

**Pembangunan Model Klasifikasi Sentimen Teks Bahasa Indonesia
Menggunakan *Library Spacy***

TUGAS AKHIR



Disusun Oleh:

Jegar Sahaduta Ginting

13 07 07567

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI INDUSTRI

UNIVERSITAS ATMA JAYA YOGYAKARTA

2018

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

Pembangunan Model Klasifikasi Sentimen Teks Bahasa Indonesia Menggunakan *Library Spacy*

Disusun Oleh :

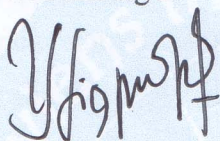
Jegar Sahaduta Ginting (13 07 07567)

Dinyatakan Telah Memenuhi Syarat

Pada Tanggal : Oktober 2018

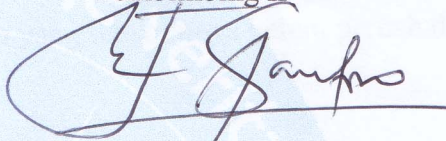
Oleh :

Pembimbing I



(Y. Sigit Purnomo W.P., S.T., M.Kom.)

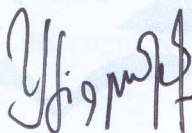
Pembimbing II



(Dr. Ir. Alb. Joko Santoso, M.T.)

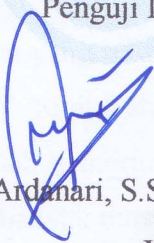
Tim Penguji :

Penguji I



(Y. Sigit Purnomo W.P., S.T., M.Kom.)

Penguji II



(Patricia Ardanari, S.Si, M.T.)

Penguji III



(Thomas Adi Purnomo Sidhi, S.T., M.T.)

Yogyakarta, Oktober 2018

Universitas Atma Jaya Yogyakarta

Fakultas Teknologi Industri

Dekan,



Dr. A. Teguh Siswantoro

KATA PENGANTAR

Penulis mengucapkan puji dan syukur kepada Tuhan Yang Maha Esa atas segala kasih dan karunia-Nya sehingga penulis dapat menyelesaikan pembuatan laporan tugas akhir ini dengan baik. Tujuan dari pembuatan laporan tugas akhir ini adalah sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta.

Penulisan tugas akhir ini, penulis sebagai mahasiswa Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta mendapat banyak pengalaman dan ilmu-ilmu baru yang sebelumnya belum pernah di dapat di kelas perkuliahan.

Dalam pengerjaan Tugas Akhir penulis banyak mendapat bantuan dari berbagai pihak. Untuk itu pada kesempatan ini, penulis ingin menyampaikan ungkapan terimakasih kepada :

1. Allah Bapa di Surga yang selalu memberikan kekuatan, kesehatan dan berkat kasih karunianya kepada penulis.
2. Keluarga tercinta yaitu bapak, ibu, kakak, adik yang selalu memberikan dukungan dalam bentuk moral dan doa. Begitu juga dengan nasihat dan semangat yang selalu diberikan kepada penulis.
3. Bapak Y. Sigit Purnomo W.P., S.T., M.Kom., selaku dosen pembimbing I yang telah menerima penulis sebagai mahasiswa bimbingan Tugas Akhir. Begitu banyak masukan dan petunjuk yang telah diberikan sehingga Tugas Akhir ini dapat selesai dengan baik.
4. Bapak Dr. Ir. Alb. Joko Santoso, M.T., selaku dosen pembimbing II yang selalu membimbing dan memberi masukan dalam penulisan tugas akhir ini.
5. Seluruh dosen dan staff Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Atma Jaya Yogyakarta yang telah membantu penulis selama kuliah.

6. Untuk Satuan Resimen Mahasiswa Universitas Atma Jaya yang selama perkuliahan menjadi tempat membentuk karakter dan belajar berorganisasi bersama rekan-rekan Senior dan Junior.
7. Untuk teman-teman Maho Class D yang selalu mendukung penulis dan menjadi teman-teman seperjuangan semasa kuliah.
8. Teman-teman seangkatan TF 2013 yang tidak bisa disebutkan satu persatu
9. Seluruh pihak yang tidak dapat disebutkan satu-persatu, yang telah mendukung dan memberikan masukan kepada penulis selama pengerjaan Tugas Akhir ini.

Penulis menyadari bahwa laporan Tugas Akhir ini masih jauh dari sempurna. Oleh sebab itu, segala kritik dan saran yang membangun sangatlah penulis harapkan. Akhir kata semoga laporan Tugas Akhir ini dapat berguna dan bermanfaat bagi semua pihak.

Yogyakarta 2018

Penulis,

Jegar Sahaduta Ginting

DAFTAR ISI

HALAMAN PENGESAHAN.....	ii
KATA PENGANTAR	iii
DAFTAR ISI.....	v
DAFTAR GAMBAR	vii
DAFTAR TABEL	viii
INTISARI.....	ix
BAB I	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian	3
1.5. Sistematika Penulisan	3
BAB II.....	5
2.1. Penelitian Terdahulu.....	5
2.2. Landasan Teori.....	8
2.3.1. Text Classification	8
2.3.2. Python.....	8
2.3.3. Natural Language Processing.....	8
2.3.4. Arsitektur Spacy	9
BAB III	13
3.1. Bahan Penelitian.....	13
3.2. Alat Penelitian.....	13
3.3.1. Perangkat Lunak	13
3.3.2. Perangkat Keras	14
3.3. Langkah-langkah Penelitian	14
3.3.1. Studi Pustaka	16
3.3.2. Pengumpulan Bahan	16
3.3.3. Pembuatan Model	16

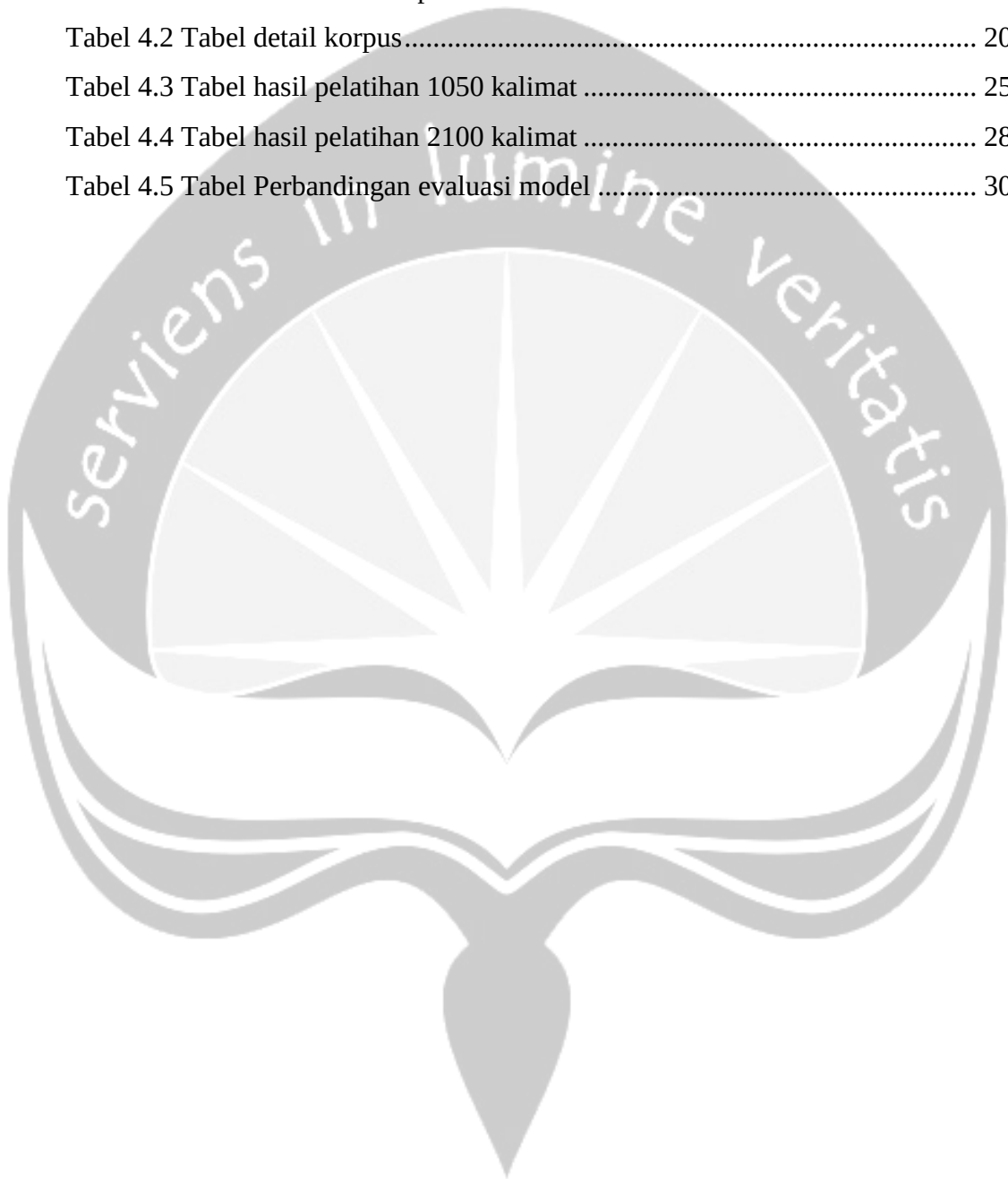
3.3.3.1. <i>Training Model Vektor</i>	16
3.3.3.2. <i>Training Model Text Classification</i>	16
3.3.4. <i>Evaluasi Model</i>	17
3.3.5. <i>Pembuatan Web Services</i>	17
BAB IV	18
4.1. <i>Training Word Vector</i>	19
4.2. <i>Training Dataset Text Classification</i>	21
4.3. <i>Penyatuan Model Text Classification dan Vector Model</i>	29
4.4. <i>Evaluasi Model</i>	30
BAB V.....	32
5.1. <i>Web Service</i>	32
BAB VI	35
6.1. <i>Kesimpulan</i>	35
6.2. <i>Saran</i>	35
DAFTAR PUSTAKA	36
LAMPIRAN.....	38

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur Spacy (https://spacy.io/api/ , 1 Juli 2018).....	10
Gambar 3.2 Langkah- Langkah Penelitian.....	15
Gambar 4.1 Hasil Pengolahan Vector.....	20
Gambar 4.2 Hasil Package Pengolahan Vector.....	21
Gambar 4.3 Pelatihan 25 iterasi 1050 kalimat	23
Gambar 4.4 Pelatihan 20 iterasi 1050 kalimat	24
Gambar 4.5 Pelatihan 15 iterasi 1050 kalimat	24
Gambar 4.6 Pelatihan 10 iterasi 1050 kalimat	25
Gambar 4.7 Pelatihan 25 iterasi 2100 kalimat	26
Gambar 4.8 Pelatihan 20 iterasi 2100 kalimat	27
Gambar 4.9 Pelatihan 15 iterasi 2100 kalimat	27
Gambar 4.10 Pelatihan 10 iterasi 2100 kalimat	28
Gambar 4.7 Hasil Package Pengolahan Model.....	29
Gambar 4.8 Hasil Penyatuan model vektor dan model texcat	30
Gambar 5.1 Kode penggunaan model pada Web Service.....	33

DAFTAR TABEL

Tabel 4.1 Tabel detail kalimat pelatihan.....	19
Tabel 4.2 Tabel detail korpus.....	20
Tabel 4.3 Tabel hasil pelatihan 1050 kalimat	25
Tabel 4.4 Tabel hasil pelatihan 2100 kalimat	28
Tabel 4.5 Tabel Perbandingan evaluasi model	30



INTISARI

Jegar Sahaduta Ginting (13 07 07567)

Teknik Informatika Universitas Atma Jaya Yogyakarta

Jl. Babarsari No. 43, Yogyakarta, Indonesia

Natural Language Processing merupakan salah satu cabang ilmu *Artificial Intelligence* yang berfokus pada pengolahan bahasa natural. *Natural Language Processing* dalam dunia teknologi informasi semakin meningkat dan berkembang. Umumnya pengembangan *NLP* terjadi secara pesat dalam pemrosesan Bahasa Inggris. Sedangkan untuk pengembangan *NLP* dalam Bahasa Indonesia masih terlalu sedikit. Salah satunya adalah tentang klasifikasi teks untuk menemukan sentimen kalimat.

Penelitian ini akan membuat sebuah model *NLP* untuk mengklasifikasi kalimat Bahasa Indonesia yang mengandung sentimen positif, negatif dan netral yang menggunakan *library Spacy*. Data yang digunakan terdiri dari korpus Universitas Leipzig yang berada di Jerman dan data kalimat Bahasa Indonesia yang diambil dari *movie reviews*, ulasan aplikasi, komentar netizen di Youtube, Facebook, Instagram dan Twitter sejumlah 2550 data kalimat. Data tersebut kemudian dibagi menjadi 2100 data pelatihan dan 450 data tes.

Dari Model yang dirancang kemudian dilakukan evaluasi dan ditemukan akurasi sebesar 83,80%. Sehingga dapat disimpulkan bahwa model yang dibuat sudah cukup baik. Model ini kemudian diterapkan pada *web service* menggunakan *framework Flask* supaya dapat digunakan di berbagai platform.

Kata Kunci: *Natural Language Processing*, *Spacy*, sentimen kalimat, Bahasa Indonesia.