



ANALISIS KINERJA *HYBRID CHI-SQUARE* DAN *BINARY PARTICLE SWARM OPTIMIZATION* DENGAN *ADAPTIVE INERTIA WEIGHTS* PADA *DISTANCE BIASED NAÏVE BAYES* UNTUK KLASIFIKASI SENTIMEN

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer**

**Oleh
MUHAMMAD IMAM SUFIAZI
NIM. 1711016310016**

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

MARET 2023



ANALISIS KINERJA *HYBRID CHI-SQUARE* DAN *BINARY PARTICLE SWARM OPTIMIZATION* DENGAN *ADAPTIVE INERTIA WEIGHTS* PADA *DISTANCE BIASED NAÏVE BAYES* UNTUK KLASIFIKASI SENTIMEN

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer**

**Oleh
MUHAMMAD IMAM SUFIAZI
NIM. 1711016310016**

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
MARET 2023**

SKRIPSI

ANALISIS KINERJA *HYBRID CHI-SQUARE* DAN *BINARY PARTICLE SWARM OPTIMIZATION* DENGAN *ADAPTIVE INERTIA WEIGHTS* PADA *DISTANCE BIASED NAÏVE BAYES* UNTUK KLASIFIKASI SENTIMEN

Oleh :

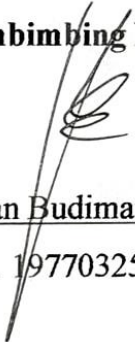
MUHAMMAD IMAM SUFIAZI

1711016310016

Telah dipertahankan di depan Dosen Penguji pada tanggal 27 Maret 2023

Susunan Dosen Penguji :

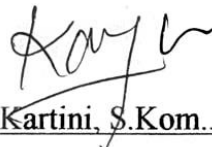
Pembimbing I



Irwan Budiman, S.T., M.Kom.

NIP. 197703252008121001

Dosen Penguji I



Dwi Kartini, S.Kom., M.Kom.

NIP. 198704212012122003

Pembimbing II



M. Itqan Mazdadi, S.Kom., M.Kom.

NIP. 199006122019031013

Dosen Penguji II



M. Reza Faisal, S.T., M.T., Ph.D.

NIP. 197612202008121001

Banjarbaru, 27 Maret 2023

Ketua Program Studi Ilmu Komputer



Irwan Budiman, S.T., M.Kom.


NIP. 197703252008121001

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu 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 disebutkan dalam daftar pustaka.

Banjarbaru, 27 Maret 2023

Yang Menyatakan,



Muhammad Imam Sufiazi

NIM. 1711016310016

ABSTRAK

ANALISIS KINERJA *HYBRID CHI-SQUARE* DAN *BINARY PARTICLE SWARM OPTIMIZATION* DENGAN *ADAPTIVE INERTIA WEIGHTS* PADA *DISTANCE BIASED NAÏVE BAYES* UNTUK KLASIFIKASI SENTIMEN

(Oleh: Muhammad Imam Sufiazi; Pembimbing: Irwan Budiman, S.T., M.Kom. dan Muhammad Itqan Mazdadi, S.Kom, M.Kom.; 2023; 103 halaman)

Media sosial adalah ruang terbaik bagi pengguna untuk dapat mengekspresikan pandangannya dengan berbagai perspektif. Twitter menjadi media sosial yang sangat populer hingga saat ini, pengguna dapat membagikan berbagai opininya dengan melakukan *tweet* dan *retweet*. Akan tetapi, dengan semakin banyaknya opini yang dibagikan oleh pengguna di Twitter setiap hari, dibutuhkan teknologi untuk dapat menganalisisnya secara akurat dan presisi, yang dikenal dengan analisis sentimen. Analisis sentimen juga dikenal dengan *opinion mining*, yang menggabungkan bidang *text mining*, dan *natural language processing*. Namun, ada tiga masalah mendasar pada analisis sentimen, yaitu identifikasi sentimen, seleksi fitur dan klasifikasi sentimen. Dengan demikian, salah satu masalah utama dalam analisis sentimen adalah untuk meningkatkan kualitas seleksi fitur. Pada penelitian ini, akan melakukan seleksi fitur menggunakan metode *Hybrid Chi-Square* dan *Binary Particle Swarm Optimization* dengan *Adaptive Inertia Weights*, serta klasifikasi sentimen menggunakan metode *Distance Biased Naïve Bayes*, pada dataset sentimen dengan topik mengenai COVID-19 berbahasa Indonesia dan *Weather* berbahasa Inggris yang bersumber dari Twitter. Dari penelitian ini, pada dataset COVID-19 model tanpa seleksi fitur menghasilkan performa terbaik dengan nilai *accuracy*, *precision*, *recall*, *f1-score* dan *time computation* secara berurutan sebesar 80,00%, 75,00%, 90,00%, 81,81%, dan 16,59 detik, sedangkan model dengan seleksi fitur menghasilkan performa terbaik dengan nilai *accuracy*, *precision*, *recall*, *f1-score* dan *time computation* secara berurutan sebesar 87,50%, 82,60%, 95,00%, 88,37%, dan 1,87 detik. Kemudian pada dataset *Weather* tanpa seleksi fitur menghasilkan performa terbaik dengan nilai *accuracy*, *precision*, *recall*, *f1-score* dan *time computation* secara berurutan sebesar 77,27%, 73,07%, 86,36%, 79,16%, dan 17,28 detik, sedangkan model dengan seleksi fitur menghasilkan performa terbaik dengan nilai *accuracy*, *precision*, *recall*, *f1-score* dan *time computation* secara berurutan sebesar 86,36%, 83,33%, 90,90%, 86,95%, dan 0,60 detik. Dapat disimpulkan bahwa model dengan seleksi fitur pada analisis sentimen dapat memberikan peningkatan yang signifikan dibandingkan dengan model tanpa seleksi fitur.

Kata kunci: *Opinion Mining*, *Text Mining*, *Natural Language Processing*, Analisis Sentimen, Seleksi Fitur, Klasifikasi, *Chi-Square*, *Binary Particle Swarm Optimization*, dan *Distance Biased Naive Bayes*.

ABSTRACT

PERFORMANCE ANALYSIS OF HYBRID CHI-SQUARE AND BINARY PARTICLE SWARM OPTIMIZATION WITH ADAPTIVE INERTIA WEIGHTS AT DISTANCE BIASED NAÏVE BAYES FOR SENTIMENT CLASSIFICATION

(By: Muhammad Imam Sufiazi; Supervisor: Irwan Budiman, S.T., M.Kom. and Muhammad Itqan Mazdadi, S.Kom, M.Kom.; 2023; 103 pages)

Social media is the best space for users to be able to express their views with various perspectives. Twitter has become a very popular social media to date, users can share various opinions by tweeting and retweeting. However, with the increasing number of opinions shared by users on Twitter every day, technology is needed to be able to analyze them with accuracy and precision, which is known as sentiment analysis. Sentiment analysis is also known as opinion mining, which combines the fields of text mining and natural language processing. However, there are three basic problems in sentiment analysis, namely sentiment identification, feature selection and sentiment classification. Thus, one of the main problems in sentiment analysis is to improve the quality of feature selection. In this study, feature selection will be carried out using the Hybrid Chi-Square and Binary Particle Swarm Optimization methods with Adaptive Inertia Weights, as well as sentiment classification using the Distance Biased Naïve Bayes methods, on sentiment datasets with the topic of COVID-19 in Indonesian and Weather in English sourced from Twitter. From this study, the COVID-19 dataset model without feature selection produces the best performance with values of accuracy, precision, recall, f1-score and time computation respectively of 80,00%, 75,00%, 90,00%, 81,81%, and 16,59 seconds, while the model with feature selection produces the best performance with values of accuracy, precision, recall, f1-score and time computation respectively of 87,50%, 82,60%, 95,00%, 88,37%, and 1,87 seconds. Then the Weather dataset without feature selection produces the best performance with values of accuracy, precision, recall, f1-score and time computation respectively of 77,27%, 73,07%, 86,36%, 79,16%, and 17,28 seconds, while the model with feature selection produces the best performance with values of accuracy, precision, recall, f1-score and time computation respectively of 86,36%, 83,33%, 90,90%, 86,95%, and 0,60 sec. It can be concluded that the model with feature selection in sentiment analysis can provide a significant improvement compared to the model without feature selection.

Keywords: Opinion Mining, Text Mining, Natural Language Processing, Sentiment Analysis, Feature Selection, Classification, Chi-Square, Binary Particle Swarm Optimization, and Distance Biased Naive Bayes.

PRAKATA

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Kuasa atas segala limpahan rahmat, kemudahan, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Kinerja *Hybrid Chi-Square* dan *Binary Particle Swarm Optimization* dengan *Adaptive Inertia Weights* pada *Distance Biased Naïve Bayes* untuk Klasifikasi Sentimen” untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut:

1. Orang tua tercinta untuk Mama dan Abah, yang sejak awal perkuliahan hingga penyelesaian skripsi ini selalu memberikan bantuan, semangat, dukungan dan doa.
2. Bapak Irwan Budiman, S.T., M.Kom. selaku dosen pembimbing utama, yang senantiasa meluangkan waktunya untuk terus membimbing dan memberi arahan selama proses penyelesaian skripsi ini.
3. Bapak Muhammad Itqan Mazdadi, S.Kom., M.Kom. selaku dosen pembimbing pendamping, yang juga turut serta meluangkan waktu untuk membimbing dan memberi arahan selama proses penyelesaian skripsi ini.
4. Bapak Irwan Budiman, S.T., M.Kom. selaku Koordinator Program Studi Ilmu Komputer FMIPA ULM dan Ibu Dwi Kartini, S.Kom., M.Kom. selaku Ketua Panitia Skripsi, atas bantuan dan izin beliau skripsi ini dapat diselesaikan.
5. Seluruh dosen dan staf Program Studi Ilmu Komputer FMIPA ULM, atas ilmu, bimbingan dan bantuan yang diberikan selama menempuh perkuliahan yang sangat bermanfaat.
6. Ibu Dwi Kartini, S.Kom, M.Kom. dan Bapak M. Reza Faisal, S.T, M.T, Ph.D., selaku tim penguji yang juga turut serta menguji, memberi saran dan arahan pada penelitian ini.

7. Kakak tercinta untuk Kak Doni dan Kak Puput, yang memberikan dorongan dan semangat untuk terus menyelesaikan skripsi ini hingga selesai.
8. Keponakan tercinta untuk Dena, Athar dan Aira, yang selalu menghibur dan menemani bermain saat penulis pulang kerumah.
9. Ratna Septia Devi, yang selalu memberikan bantuan dan dukungan saat penulis mengalami hambatan dan kendala dalam menyelesaikan skripsi.
10. Yanche Kurniawan, Nursyifa Azizah, dan Ahmad Bakeri, yang sering menjadi teman berkumpul dan bercerita dikampus dan diperpus selama penulis mengerjakan skripsi.
11. Teman-teman keluarga Ilmu Komputer angkatan 2017, terimakasih atas canda, tawa dan perjuangan bersama selama penulis menempuh perkuliahan dari awal hingga selesainya skripsi ini, sangat banyak pengalaman yang penulis dapatkan bersama kalian semua.
12. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam penyelesaian skripsi ini.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh dari sempurna. Oleh karena itu penulis mengharapkan adanya kritik dan saran yang konstruktif. Semoga tulisan ini dapat memberikan manfaat bagi ilmu pengetahuan dan bagi para pembaca.

Banjarbaru, 27 Maret 2023



Muhammad Imam Sufiazi

DAFTAR ISI

HALAMAN JUDUL	i
SKRIPSI.....	ii
PERNYATAAN.....	iii
ABSTRAK	iv
ABSTRACT	v
PRAKATA	vi
DAFTAR ISI.....	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Kajian Terdahulu	5
2.2 Keaslian Penelitian	8
2.3 Opini Publik.....	11
2.4 Analisis Sentimen	11
2.5 Text Mining	12
2.6 Ekstraksi Fitur.....	14

2.7	Seleksi Fitur	15
2.8	<i>Chi-Square</i> (CS)	15
2.9	<i>Particle Swarm Optimization</i> (PSO)	16
2.10	<i>Binary Particle Swarm Optimization</i> (BPSO).....	17
2.11	<i>Adaptive Inertia Weights</i> (AIW).....	19
2.12	<i>Hybrid CS-BPSO</i>	21
2.13	Klasifikasi.....	21
2.14	<i>Naïve Bayes</i> (NB)	22
2.15	<i>Distance Biased Naïve Bayes</i> (DBNB).....	23
	2.15.1 <i>Weighted Naïve Bayes Module</i> (WNBM).....	24
	2.15.2 <i>Distance Reinforcement Module</i> (DRM).....	25
2.16	<i>Split Validation</i>	26
2.17	Evaluasi Kinerja Algoritma	27
	BAB III METODE PENELITIAN	29
3.1	Alat Penelitian	29
3.2	Bahan Penelitian	29
3.3	Variabel Penelitian.....	29
3.4	Prosedur Penelitian	30
	BAB IV HASIL DAN PEMBAHAN.....	35
4.1	Hasil.....	35
	4.1.1 Pengumpulan dan Analisis Dataset.....	35
	4.1.2 <i>Preprocessing</i> Data.....	37
	4.1.3 Ekstraksi Fitur.....	48
	4.1.4 Pembuatan Model	56
4.2	Pembahasan	95

BAB V PENUTUP	102
5.1 Kesimpulan.....	102
5.2 Saran	102

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR TABEL

Tabel	Halaman
Tabel 1 Keaslian Penelitian.....	8
Tabel 2 Rancangan Penelitian.....	10
Tabel 3 <i>Confusion Matrix</i> Klasifikasi <i>Binary</i>	28
Tabel 4 Perbedaan Karakteristik Dataset.....	31
Tabel 5 Contoh Teks <i>Tweet</i> Dataset COVID-19.....	31
Tabel 6 Contoh Teks <i>Tweet</i> Dataset <i>Weather</i>	32
Tabel 7 Teks <i>Tweet</i> Dataset COVID-19.....	35
Tabel 8 Teks <i>Tweet</i> Dataset <i>Weather</i>	37
Tabel 9 Hasil Penerapan <i>Case Folding</i> pada Dataset COVID-19.....	38
Tabel 10 Hasil Penerapan <i>Case Folding</i> pada Dataset <i>Weather</i>	39
Tabel 11 Kamus Karakter URL.....	39
Tabel 12 Kamus Karakter <i>Hashtag</i> , <i>Mention</i> dan Elemen HTML.....	40
Tabel 13 Kamus Karakter Numerik.....	40
Tabel 14 Kamus Karakter ASCII.....	40
Tabel 15 Kamus Karakter Tanda Baca.....	40
Tabel 16 Kamus Kata <i>Stopword</i> Bahasa Indonesia.....	41
Tabel 17 Hasil Penerapan <i>Cleansing</i> pada Dataset COVID-19.....	41
Tabel 18 Kamus Kata <i>Stopword</i> Bahasa Inggris.....	42
Tabel 19 Hasil Penerapan <i>Cleansing</i> pada Dataset <i>Weather</i>	42
Tabel 20 Kamus Kata <i>Slang</i> Bahasa Indonesia dari Penelitian Terdahulu.....	43
Tabel 21 Kamus Kata <i>Slang</i> Bahasa Indonesia dari Hasil Observasi.....	43
Tabel 22 Hasil Penerapan <i>Slang Words</i> pada Dataset COVID-19.....	44
Tabel 23 Kamus Kata <i>Slang</i> Bahasa Inggris dari Situs Web <i>Internet & Text Slang Dictionary & Translator</i>	44
Tabel 24 Hasil Penerapan <i>Slang Words</i> pada Dataset <i>Weather</i>	45
Tabel 25 Kamus Kata Dasar Bahasa Indonesia.....	45
Tabel 26 Hasil Penerapan <i>Stemming</i> pada Dataset COVID-19.....	45
Tabel 27 Kamus Kata Dasar Bahasa Inggris.....	46

Tabel 28 Hasil Penerapan <i>Stemming</i> pada Dataset <i>Weather</i>	46
Tabel 29 Hasil Penerapan <i>Tokenization</i> pada Dataset COVID-19	47
Tabel 30 Hasil Penerapan <i>Tokenization</i> pada Dataset <i>Weather</i>	48
Tabel 31 Perhitungan <i>Term Frequency (TF)</i> pada Dataset COVID-19.....	48
Tabel 32 Perhitungan <i>Inverse Document Frequency (IDF)</i> pada Dataset COVID-19.....	49
Tabel 33 Perhitungan Perkalian <i>Term Frequency (TF)</i> dengan <i>Inverse Document Frequency (IDF)</i> pada Dataset COVID-19	50
Tabel 34 Perhitungan Normalisasi <i>Term Frequency–inverse Document Frequency (TF-IDF)</i> pada Dataset COVID-19.....	51
Tabel 35 Ekstraksi Fitur pada Dataset COVID-19	53
Tabel 36 Perhitungan <i>Term Frequency (TF)</i> pada Dataset <i>Weather</i>	53
Tabel 37 Perhitungan <i>Inverse Document Frequency (IDF)</i> pada Dataset <i>Weather</i>	54
Tabel 38 Perhitungan Perkalian <i>Term Frequency (TF)</i> dengan <i>Inverse Document Frequency (IDF)</i> pada Dataset <i>Weather</i>	54
Tabel 39 Perhitungan Normalisasi <i>Term Frequency–inverse Document Frequency (TF-IDF)</i> pada Dataset <i>Weather</i>	55
Tabel 40 Ekstraksi Fitur pada Dataset <i>Weather</i>	56
Tabel 41 Hasil Penerapan Pembagian Data dengan setiap Rasio pada Dataset COVID-19.....	57
Tabel 42 Hasil Pengujian Model Perbandingan Rasio pada Dataset COVID-19 .	57
Tabel 43 Hasil Penerapan Pembagian Data dengan setiap Rasio pada Dataset <i>Weather</i>	58
Tabel 44 Hasil Pengujian Model Perbandingan Rasio pada Dataset <i>Weather</i>	59
Tabel 45 Perhitungan <i>Observed</i> setiap Kelas (Rasio Pembagian Data 90:10) pada Dataset COVID-19	61
Tabel 46 Perhitungan <i>Expected</i> setiap Fitur (Rasio Pembagian Data 90:10) pada Dataset COVID-19	61
Tabel 47 Perhitungan <i>Chi-Square (CS)</i> setiap Kelas (Rasio Pembagian Data 90:10) pada Dataset COVID-19.....	61

Tabel 48 Perhitungan penjumlahan nilai <i>Chi-Square</i> (CS) setiap Kelas (Rasio Pembagian Data 90:10) pada Dataset COVID-19	62
Tabel 49 Hasil Penerapan Perangkingan Fitur (Rasio Pembagian Data 90:10) pada Dataset COVID-19	62
Tabel 50 Hasil Penerapan Seleksi Fitur <i>Chi-Square</i> (CS) setiap Persentase pada Dataset COVID-19	62
Tabel 51 Perhitungan <i>Observed</i> setiap Kelas (Rasio Pembagian Data 90:10) pada Dataset <i>Weather</i>	63
Tabel 52 Perhitungan <i>Expected</i> setiap Fitur (Rasio Pembagian Data 90:10) pada Dataset <i>Weather</i>	64
Tabel 53 Perhitungan <i>Chi-Square</i> (CS) setiap Kelas (Rasio Pembagian Data 90:10) pada Dataset <i>Weather</i>	64
Tabel 54 Perhitungan penjumlahan nilai <i>Chi-Square</i> (CS) setiap Kelas (Rasio Pembagian Data 90:10) pada Dataset <i>Weather</i>	64
Tabel 55 Hasil Penerapan Perangkingan Fitur pada (Rasio Pembagian Data 90:10) Dataset <i>Weather</i>	65
Tabel 56 Hasil Penerapan Seleksi Fitur <i>Chi-Square</i> (CS) setiap Persentase pada Dataset <i>Weather</i>	65
Tabel 57 Parameter yang digunakan pada <i>Binary Particle Swarm Optimization</i> dengan <i>Adaptive Inertia Weights</i> (BPSO-AIW)	66
Tabel 58 Skema Perhitungan untuk <i>Update</i> Nilai Bobot Inersia pada <i>Adaptive Inertia Weights</i> (AIW).....	67
Tabel 59 Inisiasi Nilai Posisi dan <i>Velocity</i> Awal setiap Partikel (Persentase 80% <i>Chi-Square</i>) pada Dataset COVID-19	67
Tabel 60 Perhitungan Nilai <i>Fitness Score</i> Iterasi Pertama setiap Partikel (Persentase 80% <i>Chi-Square</i>) pada Dataset COVID-19	68
Tabel 61 Perhitungan <i>Update</i> Nilai Posisi dan <i>Velocity</i> Iterasi Pertama setiap Partikel (Persentase 80% <i>Chi-Square</i>) pada Dataset COVID-19	69
Tabel 62 <i>Update</i> Nilai Posisi dan <i>Velocity</i> setiap Partikel (Persentase 80% <i>Chi-Square</i>) pada Dataset COVID-19.....	70

Tabel 63 Perhitungan Nilai <i>Fitness Score</i> Iterasi Kedua setiap Partikel (Persentase 80% <i>Chi-Square</i>) pada Dataset COVID-19	71
Tabel 64 Hasil Penerapan Seleksi Fitur <i>Binary Particle Swarm Optimization</i> dengan <i>Adaptive Inertia Weights</i> (BPSO-AIW) setiap Persentase <i>Chi-Square</i> (CS) pada Dataset COVID-19	72
Tabel 65 Hasil Pengujian Model setiap Persentase <i>Chi-Square</i> (CS) pada Dataset COVID-19.....	73
Tabel 66 Inisiasi Nilai Posisi dan <i>Velocity</i> Awal setiap Partikel (Persentase 80% <i>Chi-Square</i>) pada Dataset <i>Weather</i>	75
Tabel 67 Perhitungan Nilai <i>Fitness Score</i> Iterasi Pertama setiap Partikel (Persentase 80% <i>Chi-Square</i>) pada Dataset <i>Weather</i>	75
Tabel 68 Perhitungan <i>Update</i> Nilai Posisi dan <i>Velocity</i> Iterasi Pertama seluruh Partikel (Persentase 80% <i>Chi-Square</i>) pada Dataset <i>Weather</i>	76
Tabel 69 <i>Update</i> Nilai Posisi dan <i>Velocity</i> setiap Partikel (Persentase 80% <i>Chi- Square</i>) pada Dataset <i>Weather</i>	77
Tabel 70 Perhitungan Nilai <i>Fitness Score</i> Iterasi Kedua setiap Partikel (Persentase 80% <i>Chi-Square</i>) pada Dataset <i>Weather</i>	78
Tabel 71 Hasil Penerapan Seleksi Fitur <i>Binary Particle Swarm Optimization</i> dengan <i>Adaptive Inertia Weights</i> (BPSO-AIW) setiap Persentase <i>Chi-Square</i> (CS) pada Dataset <i>Weather</i>	79
Tabel 72 Hasil Pengujian Model setiap Persentase <i>Chi-Square</i> (CS) pada Dataset <i>Weather</i>	80
Tabel 73 Perhitungan Jumlah Seluruh Nilai <i>Instances</i> pada Data Latih setiap Fitur dan setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset COVID-19	82
Tabel 74 Perhitungan Nilai Probabilitas Jumlah Seluruh Nilai <i>Instances</i> pada Data Latih setiap Fitur dan setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset COVID-19	82
Tabel 75 Perhitungan Nilai Probabilitas setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset COVID-19	83

Tabel 76 Perhitungan Nilai Bobot Fitur dan Normalisasi Bobot Fitur setiap Fitur (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset COVID-19.....	83
Tabel 77 Perhitungan Nilai Probabilitas Data Uji (Data Uji Pertama) setiap Fitur dan setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset COVID-19.....	84
Tabel 78 Perhitungan Nilai <i>Belonging Score</i> (BS) (Data Uji Pertama) setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset COVID-19.....	85
Tabel 79 Perhitungan Class Center Data Latih setiap Fitur dan setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset COVID-19.....	85
Tabel 80 Perhitungan Jarak <i>Euclidean Distance</i> (Data Uji Pertama) setiap kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset COVID-19.....	86
Tabel 81 Perhitungan <i>Affiliation Degree</i> (AD) (Data Uji Pertama) setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset COVID-19.....	86
Tabel 82 Detail Evaluasi Model Klasifikasi <i>Distance Biased Naive Bayes</i> (DBNB) Tanpa Seleksi Fitur dan Dengan Seleksi Fitur pada Dataset COVID-19.....	87
Tabel 83 Hasil <i>Confussion Matrix</i> pada Dataset COVID-19.....	88
Tabel 84 Perhitungan Jumlah Nilai Data Latih Setiap Fitur dan Setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset <i>Weather</i>	89
Tabel 85 Perhitungan Nilai Probabilitas Jumlah Nilai Data Latih Setiap Fitur dan Setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset <i>Weather</i>	89
Tabel 86 Perhitungan Nilai Probabilitas Setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset <i>Weather</i>	89

Tabel 87 Perhitungan Nilai Bobot Fitur dan Normalisasi Bobot Fitur setiap Fitur (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset <i>Weather</i>	90
Tabel 88 Perhitungan Nilai Probabilitas Data Uji (Data Uji Pertama) setiap Fitur dan setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset <i>Weather</i>	91
Tabel 89 Perhitungan Nilai <i>Belonging Score</i> (BS) (Data Uji Pertama) setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset <i>Weather</i>	91
Tabel 90 Perhitungan Class Center Data Latih setiap Fitur dan setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset <i>Weather</i>	92
Tabel 91 Perhitungan Jarak <i>Euclidean Distance</i> (Data Uji Pertama) setiap kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset <i>Weather</i>	92
Tabel 92 Perhitungan <i>Affiliation Degree</i> (AD) (Data Uji Pertama) setiap Kelas (Rasio Pembagian Data 90:10) (Tanpa Seleksi Fitur) pada Dataset <i>Weather</i>	93
Tabel 93 Detail Evaluasi Model Klasifikasi <i>Distance Biased Naive Bayes</i> (DBNB) Tanpa Seleksi Fitur dan Dengan Seleksi Fitur pada Dataset <i>Weather</i>	93
Tabel 94 Hasil <i>Confussion Matrix</i> pada Dataset <i>Weather</i>	95
Tabel 95 Detail Evaluasi Model Klasifikasi <i>Distance Biased Naive Bayes</i> (DBNB) Tanpa Seleksi Fitur	99
Tabel 96 Detail Evaluasi Model Klasifikasi <i>Distance Biased Naive Bayes</i> (DBNB) Dengan Seleksi Fitur	100

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1 Diagram Venn Interaksi <i>Text Mining</i> dengan Bidang Lain	13
Gambar 2 Proses Umum <i>Text Mining</i>	13
Gambar 3 Alur Kombinasi Metode Seleksi Fitur	21
Gambar 4 Ilustrasi Teorema Bayes	23
Gambar 5 Ilustrasi Contoh Perhitungan Jarak ke-3 Kelas Target.....	26
Gambar 6 Ilustrasi Pembagian Data Split Validation	27
Gambar 7 Prosedur Penelitian.....	30
Gambar 8 Distribusi Kelas pada Dataset COVID-19	36
Gambar 9 Distribusi Kelas pada Dataset <i>Weather</i>	37
Gambar 10 Performa <i>Accuracy</i> Model Pembagian Data pada Dataset COVID-19.....	58
Gambar 11 Performa <i>Accuracy</i> Model Pembagian Data pada Dataset <i>Weather</i> ..	59
Gambar 12 Fitur Terpilih <i>Chi-Square</i> (CS) pada Dataset COVID-19.....	63
Gambar 13 Fitur Terpilih <i>Chi-Square</i> (CS) pada Dataset <i>Weather</i>	65
Gambar 14 Fitur Terpilih <i>Binary Particle Swarm Optimization</i> dengan <i>Adaptive Inertia Weights</i> (BPSO-AIW) pada Dataset COVID-19	72
Gambar 15 Evaluasi <i>Time Computation</i> Tahap Seleksi Fitur <i>Binary Particle Swarm Optimization</i> dengan <i>Adaptive Inertia Weights</i> (BPSO-AIW) pada Dataset COVID-19	73
Gambar 16 Performa <i>Accuracy</i> Model Seleksi Fitur pada Dataset COVID-19....	74
Gambar 17 Fitur Terpilih <i>Binary Particle Swarm Optimization</i> dengan <i>Adaptive Inertia Weights</i> (BPSO-AIW) pada Dataset <i>Weather</i>	79
Gambar 18 Evaluasi <i>Time Computation</i> Tahap Seleksi Fitur <i>Binary Particle Swarm Optimization</i> dengan <i>Adaptive Inertia Weights</i> (BPSO-AIW) pada Dataset <i>Weather</i>	80
Gambar 19 Performa <i>Accuracy</i> Model Seleksi Fitur pada Dataset <i>Weather</i>	81

Gambar 20 Perbandingan Evaluasi <i>Accuracy</i> , <i>Precision</i> , <i>Recall</i> , dan <i>F1-Score</i> Model Klasifikasi <i>Distance Biased Naive Bayes</i> (DBNB) Tanpa Seleksi Fitur dan Dengan Seleksi Fitur pada Dataset COVID-19	87
Gambar 21 Perbandingan Evaluasi <i>Time Computation</i> Model Klasifikasi <i>Distance Biased Naive Bayes</i> (DBNB) Tanpa Seleksi Fitur dan Dengan Seleksi Fitur pada Dataset COVID-19.....	88
Gambar 22 Perbandingan Evaluasi <i>Accuracy</i> , <i>Precision</i> , <i>Recall</i> , dan <i>F1-Score</i> Model Klasifikasi <i>Distance Biased Naive Bayes</i> (DBNB) Tanpa Seleksi Fitur dan Dengan Seleksi Fitur pada Dataset <i>Weather</i>	94
Gambar 23 Perbandingan Evaluasi <i>Time Computation</i> Model Klasifikasi <i>Distance Biased Naive Bayes</i> (DBNB) Tanpa Seleksi Fitur dan Dengan Seleksi Fitur pada Dataset <i>Weather</i>	94
Gambar 24 Perbandingan Jumlah Fitur Kedua Dataset Sebelum dan Sesudah Seleksi Fitur.....	98
Gambar 25 Perbandingan Evaluasi <i>Time Computation</i> Tahap Seleksi Fitur Kedua Dataset	99

DAFTAR LAMPIRAN

- Lampiran 1 Kamus *Stopword* Bahasa Indonesia
- Lampiran 2 Kamus *Stopword* Bahasa Inggris
- Lampiran 3 *Source Code Import Library*
- Lampiran 4 *Source Code Import Dataset*
- Lampiran 5 *Source Code Preprocessing Data*
- Lampiran 6 *Source Code Ekstraksi Fitur Term Frequency–inverse Document Frequency (TF-IDF)*
- Lampiran 7 *Source Code Pembagian Data Split Validation*
- Lampiran 8 *Source Code Seleksi Fitur Chi-Square (CS)*
- Lampiran 9 *Source Code Seleksi Fitur Binary Particle Swarm Optimization dengan Adaptive Inertia Weights (BPSO-AIW)*
- Lampiran 10 *Source Code Klasifikasi Distance Biased Naïve Bayes (DBNB)*
- Lampiran 11 *Source Code Evaluasi*