

**IMPLEMENTASI ALGORITMA SVM UNTUK KLASIFIKASI BIDANG
DOSEN BERDASARKAN JUDUL SKRIPSI**

SKRIPSI



disusun oleh

Lasmita Triana

18.11.2294

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2022

**IMPLEMENTASI ALGORITMA SVM UNTUK KLASIFIKASI BIDANG
DOSEN BERDASARKAN JUDUL SKRIPSI**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Lasmita Triana

18.11.2294

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2022

PERSETUJUAN

SKRIPSI

IMPLEMENTASI ALGORITMA SVM UNTUK KLASIFIKASI BIDANG DOSEN BERDASARKAN JUDUL SKRIPSI

yang dipersiapkan dan disusun oleh

Lasmita Triana

18.11.2294

telah disetujui oleh Dosen Pembimbing Skripsi
pada 20 Maret 2022

Dosen Pembimbing,

Windha Mega Pradnya D. M.Kom.
NIK. 190302185

PENGESAHAN

SKRIPSI

**IMPLEMENTASI ALGORITMA SVM UNTUK KLASIFIKASI
KLASIFIKASI BIDANG DOSEN BERDASARKAN JUDUL SKRIPSI**

yang dipersiapkan dan disusun oleh

Lasmata Triana

18.11.2294

telah dipertahankan di depan Dewan Penguji
pada 17 Februari 2022

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Mardhiva H.S.T M.Kom.
NIK. 190302108

Anna Balta, M.Kom.
NIK. 190302290

Windha Mega, M.Kom.
NIK. 190302185

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
02 Maret 2022

DEKAN FAKULTAS ILMU KOMPUTER

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggung jawab saya pribadi.

Yogyakarta, 21 Maret 2022

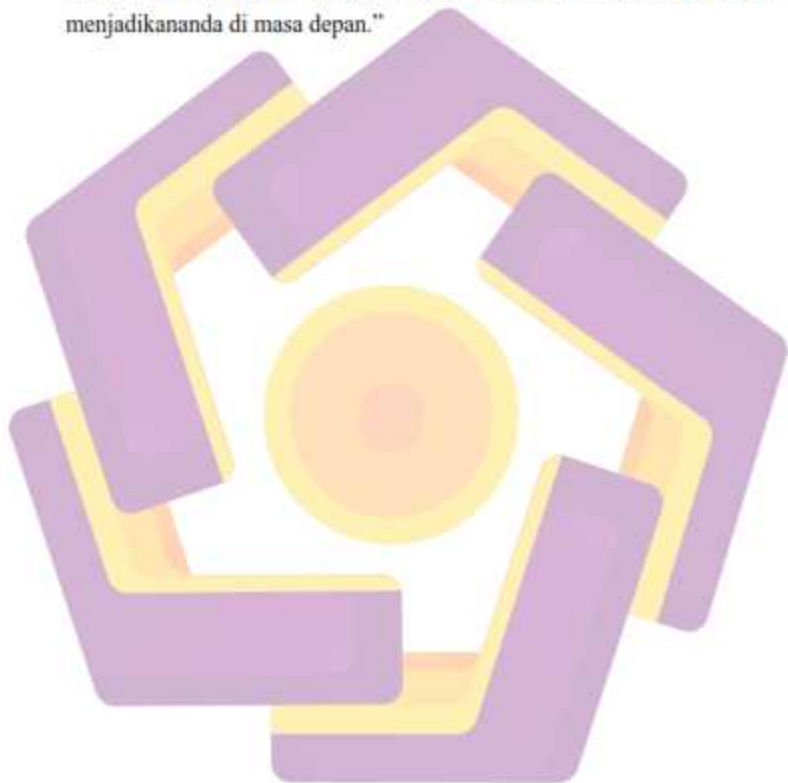


Lasmita Triana
NIM. 18.11.2294

MOTTO

“Selalu berpikir positif terhadap segala hal akan membantumu dalam menyelesaikan setiap permasalahan yang kamu alami.”

Sikap anda di masa lalu menjadikan anda hari ini, sikap anda hari ini akan menjadikan anda di masa depan.”



PERSEMBAHAN

Saya mempersembahkan skripsi ini kepada semua pihak yang terlibat secara langsung maupun tidak langsung dalam proses pembuatan skripsi.

1. Tuhan Yang Maha Esa yang memberikan segala nikmat dan kasih sayangnya sampai sejauh ini.
2. Kedua orang tua saya dan keluarga, yang selalu mendoakan, selalu menyemangati dan memberikan dukungan baik secara material maupun non material.
3. Ibu Windha Mega P.D, M.Kom. yang telah membimbing saya dari awal sampai akhir pembuatan skripsi.
4. Dosen-dosen Universitas AMIKOM Yogyakarta yang telah memberikan banyak ilmu selama kuliah.
5. Teman-teman diskusi dan sharing, Nurrofiqi Ankisqiantari, Emilia Tsiompah, Agita Putri Kusumaningtiyas, Farhan Khoirul Amri, Almuzhidul Mujhid, Fanani Amalia, dan Adi Zaenul Mustaqim.
6. Teman skripsi saya, Emilia Tsiompah dan Nurrofiqi Ankisqiantari, yang memberikan solusi, tempat bertanya ketika dimasa sulit dan tempat untuk wif-an hingga pembuatan skripsi selesai.
7. Teman-teman kelas IF07 2018 yang selalu menemani perkuliahan, mendukung dan memberikan semangat sampai saat ini. Semoga kita selalu bahagia dan menjadi pribadi yang lebih baik lagi.

KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat, hidayah dan kekuatan sehingga saya dapat menyelesaikan skripsi yang berjudul Implementasi Algoritma SVM untuk Klasifikasi Bidang Dosen Berdasarkan Judul Skripsi.

Skripsi ini saya buat guna menyelesaikan studi jenjang Strata Satu (S1) pada program studi Informatika fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kuliah jenjang program strata satu dan untuk memperoleh gelar Sarjana Komputer. Dengan selesainya skripsi ini, maka pada kesempatan ini saya mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Hanif Al Fatta, M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Windha Mega P.D, M.Kom. selaku dosen pembimbing yang selalu bijaksana memberikan bimbingan, nasehat serta waktunya selama penulisan skripsi ini.
4. Dosen Penguji dan segenap Dosen serta Karyawan Universitas AMIKOM Yogyakarta yang telah berbagi ilmu dan pengalamannya.
5. Kedua orang tua saya yang telah mendoakan, mendukung dan memberikan semangat kepada saya.

6. Orang-orang baik yang senantiasa mendoakan, memberi semangat dan bantuan kepada saya.
7. Teman-teman diskusi dan sharing, Nurrofiqi, Emilia, Agita Putri, Farhan Khoirul Amri, Almuzhidul Mujhid dan Adi Zaenul Mustaqim.
8. Teman-teman IF07 angkatan 2018 yang telah menemani selama proses perkuliahan.
9. Semua pihak yang telah membantu baik secara langsung maupun tidak langsung.

Semoga Allah subhanahu wa ta'ala memberikan balasan yang lebih kepada semua yang telah ikut membantu saya hingga menyelesaikan skripsi ini. Demi perbaikan selanjutnya, saran dan kritik yang membangun diterima dengan senang hati dan rasa terima kasih. Semoga skripsi ini dapat bermanfaat bagi saya dan kita semua.

Yogyakarta, 21 Februari 2022

Lasmita Triana

DAFTAR ISI

SKRIPSI.....	1
SKRIPSI.....	i
PERSETUJUAN	ii
PENGESAHAN.....	iii
PERNYATAAN.....	iv
MOTTO	v
PERSEMBAHAN.....	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
INTISARI.....	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN.....	1
1.1. LATAR BELAKANG.....	1
1.2. RUMUSAN MASALAH.....	2
1.3. BATASAN MASALAH.....	2
1.4. MAKSUD DAN TUJUAN PENELITIAN	3
1.5. MANFAAT PENELITIAN.....	3
1.6. METODE PENELITIAN	4
1.6.1. Metode Pengumpulan Data.....	4
1.6.2. Metode Analisis	4
1.6.3. Metode Implementasi.....	5
1.6.4. Metode Testing.....	5

1.7.	SISTEMATIKA PENELITIAN	5
BAB II LANDASAN TEORI.....		7
2.1.	KAJIAN PUSTAKA.....	7
2.2.	DASAR TEORI.....	13
2.2.1.	Data Mining	13
2.2.2.	Text Mining.....	13
2.2.3.	Klasifikasi	14
2.2.4.	Preprocessing	14
2.2.5.	SVM.....	15
2.2.6.	TF-IDF	17
2.2.7.	Confusion Matrix.....	17
2.2.8.	Black Box Testing.....	18
2.2.9.	Flask.....	19
2.2.10.	Jupyter Notebook.....	19
2.2.11.	Visual Studio Code.....	20
BAB III ANALISIS DAN PERANCANGAN		21
3.1.	GAMBARAN UMUM.....	21
3.2.	ALAT DAN BAHAN PENELITIAN.....	22
3.2.1.	Perangkat Keras	22
3.2.2.	Perangkat Lunak.....	24
3.3.	ALUR PENELITIAN.....	23
3.3.1.	Pengumpulan Data.....	25
3.4.	PREPROCESSING.....	29
3.4.1.	Case Folding	29
3.4.2.	Tokenizing.....	31
3.4.3.	Stopword Removal.....	34
3.4.4.	Data Split.....	34
3.4.5.	Pembobotan Kata	36
3.5.	HYPERPARAMETER SVM.....	54
3.6.	KLASIFIKASI ALGORITMA SVM.....	54

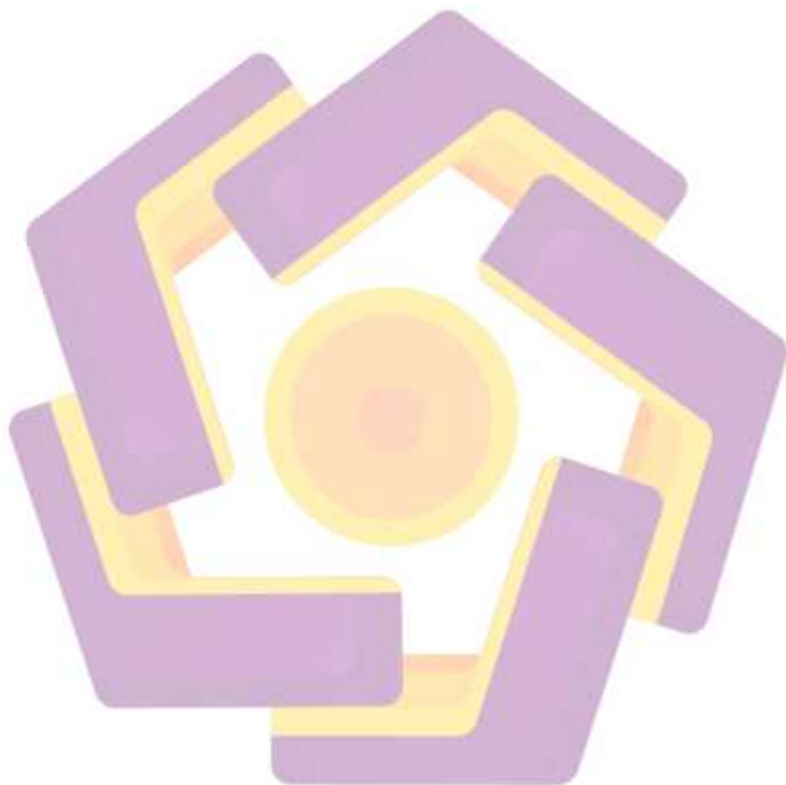
3.7. CONFUSION MATRIX	60
BAB IV IMPLEMENTASI DAN PEMBAHASAN	63
4.1. DESKRIPSI IMPLEMENTASI	63
4.2. IMPLEMENTASI PENGUMPULAN DATA	63
4.3. IMPLEMENTASI PERANCANGAN FUNGSI DAN SISTEM.....	64
4.4. IMPLEMENTASI PELABELAN DATA.....	64
4.5. IMPLEMENTASI PERANCANGAN FUNGSI SISTEM	65
4.5.1 Implementasi Fungsi Preprocessing Data	65
4.5.2 Implementasi Data Split.....	67
4.5.3 Implementasi TF-IDF.....	67
4.5.4 IMPLEMENTASI HYPERPARAMETER SVM	68
4.5.5 Implementasi Algoritma Support Vector Machine.....	69
4.5.6 Implementasi Confusion Matrix	71
4.6. IMPLEMENTASI INTERFACE	73
BAB V PENUTUP	75
5.1. KESIMPULAN.....	75
5.2. SARAN.....	75
DAFTAR PUSTAKA	77
DAFTAR TABEL	
Tabel 2.1 Perbedaan Penelitian.....	8
Tabel 2.2 Rumus Confuion Matrix	15
Tabel 3.1 Perangkat Keras	19
Tabel 3.2 Perangkat Lunak	20
Tabel 3.3 Dataset Utuh.....	22
Tabel 3.4 Labeling Data Berdasarkan Tema.....	25
Tabel 3.5 Labeling Data.....	26
Tabel 3.6 Case Folding	26

Tabel 3.7 Menghilangkan Angka dan Link.....	28
Tabel 3.8 Menghilangkan Tanda Baca.....	29
Tabel 3.9 Menghilangkan Spasi Berlebih.....	30
Tabel 3.10 Stopword.....	31
Tabel 3.11 Term Kata.....	33
Tabel 3.12 Perhitungan TF-IDF.....	38
Tabel 3.13 Hasil TF-IDF.....	44
Tabel 3.14 Transpose Data.....	55
Tabel 3.15 Confusion Matrix.....	60
Tabel 3.12 Perhitungan TF-IDF.....	61
Tabel 4.1 Hasil <i>Confusion Matrix</i> 90 % dan 10 %.....	71
Tabel 4.2 Hasil <i>Confusion Matrix</i> 80% dan 20 %.....	72

DAFTAR GAMBAR

Gambar 3.1 Diagram Alur Penelitian.....	21
Gambar 3.2 Gambar Data Split.....	32
Gambar 3.3 Interface Antar Muka.....	62
Gambar 4.1 Dataset Utuh.....	63
Gambar 4.2 File <i>xlx</i>	63
Gambar 4.3 Pelabelan Data.....	64
Gambar 4.4 Case Folding.....	64
Gambar 4.5 Menghilangkan Angka dan Link.....	66
Gambar 4.6 Menghilangkan Tanda Baca.....	66
Gambar 4.7 Menghilangkan Spasi Tidak Perlu.....	66
Gambar 4.8 Stopword Removal.....	67
Gambar 4.9 Data Split.....	67
Gambar 4.10 Implementasi TF-IDF.....	67
Gambar 4.11 Implementasi Hyperparameter SVM.....	68
Gambar 4.12 Implementasi Best Hyperparameter.....	68
Gambar 4.13 Implementasi Algoritma SVM.....	69

Gambar 4.14 Implementasi One Vs One Classifier	70
Gambar 4.15 Implementasi Testing	70
Gambar 4.16 Implementasi Confuion Matrix	71
Gambar 4.17 Implementasi Interface.....	73



INTISARI

Skripsi merupakan salah satu syarat bagi mahasiswa tingkat akhir untuk lulus dari perguruan tinggi jenjang S1. Bidang dosen merupakan kemampuan atau kemahiran yang bisa dilakukan oleh seorang dosen pada sebuah peran tertentu yang dikuasai setelah melalui pengalaman tertentu. Dalam menyusun skripsi, mahasiswa dibantu oleh dosen pembimbing.

Mahasiswa sering mengeluhkan dosen pembimbing yang tidak sesuai dengan bidang dosen hal ini mengakibatkan sulitnya dalam proses bimbingan. Dengan munculnya permasalahan tersebut peneliti melakukan klasifikasi bidang dosen yang sesuai dengan judul skripsi mahasiswa. Data diperoleh dari Universitas AMIKOM Yogyakarta Program Studi Informatika yang terdiri dari 1598 data.

Pada penelitian ini pertama melakukan proses labelisasi data, kemudian melakukan proses *Preprocessing*, diteruskan dengan proses *TF-IDF*, lalu data tersebut diklasifikasikan menggunakan algoritma *SVM*, apakah dataset tersebut sudah sesuai dengan bidang dosen berdasarkan judul skripsi mahasiswa. Bidang dosen Program studi Informatika yaitu terdiri dari *Data Science*, *Multimedia*, *Networking*, *Software Engineer*, *Game* dan *Tidak Ada Bidang Dosen*. Berdasarkan hasil penelitian hasil akurasi diperoleh sebesar 89.244%.

Kata kunci: *SVM*, *Text Classification*, *Preprocessing*, Bidang Dosen.

ABSTRACT

This thesis is one of the requirements for final year students to graduate from undergraduate tertiary institutions. The field of lecturers is an ability or skill that can be carried out by a lecturer in a certain role that is mastered after going through certain experiences. In compiling the thesis, students are assisted by a supervisor.

Students often complain about supervisors who are not in accordance with the lecturer's field, this causes difficulties in the guidance process. With the emergence of these problems, researchers classified the fields of lecturers according to the title of the student thesis. Data were obtained from AMIKOM University Yogyakarta Informatics Study Program which consisted of 1598 data.

In this study, firstly the data labeling process is carried out, then the preprocessing process is carried out, followed by the TF-IDF process, then the data is classified using the SVM algorithm, whether the dataset is in accordance with the lecturer's field based on the title of the student thesis. The field of lecturers in the Informatics study program consists of Data Science, Multimedia, Networking, Software Engineer, Game and No Lecturer. Based on the results of the study, the accuracy obtained was 89.244%.

Keyword: SVM, Text Classification, Preprocessing, Lecturer Field.