

**PENERAPAN MODEL MESIN BELAJAR *SUPPORT VECTOR  
MACHINES* PADA *AUTOMATIC SCORING* UNTUK JAWABAN  
SINGKAT**

**TUGAS AKHIR**

**Diajukan Untuk Memenuhi Sebagian Prasyarat  
Mencapai Derajat Teknik Informatika**



Disusun Oleh

**Agus Aditya Haritama W**

13 07 07474

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNOLOGI INDUSTRI  
UNIVERSITAS ATMA JAYA YOGYAKARTA  
2017**

KATA PENGANTAR

**HALAMAN PENGESAHAN**

TUGAS AKHIR BERJUDUL

**Penerapan Model Mesin Belajar *Support Vector Machines*  
Pada *Automatic Scoring* Untuk Jawaban Singkat**

Disusun Oleh


Agus Aditya Haritama W

13 07 07474

dinyatakan telah lengkap memenuhi syarat pada tanggal  
Maret 2017

Dosen Pembimbing I,

Dosen Pembimbing II,

  
(Ir. A. Djoko Budiyanto SHR,  
M.Eng., PhD)

  
(Martinus Maslim, S.T.,  
M.T.)

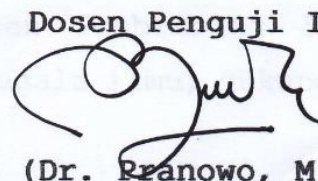
Tim Penguji  
Dosen Penguji I,

  
(Ir. A. Djoko Budiyanto SHR, M.Eng., PhD)

Dosen Penguji II,

Dosen Penguji III,

  
(Dra. Ernawati, M.T.)

  
(Dr. Pranowo, M.T.)

Yogyakarta, Maret 2017  
Universitas Atma Jaya Yogyakarta  
Fakultas Teknologi Industri  
Dekan :

  
FAKULTAS  
TEKNOLOGI INDUSTRI

  
(Dr. A. Teguh Siswanto, M.Sc.)

## KATA PENGANTAR

Puji syukur saya ucapkan kepada Ida Sang Hyang Widhi Wasa, Tuhan Yang Maha Esa. Berkat rahmat dan karunia Beliaulah penulis dapat menyelesaikan tugas akhir dengan judul "Penerapan Model Mesin Belajar *Support Vector Machines* Pada *Automatic Scoring* Untuk Jawaban Singkat". Dalam penulisan tugas akhir ini, penulis mengalami berbagai kendala dan hambatan. Namun berkat doa, semangat, dan bantuan dari keluarga, teman - teman dan berbagai pihak, tugas akhir ini dapat diselesaikan dengan baik. Pembuatan tugas akhir ini tidak lepas dari dukungan dan bantuan dari berbagai pihak. Oleh karena itu, penulis tidak lupa untuk mengucapkan banyak terimakasih kepada :

1. Ida Sang Hyang Widhi Wasa atas segala sesuatu yang telah diberikan kepada saya.
2. Bapak Dr. A. Teguh Siswanto selaku dekan Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta.
3. Bapak Martinus Maslim, S.T., M.T. selaku Kepala Program Studi Teknik Informatika Universitas Atma Jaya Yogyakarta dan Dosen Pembimbing II Tugas Akhir. Terimakasih atas segala ilmu, dukungan, dan motivasi yang diberikan.
4. Bapak Ir. A. Djoko Budiarto SHR, M.Eng., PhD selaku Dosen Pembimbing I Tugas Akhir. Terimakasih atas segala ilmu, dukungan, dan motivasi yang diberikan.
5. Mbah, Nini, Jada, I Gede Made Wiranatha, S.E., Gusti Ayu Made Widia Suharti, S.Pd.AUD dan Ayu

Riska Kristanti S selaku keluarga yang selalu memberikan doa dan motivasi.

6. Keluarga Denbantas yang sudah merawat saya sejak SMP.
7. Mathea Ayu Vanda yang telah menemani baik dalam susah maupun senang.
8. Staff dan Student Staff KSI Pak Sigit, Pak Cahyadi, Pak Irya, Mas Donny, Mas Denny, Mas Widodo, Mas Sigit, Mas Ari, Mas Johan, Mbak Lili, Mbak Rachel, Bli Agus, Kak Galih, Kak Ditto, Kak Surya, Cahya, Aji, Theo. Terimakasih atas ilmu dan pengalaman yang sudah diberikan.
9. HIMAFORKA angkatan 2013/2014 atas segala ilmu dalam berorganisasi.
10. HIMAFORKA angkatan 2014/2015 atas segala perjuangan dan kebersamaannya.

Penulisan tugas akhir ini masih jauh dari kata sempurna. Penulis memohon maaf jika ada kekurangan dalam penyajian tulisan ataupun teknik eksperimen yang digunakan. Untuk itu penulis akan sangat berterima kasih jika ada kritik maupun saran yang dapat menjadikan laporan ini lebih baik lagi.

Yogyakarta, 6 Maret 2017

Penulis

## DAFTAR ISI

<b>HALAMAN PENGESAHAN</b> .....	ii
<b>KATA PENGANTAR</b> .....	iii
<b>DAFTAR ISI</b> .....	v
<b>DAFTAR TABEL</b> .....	vii
<b>DAFTAR GAMBAR</b> .....	ix
<b>INTISARI</b> .....	xii
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah .....	3
1.4 Tujuan Penelitian .....	4
1.5 Sistematika Penulisan .....	4
<b>BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI</b> .....	6
2.1 Tinjauan Pustaka .....	6
2.2 Landasan Teori .....	13
2.2.1 <i>Automatic Scoring</i> untuk Jawaban Singkat ...	13
2.2.2 Support Vector Machines (SVM) .....	14
2.2.3 Kernel Pada SVM .....	18
2.2.4 <i>Soft Margin Classifier</i> .....	20
2.2.5 <i>Multi Class SVM</i> .....	21
2.2.6 WEKA (Waikato Environment of Knowledge Analysis ) .....	23
2.2.7 K-fold Cross Validation .....	24

2.2.10 Natural Language Processing.....	24
2.2.11 Jaccard Similarity.....	25
<b>BAB III METODOLOGI PENELITIAN.....</b>	<b>27</b>
3.1 Pembentukan <i>Dataset</i> .....	27
3.2 Pembentukan Model .....	34
3.3 Evaluasi Model .....	36
3.4 Perbandingan Penggunaan SVM dengan Metode Klasifikasi Lain .....	40
<b>BAB IV HASIL DAN PEMBAHASAN.....</b>	<b>41</b>
4.1 Karakteristik <i>Dataset</i> yang Terbentuk .....	41
4.2 Menentukan Nilai Parameter Cost dan Gamma .....	47
4.3 Pemilihan Fitur .....	51
4.4 Evaluasi Model dengan WEKA .....	55
4.5 Perbandingan Penggunaan SVM dengan Metode Klasifikasi Lain .....	67
<b>BAB V KESIMPULAN DAN SARAN.....</b>	<b>69</b>
5.1 Kesimpulan .....	69
5.2 Saran .....	70
<b>DAFTAR PUSTAKA.....</b>	<b>71</b>
<b>LAMPIRAN .....</b>	<b>75</b>

## DAFTAR TABEL

Tabel 2.1 : Kernel Pada SVM.....	20
Tabel 4.1 : Nilai rata-rata setiap fitur pada <i>dataset</i> .....	41
Tabel 4.3 : Akurasi dan waktu yang diberikan oleh nilai variabel C dengan rentangan nilai C $2^{-15} - 2^5$ ..	47
Tabel 4.3 : Hasil klasifikasi yang diberikan oleh metode SVM linier dengan rentang nilai C $2^{-5} - 2^{11}$ ...	49
Tabel 4.4 : Akurasi dan waktu yang diberikan oleh nilai variabel $\gamma$ dengan rentangan nilai $2^{-25} - 2^{13}$ ....	50
Tabel 4.5 : Akurasi yang diberikan metode SVM dengan kernel RBF dengan nilai C = $2^5$ dan rentang nilai $\gamma$ $2^{-1}$ - $2^3$ .....	51
Tabel 4.6 : Nilai Akurasi untuk masing - masing fitur pada <i>dataset</i> .....	52
Tabel 4.7 : Nilai detail akurasi untuk kombinasi fitur pada <i>dataset</i> .....	54
Tabel 4.8 : <i>Confusion matrix</i> yang terbentuk untuk metode SVM dengan kernel linier.....	58
Tabel 4.9 : <i>Confusion matrix</i> yang terbentuk untuk metode SVM dengan kernel RBF.....	59
Tabel 4.10 : Hasil detail akurasi yang diberikan oleh metode SVM linier.....	60
Tabel 4.11 : Hasil detail akurasi yang diberikan oleh metode SVM kernel RBF .....	61

Tabel 4.12 : Perbandingan penggunaan metode  
klasifikasi SVM dengan metode klasifikasi regresi  
logistik dan C4.5..... 67





## DAFTAR GAMBAR

Gambar 2.1 : Visualisasi SVM.....	15
Gambar 2.2 : Visualisasi <i>Soft Margin Classifier</i> .....	20
Gambar 2.3 : Titik data dengan tiga kelas.....	22
Gambar 2.4 : Visualisasi metode <i>one-against-all</i> .....	22
Gambar 2.5 : Ilustrasi algoritma jaccard.....	26
Gambar 3.1 : Metodologi penelitian .....	27
Gambar 3.2 : Visualisasi proses <i>Stemming</i> .....	29
Gambar 3.3 : Tampilan pada situs <a href="http://www.ranks.nl">www.ranks.nl</a> .....	30
Gambar 3.4 : Ilustrasi token pada dokumen.....	30
Gambar 3.5 : Visualisasi proses normalisasi.....	32
Gambar 3.6 : Penggunaan teknik kernel untuk merubah <i>input space</i> menjadi <i>feature space</i> .....	34
Gambar 3.7 : Visualisasi <i>Confusion Matrix</i> .....	37
Gambar 3.8 : Kurva <i>Receiver Operating Characteristic</i> ( <i>ROC</i> ) .....	39
Gambar 4.1 : Grafik sebaran nilai setiap fitur pada kelas A.....	42
Gambar 4.2 : Grafik sebaran nilai setiap fitur pada kelas B.....	43
Gambar 4.3 : Grafik sebaran nilai setiap fitur pada kelas C.....	44
Gambar 4.4 : Grafik sebaran nilai setiap fitur pada kelas D .....	45

Gambar 4.5 : Grafik sebaran nilai setiap fitur pada kelas E .....	46
Gambar 4.6 : Grafik Akurasi dan Waktu yang diberikan metode SVM linier dengan rentang nilai C $2^{-15} - 2^{25}$ ...	47
Gambar 4.7 : Grafik akurasi dan waktu yang diberikan oleh nilai variabel $\gamma$ dengan rentangan nilai $2^{-25} - 2^{13}$ .....	50
Gambar 4.8 : Pengaturan awal dari proses analisa dengan menggunakan metode SVM dengan kernel linier .	56
Gambar 4.9 : Pengaturan awal dari proses analisa dengan menggunakan metode SVM dengan kernel RBF .....	57
Gambar 4.10 : Kurva ROC kelas A dengan metode SVM linier .....	63
Gambar 4.11 : Kurva ROC kelas B dengan metode SVM linier .....	63
Gambar 4.12 : Kurva ROC kelas C dengan metode SVM linier .....	64
Gambar 4.13 : Kurva ROC kelas D dengan metode SVM linier .....	64
Gambar 4.14 : Kurva ROC kelas E dengan metode SVM linier .....	64
Gambar 4.15 : Kurva ROC kelas A dengan metode SVM kernel RBF .....	65
Gambar 4.16 : Kurva ROC kelas B dengan metode SVM kernel RBF .....	65

Gambar 4.17 : Kurva ROC kelas C dengan metode SVM  
kernel RBF ..... 65

Gambar 4.18 : Kurva ROC kelas D dengan metode SVM  
kernel RBF ..... 66

Gambar 4.19 : Kurva ROC kelas E dengan metode SVM  
kernel RBF ..... 66



Penerapan Model Mesin Belajar *Support Vector Machines*  
Pada *Automatic Scoring* Untuk Jawaban Singkat

INTISARI

Agus Aditya Haritama W (13 07 07474)

Sistem *automatic scoring* untuk jawaban singkat (ASJS) merupakan sistem yang melakukan proses penilaian jawaban dengan bantuan metode komputasi. Jawaban yang diberikan biasanya menggunakan bahasa natural. ASJS dibangun dengan harapan dapat mengoptimalkan proses pemeriksaan soal dengan tipe jawaban singkat. Permasalahan umum yang dihadapi dalam sistem adalah penggunaan bahasa natural yang membutuhkan proses *preprocessing* agar bisa dikomparasi. Permasalahan kedua adalah tidak tersedianya *dataset* jawaban dan kunci jawaban untuk bahasa Indonesia.

Dalam penelitian ini penulis akan membangun *dataset* baru dan menerapkan model klasifikasi *Support Vector Machines (SVM)* pada sistem. Input dari metode SVM adalah hasil dari proses *preprocessing* pada bahasa natural. Proses ini meliputi *stemming*, penghapusan *stopwords*, penghitungan selisih jumlah kata, penghitungan nilai kesamaan kata, penghitungan nilai kesamaan struktur kalimat dan penghitungan nilai kesamaan kalimat dengan algoritma jaccard.

Hasil dari eksperimen adalah ditentukannya empat fitur penting yang ada dalam *dataset*. Fitur tersebut adalah selisih jumlah kata, nilai kesamaan jenis kata, nilai kesamaan kalimat dan nilai kesamaan jenis kalimat. Nilai akurasi yang diberikan oleh metode SVM kernel linier dan kernel RBF sebesar 89% dengan nilai *mean absolute error* sebesar 0,044. Nilai ini lebih baik jika dibandingkan dengan metode Regresi Logistik dan C4.5.

Kata kunci : Klasifikasi, *Support Vector Machine (SVM)*,  
*Automatic Scoring*

Dosen Pembimbing 1: Ir. A. Djoko Budiyanto, M.Eng., PhD

Dosen Pembimbing 2: Martinus Maslim, S.T., M.T.