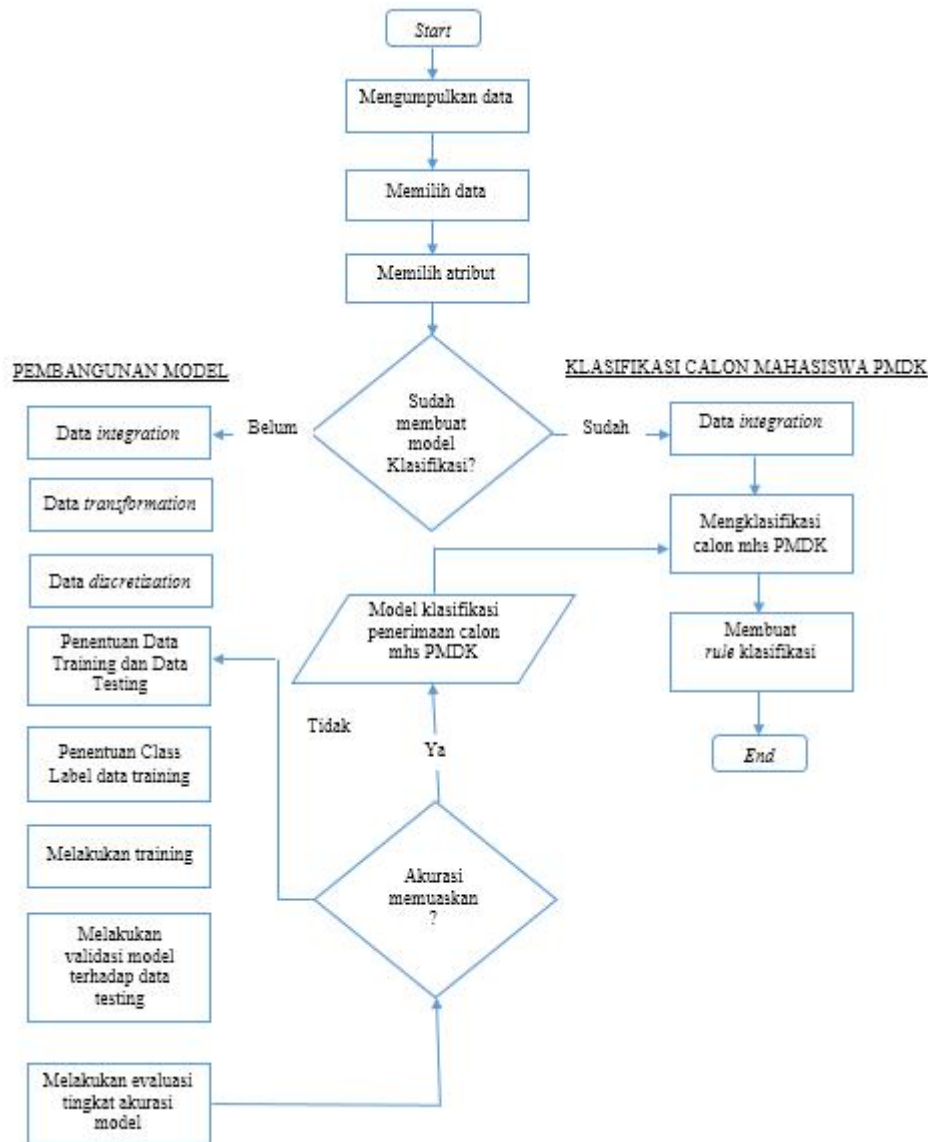
$$\text{Nilai A : } 4,0 \leq x \leq 5,83$$

Tabel 4. Data Nilai Bahasa Indonesia Mahasiswa Universitas Y dengan Range

Nama Mahasiswa	Nilai Angka	Nilai
Cika	9,5	A
Doni	4,0	C
Gino	7,8	A
Raka	9,0	A
Lidya	8,8	A
Yunita	7,8	A
Yudi	9,0	B
Rina	8,9	B

BAB III ANALISIS DAN PERANCANGAN

Bab ini memuat uraian langkah-langkah dalam menganalisis penerimaan calon mahasiswa PMDK. Proses analisis ditunjukkan pada gambar III.1



Gambar 3. Flowchart Proses Analisis dan Klasifikasi

III.1 Pengumpulan Data

Data statistik yang diharapkan adalah data yang dapat dipercaya dan tepat waktu. Oleh karena itu, pengumpulan data harus baik dan mencakup seluruh unit yang menjadi objek penelitian. Apapun tujuan pengumpulan data, terlebih dahulu harus diketahui elemen (unit

terkecil objek penelitian) dan populasi (kumpulan dari seluruh elemen sejenis tetapi dapat dibedakan satu sama lain) objek penelitian (Supranto, 2003).

Dalam analisis ini, populasi yang diteliti adalah kumpulan-kumpulan calon mahasiswa PMDK di Politeknik Negeri Batam angkatan 2012-2013, dan elemennya adalah calon mahasiswa (orang), data PMDK itu sendiri di dapat dari bagian akademik Politeknik Negeri Batam.

Metode pengumpulan data yang digunakan untuk menganalisis calon penerimaan mahasiswa PMDK adalah dengan metode sensus. Sensus adalah cara pengumpulan data di mana seluruh elemen polulasi diselidiki satu per satu. Data yang diperoleh sebagai hasil pengolahan sensus disebut data yang sebenarnya (*true value*), atau sering disebut parameter (Supranto, 2003).

III.2 Pemilihan Atribut

Atribut yang digunakan untuk angkatan 2012-2013 dipilih sesuai kebutuhan menganalisis penyeleksian calon mahasiswa PMDK Politeknik Negeri Batam. Adapun atribut data yang dipilih dari sumber data, yaitu : Prodi pilihan, jurusan SLTA, status sekolah, penghargaan, nilai rata-rata rapor semester 1-5, nilai motivasi, nilai prestasi akademik, nilai prestasi non akademik, dan nilai pemecahan masalah.

III.3 Tahapan Preprocessing Data

Data yang akan di-*mining* harus melewati data *preprocessing*, karena sumber data yang diperoleh dari *database* masih bersifat kotor, artinya terdiri dari beberapa data yaitu data *noise*, data tidak lengkap, banyak data yang kosong dan tidak konsisten. Kualitas data mempengaruhi hasil *data mining*. Preprocessing terdiri dari 4 jenis yaitu data *cleaning*, data *integration*, data *transformation*, dan data *reduction*. Jumlah data operasional PMDK 2012-2013 yang diperoleh sebanyak 605 data. Pada penelitian ini, *preprocessing* data yang akan dilakukan adalah transformasi data, *cleaning* data, dan data *discretization*.

III.3.1 Transformasi Data

Data transformation atau transformasi data adalah mengubah bentuk data menjadi bentuk yang sesuai untuk proses *mining*. Transformasi data terdiri dari beberapa jenis yaitu normalisasi, generalisasi, penghalusan, agregasi, dan konstruksi atribut. Data yang telah terkumpul akan ditransformasi secara generalisasi.

Generalisasi data adalah mengubah bentuk data mentah ke konsep yang lebih tinggi melalui penggunaan konsep hierarki. Generalisasi data dalam penelitian ini adalah mengubah data mentah yang berupa nama sekolah kemudian digolongkan ke dalam suatu jenis nama sekolah yang sama berupa status sekolah. Misal nama sekolah SMK Negeri 1 Singkep di generalisasikan menjadi SMK Negeri. Maka SMK Negeri 1 singkep tergolong kedalam status sekolah yang sudah Negeri. Contoh pengelompokan nama sekolah berdasarkan status sekolah sebagai berikut :

Tabel 5. Tranformasi Data

1	Nama Sekolah	Status Sekolah
2	MA Unit Sekolah Baru Batam	MA Negeri
3	SMK Negeri 2 Batam	SMK Negeri
4	SMA Negeri 3 Batam	SMA Negeri
5	SMK Kartini Batam	SMK Swasta
6	SMK Real Informatika Batam	SMK Swasta
7	SMA Negeri 1 Batam	SMA Negeri
8	Madrasah Aliyah Negeri 1 Batam	MA Negeri
9	Permata Harapan Batam	SMK Swasta
10	SMK Negeri 1 Singkep	SMK Negeri

III.3.2 Data Cleaning

Data cleaning digunakan untuk melengkapi atau menghilangkan data yang tidak lengkap, menghaluskan atau meniadakan data noisy, dan memperbaiki data yang tidak konsisten. Data cleaning dalam penelitian ini adalah mengubah data yang tidak lengkap dengan mengisi nilai secara manual, misal data prodi akuntansi di ubah menjadi D3 Akuntansi. Kemudian data noisy misal data prodi terdapat prodi D3 Instrumentasi maka data tersebut *outlier* karena tidak terdapat dalam data prodi yang ditetapkan di Politeknik Negeri Batam. Pada penelitian ini juga terdapat memperbaiki data yang tidak konsisten, misal pada data prodi Teknik Multimedia & Jaringan dengan D4 Teknik Multimedia dan Jaringan. Keduanya terlihat sama namun data tersebut mengandung data yang tidak konsisten. maka dari itu perlu diperbaiki sesuai dengan ketentuan nama prodi dari Politeknik Negeri Batam yaitu D4 Teknik Multimedia dan Jaringan.

III.3.3 Data Dizcretization

Melakukan perubahan data dengan menggolongkan *range* (interval) data tertentu ke level tertentu. Digunakan untuk data numerik. Level yang diperoleh dijadikan nilai interval untuk perhitungan penentuan calon mahasiswa PMDK pada *data training*. Interval yang diberikan untuk penentuan kelas rata-rata nilai rapor semester 1-5 menggunakan rumus persamaan 7:

$$c = \frac{X_n - X_1}{k}$$

Untuk mendapatkan jumlah kelas (k) maka dihitung dengan menggunakan rumus pada persamaan 6:

$$k = 1 + 3.322 \log n$$

misalnya kita akan melakukan perhitungan untuk menemukan interval rata-rata nilai rapor semester 1-5.

n = banyaknya jumlah data = 272 (jumlah data *training*)

$$k = 1 + 3.322 \log n$$

$$k = 1 + 3.322 \log 272$$

$$k = 9.08763790$$

$$k = 9 \text{ (dibulatkan)}$$

$k = 9.08763790$ (maka jumlah kelas yang akan dibentuk pada nilai rata-rata rapor semester 1-5 adalah 9)

mencari interval (c) pada nilai rata-rata rapor semester 1-5 :

$$c = \frac{X_n - X_1}{k}$$

X_n = data tertinggi pada nilai rata-rata rapor semester 1-5 = 94

X_1 = data terendah pada nilai rata-rata rapor semester 1-5 = 0

$$k = 9$$

$$c = (94-0)/9$$

$$c = 10.41$$

sehingga, range nilai rata-rata rapor semester 1-5 sebagai berikut :

Tabel 6. Range Rata-rata Sem 1-5

Interval		Kelas
0	10.41	1
10.42	20.82	2
20.83	31.23	3
31.24	41.64	4
41.65	52.05	5
52.06	62.46	6
62.47	72.87	7
72.88	83.28	8
83.29	94	9

III.4 Pemilihan Data Training dan Data Testing

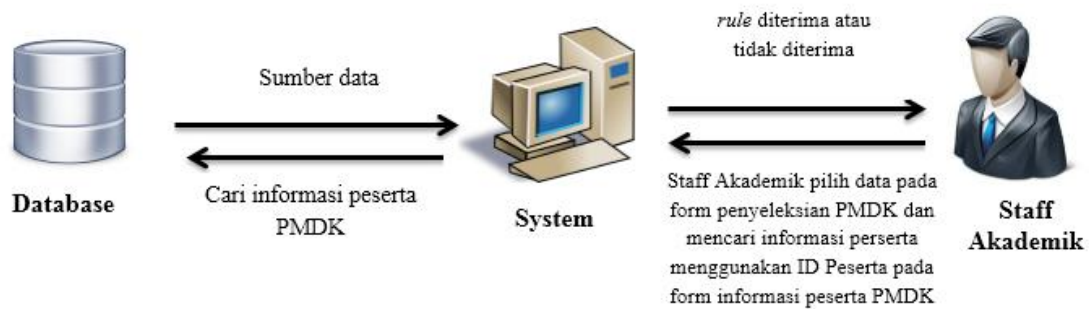
Setelah persiapan data selesai dilakukan, langkah berikutnya adalah memasukkan sebagian data kedalam algoritma *data mining*. Algoritma data mining perlu mempelajari pola data yang diberikan guna menarik informasi data tersebut. Istilah itu dikenal dengan sebutan training data mining.

Selanjutnya dilakukan validasi apakah *data mining* memberikan prediksi yang akurat. Setelah pelatihan data selesai dilakukan, *data mining* tersebut perlu diuji atau divalidasi keakuratannya terhadap *data testing*. Biasanya tidak hanya 1 algoritma *data mining* yang diimplementasikan ke dalam suatu *data mining*. Berarti data yang dipersiapkan pada persiapan data mengandung data yang digunakan untuk *training* dan data yang digunakan untuk *testing*. Pemilihan *data training* dan *data testing* dapat dilakukan dengan menggunakan metode *holdout* (Witten, 2000).

Holdout dimulai dengan membagi jumlah data yang telah melewati proses *preprocessing* menjadi data *training* dan data *testing*. Jumlah data analisis calon penerimaan mahasiswa PMDK sebanyak 340 data. Lalu 340 data dibagi menjadi (80:20). Hasil pembagian data training menjadi 272 data sementara data testing menghasilkan sebanyak 68 data.

III.5 Deskripsi Sistem

Deskripsi umum sistem analisis calon mahasiswa PMDK Politeknik Negeri Batam menggunakan metode adalah sebagai berikut :



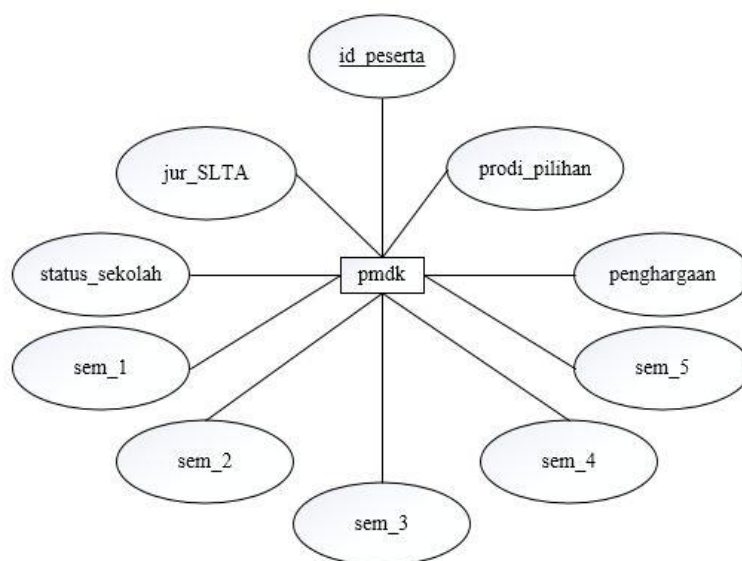
Gambar 4. Deskripsi Sistem

Adapun aktor yang berperan dalam aplikasi ini adalah Staff Akademik. Sistem akan menghasilkan rule klasifikasi berupa hasil penentuan apakah mahasiswa akan diterima atau tidak di jalur PMDK tersebut.

Database tersebut akan menjadi sumber data training dan testing, dari proses training akan di peroleh model *decision tree*, kemudian sebagian data dari *database* akan dijadikan data *testing* untuk menguji tingkat akurasi model *decision tree* yang diperoleh dari proses *training*. Model yang memiliki akurasi yang paling tinggi yang akan digunakan pada aplikasi. Sehingga Staff Akademik dapat mengklasifikasikan data baru berdasarkan model yang telah diperoleh.

III.6 Diagram ER

Menurut salah satu para ahli, Brady dan Loonam (2010), *Entity Relationship Diagram* (ERD) merupakan teknik yang digunakan untuk memodelkan kebutuhan data dari suatu organisasi, biasanya oleh *system analys* dalam tahap analisis persyaratan proyek pengembangan *system*. Sementara seolah-olah teknik diagram atau alat peraga memberikan dasar untuk desain *database relasional* yang mendasari sistem informasi yang dikembangkan. ERD bersama-sama dengan detail pendukung merupakan model data yang pada gilirannya digunakan sebagai spesifikasi untuk *database*.



Gambar 5. Diagram ER

Berdasarkan gambar 5, dapat dilihat tabel pmdk dari *database* pmdk_polibatam memiliki beberapa atribut yang akan digunakan untuk penyimpanan data-data pmdk. Atribut-atribut yang dimaksud yaitu id_peserta, prodi_pilihan, jur_SLTA, status_sekolah, penghargaan, sem_1, sem_2, sem_3, sem_4, sem_5. Adapun atribut yang digunakan sebagai *primary key* yaitu atribut id_peserta karena sifatnya yang berbeda-beda untuk tiap peserta pmdk artinya setiap peserta memiliki id yang berbeda-beda. Pada diagram ER ini tidak memiliki relasi karena hanya memiliki satu *entity* saja.

III.7 Spesifikasi Database pmdk_polibatam

Penjelasan dari tabel pmdk pada *database* pmdk_polibatam dapat dilihat pada Tabel 9.

Tabel 7. Spesifikasi Tabel PMDK

Nama Field	Deskripsi Isi	Type & Panjang	Boleh NULL	Keterangan
Id_peserta	Berisi id peserta	Int(50)	Tidak boleh	Primary key
prodi_pilihan	Berisi prodi pilihan	Varchar(200)	Tidak boleh	-
jur_SLTA	Berisi jurusan SLTA	Varchar(200)	Tidak boleh	-
status_sekolah	Berisi status sekolah	Varchar(200)	Tidak boleh	-
penghargaan	Berisi penghargaan	Varchar(50)	Tidak boleh	-
sem_1	Berisi nilai rapor sem 1	Varchar(100)	Tidak boleh	-
sem_2	Berisi nilai rapor sem 2	Varchar(100)	Tidak boleh	-
sem_3	Berisi nilai rapor sem 3	Varchar(100)	Tidak boleh	-
sem_4	Berisi nilai rapor sem 4	Varchar(100)	Tidak boleh	-
sem_5	Berisi nilai rapor sem 5	Varchar(100)	Tidak boleh	-

III.8 Lingkungan Operasional

Spesifikasi dari aplikasi ini adalah :

- a. Bahasa pemrograman : Java
- b. Perangkat keras (*hardware*).
 - Processor : AMD C-70 APU with Radeon(tm) HD Graphics 1.00 GHz.
 - Kebutuhan memori utama : Minimal 130 MB.
- c. Sistem operasi (*Operating System*): Windows 7.
- d. *Database*: XAMPP.

III.9 Spesifikasi Kebutuhan Fungsional dan Non Fungsional

III.9.1 Spesifikasi Kebutuhan Fungsional

Kebutuhan yang dibutuhkan pada sistem ini diantaranya adalah sebagai berikut:

Tabel 8. Kebutuhan Fungsional

No	Kebutuhan Fungsional
F001	Sistem dapat menyeleksi peserta berdasarkan <i>rule</i> yang dipilih
F002	Sistem dapat mencari informasi peserta berdasarkan ID peserta

III.9.2 Spesifikasi Kebutuhan Fungsional

Selain memiliki kebutuhan fungsional, sistem ini juga memiliki kebutuhan non fungsional, diantaranya terlihat pada tabel dibawah ini:

Tabel 9. Kebutuhan Non Fungsional

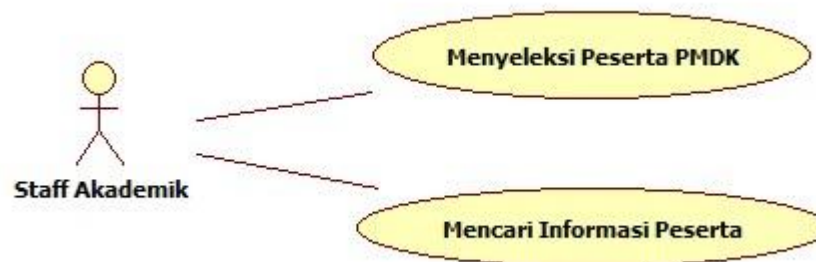
No	Kebutuhan Non Fungsional
NF001	Tampilan sistem dibuat sederhana supaya pengguna mudah menggunakannya.
NF002	Menggunakan bahasa Indonesia
NF003	Sistem memiliki rancangan antarmuka yang user friendly.

III.10 Use Case

Use case mendeskripsikan interaksi tipikal antara para pengguna sistem dengan sistem itu sendiri, dengan *member* sebuah narasi tentang bagaimana sistem tersebut digunakan. *Use Case Diagram* menampilkan *actor* mana yang menggunakan *Use Case* mana, *Use Case* mana yang memasukan *Use Case* lain dan hubungan antara *actor* dan *Use Case*. Sistem yang nantinya akan dibuat, memiliki tolak ukur dari *use case* sendiri dan penjabaran dari skenario *use case*. Diantaranya adalah sebagai berikut:

III.10.1 Diagram *Use Case*

Dalam tugas akhir ini, sistem akan menampilkan program dari dasar seperti dibawah ini:



Gambar 6. Diagram *Use Case*

III.4.1 Skenario *Use Case*

Dari gambar III.3 diatas, terdapat 1 aktor yang berperan dalam aplikasi yang nantinya akan dibangun yakni Staff Akademik. Dan berikut adalah penjabaran atau skenario dari diagram *Use Case* diatas:

3.4.1.1 Skenario *Use Case* Menyeleksi Peserta PMDK

Aktor : Staff Akademik

Kondisi Awal : Staff Akademik masuk ke halaman utama

Kondisi Akhir : Menampilkan hasil/ouput Diterima atau Tidak Diterima

Skenarion :

1. Staff akademik memilih/menginput *rule* yang akan dimasukkan kedalam sistem
2. Staff akademik akan menekan *button* hasil
3. Sistem akan memproses *rule* yang telah dimasukkan oleh staff akademik
4. Sistem akan menampilkan hasil/*output* yang menyatakan calon mahasiswa PMDK akan diterima atau tidak berdasarkan *rule* yang dimasukkan.

III.4.1.2 Skenario *Use Case* Mencari Informasi Peserta

Aktor : Staff akademik

Kondisi Awal : Staff akademik sudah masuk ke antarmuka pencarian informasi peserta

Kondisi Akhir : Menampilkan hasil informasi peserta kedalam *field*

Skenario :

1. Staff akademik menginput ID peserta kedalam sistem
2. Staff akademik menekan *button* cari

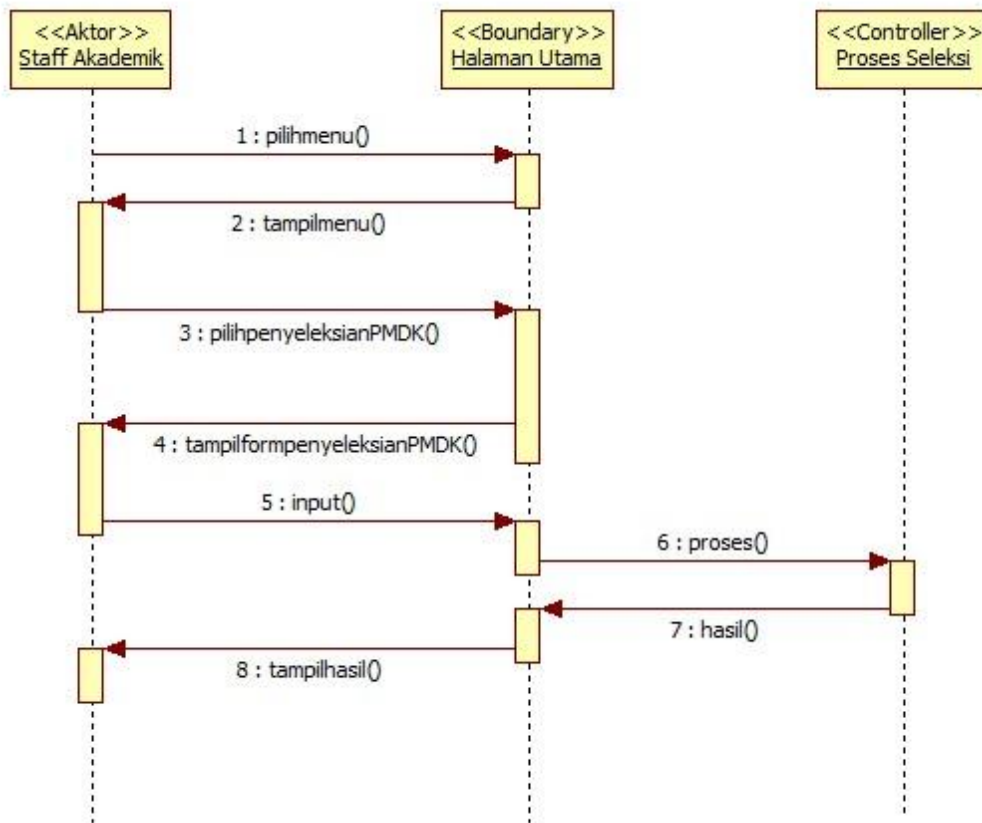
3. Sistem akan memproses dan mengecek ke dalam *database* apakah ID yang dimasukkan sesuai atau tidak dengan *database*
4. Jika ID yang dimasukkan tidak sesuai dengan yang ada didalam *database* maka sistem akan menampilkan pesan berupa data tidak ditemukan.
5. Jika ID yang dimasukkan sesuai dengan yang ada didalam *database* maka sistem akan menampilkan pesan berupa data ditemukan dan menampilkan hasil data yang dicari kedalam *field* yang telah tersedia di sistem.

III.11 Sequence Diagram

Sequence diagram secara khusus menjabarkan *behavior* sebuah skenario tunggal. *Sequence diagram* menunjukkan sebuah objek contoh dan pesan-pesan yang melewati objek-objek dalam *use case*.

III.11.1 Sequence Diagram Menyeleksi Peserta PMDK

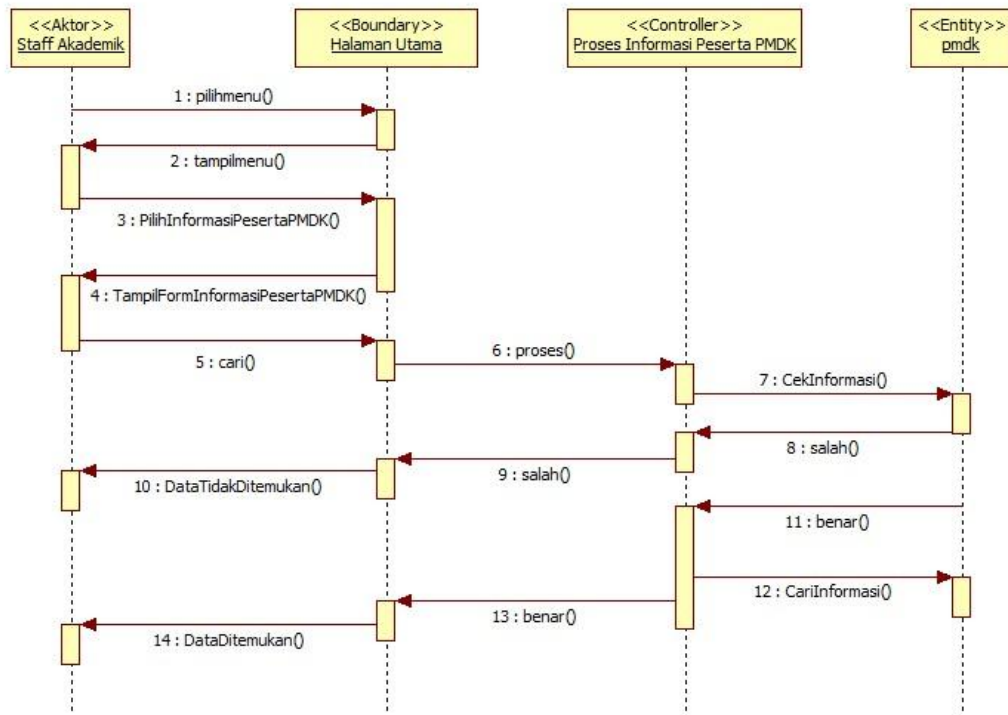
Sequence diagram ini bermula dari staff akademik memilih Penyeleksian PMDK pada menu halaman utama kemudian sistem menampilkan form penyeleksian PMDK, staff akademik menginput/memilih *rule* yang ada pada sistem dan menekan *button* hasil kemudian sistem akan memproses dan menampilkan hasil/*ouput* yang menyatakan calon mahasiswa PMDK diterima atau tidak diterima.



Gambar 7. Sequence Menyeleksi Peserta PMDK

III.11.2 Sequence Diagram Mencari Informasi Peserta

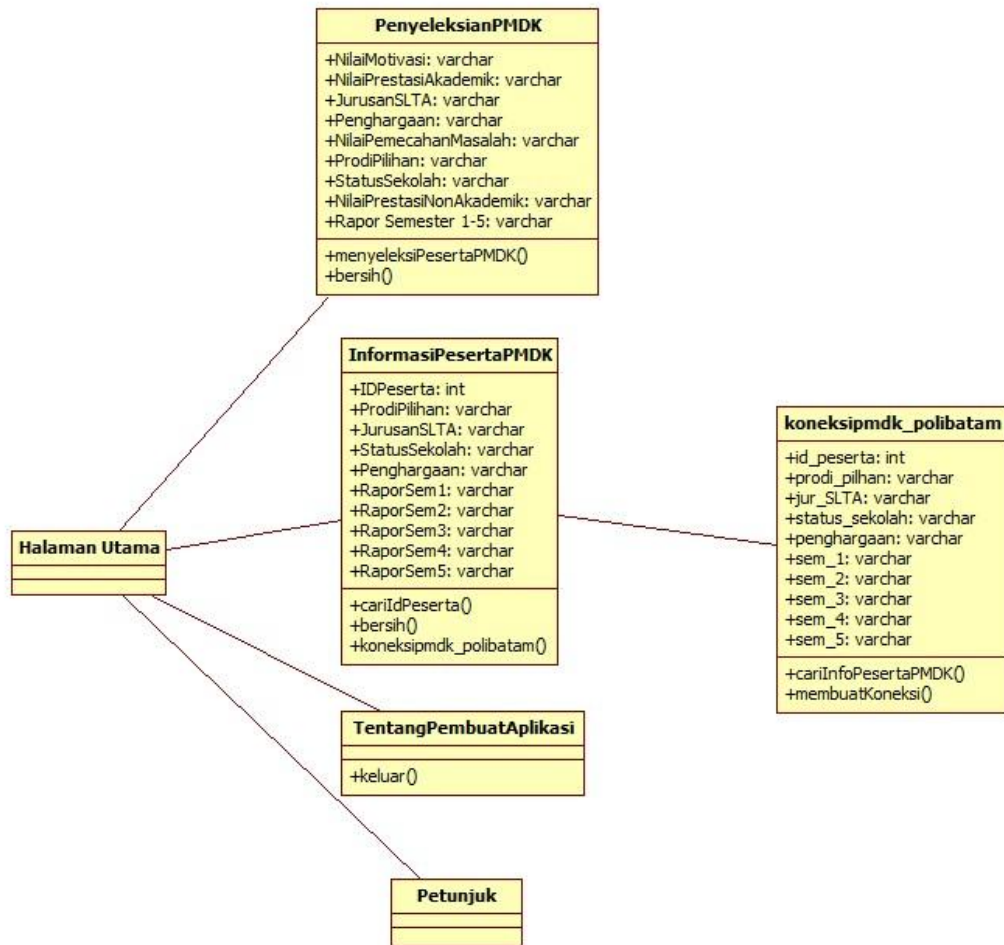
Sequence diagram ini berawal dari staff akademik memilih informasi peserta pada menu halaman utama kemudian sistem akan menampilkan form informasi PMDK. Staff akademik akan memasukkan ID peserta untuk mencari informasi peserta kemudian sistem akan mengecek apakah ID peserta yang dimasukkan sesuai dengan yang ada didalam database. Jika ID peserta tidak sesuai dengan database maka sistem akan menampilkan pesan berupa data tidak ditemukan. Jika data yang dimasukkan benar maka sistem akan menampilkan pesan bahwa data ditemukan dan menampilkan informasi peserta kedalam *field* yang ada di sistem.



Gambar 8. Sequence Diagram Mencari Informasi Peserta

III.12 Class Diagram

Class diagram mendeskripsikan jenis-jenis objek dalam sistem dan berbagai macam hubungan statis yang terdapat diantara mereka. *Class diagram* juga menunjukkan *property* dan operasi sebuah *class* dan batasan-batasan yang terdapat dalam hubungan-hubungan objek tersebut.



Gambar 9. Class Diagram

III.13 Perancangan Antarmuka

Perancangan antarmuka ini digunakan sebagai tolak ukur dari sistem aplikasi yang nantinya akan dibuat. Perancangan ini adalah gambaran umum atau contoh simple dari sistem. Adapun perancangan antarmuka yang ada adalah sebagai berikut:

III.13.1 Perancangan Antarmuka Halaman Utama



Gambar 10. Perancangan Antarmuka Halaman Utama

Berdasarkan gambar 10, dapat dilihat bahwa desain antarmuka halaman utama sistem terdiri dari *menu bar* yang terdiri dari menu dan bantuan, *menu bar* menu jika di klik memiliki 4 *submenu* yaitu penyeleksian PMDK, informasi peserta PMDK, tentang pembuat aplikasi, dan keluar. Masing-masing menu *submenu* tersebut jika diklik akan terhubung ke halaman lainnya. Jika diklik *submenu* penyeleksian PMDK, maka sistem akan menampilkan form penyeleksian PMDK, begitu juga untuk *submenu* lainnya. Sedangkan untuk *menu bar* tentang berisi *submenu* petunjuk. Jika *submenu* tersebut diklik, maka sistem akan menampilkan halaman petunjuk yang berisi tentang petunjuk penggunaan aplikasi penerimaan calon mahasiswa PMDK.

III.13.2 Perancangan Antarmuka Penyeleksian PMDK

The image shows a user interface for PMDK selection. It is divided into two main sections: 'Form Penyeleksian PMDK' and 'Output (Hasil)'. The form section contains ten dropdown menus with labels: 'Nilai Motivasi', 'Prestasi Akademik', 'Jurusan SLTA', 'Penghargaan', 'Nilai Pemecahan Masalah', 'Prodi Pilihan', 'Status Sekolah', 'Prestasi Non Akademik', and 'Rapor Semester 1-5'. Below these are two buttons labeled 'Hasil' and 'Bersih'. The output section is a text box containing the following logic: 'IF NILAI MOTIVASI = "0...0.11" ^ NILAI PRESTASI AKADEMIK = "0...0.28" v NILAI PRESTASI AKADEMIK = "1.41...1.68" v NILAI PRESTASI AKADEMIK = "1.97...2.24" THEN KETERANGAN = "TIDAK DITERIMA"'. The background is a solid teal color.

Gambar 11. Perancangan Antarmuka Penyeleksian PMDK

Berdasarkan gambar 11, dapat dilihat bahwa desain antar muka penyeleksian PMDK terdiri dari *label* form penyeleksian PMDK, *label* output/hasil, *label* nilai motivasi, *label* prestasi akademik, *label* jurusan SLTA, *label* penghargaan, *label* nilai pemecahan masalah, *label* prodi pilihan, *label* status sekolah, *label* prestasi akademik, *label* rapor semester 1-5, *combo box* nilai motivasi, *combo box* prestasi akademik, *combo box* jurusan SLTA, *combo box* nilai pemecahan masalah, *combo box* prodi pilihan, *combo box* status sekolah, *combo box* prestasi non akademik, *combo box* nilai rapor semester 1-5, *radio button* ada, *radio button* tidak ada, 2 tombol fungsi hasil dan bersih dan *text area* output/hasil. Apabila pengguna mengklik tombol hasil maka sistem akan menampilkan hasil ke *text area* hasil. Apabila pengguna menekan tombol bersih maka sistem akan menghapus isi yang ada dalam *combo box* dan *radio button*.

III.13.3 Perancangan Antarmuka Informasi Peserta PMDK

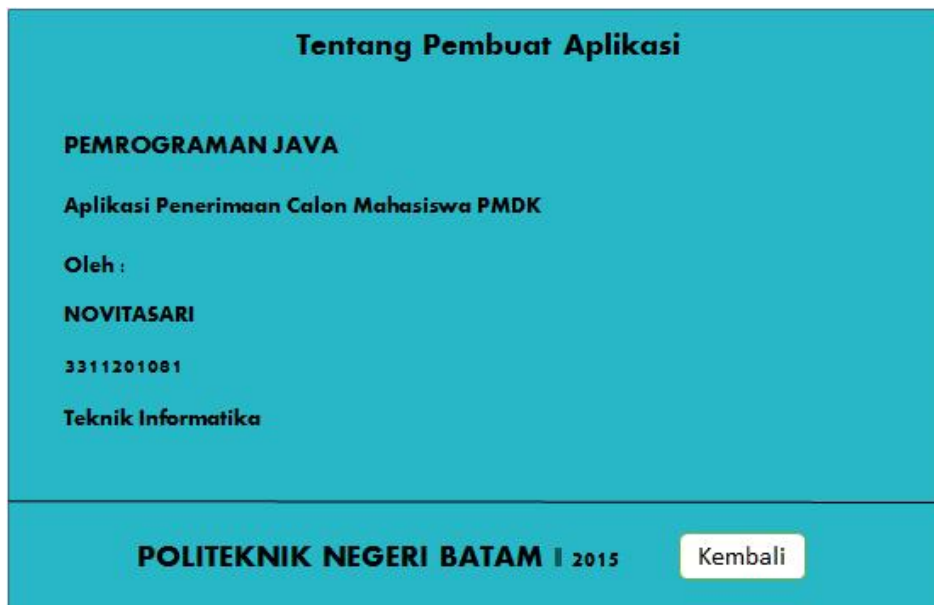
The image shows a web form titled "Form Informasi Peserta" on a teal background. The form consists of the following elements from top to bottom:

- ID Peserta:** A text input field with two buttons to its right: "Cari" and "Bersih".
- Prodi Pilihan:** A wide text input field.
- Jurusan SLTA:** A text input field.
- Status Sekolah:** A text input field.
- Penghargaan:** A text input field.
- Rapor Sem 1:** A text input field.
- Rapor Sem 2:** A text input field.
- Rapor Sem 3:** A text input field.
- Rapor Sem 4:** A text input field.
- Rapor Sem 5:** A text input field.

Gambar 12. Perancangan Antarmuka Informasi Peserta PMDK

Berdasarkan 12, dapat dilihat bahwa desain antar muka informasi peserta PMDK terdiri dari *label* ID peserta, *label* prodi pilihan, *label* jurusan SLTA, *label* status sekolah, *label* penghargaan, *label* rapor sem 1, *label* rapor sem 2, *label* rapor sem 3, *label* rapor sem 4, *label* rapor sem 5, dan 2 tombol fungsi cari dan bersih. Tombol cari digunakan untuk mencari informasi peserta PMDK berkaitan dengan ID peserta yang di input pengguna. Apabila pengguna mengklik tombol cari, sistem akan menampilkan informasi peserta berdasarkan ID peserta yang diinput ke *text field* prodi pilihan sampai *text field* nilai rapor sem 5. Apabila pengguna mengklik tombol bersih maka sistem akan menghapus seluruh isi yang ada di *text field*.

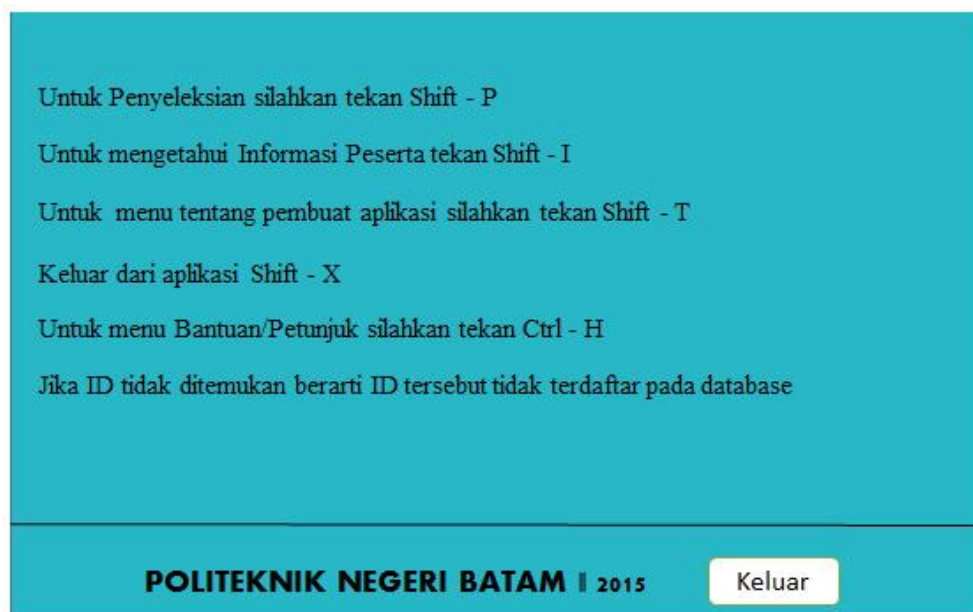
III.13.4 Perancangan Antarmuka Tentang Pembuat Aplikasi



Gambar 13. Perancangan Antarmuka Tentang Pembuat Aplikasi

Berdasarkan gambar 13, dapat dilihat bahwa desain antar muka tentang pembuat aplikasi terdiri dari *label* tentang pembuat aplikasi, *label* pemrograman, *label* aplikasi penerimaan calon mahasiswa PMDK, *label* oleh, *label* novi, *label* nim, *label* jurusan, *label* politeknik, dan tombol kembali.

III.13.5 Perancangan Antarmuka Petunjuk



Gambar 14. Perancangan Antarmuka Petunjuk

BAB IV

PROSES KLASIFIKASI

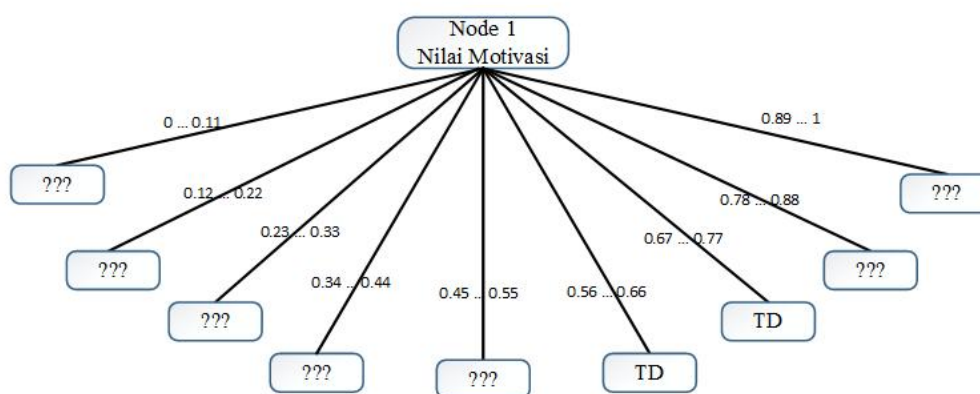
Bab ini berisi proses perhitungan *entropy* dan *gain* terhadap *data training* untuk menentukan *node* suatu *decision tree*. Perhitungan *gain* dan *entropy* dilakukan dengan menggunakan *Microsoft Excel* yang telah mengandung rumus perhitungan *gain* dan *entropy*.

IV.1 Perhitungan Entropy dan Gain Data Training

Dengan menggunakan rumus perhitungan *entropy* dan *gain*, maka diperoleh nilai *gain* tertinggi yang menjadi *root node*.

IV.1.1 Penentuan Root Node

Untuk mendapatkan *root node* suatu *decision tree*, dilakukan perhitungan *entropy* dan *gain* terhadap *data training*. Dari perhitungan tersebut maka dihasilkan nilai *gain* tertinggi yang akan menjadi *root node*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Lampiran Tabel A.1.1. Berdasarkan Tabel A.1.1 dapat diketahui bahwa atribut dengan *gain* tertinggi adalah **nilai motivasi**. Sehingga **nilai motivasi** menjadi *root node*. Ada 9 *instance* nilai dari **nilai motivasi**, yaitu **0...0.11**, **0.12...0.22**, **0.23...0.33**, **0.34...0.44**, **0.45...0.55**, **0.56...0.66**, **0.67...0.77**, **0.78...0.88**, dan **0.89...1**. Pada *instance* **0.56...0.66** dan **0.78...0.88** sudah mengklasifikasikan menjadi **TIDAK DITERIMA**. Untuk *instance* **0...0.11**, **0.12...0.22**, **0.23...0.33**, **0.34...0.44**, **0.45...0.55**, **0.78...0.88**, dan **0.89...1** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

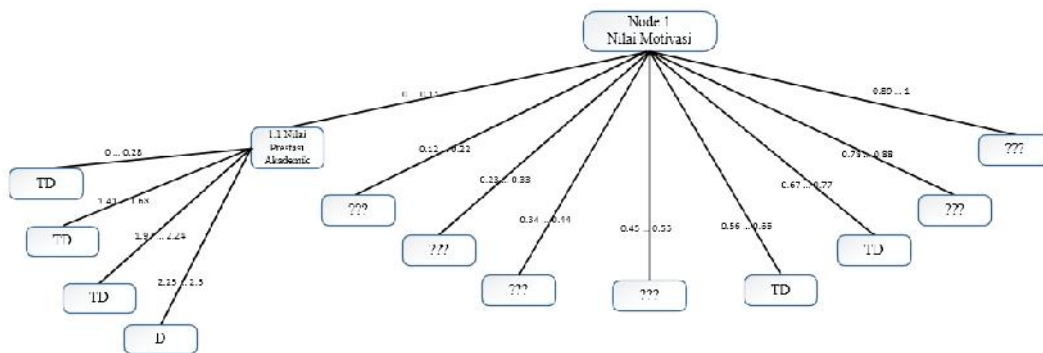


Gambar 15. Decision Tree Hasil Perhitungan Node 1

IV.1.2 Penentuan *Node* 1.1

Setelah mendapatkan *root node* (nilai motivasi) dari perhitungan *entropy* dan *gain*, untuk mendapatkan *node* selanjutnya dilakukan pula perhitungan *entropy* dan *gain*. Nilai *gain* tertinggi akan menjadi *node* 1.1. Hasil perhitungan *entropy* dan *gain*-nya dapat dilihat pada Tabel A.1.2.

Dari hasil Tabel A.1.2, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 4 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0...0.28**, **1.41...1.68**, **1.97...2.24**, dan **2.25...2.5**. Nilai *instance* **0...0.28**, **1.41...1.68**, **1.97...2.24** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **2.25...2.5** mengklasifikasikan kasus menjadi **DITERIMA**.

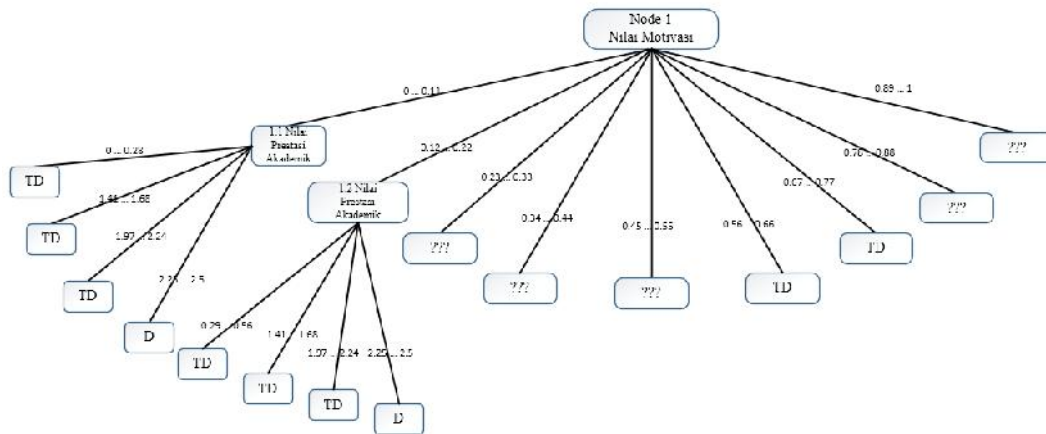


Gambar 16. *Decision Tree* Hasil Perhitungan *Node* 1.1

IV.1.3 Penentuan *Node* 1.2

Untuk mendapatkan *node* selanjutnya dilakukan pula perhitungan *entropy* dan *gain*. Nilai *gain* tertinggi akan menjadi *node* 1.2. Hasil perhitungan *entropy* dan *gain*-nya dapat dilihat pada Tabel A.1.3.

Dari hasil Tabel A.1.3, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 4 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0.29...0.56**, **1.41...1.68**, **1.97...2.24**, dan **2.25...2.5**. Nilai *instance* **0.29...0.56**, **1.41...1.68**, **1.97...2.24** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **2.25...2.5** mengklasifikasikan kasus menjadi **DITERIMA**.

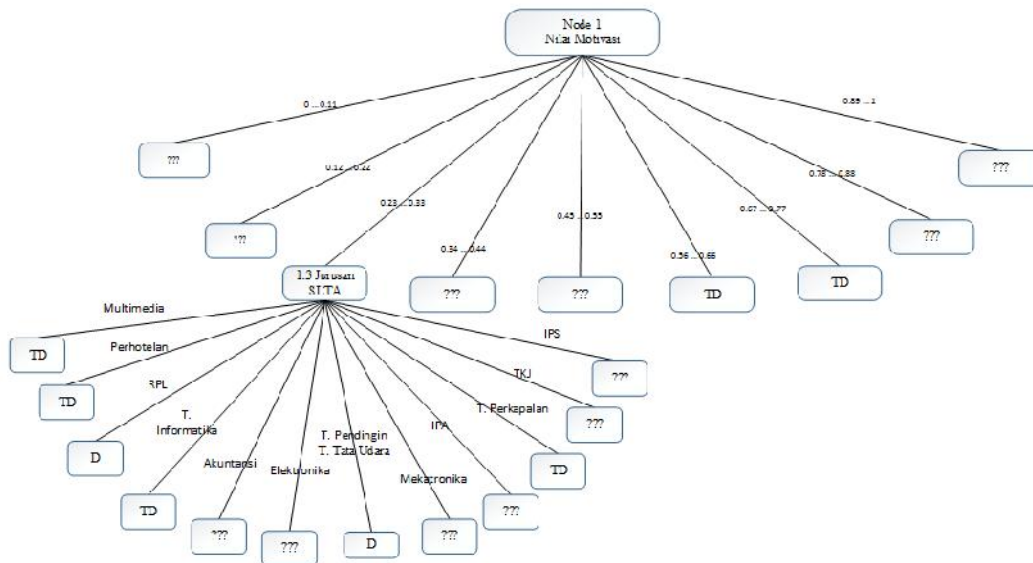


Gambar 17. Decision Tree Hasil Perhitungan Node 1.2

IV.1.4 Penentuan Node 1.3

Untuk mendapatkan *node* selanjutnya dilakukan pula perhitungan *entropy* dan *gain*. Nilai *gain* tertinggi akan menjadi *node* 1.3. Hasil perhitungan *entropy* dan *gain*-nya dapat dilihat pada Tabel A.1.4.

Dari hasil Tabel A.1.4, diketahui bahwa atribut dengan *gain* tertinggi adalah **JURUSAN SLTA**. Ada 12 *instance* dari **JURUSAN SLTA**, yaitu **Akuntansi, Elektronika, IPA, IPS, Mekatronika, Multimedia, Perhotelan, Rekayasa Perangkat Lunak, Teknik Informatika, Teknik Komputer dan Jaringan, Teknik Pendingin dan Tata Udara, dan Teknik Perkapalan**. Nilai *instance* **Multimedia, Perhotelan, Teknik Informatika, dan Teknik Perkapalan** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **Teknik Pendingin dan Tata Udara dan Teknik Perangkat Lunak** mengklasifikasikan kasus menjadi **DITERIMA**. Untuk *instance* **Akuntansi, Elektronika, IPA, IPS, Mekatronika, dan Teknik Komputer dan Jaringan** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

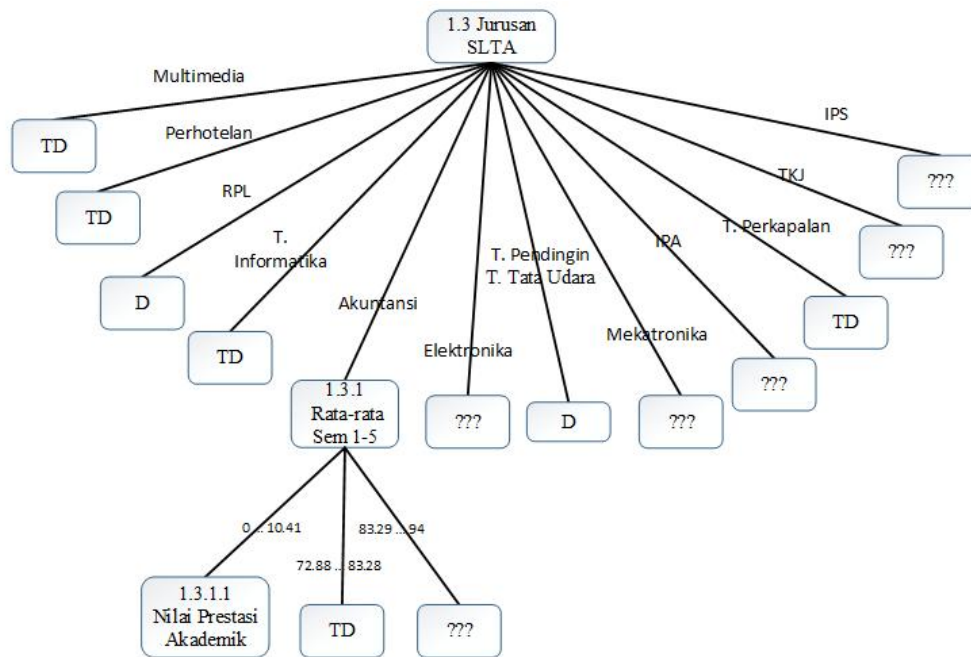


Gambar 18. Decision Tree Hasil Perhitungan Node 1.3

IV.1.5 Penentuan Node 1.3.1

Untuk menentukan *node* 1.3.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.5.

Dari hasil Tabel A.1.5, diketahui bahwa atribut dengan *gain* tertinggi adalah **RATA-RATA SEMESTER 1-5**. Ada 3 *instance* dari **RATA-RATA SEMESTER 1-5**, yaitu **0...10.41**, **72.88...83.28**, dan **83.29...94**. Nilai *instance* **72.88...83.28** dan **83.29...94** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **0...10.41**, dan **83.29...94** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

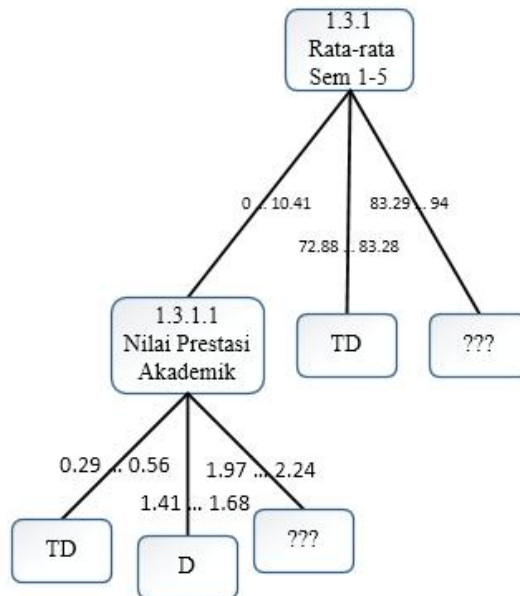


Gambar 19. *Decision Tree* Hasil Perhitungan *Node* 1.3.1

IV.1.6 Penentuan *Node* 1.3.1.1

Untuk menentukan *node* 1.3.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.6.

Dari hasil Tabel A.1.6, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 3 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0.29...0.56**, **1.41...1.68**, dan **1.97...2.24**. Nilai *instance* **0.29...0.56** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **1.41...1.68** mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **1.97...2.24** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

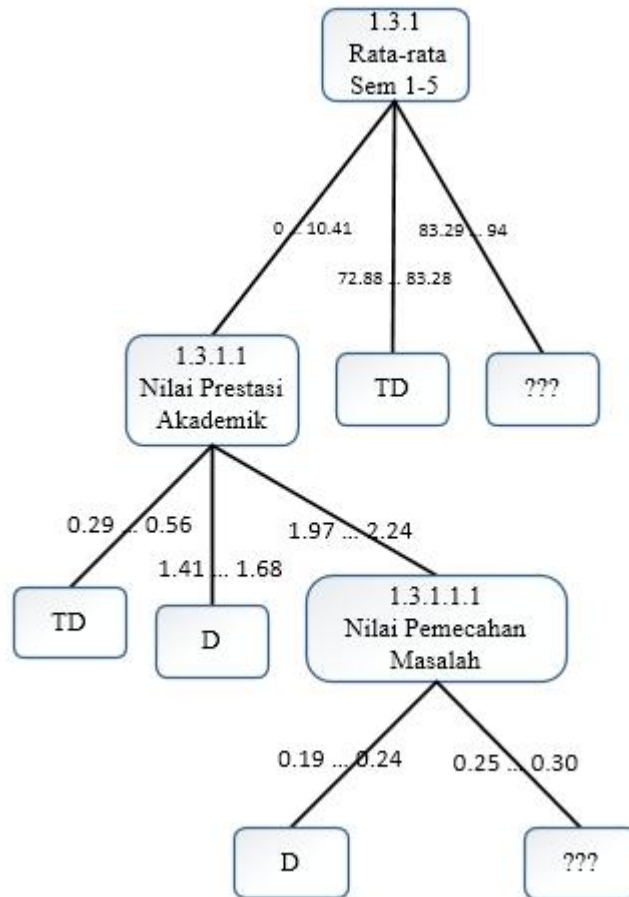


Gambar 20. Decision Tree Hasil Perhitungan Node 1.3.1.1

IV.1.7 Penentuan Node 1.3.1.1.1

Untuk menentukan *node* 1.3.1.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.7.

Dari hasil Tabel A.1.7, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PEMECAHAN MASALAH**. Ada 2 *instance* dari **NILAI PEMECAHAN MASALAH**, yaitu **0.19...0.24** dan **0.25...0.30**. Nilai *instance* **0.19...0.24** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **0.25...0.30** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

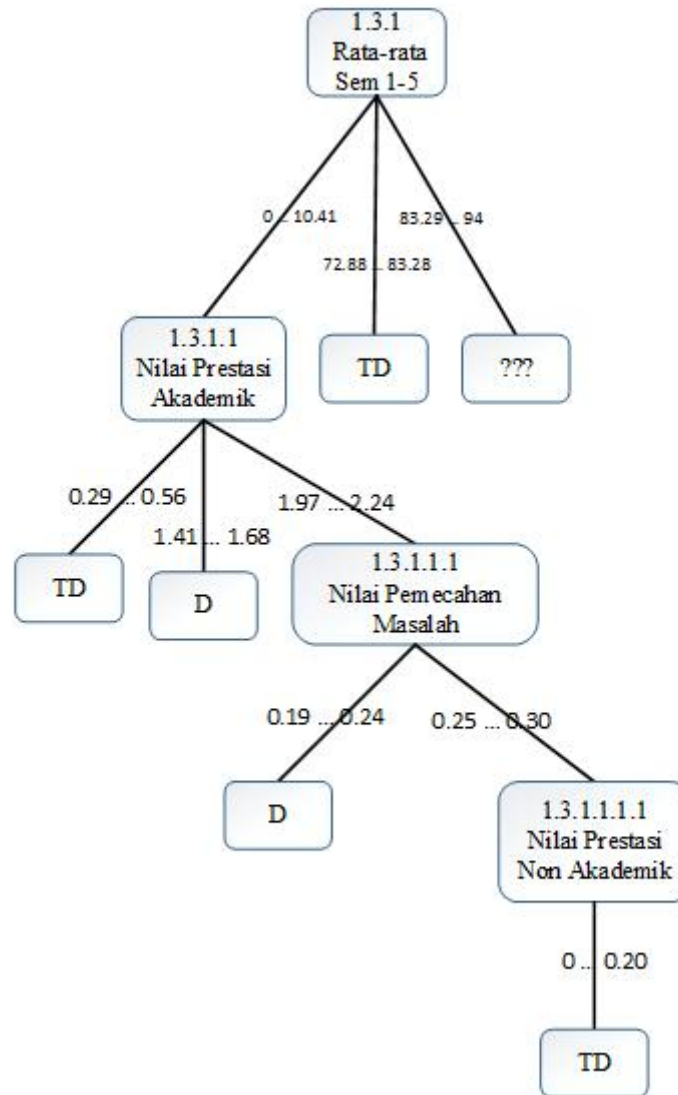


Gambar 21. Decision Tree Hasil Perhitungan Node 1.3.1.1.1

IV.1.8 Penentuan Node 1.3.1.1.1.1

Untuk menentukan *node* 1.3.1.1.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.8.

Dari hasil Tabel A.1.8, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 1 *instance* dari **NILAI PEMECAHAN MASALAH**, yaitu **0...0.20**. Nilai *instance* **0...0.20** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**.

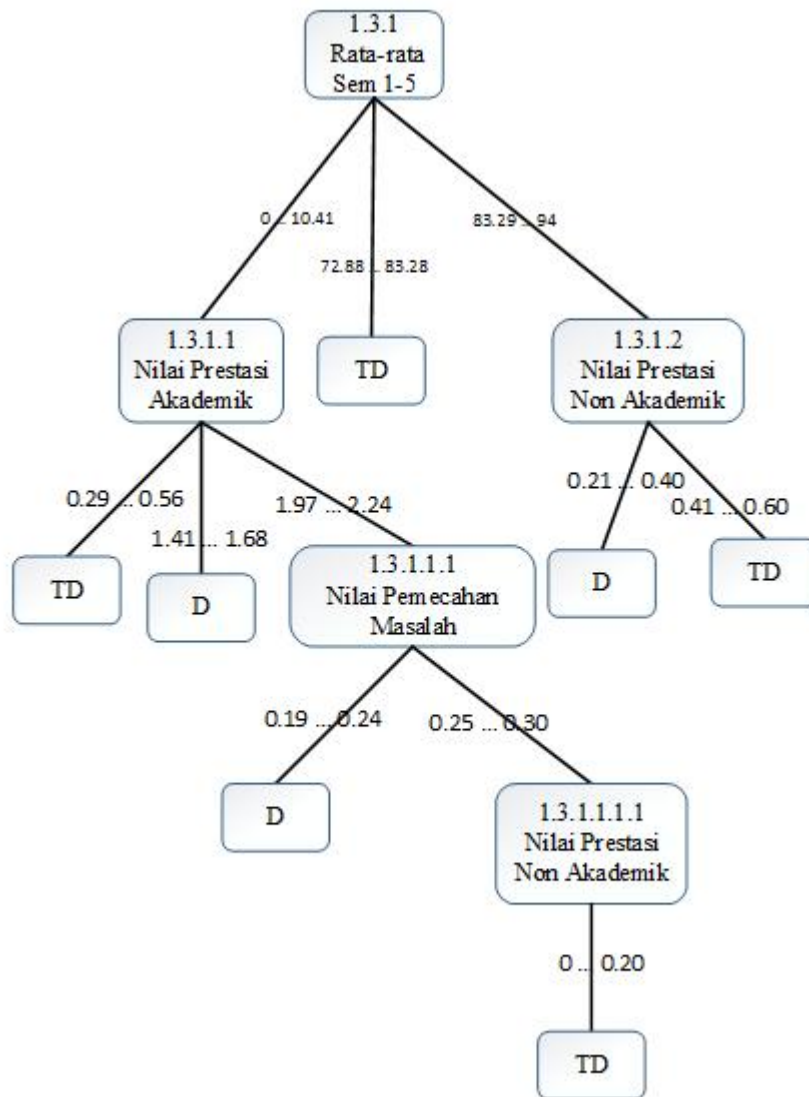


Gambar 22. Decision Tree Hasil Perhitungan Node 1.3.1.1.1.1

IV.1.9 Penentuan Node 1.3.1.2

Untuk menentukan *node* 1.3.1.2 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.9

Dari hasil Tabel A.1.9, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 2 *instance* dari **NILAI PRESTASI NON AKADEMIK**, yaitu **0.21...0.40** dan **0.41...0.60**. Nilai *instance* **0.41...0.60** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **0.21...0.40** mengklasifikasikan kasus menjadi **DITERIMA**.

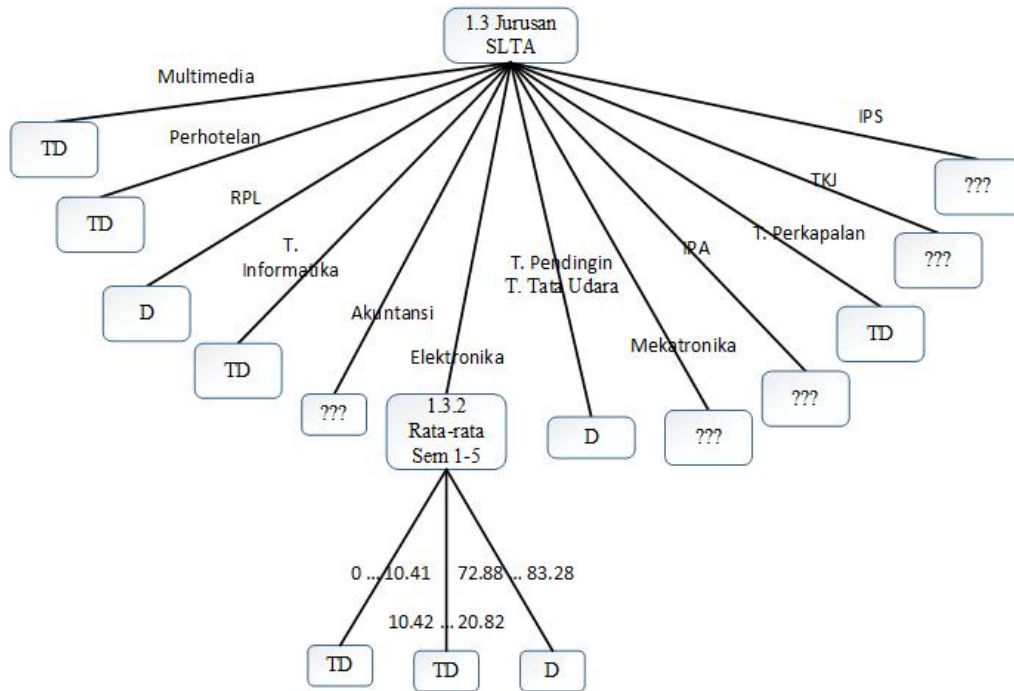


Gambar 23. Decision Tree Hasil Perhitungan Node 1.3.1.2

IV.1.10 Penentuan Node 1.3.2

Untuk menentukan *node* 1.3.2 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.10.

Dari hasil Tabel A.1.10, diketahui bahwa atribut dengan *gain* tertinggi adalah **RATA-RATA SEMESTER 1-5**. Ada 3 *instance* dari **RATA-RATA SEMESTER 1-5**, yaitu **0...10.41**, **10.42...20.82**, dan **72.88...83.28**. Nilai *instance* **0...10.41** dan **10.42...20.82** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **72.88...83.28** mengklasifikasikan kasus menjadi **DITERIMA**.

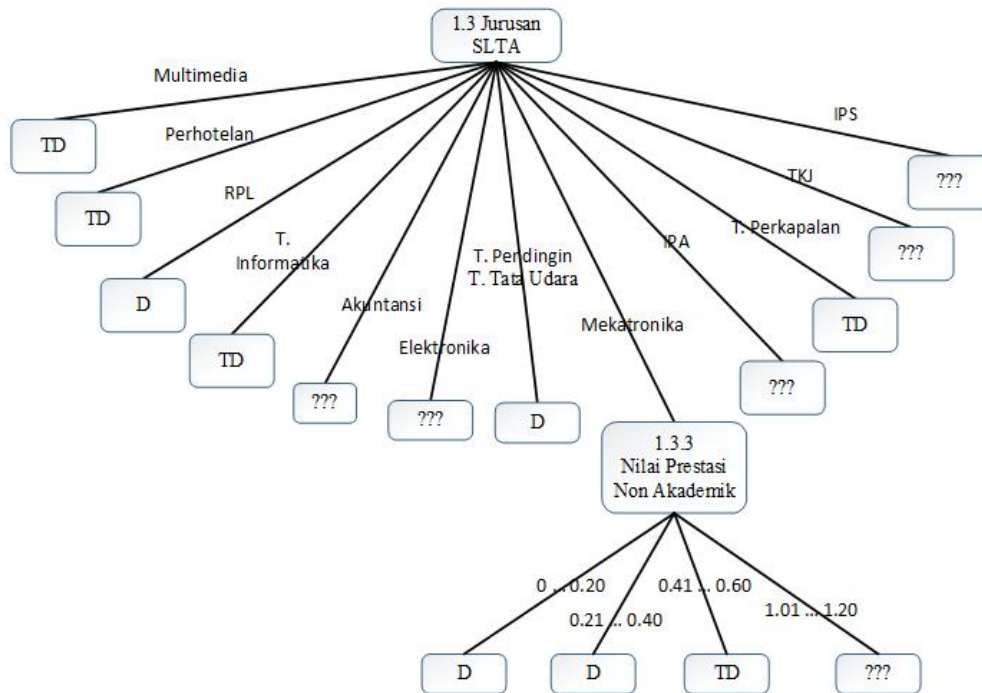


Gambar 24. Decision Tree Hasil Perhitungan Node 1.3.2

IV.1.11 Penentuan Node 1.3.3

Untuk menentukan *node* 1.3.3 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.11.

Dari hasil Tabel A.1.11, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 4 *instance* dari **NILAI PRESTASI NON AKADEMIK**, yaitu **0...0.20**, **0.21...0.40**, **0.41...0.60**, dan **1.01...1.20**. Nilai *instance* **0.41...0.60** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Dan *instance* **0...0.20** dan **0.21...0.40** mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan **1.01...1.20** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

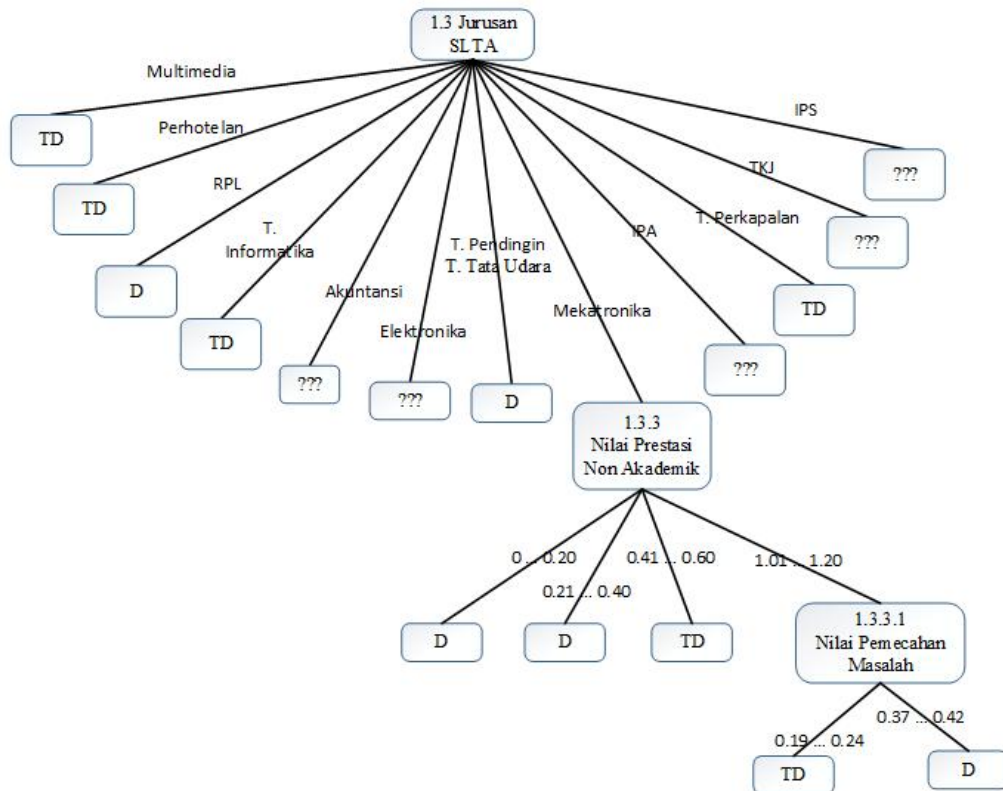


Gambar 25. Decision Tree Hasil Perhitungan Node 1.3.3

IV.1.12 Penentuan Node 1.3.3.1

Untuk menentukan node 1.3.3.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.12.

Dari hasil Tabel A.1.12, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PEMECAHAN MASALAH**. Ada 2 *instance* dari **NILAI PEMECAHAN MASALAH**, yaitu **0.19...0.24** dan **0.37...0.42**. Nilai *instance* **0.19...0.24** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Dan *instance* **0.37...0.42** mengklasifikasikan kasus menjadi **DITERIMA**.

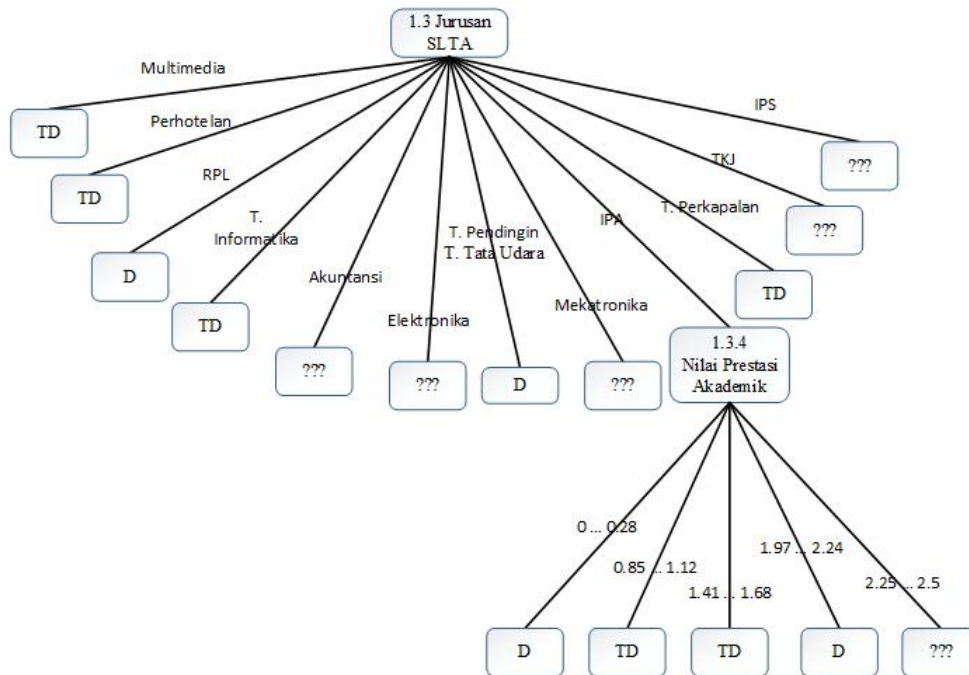


Gambar 26. Decision Tree Hasil Perhitungan Node 1.3.3.1

IV.1.13 Penentuan Node 1.3.4

Untuk menentukan *node* 1.3.4 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.13.

Dari hasil Tabel A.1.13, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 5 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0...0.28**, **0.85...1.12**, **1.41...1.68**, **1.97...2.24** dan **2.25...2.5**. Nilai *instance* **1.41...1.68**, dan **1.97...2.24** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **0...0.28** dan **1.41...1.68** mengklasifikasikan kasus menjadi **DITERIMA**.

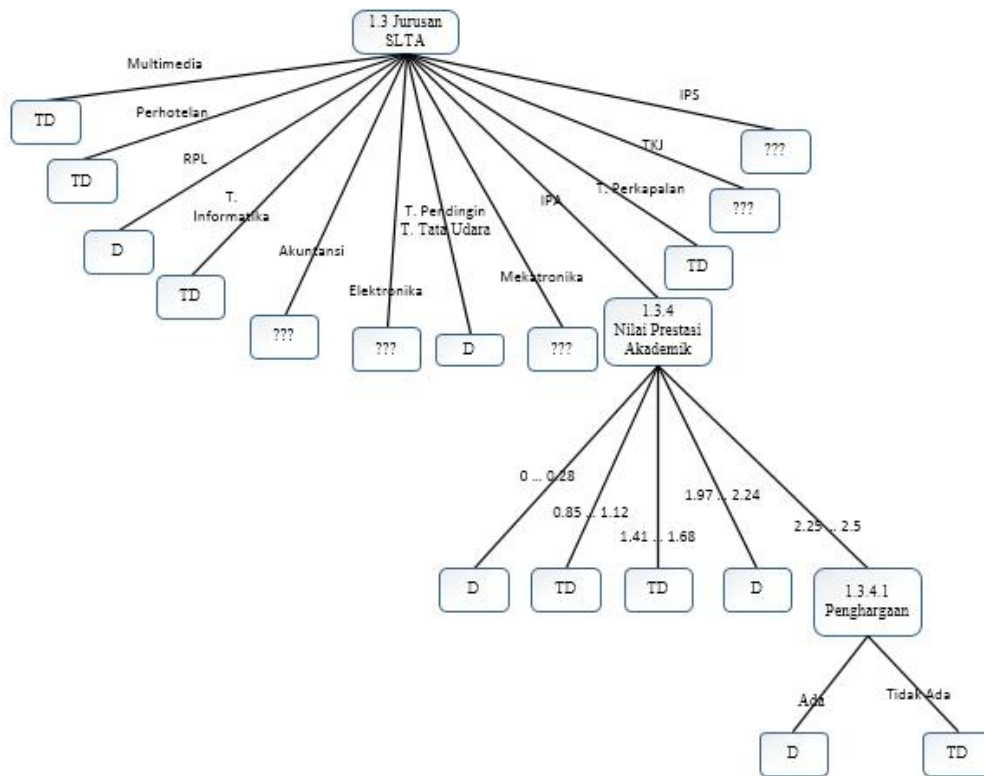


Gambar 27. Decision Tree Hasil Perhitungan Node 1.3.4

IV.1.14 Penentuan Node 1.3.4.1

Untuk menentukan *node* 1.3.4.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.14.

Dari hasil Tabel A.1.14, diketahui bahwa atribut dengan *gain* tertinggi adalah **PENGHARGAAN**. Ada 2 *instance* dari **PENGHARGAAN**, yaitu **Ada** dan **Tidak Ada**. Nilai *instance* **Tidak Ada** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **Ada** mengklasifikasikan kasus menjadi **DITERIMA**.

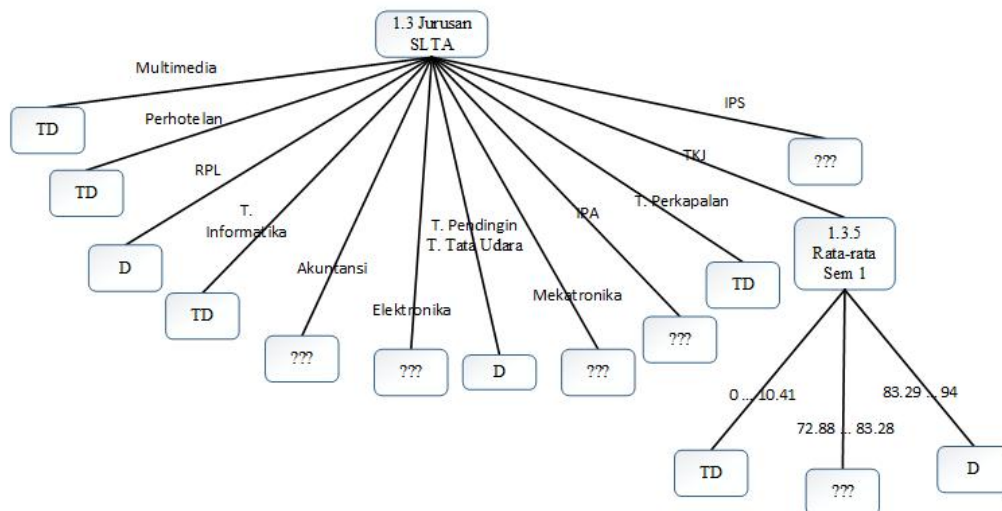


Gambar 28. Decision Tree Hasil Perhitungan Node 1.3.4.1

IV.1.15 Penentuan Node 1.3.5

Untuk menentukan node 1.3.5 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.15.

Dari hasil Tabel A.1.15, diketahui bahwa atribut dengan *gain* tertinggi adalah **RATA-RATA SEMESTER 1-5**. Ada 3 *instance* dari **RATA-RATA SEMESTER 1-5**, yaitu **0...10.41**, **72.88...83.28** dan **83.29...94**. Nilai *instance* **0...10.41** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **83.29...94** mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **72.88...83.28** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

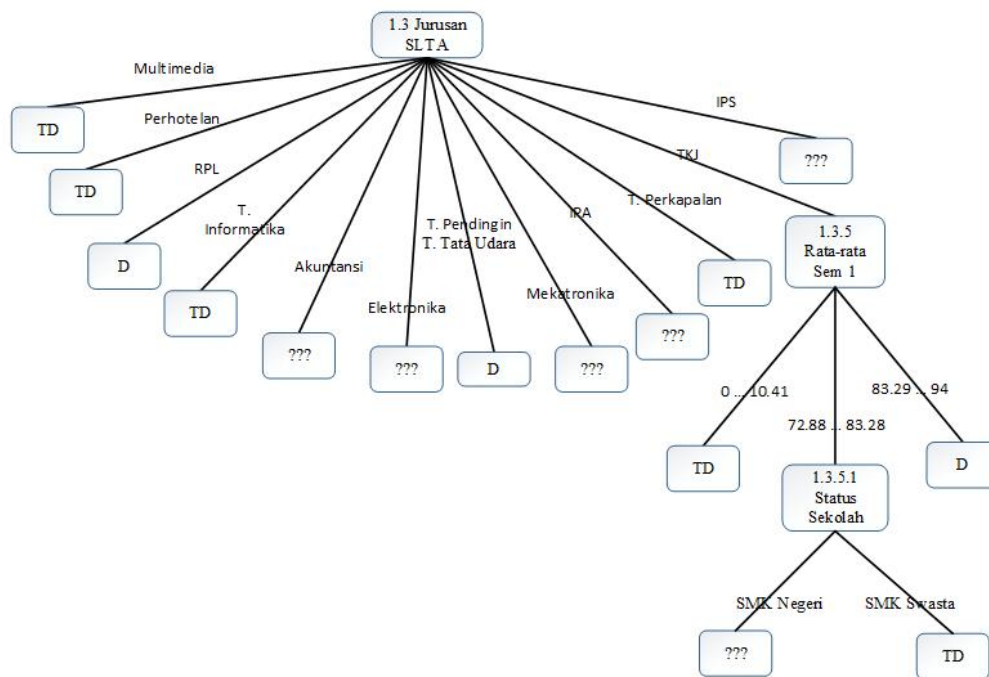


Gambar 29. Decision Tree Hasil Perhitungan Node 1.3.5

IV.1.16 Penentuan Node 1.3.5.1

Untuk menentukan *node* 1.3.5.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.16.

Dari hasil Tabel A.1.16, diketahui bahwa atribut dengan *gain* tertinggi adalah **STATUS SEKOLAH**. Ada 3 *instance* dari **STATUS SEKOLAH**, yaitu **SMK NEGERI** dan **SMK SWASTA**. Nilai *instance* **SMK SWASTA** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **SMK NEGERI** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

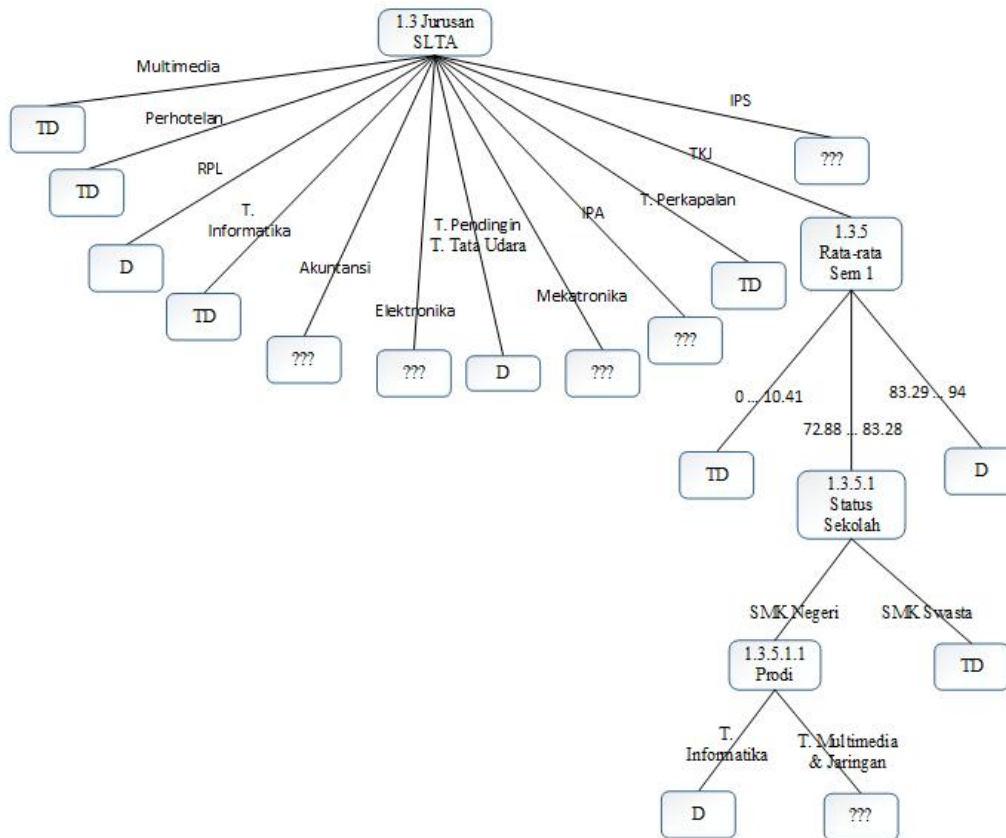


Gambar 30. *Decision Tree* Hasil Perhitungan *Node* 1.3.5.1

IV.1.17 Penentuan *Node* 1.3.5.1.1

Untuk menentukan *node* 1.3.5.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.17.

Dari hasil Tabel A.1.17, diketahui bahwa atribut dengan *gain* tertinggi adalah **PRODI**. Ada 2 *instance* dari **PRODI**, yaitu **TEKNIK INFORMATIKA** dan **TEKNIK MULTIMEDIA DAN JARINGAN**. Nilai *instance* **TEKNIK INFORMATIKA** sudah mengklasifikasikan kasus menjadi **DITERIMA** dan *instance* **TEKNIK MULTIMEDIA DAN JARINGAN** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

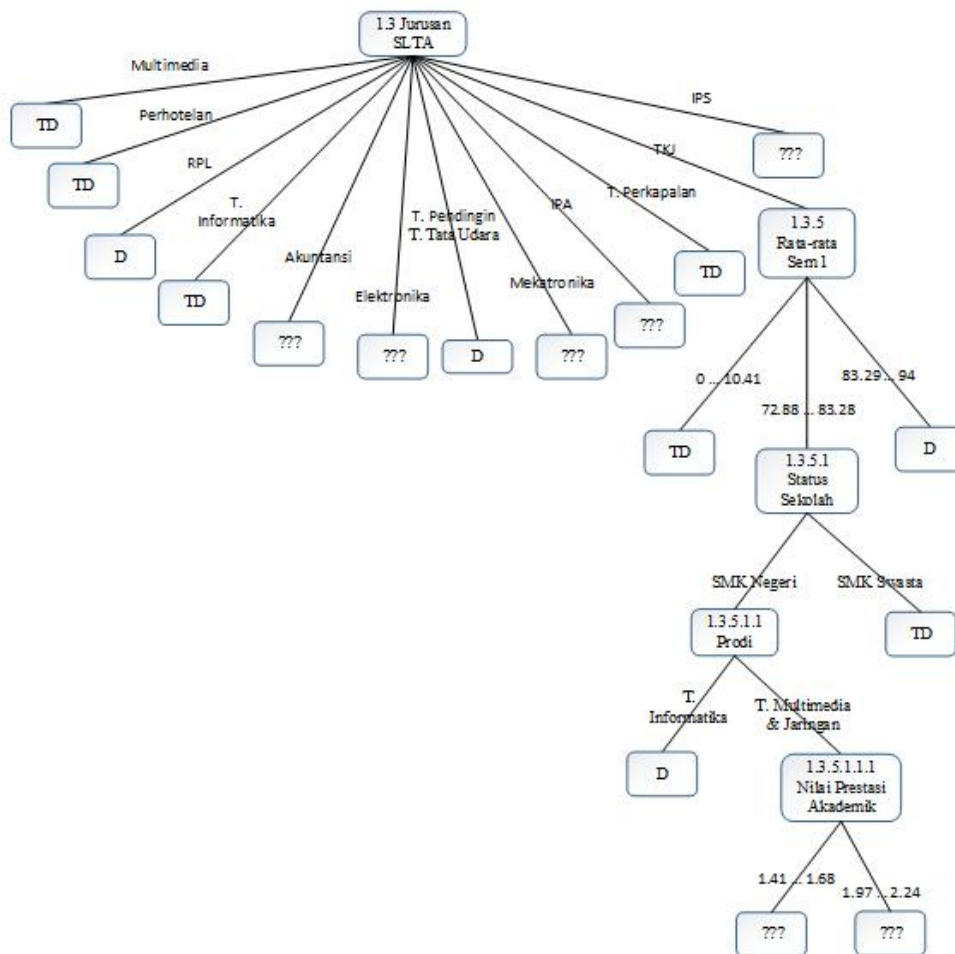


Gambar 31. Decision Tree Hasil Perhitungan Node 1.3.5.1.1

IV.1.18 Penentuan Node 1.3.5.1.1.1

Untuk menentukan node 1.3.5.1.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.18.

Dari hasil Tabel A.1.18, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 2 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **1.41...1.68** dan **1.97...2.24**. Nilai *instance* **1.41...1.68** dan **1.97...2.24** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

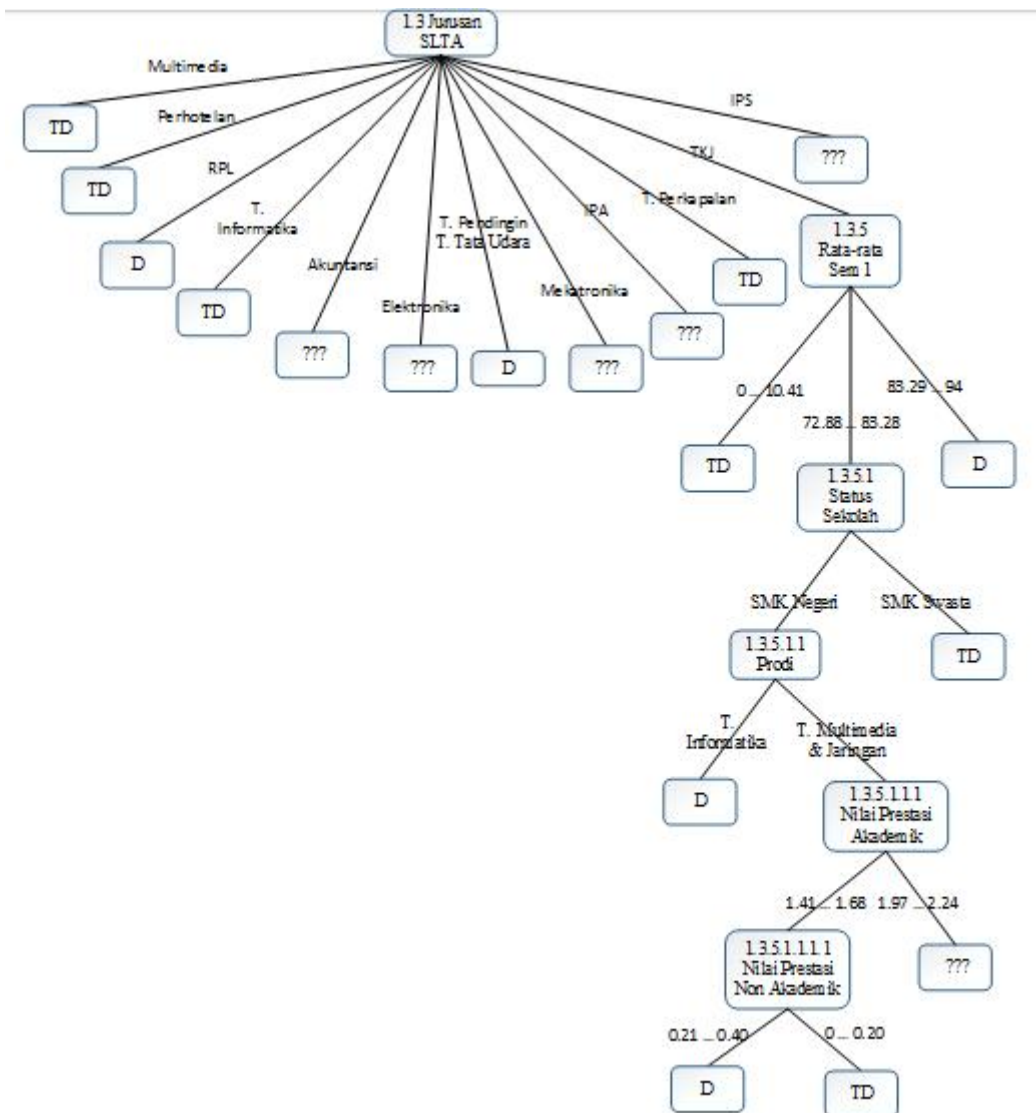


Gambar 32. Decision Tree Hasil Perhitungan Node 1.3.5.1.1.1

IV.1.19 Penentuan Node 1.3.5.1.1.1.1

Untuk menentukan *node* 1.3.5.1.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.19.

Dari hasil Tabel A.1.19, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 2 *instance* dari **NILAI PRESTASI NON AKADEMIK**, yaitu **0.21...0.40** dan **0...0.20**. Nilai *instance* **0...0.20** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **0.21...0.40** mengklasifikasikan kasus menjadi **DITERIMA**.

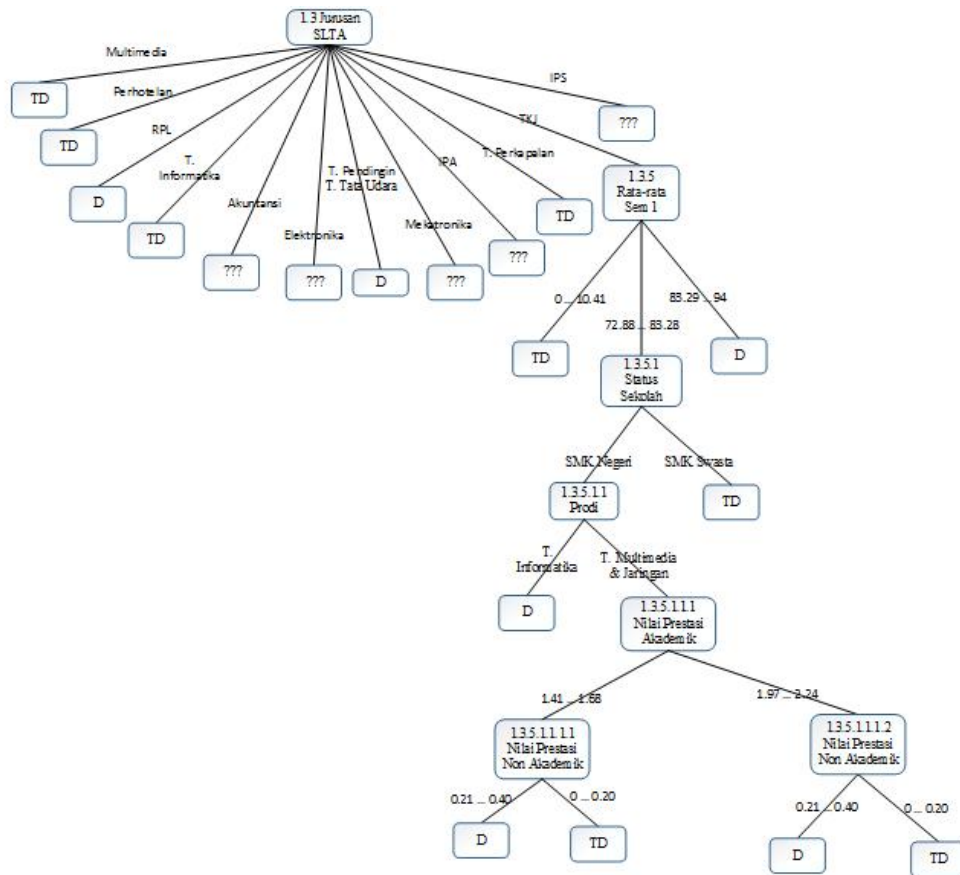


Gambar 33. Decision Tree Hasil Perhitungan Node 1.3.5.1.1.1.1

IV.1.20 Penentuan Node 1.3.5.1.1.1.2

Untuk menentukan *node* 1.3.5.1.1.1.2 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.20.

Dari hasil Tabel A.1.20, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 2 *instance* dari **NILAI PRESTASI NON AKADEMIK**, yaitu **0.21...0.40** dan **0...0.20**. Nilai *instance* **0...0.20** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **0.21...0.40** mengklasifikasikan kasus menjadi **DITERIMA**.

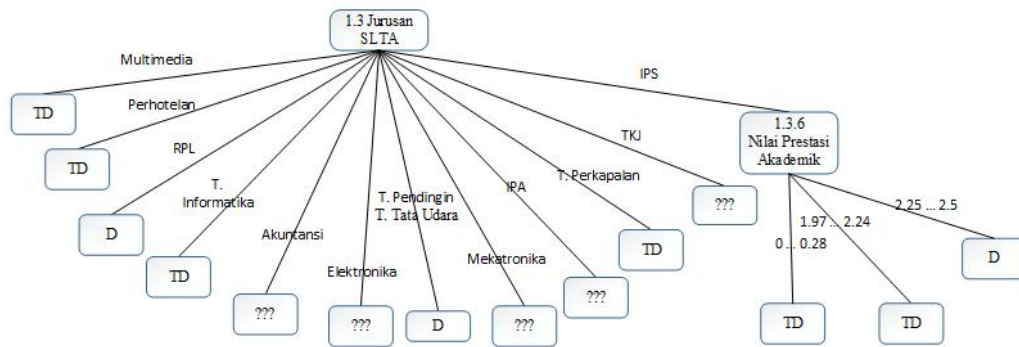


Gambar 34. Decision Tree Hasil Perhitungan Node 1.3.5.1.1.2

IV.1.21 Penentuan Node 1.3.6

Untuk menentukan *node* 1.3.6 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.21.

Dari hasil Tabel A.1.21, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 3 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0...0.28**, **1.97...2.24**, dan **2.25...2.5**. Nilai *instance* **0...0.28** dan **1.97...2.24** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**, dan *instance* **2.25...2.5** mengklasifikasikan kasus menjadi **DITERIMA**.

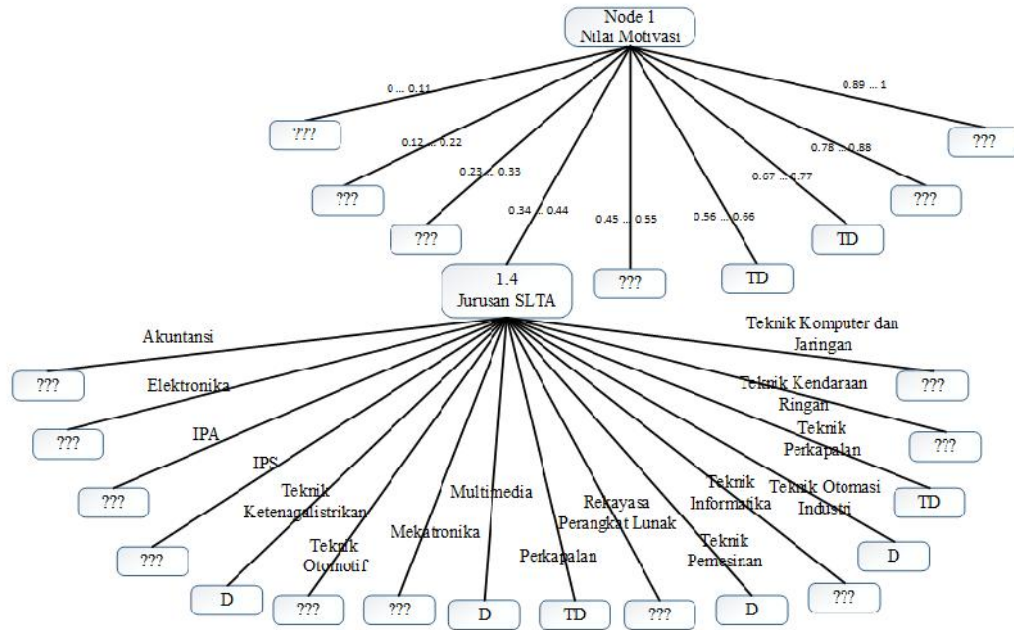


Gambar 35. Decision Tree Hasil Perhitungan Node 1.3.6

IV.1.22 Penentuan Node 1.4

Untuk mendapatkan *node* selanjutnya dilakukan pula perhitungan *entropy* dan *gain*. Nilai *gain* tertinggi akan menjadi *node* 1.4. Hasil perhitungan *entropy* dan *gain*-nya dapat dilihat pada Tabel A.1.22.

Dari hasil Tabel A.1.22, diketahui bahwa atribut dengan *gain* tertinggi adalah **JURUSAN SLTA**. Ada 16 *instance* dari **JURUSAN SLTA**, yaitu **Akuntansi, Elektronika, IPA, IPS, Teknik Ketenagalistrikan, Teknik Otomotif, Mekatronika, Multimedia, Perkapalan, Rekayasa Perangkat Lunak, Teknik Pemesinan, Teknik Informatika, Teknik Otomasi Industri, Teknik Perkapalan, Teknik Kendaraan Ringan, dan Teknik Komputer dan Jaringan**. Nilai *instance* **Perkapalan** dan **Teknik Perkapalan** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **Teknik Ketenagalistrikan, Multimedia, Teknik Pemesinan, dan Teknik Otomasi Industri** mengklasifikasikan kasus menjadi **DITERIMA**. Untuk *instance* **Akuntansi, Elektronika, IPA, IPS, Mekatronika, Teknik Otomotif, Mekatronika, Rekayasa Perangkat Lunak, Teknik Informatika, Teknik Kendaraan Ringan, dan Teknik Komputer dan Jaringan** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

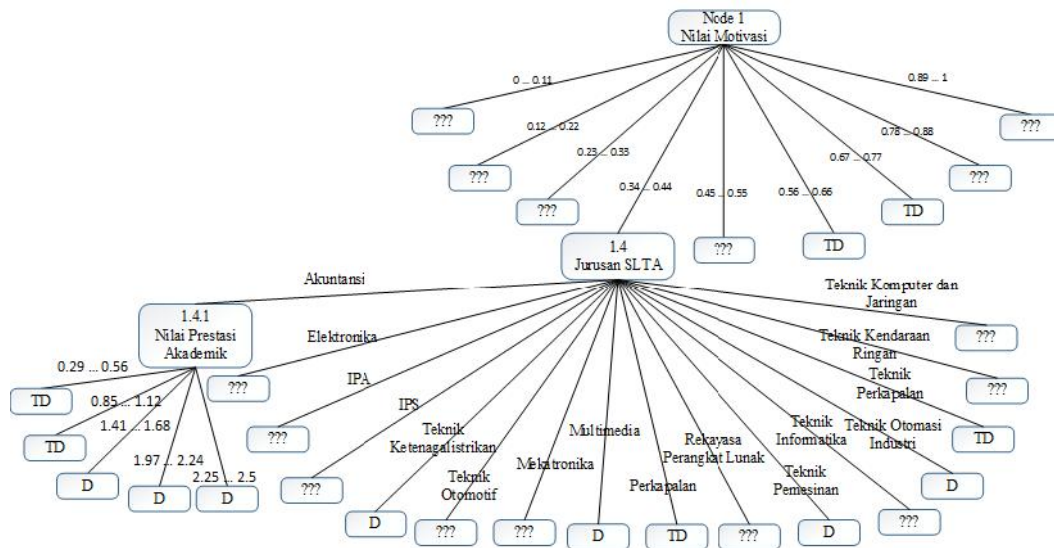


Gambar 36. Decision Tree Hasil Perhitungan Node 1.4

IV.1.23 Penentuan Node 1.4.1

Untuk menentukan *node* 1.4.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.23.

Dari hasil Tabel A.1.23, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 5 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0.29...0.56**, **0.85...1.12**, **1.41...1.68**, **1.97...2.24** dan **2.25...2.5**. Nilai *instance* **0.29...0.56**, dan **0.85...1.12** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **1.41...1.68**, **1.97...2.24** dan **2.25...2.5** mengklasifikasikan kasus menjadi **DITERIMA**.

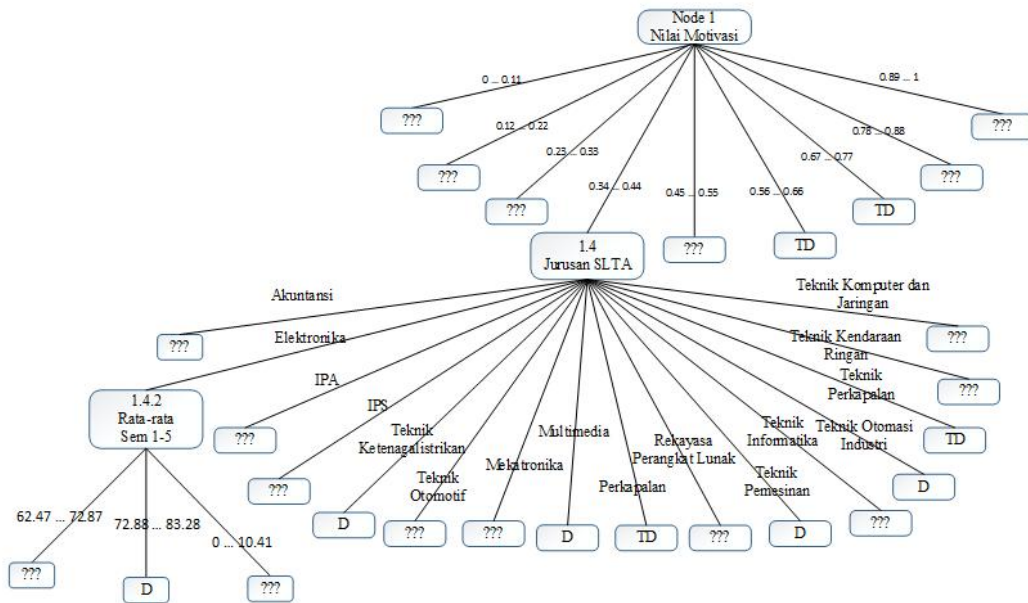


Gambar 37. Decision Tree Hasil Perhitungan Node 1.4.1

IV.1.24 Penentuan Node 1.4.2

Untuk menentukan *node* 1.4.2 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.24.

Dari hasil Tabel A.1.24, diketahui bahwa atribut dengan *gain* tertinggi adalah **RATA-RATA SEMESTER 1-5**. Ada 3 *instance* dari **RATA-RATA SEMESTER 1-5**, yaitu **62.47...72.87**, **72.88...83.28**, dan **0...10.41**. Nilai *instance* **72.88...83.28** sudah mengklasifikasikan kasus menjadi **DITERIMA** dan *instance* **62.47...72.87** dan **0...10.41** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

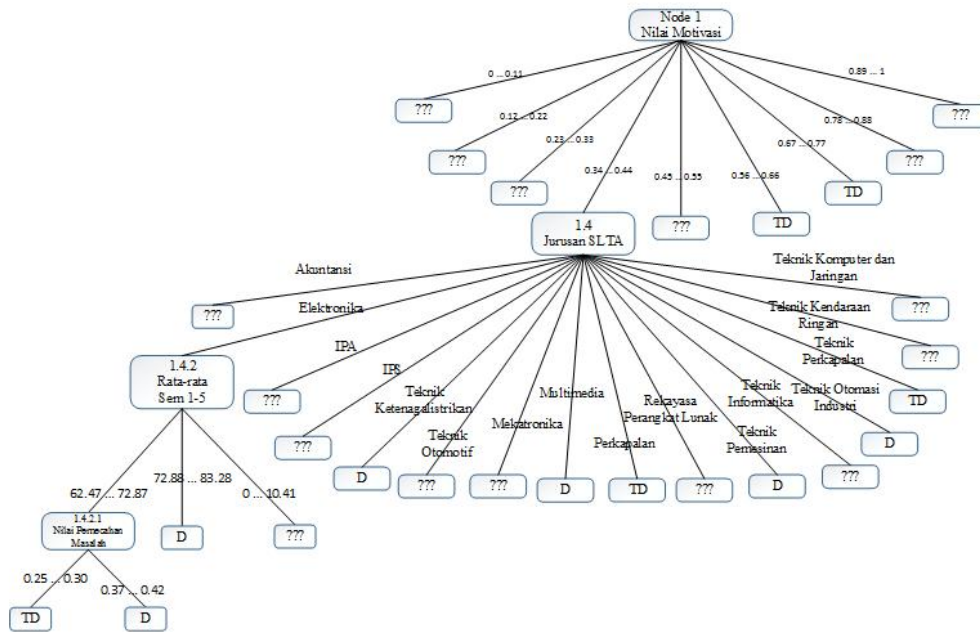


Gambar 38. Decision Tree Hasil Perhitungan Node 1.4.2

IV.1.25 Penentuan Node 1.4.2.1

Untuk menentukan *node* 1.4.2.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.25.

Dari hasil Tabel A.1.25, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PEMECAHAN MASALAH**. Ada 2 *instance* dari **NILAI PEMECAHAN MASALAH**, yaitu **0.25...0.30** dan **0.37...0.42**. Nilai *instance* **0.25...0.30** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* mengklasifikasikan kasus menjadi **DITERIMA**.

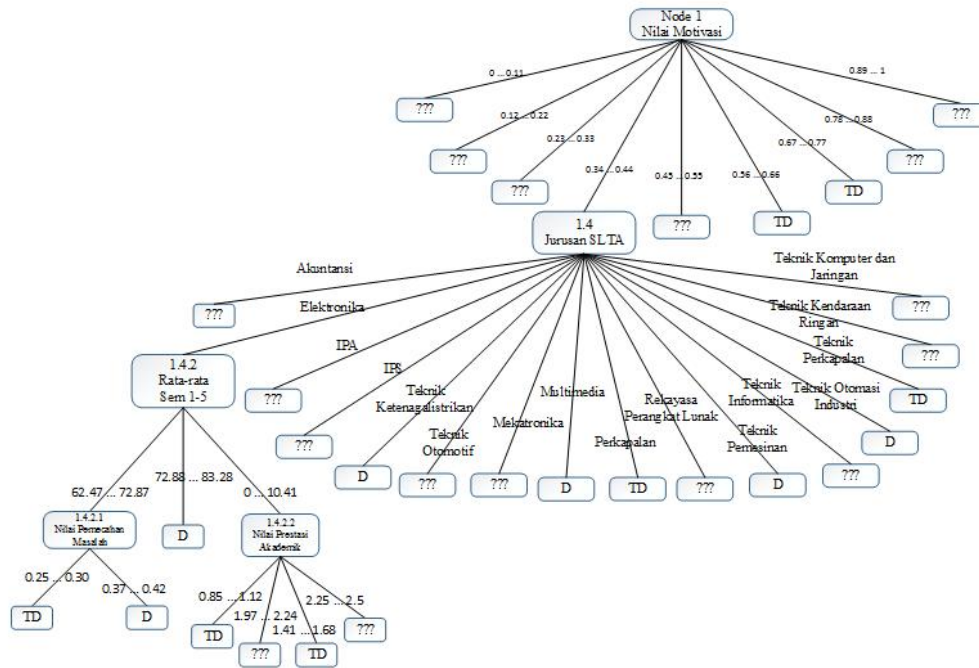


Gambar 39. Decision Tree Hasil Perhitungan Node 1.4.2.1

IV.1.26 Penentuan Node 1.4.2.2

Untuk menentukan *node* 1.4.2.2 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.26.

Dari hasil Tabel A.1.26, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 2 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0.85...1.12**, **1.97...2.24**, **1.41...1.68** dan **2.25...2.5**. Nilai *instance* **0.85...1.12** dan **1.41...1.68** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **1.97...2.24** dan **2.25...2.5** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

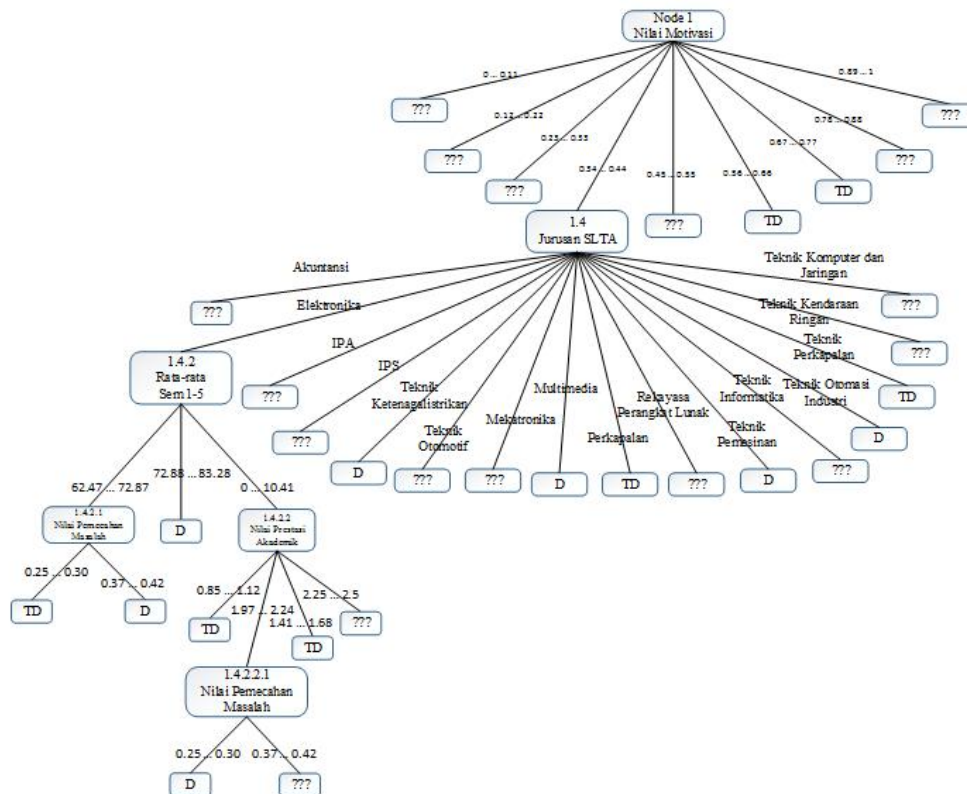


Gambar 40. Decision Tree Hasil Perhitungan Node 1.4.2.2

IV.1.27 Penentuan Node 1.4.2.2.1

Untuk menentukan *node* 1.4.2.2.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.27.

Dari hasil Tabel A.1.27, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PEMECAHAN MASALAH**. Ada 2 *instance* dari **NILAI PEMECAHAN MASALAH**, yaitu **0.25...0.30** dan **0.37...0.42**. Nilai *instance* **0.25...0.30** sudah mengklasifikasikan kasus menjadi **DITERIMA** dan *instance* **0.37...0.42** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

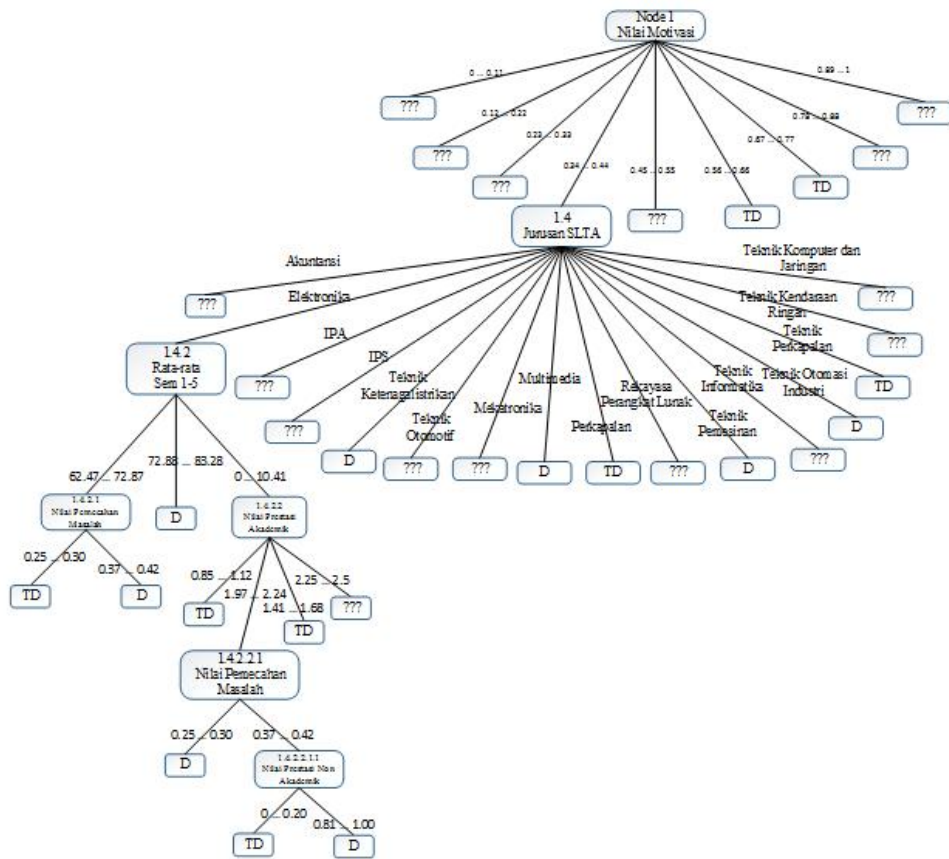


Gambar 41. Decision Tree Hasil Perhitungan Node 1.4.2.2.1

IV.1.28 Penentuan Node 1.4.2.2.1.1

Untuk menentukan *node* 1.4.2.2.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.28.

Dari hasil Tabel A.1.28, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 2 *instance* dari **NILAI PRESTASI NON AKADEMIK**, yaitu **0...0.20** dan **0.81...1.00**. Nilai *instance* **0...0.20** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **0.81...1.00** mengklasifikasikan kasus menjadi **DITERIMA**.

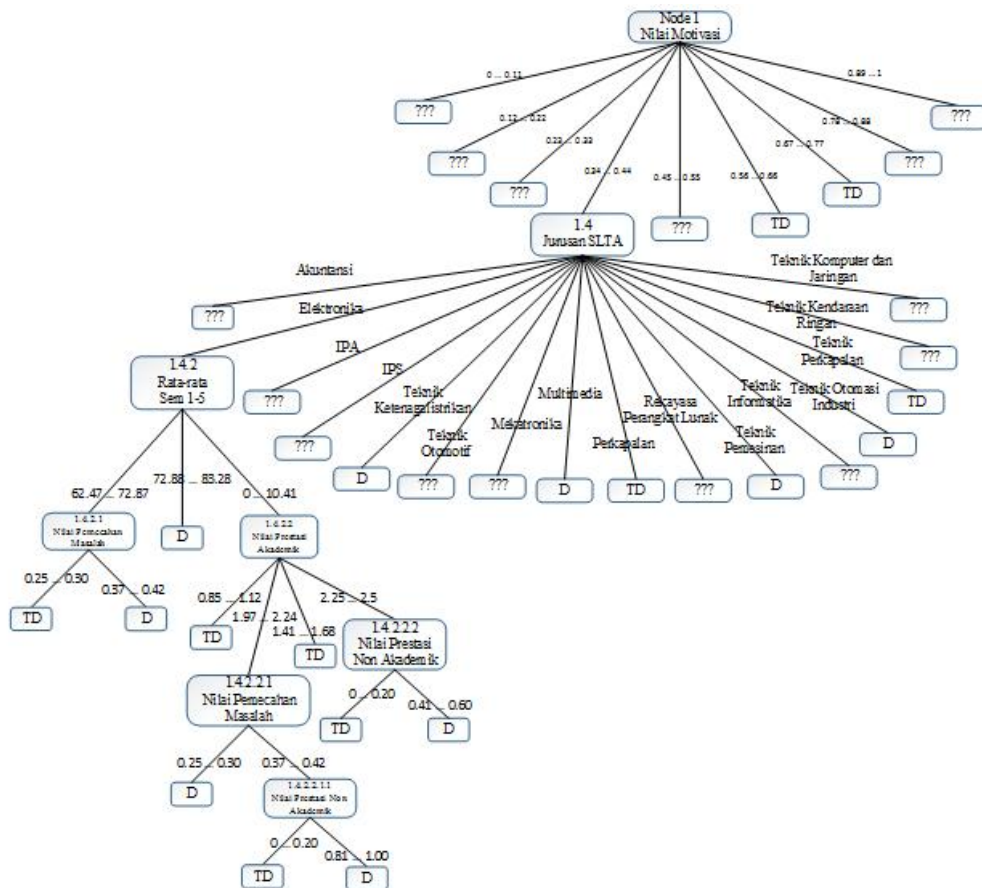


Gambar 42. Decision Tree Hasil Perhitungan Node 1.4.2.2.1.1

IV.1.29 Penentuan Node 1.4.2.2.2

Untuk menentukan node 1.4.2.2.2 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.29.

Dari hasil Tabel A.1.29, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 2 *instance* dari **NILAI PRESTASI NON AKADEMIK**, yaitu **0...0.20** dan **0.41...0.60**. Nilai *instance* **0...0.20** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **0.41...0.60** mengklasifikasikan kasus menjadi **DITERIMA**.

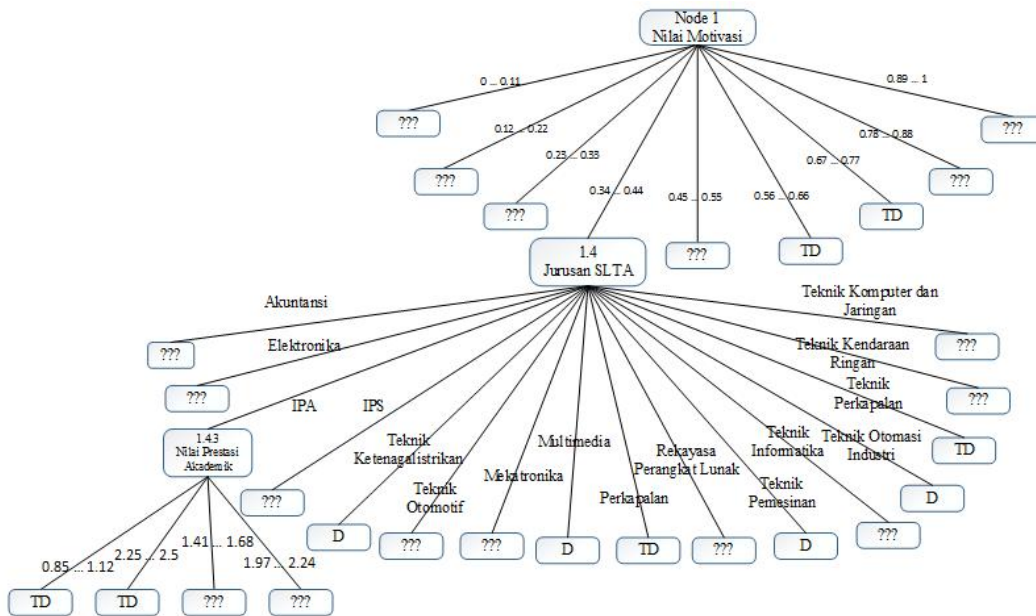


Gambar 43. Decision Tree Hasil Perhitungan Node 1.4.2.2.2

IV.1.30 Penentuan Node 1.4.3

Untuk menentukan *node* 1.4.3 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.30.

Dari hasil Tabel A.1.30, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 4 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0.85...1.12, 2.25...2.5, 1.41...1.68, dan 1.97...2.24**. Nilai *instance* **0.85...1.12, dan 2.25...2.5** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**, sedangkan *instance* **1.41...1.68, dan 1.97...2.24** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

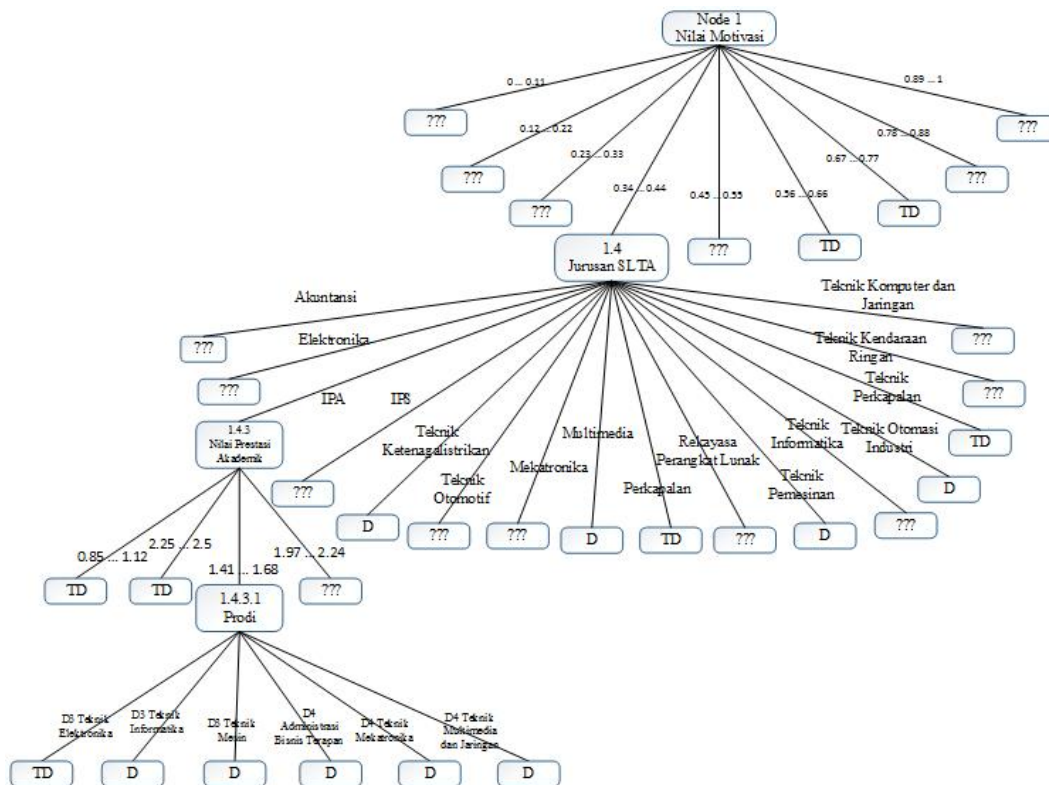


Gambar 44. Decision Tree Hasil Perhitungan Node 1.4.2.2.2

IV.1.31 Penentuan Node 1.4.3.1

Untuk menentukan node 1.4.3.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.31.

Dari hasil Tabel A.1.31, diketahui bahwa atribut dengan *gain* tertinggi adalah **PRODI**. Ada 6 *instance* dari **PRODI**, yaitu **D3 Teknik Elektronika**, **D3 Teknik Informatika**, **D3 Teknik Mesin**, **D4 Administrasi Bisnis Terapan**, **D4 Teknik Mekatronika**, dan **D4 Teknik Multimedia dan Jaringan**. Nilai *instance* **D3 Teknik Elektronika** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance* **D3 Teknik Informatika**, **D3 Teknik Mesin**, **D4 Administrasi Bisnis Terapan**, **D4 Teknik Mekatronika**, dan **D4 Teknik Multimedia dan Jaringan** mengklasifikasikan kasus menjadi **DITERIMA**.

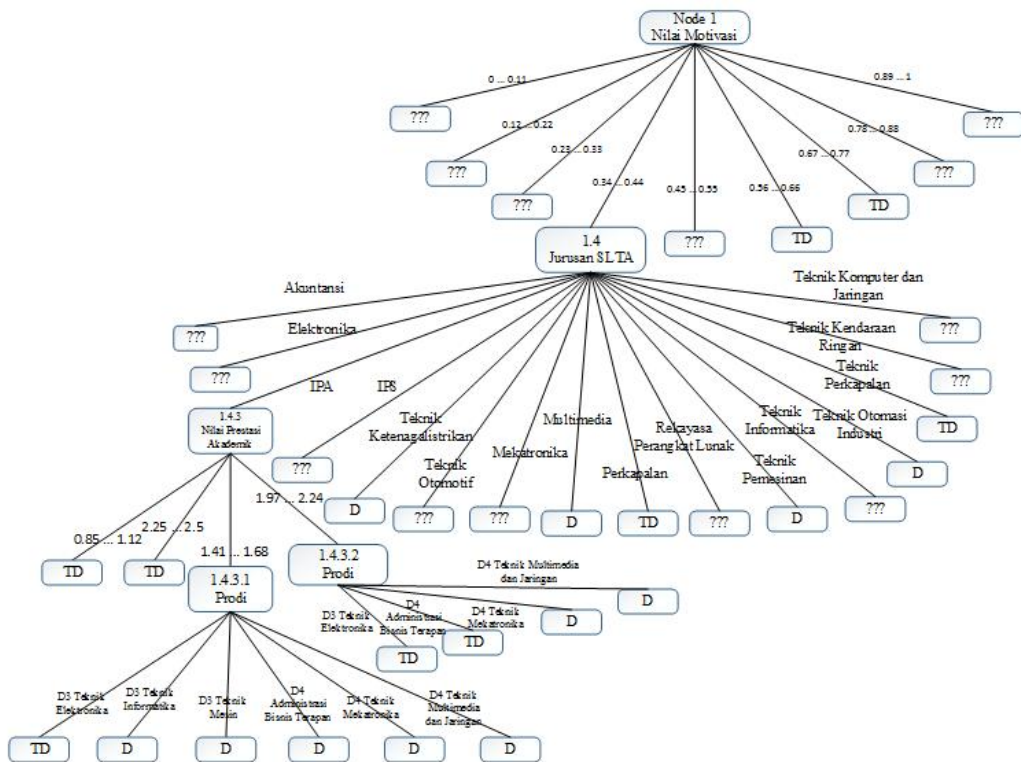


Gambar 45. *Decision Tree* Hasil Perhitungan *Node 1.4.3.1*

IV.1.32 Penentuan *Node 1.4.3.2*

Untuk menentukan *node 1.4.3.2* dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.32.

Dari hasil Tabel A.1.32, diketahui bahwa atribut dengan *gain* tertinggi adalah **PRODI**. Ada 6 *instance* dari **PRODI**, yaitu **D3 Teknik Elektronika**, **D4 Administrasi Bisnis Terapan**, **D4 Teknik Mekatronika**, dan **D4 Teknik Multimedia dan Jaringan**. Nilai *instance D3 Teknik Elektronika* dan **D4 Administrasi Bisnis Terapan** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA** dan *instance D4 Teknik Mekatronika* dan **D4 Teknik Multimedia dan Jaringan** sudah mengklasifikasikan kasus menjadi **DITERIMA**.

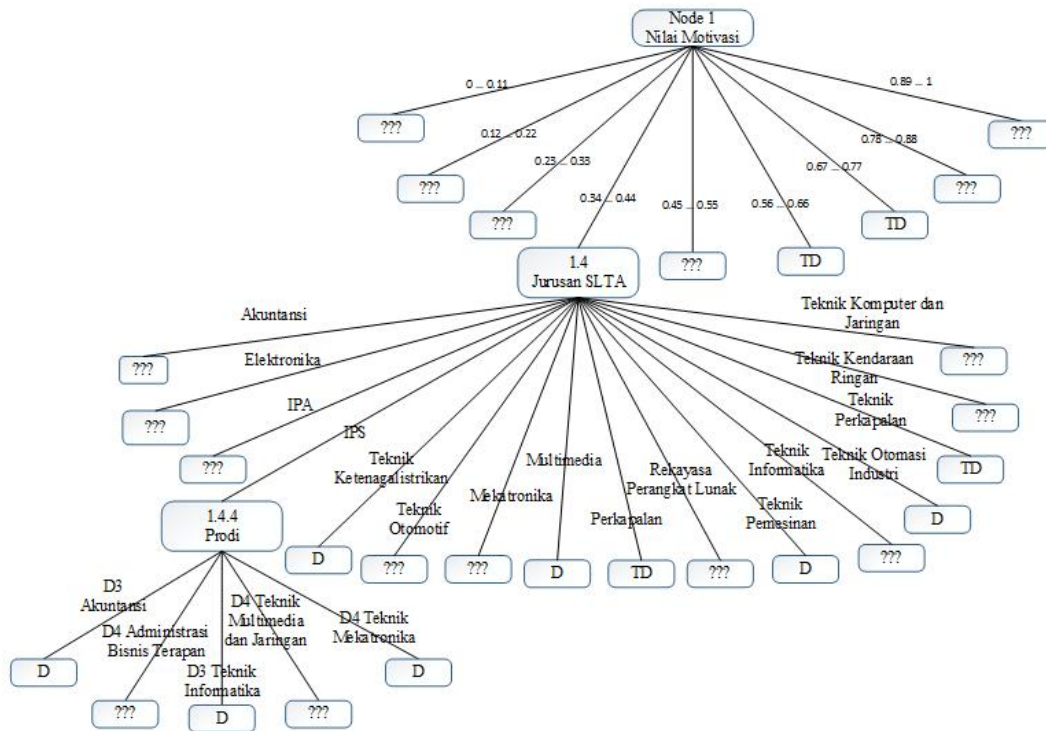


Gambar 46. Decision Tree Hasil Perhitungan Node 1.4.3.2

IV.1.33 Penentuan Node 1.4.4

Untuk menentukan *node* 1.4.4 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.33.

Dari hasil Tabel A.1.33, diketahui bahwa atribut dengan *gain* tertinggi adalah **PRODI**. Ada 5 *instance* dari **PRODI**, yaitu **D3 Akuntansi**, **D3 Teknik Informatika**, **D4 Administrasi Bisnis Terapan**, **D4 Teknik Mekatronika** dan **D4 Teknik Multimedia dan Jaringan**. Nilai *instance* **D3 Akuntansi**, **D3 Teknik Informatika**, dan **D4 Teknik Mekatronika** sudah mengklasifikasikan kasus menjadi **DITERIMA** dan *instance* **D4 Administrasi Bisnis Terapan** dan **D4 Teknik Multimedia dan Jaringan** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

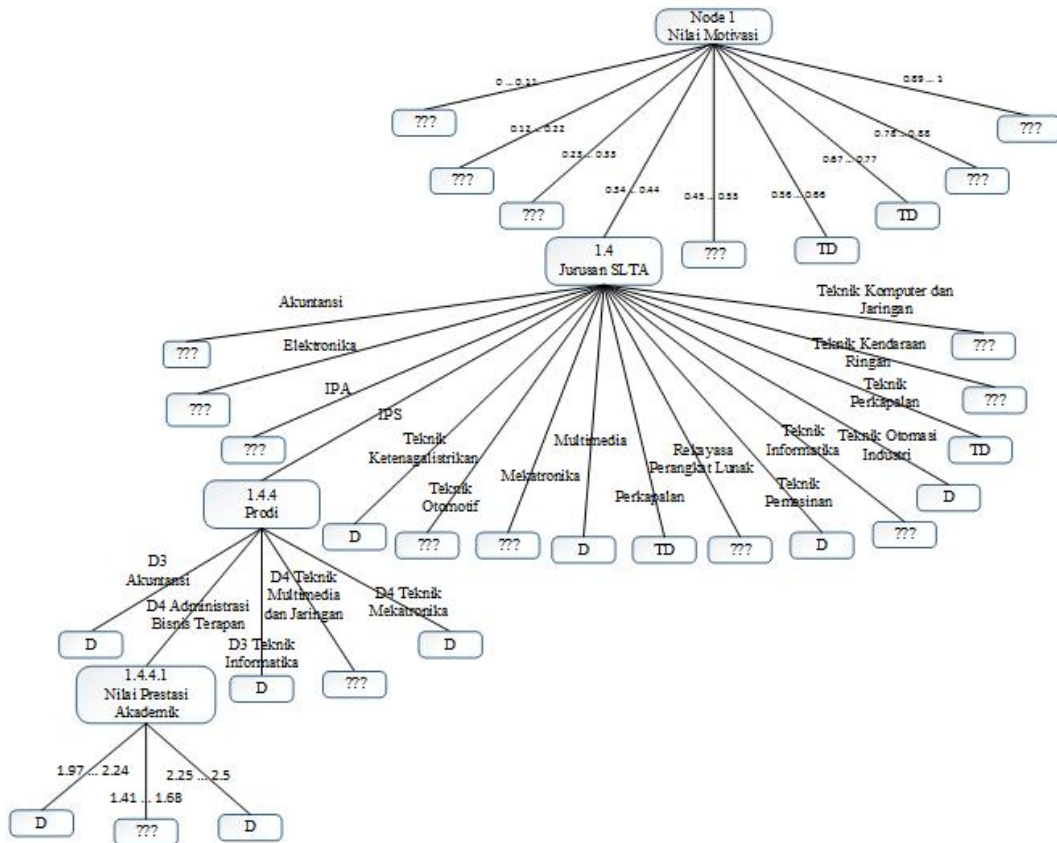


Gambar 47. Decision Tree Hasil Perhitungan Node 1.4.4

IV.1.34 Penentuan Node 1.4.4.1

Untuk menentukan *node* 1.4.4.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.34.

Dari hasil Tabel A.1.34, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 3 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **1.97...2.24**, **1.41...1.68**, dan **2.25...2.5**. Nilai *instance* **1.97...2.24** dan **2.25...2.5** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **1.41...1.68** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

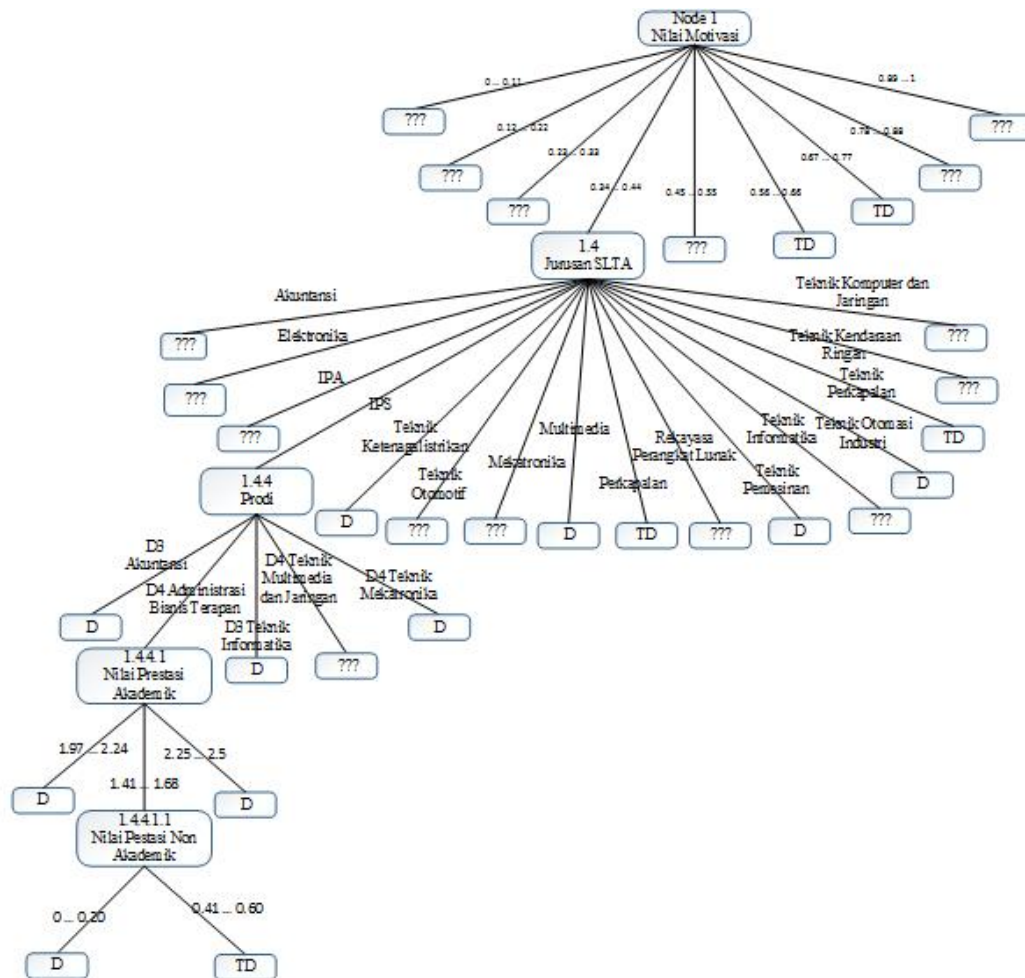


Gambar 48. Decision Tree Hasil Perhitungan Node 1.4.4.1

IV.1.35 Penentuan Node 1.4.4.1.1

Untuk menentukan *node* 1.4.4.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.35.

Dari hasil Tabel A.1.35, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 2 *instance* dari **NILAI PRESTASI NON AKADEMIK**, yaitu **0...0.20** dan **0.41...0.60**. Nilai *instance* **0.41...0.60** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **0...0.20** mengklasifikasikan kasus menjadi **DITERIMA**.

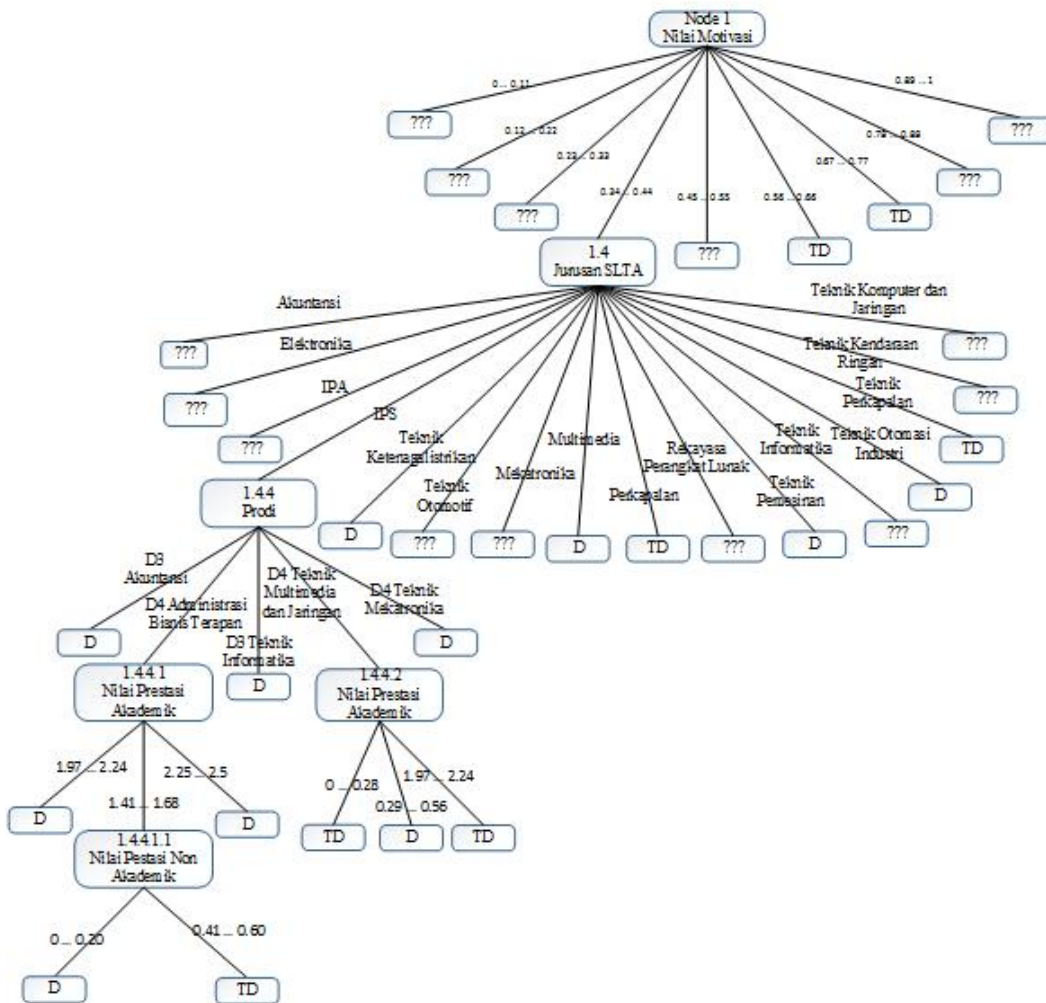


Gambar 49. Decision Tree Hasil Perhitungan Node 1.4.4.1.1

IV.1.36 Penentuan Node 1.4.4.2

Untuk menentukan *node* 1.4.4.2 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.36.

Dari hasil Tabel A.1.36, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 3 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **1.97...2.24**, **0...0.28**, dan **0.29...0.56**. Nilai *instance* **0...0.28** dan **1.97...2.24** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **0.29...0.56** mengklasifikasikan kasus menjadi **DITERIMA**.

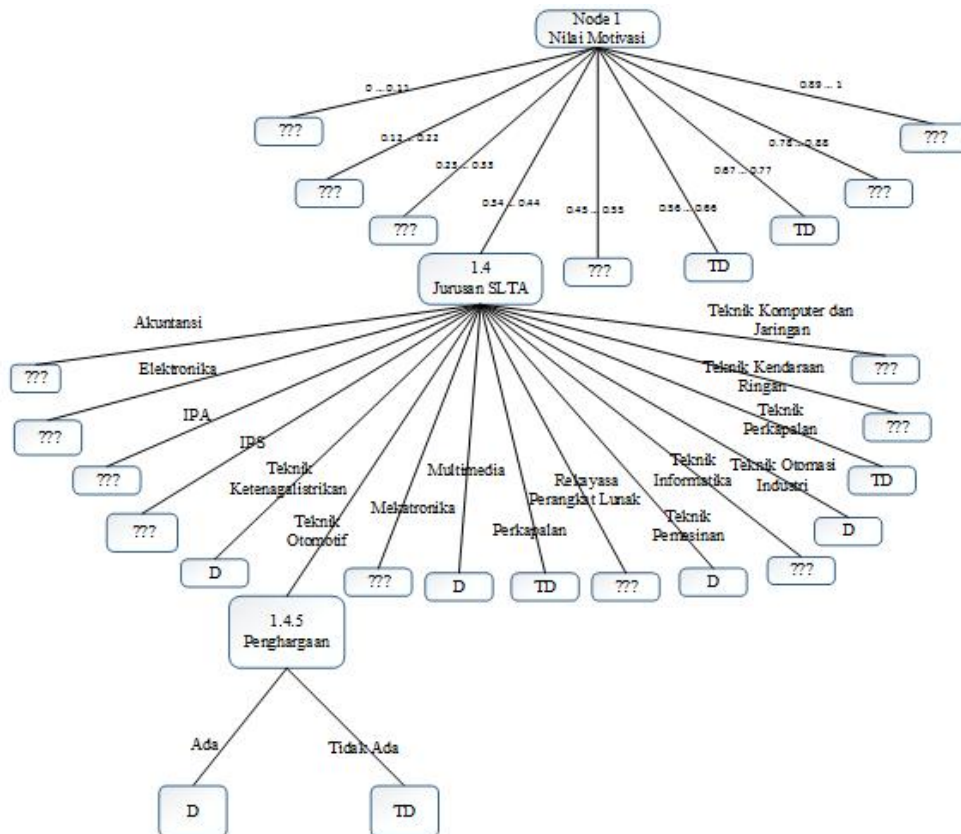


Gambar 50. Decision Tree Hasil Perhitungan Node 1.4.4.2

IV.1.37 Penentuan Node 1.4.5

Untuk menentukan *node* 1.4.5 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.37.

Dari hasil Tabel A.1.37, diketahui bahwa atribut dengan *gain* tertinggi adalah **PENGHARGAAN**. Ada 2 *instance* dari **PENGHARGAAN**, yaitu **Ada** dan **Tidak Ada**. Nilai *instance* **Tidak Ada** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **Ada** mengklasifikasikan kasus menjadi **DITERIMA**.

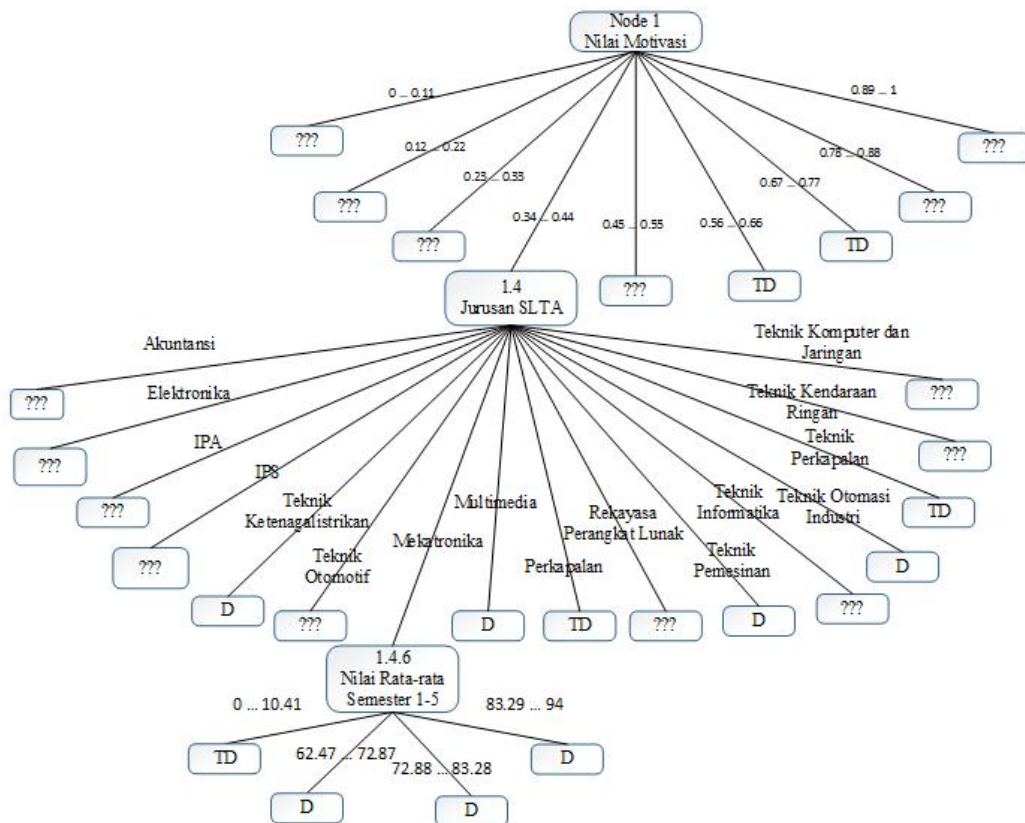


Gambar 51. Decision Tree Hasil Perhitungan Node 1.4.5

IV.1.38 Penentuan Node 1.4.6

Untuk menentukan node 1.4.6 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.38.

Dari hasil Tabel A.1.38, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI RATA-RATA SEMESTER 1-5**. Ada 4 *instance* dari **NILAI RATA-RATA SEMESTER 1-5**, yaitu **0..10.41**, **62.47...72.87**, **72.88...83.28** dan **83.29...94**. Nilai *instance* **0..10.41** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **62.47...72.87**, **72.88...83.28** dan **83.29...94** sudah mengklasifikasikan kasus menjadi **DITERIMA**.

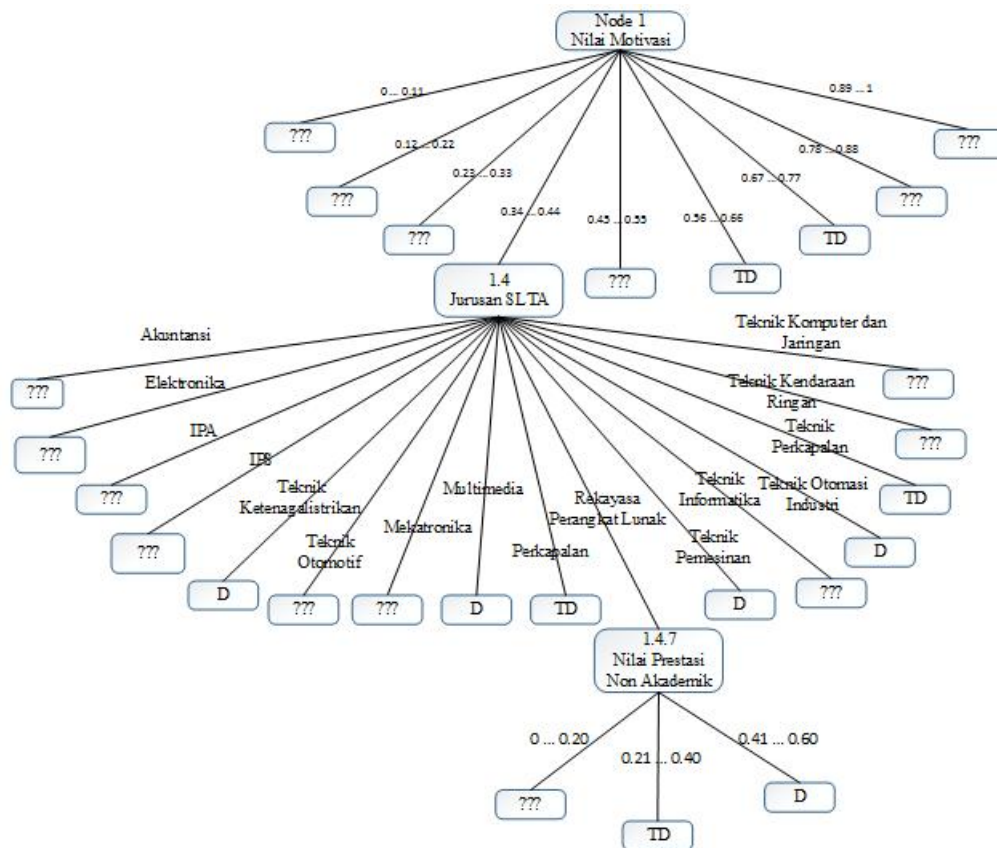


Gambar 52. Decision Tree Hasil Perhitungan Node 1.4.6

IV.1.39 Penentuan Node 1.4.7

Untuk menentukan *node* 1.4.6 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.38.

Dari hasil Tabel A.1.38, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 3 *instance* dari **NILAI PRESTASI NON AKADEMIK**, yaitu **0...0.20**, **0.21...0.40**, dan **0.41...0.60**. Nilai *instance* **0.21...0.40** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **0.41...0.60** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **0...0.20** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

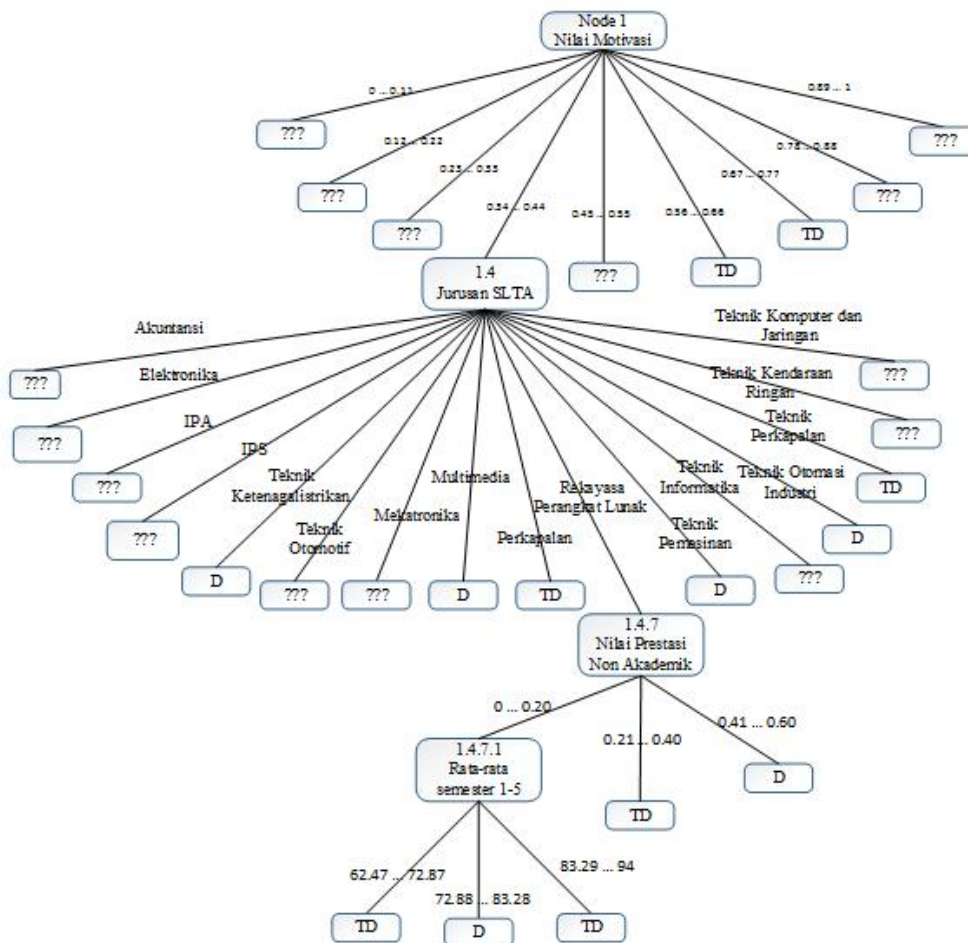


Gambar 53. Decision Tree Hasil Perhitungan Node 1.4.7

IV.1.40 Penentuan Node 1.4.7.1

Untuk menentukan node 1.4.7.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.40.

Dari hasil Tabel A.1.40, diketahui bahwa atribut dengan *gain* tertinggi adalah **RATA-RATA SEMESTER 1-5**. Ada 3 *instance* dari **RATA-RATA SEMESTER 1-5**, yaitu **62.47...72.87**, **72.88...83.28**, dan **83.29...94**. Nilai *instance* **62.47...72.87** dan **83.29...94** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **72.88...83.28** sudah mengklasifikasikan kasus menjadi **DITERIMA**.

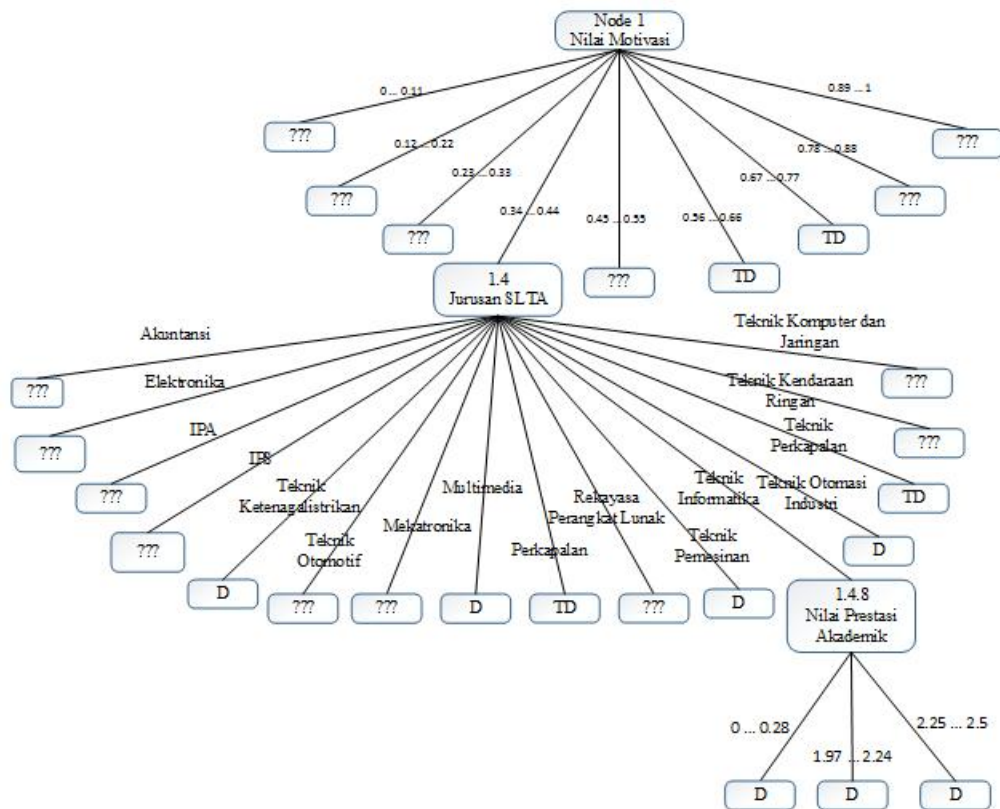


Gambar 54. Decision Tree Hasil Perhitungan Node 1.4.7.1

IV.1.41 Penentuan Node 1.4.8

Untuk menentukan *node* 1.4.8 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.41.

Dari hasil Tabel A.1.41, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 3 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0...0.28**, **1.97...2.24**, dan **2.25...2.5**. Nilai *instance* **0...0.28**, **1.97...2.24** dan **2.25...2.5** sudah mengklasifikasikan kasus menjadi **DITERIMA**.

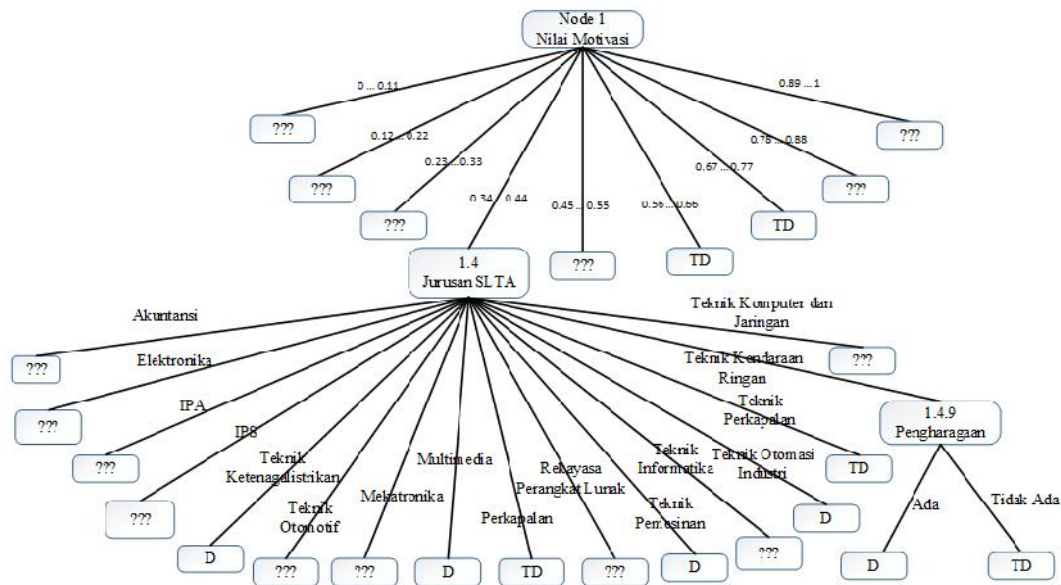


Gambar 55. Decision Tree Hasil Perhitungan Node 1.4.8

IV.1.42 Penentuan Node 1.4.9

Untuk menentukan *node* 1.4.9 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.42.

Dari hasil Tabel A.1.42, diketahui bahwa atribut dengan *gain* tertinggi adalah **PENGHARGAAN**. Ada 2 *instance* dari **PENGHARGAAN**, yaitu **Ada** dan **Tidak Ada**. Nilai *instance* **Ada** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **Tidak Ada** mengklasifikasikan kasus menjadi **TIDAK DITERIMA**.

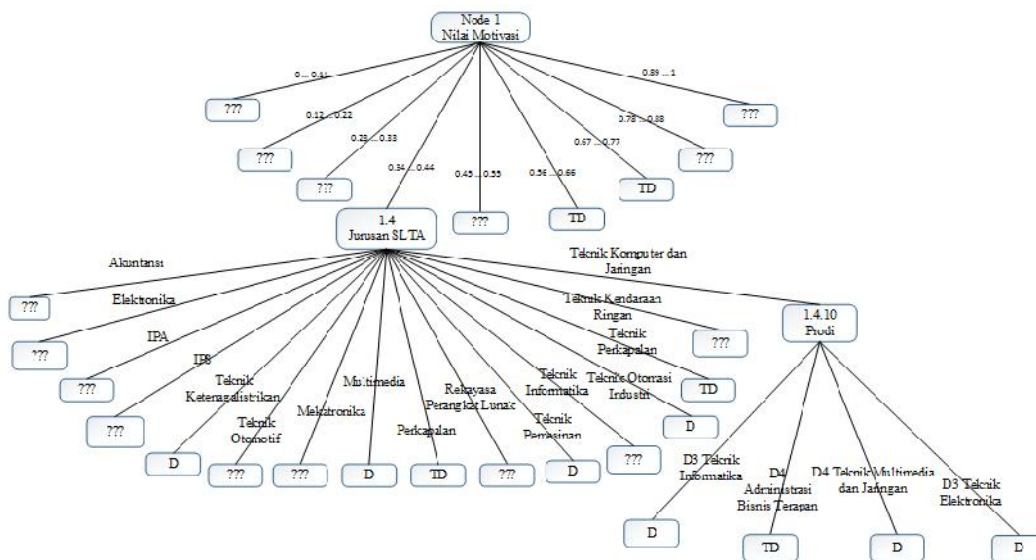


Gambar 56. Decision Tree Hasil Perhitungan Node 1.4.9

IV.1.43 Penentuan Node 1.4.10

Untuk menentukan *node* 1.4.10 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.43.

Dari hasil Tabel A.1.43, diketahui bahwa atribut dengan *gain* tertinggi adalah **PRODI**. Ada 4 *instance* dari **PRODI**, yaitu **D3 Teknik Elektronika**, **D3 Teknik Informatika**, **D4 Administrasi Bisnis Terapan** dan **D4 Teknik Multimedia dan Jaringan**. Nilai *instance* **D3 Teknik Elektronika**, **D3 Teknik Multimedia dan Jaringan** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **D4 Administrasi Bisnis Terapan** mengklasifikasikan kasus menjadi **TIDAK DITERIMA**.

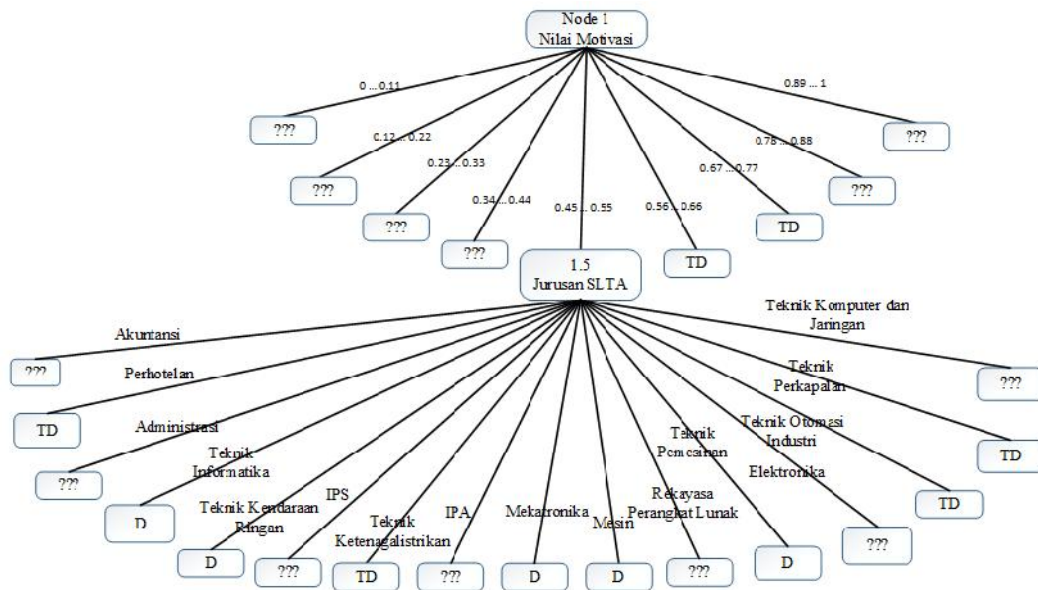


Gambar 57. Decision Tree Hasil Perhitungan Node 1.4.10

IV.1.44 Penentuan Node 1.5

Untuk menentukan *node* 1.5 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.44.

Dari hasil Tabel A.1.44, diketahui bahwa atribut dengan *gain* tertinggi adalah **JURUSAN SLTA**. Ada 16 *instance* dari **JURUSAN SLTA**, yaitu **Akuntansi, Perhotelan, Administrasi, Teknik Informatika, Teknik Kendaraan Ringan, IPS, Teknik Ketenagalistrikan, IPA, Mekatronika, Mesin, Rekayasa Perangkat Lunak, Teknik Pemesinan, Elektronika, Teknik Otomasi Industri, Teknik Perkapalan** dan **Teknik Komputer dan Jaringan**. Nilai *instance* **Perhotelan, Teknik Ketenagalistrikan, Teknik Otomasi Industri, dan Teknik Perkapalan** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **Teknik Informatika, Teknik Kendaraan Ringan, Mekatronika, Mesin, dan Teknik Pemesinan** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **Teknik Komputer dan Jaringan, Elektronika, IPA, IPS, Rekayasa Perangkat Lunak, Administrasi, dan Akuntansi** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

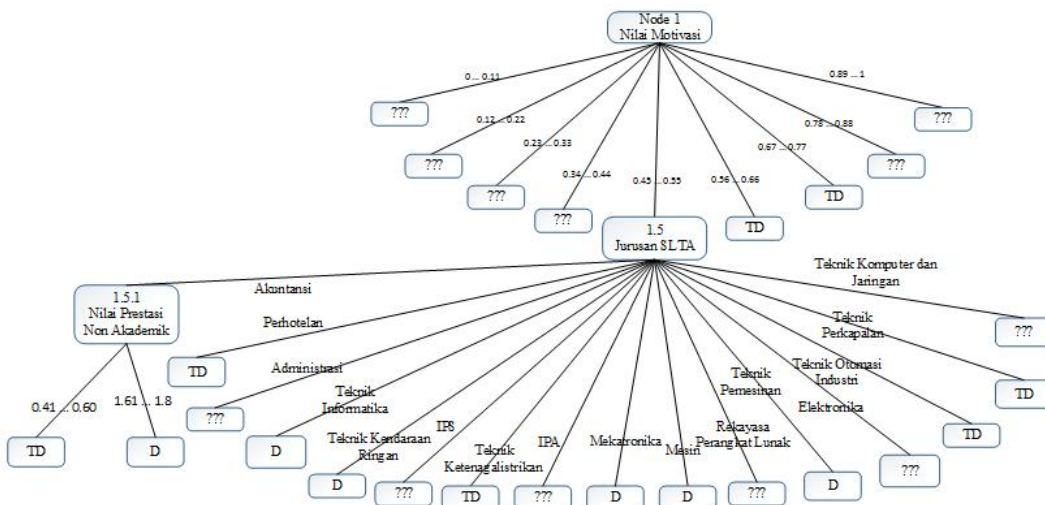


Gambar 58. Decision Tree Hasil Perhitungan Node 1.5

IV.1.45 Penentuan Node 1.5.1

Untuk menentukan *node 1.5* dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.45.

Dari hasil Tabel A.1.45, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 2 *instance* dari **NILAI PRESTASI NON AKADEMIK**, yaitu **0.41...0.60** dan **1.61...1.8**. Nilai *instance 0.41...0.60* sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance 1.61...1.8* mengklasifikasikan kasus menjadi **DITERIMA**.

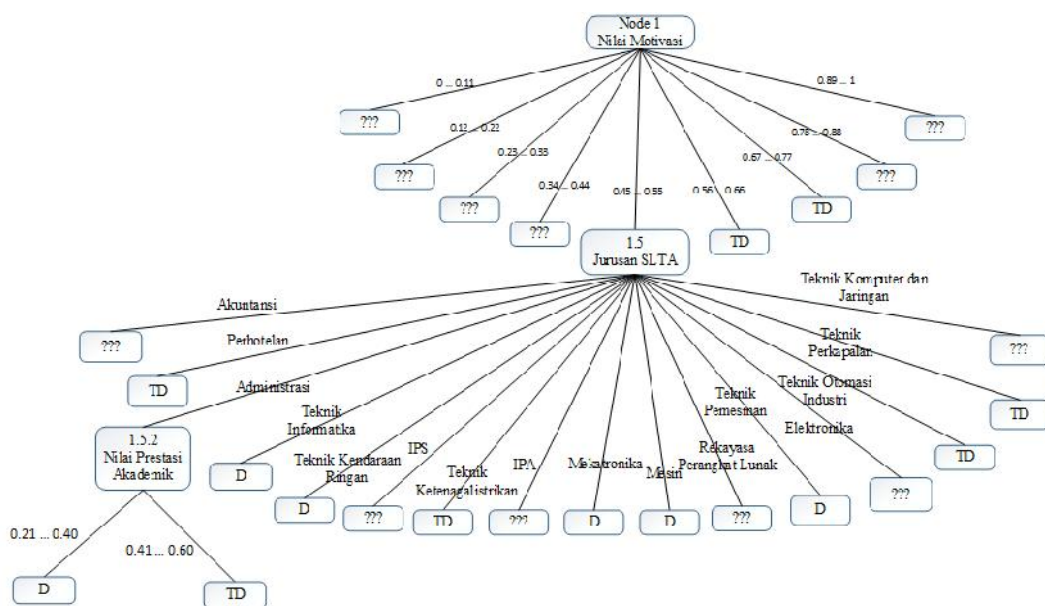


Gambar 59. Decision Tree Hasil Perhitungan Node 1.5.1

IV.1.46 Penentuan *Node* 1.5.2

Untuk menentukan *node* 1.5.2 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.46.

Dari hasil Tabel A.1.46, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 2 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0.21...0.40** dan **0.41...0.60**. Nilai *instance* **0.21...0.40** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **0.41...0.60** mengklasifikasikan kasus menjadi **TIDAK DITERIMA**.

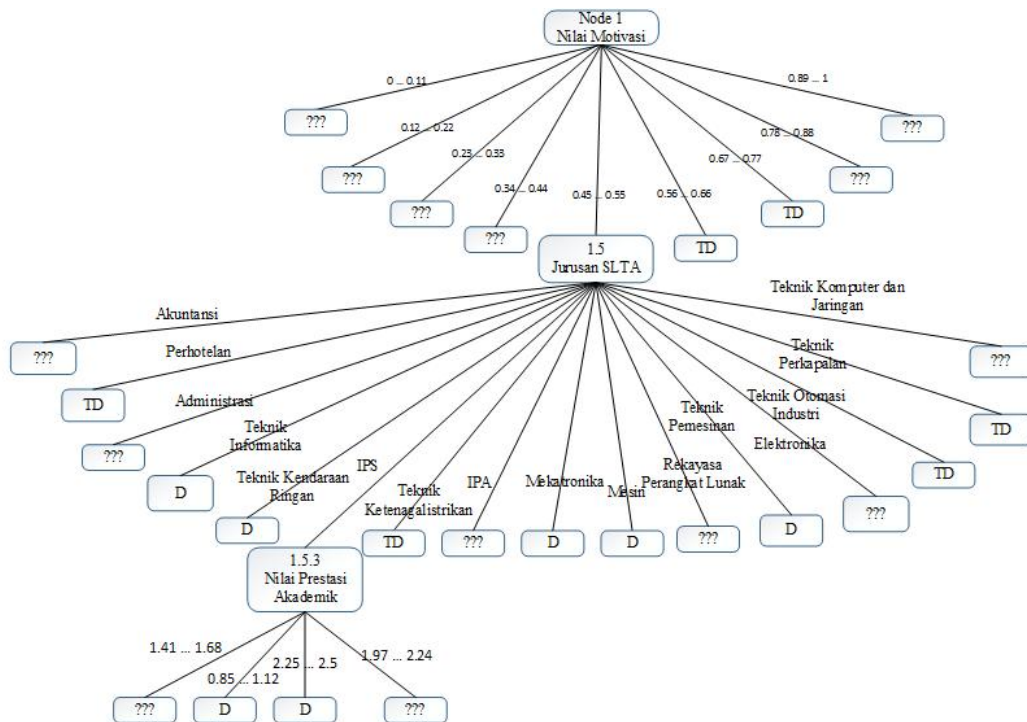


Gambar 60. *Decision Tree* Hasil Perhitungan *Node* 1.5.2

IV.1.47 Penentuan *Node* 1.5.3

Untuk menentukan *node* 1.5.3 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.47

Dari hasil Tabel A.1.47, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 4 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **1.41...1.68**, **0.85...1.12**, **2.25...2.5**, dan **1.97...2.24**. Nilai *instance* **0.85...1.12**, dan **2.25...2.5** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **1.41...1.68** dan **1.97...2.24** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

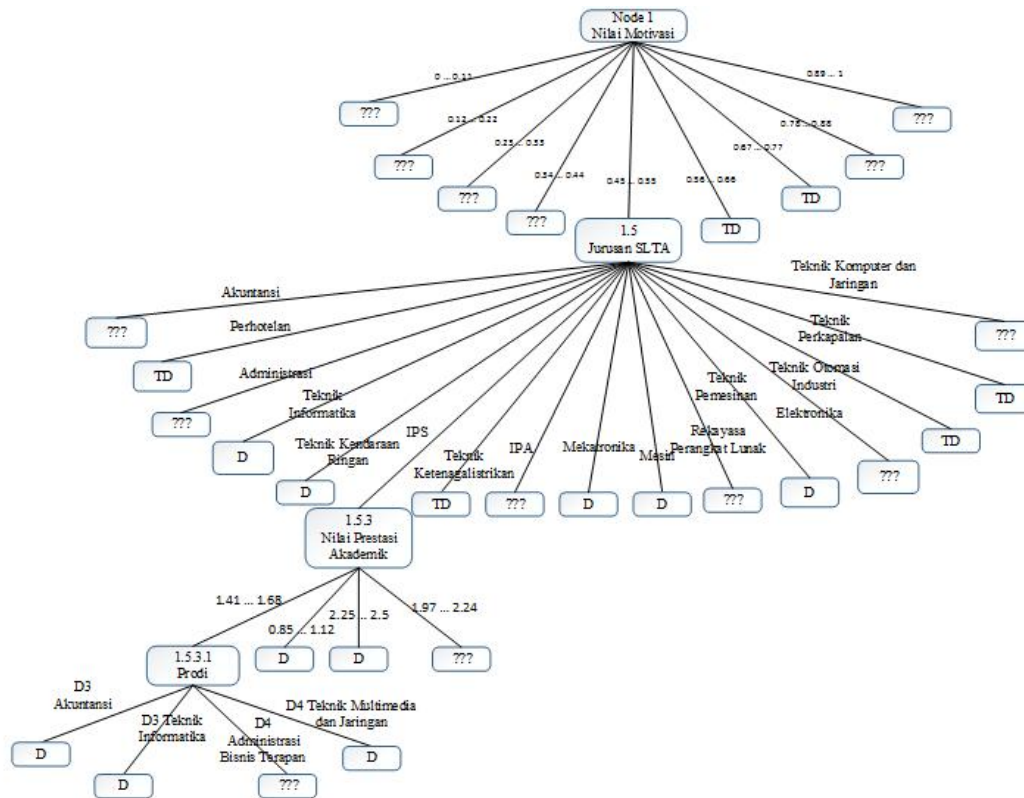


Gambar 61. Decision Tree Hasil Perhitungan Node 1.5.3

IV.1.48 Penentuan Node 1.5.3.1

Untuk menentukan node 1.5.3.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.48

Dari hasil Tabel A.1.48, diketahui bahwa atribut dengan *gain* tertinggi adalah **PRODI**. Ada 4 *instance* dari **PRODI**, yaitu **D3 Akuntansi**, **D3 Teknik Informatika**, **D4 Administrasi Bisnis Terapan**, dan **D4 Teknik Multimedia dan Jaringan**. Nilai *instance* **D3 Akuntansi**, **D3 Teknik Informatika**, dan **D4 Teknik Multimedia dan Jaringan** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **D4 Administrasi Bisnis Terapan** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

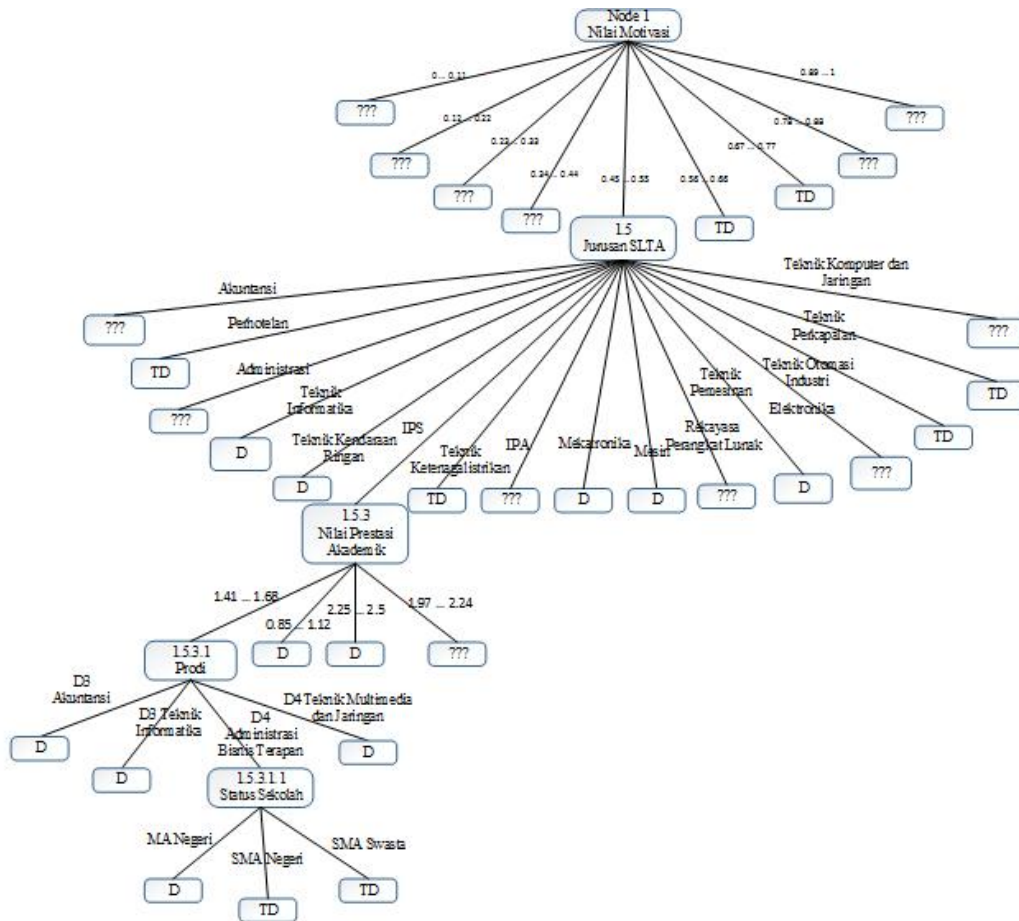


Gambar 62. Decision Tree Hasil Perhitungan Node 1.5.3.1

IV.1.49 Penentuan Node 1.5.3.1.1

Untuk menentukan *node* 1.5.3.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.49.

Dari hasil Tabel A.1.49, diketahui bahwa atribut dengan *gain* tertinggi adalah **STATUS SEKOLAH**. Ada 3 *instance* dari **STATUS SEKOLAH**, yaitu **MA Negeri**, **SMA Negeri**, dan **SMA Swasta**. Nilai *instance* **SMA Negeri** dan **SMA Swasta** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **MA Negeri** mengklasifikasikan kasus menjadi **DITERIMA**.

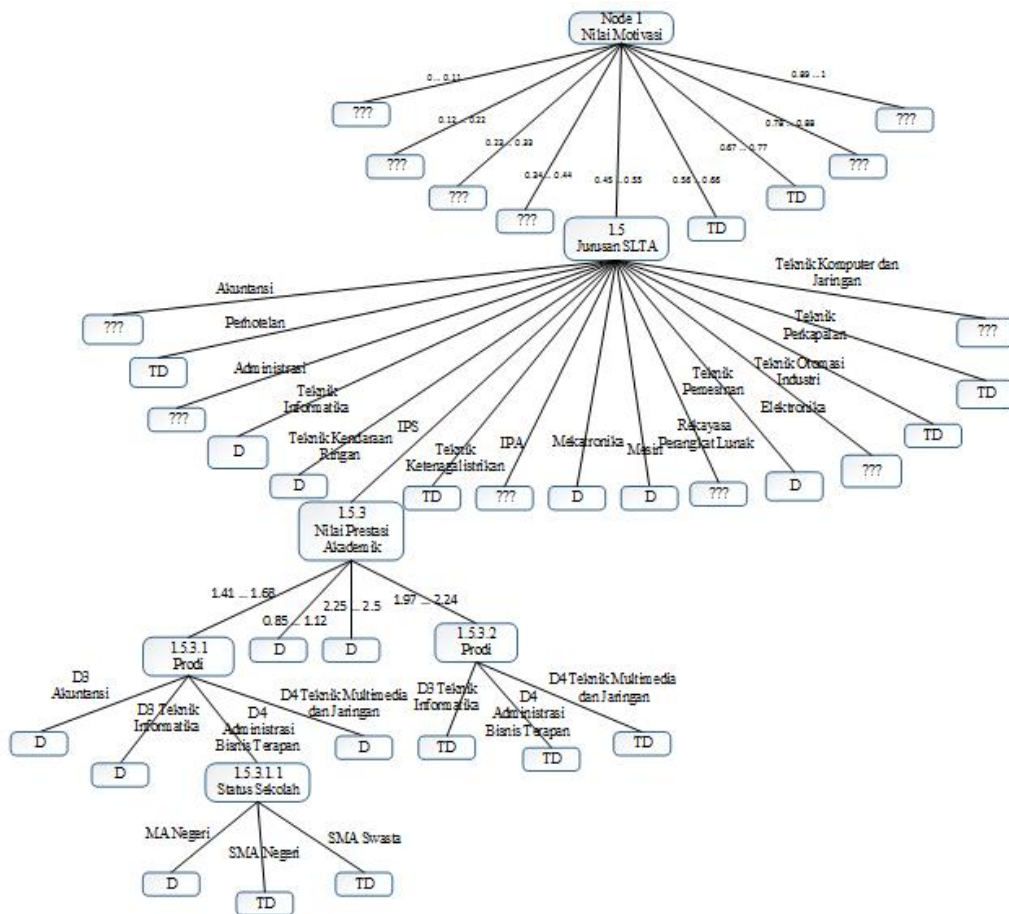


Gambar 63. Decision Tree Hasil Perhitungan Node 1.5.3.1.1

IV.1.50 Penentuan Node 1.5.3.2

Untuk menentukan *node* 1.5.3.2 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.50.

Dari hasil Tabel A.1.50, diketahui bahwa atribut dengan *gain* tertinggi adalah **PRODI**. Ada 3 *instance* dari **PRODI**, yaitu **D3 Teknik Informatika**, **D4 Administrasi Bisnis Terapan**, dan **D4 Teknik Multimedia dan Jaringan**. Nilai *instance* **D3 Teknik Informatika**, **D4 Administrasi Bisnis Terapan**, dan **D4 Teknik Multimedia dan Jaringan** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**.

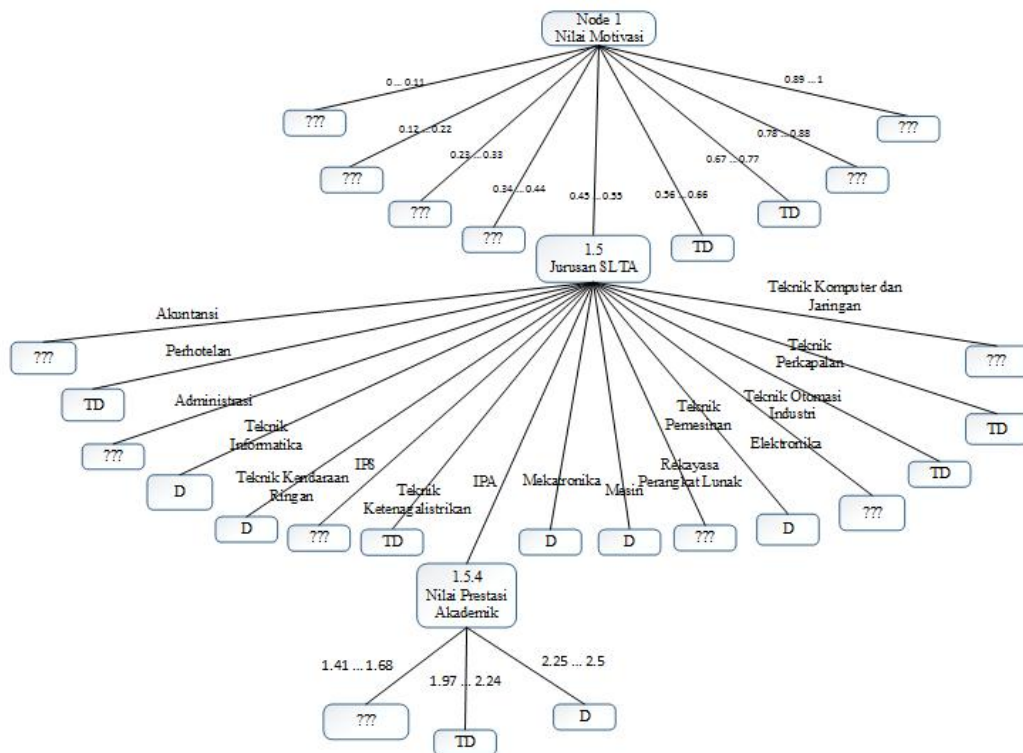


Gambar 64. Decision Tree Hasil Perhitungan Node 1.5.3.2

IV.1.51 Penentuan Node 1.5.4

Untuk menentukan *node* 1.5.4 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.51

Dari hasil Tabel A.1.51, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 3 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **1.41...1.68**, **1.97...2.24**, dan **2.25...2.5**. Nilai *instance* **1.97...2.24** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **2.25...2.5** mengklasifikasikan kasus menjadi **DITERIMA**. *Instance* **1.41...1.68** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

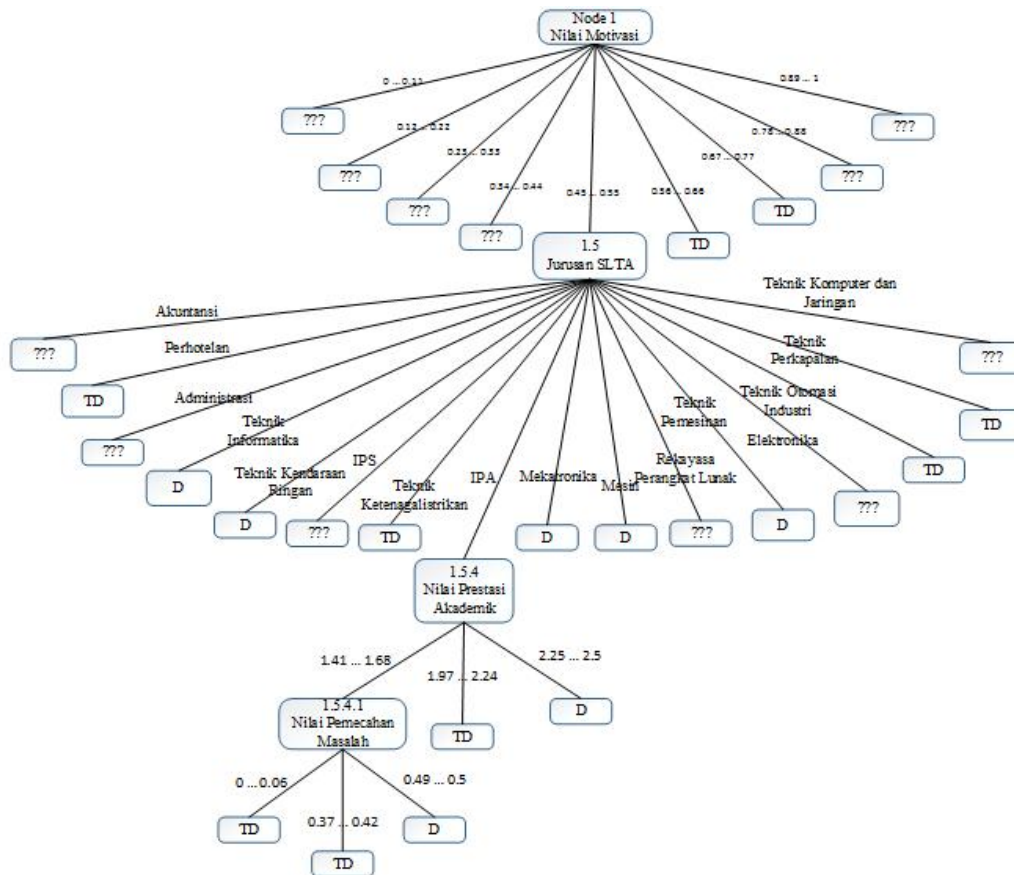


Gambar 65. *Decision Tree* Hasil Perhitungan *Node 1.5.4*

IV.1.52 Penentuan *Node 1.5.4.1*

Untuk menentukan *node 1.5.4.1* dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.52.

Dari hasil Tabel A.1.52, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PEMECAHAN MASALAH**. Ada 3 *instance* dari **NILAI PEMECAHAN MASALAH**, yaitu **0...0.06**, **0.37...0.42**, dan **0.49...0.5**. Nilai *instance* **0...0.06** dan **0.37...0.42** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **0.49...0.5** mengklasifikasikan kasus menjadi **DITERIMA**.

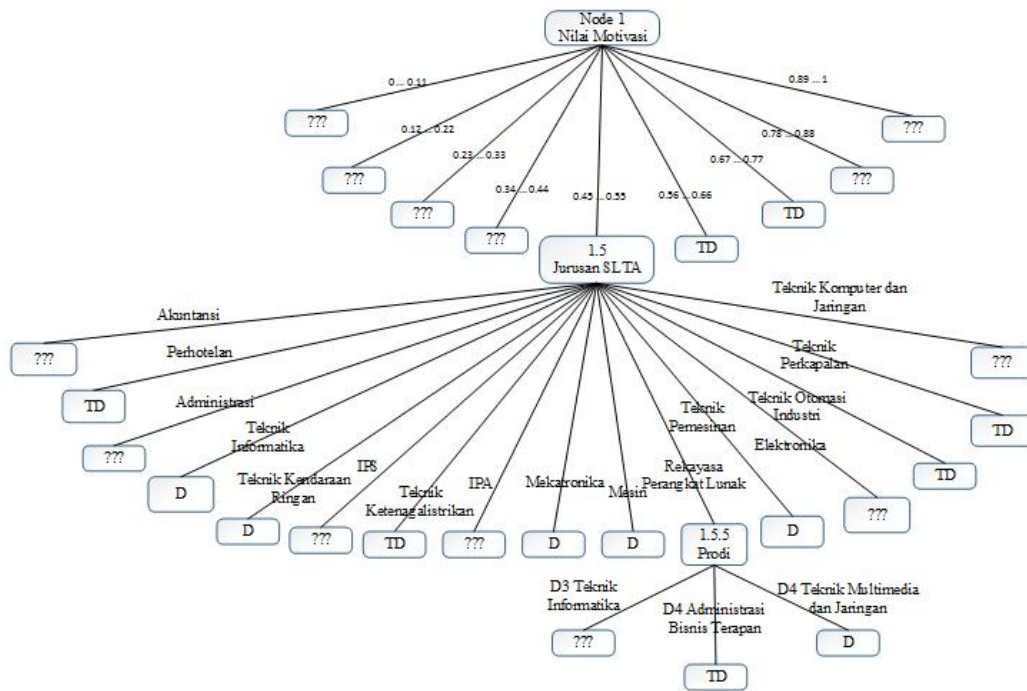


Gambar 66. Decision Tree Hasil Perhitungan Node 1.5.4.1

IV.1.53 Penentuan Node 1.5.5

Untuk menentukan *node* 1.5.5 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.53

Dari hasil Tabel A.1.53, diketahui bahwa atribut dengan *gain* tertinggi adalah **PRODI**. Ada 3 *instance* dari **PRODI**, yaitu **D3 Teknik Informatika**, **D4 Administrasi Bisnis Terapan**, dan **D4 Teknik Multimedia dan Jaringan**. Nilai *instance* **D4 Administrasi Bisnis Terapan** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **D4 Teknik Multimedia dan Jaringan** mengklasifikasikan kasus menjadi **DITERIMA**. *Instance* **D3 Teknik Informatika** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

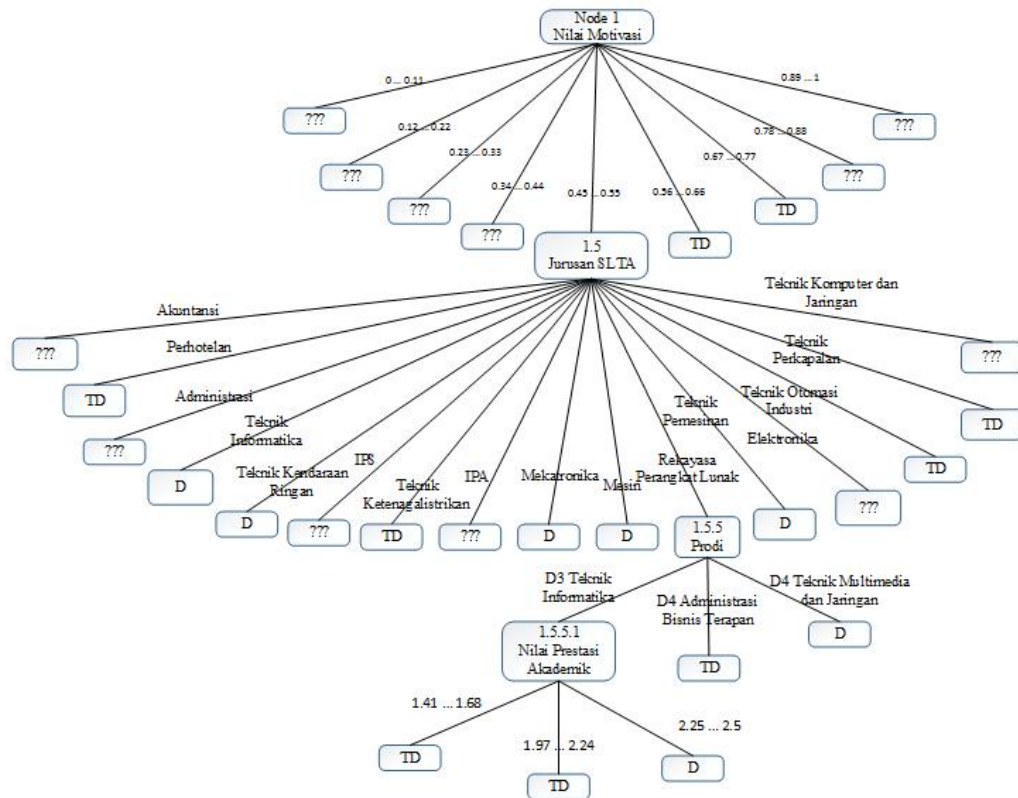


Gambar 67. Decision Tree Hasil Perhitungan Node 1.5.5

IV.1.54 Penentuan Node 1.5.5.1

Untuk menentukan *node* 1.5.5.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.54

Dari hasil Tabel A.1.54, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 3 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **1.41...1.68**, **1.97...2.24**, dan **2.25...2.5**. Nilai *instance* **1.41...1.68** dan **1.97...2.24** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **2.25...2.5** mengklasifikasikan kasus menjadi **DITERIMA**.

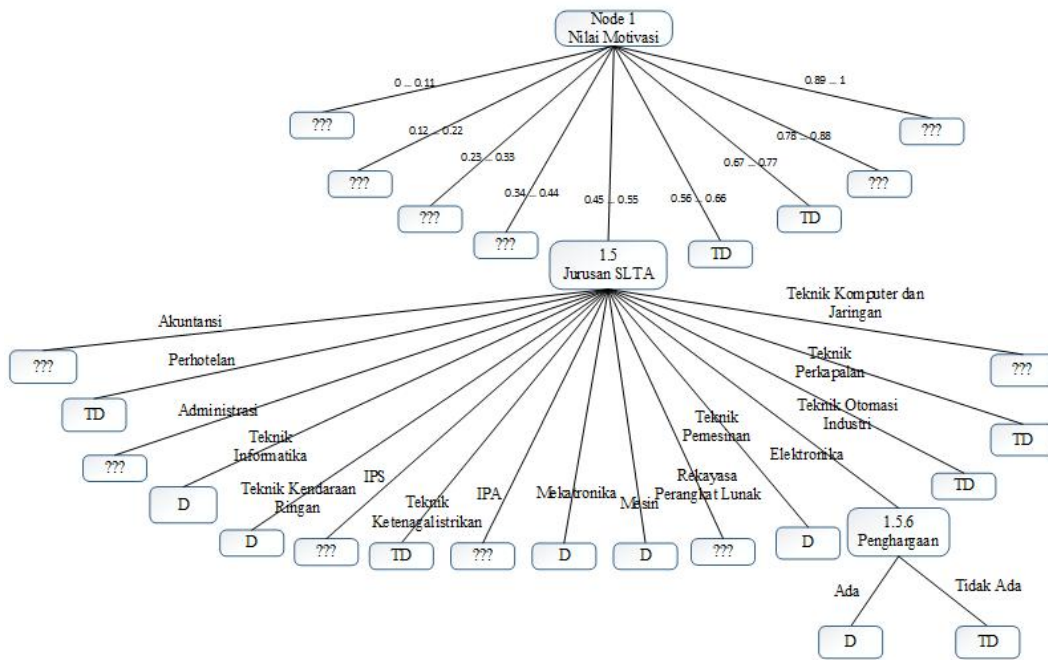


Gambar 68. Decision Tree Hasil Perhitungan Node 1.5.5.1

IV.1.55 Penentuan Node 1.5.6

Untuk menentukan *node* 1.5.6 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.55

Dari hasil Tabel A.1.55, diketahui bahwa atribut dengan *gain* tertinggi adalah **PENGHARGAAN**. Ada 2 *instance* dari **PENGHARGAAN**, yaitu **Ada** dan **Tidak Ada**. Nilai *instance* **Tidak Ada** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **Ada** mengklasifikasikan kasus menjadi **DITERIMA**.

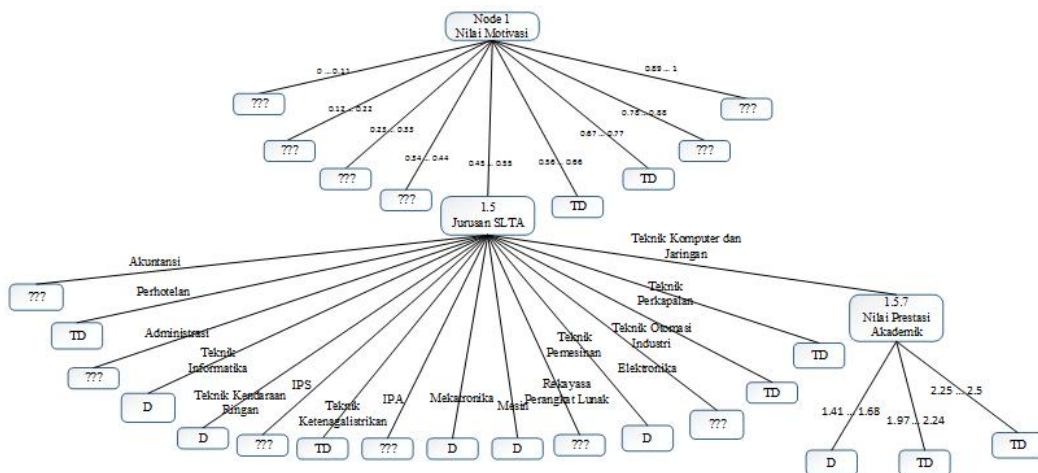


Gambar 69. Decision Tree Hasil Perhitungan Node 1.5.6

IV.1.56 Penentuan Node 1.5.7

Untuk menentukan *node* 1.5.7 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.56

Dari hasil Tabel A.1.56, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 3 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **1.41...1.68**, **1.97...2.24**, dan **2.25...2.5**. Nilai *instance* **1.97...2.24** dan **2.25...2.5** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **1.41...1.68** mengklasifikasikan kasus menjadi **DITERIMA**.

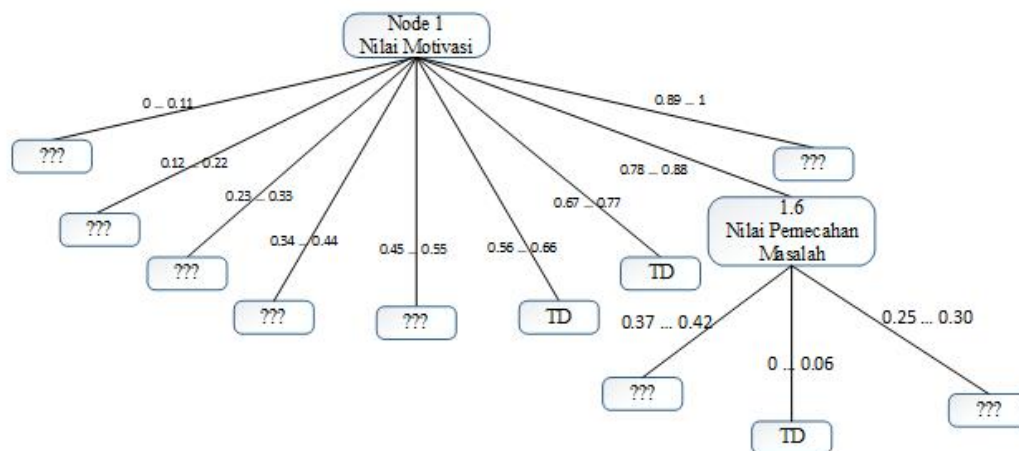


Gambar 70. Decision Tree Hasil Perhitungan Node 1.5.7

IV.1.57 Penentuan *Node* 1.6

Untuk menentukan *node* 1.6 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.57

Dari hasil Tabel A.1.57, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PEMECAHAN MASALAH**. Ada 3 *instance* dari **NILAI PEMECAHAN MASALAH**, yaitu **0.37...0.42**, **0...0.06**, dan **0.25...0.30**. Nilai *instance* **0...0.06** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **0.37...0.42** dan **0.25...0.30** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

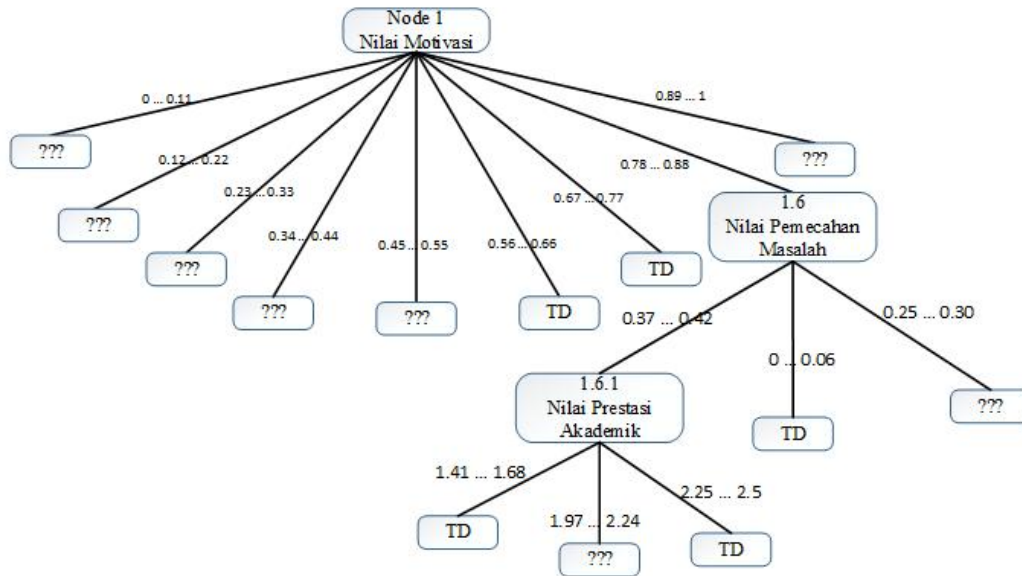


Gambar 71. *Decision Tree* Hasil Perhitungan *Node* 1.6

IV.1.58 Penentuan *Node* 1.6.1

Untuk menentukan *node* 1.6.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.58

Dari hasil Tabel A.1.58, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 3 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **1.41...1.68**, **1.97...2.24**, dan **2.25...2.5**. Nilai *instance* **1.41...1.68** dan **2.25...2.5** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **1.97...2.24** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

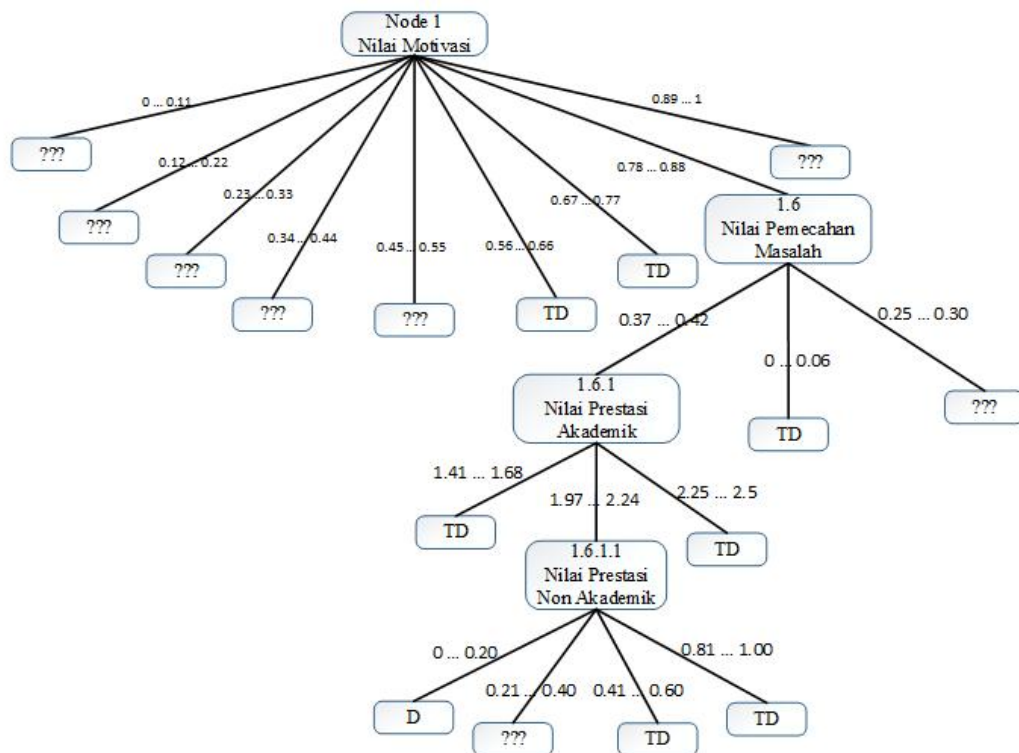


Gambar 72. Decision Tree Hasil Perhitungan Node 1.6.1

IV.1.59 Penentuan Node 1.6.1.1

Untuk menentukan *node* 1.6.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.59

Dari hasil Tabel A.1.59, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 4 *instance* dari **NILAI PRESTASI NON AKADEMIK**, yaitu **0...0.20**, **0.21...0.40**, **0.41...0.60**, dan **0.81...1.00**. Nilai *instance* **0.41...0.60**, dan **0.81...1.00** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **0...0.20** mengklasifikasikan kasus menjadi **DITERIMA**. *Instance* **0.21...0.40** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

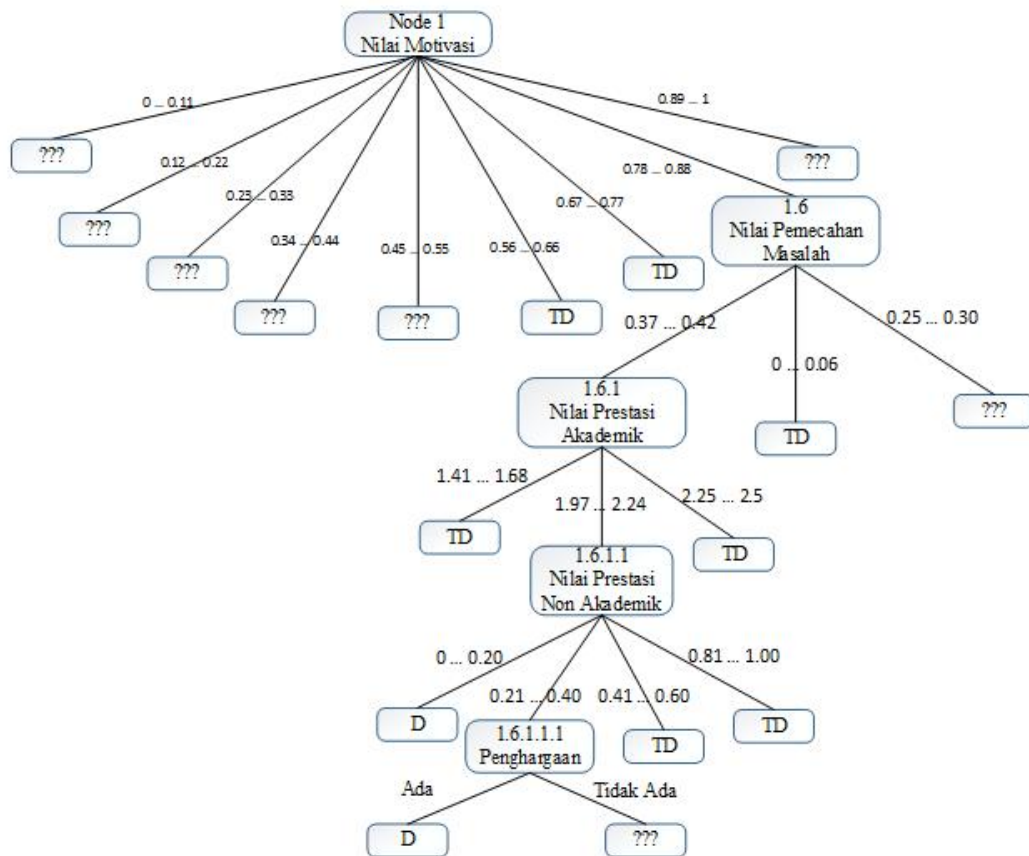


Gambar 73. Decision Tree Hasil Perhitungan Node 1.6.1.1

IV.1.60 Penentuan Node 1.6.1.1.1

Untuk menentukan *node* 1.6.1.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.60.

Dari hasil Tabel A.1.60, diketahui bahwa atribut dengan *gain* tertinggi adalah **PENGHARGAAN**. Ada 2 *instance* dari **PENGHARGAAN**, yaitu **Ada** dan **Tidak Ada**. Nilai *instance* **Ada** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **Tidak Ada** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

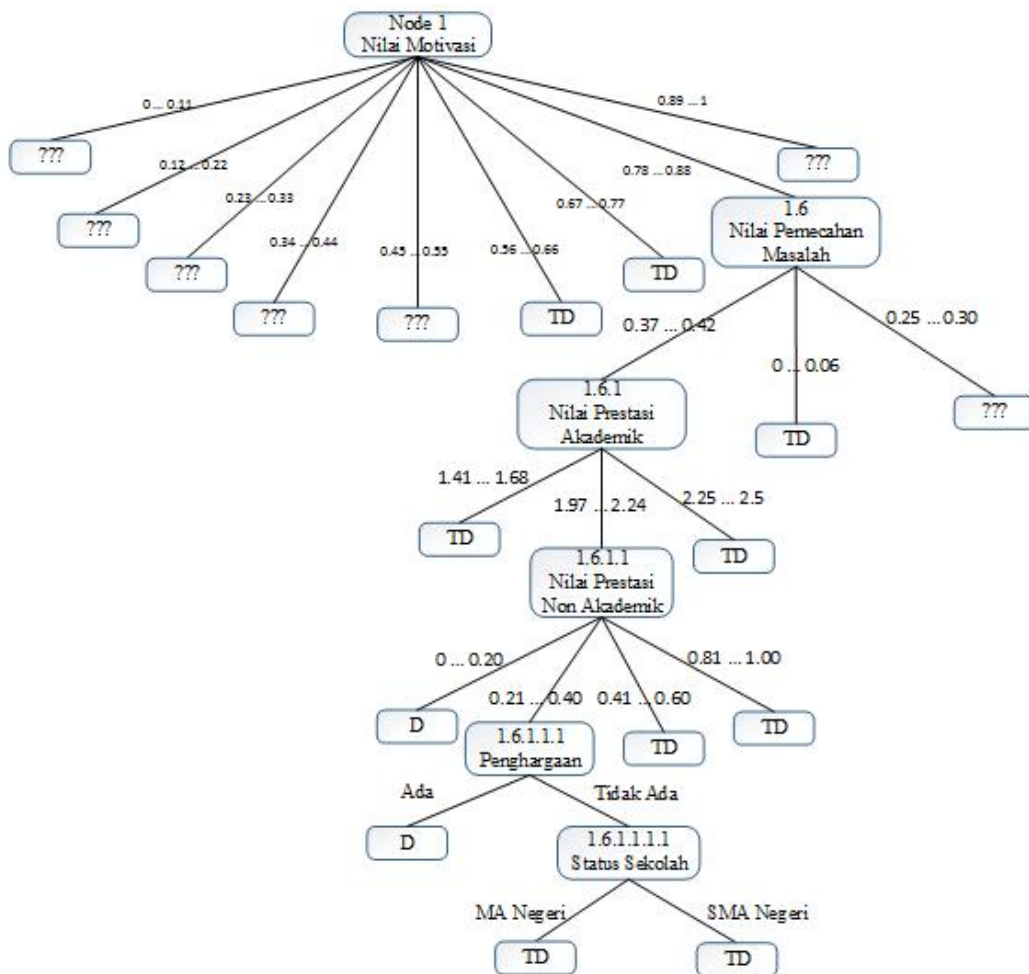


Gambar 74. Decision Tree Hasil Perhitungan Node 1.6.1.1.1

IV.1.61 Penentuan Node 1.6.1.1.1.1

Untuk menentukan *node* 1.6.1.1.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.61

Dari hasil Tabel A.1.61, diketahui bahwa atribut dengan *gain* tertinggi adalah **STATUS SEKOLAH**. Ada 2 *instance* dari **STATUS SEKOLAH**, yaitu **MA Negeri** dan **SMA Negeri**. Nilai *instance* **MA Negeri** dan **SMA Negeri** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**.

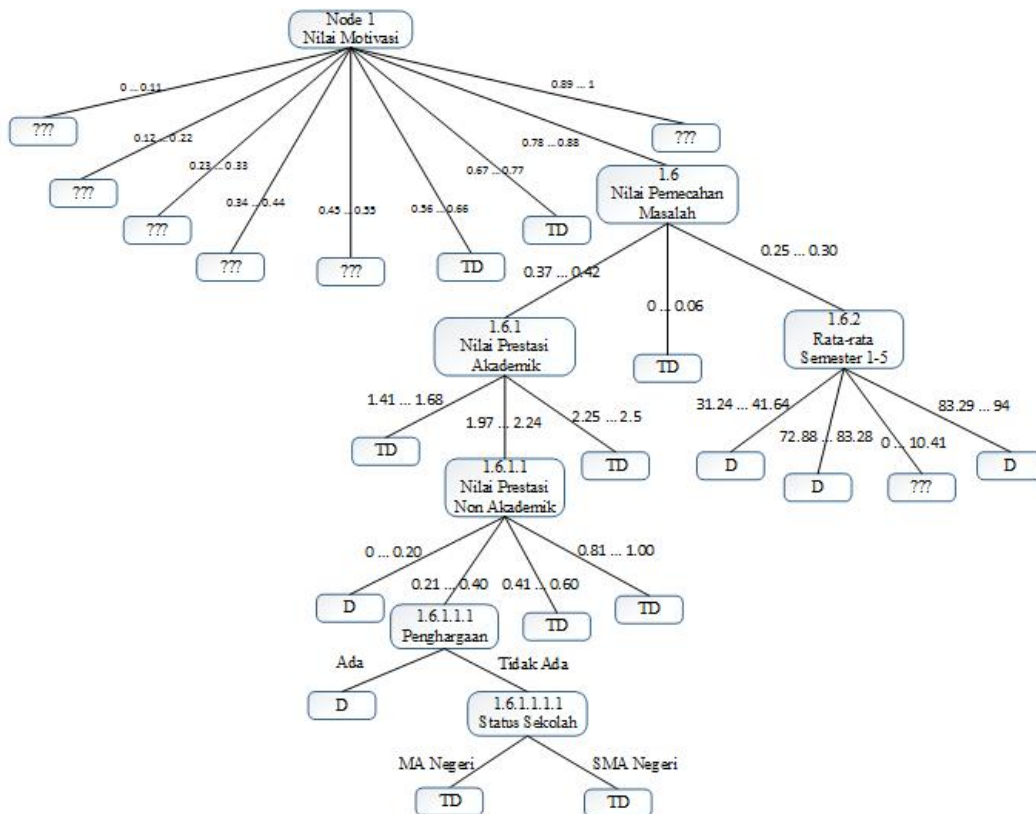


Gambar 75. Decision Tree Hasil Perhitungan Node 1.6.1.1.1.1

IV.1.62 Penentuan Node 1.6.2

Untuk menentukan node 1.6.2 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.62

Dari hasil Tabel A.1.62, diketahui bahwa atribut dengan *gain* tertinggi adalah **RATA-RATA SEMESTER 1-5**. Ada 4 *instance* dari **RATA-RATA SEMESTER 1-5**, yaitu **31.24...41.64**, **72.88...83.28**, **0...10.41**, dan **83.29...94**. Nilai *instance* **31.24...41.64**, **72.88...83.28** dan **83.29...94** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **0...10.41** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

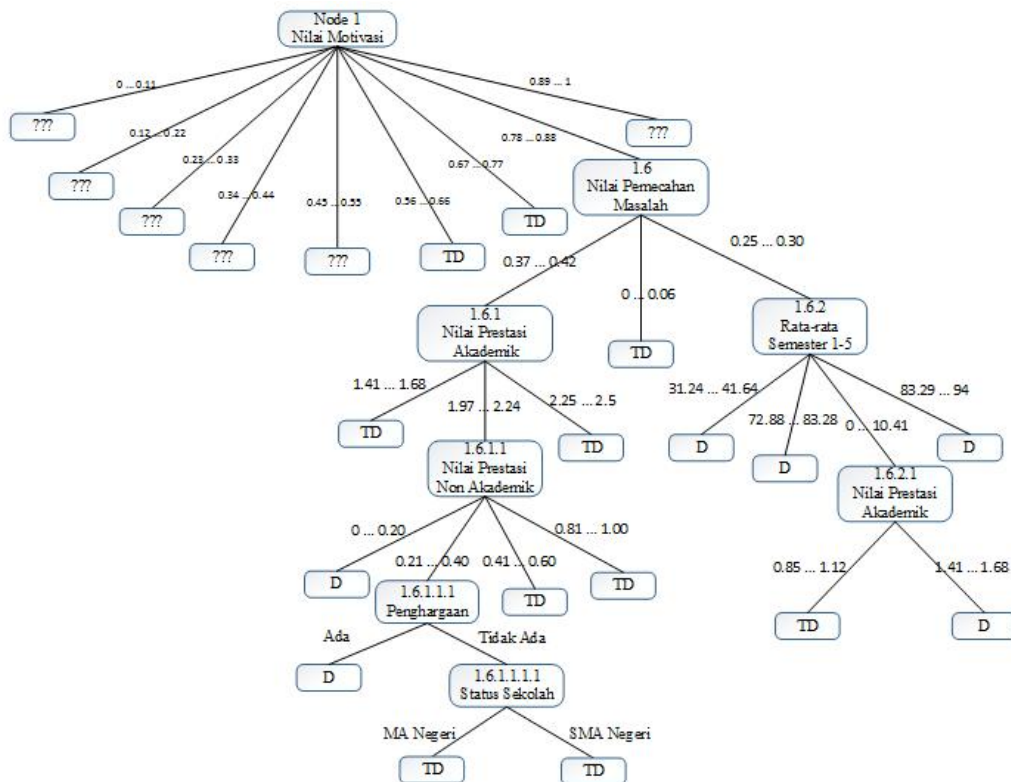


Gambar 76. Decision Tree Hasil Perhitungan Node 1.6.2

IV.1.63 Penentuan Node 1.6.2.1

Untuk menentukan *node* 1.6.2.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.63

Dari hasil Tabel A.1.63, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 2 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0.85...1.12** dan **1.41...1.68**. Nilai *instance* **0.85...1.12** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. Sedangkan *instance* **1.41...1.68** mengklasifikasikan kasus menjadi **DITERIMA**.

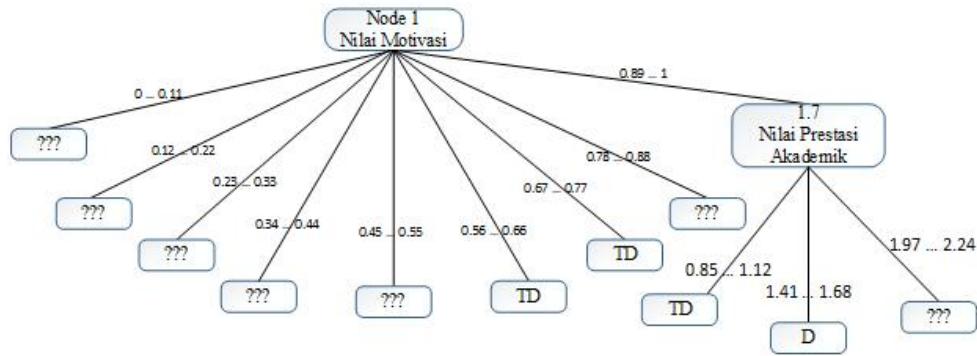


Gambar 77. Decision Tree Hasil Perhitungan Node 1.6.2.1

IV.1.64 Penentuan Node 1.7

Untuk menentukan *node 1.7* dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.64

Dari hasil Tabel A.1.64, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI AKADEMIK**. Ada 3 *instance* dari **NILAI PRESTASI AKADEMIK**, yaitu **0.85...1.12**, **1.41...1.68**, dan **1.97...2.24**. Nilai *instance* **0.85...1.12** sudah mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. *instance* **1.41...1.68** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **1.97...2.24** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

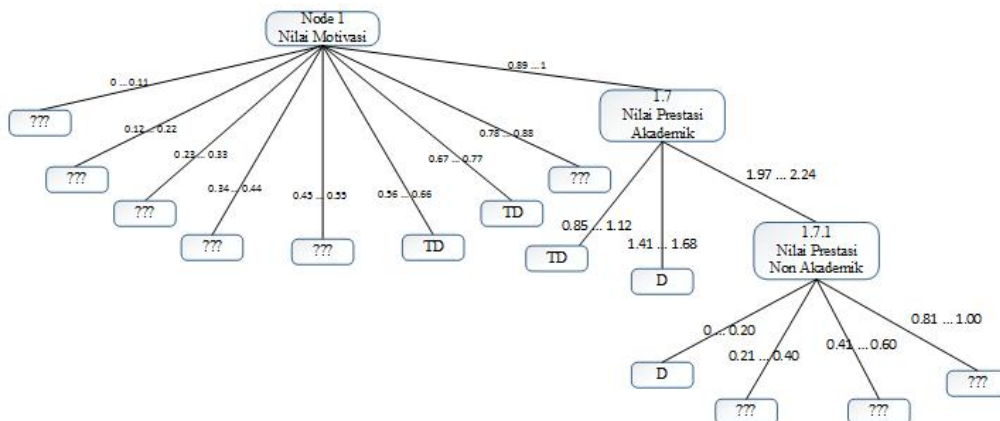


Gambar 78. Decision Tree Hasil Perhitungan Node 1.7

IV.1.65 Penentuan Node 1.7.1

Untuk menentukan node 1.7.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.65

Dari hasil Tabel A.1.65, diketahui bahwa atribut dengan *gain* tertinggi adalah **NILAI PRESTASI NON AKADEMIK**. Ada 4 *instance* dari **NILAI PRESTASI NON AKADEMIK**, yaitu **0...0.20**, **0.21...0.40**, **0.41...0.60**, dan **0.81...1.00**. Nilai *instance* **0...0.20** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **0.21...0.40**, **0.41...0.60**, dan **0.81...1.00** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

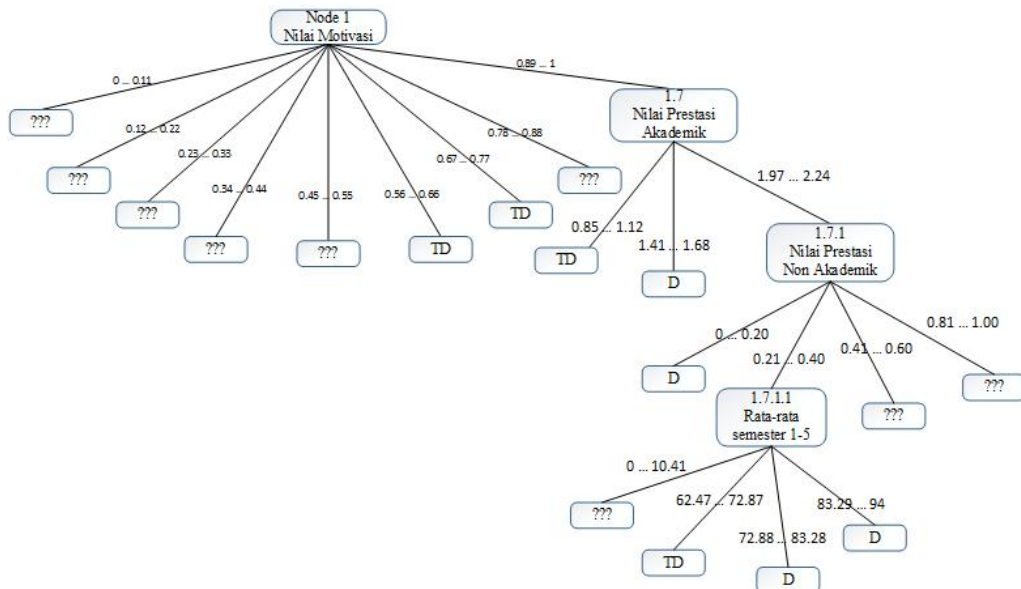


Gambar 79. Decision Tree Hasil Perhitungan Node 1.7.1

IV.1.66 Penentuan Node 1.7.1.1

Untuk menentukan *node* 1.7.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.66

Dari hasil Tabel A.1.66, diketahui bahwa atribut dengan *gain* tertinggi adalah **RATA-RATA SEMESTER 1-5**. Ada 4 *instance* dari **RATA-RATA SEMESTER 1-5**, yaitu **0...10.41**, **62.47...72.87**, **72..88...83.28**, dan **83.29...94**. Nilai *instance* **72..88...83.28**, dan **83.29...94** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **62.47...72.87** mengklasifikasikan kasus menjadi **TIDAK DITERIMA**. *Instance* **0...10.41** masih perlu dilakukan perhitungan lagi karena belum bisa mengklasifikasikan kasus.

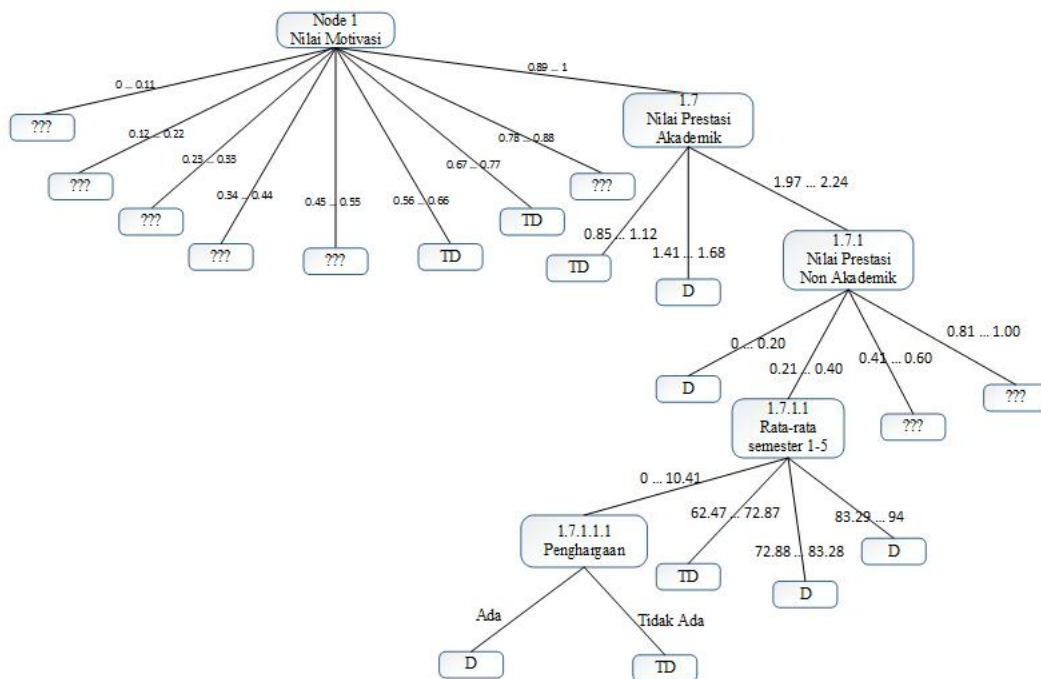


Gambar 80. Decision Tree Hasil Perhitungan Node 1.7.1.1

IV.1.67 Penentuan Node 1.7.1.1.1

Untuk menentukan *node* 1.7.1.1.1 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.67

Dari hasil Tabel A.1.67, diketahui bahwa atribut dengan *gain* tertinggi adalah **PENGHARGAAN**. Ada 2 *instance* dari **PENGHARGAAN**, yaitu **Ada** dan **Tidak Ada**. Nilai *instance* **Ada** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **Tidak Ada** mengklasifikasikan kasus menjadi **TIDAK DITERIMA**.

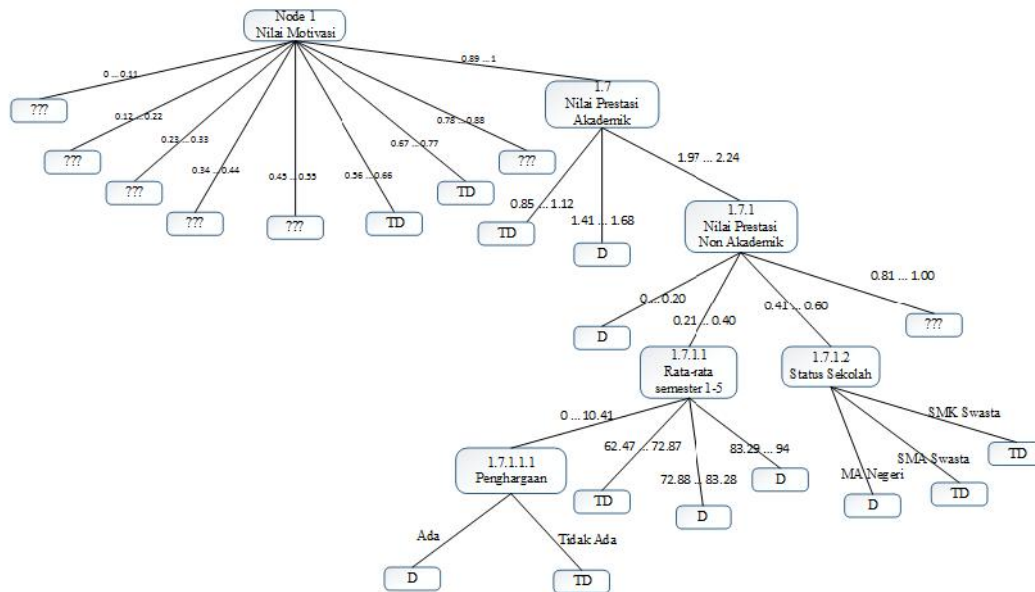


Gambar 81. Decision Tree Hasil Perhitungan Node 1.7.1.1.1

IV.1.68 Penentuan Node 1.7.1.2

Untuk menentukan node 1.7.1.2 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.68

Dari hasil Tabel A.1.68, diketahui bahwa atribut dengan *gain* tertinggi adalah **STATUS SEKOLAH**. Ada 3 instance dari **STATUS SEKOLAH**, yaitu **MA Negeri**, **SMA Swasta** dan **SMK Swasta**. Nilai instance **MA Negeri** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan instance **SMA Swasta** dan **SMK Swasta** mengklasifikasikan kasus menjadi **TIDAK DITERIMA**.

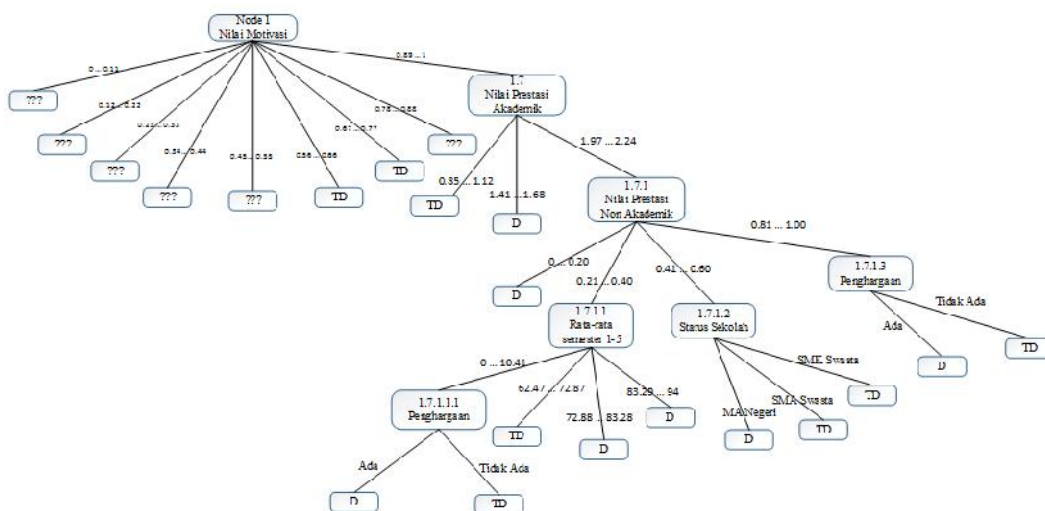


Gambar 82. Decision Tree Hasil Perhitungan Node 1.7.1.2

IV.1.69 Penentuan Node 1.7.1.3

Untuk menentukan *node* 1.7.1.3 dilakukan perhitungan *entropy* dan *gain*. Hasil perhitungan *entropy* dan *gain* dapat dilihat pada Tabel A.1.69

Dari hasil Tabel A.1.69, diketahui bahwa atribut dengan *gain* tertinggi adalah **PENGHARGAAN**. Ada 2 *instance* dari **PENGHARGAAN**, yaitu **Ada** dan **Tidak Ada**. Nilai *instance* **Ada** sudah mengklasifikasikan kasus menjadi **DITERIMA**. Sedangkan *instance* **Tidak Ada** mengklasifikasikan kasus menjadi **TIDAK DITERIMA**.



Gambar 83. Decision Tree Hasil Perhitungan Node 1.7.1.3

IV.2 Tingkat Error Rate Data

Setelah menghitung *entropy* dan *gain* maka akan menghasilkan *decision tree*. Decision tree kemudian diuji untuk mengetahui estimasi keakuratannya. Semakin sedikit *error rate* (kesalahan) yang dihasilkan dari *decision tree* maka semakin akurat *decision tree* yang dihasilkan. Untuk menghitung *error rate* digunakan persamaan 3. Untuk mengetahui berapa banyak *error data* silahkan lihat lampiran F.

$$E \quad r_i \quad (E) = \frac{b}{j_i} \frac{d}{h d} \frac{e}{h d} * 100\%$$

Penyelesaian:

Jumlah data = 68

Jumlah data error = 39

$$E \quad r_i \quad (E) = \frac{b}{j_i} \frac{d}{h d} \frac{e}{h d} * 100\%$$

$$Error \ rate \ (ER) = \frac{3}{6} * 100\%$$

$$Error \ rate \ (ER) = 57,35 \ %$$

BAB V

IMPLEMENTASI DAN PENGUJIAN

V.1 Algoritma

V.1.1 Algoritma Menyeleksi Peserta PMDK

Pemrograman Menyeleksi Peserta PMDK
{*untuk menyeleksi calon mahasiswa PMDK*}

Deklarasi:

Prodi Pilihan, Jurusan SLTA, Status Sekolah, Penghargaan, Rapor Semester, Nilai Motivasi, Nilai Prestasi Akademik, Nilai Prestasi Non Akademik, Nilai Pemecahan Masalah : varchar

Algoritma:

Input (Prodi Pilihan)

Input (Jurusan SLTA)

Input (Status Sekolah)

Input (Penghargaan)

Input (Rapor Semester)

Input (Nilai Motivasi)

Input (Nilai Prestasi Akademik)

Input (Nilai Prestasi Non Akademik)

Input (Nilai Pemecahan Masalah)

if Nilai Motivasi = 0.56-0.66 then
output (“Tidak Diterima”)

else if Nilai Motivasi = 0.67-0.77 then
output (“Tidak Diterima”)

else if Nilai Motivasi = 0-0.11 && Nilai Prestasi Akademik = 0-0.28 then
output (“Tidak Diterima”)

else if Nilai Motivasi = 0-0.11 && Nilai Prestasi Akademik = 1.41-1.68 then
output (“Tidak Diterima”)

else if Nilai Motivasi = 0-0.11 && Nilai Prestasi Akademik = 1.97-2.24 then
output (“Tidak Diterima”)

```

else if Nilai Motivasi = 0-0.11 && Nilai Prestasi Akademik = 2.25-2.5 then
    output (“Diterima”)

else if Nilai Motivasi = 0.12-0.22 && Nilai Prestasi Akademik = 0.29-0.56 then
    output (“Tidak Diterima”)

else if Nilai Motivasi = 0.12-0.22 && Nilai Prestasi Akademik = 1.41-1.68 then
    output (“Tidak Diterima”)

else if Nilai Motivasi = 0.12-0.22 && Nilai Prestasi Akademik = 1.97-2.24 then
    output (“Tidak Diterima”)

else if Nilai Motivasi = 0.12-0.22 && Nilai Prestasi Akademik = 2.25-2.5 then
    output (“Diterima”)

else if Nilai Motivasi = 0.23-0.33 && Jurusan SLTA = Multimedia then
    output (“Tidak Diterima”)

else if Nilai Motivasi = 0.23-0.33 && Jurusan SLTA = Rekayasa Perangkat Lunak then
    output (“Diterima”)

else if Nilai Motivasi = 0.23-0.33 && Jurusan SLTA = Teknik Informatika then
    output (“Tidak Diterima”)

else if Nilai Motivasi = 0.23-0.33 && Jurusan SLTA = Teknik Pendingin dan Tata
Udara then
    output (“Diterima”)

else if Nilai Motivasi = 0.23-0.33 && Jurusan SLTA = Teknik Perkapalan then
    output (“Tidak Diterima”)

else if Nilai Motivasi = 0.23-0.33 && Jurusan SLTA = Akuntansi && Rapor Semester
= 72.88-83.28 then
    output (“Tidak Diterima”)

.....// Lengkapnya mengikuti di lampiran

else if Nilai Motivasi = 0.89-1 && Nilai Prestasi Akademik = 1.97-2.24 && Nilai
Prestasi Non Akademik && Penghargaan = Tidak Ada then
    output (“Tidak Diterima”)

end if

```

V.1.2 Algoritma Mencari Informasi Peserta PMDK

Pemrograman Mencari Informasi Peserta PMDK

{untuk mencari informasi peserta pmdk berdasarkan id peserta yang dimasukkan pengguna}

Deklarasi:

tf_IdPeserta: int

tf_ProdiPilihan, tf_JurusanSLTA, tf_StatusSekolah, tf_Penghargaan, tf_RapotSem1,

tf_RapotSem2, tf_RapotSem3, tf_RapotSem4, tf_RapotSem5: String

Algoritma:

if

tf_IdPeserta == empty

then

output (“Silahkan masukkan ID peserta PMDK”)

else if

txtJudulBuku == true

then

output (tf_IdPeserta, tf_ProdiPilihan, tf_JurusanSLTA, tf_StatusSekolah, tf_Penghargaan,

tf_RapotSem1, tf_RapotSem2, tf_RapotSem3, tf_RapotSem4, tf_RapotSem5)

output (“Data Ditemukan!”)

else

then

output (“ID yang Anda masukkan tidak terdaftar”)

V.2 Implementasi Perancangan AntarMuka

V.2.1 Hasil Implementasi Perancangan Halaman Utama

Berikut adalah implementasi klasifikasi penerimaan calon mahasiswa PMDK:



Gambar 84. Hasil Implementasi Perancangan Halaman Utama

Berdasarkan gambar 84, bahwa halaman utama dari sistem yang memiliki *menu bar* yang terdiri dari menu dan bantuan. Untuk menu dan tentang memiliki *submenu* masing-masing. Untuk menu memiliki submenu yang terlihat pada gambar 85.



Gambar 85. Hasil Implementasi Halaman Utama Submenu Menu

Pada gambar 85, dapat dilihat bahwa menu terdiri dari 4 submenu yang jika di klik akan berpindah ke halaman yang diklik. Misalnya jika pengguna mengklik submenu informasi peserta PMDK, maka sistem akan menampilkan form yang berisi informasi peserta PMDK. Menu bantuan juga memiliki submenu sama halnya dengan menu seperti yang terlihat pada gambar 86.



Gambar 86. Hasil Implementasi Halaman Utama Submenu Tentang

Pada gambar 86, dapat dilihat bahwa menu tentang memiliki *submenu* petunjuk yang jika di klik pengguna, akan menampilkan antarmuka petunjuk penggunaan aplikasi.

V.2.2 Hasil Implementasi Perancangan Penyeleksian PMDK

Hasil implementasi *input* penyeleksian PMDK dapat dilihat seperti pada gambar 87.

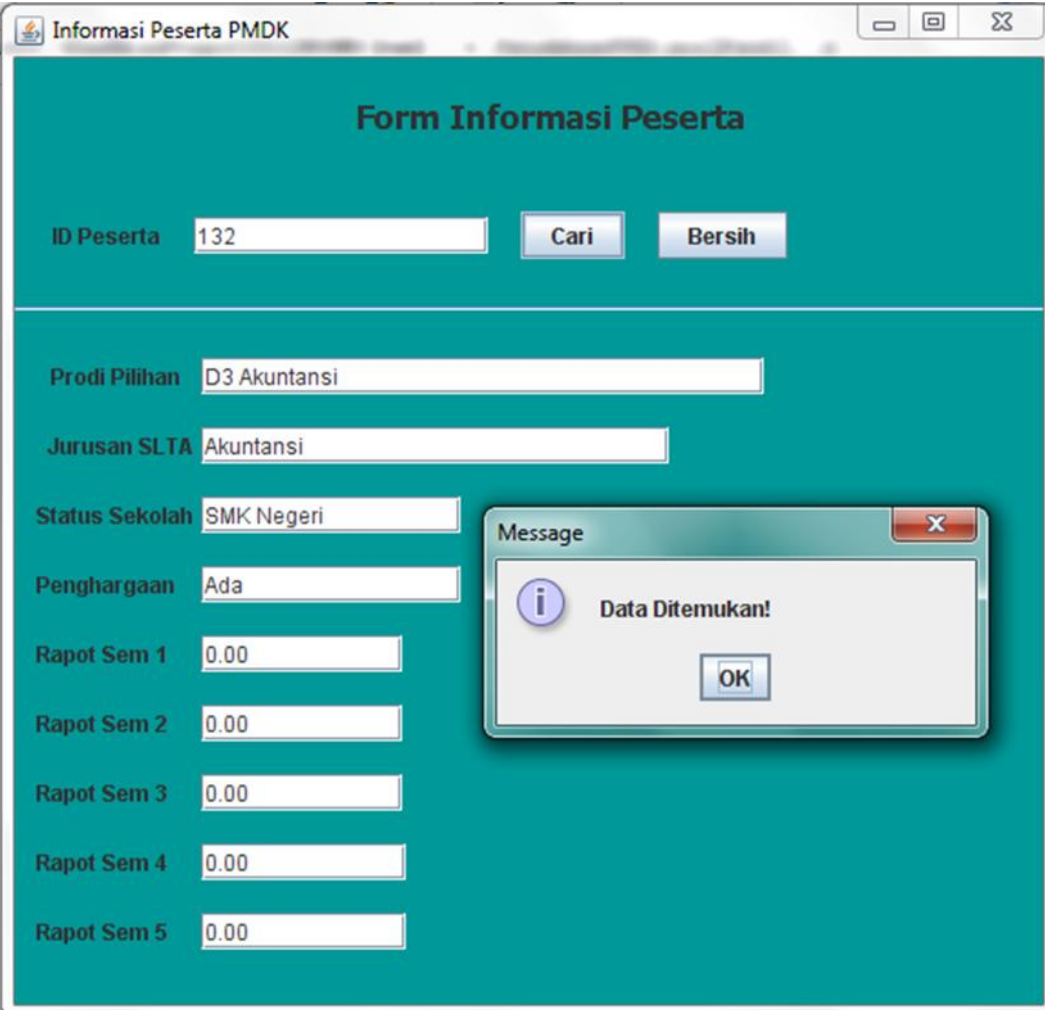
Form Penyeleksian PMDK	
Nilai Motivasi	0.23-0.33
Nilai Prestasi Akademik	Pilih
Jurusan SLTA	Akuntansi
Penghargaan	Pilih
Nilai Pemecahan Masalah	Pilih
Prodi Pilihan	Pilih
Status Sekolah	Pilih
Nilai Prestasi Non Akademik	Pilih
Rapor Semester 1.5	72.88-83.28
<input type="button" value="Hasil"/> <input type="button" value="Bersih"/>	
Output (Hasil)	Tidak Diterima
Aturan Klasifikasi F Nilai Motivasi = 0.23-0.33 * Nilai Prestasi Akademik = Akuntansi * Nilai Rata-rata Rapor Semes	

Gambar 87. Hasil Implementasi Perancangan Penyeleksian PMDK

Pada gambar 87, dapat dilihat bahwa apabila pengguna memilih salah satu nilai motivasi yang sudah ada di *combo box*. Jika *combo box* nilai motivasi sudah terpilih maka akan muncul pesan berwarna merah, gunanya supaya pengguna dapat mengikuti langkah selanjutnya. Setelah pengguna mengikuti langkah sampai selesai biasanya akan muncul pesan berwarna merah supaya pengguna menekan tombol hasil, maka pengguna menekan tombol hasil, maka sistem akan menampilkan hasil *rule* klasifikasi yang menyatakan apakah calon mahasiswa PMDK diterima atau tidak diterima.

V.2.3 Hasil Implementasi Perancangan Informasi Peserta PMDK

Hasil implementasi perancangan *input* informasi peserta dapat dilihat pada gambar 88.



Gambar 88. Hasil Implementasi Perancangan Informasi Peserta PMDK

Berdasarkan pada gambar 88, dapat dilihat pengguna memasukkan salah satu Id Peserta “132” pada *text field* id peserta. Selanjutnya pengguna menekan tombol cari pada halaman informasi peserta sehingga informasi peserta tampil seperti yang terlihat pada gambar 88.

V.2.4 Hasil Implementasi Perancangan Tentang Pembuat Aplikasi

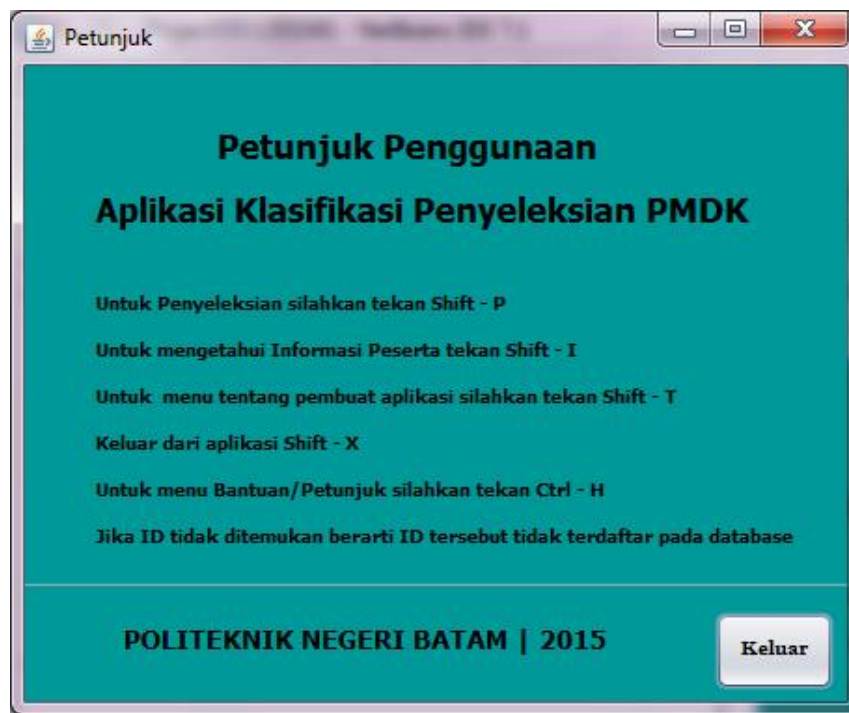
Hasil implementasi perancangan tentang pembuat aplikasi dapat dilihat pada gambar 89.



Gambar 89. Hasil Implementasi Perancangan Tentang Pembuat Aplikasi

Berdasarkan gambar 89, dapat dilihat bahwa pengguna mengklik *submenu* tentang pembuatan aplikasi maka akan tampil antarmuka tentang pembuat aplikasi. Jika pengguna menekan tombol keluar maka sistem akan keluar dari aplikasi penerimaan calon mahasiswa PMDK.

V.2.5 Hasil Implementasi Perancangan Tentang Pembuat Aplikasi



Gambar 90. Perancangan Tentang Pembuat Aplikasi

V.3 Pengujian

V.3.1 Strategi Pengujian

Strategi pengujian sistem yang menerapkan analisis data *mining* ini menggunakan metode *black box* yaitu berfokus pada kebutuhan deskripsi fungsional guna untuk mengetahui jalannya sistem secara lengkap.

V.3.2 Skenario Pengujian

Skenario pengujian sistem yang menerapkan data *mining* ini dijelaskan seperti pada gambar 90.



Gambar 91. Skenario Pengujian

Berdasarkan gambar 91, dapat dilihat bahwa skenario pengujian yang akan dilakukan pada sistem ada 2 yaitu pengujian menyeleksi peserta PMDK, mencari informasi peserta.

V.3.3 Hasil Pengujian

Pengujian dilaksanakan dengan baik sesuai metode *black box* yang digunakan dan hasilnya dapat dilihat pada tabel 10.

Tabel 10. Hasil Pengujian Sistem

No	Use Case	Fungsi	Skenario		Data Uji	Target	Pengujian	Penguji
1	Menyeleksi PMDK	Untuk menyeleksi peserta PMDK	1. Memilih nilai motivasi pada <i>combo box</i> 2. Mengikuti langkah-langkah pesan berwarna merah sampai muncul pesan "silahkan untuk menekan tombol hasil" 3. Menekan tombol hasil		Nilai motivasi : 0.23-0.33 Jurusan SLTA : Akuntansi Rapor sem 1-5 : 72.88-83.28	Sistem akan menampilkan hasil "tidak diterima" beserta dengan <i>rule</i> klasifikasinya.	OK	Nama: Febriandika Umur: 21 Tahun Jenis Kelamin: Perempuan
2.	Mencari informasi peserta	Untuk mencari informasi peserta	1. Pengguna memasukkan Id Peserta ke dalam <i>text field</i> id peserta 2. Menekan tombol cari		Data Benar: 132 Data salah: 001	Sistem akan menampilkan "data ditemukan" Sistem akan menampilkan "ID yang anda masukkan tidak terdaftar"	OK	Nama: Febriandika Umur: 21 Tahun Jenis Kelamin: Perempuan

BAB VI

KESIMPULAN DAN SARAN

VI.1 Kesimpulan

Berdasarkan analisis yang dilakukan dalam mengklasifikasikan penerimaan calon mahasiswa PMDK, ada beberapa hal yang disimpulkan, yaitu:

1. Tugas Akhir ini menemukan *rule* klasifikasi sebanyak 113 *rule* yang sesuai dengan dengan *decision tree*.
2. Tingkat *correct rate* dari *rule* yang dihasilkan sebesar 47.06%. hal ini disebabkan kurang tepat dalam pemilihan atribut.
3. Berdasarkan *rule* yang didapat maka dapatlah diimplementasikan kedalam bentuk visual berupa aplikasi penyeleksian calon mahasiswa PMDK. Dengan fungsi yang dapat menghasilkan *rule* berupa *output* diterima atau tidak diterima.

VI.2 Saran

Adapun saran pengembangan dari penelitian yang dibuat dan aspek yang belum terselesaikan, yaitu:

1. Pada penelitian selanjutnya diharapkan tepat dalam pemilihan atribut maupun *class label* pada data operasional sehingga untuk proses *training* dan *testing* lebih meningkat dan akurat.
2. Sebaiknya menggunakan WEKA dalam melakukan klasifikasi data berikutnya.
3. Pengembangan aplikasi selanjutnya bisa menangani men-*generate* semua data, sehingga data yang dimasukkan tidak satu-persatu diinput kedalam aplikasi.
4. Sebaiknya menerapkan *pruning trees* (pemangkasan pohon).

DAFTAR PUSTAKA

- Adhatrao, K., dkk, 2013, *PREDICTING STUDENTS' PERFORMANCE USING ID3 AND C4.5 CLASSIFICATION ALGORITHMS*, International Journal of Data Mining & Knowledge Management Process (IJDKP).
- Ahmed A. B. E. D. dan Elaraby, E. S., 2013, "Using Data Mining Technique to Analyze Student's Performance," *International Journal of Research in Education Methodolgy*, vol. 5, no. 2, p. 7.
- Ananda, S. dan Rouf, I., 2011, *Buku Pintar Menguasai Microsoft Excel Untuk Pemula*, MediaKita, Jakarta.
- Arbie, 2004, *Managemen database dengan MySQL*, Yogyakarta, Andi.
- Han, J., & Kamber, M., 2006, *Data Mining: Concepts and Techniques*, San Francisco, Services Manager Simon Crump.
- Hastuti, K., 2012, *ANALISIS KOMPARASI ALGORITMA KLASIFIKASI DATA MINING UNTUK PREDIKSI MAHASISWA NON AKTIF*, Seminar Nasional Teknologi Informasi & Komunikasi Terapan.
- <http://www.polibatam.ac.id/penerimaan-mahasiswa-baru/pmdk/>. Diakses 17 September 2014.
- Jogiyanto. 2001. *Analisis & Disain Sistem Informasi: Pendekatan Terstruktur Teori dan Praktek Aplikasi Bisnis*. Yogyakarta: Andi.
- Kusrini dan Taufiq, Emha. 2009. *Algoritma Data mining*. Yogyakarta: Andi Offset.
- Larose, D. T., 2005, *Discovering Knowledge in Data: An Introduction to Data Mining*, New Jersey, John Wiley & Sons, Inc.
- Madcoms, 2011, *Pasti Bisa!! Belajar Sendiri Microsoft Excel 2010*, Yogyakarta, Andi Offset.
- Pascal, S. A., 2007, *Tip & Trik Microsoft Office 2007*, Elex Media Komputindo, Jakarta.
- Supranto, J. 2003. *Metode Riset: Aplikasinya dalam Pemasaran*. Jakarta: Rineka Cipta.
- Triandini, E. dan Suardika, I. G., 2012. *Step By Step Desain Proyek Menggunakan UML*, Andi Offset, Yogyakarta.

Turban, E., dkk. 2005. *Decicion Support System and Intelegant Systems*. Yogyakarta: Andi Offset.

Whitten, J. L., dkk, 2002, *Metode Desain & Analis Sistem Ed. 6*, McGraw Hill Education.

Witten, Ian H., and Eibe Frank. 2000. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. San Diego, CA: Morgan Kaufmann.

Zega, A. S., & Agniesari, A., 2011, *Analisis Kualitas Mahasiswa Politeknik Negeri Batam Program Studi Teknik Informatika Berdasarkan Jalur Masuk Kuliah*, Batam