

**ANALISIS KLASIFIKASI MASA STUDI MAHASISWA PRODI TEKNIK
INFORMATIKA UNILAK DENGAN METODE SUPPORT VECTOR
MACHINE (SVM), ITERATIVE DICHOTOMISER 3 (ID3), RANDOM
FOREST DAN K-NEAREST NEIGHBORS (KNN)**

SKRIPSI



OLEH:

HADI SUNARYANTO

NIM : 1655201029

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS LANCANG KUNING

2022

HALAMAN PERSETUJUAN

Nama : Hadi Sunaryanto
NIM : 1655201029
Program Studi : Teknik Informatika
Judul : Analisis Klasifikasi Masa Studi Mahasiswa Prodi Teknik Informatika UNILAK dengan metode *Support Vector Machine (SVM)*, *Iterative Dichotomiser 3 (ID3)*, *Random Forest* dan *K-Nearest Neighbor (KNN)*

Pekanbaru, 12 Februari 2022

Disetujui oleh,
Pembimbing

MHD Arief Hasan M.Kom., MTA,
NIDN. 1021068701

Guntoro, ST., M.Kom., MTA
NIDN. 1012018803

Mengetahui,

Dekan
Fakultas Ilmu Komputer



H. Fajrizal, S.P., M.Kom
NIDN. 1024076201

Ketua Program Studi
Teknik Informatika



Dr. Yogi Yuncetri, M.Kom., MTA., MCE.
NIDN. 102206880

HALAMAN PENGESAHAN

Nama : Hadi Sunaryanto
NIM : 1655201029
Program Studi : Teknik Informatika
Judul : Analisis Klasifikasi Masa Studi Mahasiswa Prodi Teknik Informatika UNILAK dengan Metode *Support Vector Machine (SVM)*, *Iterative Dichotomiser 3 (ID3)*, *Random Forest* dan *K-Nearest Neighbors (KNN)*

Skripsi ini telah dipertahankan di depan Dewan Penguji pada tanggal 12 Februari 2022 Ujian Komprehensif sebagai persyaratan untuk mendapatkan gelar Sarjana Komputer (S.Kom)

Dewan Penguji,

Ketua



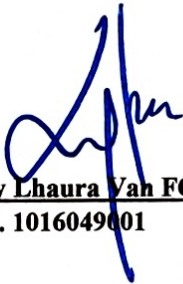
MHD Arief Hasan, M.Kom., MTA.
NIDN. 1021068701

Sekretaris



Guntoro, ST., M.Kom., MTA.
NIDN. 1012018803

Anggota



Lucky Lhaura Van FC, M.Kom., MTA.
NIDN. 1016049001

Anggota



Sutejo, M.Kom., MTA.
NIDN. 1002068801

Anggota



Ahmad Zamsuri, M.Kom., MTA.
NIDN. 1007078602

Mengetahui,

Dekan,



H. Farizal, S.P., M.Kom., MTA.
NIDN. 1024076201

Kepala Program Studi



Dr. Yosi Yusefri, M.Kom., MTA., MCE.
NIDN. 102206880

PERNYATAAN KEASLIAN KARYA ILMIAH

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepengetahuan saya juga tidak terdapat Karya/Penndapat yang pernah ditulis oleh orang lain, kecuali bagian-bagian secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Apabila ditemukan Jiplakan/Plagiat maka saya bersedia menerima akibat berupa sanksi Akademis dan sanksi lain diberikan oleh orang yang berwenang sesuai ketentuan peraturan dan perundang-undangan yang berlaku.

Pekanbaru, 03 Februari 2022

Pembuat Pernyataan

Hadi Sunaryanto
1655201029

ANALISIS KLASIFIKASI MASA STUDI MAHASISWA PRODI TEKNIK INFORMATIKA UNILAK DENGAN METODE SUPPORT VECTOR MACHINE (SVM), ITERATIVE DICHOTOMISER 3 (ID3), RANDOM FOREST DAN K-NEAREST NEIGHBORS (KNN)

Hadi Sunaryanto¹⁾, MHD. Arif Hasan, M.Kom²⁾, Guntoro ST., M.Kom³⁾

Email : sunaryantohadi2@gmail.com¹⁾, m.arif@unilak.ac.id²⁾, guntoro@unilak.ac.id³⁾

ABSTRAK

Pemelitian ini berjudul “Analisis Klasifikasi Masa Studi Mahasiswa Prodi Teknik Informatika Unilak Dengan Metode Support Vector Machine (SVM), Iterative Dichotomiser 3 (ID3), Random Forest dan K-Nearest Neighbors (KNN)” dalam menyikapi kecenderungan lebih besarnya jumlah masuknya mahasiswa dibanding angka kelulusan mahasiswa maka diperlukan sebuah usaha untuk memahami apakah ada faktor-faktor yang mempengaruhi lama masa studi mahasiswa. Pada dasarnya lama masa studi bukanlah sebagai tolak ukur kemampuan akademik non akademik seorang mahasiswa namun sebagian besar orang menilai mahasiswa dengan masa studi lebih dari 8 semester atau tergolong lama itu tidak bagus. Oleh karena itu peneliti memilih untuk mengklasifikasi faktor-faktor yang mempengaruhi lama masa studi mahasiswa di Fakultas Ilmu Komputer Universitas Lancang Kuning. Penelitian ini menggunakan 4 (empat) metode perhitungan. Dengan adanya beberapa metode yang digunakan penulis dapat membandingkan hasil dari empat metode perhitungan sehingga dapat menentukan mana metode dengan perhitungan lebih baik. Hasil dari penelitian ini adalah perbandingan antara 4 (empat) metode perhitungan dalam menentukan metode mana yang memiliki kemampuan klasifikasi baik.

Kata Kunci : Klasifikasi, Masa Studi, *Support Vector Machine* (SVM), *Iterative Dichotomiser 3*, *Random Forest*, *K-Nearest Neighbor* (KNN)

**CLASSIFICATION ANALYSIS OF UNILAK INFORMATICS
ENGINEERING STUDENTS USING THE SUPPORT VECTOR
MACHINE (SVM) METHOD, ITERATIVE DICHOTOMISER 3 (ID3),
RANDOM FOREST AND K-NEAREST NEIGHBORS (KNN) SVM),
ITERATIVE DICHOTOMISER 3 (ID3), RANDOM FOREST AND
K-NEAREST NEIGHBORS (KNN)**

Hadi Sunaryanto¹⁾, Muh. Arief Hasan, M.Kom²⁾, Guntoro ST., M.Kom³⁾

Email : sunaryantohadi2@gmail.com¹, m.arif@unilak.ac.id², guntoro@unilak.ac.id³

ABSTRACT

This research is entitled "Analysis of Classification of Study Periods of Students in the Informatics Engineering Study Program of Unilak with the Support Vector Machine (SVM), Iterative Dichotomiser 3 (ID3), Random Forest and K-Nearest Neighbors (KNN) method" in responding to the tendency for the larger number of students to enter than the number. After graduation, an effort is needed to understand whether there are factors that affect the length of the student's study period. Basically, the length of the study period is not a measure of a student's non-academic academic ability, but most people judge that students with a study period of more than 8 semesters or long are not good. Therefore, the researcher chose to classify the factors that affect the length of the student's study period at the Faculty of Computer Science, Lancang Kuning University. This study uses 4 (four) calculation methods. With the several methods used, the authors can compare the results of the four calculation methods so that they can determine which method is better calculated. The result of this research is a comparison between 4 (four) calculation methods in determining which method has good classification ability.

Keywords : Classification, Study Period, Support Vector Machine (SVM), Iterative Dichotomiser 3, Random Forest, K-Nearest Neighbor (KNN)

KATA PENGANTAR

Syukur Alhamdulillah kehadiran Allah SWT yang selalu melimpahkan rahmat dan hidayah-Nya, karena berkat rahmat dan hidayah-Nya berupa iman, ilmu, kesabaran, kesehatan, keikhlasan serta optimisme yang diberikan-Nya kepada penulis sehingga penulis mampu untuk menyelesaikan penelitian. Selanjutnya shalawat beriring salam penulis haturkan kepada Rasulullah SAW, yang telah berhasil membawa umatnya ke dalam yang penuh ilmu pengetahuan ini.

Dalam penyusunan penulisan Skripsi ini banyak pihak yang dilibatkan baik dalam pengumpulan data maupun penyusunan data serta informasi bahkan semangat dan motivasi yang penulis butuhkan, oleh karena itu sepantasnya penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada :

1. Bapak Dr. Junaidi. SS. M.Hum, selaku Rektor Universitas Lancang Kuning Pekanbaru.
2. Bapak H. Fajrizal, SP., M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Lancang Kuning Pekanbaru, tempat di mana penulis masih menimba banyak ilmu di sana.
3. Bapak Dr. Yogi Yunefri, M.Kom, MTA, MCE selaku Ketua Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Lancang Kuning Pekanbaru, tempat di mana penulis masih menimba banyak ilmu di sana.
4. Bapak MHD. Arief Hasan, M.Kom, selaku Pembimbing I penulis dalam melaksanakan skripsi.
5. Bapak Guntoro, S.T., M. Kom, selaku Pembimbing II penulis dalam melaksanakan skripsi.
6. Ibu tercinta, Surjanati, yang tiada henti-hentinya mendoakan, memberi dorongan moril maupun materi selama kuliah dan melaksanakan penelitian di Universitas Lancang Kuning.

Penulis menyadari, tanpa bantuan, bimbingan serta dorongan dari berbagai pihak, penelitian ini tidak dapat diselesaikan dengan baik. Hanya Allah SWT yang bisa membalas kebaikan yang telah diberikan, amiiin.

Penulis berharap adanya sumbangan pikiran, kritik, dan saran untuk perbaikan dan kesempurnaan penelitian ini. Harapan penulis semoga hasil dari Proposal Skripsi ini dapat bermanfaat bagi kemajuan ilmu pengetahuan, khususnya di bidang ilmu sistem dan teknik komputer.

Pekanbaru, 12 Februari 2022

Penulis,

Hadi Sunaryanto

DAFTAR ISI

HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
PERNYATAAN KEASLIAN KARYA ILMIAH	iv
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN	13
A. Latar Belakang	13
B. Rumusan Masalah	14
C. Batasan Masalah	15
D. Tujuan Penelitian	15
E. Manfaat Penelitian	16
F. Sistematika Penulisan	17
BAB II TINJAUAN PUSTAKA	19
A. Tinjauan Pustaka	19
B. Teori Dasar	22
BAB III METODOLOGI PENELITIAN	31
BAB IV GAMBARAN UMUM TEMPAT PENELITIAN	37
2. Misi Fakultas Ilmu Komputer	37
BAB V	39
PEMBAHASAN	39
A. Tahapan Pengumpulan Data	39
B. Metode Support Vector Machine	43
45	
C. Metode Iterative Dichotomiser 3	46
D. Metode Random Forrest	49

E.	Metode K-Nearest Neighbor	50
BAB VI		52
KESIMPULAN DAN SARAN		52
DAFTAR PUSTAKA		53

DAFTAR TABEL

Tabel 5.1 Isi Pertanyaan Kuesioner	38
Tabel 5.2 Nilai Kuesioner	38
Tabel 5.3 Jumlah mahasiswa yang mengikuti organisasi, beasiswa dan part time	45
Tabel 5.4 Nilai kuesioner	45
Tabel 5.5 Nilai entropy dari data kuesioner	46
Tabel 5.6 Nilai gain dari data kuesioner	46
Tabel 5.7 Perbandingan nilai gain tiap atribut	46
Tabel 5.8 Data kuesioner	50
Tabel 5.9 Data test untuk mencari nilai label tiap data	50
Tabel 5.10 Data training 1	51
Tabel 5.11 Data training 2	51
Tabel 5.12 Data training 3	52
Tabel 5.13 Data training 4	52
Tabel 5.14 Data training 5	53
Tabel 5.15 Data training 6	53
Tabel 5.16 Data training 7	54
Tabel 5.17 Data training 8	54
Tabel 5.18 Data training 9	55
Tabel 5.19 Data training 10	55
Tabel 5.20 Data nilai label	56

DAFTAR GAMBAR

Gambar 5.1	44
Gambar 5.2	44
Gambar 5.3	49
Gambar 5.4	57

BAB I

PENDAHULUAN

A. Latar Belakang

Perguruan tinggi merupakan jenjang pendidikan setelah pendidikan menengah yang mencakup program pendidikan diploma, sarjana, magister, spesialis dan doktor yang diselenggarakan oleh perguruan tinggi. Kelulusan adalah hasil akhir dari proses kegiatan belajar mengajar selama mengikuti perkuliahan di perguruan tinggi. Universitas Lancang Kuning (UNILAK) adalah salah satu Universitas Swasta yang berada di Riau dengan memiliki 9 (sembilan) fakultas. Salah satu diantaranya adalah Fakultas Ilmu Komputer, yang disingkat Fasilkom. Fasilkom terdiri dari program Studi S1 yaitu Teknik Informatika dan Sistem Informasi.

Kurikulum program S-1 Fakultas Ilmu Komputer dirancang berdasarkan Kurikulum Nasional, tanpa mengabaikan standar internasional. Pengembangan kurikulum yang dilakukan setiap 2 tahun sekali selalu memperhatikan kebutuhan *users* dan perubahan lingkungan yang begitu cepat dalam menghadapi era Millenium baru.

Kurikulum Fakultas Ilmu Komputer Universitas Lancang Kuning juga dilengkapi dengan berbagai praktikum dari berbagai Laboratorium, sedangkan dalam rangka *Link and Match* serta pengembangan wawasan, mahasiswa diberi kesempatan mengikuti praktek kerja magang di perusahaan-perusahaan dalam / luar negeri dan mengikuti program pertukaran mahasiswa dengan Perguruan Tinggi luar negeri agar lulusan program studi S-1 Fakultas Ilmu Komputer lebih

mampu bersaing di pasar global, atas kerjasama dengan Industri-Industri Besar di Indonesia.

Lama studi program sarjana menurut peraturan akademik Universitas Lancang Kuning dijadwalkan dalam 8 semester (4 tahun) atau dapat ditempuh kurang dari 8 semester (4 tahun) dan selambat-lambatnya 14 semester (7 tahun). Setiap tahun Universitas Lancang Kuning menyelenggarakan upacara wisuda dalam 2 periode yaitu periode April dan Oktober. Dalam 2 periode kelulusan, jumlah lulusan dengan jumlah mahasiswa baru tidak sebanding. Sehingga menimbulkan jumlah mahasiswa meningkat. Lama masa studi mahasiswa kemungkinan dapat dipengaruhi oleh banyak faktor. Faktor-faktor yang diperkirakan mempengaruhi dalam kelulusan tepat waktu antara lain Indeks Prestasi Kelulusan (IPK), jenis kelamin, beasiswa, bekerja *part time*, organisasi, dan jalur masuk universitas.

B. Rumusan Masalah

Dari uraian yang telah dipaparkan pada latar belakang diatas, maka dapat dirumuskan beberapa permasalahan antara lain :

1. Analisis seberapa besar pengaruh faktor yang mempengaruhi masa studi mahasiswa Fasilkom Universitas Lancang Kuning disetiap metode yang digunakan.
2. Mengetahui metode mana yang memiliki tingkat ke akuratan paling tinggi.

C. Batasan Masalah

Untuk menghindari pembahasan diluar permasalahan penulis pun melakukan pembatasan terhadap masalah yang diteliti, berikut ini adalah batasan permasalahannya :

1. Penelitian ini di dilaksanakan di Fakultas Ilmu Komputer Universitas Lancang Kuning program studi Teknik Informatika.
2. Pengambilan data adalah dari mahasiswa angkatan 2016 - 2019.
3. Metode yang digunakan adalah *Support Vector Machine* (SVM), *Iterative Dichotomiser 3* (ID3), *Random Forrest* dan *K- Nearest Neighbor* (KNN).

D. Tujuan Penelitian

Tujuan dari penelitian ini adalah :

1. Memperoleh gambaran secara deskripsi faktor-faktor yang mempengaruhi masa studi mahasiswa prodi Teknik Informatika.
2. Memperoleh model dan ketepatan klasifikasi lama studi mahasiswa kelulusan mahasiswa prodi Teknik Informatika dengan menggunakan algoritma *Support Vector Machine* (SVM), *Iterative Dichotomiser 3* (ID3), *K- Nearest Neighbor* (KNN), *Random Forrest*.
3. Membandingkan ketepatan klasifikasi algoritma SVM, ID3, KNN dan Random Forest pada klasifikasi lama studi mahasiswa prodi Teknik Informatika Universitas Lancang Kuning.

4. Menentukan algoritma terbaik dalam klasifikasi lama masa studi mahasiswa prodi Teknik Informatika Universitas Lancang Kuning.

E. Manfaat Penelitian

Manfaat dalam penelitian ini sebagai berikut:

1. Manfaat bagi Penulis.
 - a. Penelitian ini digunakan untuk menambah wawasan, pengetahuan dan pengalaman peneliti yang mana terkait permasalahan yang diteliti tentang analisis klasifikasi masa studi mahasiswa Teknik Informatika Universitas Lancang Kuning
 - b. Dapat mengetahui perbedaan antara beberapa algoritma dalam penyelesaian perhitungan klasifikasi lama masa studi mahasiswa Teknik Informatikaa Universitas Lancang Kuning.
2. Manfaat bagi instansi.
 - a. Mengetahui apa-apa saja faktor yang dapat mempengaruhi lama masa studi mahasiswa Teknik Informatika Universitas Lancang Kuning.
 - b. Dapat menggunakan informasi yang didapat untuk memperbaiki atau mencari solusi dari masalah yang berkaitan dengan lama masa studi mahasiswa Teknik Informatika Universitas Lancang Kuning
3. Manfaat bagi Pembaca.

- a. Sebagai salah satu referensi untuk peneliti lain yang mengambil permasalahan yang serupa dimasa yang akan datang.
- b. Menyumbangkan isi pemikiran penulis dalam menerapkan metode *Support Vector Machine* (SVM), algoritma *Iterative Dichotomiser 3* (ID3), *Random Forrest* dan *K- Nearest Neighbor* (KNN).

F. Sistematika Penulisan

Agar lebih jelas dan mudah dimengerti, maka penulis akan memberikan rangkuman sebagai berikut :

BAB I	Pendahuluan
	Dalam bab ini menjelaskan tentang latar belakang , ruang lingkup, rumusan masalah, batasan masalah, waktu dan tempat pelaksanaan, tujuan penelitian, manfaat penelitian dan sistematika penulisan.
BAB II	Tinjauan Pustaka
	Dalam bab ini berisikan mengenai tinjauan pustaka yang berisikan penelitian terdahulu serta teori dasar yang mendukung tentang klasifikasi masa studi mahasiswa.
BAB III	Metodologi Penelitian
	Dalam bab ini berisi tahap-tahapan penelitian, lokasi dan waktu penelitian, data yang digunakan, teknik pengumpulan data dan metode / pemodelan yang digunakan.
BAB IV	Jadwal Pelaksanaan Penelitian

Bab ini membahas tentang jadwal penelitian proposal.

DAFTAR PUSTAKA

BAB VI

KESIMPULAN DAN SARAN

A. Kesimpulan

Kesimpulan dari penelitian ini adalah masing-masing dari algoritma Support vector machine, Iterative dichotomiser 3, Random forest dan K-nearest neighbor mampu melakukan klasifikasi terhadap data yang ada. Algoritma Random forest menjadi algoritma yang paling baik tingkat keakuratannya dibanding ketiga algoritma lain dengan persentase keakuratan data senilai 100%. Mengikuti algoritma Support vector machine dengan 94,4% serta Iterative dichotomiser dan K-nearest neighbor dengan nilai persentase masing-masing 90%.

B. Saran

1. Penelitian klasifikasi lama masa studi ini hanya analisis terhadap data yang ada dilapangan maka diperlukan beberapa faktor luar (eksternal) yang dapat dijadikan variabel dalam menentukan lama masa studi seperti kondisi psikologis, kondisi finansial, keluarga dll.
2. Diperlukan penambahan dataset agar hasil akurasi lebih baik

DAFTAR PUSTAKA

- [1] A. Redaksi, P. Jawab, and K. Panitia, "PROCEEDING SENTIKA 2015."
- [2] B. Santosa, "1 . Ide Dasar Support Vector Machine," no. x, 2015.
- [3] D. Fransiska Amalia Kurniawan, "Analisis dan Implementasi Random Forest dan Regression Tree (CART) Untuk Klasifikasi pada Misuse Intrusion Detection System," *Fak. Tek. Inform.*, no. Data Mining, pp. 1–7, 2011.
- [4] D. Ispriyanti and A. Hoyyi, "Analisis klasifikasi masa studi mahasiswa prodi statistika undip dengan metode support vector machine (svm) dan id3 (iterative dichotomiser 3) 1,2," vol. 9, no. 1, pp. 15–29, 2016, doi: 10.14710/medstat.9.1.15-29.
- [5] Paul M. Muchinsky, "濟無No Title No Title," *Psychol. Appl. to Work An Introd. to Ind. Organ. Psychol. Tenth Ed. Paul*, vol. 53, no. 9, pp. 1689–1699, 2012, doi: 10.1017/CBO9781107415324.004.
- [6] R. A. Permana and S. Sahara, "Metode Support Vector Machine Sebagai Penentu Kelulusan Mahasiswa pada Pembelajaran Elektronik," *J. Khatulistiwa Inform.*, vol. 7, no. 1, pp. 50–58, 2019, doi: 10.31294/jki.v7i1.5743.
- [7] R. A. Saputra *et al.*, "PENERAPAN METODE ITERATIVE DICHOTOMIZER 3 (ID 3) UNTUK MENENTUKAN BEASISWA BERPRESTASI PADA SMP PGRI," vol. 15, no. 1, pp. 35–40, 2019.