

## DAFTAR ISI

<b>PERNYATAAN.....</b>	<b>I</b>
<b>KATA PENGANTAR.....</b>	<b>II</b>
<b>UCAPAN TERIMA KASIH .....</b>	<b>IV</b>
<b>ABSTRAK .....</b>	<b>VI</b>
<b>ABSTRACT .....</b>	<b>VII</b>
<b>DAFTAR ISI.....</b>	<b>VIII</b>
<b>DAFTAR TABEL .....</b>	<b>XI</b>
<b>DAFTAR GAMBAR.....</b>	<b>XII</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 LATAR BELAKANG.....	1
1.2 PERUMUSAN MASALAH .....	5
1.3 BATASAN MASALAH.....	5
1.4 TUJUAN PENELITIAN .....	5
1.5 MANFAAT PENELITIAN .....	6
1.6 SISTEMATIKA PENULISAN .....	6
<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>8</b>
2.1 DEMOGRAFI .....	8
2.1.1 Demografi dan Pasar.....	10

2.1.2	Segmentasi Pasar ( <i>Market Segmentation</i> ).....	11
2.1.3	Keuntungan mensegmentasikan pelanggan .....	12
2.2	TWITTER .....	12
2.2.1	Statistika Pengguna Twitter .....	14
2.2.2	Twitter API ( <i>Application Programming Interface</i> ).....	15
2.3	KLASIFIKASI ( <i>CLASSIFICATION</i> ) .....	16
2.3.1	Praproses ( <i>Preprocessing</i> ) .....	18
2.4	METODE POHON KEPUTUSAN ( <i>DECISION TREE</i> ) DAN ALGORITMA C4.5 .....	19
2.4.1	Pembentukan Pohon Keputusan ( <i>Decision Tree</i> ) .....	20
2.4.2	Perhitungan <i>Entropy</i> dan <i>Information Gain</i> .....	21
2.4.3	Percabangan Atribut Dengan Nilai Kontinyu .....	24
2.4.4	<i>Decision Rules</i> .....	25
2.4.5	Perhitungan Akurasi ( <i>Accuracy</i> ) dan Kesalahan ( <i>Error</i> ).....	26
2.4.6	Evaluasi Pohon Keputusan Dengan Menggunakan <i>k-folds Cross Validation</i> .....	27
2.4.7	Menyederhanakan Pohon Keputusan ( <i>Pruning</i> ).....	28
<b>BAB III METODE PENELITIAN .....</b>		<b>32</b>
3.1	DESAIN PENELITIAN.....	32
3.2	METODE PENELITIAN .....	34
3.2.1	Metode Pengumpulan Data.....	34
3.2.2	Metode Pengembangan Perangkat Lunak.....	35
3.3	ALAT DAN BAHAN PENELITIAN.....	37
3.3.1	Alat Penelitian.....	37

3.3.2	Bahan Penelitian .....	37
<b>BAB IV HASIL PENELITIAN DAN PEMBAHASAN .....</b>		<b>39</b>
4.1	EKSPERIMEN .....	39
4.1.1	Membuat daftar kata .....	39
4.1.2	Menghitung frekuensi penggunaan kata .....	41
4.1.3	Menentukan kata kunci dan menghitung akurasi kata kunci .....	41
4.2	IMPLEMENTASI.....	44
4.2.1	Analisa Kebutuhan Perangkat Lunak.....	44
4.2.2	Desain Perangkat Lunak .....	45
4.2.3	Pengujian Perangkat Lunak Secara <i>Blackbox</i> .....	47
4.3	ANALISA HASIL PEMBENTUKAN MODEL KLASIFIKASI <i>DECISION TREE C4.5</i>	
	49	
4.3.1	Simulasi Pembentukan Model Pohon Klasifikasi .....	49
4.3.2	Hasil Model Pohon Klasifikasi Demografi .....	73
4.4	PENGUJIAN MODEL KLASIFIKASI .....	76
4.5	STUDI KASUS HASIL DEMOGRAFI PELANGGAN .....	77
<b>BAB V KESIMPULAN DAN SARAN.....</b>		<b>81</b>
5.1	KESIMPULAN.....	81
5.2	SARAN .....	82
<b>DAFTAR PUSTAKA.....</b>		<b>84</b>

## DAFTAR TABEL

TABEL 4.1 JUMLAH DAFTAR KATA YANG TERSIMPAN DALAM DATABASE .....	40
TABEL 4.2 POLA KATA KUNCI UNTUK MENGLASIFIKASIKAN GENDER .....	43
TABEL 4.3 POLA KATA KUNCI UNTUK MENGLASIFIKASIKAN USIA.....	43
TABEL 4.4 POLA KATA KUNCI UNTUK MENGLASIFIKASIKAN STATUS PEKERJAAN .	44
TABEL 4.5 HASIL PENGUJIAN MODEL KLASIFIKASI 10 DATA TEST .....	76

## DAFTAR GAMBAR

GAMBAR 1.1 PERKEMBANGAN UMKM PERIODE 1997 - 2012 (SUMBER: HTTP://WWW.BPS.GO.ID/) .....	1
GAMBAR 1.2 CONTOH DIAGRAM POHON SEDERHANA UNTUK MENGLASIFIKAN MAMALIA .....	4
GAMBAR 2.1 DEMOGRAFI PENDUDUK INDONESIA TAHUN 2012/2013 (SUMBER : HTTPS://WWW.CIA.GOV/LIBRARY/PUBLICATIONS/THE-WORLD- FACTBOOK/GEOS/ID.HTML, 2013).....	9
GAMBAR 2.2 DEMOGRAFI JENIS KELAMIN PENDUDUK INDONESIA TAHUN 2012/2013 (SUMBER : HTTPS://WWW.CIA.GOV/LIBRARY/PUBLICATIONS/THE- WORLD-FACTBOOK/GEOS/ID.HTML, 2013).....	9
GAMBAR 2.3 DAFTAR 20 NEGARA DENGAN AKUN TWITTER TERTINGGI (SUMBER : HTTP://SEMIICAST.COM/PUBLICATIONS/2012_07_30_TWITTER_REACHES_HAL F_A_BILLION_ACCOUNTS_140M_IN_THE_US, 2013) .....	14
GAMBAR 2.4 DAFTAR 20 KOTA DENGAN JUMLAH TWEET TERBANYAK (SUMBER : HTTP://SEMIICAST.COM/PUBLICATIONS/2012_07_30_TWITTER_REACHES_HAL F_A_BILLION_ACCOUNTS_140M_IN_THE_US, 2013) .....	15
GAMBAR 2.5 METODE PEMBENTUKAN MODEL KLASIFIKASI .....	17
GAMBAR 2.6 STRUKTUR POHON KEPUTUSAN .....	19
GAMBAR 2.7 CONTOH DATASET UNTUK MENENTUKAN HIPERTENSI.....	23
GAMBAR 2.8 CONTOH BENTUK POHON KEPUTUSAN .....	25
GAMBAR 2.9 CONTOH DATA HASIL PREDIKSI MENENTUKAN HIPERTENSI .....	27

**Naufal Faruqi, 2018**

*SISTEM PENENTU DEMOGRAFI PELANGGAN BERDASARKAN DATA TWEET DENGAN MENGGUNAKAN ALGORITMA DECISION TREE C4.5*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

GAMBAR 2.10 CONTOH DATA UNTUK MENENTUKAN HIPERTENSI .....	29
GAMBAR 2.11 BENTUK POHON KEPUTUSAN UNTUK MENENTUKAN HIPERTENSI.....	30
GAMBAR 2.12 HASIL PREDIKSI BERDASARKAN POHON KEPUTUSAN YANG TELAH TERBENTUK.....	30
GAMBAR 2.13 HASIL PEMANGKASAN POHON KEPUTUSAN UNTUK MENENTUKAN HIPERTENSI .....	31
GAMBAR 3.1 DESAIN PENELITIAN.....	32
GAMBAR 3.2 MODEL PENGEMBANGAN SEKUENSIAL LINIER (PRESSMAN, 2002)....	35
GAMBAR 4.1 WEKA EXPLORER DATA .....	42
GAMBAR 4.2 HASIL AKURASI WEKA .....	42
GAMBAR 4.3 ANTAR MUKA PERANGKAT LUNAK TAB EKSPLORASI DATA .....	46
GAMBAR 4.4 ANTAR MUKA PERANGKAT LUNAK TAB KLASIFIKASI .....	47
GAMBAR 4.5 MODEL POHON UNTUK KLASIFIKASI GENDER.....	74
GAMBAR 4.6 MODEL POHON UNTUK KLASIFIKASI USIA .....	75
GAMBAR 4.7 MODEL POHON UNTUK KLASIFIKASI STATUS PEKERJAAN.....	76
GAMBAR 4.8 GRAFIK DEMOGRAFI PELANGGAN PERUSAHAAN X BERDASARKAN GENDER.....	77
GAMBAR 4.9 DEMOGRAFI PELANGGAN PERUSAHAAN X BERDASARKAN USIA .....	78
GAMBAR 4.10 DEMOGRAFI PELANGGAN PERUSAHAAN X BERDASARKAN STATUS PEKERJAAN .....	78
GAMBAR 4.11 DEMOGRAFI PELANGGAN PERUSAHAAN X BERDASARKAN GENDER DAN USIA.....	79

GAMBAR 4.12 DEMOGRAFI PELANGGAN PERUSAHAAN X BERDASARKAN GENDER DAN STATUS PEKERJAAN .....	79
GAMBAR 4.13 DEMOGRAFI PELANGGAN PERUSAHAAN X BERDASARKAN USIA DAN STATUS PEKERJAAN.....	80