

**PREDIKSI KETERLAMBATAN BIAYA KULIAH DI UNIVERSITAS
MUHAMMADIYAH KALIMANTAN TIMUR MENGGUNAKAN METODE
*LOGISTIC REGRESSION***

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar
Sarjana Komputer

DISUSUN OLEH:

TAUFIQURRAHMAN

1811102441094



**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR
SAMARINDA
2022**

**Prediksi Keterlambatan Biaya Kuliah di Universitas Muhammadiyah
Kalimantan Timur Menggunakan Metode *Logistic Regression***

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar
Sarjana Komputer

Disusun Oleh:

Taufiqurrahman

1811102441094



**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR
SAMARINDA
2022**

HALAMAN PENGESAHAN

PREDIKSI KETERLAMBATAN BIAYA KULIAH DI UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR MENGGUNAKAN METODE *LOGISTIC REGRESSION*

DISUSUN OLEH :

TAUFIQURRAHMAN

1811102441094

Telah melaksanakan ujian skripsi dan dinyatakan lulus,

Pada tanggal 2 Juli 2022

Dosen Pembimbing

Taghfirul Azhima Yoga Siswa, S.kom., M.kom
NIDN : 1118038805

Penguji

Rofilde Hasudungan, Skom., M.sc
NIDN : 1107048601

Dekan



Prof. Ir. Sarjito, MT., Ph.D.
NIDN : 0610116204

Ketua Program Studi



Asslia Johar Latipah, S.Kom M.Cs
NIDN : 1124098902

PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Taufiqurrahman
NIM : 1811102441094
Konsentrasi : Sistem Informasi

Menyatakan bahwa Skripsi dengan judul berikut:

Prediksi Keterlambatan Biaya Kuliah Di Universitas Muhammadiyah Kalimantan Timur Menggunakan Metode Logistic Regression


Dosen Pembimbing : Taghfirul Azhima Yoga Siswa, S.Kom., M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar Sarjana Komputer, baik di Universitas Muhammadiyah Kalimantan Timur(UMKT) maupun di Perguruan Tinggi lainnya
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas Muhammadiyah Kalimantan Timur (UMKT)
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi Muhammadiyah.

Samarinda, 10 Oktober 2022

Yang Menyatakan,




Taufiqurrahman
1811102441094

PRAKATA

Puji dan syukur dipanjatkan kepada Tuhan Yang Maha Esa, karena atas rahmatnya Skripsi yang berjudul "Prediksi Keterlambatan Biaya Kuliah Di Universitas Muhammadiyah Kalimantan Timur Menggunakan Metode *Logistic Regression*" ini diajukan sebagai salah satu syarat untuk memenuhi persyaratan mencapai gelar Sarjana Komputer di Universitas Muhammadiyah Kalimantan Timur.

Penulis menyadari bahwa dalam penyelesaian Proposal ini mendapat bantuan dari berbagai pihak baik langsung maupun tidak langsung, Oleh karena itu pada kesempatan ini penulis menyampaikan terima kasih kepada:

1. Orang tua penulis Bapak Riduan Gani dan Ibu Asma yang selalu memberikan doa dan dukungan.
2. Bapak Taghfirul Azhima Yoga Siswa, S.Kom, M.Kom, selaku Dosen Pembimbing yang telah memberikan bimbingan dan masukkan kepada penulis.
3. Ibu Asslia Johar Latipah, M.Cs., ketua Program Studi S1 Teknik Informatika.
4. Prof. Ir. Sarjito, M.T.,Ph.D., selaku Dekan Fakultas Sains & Teknologi.
5. Prof. Dr. H. Bambang Setiaji, selaku Rektor Universitas Muhammadiyah Kalimantan Timur.
6. Teman-teman mahasiswa yang telah membantu dan memberi dukungan dalam penyelesaian laporan ini.

Dengan segala kekurangan dan keterbatasan dalam pembahasan, penulis berharap semoga proposal ini dapat memberikan manfaat bagi semua pihak yang memerlukan. Demi kesempurnaan penulisan laporan ini maka segala kritik dan saran yang membangun akan penulis perhatikan. Terima kasih.

Samarinda, 12 Februari 2022



Taufiqurrahman

ABSTRAK

Universitas Muhammadiyah Kalimantan Timur dalam melaksanakan kegiatan operasionalnya mengandalkan dana dari mahasiswa, salah satunya adalah biaya kuliah (SPP). Hal ini menimbulkan masalah jika mahasiswa terlambat dalam melakukan pembayaran kuliah, karena dari SPP tersebut dapat meningkatkan kualitas pendidikan dan sarana fasilitas. Maka itu tujuan dari penelitian ini adalah menentukan indikator, kemudian mengimplementasikan algoritma *logistic regression* dengan membagi data menjadi 70:30 dan mengevaluasi hasil kinerja dari algoritma tersebut menggunakan *confusion matrix*. Data yang didapat bersumber dari Bagian Administrasi Akademik dan Bagian Administrasi Keuangan yaitu sebanyak 12.408 data dengan beberapa atribut seperti fakultas, prodi, angkatan, gender, penghasilan ayah, penghasilan ibu, pendidikan ayah, pendidikan ibu dan label (terlambat atau tepat). Dari hasil uji data yang telah dilakukan maka mendapatkan hasil akurasi sebesar 55,89%.

Kata Kunci : Prediksi, *Logistic Regression*, *Confusion Matrix*

ABSTRACT

Muhammadiyah University of East Kalimantan in carrying out its operational activities relies on funds from students, one of which is tuition fees (SPP). This creates problems if students are late in making tuition payments, because tuition payments can improve the quality of education and facilities. Therefore, the purpose of this study is to determine indicators, then implement the logistic regression algorithm by dividing the data into 70:30 and evaluating the performance results of the algorithm using the confusion matrix. The data obtained are sourced from the Academic Administration Section and the Financial Administration Section as many as 12,408 data with several attributes such as the faculty, study program, class, gender, father's income, mother's income, father's education, mother's education and label (late or not late). From the results of the data test that has been carried out, it gets the results of an accuracy of 55.89%.

Keywords : *Prediction, Logistics Regression, Confusion Matrix, Accuracy*

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN JUDUL	ii
HALAMAN PENGESAHAN	iii
PERNYATAAN KEASLIAN SKRIPSI	iv
PRAKATA.....	v
ABSTRAK.....	vi
<i>ABSTRACT</i>	vii
DAFTAR ISI	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN	xii
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	4
1.4 Batasan Masalah.....	4
BAB 2 TINJAUAN PUSTAKA.....	5
2.1 Prediksi Keterlambatan Biaya Kuliah.....	5
2.2 <i>Data Mining</i>	6
2.3 Teknik Resampling.....	8
2.4 <i>Logistic Regression</i>	8
2.5 <i>Confusion Matrix</i>	10
BAB 3 METODE PENELITIAN	12
3.1 Tahapan Penelitian	12
3.2 Pengumpulan Data (Dataset)	12
3.3 Persiapan Data.....	13

3.4 Pembagian Data	14
3.5 Modeling.....	14
3.6 Evaluasi	15
BAB 4 HASIL DAN PEMBAHASAN	16
4.1 Data Penelitian	16
4.2 Persiapan Data.....	18
4.3 Permodelan Algoritma Logistic Regression	21
4.4 Evaluasi	24
BAB 5 PENUTUP	26
5.1 Kesimpulan	26
5.2 Saran	26
DAFTAR PUSTAKA.....	27
LAMPIRAN	30

DAFTAR TABEL

Tabel 2.1 Penelitian Relevan	5
Tabel 2.2 Penelitian Prediksi	10
Tabel 2.3 Confusion Matrix	11
Tabel 3.1 Atribut Data Mahasiswa Dari Biro Administrasi Keuangan.....	13
Tabel 3.2 Atribut Data Mahasiswa Dari Biro Administrasi Akademik.....	13
Tabel 3.1 Atribut Data Mahasiswa Dari Biro Administrasi Keuangan.....	13
Tabel 3.2 Atribut Data Mahasiswa Dari Biro Administrasi Akademik.....	13
Tabel 4.1 Data Bagian Administrasi Keuangan	16
Tabel 4.2 Data Bagian Administrasi Akademik	17
Tabel 4.3 Transformasi Data Fakultas	19
Tabel 4.4 Transformasi Data Prodi.....	19
Tabel 4.5 Transformasi Angkatan	20
Tabel 4.6 Transformasi Data Gender	20
Tabel 4.7 Transformasi Data Penghasilan Ayah.....	20
Tabel 4.8 Transformasi Data Penghasilan Ibu.....	20
Tabel 4.9 Transformasi Data Pendidikan Ayah	20
Tabel 4.10 Transformasi Pendidikan Ibu.....	21
Tabel 4.11 Transformasi Label	21
Tabel 4.12 Perhitungan Confusion Matrix	23
Tabel 4.13 Import Dataset	24
Tabel 4.14 Pembagian data 70:30.....	24
Tabel 4.15 Confusion Matrix 70:30	25

DAFTAR GAMBAR

Gambar 1.1 Data Keterlambatan Pembayaran.....	2
Gambar 2.1 Pola Kurva.....	9
Gambar 3.1 Proses Analisis Data	12

DAFTAR LAMPIRAN

- Lampiran 1. Riwayat Hidup
- Lampiran 2. Data Bagian Biro Administrasi Keuangan²
- Lampiran 3. Data Bagian Biro Administrasi Akademik
- Lampiran 4. Integrasi Data
- Lampiran 5. Pembersihan Data
- Lampiran 6. Reduksi Data
- Lampiran 7. Coding
- Lampiran 8. Surat Ijin Penelitian Biro Administrasi Keuangan
- Lampiran 9. Surat Ijin Penelitian Biro Administrasi Akademik
- Lampiran 10. Surat Keterangan Tidak Uji Validitas
- Lampiran 11. Lembar Bimbingan Skripsi
- Lampiran 12. Hasil Uji Plagiasi