

## DAFTAR ISI

Halaman Judul .....	i
Halaman Pengesahan Pembimbing.....	ii
Halaman Pengesahan Penguji.....	iii
Surat Pernyataan .....	iv
Pernyataan Bebas Plagiat.....	v
Halaman Persembahan.....	vi
Abstrak.....	vii
Kata pengantar .....	viii
Daftar Isi .....	ix
Daftar Gambar .....	xi
Daftar Tabel .....	xii
Daftar Modul.....	xiii
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah .....	3
1.4 Tujuan Penelitian .....	3
1.5 Manfaat Penelitian .....	3
1.6 Metodologi Penelitian.....	3
1.6.1 Metodologi Pengumpulan Data.....	3
1.6.2 Metodologi Pengembangan Sistem .....	4
1.7 Sistematika Penulisan .....	5
<b>BAB II TINJAUAN LITERATUR.....</b>	<b>6</b>
2.1 Berita.....	6
2.2 <i>Hoax</i> .....	6
2.3 <i>Text Mining</i> .....	6
2.4 Praproses Dokumen ( <i>Preprocessing</i> ) .....	7
2.4.1 <i>Cleansing</i> .....	7
2.4.2 <i>Case Folding</i> .....	8
2.4.3 <i>Tokenizing</i> .....	8
2.4.4 <i>Stopword Removal</i> .....	9
2.4.5 <i>Stemming</i> .....	10
2.5 Klasifikasi .....	12
2.6 Pembobotan <i>Term Frequency-Inverse Document Frequency</i> (TFIDF).....	13
2.7 <i>K-Nearest Neighbor</i> (KNN).....	14
2.8 <i>Naïve Bayes</i> .....	15
2.9 <i>Confusion Matrix</i> .....	17
2.10 Penelitian Sebelumnya.....	18
<b>BAB III METODOLOGI PENELITIAN DAN PENGEMBANGAN SISTEM.....</b>	<b>23</b>
3.1 Metodologi Penelitian.....	23
3.1.1 Pengumpulan Data .....	23
3.1.2 <i>Labeling</i> .....	25
3.1.3 <i>Data Preprocessing</i> .....	26
3.1.4 <i>Cleansing</i> .....	27
3.1.5 <i>Case Folding</i> .....	28

3.1.6 <i>Tokenizing</i> .....	29
3.1.7 <i>Stopword Removal</i> .....	30
3.1.8 <i>Stemming</i> .....	31
3.1.9 Pembobotan Fitur TFIDF .....	32
3.1.10 Klasifikasi Metode <i>K-Nearest Neighbor</i> (KNN).....	36
3.1.11 Klasifikasi Metode <i>Naïve Bayes Classifier</i> .....	41
3.2 Metodologi Pengembangan Sistem.....	45
3.2.1 Pengumpulan Kebutuhan .....	46
3.2.2 Proses Desain .....	47
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	<b>53</b>
4.1 Hasil Penelitian .....	53
4.1.1 Modul <i>Preprocessing</i> .....	53
4.1.2 Modul TFIDF ( <i>Term Frequency-Inverse Document Frequency</i> ).....	55
4.1.3 Modul KNN.....	56
4.1.4 Modul <i>Naïve Bayes</i> .....	56
4.1.5 Implementasi Aplikasi.....	57
4.2 Pengujian.....	60
4.2.1 <i>K-Fold Cross Validation</i> .....	60
4.2.2 <i>Confusion Matrix</i> .....	60
4.2.3 Grafik Pengujian.....	66
4.2.4 Pengujian Implementasi Sistem .....	67
4.3 Pembahasan.....	69
<b>BAB V PENUTUP</b> .....	<b>71</b>
5.1 Kesimpulan .....	71
5.2 Saran .....	71
<b>DAFTAR PUSTAKA</b> .....	<b>72</b>

## DAFTAR GAMBAR

<b>Gambar 2.1</b> Tahapan <i>Text Mining</i> .....	7
<b>Gambar 2.2</b> Arsitektur Pada Algoritma <i>Naïve Bayes</i> .....	16
<b>Gambar 3.1</b> Metodologi Penelitian .....	23
<b>Gambar 3.2</b> Tampilan Situs covid19.go.id.....	24
<b>Gambar 3.3</b> Proses <i>Web Scraping</i> Menggunakan <i>Data Miner</i> .....	25
<b>Gambar 3.4</b> <i>Flowchart Preprocessing</i> .....	26
<b>Gambar 3.5</b> <i>Flowchart</i> proses <i>cleansing</i> .....	27
<b>Gambar 3.6</b> <i>Flowchart</i> proses <i>Case Folding</i> .....	28
<b>Gambar 3.7</b> <i>Flowchart</i> proses <i>Tokenizing</i> .....	29
<b>Gambar 3.8</b> <i>Flowchart</i> proses <i>Stopword Removal</i> .....	30
<b>Gambar 3.9</b> <i>Gambar Flowchart</i> proses <i>Stemming</i> .....	31
<b>Gambar 3.10</b> <i>Flowchart</i> proses TFIDF .....	33
<b>Gambar 3.11</b> <i>Flowchart</i> klasifikasi metode KNN.....	37
<b>Gambar 3.12</b> <i>Flowchart</i> klasifikasi metode <i>Naïve Bayes</i> .....	41
<b>Gambar 3.13</b> Siklus <i>Prototype</i> .....	45
<b>Gambar 3.14</b> Arsitektur Sistem .....	47
<b>Gambar 3.15</b> Struktur Menu.....	49
<b>Gambar 3.16</b> Rancangan <i>Interface</i> Halaman <i>Home</i> .....	49
<b>Gambar 3.17</b> Rancangan <i>Interface</i> Halaman <i>Prediksi</i> .....	50
<b>Gambar 3.18</b> Rancangan <i>Interface</i> Halaman <i>Dataset</i> .....	50
<b>Gambar 3.19</b> Rancangan <i>Interface</i> Halaman <i>Model KNN-Naïve Bayes</i> .....	51
<b>Gambar 3.20</b> Rancangan <i>Interface</i> Halaman <i>Grafik</i> .....	51
<b>Gambar 4.1</b> Halaman <i>Home</i> .....	58
<b>Gambar 4.2</b> Halaman <i>Prediksi</i> .....	58
<b>Gambar 4.3</b> Halaman <i>Dataset</i> .....	59
<b>Gambar 4.4</b> Halaman <i>Model KNN-Naïve Bayes</i> .....	59
<b>Gambar 4.5</b> Halaman <i>Grafik</i> .....	60
<b>Gambar 4.6</b> Halaman <i>Grafik</i> Pada Metode <i>Naïve Bayes</i> .....	66
<b>Gambar 4.7</b> Halaman <i>Grafik</i> Pada Metode <i>K-Nearest Neighbor</i> .....	67
<b>Gambar 4.8</b> Contoh Pengujian Sistem Dengan Data Uji Baru.....	68

## DAFTAR TABEL

<b>Tabel 2.1</b> Kombinasi Awalan dan Akhiran Tidak Diiijinkan (Nazief dan Adriani, 1996) ....	11
<b>Tabel 2.2</b> Aturan Peluruhan Kata Dasar (Nazief dan Adriani, 1996) .....	11
<b>Tabel 2.3</b> Tabel <i>confusion matrix</i> .....	18
<b>Tabel 2.4</b> Rangkuman Penelitian Terkait.....	20
<b>Tabel 2.5</b> Lanjutan Rangkuman Penelitian Terkait.....	21
<b>Tabel 2.6</b> Perbedaan <i>Preprocessing</i> .....	22
<b>Tabel 2.7</b> Perbedaan Jumlah Kelas dan Nilai K pada KNN.....	22
<b>Tabel 3.1</b> Contoh Data Hasil <i>Labeling</i> .....	25
<b>Tabel 3.2</b> Rincian Dataset .....	26
<b>Tabel 3.3</b> Tabel contoh proses perhitungan TF.....	34
<b>Tabel 3.4</b> Tabel contoh proses perhitungan DF dan IDF .....	34
<b>Tabel 3.5</b> Lanjutan Tabel contoh proses perhitungan DF dan IDF .....	35
<b>Tabel 3.6</b> Tabel Contoh proses perhitungan TFIDF.....	36
<b>Tabel 3.7</b> Hasil Perhitungan Pembobotan TFIDF Dokumen Uji Dengan Dokumen Latih ..	38
<b>Tabel 3.8</b> Hasil Perkalian Pembobotan .....	39
<b>Tabel 3.9</b> Hasil Perhitungan Kuadrat Pembobotan .....	40
<b>Tabel 3.10</b> Hasil Perhitungan Jumlah Total Pembobotan .....	40
<b>Tabel 3.11</b> Hasil Voting Data Uji.....	41
<b>Tabel 3.12</b> Hasil Probabilitas Data Latih Kategori Fakta .....	42
<b>Tabel 3.13</b> Lanjutan Hasil Probabilitas Data Latih Kategori Fakta .....	43
<b>Tabel 3.14</b> Hasil Probabilitas Data Latih Kategori <i>Hoax</i> .....	43
<b>Tabel 3.15</b> Lanjutan Hasil Probabilitas Data Latih Kategori <i>Hoax</i> .....	44
<b>Tabel 3.16</b> Hasil Probabilitas Data Uji.....	44
<b>Tabel 3.17</b> Struktur Tabel Dataset.....	48
<b>Tabel 3.18</b> Struktur Tabel Pengujian.....	48
<b>Tabel 3.19</b> Pengujian <i>confusion matrix</i> .....	52
<b>Tabel 3.20</b> Pengujian <i>K-fold validation</i> .....	52
<b>Tabel 4.1</b> <i>Confusion Matrix K-Nearest Neighbor</i> dengan nilai K=1 dengan <i>fold</i> =1 .....	61
<b>Tabel 4.2</b> <i>Confusion Matrix K-Nearest Neighbor</i> dengan nilai K=1 .....	61
<b>Tabel 4.3</b> Pengujian <i>Cross Validation K-Nearest Neighbor</i> dengan nilai K=1 .....	61
<b>Tabel 4.4</b> <i>Confusion Matrix K-Nearest Neighbor</i> dengan nilai K=3 dengan <i>fold</i> =1 .....	62
<b>Tabel 4.5</b> <i>Confusion Matrix K-Nearest Neighbor</i> dengan nilai K=3 .....	62
<b>Tabel 4.6</b> Pengujian <i>Cross Validation K-Nearest Neighbor</i> dengan nilai K=3 .....	63
<b>Tabel 4.7</b> <i>Confusion Matrix K-Nearest Neighbor</i> dengan nilai K=5 dengan <i>fold</i> =1 .....	63
<b>Tabel 4.8</b> <i>Confusion Matrix K-Nearest Neighbor</i> dengan nilai K=5 .....	63
<b>Tabel 4.9</b> Pengujian <i>Cross Validation K-Nearest Neighbor</i> dengan nilai K=5 .....	64
<b>Tabel 4.10</b> <i>Confusion Matrix K-Nearest Neighbor</i> dengan nilai K=7 dengan <i>fold</i> =1 .....	64
<b>Tabel 4.11</b> <i>Confusion Matrix K-Nearest Neighbor</i> dengan nilai K=7 .....	64
<b>Tabel 4.12</b> Pengujian <i>Cross Validation K-Nearest Neighbor</i> dengan nilai K=7 .....	65
<b>Tabel 4.13</b> <i>Confusion Matrix Naïve Bayes</i> dengan <i>fold</i> =1 .....	65
<b>Tabel 4.14</b> <i>Confusion Matrix Naïve Bayes</i> .....	65
<b>Tabel 4.15</b> Pengujian <i>Cross Validation Naïve Bayes</i> .....	66
<b>Tabel 4.16</b> Pengujian Implementasi Sistem .....	66

## DAFTAR MODUL

<b>Modul 2.1</b> Algoritma <i>Cleansing</i> (Alenzi et.al.,2019) .....	7
<b>Modul 2.2</b> Algoritma <i>Case Folding</i> (Hakim, 2020).....	8
<b>Modul 2.3</b> Algoritma <i>Tokenizing</i> (Alenzi et.al.,2019) .....	9
<b>Modul 2.4</b> Algoritma <i>Stopword Removal</i> (Alenzi et.al.,2019) .....	9
<b>Modul 2.5</b> Algoritma <i>Stemming</i> (Soyusiawaty et.al., 2020) .....	12
<b>Modul 2.6</b> Algoritma TFIDF (Jiang et.al.,2018).....	13
<b>Modul 2.7</b> Algoritma <i>K-Nearest Neighbor</i> (Wandabwa et.al.,2016).....	15
<b>Modul 2.8</b> Algoritma <i>Naïve Bayes</i> (Jiang et.al.,2018) .....	17
<b>Modul 4.1</b> <i>Source Code</i> Proses <i>Cleansing</i> .....	53
<b>Modul 4.2</b> <i>Source Code</i> Proses <i>Case Folding</i> .....	54
<b>Modul 4.3</b> <i>Source Code</i> Proses <i>Tokenizing</i> .....	54
<b>Modul 4.4</b> <i>Source Code</i> Proses <i>Stopword Removal</i> .....	54
<b>Modul 4.5</b> <i>Source Code</i> Proses <i>Stemming</i> .....	55
<b>Modul 4.6</b> <i>Source Code</i> Algoritma TFIDF .....	55
<b>Modul 4.7</b> <i>Source Code</i> Algoritma <i>K-Nearest Neighbor</i> .....	56
<b>Modul 4.8</b> <i>Source Code</i> Algoritma <i>Naïve Bayes</i> .....	57