

DAFTAR ISI

LEMBAR PENGESAHAN	i
PERNYATAAN.....	ii
KATA PENGANTAR	iii
UCAPAN TERIMA KASIH.....	iv
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Tujuan.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	5
BAB II KAJIAN PUSTAKA	6
2.1 Pengertian <i>Natural Language Processing</i>	6
2.1.1 Area <i>Natural Language Processing</i>	7
2.1.2 Terminologi <i>Natural Language Processing</i>	8
2.2 Pengertian <i>Natural Language Generation</i>	12
2.3 Pengertian dan Arsitektur <i>Data-to-text</i>	12
2.4 Penelitian Terkait <i>Data-to-text</i>	18
2.4.1 <i>SoftwareForecast Generator (FOG)</i>	20
2.4.2 Pembangkitan Berita Olah Raga dari <i>Twitter</i>	23
2.4.3 Perangkat Lunak <i>BabyTalk</i>	25

Brahma Putra, 2017

**PENGEMBANGAN SISTEM DATA-TO-TEXT UNTUK MEMBANGKITKAN BERITA CUACA DENGAN
PENDEKATAN TIME-SERIES DALAM R**

2.5	Pengertian dan Sejarah <i>Machine Learning</i>	26
-----	--	----

Brahma Putra, 2017

**PENGEMBANGAN SISTEM DATA-TO-TEXT UNTUK MEMBANGKITKAN BERITA CUACA DENGAN
PENDEKATAN TIME-SERIES DALAM R**

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

2.5.1	<i>Supervised Learning</i>	27
2.5.2	<i>Unsupervised Learning</i>	28
2.5.3	<i>Algoritma Gradient Descent</i>	29
2.6	<i>Time-series Data</i>	30
2.7	<i>Exponential Smoothing</i>	31
2.7.1	<i>Single Exponential Smoothing</i>	32
2.7.2	<i>Double Exponential Smoothing</i>	32
2.7.3	<i>Triple Exponential Smoothing</i>	34
2.8	<i>Logika Fuzzy</i>	34
2.9	<i>R Programming</i>	36
2.9.1	Model data dalam R	38
2.9.2	Contoh kode program bahasa R	39
2.9.3	Contoh visualisasi data dalam R	41
2.9.4	<i>Package</i> alam bahasa R.....	43
BAB III METODOLOGI PENELITIAN.....		44
3.1	Desain Penelitian	44
3.2	Metode Penelitian.....	47
3.3	Alat dan Bahan Penelitian	47
3.3.1	Alat Penelitian.....	47
3.3.2	Bahan Penelitian.....	48
3.4	Proses Pengembangan Perangkat Lunak	49
BAB IV HASIL PENELITIAN DAN PEMBAHASAN		52
4.1	Pengumpulan Data	52
4.2	Pengembangan Model <i>Data-to-text</i> dengan <i>Machine Learning</i> dan <i>Time-series</i>	57
4.2.1	Model Proses <i>Signal Analysis</i>	59
4.2.2	Model Proses <i>Data Interpretation</i>	61
4.2.3	Model Proses <i>Document Planning</i>	71

4.2.4	Model Proses <i>Microplanning</i> dan <i>Realisation</i>	79
4.2.5	Pemanfaatan <i>Machine Learning</i>	87
4.2.6	Model Antarmuka Perangkat Lunak	88
4.3	Implementasi Sistem <i>Data-to-text</i>	89
4.3.1	Instalasi R <i>Packages</i> yang dibutuhkan.....	89
4.3.2	Implementasi Proses <i>Signal Analysis</i>	92
4.3.3	Implementasi Proses <i>Data Interpretation</i>	102
4.3.4	Implementasi Proses <i>Document Planning</i>	107
4.3.5	Implementasi Proses <i>Microplanning</i> dan <i>Realization</i>	108
4.3.6	Implementasi <i>Machine Learning</i> untuk mengatasi missing value	110
4.4	Hasil Implementasi.....	111
4.4.1	<i>Output</i> sistem	111
4.4.2	Perbandingan <i>Output</i> dengan Penelitian Lain.....	113
4.5	Panduan Penggunaan Aplikasi	114
4.6	Rancangan Eksperimen.....	120
4.7	<i>Output</i> Hasil Eksperimen	126
4.8	Analisis dan Hasil Eksperimen.....	131
BAB V KESIMPULAN DAN SARAN.....		137
5.1	Kesimpulan.....	137
5.2	Saran.....	137
DAFTAR PUSTAKA		138
LAMPIRAN.....		144

Brahma Putra, 2017

***PENGEMBANGAN SISTEM DATA-TO-TEXT UNTUK MEMBANGKITKAN BERITA CUACA DENGAN
PENDEKATAN TIME-SERIES DALAM R***

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

DAFTAR TABEL

Tabel 2.1 Penelitian terkait <i>Data-to-text</i> dan <i>Natural Language Generation</i>	18
Tabel 4.1 Kutipan data cuaca dari MeteoGalicia	55
Tabel 4.2 Kutipan data kualitas udara dari MeteoGalicia.....	56
Tabel 4.3 Detail variabel data cuaca	56
Tabel 4.4 Detail variabel data kualitas udara	57
Tabel 4.5 Klasifikasi kecepatan angin Russ Rowlet	63
Tabel 4.6 Klasifikasi untuk menginterpretasikan arah angin menurut <i>Zandlo</i> dkk. ...	64
Tabel 4.7 Klasifikasi cakupan awan Jeff Huby.....	65
Tabel 4.8 Index kualitas udara dalam <i>Pollutant Standard Index (PSI)</i>	67
Tabel 4.9 Keterangan rumus <i>Pollutant Standard Index (PSI)</i>	68
Tabel 4.10 Kutipan data kualitas udara.....	80
Tabel 4.11 Index kualitas udara	81
Tabel 4.12 Nilai kontras dalam proses agregasi dengan <i>simple conjunction</i>	85
Tabel 4.13 data iklim hasil prediksi.	93
Tabel 4.14 Hasil prediksi data kualitas udara.	93
Tabel 4.15 Seluruh fungsi <i>Statistical Event Detection</i> dan keterangannya.....	97
Tabel 4.16 Hasil proses <i>Statistical Event Detection</i>	101
Tabel 4.17 Perbandingan <i>output</i> yang dibangun dengan <i>output</i> dari penelitian terkait lainnya	113
Tabel 4.18 Penilaian kualitas <i>Readability</i> teks dengan menggunakan <i>Flesch Reading Ease</i>	122
Tabel 4.19 Keterangan rentang kesesuaian kuisisioner.	123
Tabel 4.20 Hasil evaluasi dengan menggunakan <i>Flesch Reading Ease</i>	131
Tabel 4.21 Hasil perhitungan waktu komputasi dengan <i>System.time()</i>	132
Tabel 4.22 Hasil evaluasi <i>experts</i>	133
Tabel 4.23 Hasil evaluasi user terkait aspek <i>comprehensability</i> dan <i>importance</i>	135

DAFTAR GAMBAR

Gambar 2.1 Contoh skema <i>input-output</i> sistem <i>Data-to-text</i>	13
Gambar 2.2 Arsitektur <i>Data-to-text</i> yang diusulkan Reiter	14
Gambar 2.3 Contoh <i>output software Forecast Generator</i> (FOG).....	21
Gambar 2.4 Diagram skematik <i>software Forecast Generator</i> (FOG)	22
Gambar 2.5 Contoh <i>output</i> ringkasan berita olah raga	24
Gambar 2.6 Perbandingan <i>output</i> sistem dan ringkasan manual.	24
Gambar 2.7 Skema <i>input-output</i> aplikasi <i>BabyTalk</i>	25
Gambar 2.8 Contoh <i>Supervised</i> learning pada pengenalan koin.....	28
Gambar 2.9 Contoh <i>Unsupervised</i> Learning dalam pengenalan koin.....	29
Gambar 2.10 Contoh himpunan <i>Crisp</i> pada kasus umur	35
Gambar 2.11 Contoh himpunan fuzzy pada kasus umur	36
Gambar 2.12 Logo bahasa pemrograman R.....	36
Gambar 2.13 Antarmuka R <i>Graphical User Interface</i> (RGui)	38
Gambar 2.14 Model data dalam pemrograman R	39
Gambar 2.15 Operator <i>concatenate</i> dalam R.....	39
Gambar 2.16 Menampilkan dua data pertama dalam R.....	40
Gambar 2.17 Penggunaan fungsi <i>summarize</i> dalam R.....	40
Gambar 2.18 Penggabungan data dengan <i>concatenate</i> dalam R	40
Gambar 2.19 Pembuatan <i>matriks</i> dalam R.....	41
Gambar 2.20 Contoh visualisasi grafis dalam R.....	41
Gambar 2.21 Contoh perulangan dalam R.....	42
Gambar 2.22 Contoh implementasi <i>decision</i> dalam R.....	42
Gambar 2.23 Contoh fungsi dalam R.....	43
Gambar 3.1 Metodologi Penelitian Pengembangan Sistem <i>Data-to-text</i> dengan Pendekatan <i>Time-series</i>	46
Gambar 3.2 Pengembangan perangkat lunak model <i>Waterfall</i>	49
Gambar 4.1 Model <i>Data-to-text</i> dengan pendekatan <i>Machine Learning</i> dan <i>Time Series</i>	58

Gambar 4.2 Model proses <i>Signal Analysis</i> prediksi dengan <i>exponential smoothing</i> dalam membangun pesan deskripsi cuaca dan kualitas udara.	60
Gambar 4.3 Model proses <i>Signal Analysis</i> pencarian event dalam membangun pesan deskripsi cuaca	61
Gambar 4.4 Model Proses <i>Data Interpretation</i> dengan menggunakan <i>Expert Rules</i> . 62	
Gambar 4.5 Fungsi keanggotaan suhu	65
Gambar 4.6 Fungsi keanggotaan curah hujan	66
Gambar 4.7 Perhitungan <i>Pollutant Standard Index (PSI)</i>	68
Gambar 4.8 Hasil perhitungan <i>Pollutant Standard Index (PSI)</i>	70
Gambar 4.9 Hasil perhitungan <i>Pollutant Standard Index (PSI)</i>	70
Gambar 4.10 Model <i>Document Planning</i>	71
Gambar 4.11 <i>Routine message</i> untuk pesan prediksi.....	73
Gambar 4.12 <i>Significant Event Message</i> untuk pesan prediksi	73
Gambar 4.13 <i>Target corpus</i> hasil prediksi.....	74
Gambar 4.14 Skema untuk pesan prediksi.....	75
Gambar 4.15 Event Tree untuk pesan prediksi.	75
Gambar 4.16 Target Teks untuk <i>weather summary</i>	76
Gambar 4.17 Target teks untuk <i>weather summary</i>	76
Gambar 4.18 Content Determination untuk <i>weather summary</i>	77
Gambar 4.19 <i>Rhetorical Structured Text</i> untuk <i>weather summary</i>	77
Gambar 4.20 Model <i>Microplanning</i> dan Realisation untuk proses prediksi dan <i>weather summary</i>	78
Gambar 4.21 Kutipan inisial target korpus	80
Gambar 4.22 Trend Description untuk mendeskripsikan persepsi perubahan data kualitas udara.	82
Gambar 4.23 <i>Progressive Change word bank</i>	83
Gambar 4.24 Contoh Phrase Aggregation	84
Gambar 4.25 Struktur file aplikasi D2T_Apps.	88
Gambar 4.26 Purwa rupa antarmuka sistem <i>Data-to-text Weather Prediction (DWP)</i>	88

Gambar 4.27 Perintah untuk instalasi <i>packages Shiny R</i> yang berfungsi dalam membangun antarmuka aplikasi berbasis web	90
Gambar 4.28 Perintah untuk instalasi <i>packages Smooth</i> yang berfungsi dalam membangun model prediksi berbasis <i>time-series</i> dengan algoritma <i>exponential smoothing</i>	90
Gambar 4.30 Perintah untuk instalasi package <i>xts</i> untuk mendefinisikan dataset kedalam objek <i>time-series</i> berdasarkan tanggal.....	91
Gambar 4.29 Perintah untuk instalasi <i>package gradDesc</i> untuk melakukan prediksi dengan pendekatan <i>machine Learning</i> menggunakan algoritma <i>gradient descet</i>	91
Gambar 4.31 Prediksi dengan <i>exponential smoothing</i> dalam R	92
Gambar 4.33 Memuat data kualitas udara kedalam sistem.....	94
Gambar 4.34 Mengambil data satu bulan terakhir	94
Gambar 4.32 Memuat data kedalam sistem	94
Gambar 4.36 Mencari nilai minimum setiap variabel.....	95
Gambar 4.37 Mencari nilai rata-rata	95
Gambar 4.35 Mencari nilai maximum setiap variabel.	95
Gambar 4.38 Mencari nilai total setiap variabel	96
Gambar 4.39 <i>Output</i> standar statistik.....	96
Gambar 4.40 Pendeteksi hari hujan berturut-turut.....	99
Gambar 4.41 <i>MonthlyTempMessage</i> function	100
Gambar 4.42 Fungsi mendeteksi keadaan kualitas udara bulan terakhir	100
Gambar 4.43 Pseudocode proses interpretasi data	102
Gambar 4.44 <i>Data Interpretation</i> dalam R.....	102
Gambar 4.45 <i>Input</i> kecepatan angin.....	103
Gambar 4.46 <i>Input</i> arah angin.....	103
Gambar 4.47 Fungsi <i>AirQuality_Interpreter</i> untuk menginterpretasikan kualitas udara.....	104
Gambar 4.48 Fungsi untuk menginterpretasikan curah hujan dan suhu.	105
Gambar 4.49 Hasil proses data interpretation	106

Gambar 4.50 <i>Content Selection</i> saat menggabungkan pesan yang berelasi (<i>Wind State</i>)	106
Gambar 4.51 <i>Source code Trend Description</i>	107
Gambar 4.52 Pseudocode Aggregation	108
Gambar 4.53 <i>Referring Expression generation</i>	108
Gambar 4.54 <i>Structure Realization</i>	109
Gambar 4.55 <i>Machine Learning</i> dalam R untuk mengatasi <i>missing value</i>	109
Gambar 4.56 Hasil keluaran sistem.	110
Gambar 4.57 Jendela saat proses instalasi aplikasi <i>R for Windows</i>	114
Gambar 4.58 Jendela antarmuka R <i>Graphical User Interface (RGui)</i>	115
Gambar 4.60 Instalasi packages ShinyR	116
Gambar 4.61 Contoh <i>output</i> dari <i>console</i> jika instalasi <i>packages</i> ShinyR berhasil. 116	
Gambar 4.62 Keluaran dari <i>console</i> setelah server dijalankan pada port 6632	116
Gambar 4.63 Contoh jendela browser yang muncul ketika menjalankan perintah <code>runApp("D2T_Apps")</code>	117
Gambar 4.64 Isi dari direktori <i>D2T_Apps</i>	117
Gambar 4.65 Contoh isi dari file <i>D2T_Machine.R</i>	118
Gambar 4.66 Folder <i>Datasets</i> , berisi seluruh dataset yang tersedia.	119
Gambar 4.67 Cara mengganti dataset yang digunakan dalam file <i>D2T_Machine.R</i> 119	
Gambar 4.68 Dataset eksperimen	121
Gambar 4.69 Evaluasi aspek <i>relevance</i> dan <i>truthfulness</i> oleh <i>expert</i>	124
Gambar 4.70 Evaluasi aspek <i>comprehensability</i> dan <i>importance</i> oleh <i>expert</i>	125
Gambar 4.72 Hasil keluaran dari eksperimen ke-dua dengan data real 2016-2017. 126	
Gambar 4.71 Hasil keluaran dari eksperimen pertama dengan dataset <i>Modified1</i> ... 126	
Gambar 4.73 Hasil keluaran dari eksperimen ke-tiga dengan data <i>Modified2</i>	127
Gambar 4.74 Hasil keluaran dari eksperimen ke-empat dengan data real 2014-2015	127
Gambar 4.75 Hasil keluaran dari eksperimen ke-lima dengan data real 2015-2016 128	
Gambar 4.76 Hasil keluaran dari eksperimen ke-enam dengan data real 2013-2014	128

Gambar 4.77 Hasil keluaran dari eksperimen ke-tujuh dengan data Modified 3.....	129
Gambar 4.78 Hasil keluaran dari eksperimen ke-delapan dengan data Modified 4.	129
Gambar 4.79 Hasil keluaran dari eksperimen ke-sembilan dengan data Modified 5.	130
Gambar 4.80 Hasil keluaran dari eksperimen ke-sepuluh dengan data Modified 6.	130
Gambar 4.81 Contoh kuisisioner yang diisi oleh pakar klimatologi	133
Gambar 4.82 Contoh kuisisioner yang diisi oleh user.	135