

SKRIPSI

**PREDIKSI INDEKS PRESTASI SEMESTER (IPS) MAHASISWA
PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER
UNIVERSITAS NEGERI JAKARTA SELAMA PEMBELAJARAN
DARING MENGGUNAKAN ALGORITMA NAIVE BAYES**



*Mencerdaskan dan
Memartabatkan Bangsa*

MIFTAHUL JANNAH

1512619027

Disusun untuk Memenuhi Persyaratan
Guna Memperoleh Gelar Sarjana Pendidikan

**PROGRAM STUDI PENDIDIKAN TEKNIK INFORMATIKA DAN
KOMPUTER**

FAKULTAS TEKNIK

UNIVERSITAS NEGERI JAKARTA

2023

**PREDIKSI INDEKS PRESTASI SEMESTER (IPS) MAHASISWA
PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER
UNIVERSITAS NEGERI JAKARTA SELAMA PEMBELAJARAN
DARING MENGGUNAKAN ALGORITMA NAIVE BAYES**

MIFTAHUL JANNAH

ABSTRAK

Penelitian ini bertujuan untuk mengetahui performa model prediksi dengan menggunakan teknik klasifikasi data mining dalam memprediksi Indeks Prestasi Semester (IPS) mahasiswa Pendidikan Teknik Informatika dan Komputer Universitas Negeri Jakarta selama melaksanakan pembelajaran daring sehingga dapat bermanfaat bagi mahasiswa dan perguruan tinggi dalam merencanakan strategi pembelajaran yang baik pada semester berikutnya. Atribut yang digunakan dalam penelitian ini adalah aspek psikologis dan beberapa atribut lainnya yang mempengaruhi prestasi akademik mahasiswa dalam melaksanakan pembelajaran daring. Prediksi akan menggunakan algoritma Naive Bayes karena merupakan algoritma klasifikasi yang sering digunakan dalam melakukan prediksi. *Dataset* yang digunakan pada penelitian ini diperoleh dari pengumpulan kuesioner terhadap mahasiswa Pendidikan Teknik Informatika dan Komputer Universitas Negeri Jakarta angkatan 2020 dan 2021. Jumlah data yang diperoleh adalah sebanyak 155 *record* dengan 13 (tiga belas) atribut dengan rincian 1 (satu) atribut ID yaitu NIM, 11 (sebelas) atribut reguler yaitu jenis kelamin, jalur masuk, fasilitas smartphone, kondisi jaringan di daerahnya, aplikasi daring yang disukai, minat belajar saat pembelajaran daring, sikap belajar, kreativitas belajar, dukungan orang tua, kelompok belajar (diskusi online), dan kegiatan lain di luar perkuliahan selama pembelajaran daring, serta 1 (satu) atribut label kelas yaitu Indeks Prestasi Semester mahasiswa di semester 117. Penelitian ini memperoleh evaluasi kinerja algoritma dengan pengujian menggunakan *confusion matrix* dan kurva ROC (*Receiver Operating Characteristic*) menggunakan *tools* RapidMiner. Hasil evaluasi *confusion matrix* yaitu *accuracy* sebesar 75 %, *precision* sebesar 28,33 %, dan *recall* sebesar 26,43 %. Sedangkan evaluasi dengan menggunakan kurva ROC menghasilkan nilai AUC sebesar 0,679 yang dikategorikan sebagai klasifikasi yang buruk. Penelitian ini juga menerapkan teknik penyeimbang data SMOTE guna menghasilkan data yang seimbang sehingga dapat dihasilkan akurasi yang lebih baik. Hasil evaluasi yang diperoleh dengan *confusion matrix* menghasilkan *accuracy* sebesar 88,46 %, *precision* sebesar 57,43 %, dan *recall* sebesar 52,14 %. Sedangkan evaluasi kurva ROC dengan SMOTE menghasilkan nilai AUC sebesar 0,809 yang dikategorikan sebagai klasifikasi yang baik.

Kata kunci: Prediksi, Data Mining, Naive Bayes, Pembelajaran Daring, Indeks Prestasi

**PREDICTION OF GRADE POINT AVERAGE (GPA) STUDENTS OF
INFORMATICS AND COMPUTER ENGINEERING JAKARTA STATE
UNIVERSITY DURING ONLINE LEARNING USING NAIVE BAYES
ALGORITHM**

MIFTAHUL JANNAH

ABSTRACT

This study aims to determine the performance of the prediction model using data mining classification techniques in predicting the Semester Grade Point Average (GPA) of Informatics and Computer Engineering Education students at Jakarta State University while carrying out online learning so that it can be useful for students and universities in planning good learning strategies in the next semester. The attributes used in this study are psychological aspects and several other attributes that affect students' academic achievement in carrying out online learning. The prediction will use the Naive Bayes algorithm because it is a classification algorithm that is often used in making predictions. Dataset used in this study was obtained from collecting questionnaires from Informatics and Computer Engineering Education students at Jakarta State University class of 2020 and 2021. The total data obtained was 155 records with 13 (thirteen) attributes with details of 1 (one) ID attribute, namely NIM, 11 (eleven) regular attributes, namely gender, university entrance selection, smartphone facilities, network conditions in the area, preferred online applications, interest in learning when online learning, learning attitudes, learning creativity, parental support, study groups (online discussion) , and other activities outside of lectures during online learning, as well as 1 (one) class label attribute, namely the Semester Grade Point Average (GPA) for students in semester 117. This study obtained an algorithm performance evaluation by testing using the confusion matrix and the ROC (Receiver Operating Characteristic) curve using RapidMiner tools. The results of the evaluation of the confusion matrix are accuracy of 75%, precision of 28.33%, and recall of 26.43%. The evaluation using the ROC curve produces an AUC value of 0.679 which is categorized as a bad classification. This study also applies the SMOTE data balancing technique to produce balanced data so that better accuracy can be produced. The evaluation results obtained with the confusion matrix produce an accuracy of 88.46%, a precision of 57.43%, and a recall of 52.14%. Meanwhile, the evaluation of the ROC curve with SMOTE produced an AUC value of 0.809 which was categorized as a good classification.

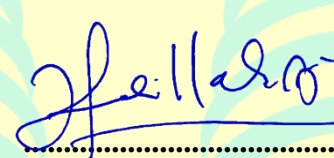
Keywords: Prediction, Data Mining, Naive Bayes, Online Learning, Grade Point Average

LEMBAR PENGESAHAN
PREDIKSI INDEKS PRESTASI SEMESTER (IPS) MAHASISWA
PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER
UNIVERSITAS NEGERI JAKARTA SELAMA PEMBELAJARAN
DARING MENGGUNAKAN ALGORITMA NAIVE BAYES

Miftahul Jannah, NIM. 1512619027

NAMA DOSEN	TANDA TANGAN	TANGGAL
------------	--------------	---------

Dr. Widodo, S.Kom, M.Kom (Dosen Pembimbing I)		24 Juli 2023
--	---	--------------

Hamidillah Ajie, S.Si, M.T (Dosen Pembimbing II)		24 Juli 2023
---	--	--------------

PENGESAHAN PANITIA UJIAN SKRIPSI

NAMA DOSEN	TANDA TANGAN	TANGGAL
------------	--------------	---------

Diat Nurhidayat, S.Pd., MTI (Ketua Penguji)		14 Juli 2023
--	---	--------------

Murien Nugraheni, M.Cs (Dosen Penguji I)		13 Juli 2023
---	---	--------------

Ali Idrus, S.Kom, M.Kom Dosen Penguji II		17 Juli 2023
---	--	--------------

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Skripsi ini merupakan Karya asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di Perguruan Tinggi lain.
2. Skripsi ini belum dipublikasikan, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
3. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan tidak benaran, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, 27 Juni 2023

Yang membuat pernyataan



Miftahul Jannah

No. Reg. 1512619027



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS NEGERI JAKARTA
UPT PERPUSTAKAAN

Jalan Rawamangun Muka Jakarta 13220
Telepon/Faksimili: 021-4894221
Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Miftahul Jannah
NIM : 1512619027
Fakultas/Prodi : Teknik/Pendidikan Teknik Informatika dan Komputer
Alamat email : miftahuljnnh294@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada UPT Perpustakaan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

Skripsi Tesis Disertasi Lain-lain (.....)

yang berjudul :

Prediksi Indeks Prestasi Semester (IPS) Mahasiswa Pendidikan Teknik Informatika dan Komputer Universitas Negeri Jakarta Selama Pembelajaran Daring Menggunakan Algoritma Naive bayes

Dengan Hak Bebas Royalti Non-Eksklusif ini UPT Perpustakaan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 25 Agustus 2023

Penulis

(Miftahul Jannah)
nama dan tanda tangan

KATA PENGANTAR

Puji dan syukur saya panjatkan atas kehadiran Allah SWT, atas rahmat dan karunia-Nya sehingga skripsi ini yang berjudul “Prediksi Indeks Prestasi Semester (IPS) Mahasiswa Pendidikan Teknik Informatika Dan Komputer Universitas Negeri Jakarta Selama Pembelajaran Daring Menggunakan Algoritma Naive Bayes” dapat terselesaikan dengan lancar sebagai salah satu persyaratan dalam menyelesaikan studi Strata Satu (S-1) dan memperoleh gelar sarjana pada Program Studi Pendidikan Teknik Informatika dan Komputer Universitas Negeri Jakarta.

Skripsi ini tidak dapat terselesaikan dengan baik tanpa adanya bantuan dan bimbingan dari berbagai pihak yang memberikan banyak masukan serta motivasi kepada peneliti. Maka pada kesempatan ini, peneliti menyampaikan ucapan terima kasih kepada:

1. Bapak Widodo, S.Kom, M.Kom selaku Koordinator Program Studi Pendidikan Teknik Informatika dan Komputer, Fakultas Teknik, Universitas Negeri Jakarta.
2. Bapak Widodo, S.Kom, M.Kom selaku pembimbing I dan Bapak Hamidillah Ajie, S.Si, M.T selaku pembimbing II yang telah memberikan arahan, bimbingan, dan kepercayaan kepada peneliti dalam menyelesaikan skripsi ini.
3. Ibu Yusmawati, sebagai sumber motivasi bagi peneliti, yang senantiasa selalu memanjatkan doa serta memberikan motivasi dan semangatnya dalam menyelesaikan skripsi ini.
4. Bapak Hendaya Tanupus yang selalu memberikan dukungannya baik moral dan materiil bagi peneliti dalam menyelesaikan skripsi ini.
5. Najwa Ramadhani dan Lukmanul Hakim yang selalu memberikan dukungan, bantuan dan motivasinya.
6. Siti Nabila Suryaningsih, Intan Ramadhanti, dan Fatih Abhipraya yang selalu bersedia memberikan dukungan, semangat, dan dorongannya kepada peneliti selama menyelesaikan skripsi ini.

7. Anggota NCT Dream, Day6 dan Zerobaseone atas karya-karyanya yang telah menemani selama pengerjaan skripsi dan menjadi inspirasi untuk berani menuju masa depan.
8. Mahasiswa PTIK UNJ angkatan 2019, 2020, dan 2021 yang telah memberikan bantuan dan kesediaannya untuk ikut serta dalam penyelesaian skripsi ini.
9. Seluruh pihak yang telah membantu peneliti selama proses penyusunan skripsi tetapi tidak dapat disebutkan satu persatu.

Peneliti menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, peneliti mohon maaf apabila terdapat kekurangan serta mengharapkan kritik dan saran yang dapat membangun agar skripsi ini dapat menjadi lebih baik.

Jakarta, 26 Juni 2023

Peneliti,

Miftahul Jannah

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN	iii
KATA PENGANTAR.....	iv
ABSTRAK	vi
ABSTRACT.....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Batasan Masalah.....	4
1.4 Rumusan Masalah	5
1.5 Tujuan Penelitian.....	5
1.6 Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Landasan Teori	6
2.1.1 Prestasi Akademik Mahasiswa	6
2.1.2 Indeks Prestasi Semester (IPS)	8
2.1.3 Pembelajaran Daring.....	9
2.1.4 Prediksi	10
2.1.5 <i>Data Mining</i>	11
2.1.6 <i>Knowledge Discovery in Database (KDD)</i>	12
2.1.7 Klasifikasi	14
2.1.8 Naive Bayes	15
2.1.9 <i>Information Gain</i>	18
2.1.10 <i>Confusion Matrix</i>	18
2.1.11 Kurva ROC (<i>Receiver Operating Characteristic</i>).....	20

2.1.12 RapidMiner	21
2.1.13 <i>Imbalance</i> Data	22
2.1.14 SMOTE.....	23
2.2 Penelitian yang Relevan	23
2.3 Kerangka Berpikir	29

BAB III METODOLOGI PENELITIAN 31

3.1 Tempat dan Waktu Penelitian	31
3.2 Alat dan Bahan Penelitian	31
3.2.1 Alat Penelitian.....	31
3.2.2 Bahan Penelitian	31
3.3 Diagram Alir Penelitian.....	36
3.3.1 Merapikan Data Mahasiswa.....	36
3.3.2 Seleksi Atribut	36
3.3.3 Pembersihan Data	36
3.3.4 Penyeimbang Data	37
3.3.5 Pemodelan.....	37
3.3.6 Evaluasi dan Validasi hasil.....	37
3.4 Teknik dan Prosedur Pengumpulan Data	38
3.5 Teknik Analisis Data	38

BAB IV HASIL PENELITIAN..... 41

4.1 Deskripsi Hasil Penelitian	41
4.1.1 Merapikan Data Mahasiswa.....	42
4.1.2 Seleksi Atribut	45
4.1.3 Pembersihan Data	48
4.1.4 Pemodelan.....	50
4.1.5 Evaluasi dan Validasi.....	53
4.1.6 Pengujian Menggunakan Teknik SMOTE.....	61
4.2 Analisis Data Penelitian	67
4.3 Pembahasan	70
4.4 Aplikasi Hasil Penelitian	71

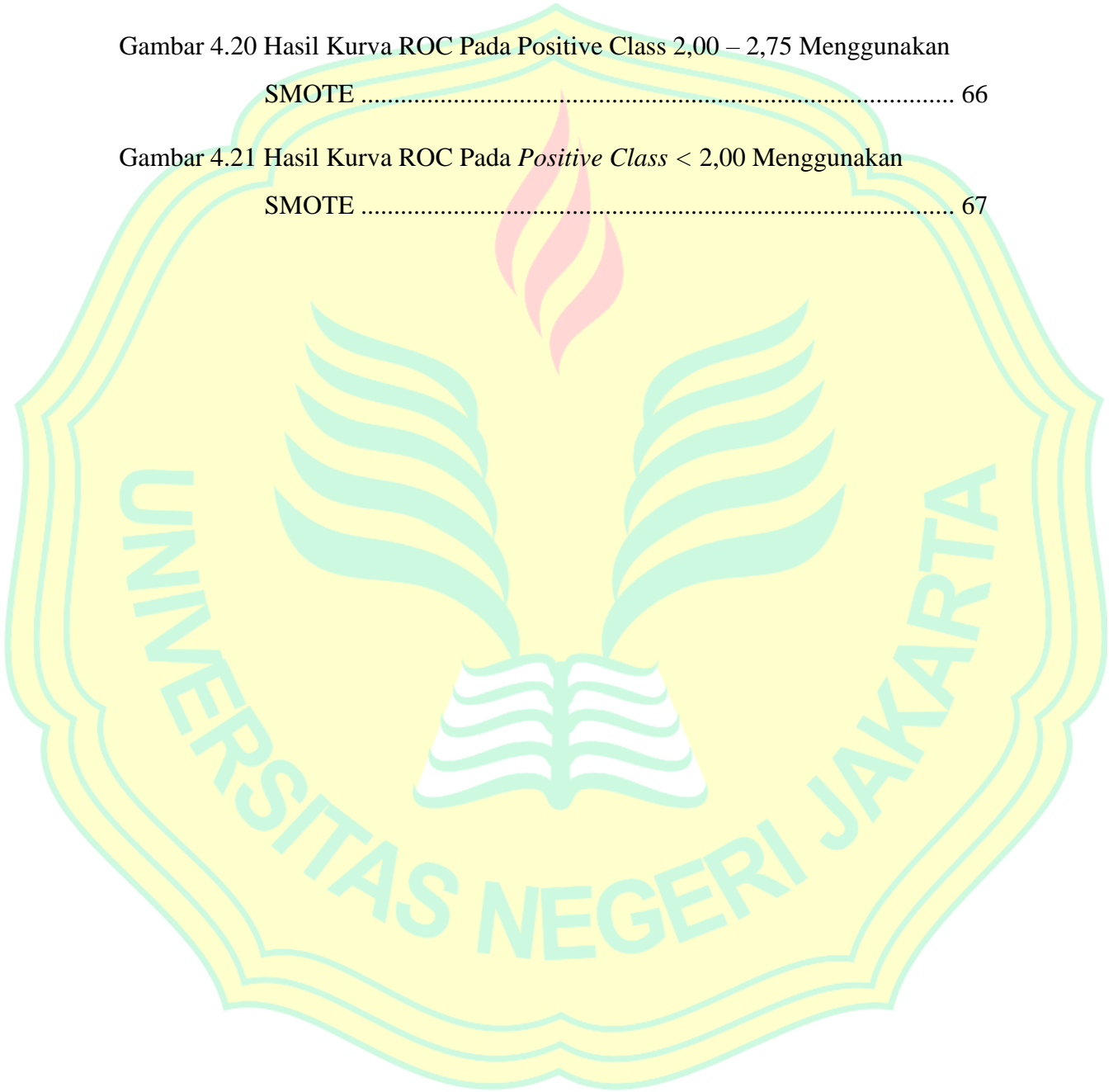
BAB V KESIMPULAN DAN REKOMENDASI.....	72
5.1 Kesimpulan.....	72
5.2 Saran.....	72
DAFTAR PUSTAKA	73
LAMPIRAN.....	79



DAFTAR GAMBAR

Gambar 2.1 Proses <i>Knowledge Discovery Data</i> (KDD)	12
Gambar 2.2 Alur Naive Bayes	17
Gambar 2.3 Kurva ROC untuk kasus diskrit dan kontinu	20
Gambar 2.4 Kerangka Pemikiran	30
Gambar 3.1 Diagram Alir Penelitian	38
Gambar 4.1 Hasil Merapikan Data di Excel	43
Gambar 4.2 Mendeklarasikan Tipe Data dan <i>Role</i> di RapidMiner	44
Gambar 4.3 Tampilan Data Tersimpan di RapidMiner	45
Gambar 4.4 Pemodelan Seleksi Atribut dengan <i>Information Gain</i>	46
Gambar 4.5 Hasil <i>Information Gain</i> Atribut	46
Gambar 4.6 Pemodelan Pembersihan Data	49
Gambar 4.7 Hasil Pembersihan Data	49
Gambar 4.8 Pemodelan Naive Bayes	51
Gambar 4.9 Hasil Prediksi	51
Gambar 4.10 Desain Pemodelan Evaluasi dan Validasi	53
Gambar 4.11 Hasil <i>Confusion Matrix</i> pada RapidMiner	57
Gambar 4.12 Hasil Kurva ROC pada <i>Positive Class</i> > 3,50	59
Gambar 4.13 Hasil Kurva ROC pada <i>Positive Class</i> 2,75 – 3,50	59
Gambar 4.14 Hasil Kurva ROC pada <i>Positive Class</i> 2,00 – 2,75	60
Gambar 4.15 Hasil Kurva ROC Pada <i>Positive Class</i> < 2,00.....	61
Gambar 4.16 Desain Pemodelan Menggunakan SMOTE	61
Gambar 4.17 Hasil <i>Confusion Matrix</i> Menggunakan SMOTE	64

Gambar 4.18 Hasil Kurva ROC Pada <i>Positive Class</i> > 3,50 Menggunakan SMOTE	65
Gambar 4.19 Hasil Kurva ROC Pada <i>Positive Class</i> 2,75 – 3,50 Menggunakan SMOTE	65
Gambar 4.20 Hasil Kurva ROC Pada <i>Positive Class</i> 2,00 – 2,75 Menggunakan SMOTE	66
Gambar 4.21 Hasil Kurva ROC Pada <i>Positive Class</i> < 2,00 Menggunakan SMOTE	67



DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i>	19
Tabel 2.2 Penelitian yang Relevan	26
Tabel 3.1 Daftar Alat Penelitian	31
Tabel 3.2 Atribut Dataset	32
Tabel 3.3 Hasil Klasifikasi	39
Tabel 3.4 Hasil <i>Confusion Matrix</i>	39
Tabel 3.5 Hasil Pengujian.....	40
Tabel 4.1 Tipe Data dan Role	43
Tabel 4.2 Daftar Atribut Setelah Seleksi Atribut	47
Tabel 4.3 Hasil Prediksi	52
Tabel 4.4 Hasil Klasifikasi <i>Confusion Matrix</i>	53
Tabel 4.5 Hasil <i>Confusion Matrix</i>	54
Tabel 4.7 Hasil Prediksi Menggunakan SMOTE	62
Tabel 4.8 <i>Confusion Matrix</i> Menggunakan SMOTE	63
Tabel 4.9 Hasil Pengujian <i>Confusion Matrix</i>	68
Tabel 4.10 Akumulasi Nilai AUC	68
Tabel 4.11 Hasil Pengujian <i>Confusion Matrix</i> Menggunakan SMOTE	69
Tabel 4.12 Akumulasi Nilai AUC Menggunakan SMOTE	70

DAFTAR LAMPIRAN

Lampiran 1. <i>Dataset</i> Prediksi Indeks Prestasi Semester (IPS) Mahasiswa PTIK UNJ Angkatan 2020 dan 2021	79
Lampiran 2. <i>Dataset</i> Prediksi Indeks Prestasi Semester (IPS) Mahasiswa PTIK UNJ Angkatan 2020 dan 2021 Menggunakan SMOTE.....	81
Lampiran 3. Kuesioner Penelitian	84
Lampiran 4. Data Hasil Kuesioner Penelitian	90

