HALAMAN PERSETUJUAN PEMBIMBING

TUGAS AKHIR

Judul : INISIASI *NLP (NATURAL LANGUAGE PROCESSING)* DALAM IMPLEMENTASI *CHATBOT* BERDASARKAN SISTEM PERTANYAAN DAN JAWABAN INTERAKTIF (Studi Kasus: *FAQ* Operator Seluler Telkomsel)

Nama : Muhamad Hary Reynaldi

Nomor Mahasiswa : 14 611 235

TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK DIUJIKAN

Yogyakarta, 28 Oktober 2018

Pembimbing

(Ayundyah Kesumawati, M.Si)

HALAMAN PENGESAHAN

TUGAS AKHIR

INISIASI *NLP (NATURAL LANGUAGE PROCESSING)* DALAM IMPLEMENTASI *CHATBOT* BERDASARKAN SISTEM PERTANYAAN DAN JAWABAN INTERAKTIF

(Studi Kasus : FAQ Operator Seluler Telkomsel)

Nama Mahasiswa : Muhamad Hary Reynaldi

Nomor Mahasiswa : 14 611 235

TUGAS AKHIR INI TELAH DIUJIKAN

PADA TANGGAL
26 NOVEMBER 2018

Nama Penguji :	**Tanda Tangan**
1. Ir. Sukirman, M.M	.....
2. Dr. Jaka Nugraha, M.Si	.....
3. Ayundyah Kesumawati, M.Si	.....

Mengetahui,
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Prof. Riyanto, S.Pd.,M.Si., Ph.D

KATA PENGANTAR



Assalamualaikum Wr.Wb

Puji syukur penulis panjatkan atas kehadirat Allah SWT karena atas berkat, rahmat, kesehatan dan kekuatan yang diberikan oleh- Nya tugas akhir ini dapat berjalan dengan lancar. Tugas akhir ini memberikan begitu banyak pembelajaran yang kemudian dapat di kembangkan pada penelitian selanjutnya.

Keberhasilan pembuatan tugas akhir ini tentunya tidak terlepas dari pihak-pihak yang membantu, memberi semangat dan turut memberi dukungannya selama kegiatan ini berlangsung. Ucapan terima kasih ini disampaikan kepada :

1. Bapak Prof. Riyanto, S.pd., M.Si., Ph.D., selaku dekan Fakultas Matematika dan Ilmu Pengetahuan Alam.
2. Bapak Dr. Edy Widodo, M.Si selaku Ketua Jurusan Statistika beserta jajarannya.
3. Ibu Ayundyah Kesumawati M.Si selaku Dosen Pembimbing penelitian yang selalu sabar membimbing dan mengarahkan penulis.
4. Bapak RB Fajriya Hakim, M.Si selaku Dosen Pembimbing Akademik penulis yang sangat penulis kagumi dan memberikan ilmu yang sangat bermanfaat untuk penulis.
5. Orang tua penulis, Ibu Rohayati dan Bapak Dwinanto Daru Priyono yang telah memberi dukungan dan motivasi penulis dalam menulis penelitian ini.
6. Bapak Suhadi selaku ayah kandung penulis yang memberikan semangat untuk penulis.
7. Keluarga penulis Agesty Lidya Pangestika, Pritha Erika Fiharsih, Oky Subandi, serta keponakan tercinta penulis Zayden Alfarezel Subandi dan Zania Aranaomi Subandi.
8. Sahabat tercinta Frizia Amanda Josifasyah, Dalliasari Rahmani, Imam Maliki Akbar, dan Gemilang Kuswandi Putra.

9. Teman seperjuangan *Gang Godel*, Davien Alcarlie Yudistira Antoni, Faiz Dzulfaqor, Gita Sandy Ariadne, Syafiq, Aji Wira Buana, dan Muhammad Farhan Abdul Fattah.
10. Teman sekaligus rekan bekerja dalam band *Meneer Roem*, Fachri Dwi Kurniawan dan Jotro Perta Ginting.
11. Lutfi Apriza, Fauzella Sahara, Ade Muhammad Irfan dan Egy Zulwandi selaku teman terdekat penulis untuk saat ini.
12. Keluarga Besar XIX, Statistika UII angkatan 2014.
13. Seluruh pihak yang mendoakan penulis dalam diam yang tidak bisa disebutkan satu-persatu.

Semoga segala bentuk kegiatan yang dilakukan selama tugas akhir ini dapat bermanfaat. Penulis menyadari sepenuhnya bahwa tugas akhir ini masih jauh dari sempurna, oleh karena itu segala kritik dan saran yang sifatnya membangun selalu penulis harapkan. Semoga tugas akhir ini dapat bermanfaat bagi penulis khususnya dan bagi semua yang membutuhkan umumnya. Akhir kata, semoga Allah SWT selalu melimpahkan rahmat serta hidayah-Nya kepada kita semua, Amin amin ya robbal ‘alamiin

Wassalamualaikum Wr.Wb

Yogyakarta, 21 Oktober 2018

Penulis



Muhamad Hary Reynaldi

DAFTAR ISI

HALAMAN PERSETUJUAN PEMBIMBING ii
HALAMAN PENGESAHAN iii
KATA PENGANTAR iv
DAFTAR ISI vi
DAFTAR TABEL ix
DAFTAR GAMBAR x
DAFTAR LAMPIRAN xi
PERNYATAAN xii
INTISARI xiii
ABSTRACT xiv
BAB I 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 3
1.3 Batasan Masalah 4
1.4 Tujuan Penelitian 4
1.5 Manfaat Penelitian 4
BAB II 5
BAB III 8
3.1 Telkomsel 8
3.1.1 Sejarah Telkomsel 8
3.1.2 Kategori Isi Ulang Pulsa 8
3.1.3 Kategori Tidak Terhubung Jaringan *Internet* 9
3.1.4 Kategori Tidak Bisa *Browsing* 9
3.1.5 Kategori Internet Lambat 9
3.1.6 Kategori Tidak Ada Layanan 9
3.1.7 Kategori Nomor Tidak Bisa Digunakan 9
3.1.8 Kategori *Upgrade* Kartu SIM *4G* 9
3.1.9 Kategori Informasi Profil dan Nomor *PUK* 10
3.1.10 Kategori Tagihan *Online* KartuHALO 10

3.1.11 Kategori Tagihan dan Pemakaian Pascabayar 10
3.2 *Frequently Asked Question* 10
3.3 Analisis Deskriptif 10
3.4 *Machine Learning* 11
3.5 *Natural Language Processing (NLP)* 12
3.6 *Chatbot* 13
3.7 *Natural Language Tool Kit (NLTK)* 14
3.8 Pembobotan Kata *TF-IDF* 14
3.9 Data *Training* dan Data *Testing* 16
3.10 *Support Vector Machine* 16
3.11 Klasifikasi 18
3.12 *Preprocessing Data* 20
BAB IV 24
4.1 Populasi dan Sampel 24
4.2 Sumber Data 24
4.3 Variabel Penelitian 24
4.4 Metode Pengambilan Data 23
4.5 Metode Analisis Data 23
4.6 Tahapan Penelitian 23
BAB V 27
5.1 Analisis Deskriptif 27
5.2 Klasifikasi 29
5.2.1 *Preprocessing Data* 29
5.2.2 *Lower Casing* 29
5.2.3 *Remove Punctuation* 30
5.2.4 *Stopwords Removal* 31
5.2.5 *Tokenizing* 31
5.3 *TF-IDF* 32
5.4 *Randomize Dataset* 33
5.5 Pembuatan Data Latih dan Data Uji 35
5.6 Klasifikasi *Support Vector Machine (SVM)* 36

5.7 Memprediksi kategori untuk data pertanyaan baru 38
5.8 *Chatbot* 39
BAB VI 41
6.1 Kesimpulan 41
6.2 Saran 41
DAFTAR PUSTAKA 42
LAMPIRAN 47

DAFTAR TABEL

Tabel 4.1 Variabel Penelitian 22

Tabel 5.1 Input Pertanyaan ke dalam Format *Microsoft Excel* 27

Tabel 5.2 *Lower Casing* 30

Tabel 5.3 *Remove Punctuation* 30

Tabel 5.4 *Stopwords Removal* 31

Tabel 5.5 *Tokenizing* 32

Tabel 5.6 Sampel Pembobotan Kata *TF-IDF* 33

Tabel 5.7 Sampel Hasil *Randomize Dataset* 34

Tabel 5.8 Data Latih dan Data Uji 35

Tabel 5.9 Perancangan Respon dari *Chatbot* 39

DAFTAR GAMBAR

Gambar 3.1 Optimal *Hyperplane SVM*..... 17

Gambar 3.2 Ilustrasi *Problem Non-Linier* menjadi *Linier* dengan *Kernel SVM*..... 18

Gambar 3.3 Fase *Learning* 19

Gambar 3.4 Fase *Testing* 19

Gambar 4.1 Tahapan Penelitian..... 23

Gambar 5.1 Kategori dalam Halaman *FAQ Website* Telkomsel..... 25

Gambar 5.2 Variabel Pertanyaan dan Jawaban Kategori Isi Ulang Pulsa..... 26

Gambar 5.3 Banyaknya Jumlah Pertanyaan Berdasarkan Kategori..... 27

Gambar 5.4 Banyaknya Jumlah Inisiasi Pertanyaan Berdasarkan Kategori 28

Gambar 5.5 *Confussion Matrix Support Vector Machine* Data Uji 37

Gambar 5.6 Hasil Prediksi Kategori Pertanyaan dari Pertanyaan Baru 38

Gambar 5.7 Implementasi Percakapan Aplikasi *Chatbot*..... 40

DAFTAR LAMPIRAN

Lampiran 1 *Dataset FAQ* Operator Seluler Telkomsel 46

Lampiran 2 Ulasan *Syntax* Mesin Klasifikasi 67

Lampiran 3 Sintaks Pembuatan Respon Jawaban *Chatbot* 73

PERNYATAAN

Dengan ini menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan disuatu Perguruan Tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan diterbitkan dalam daftar pustaka.

Yogyakarta, 26 Oktober 2018



Muhamad Hary Reynaldi