

**IMPLEMENTASI ALGORITMA *RANDOM FOREST* DALAM  
MEMPREDIKSI KETERLAMBATAN PEMBAYARAN SPP DI  
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR**

**SKRIPSI**

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar  
Sarjana Komputer

**DISUSUN OLEH :**

**ILHAM QASTHALANI**

**1811102441042**



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

**2022**

**Implementasi Algoritma *Random Forest* dalam Memprediksi  
Keterlambatan Pembayaran SPP di  
Universitas Muhammadiyah Kalimantan Timur**

**SKRIPSI**

Diajukan untuk memenuhi sebagian persyaratan pencapaian gelar  
Sarjana Komputer

**Disusun Oleh :**

**Ilham Qasthalani**

**1811102441042**



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

**HALAMAN PENGESAHAN**  
**IMPLEMENTASI ALGORITMA *RANDOM FOREST* DALAM**  
**MEMPREDIKSI KETERLAMBATAN PEMBAYARAN SPP DI**  
**UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR**

DISUSUN OLEH :

**ILHAM QASTHALANI**

**1811102441042**

Telah melaksanakan ujian skripsi dan dinyatakan lulus,

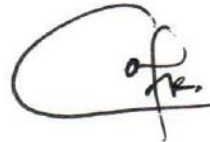
Pada tanggal 2 Juli 2022

Dosen Pembimbing



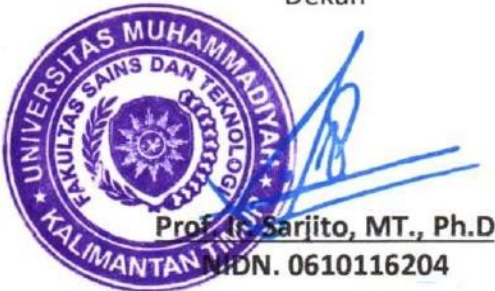
Taghfirul Azhima Yoga Siswa, S.Kom., M.Kom  
NIDN. 1118038805

Penguji



Rofilde Hasudungan, S.Kom., M.Sc  
NIDN. 1107048601

Dekan



Prof. Dr. 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 : Ilham Qasthalani  
NIM : 1811102441042  
Konsentrasi : Sistem Informasi

Menyatakan bahwa Skripsi dengan judul berikut:

**Implementasi Algoritma *Random Forest* Dalam Memprediksi Keterlambatan Pembayaran SPP di Universitas Muhammadiyah Kalimantan Timur.**

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 Agustus 2022

Yang Menyatakan,



Ilham Qasthalani

1811102441042

## PRAKATA

Puji dan syukur atas kehadiran Allah SWT karena dengan Rahmat dan Hidayah-Nya penulis dapat menyelesaikan skripsi sebagai bukti pelaksanaan skripsi dengan judul “Implementasi Algoritma *Random Forest* Dalam Memprediksi Keterlambatan Pembayaran SPP Di Universitas Muhammadiyah Kalimantan Timur” dapat penulis selesaikan dengan sebaik – baiknya.

Penyelesaian skripsi yang dilaksanakan tidak dapat terlaksana dengan baik apabila tidak didukung dan dibantu oleh pihak yang telah mendorong, membimbing dan mengarahkan penulis. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Bapak Mursidi dan Ibu Dra. Mu’Minah selaku orang tua, terima kasih atas do’a serta dukungan moral dan materi yang diberikan selama ini.
2. Bapak Taghfirul Azhima Yoga Siswa, S.Kom, M.Kom selaku Dosen Pembimbing yang telah memberikan bimbingan dan masukkan kepada penulis
3. Asslia Johar Latipah, M.Cs., selaku 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. Bapak dan Ibu Dosen Prodi Teknik Informatika, Universitas Muhammadiyah Kalimantan Timur dan para staf karyawan yang saya banggakan dan hormati.
7. Teman-teman seperjuangan proposal, khususnya Agustya Nanda Pratiwi, Anton Saputra, Asnur Karima, Muhammad Norhalimi, Muhammad Rhosyid Akhmad, Rizky Aspiah, Taufiqurrahman, dan Yoga Priantama yang telah membantu dan memberi dukungan dalam penyelesaian skripsi ini.

Penulis menyadari bahwa dalam Skripsi ini masih banyak kekurangan, oleh karena itu kritik dan saran dari semua pembaca sangat penulis harapkan demi perbaikan dimasa mendatang. Penulis berharap semoga apa yang telah tertulis dalam Skripsi ini bermanfaat bagi yang memerlukan. Akhir kata, penulis

menyampaikan terima kasih kepada semua pihak yang telah membantu dalam membuat Skripsi ini.

Samarinda, 10 Mei 2022

A handwritten signature in black ink, appearing to read 'Ilham Qasthalani', with a horizontal line underneath.

Ilham Qasthalani

## ABSTRAK

Universitas Muhammadiyah Kalimantan Timur (UMKT) merupakan salah satu lembaga pendidikan yang dimiliki oleh persyarikatan muhammadiyah di Provinsi Kalimantan Timur yang pembiayaannya operasionalnya dibebankan kepada mahasiswa, terutama pada pembayaran Sumbangan Pembangunan Pendidikan (SPP). Oleh karena itu, keterlambatan dari pembayaran SPP tentunya sangat berpengaruh bagi perguruan tinggi. Penelitian ini bertujuan untuk menentukan indikator, mengimplementasikan algoritma *random forest* dan mengevaluasi hasil kinerja algoritma *random forest* dengan pembagian data 90% data *training* dan 10% *testing* menggunakan *confusion matrix*. Penelitian ini menggunakan aplikasi *rapidminer* untuk melakukan prediksi melalui klasifikasi dengan metode *random forest*. Data yang digunakan dalam penelitian ini yaitu berupa data pembayaran SPP mahasiswa pada tahun 2019-2021 yang berasal dari Biro Administrasi Akademik dan Administrasi Keuangan yang berjumlah 12.408 data. Hasil dari implementasi algoritma *random forest* dengan pembagian data data 90% data *training* dan 10% *testing* menghasilkan akurasi sebesar 78,47%.

**Kata kunci:** *confusion matrix*, prediksi, *random forest*

## **ABSTRACT**

*Muhammadiyah University of East Kalimantan (UMKT) is one of the educational institutions owned by the Muhammadiyah association in East Kalimantan Province whose operational financing is charged to students, especially the payment of Educational Development Contributions (SPP). Therefore, the delay in paying tuition fees is certainly very influential for universities. This study aims to determine indicators, implement the random forest algorithm and evaluate the performance results of the random forest algorithm with 90% data sharing of training data and 10% testing using a confusion matrix. This study uses the rapidminer application to make predictions through classification using the random forest method. The data used in this study is in the form of student tuition payment data in 2019-2021 originating from the Academic Administration and Financial Administration Bureau, totaling 12,408 data. The results of the implementation of the random forest algorithm with data sharing of 90% of training data and 10% of testing resulted in an accuracy of 78,47%.*

**Keywords:** *confusion matrix, prediction, random forest*



## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN JUDUL .....	ii
HALAMAN PENGESAHAN .....	iii
PERNYATAAN KEASLIAN SKRIPSI .....	iv
PRAKATA.....	v
ABSTRAK.....	vii
<i>ABSTRACT</i> .....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR .....	xii
DAFTAR LAMPIRAN .....	xiii
BAB 1 PENDAHULUAN.....	1
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	3
1.3. Tujuan Penelitian .....	3
1.4. Batasan Masalah.....	4
BAB 2 TINJAUAN PUSTAKA.....	5
2.1. Prediksi Keterlambatan Pembayaran SPP .....	5
2.2. <i>Data Mining</i> .....	8
2.3. Prapemrosesan Data .....	9
2.4. <i>Random Under Sampling</i> .....	11
2.5. <i>Random Forest</i> .....	12
2.6. Seleksi Atribut <i>Information Gain Ratio</i> .....	15
2.7. <i>Confusion Matrix</i> .....	16
BAB 3 METODE PENELITIAN .....	18
3.1 Tahapan Penelitian .....	18
3.2 Pengumpulan Data .....	18
3.3 Persiapan Data.....	19
3.4 Permodelan .....	20

3.5	Evaluasi .....	21
BAB 4 HASIL DAN PEMBAHASAN .....		22
4.1	Data Penelitian .....	22
4.2	Persiapan Data.....	25
4.3	Permodelan Algoritma <i>Random Forest</i> .....	26
4.3.1	Menyiapkan Dataset .....	26
4.3.2	Menghitung <i>Entropy</i> .....	28
4.3.3	Menghitung <i>Gain</i> .....	29
4.4	Evaluasi <i>Performance Random Forest</i> .....	33
BAB 5 PENUTUP .....		35
5.1	Kesimpulan .....	35
5.2	Saran .....	35
DAFTAR PUSTAKA.....		36
LAMPIRAN .....		40

## DAFTAR TABEL

Tabel 2.1 Prediksi Keterlambatan Pembayaran SPP .....	5
Tabel 2.2 Penelitian <i>Random Forest</i> .....	12
Tabel 2.2 Penelitian <i>Random Forest</i> (Lanjutan).....	13
Tabel 2.2 Penelitian <i>Random Forest</i> (Lanjutan).....	14
Tabel 2.3 <i>Confusion Matrix</i> .....	17
Tabel 3.1 Atribut Data Mahasiswa dari Biro Administrasi Keuangan .....	18
Tabel 3.1 Atribut Data Mahasiswa dari Biro Administrasi Keuangan (Lanjutan)..	19
Tabel 3.2 Atribut Data Mahasiswa dari Biro Administrasi Akademik .....	19
Tabel 4.1 Data Bagian Administrasi Keuangan .....	22
Tabel 4.2 Data Bagian Administrasi Akademik Tahun 2017-2021 .....	23
Tabel 4.2 Data Bagian Administrasi Akademik Tahun 2017-2021 (Lanjutan).....	24
Tabel 4.3 Hasil Pembobotan <i>Information Gain Ratio</i> .....	25
Tabel 4.4 Data Sampel <i>Training</i> .....	27
Tabel 4.4 Data Sampel <i>Training</i> (Lanjutan).....	28
Tabel 4.5 Perhitungan Nilai <i>Entropy</i> Fakultas .....	28
Tabel 4.5 Perhitungan Nilai <i>Entropy</i> Fakultas (lanjutan) .....	29
Tabel 4.6 Perhitungan <i>Entropy</i> dan <i>Gain</i> .....	29
Tabel 4.6 Perhitungan <i>Entropy</i> dan <i>Gain</i> (Lanjutan).....	30
Tabel 4.6 Perhitungan <i>Entropy</i> dan <i>Gain</i> (Lanjutan).....	31
Tabel 4.7 <i>Entropy</i> dan <i>Gain</i> Ekonomi Bisnis dan Politik.....	32
Tabel 4.8 Evaluasi <i>Confusion Matrix</i> .....	33

## DAFTAR GAMBAR

Gambar 1.1 Pembayaran SPP Di UMKT Tahun 2019-2021 .....	1
Gambar 2.1 Flowchart Algoritma <i>Random Under sampling</i> .....	11
Gambar 2.2 Contoh <i>Random Forest</i> .....	14
Gambar 3.1 Tahapan Penelitian.....	18
Gambar 4.1 Pohon Keputusan Sementara.....	31
Gambar 4.2 Pohon Keputusan Akhir.....	33

## DAFTAR LAMPIRAN

- Lampiran 1. Riwayat Hidup
- Lampiran 2. Data Keseluruhan Mahasiswa
- Lampiran 3. Integrasi Data
- Lampiran 4. Pembersihan Data
- Lampiran 5. Reduksi Data
- Lampiran 6. *Rapidminer*
- Lampiran 7. Surat ijin Penelitian Biro Administrasi Keuangan
- Lampiran 8. Surat ijin Penelitian Biro Administrasi Akademik
- Lampiran 9. Surat Keterangan Tidak Uji Validitas
- Lampiran 10. Lembar Bimbingan Skripsi
- Lampiran 11. Hasil Uji Plagiasi