



IMPLEMENTASI EKSTRAKSI INFORMASI PADA SURAT KEPUTUSAN MENGUNAKAN METODE SUPPORT VECTOR MACHINE

Ketua Penguji

Kania Evita Dewi, S.Pd., M.Si.

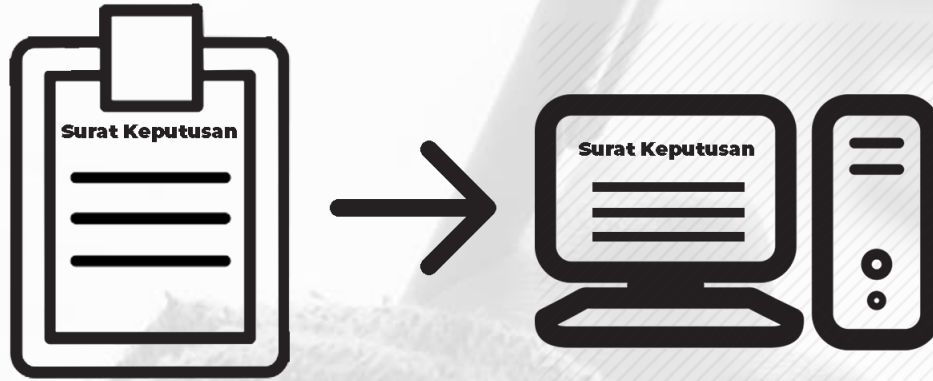
Penguji 1

Nelly Indriani W, S.Si., M.T.

Penguji 2

Iskandar Ikbal, S.T., M.Kom.

LATAR BELAKANG MASALAH



**Support Vector
Machines**



1

IDENTIFIKASI MASALAH



Karakteristik pada dokumen surat keputusan tiap instansi berbeda sehingga di butuhkan klasifikasi untuk mengekstraksi informasi di dalamnya. Metode SVM belum pernah di terapkan pada ekstraksi informasi suatu dokumen surat keputusan.

2

MAKSUD DAN TUJUAN



Maksud

- Membuat sistem ekstraksi informasi pada surat.
- Mengimplementasi SVM dalam ekstraksi informasi pada dokumen surat keputusan.



Tujuan

- Mengetahui informasi yang terstruktur pada dokumen surat keputusan setelah diklasifikasi.
- Mengukur performa SVM dalam ekstraksi informasi pada dokumen surat keputusan dengan baik.

3

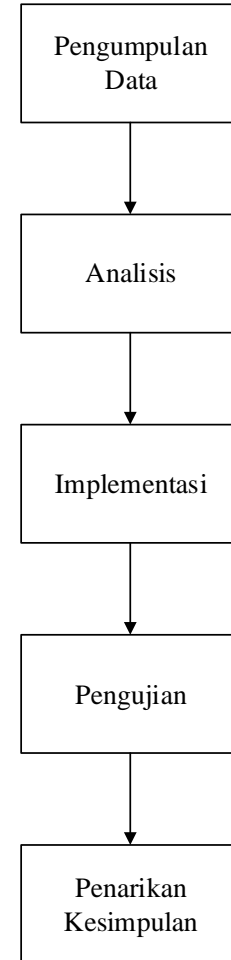
BATASAN MASALAH



- Data yang dipakai dalam bentuk teks dengan ekstensi pdf (*.pdf).
- Data training dan data testing diperoleh dari internet ataupun umum.
- Data merupakan surat keputusan yang ditujukan hanya kepada satu orang.
- Data minimal terdiri dari nomor surat, nama penerima surat keputusan dan waktu ditetapkan nya surat keputusan
- Data harus merupakan gambar dengan kualitas yang baik.

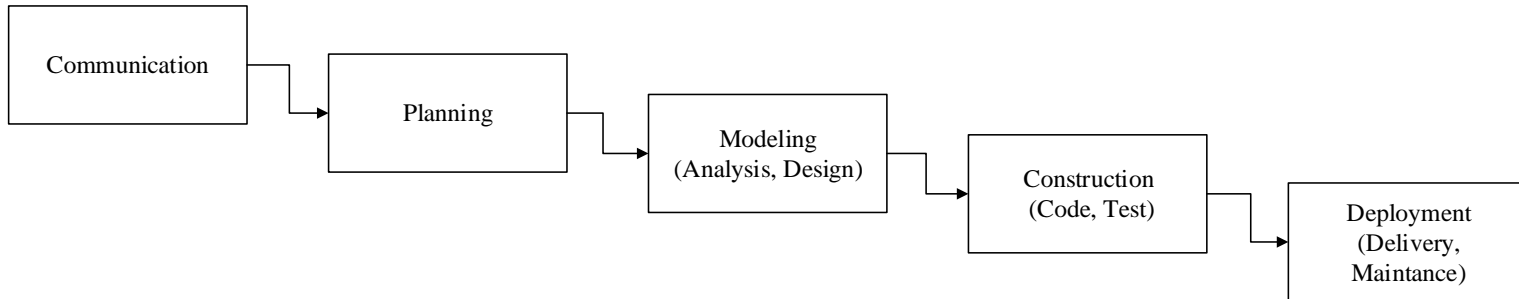
4 Metode Penelitian

Metode Penelitian yang digunakan adalah metode deskriptif. Metode deskriptif merupakan metode yang menggambarkan fakta - fakta dan informasi dalam situasi atau kejadian sekarang secara sistematis, faktual dan akurat.



Tahap Implementasi

Metode *waterfall* adalah metode yang melakukan pendekatan secara sistematis dan berurutan. Metode ini disebut dengan *waterfall* karena tahap demi tahap yang dilalui harus menunggu selesainya tahap sebelumnya dan berjalan berurutan.



ANALISIS MASALAH

Agni Ismaya



Rule Base
Classification



98,27%

Data Latih



89,77%

Data Uji



“Wordmatch”



PEMERINTAH KABUPATEN GARUT
BADAN KEPEGAWAIAN DAN DIKLAT
 Jl. Pahlawan No. 47 Tlp. (0262) 243003 Garut

Memiliki
header

PETIKAN
KEPUTUSAN BUPATI GARUT
 Nomor : 823.2/KEP.030-BKD/TV/2014
 TENTANG

KENAikan PANGKAT PEGAWAI NEGERI SIPIL

Menimbang : ds;
 Mengingat : ds;
 Memperhatikan : Penetapan teknis Kepala Kantor Regional III BKN Nomor DG-23205000929 tanggal 27 Maret 2014

MEMUTUSKAN

Menetapkan
 PERTAMA : Pegawai Negeri Sipil, nomor urut : 16
Nama : KAMALUDIN
Tempat/Tanggal Lahir : GARUT/08-12-1978
NIP : 1978120820010011016
Pendidikan : MADRASAH ALYAH Tahun : 1998
Pangkat lama/golongan Ruang/TMT : Pengantar Muda / IIIa / 01-01-2010
Jabatan : SEKRETARIS DESA
Masa Kerja Golongan : 05 tahun 07 bulan
Gaji Pokok : Rp. 1.971.000
Unit Kerja : DESA KARYAMEKAR KEC. CLAWU
 PEMERINTAH KABUPATEN GARUT

Terhitung mulai tanggal **01-04-2014** berdasarkan Kenaikan Pangkat **Reguler** dinaikkan pangkatnya menjadi **Pengantar Muda Tingkat I** golongan ruang **IIIa**, dengan masa kerja golongan **09 tahun 07 bulan**, dan diberikan gaji pokok sebesar **Rp. 2.054.400** ditambah dengan penghasilan lain berdasarkan ketentuan peraturan perundang-undangan yang berlaku.

KEDUA : Apabila dikemudian hari ternyata terdapat kekeliruan dalam keputusan ini, akan diadakan perbaikan dan penghitungan kembali sebagaimana mestinya.

KETIGA : Petikan keputusan ini diberikan kepada yang bersangkutan, dan yang berkepentingan untuk dapat diketahui dan dipergunakan sebagaimana mestinya.

Ditetapkan di : GARUT
 Pada tanggal : 28 Maret 2014

BUPATI GARUT

ms

H. RUDY GUNAWAN, SH, MH, MP



6844329d44ff.4a801450c12968f4f69



MUHLIS, S.Ses
 Pembina
 NIP. 196507171990031007

Memiliki barcode surat
 tapi tidak memiliki
 tembusan

KEPUTUSAN MENTERI PENDIDIKAN NASIONAL,
 REPUBLIK INDONESIA
 NOMOR : 77801/A4.3/PP/2010
 MENTERI PENDIDIKAN NASIONAL,

Tidak
Memiliki
header

Susunan

dimulai dari
kata
membaca

Membaca : Surat Rektor Universitas Negeri Yogyakarta nomor 557/H34/KP/2010 tanggal 10 Mei 2010, hal usul pemberian tugas belajar a.n. Sdr. Drs. Arief Rohman, M.Si.
 Menimbang : a. bahwa sesuai dengan rencana kebutuhan Fakultas Ilmu Pendidikan Universitas Negeri Yogyakarta perlu dilakukan peningkatan kompetensi sumber daya manusia melalui program pemberian tugas belajar;
 b. bahwa berdasarkan surat Direktur Program Pascasarjana Universitas Negeri Yogyakarta nomor 2879/H34.17/PP/2009 tanggal 1 September 2009, Sdr. Drs. Arief Rohman, M.Si., dinyatakan diterima sebagai mahasiswa program studi Ilmu Pendidikan;
 c. bahwa sehubungan dengan huruf a dan b di atas, dipandang perlu memberikan tugas belajar kepada yang bersangkutan.
 Mengingat : 1. Undang-undang Nomor 6 Tahun 1974 jo Nomor 43 Tahun 1999,
 2. Undang-undang Nomor 14 Tahun 2005,
 3. Peraturan Presiden RI Nomor 12 Tahun 1991;
 4. Keputusan Presiden RI Nomor 84/P Tahun 2009;
 5. Keputusan Menteri Pertama Nomor 224/MP/1961;
 6. Peraturan Menteri Pendidikan Nasional Nomor 48 Tahun 2009,
 7. Peraturan Menteri Pendidikan Nasional Nomor 61 Tahun 2009.

Memperhatikan : Surat Edaran Menteri Negara Pendayagunaan Aparatur Negara nomor SE/ISM/PAN/S/2004 tanggal 14 Mei 2004.

MEMUTUSKAN

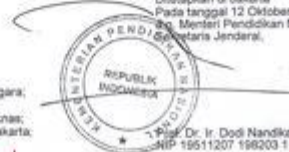
Menetapkan
 Pertama : Memberikan tugas belajar kepada Pegawai Negeri Sipil,
Nama : Drs. Arief Rohman, M.Si.
NIP : 196703291994121002
Tempat, tanggal lahir : Demak, 29 Maret 1967
Pangkat, golongan ruang : Pembina, IV/a
Jabatan : Lektor Kepala
Unit kerja : Universitas Negeri Yogyakarta
 untuk mengikuti Program S3 di bidang Ilmu Pendidikan pada Program Pascasarjana Universitas Negeri Yogyakarta.

Data diri
memiliki
susunan
yang
berbeda

Kedua : Tugas belajar sebagaimana tersebut pada dikum Pertama, dilaksanakan mulai bulan September 2009 sampai dengan Agustus 2012.
 Ketiga : Wajib melaksanakan iktan dinas selama 1n+1 dari masa belajarnya.

Keputusan ini diberikan kepada yang bersangkutan untuk diketahui dan dilaksanakan.

Ditetapkan di Jakarta
 Pada tanggal 12 Oktober 2010
 Dr. Menteri Pendidikan Nasional
 Sekretaris Jenderal,



Tembusan :
 1. Menteri Pendidikan Nasional;
 2. Kepala Badan Kepegawaian Negara;
 3. Inspektur Jenderal Kemdiknas;
 4. Dirjen Pendidikan Tinggi Kemdiknas;
 5. Rektor Universitas Negeri Yogyakarta;
 6. Kepala KPPN Yogyakarta.

01/10/2010 09:10:58

Memiliki tembusan tapi tidak memiliki barcode

ANALISIS MASALAH

Support Vector Machines



Yaoyong Li **Kuspriyanto**

“SVM Based Learning
System For Information
Extraction”

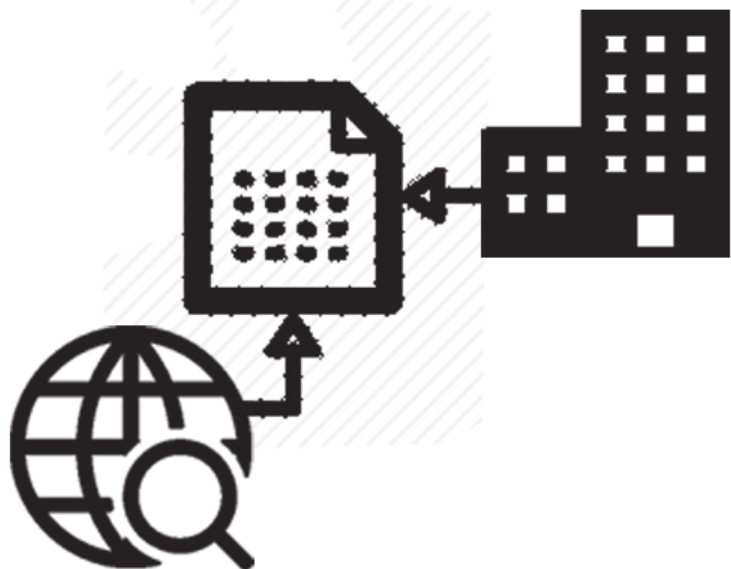
“Performance Evaluation of
SVM-Based Information
Extraction using τ
Margin Values”

Kernel Linear



2

Analisis Data Masukan



BATASAN DATA MASUKAN :

1

Memiliki

Nomor : 823.2/KEP.030-BKD/TV /2014

2

Nomor : 823.2/KEP.030-BKD/TV /2014

Ditetapkan di : GARUT
Pada tanggal : 28 Maret 2014

Pada 1 Halaman

3



4



**Tidak
Termasuk**

“

Contoh Masukan Untuk Data Latih



PEMERINTAH KABUPATEN GARUT
BADAN KEPEGAWAIAN DAN DIKLAT
Jl. Pahlawan No. 47 Tlp. (0262) 243003 Garut

**PETIKAN
KEPUTUSAN BUPATI GARUT
Nomor : 823.2/KEP.030-BKD/IV /2014
TENTANG
KENAIKAN PANGKAT PEGAWAI NEGERI SIPIL**

Menimbang : dst;
Mengingat : dst;
Memperhatikan : Persetujuan teknis Kepala Kantor Regional III BKN
Nomor **DG-23205000929** tanggal 27 Maret 2014

MEMUTUSKAN

Menetapkan :
PERTAMA : Pegawai Negeri Sipil, nomor urut : 16
Nama : **KAMALUDIN**
Tempat/Tanggal Lahir : GARUT/08-12-1978
NIP : 197812082010011016
Pendidikan : MADRASAH ALYIAH Tahun : 1998
Pangkat lama/golongan Ruang/TMT : Pengatur Muda / II/a / 01-01-2010
Jabatan : SEKRETARIS DESA
Masa Kerja Golongan : 09 tahun 07 bulan
Gaji Pokok : Rp. 1.971.000
Unit Kerja : DESA KARYAMEKAR KEC. CILAWU
PEMERINTAH KABUPATEN GARUT

Terhitung mulai tanggal **01-04-2014** berdasarkan Kenaikan Pangkat **Reguler** dinaikkan pangkatnya menjadi **Pengatur Muda Tingkat I** golongan ruang **II/b**, dengan masa kerja golongan **09 tahun 07 bulan**, dan diberikan gaji pokok sebesar **Rp. 2.054.400** ditambah dengan penghasilan lain berdasarkan ketentuan peraturan perundang-undangan yang berlaku.

KEDUA : Apabila dikemudian hari ternyata terdapat kekeliruan dalam keputusan ini, akan diadakan perbaikan dan penghitungan kembali sebagaimana mestinya.
KETIGA : Petikan keputusan ini diberikan kepada yang bersangkutan, dan yang berkepentingan untuk dapat diketahui dan dipergunakan sebagaimana mestinya.

Ditetapkan di : GARUT
Pada tanggal : 28 Maret 2014

BUPATI GARUT

ttd

H. RUDY GUNAWAN, SH.,MH.,MP



MUHLIS, S.Sos
Pembina
NIP. 196507171990031007



5ae4829d44ff14aa01450c1296bf4f69



Contoh Masukan Untuk Data Uji

KEPUTUSAN MENTERI PENDIDIKAN NASIONAL
REPUBLIK INDONESIA
NOMOR : 77601/A4.3/KP/2010

MENTERI PENDIDIKAN NASIONAL

- Membaca : Surat Rektor Universitas Negeri Yogyakarta nomor 557/H34/KP/2010 tanggal 10 Mei 2010, hal usul pemberian tugas belajar a.n. Sdr. Drs. Arief Rohman, M.Si.
- Menimbang : a. bahwa sesuai dengan rencana kebutuhan Fakultas Ilmu Pendidikan Universitas Negeri Yogyakarta perlu dilakukan peningkatan kompetensi sumber daya manusia melalui program pemberian tugas belajar;
- b. bahwa berdasarkan surat Direktur Program Pascasarjana Universitas Negeri Yogyakarta nomor 2879/H34.17/PP/2009 tanggal 1 September 2009, Sdr. Drs. Arief Rohman, M.Si., dinyatakan diterima sebagai mahasiswa program studi Ilmu Pendidikan;
- c. bahwa sehubungan dengan huruf a dan b di atas, dipandang perlu memberikan tugas belajar kepada yang bersangkutan.
- Mengingat : 1. Undang-undang Nomor 8 Tahun 1974 jo Nomor 43 Tahun 1999;
2. Undang-undang Nomor 14 Tahun 2005;
3. Peraturan Presiden RI Nomor 12 Tahun 1961;
4. Keputusan Presiden RI Nomor 84/P Tahun 2009;
5. Keputusan Menteri Pertama Nomor 224/MP/1961;
6. Peraturan Menteri Pendidikan Nasional Nomor 48 Tahun 2009;
7. Peraturan Menteri Pendidikan Nasional Nomor 61 Tahun 2009;
- Memperhatikan : Surat Edaran Menteri Negara Pendayagunaan Aparatur Negara nomor SE/18/M.PAN/5/2004 tanggal 14 Mei 2004;

MEMUTUSKAN

- Menetapkan :
Pertama : Memberikan tugas belajar kepada Pegawai Negeri Sipil,
- N a m a : Drs. Arief Rohman, M.Si.
N I P : 19670329 199412 1 002
Tempat, tanggal lahir : Demak, 29 Maret 1967
Pangkat, golongan ruang : Pembina, IV/a
Jabatan : Lektor Kepala
Unit kerja : Universitas Negeri Yogyakarta
- untuk mengikuti Program S3 di bidang Ilmu Pendidikan pada Program Pascasarjana Universitas Negeri Yogyakarta.
- Kedua : Tugas belajar sebagaimana tersebut pada diktum Pertama, dilaksanakan mulai bulan September 2009 sampai dengan Agustus 2012.
- Ketiga : Wajib melaksanakan ikatan dinas selama 1n+1 dari masa belajarnya.
- Keputusan ini diberikan kepada yang bersangkutan untuk diketahui dan dilaksanakan.

Ditetapkan di Jakarta
Pada tanggal 12 Oktober 2010
a.n. Menteri Pendidikan Nasional
Sekretaris Jenderal,



- Tembusan :
1. Menteri Pendidikan Nasional;
 2. Kepala Badan Kepegawaian Negara;
 3. Inspektur Jenderal Kemdiknas;
 4. Dirjen Pendidikan Tinggi Kemdiknas;
 5. Rektor Universitas Negeri Yogyakarta;
 6. Kepala KPPN Yogyakarta.

Prof. Dr. Ir. Dodi Nandika, M.S.
NIP 19511207 198203 1 001*

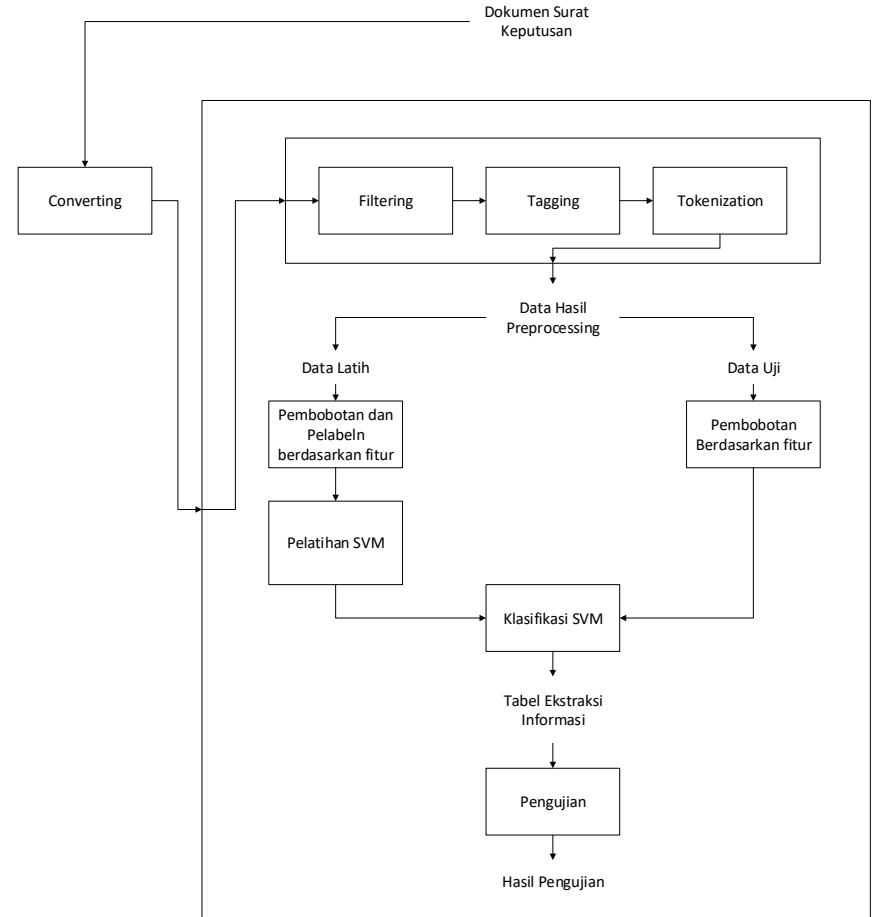
2

Analisis Sistem

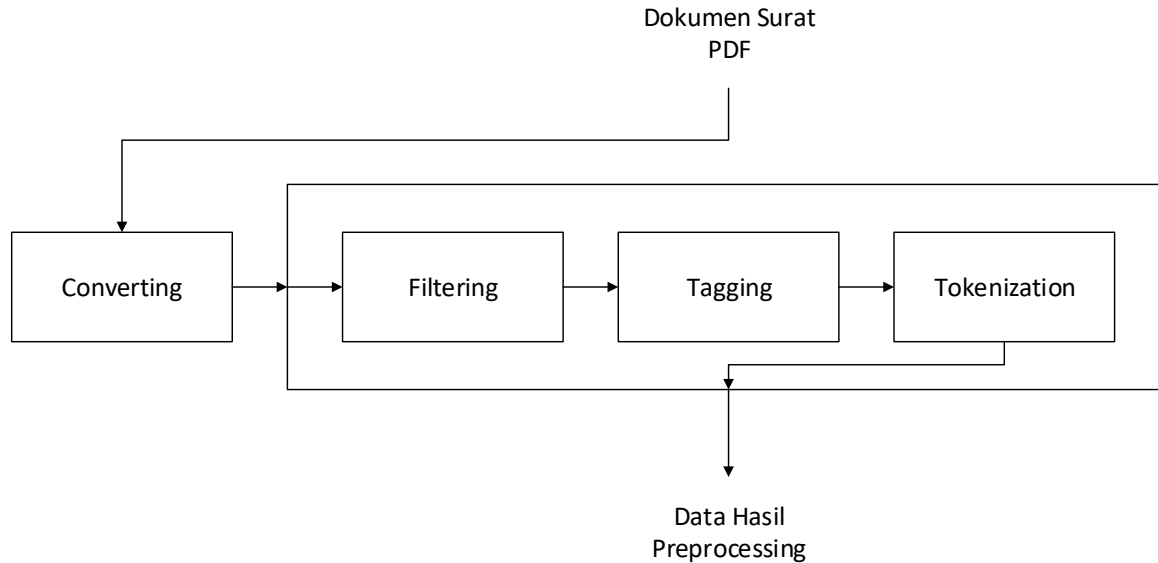
1. Analisis Perprocessing
2. Analisis Pembobotan
3. Analisis Algoritma SVM



Gambaran Analisis Sistem

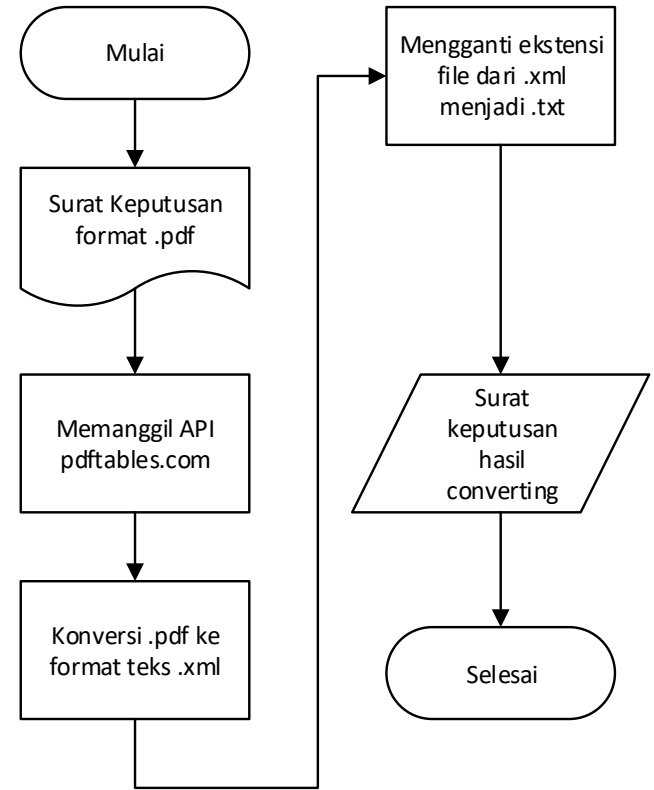


1 Analisis Preprocessing



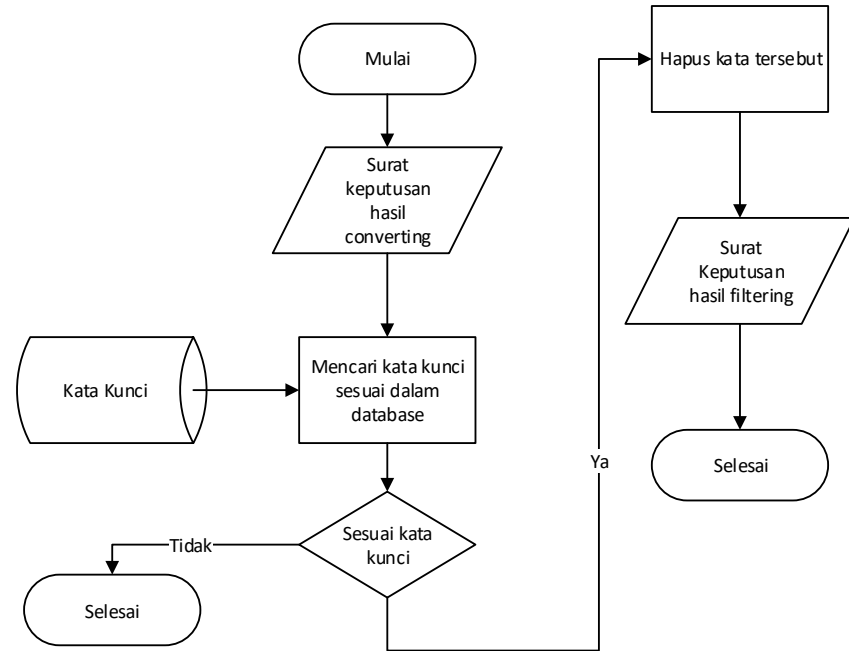
Converting

Converting merupakan suatu tahapan mengubah dokumen dengan format .pdf menjadi format .xml. Hal ini dapat memudahkan saat pembacaan *string* oleh sistem.



Filtering

Tahap *filtering* pada penelitian ini dilakukan untuk menghilangkan karakter karakter yang muncul akibat konversi seperti tag html dan lain nya, sehingga hanya menyisakan teks surat yang akan diekstraksi dan tanda yang ingin dipertahankan



Aturan *Filtering*

Tanda/Kata	Aksi
p, td, tr, </td> </tr>	Dihapus
colspan="2",..., colspan="12"	
data-page="1" data-table="1"	
Style="text-align: right"	
Selain tanda/kata diatas	Tidak ada aksi

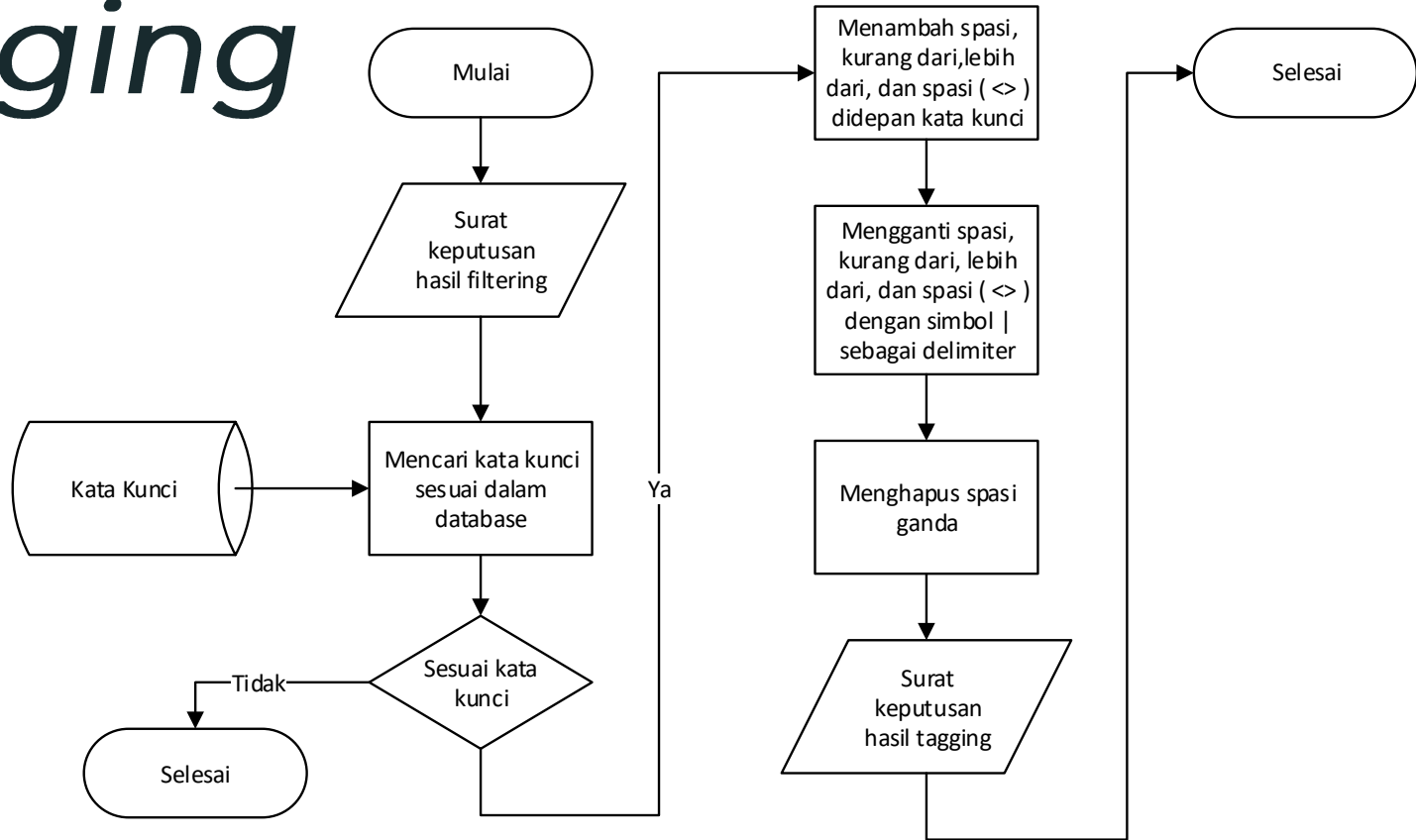
Hasil *Filtering*

<><><><><><>PEMERINTAH KABUPATEN GARUT<><><><><>BADAN
KEPEGAWAIAN DAN DIKLAT<><><><><>JI. Pahlawan No. 47 Tip. (0262) 243003
Garut<><><><><><><>PETIKAN<><><><><><><>KEPUTUSAN BUPATI
GARUT<><><><><><><>Nomor : 823.2/KEP.030-BKD/IV
/2014<><><><><><><><>TENTANG<><><><><><><>KENAIKAN
PANGKAT PEGAWAI NEGERI
SIPIL<><><><><>Menirnbang<>:<>dst;<><><><><><><>Mengingat<>:<>dst;<><>
><><><><><><>Memperhatikan<>:<>Persetujuan teknis Kepala Kantor<>Regional III
BKN<><><><><><>Nomor DG-2320S000929 tanggal 27 Maret
2014<><><><><><><><>MEMUTUSKAN<><><><><>

Tagging

Tahapan tagging bertujuan untuk memberikan tanda pada data hasil filtering yang akan menjadi penanda awal atau akhir suatu kalimat yang akan di ekstraksikan. Pada tahapan ini, ada beberapa aturan proses agar hasil *tagging* kepada kata kunci dapat sesuai dengan yang di inginkan. Adapun tag yang akan diberikan pada kata kunci adalah (< >).

Tagging



Aturan *Tagging*

Kata kunci yang diberikan tag (<>)		
Keputusan	Nomor :	Tentang
Menimbang	Mengingat	Memperhatikan
Menetapkan	Pertama	Kedua
Ketiga	Ditetapkan	Pada tanggal :
Memutuskan	Pada tanggal dengan 2 digit/1 digit angka yang di ikuti nama Bulan dan 4 digit angka	Nama
Tempat/Tanggal Lahir	NIP	Pendidikan
Pangkat	Jabatan	Masa Kerja
Gaji	Unit Kerja	Golongan

Kondisi Awal Sebelum *Tagging* Tahap Pertama

<><><><><>PEMERINTAH KABUPATEN GARUT<><><><><>BADAN
KEPEGAWAIAN DAN DIKLAT<><><><><>Jl. Pahlawan No. 47 Tip. (0262) 243003
Garut<><><><><><><><>PETIKAN<><><><><><><><>KEPUTUSAN BUPATI
GARUT<><><><><><><><>Nomor : 823.2/KEP.030-BKD/IV
/2014<><><><><><><><><>TENTANG<><><><><><><><>KENAIKAN
PANGKAT PEGAWAI NEGERI
SIPIL<><><><><>Menirnbang<>:<>dst;<><><><><><><><>Mengingat<>:<>dst;<><>
><><><><><><>Memperhatikan<>:<>Persetujuan teknis Kepala Kantor<>Regional III
BKN<><><><><><><><>Nomor DG-2320S000929 tanggal 27 Maret
2014<><><><><><><><>MEMUTUSKAN<><><><><>

Hasil *Tagging* Tahap Pertama

<><><><><><>PEMERINTAH KABUPATEN GARUT<><><><><>BADAN
KEPEGAWAIAN DAN DIKLAT<><><><><>Jl. Pahlawan No. 47 Tip. (0262) 243003
Garut<><><><><><><>PETIKAN<><><><><><><><> <> KEPUTUSAN BUPATI
GARUT<><><><><><><> <> Nomor : 823.2/KEP.030-BKD/IV
/2014<><><><><><><><> <> TENTANG<><><><><><><><><><>KENAIKAN
PANGKAT PEGAWAI NEGERI SIPIL<><><><><> <>
Menirnbang<>:<>dst;<><><><><><><><><> <>
Meningat<>:<>dst;<><><><><><><><><> <> Memperhatikan<>:<>Persetujuan teknis
Kepala Kantor<>Regional III BKN<><><><><><><><>Nomor DG-2320S000929 tanggal <>
27 Maret 2014<><><><><><><><> <> MEMUTUSKAN<><><><><>

Aturan *Tagging* Tahap Terakhir

Kata Kunci	Aksi
<>	Diganti delimiter
<>	Diganti dengan 1 spasi
Spasi ganda	Diganti dengan 1 spasi
Selain tanda diatas	Tidak ada aksi

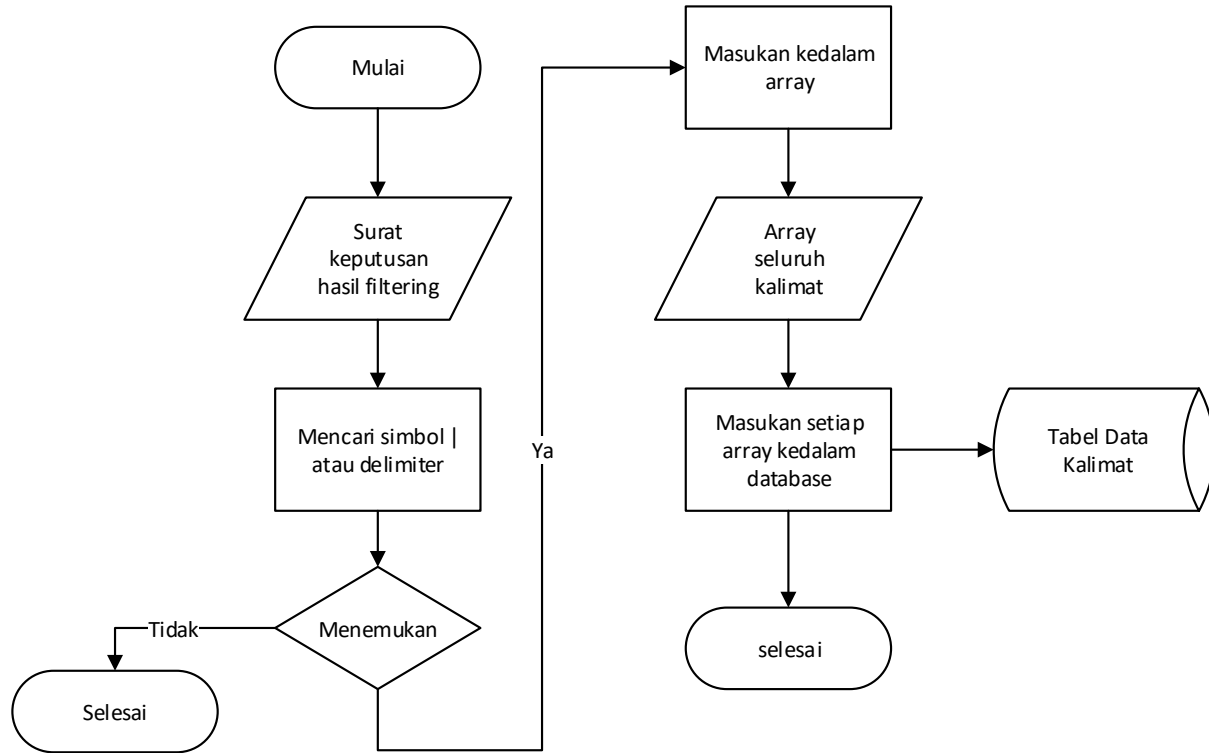
Hasil Akhir Tahap *Tagging*

|PEMERINTAH KABUPATEN GARUT BADAN KEPEGAWAIAN DAN DIKLAT JI.
Pahlawan No. 47 Tip. (0262) 243003 Garut PETIKAN |KEPUTUSAN BUPATI
GARUT |Nomor : 823.2/KEP.030-BKD/IV /2014 |TENTANG KENAIKAN PANGKAT
PEGAWAI NEGERI SIPIL |Menirnbang : dst; |Mengingat : dst; |Memperhatikan :
Persetujuan teknis Kepala Kantor Regional III BKN Nomor DG-2320S000929
tanggal |27 Maret 2014|MEMUTUSKAN

Tokenization

Tahap *tokenization* atau tokenisasi dalam penelitian ini adalah untuk membagi-bagi *string* dari data surat hasil tagging menjadi kalimat atau kata yang di bagi bagi dalam *array-arrary* yang akan dimasukkan ke dalam tabel

Tokenization



Aturan *Tokenzation*

Kondisi	Aksi
Memiliki delimiter	Pisahkan menjadi array
Selain kata diatas	Tidak ada aksi

Kondisi Sebelum *Tokenization*

|PEMERINTAH KABUPATEN GARUT BADAN KEPEGAWAIAN DAN DIKLAT Jl. Pahlawan
No. 47 Tip. (0262) 243003 Garut PETIKAN |KEPUTUSAN BUPATI GARUT |Nomor :
823.2/KEP.030-BKD/IV /2014 |TENTANG KENAIKAN PANGKAT PEGAWAI NEGERI SIPIL
|Menirnbang : dst; |Mengingat : dst; |Memperhatikan : Persetujuan teknis Kepala Kantor
Regional III BKN Nomor DG-2320S000929 tanggal |27 Maret 2014 |MEMUTUSKAN

Hasil *Tokenization*

No	Kalimat
1	PEMERINTAH KABUPATEN GARUT BADAN KEPEGAWAIAN DAN DIKLAT Jl. Pahlawan No. 47 Tip. (0262) 243003 Garut PETIKAN
2	KEPUTUSAN BUPATI GARUT
3	Nomor : 823.2/KEP.030-BKD/IV /2014
4	TENTANG KENAIKAN PANGKAT PEGAWAI NEGERI SIPIL
5	Menirnbang : dst;
6	Mengingat : dst;
7	Nomor : 823.2/KEP.030-BKD/IV /2014
8	Memperhatikan : Persetujuan teknis Kepala Kantor Regional III BKN Nomor DG-2320S000929 tanggal
9	27 Maret 2014
10	KENAIKAN PANGKAT PEGAWAI NEGERI SIPIL
11	MEMUTUSKAN



Analisis Pembobotan

Pembobotan atau ekstraksi fitur dilakukan menggunakan algoritma rule-based. Dimana menggunakan fitur lokal yang dikembangkan menjadi parameter pembobotan.

Fitur

FITUR	
X1	INITCAPS
X2	ALLCAPS
X3	CONTAINSDIGIT
X4	CONTAINSDASH
X5	CONTAINSLASH
X6	CONTAINCOLON
X7	STRING_LENGTH 1
X8	STRING_LENGTH 2
X9	STRING_LENGTH 3
X10	INITIAL_KEYWORD
X11	DATE
X12	LINE_STARTS
X13	LINE_IN
X14	LINE_END

Penjelasan Fitur

FITUR	
X1	Bobot bernilai 1 jika data uji yang dimasukan dimulai dengan kapital
X2	Bobot bernilai 1 jika data uji yang dimasukan seluruh karakter adalah huruf kapital
X3	Bobot bernilai 1 jika data uji yang dimasukan mengandung digit angka
X4	Bobot bernilai 1 jika uji yang dimasukan memiliki sedikitnya 1 karakter garis (-)
X5	Bobot bernilai 1 jika data uji yang dimasukan memiliki sedikitnya 1 karater garis miring (/)
X6	Bobot bernilai 1 jika data uji yang dimasukan memiliki sedikitnya 1 karakter titik dua (:)
X7	Bobot bernilai 1 jika panjang karakter dari data uji yang dimasukan lebih dari 50 karakter
X8	Bobot bernilai 1 jika panjang karakter dari data uji yang dimasukan lebih dari 7 karakter
X9	Bobot bernilai 1 jika panjang karakter dari data uji yang dimasukan kurang dari 10 karakter
X10	Bobot bernilai 1 jika data uji yang dimasukan memiliki sedikitnya 1 kata kunci (Tabel)
X11	Bobot bernilai 1 jika data uji yang dimasukan memiliki identitas tanggal
X12	Bobot bernilai 1 jika letak data uji berada di awal baris
X13	Bobot bernilai 1 jika letak data uji berada di tengah baris
X14	Bobot bernilai 1 jika letak data uji berada di akhir baris

Contoh Pembobotan

No	Kalimat	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14
2	KEPUTUSAN BUPATI GARUT	1	1	0	0	0	0	0	1	0	1	0	1	0	0



Pemberian Label

Pemberian label dilakukan dengan cara memasukan data label secara manual ke tabel data latih

No	KELAS	LABEL KELAS
1	1	1. Keputusan Siapa
2	2	2. Nomor surat
3	3	3. Nama
4	4	4. NIP
5	5	5. Waktu ditetapkan
6	6	6. Lain Lain

Hasil Akhir *Preprocessing* Dan Pembobotan

No	Kalimat	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	Label
1	PEMERINTAH KABUPATEN GARUT BADAN KEPEGAWAIAN DAN DIKLAT Jl. Pahlawan No. 47 Tip. (0262) 243003 Garut PETIKAN	1	0	1	0	0	0	1	1	0	0	0	1	0	0	6
2	KEPUTUSAN BUPATI GARUT	1	1	0	0	0	0	0	1	0	1	0	1	0	0	1
3	Nomor : 823.2/KEP.030- BKD/IV /2014	1	0	1	1	1	1	0	1	0	1	0	1	0	0	2
4	TENTANG KENAIKAN PANGKAT PEGAWAI NEGERI SIPIL	1	1	0	0	0	0	0	1	0	0	0	1	0	0	6



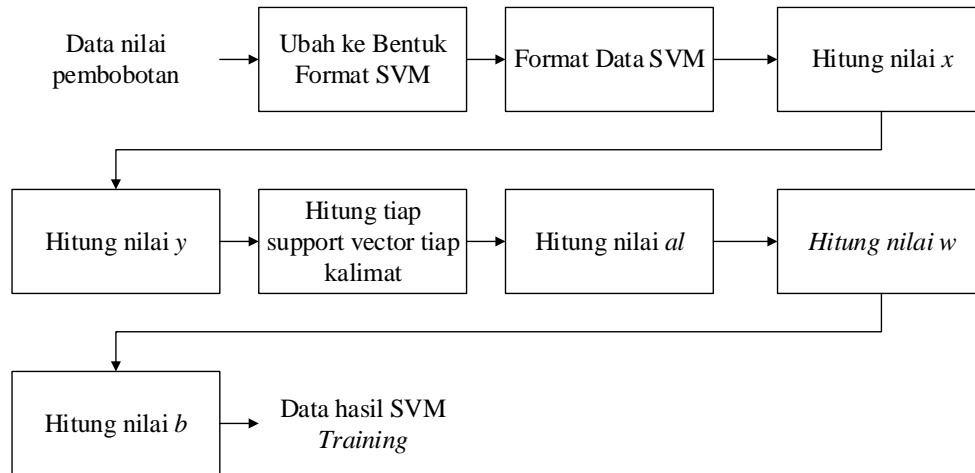
Analisis Algoritma

SVM

1. Analisis Proses Training
2. Analisis Proses Testing

Analisis Proses *Training*

Hasil dari pembobotan akan digunakan sebagai nilai fitur. Jadi setiap nilai fitur yang berjumlah 13 akan di proses menggunakan metode SVM





Mengubah Nilai Bobot kedalam Format SVM

Dalam penelitian ini untuk representasi data menggunakan format *sparse data representation* dengan vektor sebagai inputnya dengan bobot sebagai inputnya. Format data input untuk klasifikasi SVM dalam penelitian ini adalah [+1 1:1 2:0 3:1 4:0 5:0 6:0 7:1 8:1 9:0 10:0 11:0 12:1 13:0 14:0] dengan masukan pertama +1 atau -1 menyatakan dua kelas. Angka kedua menyatakan dimensi (row_id) dan angka ketiga setelah tanda ":" menyatakan nilai bobot dari kalimat tersebut

Kelas	Row_id	Bobot fitur
-1	1	1
		0
		1
		0
		0
		0
		1
		1
		0
		0
		0
		0
		1
		0
		0

Mengubah Nilai Bobot kedalam Format SVM

Data ke-1	Data ke-2	Data ke3
[+1 1:1	[-1 1:1	[+1 1:1
2:0	2:1	2:0
3:1	3:0	3:1
4:0	4:0	4:1
5:0	5:0	5:1
6:0	6:0	6:1
7:1	7:0	7:0
8:1	8:1	8:1
9:0	9:0	9:0
10:0	10:1	10:1
11:0	11:0	11:0
12:1	12:1	12:1
13:0	13:0	13:0
14:0]	14:0]	14:0]

2

Menghitung kernel

Kernel RBF



$$K(x_i, x_j) = \exp(-\gamma \|x - x_i\|^2)$$

Dilakukan kernelisasi pada set data dari fitur dimensi lama sehingga mendapatkan set data dengan fitur baru berdimensi tinggi.

$$\gamma = 0,125$$

Support Vector Machines



Rimah Amami

“Practical Selection of SVM Supervised Parameters with Different Feature Representations for Vowel Recognition”

Supriya Pahwa

“Comparison Of Various Kernels Of Support Vector Machine”



 **15,86%**
Sigmoid

 **40,10%**
Polynomial

 **84,53%**
Sigmoid

 **5,66 sec**
Sigmoid

 **44,44%**
RBF

 **76,11%**
Polynomial

 **25,64 sec**
Polynomial

 **88,20%**
Linear

 **10,35 sec**
Linear

 **80,17%**
RBF

 **4,92 sec**
RBF

Support Vector Machines



Dimass Prambudi

“IMPLEMENTASI OPTICAL CHARACTER (OCR) UNTUK PENGENALAN TEKS PADA JUDUL NOVEL MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE”

Reza Yogi

“OPTICAL CHARACTER RECOGNITION (OCR) MENGGUNAKAN SUPPORT VECTOR MACHINE (SVM) DAN ZONING PADA SERTIFIKAT”



55,62%

Quadratic Programming



83,27%

Kernel Trick



Proses Perhitungan $x_i - \bar{x}$

$x_i - \bar{x}$		Perhitungan $x_i - \bar{x}$	Hasil
Data ke-1	$x_2 - x_1$	$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$



Memasukkan Panjang Vektor Kedalam Persamaan kernel RBF

$$\begin{aligned}K(x_2, x_1) &= \exp(-\gamma \|x_2 - x_1\|^2) \\ &= \exp(-0,125(2^2)) \\ &= \exp(-0,5) = 0,606530\end{aligned}$$

$$\begin{aligned}K(x_2, x_2) &= \exp(-\gamma \|x_2 - x_2\|^2) \\ &= \exp(-0,125(1^2)) \\ &= \exp(0) = 1\end{aligned}$$

$$\begin{aligned}K(x_2, x_3) &= \exp(-\gamma \|x_2 - x_3\|^2) \\ &= \exp(-0,125(2,236067^2)) \\ &= \exp(-0,625) = 0,535261\end{aligned}$$



Hasil perhitungan persamaan kernel RBF

$K(x_1, x_1)$	$K(x_1, x_2)$	$K(x_1, x_3)$
1	0.606530	0.535261
$K(x_2, x_1)$	$K(x_2, x_2)$	$K(x_2, x_3)$
0,687289	1	0.535261
$K(x_3, x_1)$	$K(x_3, x_2)$	$K(x_3, x_3)$
0.535261	0.535261	1



$$K = \begin{bmatrix} 1 & 0.606530 & 0.535261 \\ 0.606530 & 1 & 0.535261 \\ 0.535261 & 0.535261 & 1 \end{bmatrix}$$

3 Menghitung y

Rumus menghitung y



$y_i y_j$

Nilai y

y_1	y_2	y_3
-1	1	-1



Contoh Perhitungan y

$$y_1 y_1 = (-1) (-1) = 1$$

$$y_1 y_2 = (-1) (1) = -1$$

$$y_1 y_3 = (-1) (-1) = 1$$



Hasil perhitungan y

y_1	y_2	y_3
1	-1	1
y_1	y_2	y_3
-1	1	-1
y_1	y_2	y_3
1	-1	1



$$y_i y_j = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$$



Hasil perkalian dot-product $x_i x_j$
dan $y_i y_j$

1	-0,606530	0,535261429
-0,606530	1	-0,5352614
0,5352614	-0,5352614	1

4

Menghitung Lagrange multiplayer

Maksimalkan

$$LD_{max} = a_1 + a_2 + a_3 - \frac{1}{2} (a_1^2 + a_1a_2(-1,213061319) + a_1a_31,070522857 + a_2^2 + a_2a_3(-1,070522857) + a_3^2)$$

Dengan Syarat:

1. $a_1y_1 - a_2y_2 - a_3y_3 = 0$
2. $a_1, a_2, a_3 \geq 0$



Didapatkan Nilai Alpha pada kelas ke-1

$$a_1 = 0,291549$$

$$a_2 = 0,688715$$

$$a_3 = 0,397166$$

$$b = 0,996323$$

Proses mendapatkan nilai alpha kelas lain pun memiliki cara yang sama



Hasil Perhitungan Support Vector

Kalimat	Data ke-1	Data ke-2	Data ke-3
<i>Support Vector</i>	$\begin{bmatrix} 1,363741 \\ 2,505533 \end{bmatrix}$	$\begin{bmatrix} 3,437157 \\ 6,659708 \end{bmatrix}$	$\begin{bmatrix} 1,299363 \\ 2,369887 \end{bmatrix}$



Penambahan Bias Pada Hasil Support Vector

Kalimat	Data ke-1	Data ke-2	Data ke-3
<i>Support Vector</i>	$\begin{bmatrix} 1,363741 \\ 2,505533 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 3,437157 \\ 6,659708 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1,299363 \\ 2,369887 \\ 1 \end{bmatrix}$

5

Menggunakan fungsi pemisah

Rumus



$$f(x_{testing}) = \sum_{i=1}^n a_i y_i K(x_{training}, x_{testing}) + b$$



Hasil Training SVM

Kelas	Model
6	$f(x_{testing}) = \begin{bmatrix} 0,688715 \\ 0,291549 \\ 0,397166 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \exp(-0.125 \ x_{training} - x_{testing}\ ^2) + 0,851913$
1	$f(x_{testing}) = \begin{bmatrix} 0,291549 \\ 0,688715 \\ 0,397166 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \exp(-0.125 \ x_{training} - x_{testing}\ ^2) + 0,996323$
2	$f(x_{testing}) = \begin{bmatrix} 0,347972 \\ 0,347972 \\ 0,695945 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \exp(-0.125 \ x_{training} - x_{testing}\ ^2) + 0,873477$

2 Analisis Proses *Testing*



Mengubah data ke-1 pada data uji ke format SVM

Kalimat	Nilai Fitur
KEPUTUSAN MENTERI	$x_{1\text{test}} = [-1 \ 1:1$
PENDIDIKAN NASIONAL	2:1
REPUBLIK INDONESIA	3:0
	4:0
	5:0
	6:0
	7:0
	8:1
	9:0
	10:1
	11:0
	12:1
	13:0
	14:0]

2

Perhitungan kernel data training ke-1

$$\begin{aligned}K(x_{\text{training1}}, x_{\text{testing}}) &= \exp\left(-\gamma\|x_{\text{training1}} - x_{\text{testing}}\|^2\right) \\ &= \exp(-0,125(2)) \\ &= \exp(-0,5) = 0,606530\end{aligned}$$

3

Perhitungan Testing SVM

Model	Perhitungan Uji
6(-1)	$f(x_{testing}) = \left(\begin{bmatrix} 0,688715 \\ 0,291549 \\ 0,397166 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \right) \begin{bmatrix} 0,606530 \\ 1 \\ 0,535261 \end{bmatrix} + 0,851913$ $= 0,513147$
1(1)	$f(x_{testing}) = \left(\begin{bmatrix} 0,291549 \\ 0,688715 \\ 0,397166 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \right) \begin{bmatrix} 0,606530 \\ 1 \\ 0,535261 \end{bmatrix} + 0,851913$ $= 1,295616$
2(-1)	$f(x_{testing}) = \left(\begin{bmatrix} 0,347972 \\ 0,347972 \\ 0,695945 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \right) \begin{bmatrix} 0,606530 \\ 1 \\ 0,535261 \end{bmatrix} + 0,851913$ $= 0,637880$

4

Memprediksi kelas

Persamaan
untuk prediksi



$$\max (f(x_{testing}))$$



Contoh Prediksi

$$\max(f(x_{testing})) = [(0,513147), (1,295616), (0,637880)]$$

Dari ketiga model yang digunakan nilai terbesar ada pada Kelas ke-1 sebesar 1,295616. Sehingga dapat disimpulkan bahwa kalimat yang diuji masuk ke dalam kelas ke-1.

3 Analisis Kebutuhan Non Fungsional

1. Kebutuhan Perangkat Keras
2. Kebutuhan Perangkat Lunak
3. Analisis Pengguna



Kebutuhan Perangkat Keras

- a. CPU : Intel Pentium Core 2 Duo @ 1.4 GHz
- b. RAM : 2 GB
- c. VGA : 1 GB
- d. Harddisk : 10 GB
- e. Peralatan Input : Keyboard dan Mouse
- f. Peralatan Output : Monitor 14 Inc

2

Kebutuhan Perangkat Lunak

- a. Sistem Operasi : Windows 7
- b. Sistem Manajemen Basis Data : MySQL
- c. Server Basis Data : XAMPP
- d. Pengembangan Aplikasi : Sublime 2

3

Analisis Pengguna

Pengguna untuk perangkat lunak ini adalah seseorang yang dapat mengoperasikan komputer dan mengerti alur kerja daripada aplikasi tersebut.

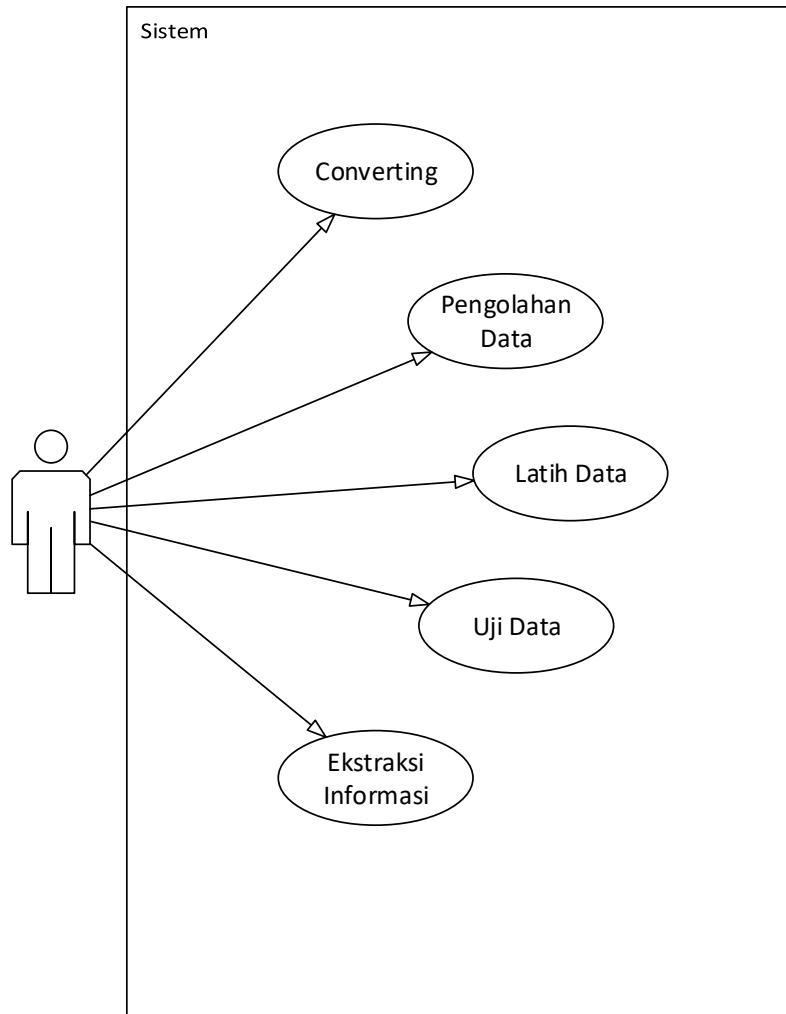


Analisis Kebutuhan Fungsional

1. Usecase Diagram



Usecase Diagram



4 Pengujian Akurasi

Support Vector Machines



Perhitungan Akurasi

Metode Confusion Matrix



Data Latih 40

4:1



Data Uji 10

Pembagian antara data latih dan data uji dilakukan secara acak untuk digunakan pada metode K-fold.





Confussion Matrix dengan K-fold percobaan ke-1

		Kelas Hasil Klasifikasi					
		1	2	3	4	5	6
Kelas Sebenarnya	1	10	0	0	0	0	0
	2	0	10	0	0	0	0
	3	0	0	10	0	0	0
	4	0	0	0	5	0	0
	5	0	0	0	0	10	0
	6	0	0	0	0	0	258

K-fold percobaan ke-1 menghasilkan 1256 data kelas pada tabel data latih dan 303 data kelas pada tabel data uji.

Akurasi : 100%
Precision : 100%
Recall : 100%
F-score : 100%

2

Confussion Matrix dengan K-fold percobaan ke-2

		Kelas Hasil Klasifikasi					
		1	2	3	4	5	6
Kelas Sebenarnya	1	10	0	0	0	0	0
	2	0	10	0	0	0	0
	3	0	0	10	0	0	0
	4	0	0	0	4	0	0
	5	0	0	0	0	10	0
	6	0	1	0	0	1	297

K-fold percobaan ke-2 menghasilkan 1216 data kelas pada tabel data latih dan 343 data kelas pada tabel data uji.

Akurasi : 99.81%
Precision : 99.89%
Recall : 96.97%
F-score : 98.36%

3

Confussion Matrix dengan K-fold percobaan ke-3

		Kelas Hasil Klasifikasi					
		1	2	3	4	5	6
Kelas Sebenarnya	1	10	0	0	0	0	0
	2	0	10	0	0	0	0
	3	0	0	8	0	0	0
	4	0	0	0	6	0	0
	5	0	0	0	0	10	0
	6	0	0	0	0	2	247

K-fold percobaan ke-3 menghasilkan 1266 data kelas pada tabel data latih dan 293 data kelas pada tabel data uji.

Akurasi : 99.77%
Precision : 99.87%
Recall : 97.22%
F-score : 98.42%

4

Confussion Matrix dengan K-fold percobaan ke-4

		Kelas Hasil Klasifikasi					
		1	2	3	4	5	6
Kelas Sebenarnya	1	10	0	0	0	0	0
	2	0	10	0	0	0	0
	3	0	0	10	0	0	0
	4	0	0	0	6	0	0
	5	0	0	0	0	10	0
	6	0	0	0	0	0	254

K-fold percobaan ke-4 menghasilkan 1259 data kelas pada tabel data latih dan 300 data kelas pada tabel data uji.

Akurasi : 100%
Precision : 100%
Recall : 100%
F-score : 100%

5

Confussion Matrix dengan K-fold percobaan ke-5

		Kelas Hasil Klasifikasi					
		1	2	3	4	5	6
Kelas Sebenarnya	1	10	0	0	0	0	0
	2	0	10	0	0	0	0
	3	0	0	10	0	0	0
	4	0	0	0	5	0	0
	5	0	0	0	0	10	0
	6	0	0	0	0	0	274

K-fold percobaan ke-5 menghasilkan 1240 data kelas pada tabel data latih dan 319 data kelas pada tabel data uji.

Akurasi : 100%
Precision : 100%
Recall : 100%
F-score : 100%

6

Hasil Perhitungan Pengujian

No	Iterasi	Precision	Recall	F-Score	Akurasi
1	1	100%	100%	100%	100%
2	2	99.89%	96.97%	98.36%	99.81%
3	3	99.87%	97.22%	98.42%	99.77%
4	4	100%	100%	100%	100%
5	5	100%	100%	100%	100%

Didapatkan nilai rata rata pada *precision*, *recall*, *f-score*, dan akurasi untuk klasifikasi metode SVM adalah bernilai 99,95%, 98,83%,99,35%. Dan 99,91%.

5 Kesimpulan dan Saran



Kesimpulan

Berdasarkan hasil pengujian sistem ekstraksi informasi menggunakan *Support Vector Machine* (SVM) pada surat keputusan maka diperoleh rata rata akurasi sebesar 99.91%. Maka dapat disimpulkan bahwa metode SVM layak digunakan pada sistem ekstraksi informasi surat keputusan.

2 Saran

Saran terhadap penelitian yang telah dilakukan adalah dibutuhkan API untuk mengkonversi file pdf yang lebih baik dalam segi merekognisi tulisan dalam surat keputusan sehingga meminimalisir kesalahan yang ada dalam konversi.

Demo Program 😊



THANKS!