



UNIVERSITAS
Dinamika

**ANALISIS SENTIMEN *REVIEW* PELANGGAN PADA LAYANAN
INDIHOME MENGGUNAKAN ALGORITMA *NAÏVE BAYES*
*CLASSIFIER***

TUGAS AKHIR



**Program Studi
S1 Sistem Informasi**

UNIVERSITAS
Dinamika

Oleh:

Arganata Alif Fani

15410100053

FAKULTAS TEKNOLOGI DAN INFORMATIKA

UNIVERSITAS DINAMIKA

2020

**ANALISIS SENTIMEN *REVIEW* PELANGGAN PADA LAYANAN
INDIHOME MENGGUNAKAN ALGORITMA *NAÏVE BAYES*
*CLASSIFIER***

TUGAS AKHIR



Diajukan sebagai salah satu syarat untuk menyelesaikan Program Sarjana

**UNIVERSITAS
Dinamika**

Oleh:

Nama : Arganata Alif Fani

NIM : 15410100053

Program Studi : S1 Sistem Informasi

FAKULTAS TEKNOLOGI DAN INFORMATIKA

UNIVERSITAS DINAMIKA

2020

Tugas Akhir
ANALISIS SENTIMEN *REVIEW* PELANGGAN PADA LAYANAN
INDIHOME MENGGUNAKAN ALGORITMA *NAÏVE BAYES*
CLASSIFIER

dipersiapkan dan disusun oleh

Arganata Alif Fani
NIM : 15410100053

Telah diperiksa, diuji dan disetujui oleh Dewan Penguji
Pada : Selasa, 25 Februari 2020

Susunan Dewan Pembahas

Pembimbing:

I. Pantjawati Sudarmaningtyas, S.Kom., M.Eng.
NIDN. 0712066801

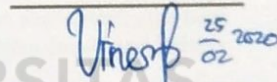
II. Vivine Nurcahyawati, M.Kom.
NIDN. 0723018101

Pembahas:

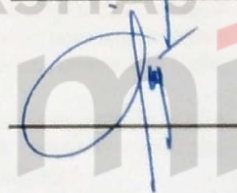
Dr. Jusak
NIDN. 0708017101



25/02/2020



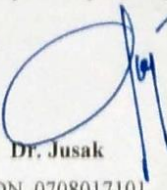
25/02/2020



Tugas Akhir ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana



Fakultas Teknologi dan Informatika
UNIVERSITAS
Dinamika



27/2/20

Dr. Jusak
NIDN. 0708017101

Dekan Fakultas Teknologi dan Informatika
UNIVERSITAS DINAMIKA

PERNYATAAN

PERSETUJUAN PUBLIKASI DAN KEASLIAN KARYA ILMIAH

Sebagai mahasiswa Universitas Dinamika, saya :

Nama : Arganata Alif Fani

NIM : 15410100053

Program Studi : S1 Sistem Informasi

Fakultas : Fakultas Teknologi dan Informatika

Jenis Karya : Tugas Akhir

Judul Karya : **ANALISIS SENTIMEN *REVIEW* PELANGGAN PADA LAYANAN
INDIHOME MENGGUNAKAN ALGORITMA *NAÏVE BAYES*
CLASSIFIER**

Menyatakan dengan sesungguhnya bahwa:

1. Demi pengembangan Ilmu Pengetahuan, Teknologi dan Seni, saya menyetujui memberikan kepada Universitas Dinamika Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalti Free Right*) atas seluruh isi/ sebagian karya ilmiah saya tersebut di atas untuk disimpan, dialihmediakan dan dikelola dalam bentuk pangkalan data (*database*) untuk selanjutnya didistribusikan atau dipublikasikan demi kepentingan akademis dengan tetap mencantumkan nama saya sebagai penulis atau pencipta dan sebagai pemilik Hak Cipta.
2. Karya tersebut di atas adalah karya asli saya, bukan plagiat baik sebagian maupun keseluruhan. Kutipan, karya atau pendapat orang lain yang ada dalam karya ilmiah ini adalah semata hanya rujukan yang dicantumkan dalam Daftar Pustaka saya .
3. Apabila dikemudian hari ditemukan dan terbukti terdapat tindakan plagiat pada karya ilmiah ini, maka saya bersedia untuk menerima pencabutan terhadap gelar kesarjanaan yang telah diberikan kepada saya.

Demikian surat pernyataan ini saya buat dengan sebenarnya.

Surabaya, 25 Februari 2020

Yang menyatakan



Arganata Alif Fani

Nim: 15410100053



“Jangan lupa bahagia 😊”

UNIVERSITAS
Dinamika



“Tugas Akhir ini saya persembahkan untuk semua orang yang selalu nanya kapan
saya wisuda”

UNIVERSITAS
Dinamika

ABSTRAK

PT. Telkom Indonesia sebagai perusahaan yang berorientasi kepada pelanggan (*customer-oriented*) tentunya harus mengerti dan mengetahui betapa pentingnya penilaian pelanggan terhadap layanan, instansi atau organisasi. pada PT. Telkom Indonesia terdapat permasalahan yaitu Telkom Indonesia belum mengetahui tentang pendapat atau *review* pelanggan terhadap layanan IndiHome pada *website* dan media sosial. Saat ini Telkom Indonesia hanya mengetahui *review* atau pendapat pelanggan melalui komplain secara langsung atau melalui telepon *Customer Service (CS)*.

Berdasarkan permasalahan yang ada, maka didapatkan solusi berupa sebuah aplikasi yang dapat menampilkan *review* pelanggan apakah positif atau negatif. *Review* pelanggan dapat diproses kata/kalimatnya dengan menggunakan metode *Text Mining* dan kata/kalimat dapat diklasifikasi kelas positif dan negatif menggunakan metode *Naïve Bayes*. Penelitian ini menggunakan metode *System Development Life Cycle (SDLC)* dalam mengembangkan aplikasi dan untuk menguji akurasi pada penelitian ini menggunakan metode *Confusion Matrix*.

Dari hasil penelitian analisis sentimen dapat disimpulkan bahwa analisis sentimen pada penelitian ini dapat menampilkan visualisasi dari hasil klasifikasi analisis sentimen dalam bentuk diagram lingkaran (*Pie Chart*) dan menghasilkan klasifikasi analisis sentimen yang dibuktikan pada pengujian fungsi *Crawling Data*, Ekstraksi, Klasifikasi dan Visualisasi dengan nilai sebesar 100%, pengujian perhitungan dengan nilai yang berbeda dengan klasifikasi yang sama, dan pengujian akurasi pada data tahun 2017 sampai dengan 2020 dari jumlah data training sebanyak 40 data dan data testing sebanyak 13 data didapatkan nilai sebesar 69%.

Kata Kunci: Analisis Sentimen, Telkom Indonesia, Naïve Bayes, Text Mining

KATA PENGANTAR

Segala puji bagi Tuhan yang Maha Esa, Allah SWT atas segala rahmat, karunia dan hidayah-Nya, sehingga dapat menyelesaikan Laporan Tugas Akhir dengan judul “**Analisis Sentimen *Review* Pelanggan Pada Layanan Indihome Menggunakan Algoritma *Naïve Bayes Classifier***” dengan baik meskipun penulis sadari bahwa masih ada banyak kekurangan yang ada didalamnya, karena kesempurnaan hanya milik Tuhan semata.

Tugas Akhir ini adalah salah satu syarat untuk menyelesaikan Program Studi Strata Satu (S1) pada jurusan Sistem Informasi Universitas Dinamika Surabaya. Laporan ini disusun sebagai bukti bahwa pengerjaan Tugas Akhir telah diselesaikan.

Dalam penyelesaian Tugas Akhir ini tentu mendapatkan dukungan dari berbagai pihak dan secara pribadi penulis mengucapkan banyak terima kasih kepada:

1. Tuhan yang Maha Esa, Allah SWT karena selalu mendengarkan doa yang penulis selalu panjatkan.
2. Mama yang selalu mendoakan, memberi kasih sayang, dan mengajarkan kesabaran.
3. Ibu Pantjawati Sudarmaningtyas, S.Kom., M.Eng. dan Ibu Vivine Nurcahyawati, M.Kom., OCP selaku dosen pembimbing yang selalu sabar dalam memberikan bimbingan dan mendukung dalam menyelesaikan Tugas Akhir.
4. Ibu Jayanti Sukma Maulani selaku Penyelia pada PT. Telekomunikasi Indonesia Regional V Surabaya yang memberikan masukan terhadap penelitian yang dilakukan.
5. Bapak Dr. Jusak selaku dosen pembahas topik penelitian Tugas Akhir.
6. Dirga Ambara Selaku teman yang sabar mengajari, bisa menyempatkan waktu untuk membantu dalam mengerjakan Tugas Akhir.
7. Kepada Sahabat Soponyono, Teman OKK Bill Gates 1.4. angkatan 2015, teman-temanku satu bimbingan Tugas Akhir, Sahabat Om Karyo, Teman SMK, Teman SMP, Teman SD, Teman TK, Night Workout Club Surabaya

dan semua pihak yang tidak dapat disebutkan satu-persatu, terima kasih banyak.

8. Untuk wanita yang harusnya kutulis Namanya disini, semoga kamu berbahagia dengan senyumanmu yang hangat dan lucu.

Surabaya, 25 Februari 2020

Penulis

Arganata Alif Fani



UNIVERSITAS
Dinamika

DAFTAR ISI

	Halaman
ABSTRAK	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN	xvi
BAB 1. PENDAHULUAN	1
1.1. Latar Belakang Penelitian	1
1.2. Rumusan Masalah Penelitian	2
1.3. Batasan Masalah Penelitian	3
1.4. Tujuan Penelitian	3
1.5. Manfaat Penelitian	3
1.6. Sistematika Penulisan Penelitian	4
BAB 2. LANDASAN TEORI	5
2.1. Analisis Sentimen	5
2.2. <i>Web Crawler</i>	6
2.3. <i>Text Mining</i>	6
2.3.1. <i>Preprocessing Task</i>	6
2.4. <i>Naïve Bayes</i>	8
2.5. Visualisasi	9
2.6. <i>System Development Life Cycle (SDLC)</i>	10
2.7. <i>Confusion Matrix</i>	10
2.7.1. Akurasi	11
BAB 3. METODE PENELITIAN	12
3.1. Tahap Awal	13
3.1.1. Pengumpulan Data	13
3.1.2. Ruang Lingkup	15
3.1.3. Jadwal Kerja	15
3.2. Tahap Pengembangan	16
3.2.1. Analisis	16

	Halaman
3.2.2. Desain.....	20
3.2.3. Pengkodean.....	25
3.2.4. Pengujian.....	27
3.3. Tahap Akhir	27
3.3.1. Evaluasi Uji Fungsi.....	27
3.3.2. Evaluasi Uji Perhitungan.....	28
3.3.3. Evaluasi Uji Akurasi.....	28
BAB 4. HASIL DAN PEMBAHASAN	29
4.1. Implemenstasi	29
4.1.1. Halaman <i>Login</i>	29
4.1.2. Halaman Utama.....	29
4.1.3. Halaman Hasil Data <i>Crawling</i>	30
4.1.4. Halaman Hasil Ekstraksi Data (<i>Case Folding</i>).....	31
4.1.5. Halaman Hasil Ekstraksi (<i>Tokenizing</i>)	31
4.1.6. Halaman Hasil Ekstraksi (<i>Filtering</i>).....	31
4.1.7. Halaman Hasil Ekstraksi (<i>Stemming</i>)	32
4.1.8. Halaman Hasil Klasifikasi Data	32
4.2. Pengujian.....	33
4.2.1. Pengujian Fungsi.....	34
4.2.2. Pengujian Perhitungan.....	35
4.2.3. Pengujian Akurasi	45
4.3. Evaluasi.....	45
4.3.1. Evaluasi Hasil Uji Fungsi.....	46
4.3.2. Evaluasi Hasil Uji Hitungan	46
4.3.3. Evaluasi Hasil Uji Akurasi.....	47
BAB 5. PENUTUP	48
5.1. Kesimpulan	48
5.2. Saran	48
DAFTAR PUSTAKA	49
LAMPIRAN	52

DAFTAR GAMBAR

	Halaman
Gambar 2.1. Proses <i>System Development Life Cycle</i> (SDLC).....	10
Gambar 2.2. <i>Confusion Matrix</i>	11
Gambar 3.1. Metode Penelitian.....	12
Gambar 3.2. Tahap Awal	13
Gambar 3.3. Tahap Pengembangan	16
Gambar 3.25. Tahap Akhir.....	27
Gambar L.3.4. <i>System Flow</i> Penambangan Data	58
Gambar L.3.5. <i>System Flow</i> Ekstraksi Data.....	58
Gambar L.3.6. <i>System Flow</i> Klasifikasi Data	59
Gambar L.3.7. <i>Flowchart</i> Kamus Liu.....	59
Gambar L.3.8. <i>Flowchart</i> <i>Naïve Bayes</i>	60
Gambar L.3.9. <i>Conceptual Data Model</i> (CDM).....	60
Gambar L.3.10. <i>Physical Data Model</i> (PDM)	61
Gambar L.3.11. Desain <i>Interface Login</i>	63
Gambar L.3.12. Desain <i>Interface</i> Halaman Utama.....	63
Gambar L.3.13. Desain <i>Interface</i> Halaman Data Hasil <i>Crawling</i>	63
Gambar L.3.14. Desain <i>Interface</i> Data Hasil Ekstraksi (<i>Case Folding</i>).....	64
Gambar L.3.15. Desain <i>Interface</i> Data Hasil Ekstraksi (<i>Tokenizing</i>).....	64
Gambar L.3.16. Desain <i>Interface</i> Data Hasil Ekstraksi (<i>Filtering</i>).....	64
Gambar L.3.17. Desain <i>Interface</i> Data Hasil Ekstraksi (<i>Stemming</i>).....	65
Gambar L.3.18. Desain <i>Interface</i> Data Klasifikasi	65
Gambar L.3.19. Kode Penambangan Data.....	66
Gambar L.3.20. Kode Ekstraksi Data <i>Case Folding</i>	66
Gambar L.3.21. Kode Ekstraksi Data <i>Tokenizing</i>	66
Gambar L.3.22. Kode Ekstraksi Data <i>Filtering</i>	66
Gambar L.3.23. Kode Ekstraksi Data <i>Stemming</i>	67
Gambar L.3.24. Kode Klasifikasi Data	67
Gambar L.4.1. Halaman <i>Login</i>	69
Gambar L.4.2. Halaman Utama	69
Gambar L.4.3. Halaman <i>Pop Up</i> Penambangan Data.....	69

	Halaman
Gambar L.4.4. Halaman <i>Pop Up</i> Ekstraksi Data	70
Gambar L.4.5. Halaman <i>Pop Up</i> Ekstraksi Data	70
Gambar L.4.6. Halaman Hasil Data <i>Crawling</i>	70
Gambar L.4.7. Halaman <i>Pop Up</i> Detail Hasil Data <i>Crawling</i>	71
Gambar L.4.8. Halaman Hasil Ekstraksi Data (<i>Case Folding</i>).....	71
Gambar L.4.9. Halaman Detail Hasil Ekstraksi Data (<i>Case Folding</i>).....	71
Gambar L.4.10. Halaman Hasil Ekstraksi Data (<i>Tokenizing</i>).....	72
Gambar L.4.11. Halaman Detail Hasil Ekstraksi Data (<i>Tokenizing</i>).....	72
Gambar L.4.12. Halaman Hasil Ekstraksi Data (<i>Filtering</i>)	72
Gambar L.4.13. Halaman Detail Hasil Ekstraksi Data (<i>Filtering</i>)	73
Gambar L.4.14. Halaman Hasil Ekstraksi Data (<i>Stemming</i>).....	73
Gambar L.4.15. Halaman Detail Hasil Ekstraksi Data (<i>Stemming</i>).....	73
Gambar L.4.16. Halaman Hasil Klasifikasi Data.....	74
Gambar L.4.17. Halaman Perhitungan Klasifikasi Data.....	74
Gambar L.4.18. Halaman Hasil Perhitungan Klasifikasi Data	75
Gambar L.4.19. Halaman Detail Klasifikasi Data	75
Gambar L.4.20. Tabel Data Master.....	81
Gambar L.4.21. Tabel Data <i>Training</i> Positif	81
Gambar L.4.22. Tabel Data <i>Training</i> Negatif.....	81
Gambar L.4.23. Hasil Perhitungan <i>Testing</i> mediakonsumen.com.....	82
Gambar L.4.24. Hasil Perhitungan <i>Testing</i> twitter.com.....	82
Gambar L.4.25. Hasil Perhitungan <i>Testing</i> detik.com	82
Gambar L.4.26. <i>Confusion Matrix</i> Data <i>Testing</i>	104

DAFTAR TABEL

	Halaman
Tabel L.2.1. Proses <i>Case Folding</i>	52
Tabel L.2.2. Proses <i>Tokenizing</i>	52
Tabel L.2.3. Proses <i>Filtering</i>	52
Tabel L.2.4. Proses <i>Stemming</i>	53
Tabel 3.1. Jadwal Kerja.....	15
Tabel 3.14. Desain <i>Interface</i>	23
Tabel L.3.2. Data Observasi.....	54
Tabel L.3.3. Identifikasi Permasalahan.....	56
Tabel L.3.4. Analisis Kebutuhan Pengguna.....	56
Tabel L.3.5. Analisis Kebutuhan Data dan Informasi.....	56
Tabel L.3.6. Analisis Kebutuhan Fungsional.....	56
Tabel L.3.7. Analisis Kebutuhan Non Fungsional.....	57
Tabel L.3.8. Struktur Basis Data <i>data_crawling</i>	61
Tabel L.3.9. Struktur Basis Data <i>case_folding</i>	61
Tabel L.3.10. Struktur Basis Data <i>tokenize</i>	62
Tabel L.3.11. Struktur Basis Data <i>stopword</i>	62
Tabel L.3.12. Struktur Basis Data <i>stemmer</i>	62
Tabel L.3.13. Struktur Basis Data klasifikasi	62
Tabel L.3.15. Uji Coba Penambangan Data.....	68
Tabel L.3.16. Uji Coba Ekstraksi Data	68
Tabel L.3.17. Uji Coba Klasifikasi Data.....	68
Tabel L.3.18. Uji Coba Visualisasi Data.....	68
Tabel L.4.1. Pengujian Fungsi <i>Crawling</i> Data (mediakonsumen.com).....	76
Tabel L.4.2. Pengujian Fungsi <i>Crawling</i> Data (twitter.com).....	76
Tabel L.4.3. Pengujian Fungsi <i>Crawling</i> Data (detik.com)	76
Tabel L.4.4. Pengujian Fungsi Ekstraksi Data (mediakonsumen.com)	77
Tabel L.4.5. Pengujian Fungsi Ekstraksi Data (twitter.com).....	77
Tabel L.4.6. Pengujian Fungsi Ekstraksi Data (detik.com)	78
Tabel L.4.7. Pengujian Fungsi Klasifikasi Data (mediakonsumen.com).....	79
Tabel L.4.8. Pengujian Fungsi Klasifikasi Data (twitter.com)	79

	Halaman
Tabel L.4.9. Pengujian Fungsi Klasifikasi Data (detik.com).....	79
Tabel L.4.10. Pengujian Fungsi Visualisasi Data	80
Tabel L.4.11. Tabel Data <i>Training</i> Dan Data <i>Testing</i>	82
Tabel L.4.12. Tabel Data Master	86
Tabel L.4.13. Tabel Data <i>Training</i> Positif	93
Tabel L.4.14. Tabel Data <i>Training</i> Negatif.....	94
Tabel L.4.15. Perhitungan Data <i>Training</i> Positif dan Negatif.....	99
Tabel L.4.16. Tabel Data Pengujian Akurasi Data <i>Testing</i>	102
Tabel L.4.17. Tabel Data Perbandingan Hasil Uji Aplikasi Dan Uji Manual	105



UNIVERSITAS
Dinamika

DAFTAR LAMPIRAN

	Halaman
LAMPIRAN 1. <i>PREPROCESSING TASK</i>	52
LAMPIRAN 2. PENGUMPULAN DATA.....	54
LAMPIRAN 3. ANALISIS.....	56
LAMPIRAN 4. DESAIN	58
LAMPIRAN 5. PENGKODEAN.....	66
LAMPIRAN 6. PENGUJIAN	68
LAMPIRAN 7. IMPLEMENTASI	69
LAMPIRAN 8. PENGUJIAN FUNGSI	76
LAMPIRAN 9. PENGUJIAN PERHITUNGAN	81
LAMPIRAN 10. PENGUJIAN AKURASI.....	102
LAMPIRAN 11. EVALUASI HASIL UJI HITUNGAN	105



UNIVERSITAS
Dinamika

BAB I

PENDAHULUAN

Bab pertama yang ditulis pada penelitian ini adalah pendahuluan. Pada bab ini menjelaskan mengenai gambaran dan pengantar dari penelitian yang dilakukan. Bab ini berisikan 6 (enam) subbab, yaitu Latar Belakang Penelitian, Rumusan Masalah Penelitian, Batasan Masalah Penelitian, Tujuan Penelitian, Manfaat Penelitian, dan Sistematika Penulisan Penelitian.

1.1. Latar Belakang Penelitian

PT. Telekomunikasi Indonesia (Persero) Tbk, (Telkom Indonesia) memiliki layanan yang paling banyak digunakan yaitu IndiHome. Dikutip dari *website* IndiHome (2019) IndiHome adalah layanan digital berupa internet rumah, telepon rumah dan TV interaktif (*UseeTV*) dengan teknologi fiber optik. Dikutip dari *website detik.com* (2019) tercatat ada 5 juta pelanggan hingga awal desember 2018.

Saat ini Telkom Indonesia mengetahui *review* atau pendapat pelanggan melalui komplain secara langsung atau melalui telepon *Customer Service* (CS). Sedangkan, pada Telkom Indonesia terdapat permasalahan yaitu Telkom Indonesia belum mengetahui tentang pendapat atau *review* pelanggan terhadap layanan IndiHome pada *website* dan media sosial. Menurut situs layanan manajemen konten (*content management*) *HootSuite*, data tren internet dan media sosial 2019 di Indonesia sebagai berikut: 1.) Total Populasi (jumlah penduduk): 268,2 juta (naik 1% dari tahun 2018), 2.) Pengguna *Mobile* (*Smartphone*): 355,5 juta (naik 50 % dari tahun 2018), 3.) Pengguna Internet: 150 juta (naik 12% dari tahun 2018), 4.) Pengguna Media Sosial Aktif: 150 juta (naik 13% dari tahun 2018), 5.) Pengguna

Media Sosial *Mobile*: 130 juta (naik 8% dari tahun 2018). Menurut *HootSuite* 4 platform media sosial yang digunakan di Indonesia dengan angka di atas 80% sebagai berikut: 1. *youtube*: 88% (naik 45% dari tahun 2018), 2. *whatsapp*: 83% (naik 43% dari tahun 2018), 3. *facebook*: 81% (naik 40% dari tahun 2018), 4. *instagram*: 80% (naik 42% dari tahun 2018).

Untuk mengoptimalkan sumber data mengenai pendapat pelanggan, Telkom Indonesia bisa mendapatkannya melalui internet dan media sosial. Internet dan media sosial adalah media dimana pelanggan dapat mengungkapkan pendapatnya mengenai layanan dan instansi yang digunakan secara jujur. Berdasarkan latar belakang dan temuan hasil dari aktivitas observasi dan wawancara didapatkan bahwa dibutuhkan suatu aplikasi untuk dapat melakukan analisis sentimen pada *review* atau pendapat pelanggan yang tertulis pada media sosial maupun *website*. Analisis sentimen yang dilakukan akan menghasilkan sebuah informasi tentang bagaimana *review* pelanggan terhadap layanan perusahaan. Visualisasi dari hasil analisis sentimen akan dituangkan dalam bentuk grafik pada *website* sehingga pihak perusahaan dapat membaca hasil analisis sentimen yang telah dilakukan.

1.2. Rumusan Masalah Penelitian

Berdasarkan permasalahan di atas maka dapat dirumuskan bahwa masalah yang ada yaitu, bagaimana melakukan analisis sentimen pada *review* pelanggan terhadap layanan IndiHome pada Telkom Indonesia.

1.3. Batasan Masalah Penelitian

Batasan masalah yang digunakan dalam melakukan analisis sentimen adalah:

1. Data sekunder yang digunakan dalam penelitian ini adalah media sosial (*twitter.com*) dan *website* (*mediakonsumen.com* dan *detik.com*).
2. Data pada penelitian ini menggunakan *dataset* pada awal bulan januari tahun 2017 sampai akhir bulan mei tahun 2019, kemudian data simulasi dibagi menjadi dua bagian yaitu data *training* dan data *testing*.
3. Kamus kata-kata positif dan menggunakan kamus Liu (2012).
4. Dalam melakukan ekstraksi data, penelitian ini hanya menggunakan tahapan *preprocessing task* pada *text mining*.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah menghasilkan sebuah klasifikasi analisis sentimen pada *review* pelanggan terhadap layanan IndiHome pada Telkom Indonesia dan visualisasi dari hasil analisis sentimen dapat ditampilkan dalam bentuk grafik pada aplikasi *website*, sehingga hasil dari analisis sentimen yang telah dilakukan dapat dibaca oleh pihak perusahaan.

1.5. Manfaat Penelitian

Berdasarkan analisis sentimen yang dilakukan, diharapkan hasil analisis sentimen dapat digunakan sebagai tolak ukur dalam melakukan evaluasi pada strategi bisnis dan kegiatan operasional perusahaan yang sedang berjalan serta Telkom Indonesia dapat mengetahui bagaimana *review* atau pendapat pelanggan terhadap layanan IndiHome yang digunakan saat ini.

1.6. Sistematika Penulisan Penelitian

Pada Laporan Tugas Akhir ditulis berdasarkan dengan sistematika pada setiap bagian penulisannya sebagai berikut:

BAB I: PENDAHULUAN

Pada bab ini membahas tentang pendahuluan berupa Latar Belakang Penelitian, Rumusan Masalah Penelitian, Batasan Masalah Penelitian, Tujuan Penelitian, Manfaat Penelitian, Dan Sistematika Penulisan.

BAB II: LANDASAN TEORI

Pada bab ini menjelaskan tentang landasan teori apa saja yang sudah dipelajari dan digunakan dalam penelitian ini, teori-teori yang digunakan yaitu Analisis Sentimen, *Web Crawler*, *Text Mining*, *Naïve Bayes*, Visualisasi, tentang Telkom Indonesia, *System Development Life Cycle (SDLC)*, dan *Confusion Matrix*.

BAB III: METODOLOGI PENELITIAN

Pada bab ini membahas tentang metodologi penelitian berupa tahapan-tahapan apa saja yang akan digunakan dan diimplementasikan pada penelitian yang dilakukan.

BAB IV: HASIL DAN PEMBAHASAN

Pada bab ini membahas tentang deskripsi hasil dan pembahasan dari penelitian yang dilakukan berupa implementasi, pengujian, dan evaluasi yang dilakukan.

BAB V: PENUTUP

Pada bab ini membahas tentang penutup berupa kesimpulan dan saran dari hasil penelitian yang sudah dilakukan.

BAB II

LANDASAN TEORI

Bab kedua yang ditulis pada penelitian ini adalah landasan teori. Pada bab ini menjelaskan mengenai literatur apa saja yang digunakan pada penelitian ini. Bab ini berisikan 8 (delapan) subbab, yaitu Analisis Sentimen, *Web Crawler*, *Text Mining*, *Naïve Bayes*, Visualisasi, Tentang Telkom Indonesia, *System Development Life Cycle (SDLC)*, dan *Confusion Matrix*.

2.1. Analisis Sentimen

Menurut Liu (2012), Analisis sentimen adalah bidang studi dengan melakukan analisis pada opini, sentimen, evaluasi, penilaian, sikap, dan emosi orang terhadap produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atributnya. Analisis sentimen memiliki banyak nama penelitian seperti penambangan opini, ekstraksi pendapat, penambangan sentimen, analisis subjektivitas, analisis pengaruh, analisis emosi, penambangan review, dll.

Sedangkan, menurut Baj-Rogowska (2017), analisis sentimen adalah suatu arahan dalam melakukan analisis kualitatif dengan mengidentifikasi pendapat emosional yang disuarakan oleh pelanggan atau orang. Informasi yang diidentifikasi dengan cara ini memiliki nilai bisnis yang signifikan dan dasar dalam pengambilan keputusan. Sedangkan, menurut Saragih (2017), analisis sentimen adalah studi tentang pendapat, perasaan, dan emosi yang diekspresikan dalam tulisan. Tujuan utama analisis sentimen adalah mengekstraksi atribut pada dokumen atau teks sehingga pendapat tersebut dapat dikategorikan dalam kelas tertentu.

2.2. *Web Crawler*

Menurut Shrivastava (2018), *Web crawler* (*spider* atau *robot website*) adalah program yang otomatis melintasi sejumlah halaman *website* dengan mengikuti acuan (*hyperlink* atau pranala), menentukan urutan unit-unit atau bagian-bagian dari kata (mengindeks), dan menyimpan sumber halaman *website* yang dilalui. Data yang tersedia pada *website* berupa data tidak terstruktur dan terstruktur. Mesin pencarian yang berbeda-beda menggunakan berbagai jenis *web crawler* seperti Google menggunakan GoogleBot, *web crawler* Microsoft menggunakan Msnbot/Bingbot, dan Baidu menggunakan Baidu Spider.

2.3. *Text Mining*

Sedangkan, Menurut Feldman & Sanger (2007), *Text mining* adalah interaksi antara pengguna dengan kumpulan dokumen dari waktu ke waktu dengan menggunakan alat atau metode dalam melakukan analisis. Dengan cara yang serupa dengan *data mining*, *text mining* melakukan ekstraksi informasi yang berguna dari sumber data melalui identifikasi dan eksplorasi pola yang menarik. Dalam kasus *text mining*, sumber data adalah kumpulan dokumen dan pola yang menarik serta tidak ditemukan diantara catatan basis data secara umum, tetapi dalam dokumen dengan data tesktual yang tidak terstruktur. Salah satu langkah yang dilakukan dalam *text mining* adalah *preprocessing task*.

2.3.1. *Preprocessing Task*

Menurut Feldman & Sanger (2007), *Preprocessing task* atau *preprocessing text* adalah proses yang diperlukan untuk menyiapkan data dalam penemuan kata inti pada sistem *text mining*. *Preprocessing task* pada umumnya mengubah

informasi dari masing-masing sumber data asli ke dalam format kata dasar dengan menerapkan berbagai jenis metode fitur ekstraksi. Sedangkan, menurut Kadhim, Cheah, & Ahamed (2014), *Preprocessing task* setelah membaca masukan dokumen dalam bentuk teks, *preprocessing task* membagi dokumen dalam fitur yang disebut pemotongan kata (*tokenization*), menghilangkan kata noninformatif (*stop word*) dan fitur yang tersisa adalah mengembalikan kata menjadi kata dasar (*stemming*).

A. *Case Folding*

Menurut Langgeni (2010), *Case folding* adalah proses mengubah huruf yang ada di dalam dokumen menjadi *lowercase* (huruf kecil). *Case folding* hanya menerima huruf 'a' sampai 'z'. Contoh proses *case folding* terdapat pada tabel

L.2.1. Proses *Case Folding* Lampiran 1. Landasan Teori.

B. *Tokenizing*

Menurut Weiss, Indurkha, Zhang, & Damerou (2005), *Tokenization* (*Tokenizing* atau *Parsing*) adalah proses pemotongan dokumen atau kalimat menjadi kata dan menghilangkan delimiter atau pemisah seperti tanda titik (.), koma (,), spasi, angka dan karakter lain yang ada pada kalimat. Contoh proses *tokenizing* terdapat pada tabel L.2.2. Proses *Tokenizing* Lampiran 1. Landasan Teori.

C. *Filtering*

Menurut Srividhya (2010), *filtering (Stopword Removal)* adalah menghilangkan kata tidak berguna yang ada pada dokumen atau kalimat. Kata yang tidak berguna seperti kata ganti, kata preposisi, dan kata sambung. Menurut Weiss, Indurkha, Zhang, & Damerou (2009), *stop list* atau *stop words* atau *stop word removal* merupakan Adalah proses menghilangkan kata yang tidak memiliki ciri (kata unik) pada dokumen. Contoh proses *filtering* terdapat pada tabel L.2.3. Proses *Tokenizing* Lampiran 1. Landasan Teori

D. *Stemming*

Menurut Gurusamy (2014), *stemming* (*Lemmatizing*) adalah proses mengembalikan kata menjadi kata dasar. Contoh proses *stemming* seperti kata "mempresentasikan", "dipresentasikan", dan "berpresentasi" diubah menjadi kata dasar "presentasi". Contoh proses *stemming* terdapat pada tabel L.2.4. Proses *Stemming* Lampiran 1. Landasan Teori.

2.4. *Naïve Bayes*

Menurut Saleh (2015), *naïve bayes* adalah algoritma klasifikasi probabilistik yang menghitung sekumpulan probabilitas dengan menjumlahkan kelas dan kombinasi nilai dari data training. Teorema *bayes* mengasumsikan semua kata berdiri sendiri dan tidak bergantung oleh variabel kelas. Berikut adalah rumus dari

naïve bayes:

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (1)$$

Dimana:

X : Data dengan *class* yang belum diketahui

H : Hipotesis data merupakan suatu *class* spesifik

$P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (*posteriori* probabilitas)

$P(H)$: Probabilitas hipotesis H (*prior* probabilitas)

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: Probabilitas X

Pada penelitian ini, algoritma *naïve bayes* disesuaikan dengan kelas apa yang cocok untuk data yang akan diklasifikasi. Berikut adalah rumus *naïve bayes* yang sudah disesuaikan dengan penelitian ini:

$$P(Wk|Ci) = \frac{n_k + 1}{n + |Vocabulary|} \quad (2)$$

Dimana:

$P(Wk)$: Probabilitas (peluang) kemunculan kata

Ci : Kategori kelas

$P(Wk|Ci)$: Probabilitas kemunculan kata (Wk) pada suatu dokumen dengan kategori kelas (Ci)

n_k : nilai kemunculan kata pada kategori c_i

n : Jumlah keseluruhan kata pada kategori c_i

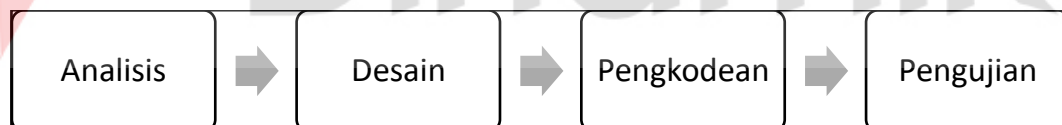
$Vocabulary$: Jumlah keseluruhan kata

2.5. Visualisasi

Menurut Kosara (2007), Visualisasi dibagi menjadi 3 (tiga) kriteria, yaitu visualisasi didasarkan pada (non-visual) data, visualisasi menghasilkan gambar, dan visualisasi yang hasilnya mudah dibaca dan dikenali. visualisasi didasarkan pada (non-visual) data yaitu data yang divisualisasikan berdasarkan *dataset* yang berbeda. visualisasi menghasilkan gambar yaitu menampilkan satu atau lebih gambar dari data dan visualisasi harus menjadi sarana utama untuk mengomunikasikan data. visualisasi yang hasilnya mudah dibaca dan dikenali yaitu visualisasi data yang menjadi gambar dapat dikenali dan dibaca oleh orang lain dengan pelatihan dan pelatihan.

2.6. *System Development Life Cycle (SDLC)*

Menurut Sukamto dan Shalahuddin (2013) *Software Development Life Cycle (SDLC)* atau *System Development Life Cycle* adalah proses mengembangkan sistem perangkat lunak dengan menggunakan model-model pengembangan. Pengembangan sistem dikerjakan secara berurutan mulai dari analisis, desain, pengkodean, dan pengujian. Tahapan dari SDLC terdapat pada gambar 2.1. Proses *System Development Life Cycle (SDLC)*. Tahapan SDLC meliputi Analisis, Desain, Pengkodean, dan Pengujian. Tahap analisis melakukan analisis apa saja kebutuhan dari sistem, mulai dari kebutuhan fungsional maupun kebutuhan non fungsional. Tahap Desain menyajikan desain-desain dari aplikasi seperti desain antar muka, dan desain basis data yang akan diterapkan kedalam aplikasi. Tahap Pengkodean menerapkan desain basis data serta desain antar muka ke dalam aplikasi *website*. Tahap Pengujian merupakan tahap akhir dimana dalam tahap pengujian ini dilakukan pengujian pada fungsi aplikasi, perhitungan, dan akurasi.



Gambar 2.1. Proses *System Development Life Cycle (SDLC)*
(Sumber: SDLC Sukamto dan Salahudin)

2.7. *Confusion Matrix*

Menurut Manning, Raghavan, & Schütze (2009) *confusion matrix* atau disebut juga matriks klasifikasi adalah suatu metode perhitungan yang biasa digunakan dalam *supervised learning*. Dikutip dari *website Toward Data Science* (2019) *confusion matrix* adalah rangkuman perhitungan prediksi dalam menyelesaikan klasifikasi. Prediksi dengan nilai yang benar dan nilai yang salah

dirangkum dalam sebuah tabel (matriks) dengan nilai dan dibedakan berdasarkan kelas masing-masing. Sebagai contoh, pada penelitian terdapat *dataset* sebanyak 100 pasien rumah sakit di mana 5 menderita diabetes dan 95 sehat. Didapatkan nilai akurasi dari klasifikasi 100 pasien rumah sakit dengan nilai 95%. Gambar *confusion matrix* terdapat pada gambar 2.2. *Confusion Matrix*.

		<i>Predictive Value</i>	
		<i>Positive (1)</i>	<i>Negative (0)</i>
<i>Actual values</i>	<i>Positive (1)</i>	<i>TP</i>	<i>FN</i>
	<i>Negative (0)</i>	<i>FP</i>	<i>TN</i>

Gambar 2.2. Confusion Matrix

(Sumber: *confusion matrix website Toward Data Science*)

Dimana:

True Positive (TP): Ketika prediksi positif dan ternyata hasilnya benar positif.

True Negative (TN): Ketika prediksi negatif dan ternyata hasilnya benar negatif.

False Positive (FP): Ketika prediksi positif dan ternyata hasilnya salah.

False Negative (FN): ketika prediksi negatif dan hasilnya ternyata salah.

2.7.1. Akurasi

Akurasi digunakan untuk mengetahui nilai akurasi dari hasil klasifikasi pada penelitian. Berikut adalah rumus dari akurasi *confusion matrix*:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

atau

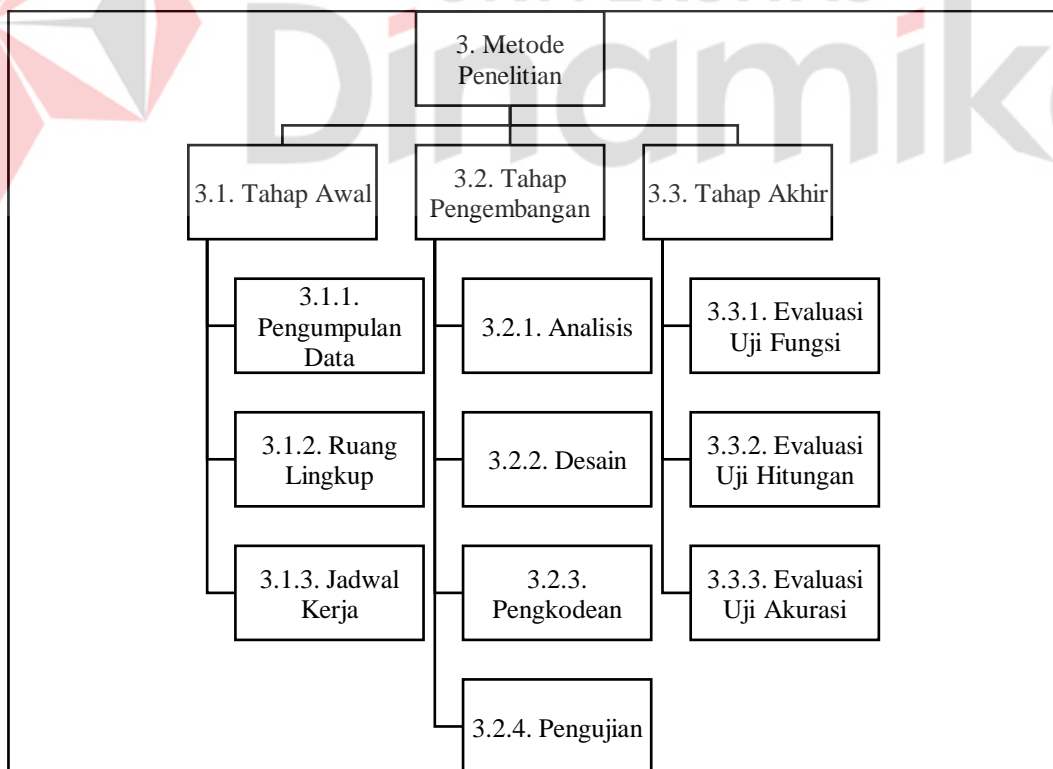
(3)

$$Akurasi = \frac{\text{Jumlah Prediksi Yang Hasilnya Benar}}{\text{Total Keseluruhan Jumlah Prediksi}}$$

BAB III

METODE PENELITIAN

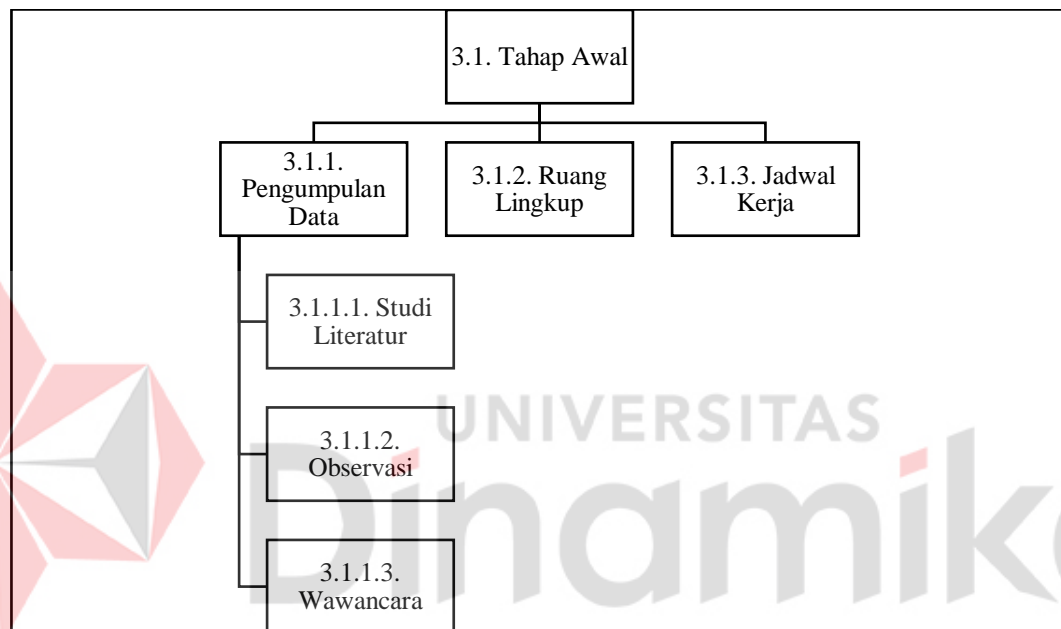
Pada bab ini menjelaskan tentang tahapan-tahapan yang digunakan untuk menyelesaikan penelitian. Menurut wahono (2019) tahapannya terdiri atas tahap awal, tahap pengembangan dan tahap akhir. Pada penelitian ini dilakukan 3 tahapan, yaitu Tahap Awal, Tahap Pengembangan, dan Tahap Akhir. Tahap Awal meliputi 3 (tiga) subbab, yaitu Pengumpulan Data, Ruang Lingkup, Dan Jadwal Kerja. Selanjutnya, Tahap Pengembangan meliputi 4 (empat) subbab, yaitu Analisis, Desain, Pengkodean, dan Pengujian. Selanjutnta, Tahap Akhir meliputi 3 (tiga) subbab, yaitu Evaluasi Uji Fungsi, Evaluasi Uji Hitungan, dan Uji Akurasi. Gambaran dari metodologi penelitian terdapat pada gambar 3.1. Metode Penelitian.



Gambar 3.1. Metode Penelitian
(Sumber: Olahan Penulis)

3.1. Tahap Awal

Bagian ini menjelaskan tentang Tahap Awal dalam menyelesaikan penelitian. Tahap Awal meliputi 3 (tiga) subbab, yaitu Pengumpulan Data, Ruang Lingkup, dan Jadwal Kerja. Pada proses Pengumpulan Data meliputi beberapa bagian yaitu Studi Literatur, Observasi, Dan Wawancara. Untuk gambaran tahap awal terdapat pada gambar 3.2. Tahap Awal.



Gambar 3.2. Tahap Awal
(Sumber: Olahan Penulis)

3.1.1. Pengumpulan Data

Selanjutnya pada tahap awal metode penelitian, yang dilakukan adalah melakukan pengumpulan data. Pada pengumpulan data, yang dilakukan adalah mengumpulkan data yang ada demi memperkuat alasan mengapa penelitian yang dilakukan sangat penting. Pengumpulan data diambil pada media sosial (*twitter.com*) dan *website* (*detik.com* dan *mediakonsumen.com*). Pengumpulan data

dilakukan dalam penelitian ini selain melakukan observasi dan wawancara adalah melakukan studi literatur.

A. Studi Literatur

Pada tahap awal metode penelitian, hal yang pertama kali dilakukan adalah melakukan studi literatur. Pada studi literatur, yang dilakukan adalah *review* terhadap penelitian-penelitian sebelumnya dalam bentuk Tugas Akhir (TA), Jurnal, dsb. Setelah melakukan *review* pada penelitian yang sudah ada, penelitian ini dapat ditentukan studi kasus dan tahapan-tahapan penyelesaian yang digunakan untuk melakukan penelitian.

B. Observasi

Observasi dilakukan untuk mengamati objek yang digunakan sebagai sumber data penelitian. Objek yang diamati adalah *tweet* atau *reply* pada *twitter.com* dan halaman berita pada *detik.com* dan *mediakonsumen.com*. Pada beberapa objek yang diamati terdapat berbagai macam konteksnya, seperti macam-macam opini netizen atau warganet pada *twitter.com* dan macam-macam halaman berita dari *website*.

Tabel dari hasil observasi terdapat pada tabel L.3.1. Data Observasi Lampiran 2. Pengumpulan Data.

C. Wawancara

Pada pengumpulan data, yang dilakukan selanjutnya adalah melakukan wawancara. Wawancara dilakukan dengan pegawai PT. Telkom Indonesia yang ada pada divisi *Officer Digital Consumer*. Pelaksanaan wawancara berlokasi di PT. Telkom Indonesia Divisi Regional V Surabaya pada tanggal 02/07/2019. Pada wawancara ditemukan hasil bahwa Telkom Indonesia belum memiliki aplikasi untuk menganalisis tentang *review* pelanggan terhadap IndiHome seperti apa saja

laporan yang sering muncul dari pelanggan dan bagaimana performa Telkom Indonesia dalam menangani laporan pelanggan.

3.1.2. Ruang Lingkup

Ruang Lingkup menjelaskan tentang lingkup penelitian yang akan dibahas dan yang tidak dibahas. Penelitian ini membahas tentang analisis sentimen mengenai *review* pelanggan dengan konotasi positif dan negatif yang ada di media sosial dan *website*. Penelitian ini tidak membahas tentang langkah yang harus diambil perusahaan dalam mengembangkan bisnis setelah mendapatkan hasil dari analisis sentimen.

3.1.3. Jadwal Kerja

Bagian ini menjelaskan penjadwalan pengerjaan penelitian dari tahap awal, tahap pengembangan, dan tahap akhir. Untuk lebih jelasnya, jadwal kerja terdapat pada tabel 3.2. Jadwal Kerja.

Tabel 3.1. Jadwal Kerja

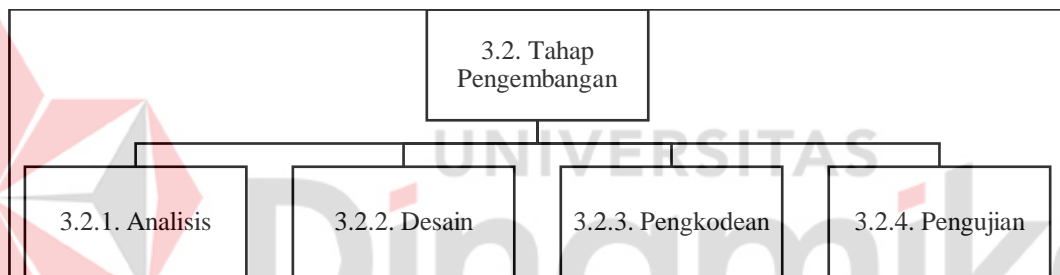
No	Kegiatan	Bulan Ke-1				Bulan Ke-2				Bulan Ke-3				Bulan Ke-4				Bulan Ke-5			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1.	Tahap Awal: Pengumpulan data (Studi Literatur, Observasi dan Wawancara), Ruang Lingkup, dan Jadwal Kerja.																				
2.	Tahap Pengembangan: Analisis, Desain, Pengkodean, dan Pengujian.																				
3.	Tahap Akhir: Evaluasi Uji Fungsi,																				

No	Kegiatan	Bulan Ke-1				Bulan Ke-2				Bulan Ke-3				Bulan Ke-4				Bulan Ke-5			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
	Evaluasi Hitungan, dan Evaluasi Uji Akurasi.																				

3.2. Tahap Pengembangan

Bagian ini menjelaskan tentang Tahap Pengembangan untuk menyelesaikan penelitian ini. Tahap Pengembangan meliputi 4 proses, Yaitu Analisis, Desain, Pengkodean, dan Pengujian. Gambaran tahap pengembangan terdapat pada gambar

3.3. Tahap Pengembangan.



Gambar 3.3. Tahap Pengembangan
(Sumber: Olahan Penulis)

3.2.1. Analisis

Tahapan selanjutnya yang dilakukan pada Tahap Pengembangan yaitu Analisis. Tahap Analisis bertujuan untuk mengetahui kebutuhan apa saja yang dibutuhkan pada penelitian ini. Analisis yang dilakukan terdiri dari 6 (enam) subbab, yaitu Identifikasi Masalah, Analisis Kebutuhan Pengguna, Analisis Kebutuhan Data dan Informasi, Analisis Kebutuhan Fungsional, Analisis Kebutuhan Non Fungsional, dan Analisis Kebutuhan Sistem.

A. Identifikasi Masalah

Tahap identifikasi masalah digunakan sebagai dasar dalam melakukan penelitian. Identifikasi masalah dilakukan dengan mempelajari studi literatur yang serupa dengan penelitian yang dikerjakan, melakukan observasi pada media sosial dan *website*, dan melakukan wawancara dengan pihak Telkom Indonesia sehingga didapatkan informasi yang berkaitan dengan penelitian. Berdasarkan Studi Literatur, Observasi, dan Wawancara yang sudah dilakukan, Telkom Indonesia melakukan rapat tahunan untuk mengetahui *review* atau pendapat pada layanan yang digunakan oleh pelanggan. Sedangkan dalam mengatasi *review* atau pendapat pelanggan, Telkom Indonesia mengatasi secara langsung melalui komplain secara langsung atau melalui telepon *Customer Service* (CS) dan media online *twitter.com*. Dari masalah tersebut, Telkom Indonesia yang berada pada Divisi Regional V Jawa Timur membutuhkan sebuah aplikasi yang dapat membantu untuk mengetahui *review* atau pendapat pelanggan terhadap layanan IndiHome. Untuk lebih jelasnya, terdapat pada tabel L.3.3. Identifikasi Permasalahan Lampiran 3. Analisis.

B. Analisis Kebutuhan Pengguna

Analisis Kebutuhan Pengguna dilakukan dengan tujuan untuk mengetahui kebutuhan pengguna yang digunakan dan/atau dibutuhkan. Berdasarkan hasil Wawancara dan Identifikasi Masalah maka dapat dilakukan analisis kebutuhan pengguna untuk penelitian yang akan dibuat, pengguna tersebut dapat diidentifikasi, yaitu *Consumer Marketing*. *Consumer Marketing* bertugas untuk memberikan informasi mengenai pengguna yang menggunakan IndiHome. Analisis kebutuhan pengguna dari *Consumer Marketing* terdapat pada tabel L.3.4. Analisis Kebutuhan Pengguna Lampiran 3. Analisis.

C. Analisis Kebutuhan Data dan Informasi

Analisis kebutuhan Data dan Informasi dilakukan dengan tujuan untuk mengetahui data dan informasi yang digunakan dan/atau dibutuhkan oleh pengguna sistem (perangkat lunak) yang akan dibuat. Selain itu juga untuk menganalisis *output* yang diperoleh dari pengguna tersebut. Berdasarkan hasil analisis kebutuhan pengguna, kebutuhan data dan informasi yang terdapat pada tabel L.3.5. Analisis Kebutuhan Data dan Informasi Lampiran 3. Analisis.

D. Analisis Kebutuhan Fungsional

Analisis Kebutuhan Fungsional dilakukan dengan tujuan untuk mengetahui kebutuhan fungsi yang akan digunakan. Berdasarkan kebutuhan pengguna, dibutuhkan kebutuhan fungsional berupa fungsi penambahan data untuk mengumpulkan data, fungsi untuk melakukan ekstraksi data, dan fungsi visualisasi hasil *review* dari pelanggan. Kebutuhan fungsional terdapat pada tabel L.3.6. Analisis Kebutuhan Fungsional dan Informasi Lampiran 3. Analisis.

E. Analisis Kebutuhan Non Fungsional

Analisis Kebutuhan Non Fungsional dilakukan dengan tujuan untuk mengetahui kebutuhan non fungsi yang akan digunakan. Kebutuhan Non Fungsional Yang dibutuhkan adalah kebutuhan akses dan kebutuhan dalam menyajikan data dalam bentuk visualisasi dengan waktu yang efektif. Kebutuhan Non Fungsional terdapat pada tabel L.3.7. Analisis Kebutuhan Non Fungsional dan Informasi Lampiran 3. Analisis.

F. Analisis Kebutuhan Sistem

1. Kebutuhan Perangkat Lunak

Kebutuhan perangkat lunak atau *software requirement* adalah komponen perangkat lunak yang digunakan untuk mengembangkan dan membangun aplikasi pada penelitian ini. Adapun perangkat lunak yang digunakan sebagai berikut:

- a. Sistem operasi yang digunakan pada penelitian ini menggunakan sistem operasi *windows 10* karena performa yang lebih cepat daripada sistem operasi *windows* yang lainnya dan lebih familier untuk digunakan.
- b. *XAMPP* digunakan untuk membuat *web server localhost* pada penelitian ini karena penggunaan untuk menyimpan database pada folder lebih mudah.
- c. Bahasa pemrograman yang digunakan pada penelitian ini yaitu *Hypertext Preprocessor* (PHP) karena PHP banyak digunakan untuk membuat aplikasi berbasis *website* yang membutuhkan fungsionalitas maksimal dengan minimal pengkodean.
- d. Basis data yang digunakan pada penelitian ini adalah *My Structured Query Language* (MYSQL) karena basis data (*database*) ini bersifat *Open Source* sehingga dapat digunakan oleh siapapun tanpa dibebani biaya lisensi yang lumayan tinggi.
- e. Alat pengkodean yang digunakan pada penelitian ini adalah *Visual Code* karena tidak dibebankan biaya.

2. Kebutuhan Perangkat Keras

Berdasarkan analisis kebutuhan perangkat lunak yang telah dibuat maka dapat dilanjutkan dengan analisis kebutuhan perangkat keras atau *hardware*. Kebutuhan perangkat keras merupakan komponen peralatan fisik yang membentuk

sistem komputer terstruktur, serta perangkat keras lain yang mendukung komputer dalam menjalankan fungsinya. Kebutuhan perangkat keras yang dibutuhkan adalah sebagai berikut:

- a. Pada penelitian ini menggunakan *Central Processor Unit (CPU) intel core i5-8250U*.
- b. Pada penelitian ini menggunakan *Random Access Memory (RAM)* sebesar 8GB.
- c. Pada penelitian ini menggunakan *Monitor* dengan resolusi 1920x1080 *pixel*.
- d. Pada penelitian ini menggunakan *Graphic Processing Unit (GPU) NVIDIA GeForce MX150*.

3.2.2. Desain

Tahapan selanjutnya yang dilakukan pada Tahap Pengembangan yaitu Desain. Tahap Desain bertujuan untuk mengetahui desain apa saja yang dibutuhkan pada penelitian ini. Desain yang dibutuhkan pada penelitian ini adalah *System Flow*, *Data Modelling (Conceptual Data Model (CDM), Physical Data Model (PDM), Struktur Basis Data (Desain Database))*, dan *Desain interface*.

A. *System Flow*

Bagian ini menjelaskan tentang *system flow* yang digunakan pada penelitian. *System Flow* pada penelitian ini merupakan alur pengerjaan sistem yang ada pada penelitian ini. Pada penelitian ini terdiri dari 3 (Tiga) *System Flow*, yaitu: *System Flow Penambangan Data*, *System Flow Ekstraksi Data*, dan *System Flow Klasifikasi Data*.

1. *System Flow* Penambangan Data

System Flow Penambangan Data adalah alur sistem yang berfungsi untuk melakukan pengambilan/penambangan (*crawling*) data. Proses dimulai ketika *Consumer Marketing* menjalankan aplikasi dengan menekan tombol '*generate*'. Setelah *Consumer Marketing* menekan tombol '*generate*', sistem pada aplikasi ini akan mencari data sesuai dengan sumber yang sudah ditentukan. Setelah selesai melakukan pencarian data, data diambil dan disimpan pada *database*. *System Flow* penambangan data terdapat pada gambar L.3.4. *System Flow* Penambangan Data Lampiran 4. Desain.

2. *System Flow* Ekstraksi Data

System Flow Ekstraksi Data adalah alur sistem yang berfungsi untuk melakukan proses mengekstraksi data dari hasil penambangan data. Proses dimulai ketika *Consumer Marketing* menjalankan aplikasi dengan menekan tombol '*Ekstraksi*'. Setelah *Consumer Marketing* menekan tombol '*Ekstraksi*', sistem pada aplikasi ini akan melakukan ekstraksi data pada hasil penambangan data. Proses ekstraksi data sendiri terdiri dari 4 (empat) proses, yaitu: proses *case folding*, proses *tokenizing*, proses *filtering*, dan proses *stemming*. Pada proses *tokenizing*, proses *filtering*, dan proses *stemming* menggunakan *plugin sastrawi*. Setelah selesai melakukan ekstraksi data, data disimpan pada *database*. *System Flow* ekstraksi data terdapat pada gambar L.3.5. *System Flow* Ekstraksi Data Lampiran 4. Desain.

3. *System Flow* Klasifikasi Data

System Flow Klasifikasi Data adalah alur sistem yang berfungsi untuk melakukan proses menentukan *class* dari data dari hasil ekstraksi data. Proses dimulai ketika *Consumer Marketing* menjalankan aplikasi dengan menekan tombol

'klasifikasi'. Setelah *Consumer Marketing* menekan tombol 'klasifikasi', sistem pada aplikasi ini akan mengambil data pada hasil ekstraksi data. Data yang digunakan pada ekstraksi data yaitu data hasil *stemming*. Setelah sistem mengambil data hasil ekstraksi data, sistem melakukan *filter* menggunakan kamus liu (untuk menentukan positif dan negatif) dan melakukan perhitungan menggunakan *naïve bayes*. Pada penelitian ini *naïve bayes* digunakan untuk menentukan *class* pada data yang diolah. Setelah selesai melakukan klasifikasi data, data disimpan pada *database*. *System Flow* klasifikasi data terdapat pada gambar L.3.6. *System Flow* Klasifikasi Data Lampiran 4. Desain. Pada proses *filtering* menggunakan kamus liu dan perhitungan *naïve bayes* dijelaskan pada *flowchart* kamus liu dan *flowchart* *naïve bayes* terdapat pada gambar L.3.7. *Flowchart* Kamus Liu dan L.3.8. *Flowchart* *Naïve Bayes* Lampiran 4. Desain.

B. *Data Modelling*

Data Modelling berisikan *Conceptual Data Model* (CDM), *Physical Data Model* (PDM), dan Struktur Basis Data. *Conceptual Data Model* (CDM) berguna sebagai gambaran dari struktur basis data yang digunakan berdasarkan analisis dan pengumpulan data. *Physical Data Model* (PDM) berguna sebagai gambaran dari struktur basis data yang digunakan bersifat lebih spesifik dan representasi bentuk fisik dari basis data yang digunakan.

CDM pada penelitian ini terdiri dari 6 (enam) tabel yaitu tabel *data_crawling*, tabel *case_folding*, tabel *tokenize*, tabel *stopword*, tabel *stemmer*, dan tabel klasifikasi. PDM pada penelitian ini terdiri dari 6 (enam) tabel yaitu tabel *data_crawling*, tabel *case_folding*, tabel *tokenize*, tabel *stopword*, tabel *stemmer*, dan tabel klasifikasi. Struktur Basis Data berguna sebagai gambaran dari struktur

basis data yang digunakan pada aplikasi. Struktur Basis Data dibuat berdasarkan CDM dan PDM. Struktur Basis Data yang digunakan adalah tabel *data_crawling*, tabel *case_folding*, tabel *tokenize*, tabel *stopword*, tabel *stemmer*, dan tabel klasifikasi. Tabel *data_crawling* berfungsi untuk menyimpan hasil penambangan data, tabel *case_folding* berfungsi untuk menyimpan hasil pengubahan data menjadi huruf kecil, tabel *tokenize* berfungsi untuk menyimpan hasil pemotongan kata, tabel *stopword* berfungsi untuk menyimpan hasil penghilangan kata yang tidak digunakan, tabel *stemmer* berfungsi untuk menyimpan hasil pengembalian ke kata dasar, dan tabel klasifikasi berfungsi untuk menyimpan hasil penentuan *class* pada data. CDM dan PDM pada penelitian ini terdapat pada L.3.9. *Conceptual Data Model* (CDM) dan gambar L.3.10. *Physical Data Model* (PDM) Lampiran 4. Desain. Sedangkan, Struktur Basis Data pada penelitian ini terdapat pada tabel L.3.8. Struktur Basis Data *data_crawling*, tabel L.3.9. Struktur Basis Data *case_folding*, tabel L.3.10. Struktur Basis Data *tokenize*, tabel L.3.11. Struktur Basis Data *stopword*, tabel L.3.12. Struktur Basis Data *stemmer*, dan tabel L.3.13. Struktur Basis Data klasifikasi Lampiran 4. Desain.

C. Desain Interface

Desain *Interface* adalah gambaran dari aplikasi yang dibuat. Desain *Interface* dibuat berdasarkan kebutuhan fungsi. Berikut ini adalah gambaran Desain *Interface* yang akan diimplementasikan pada penelitian ini:

Tabel 3.2. Desain *Interface*

Desain <i>Interface</i>	Fungsi	Gambar
Desain <i>Interface Login</i>	berfungsi sebagai penghubung antar pengguna dengan aplikasi.	Terdapat pada gambar L.3.11. Desain <i>Interface Login</i> Lampiran 4. Desain.

Desain <i>Interface</i>	Fungsi	Gambar
Desain <i>Interface</i> Halaman Utama <i>Dashboard</i>	berfungsi melakukan penambangan data dan menampilkan hasil dari seluruh proses yang sudah dilakukan. pada desain <i>interface</i> halaman utama terdapat tombol 'perbarui data' yang berfungsi untuk melakukan penambangan data.	Terdapat pada gambar L.3.12. Desain <i>Interface</i> Halaman Utama Lampiran 4. Desain.
Desain <i>Interface</i> Halaman Hasil Data <i>Crawling</i>	berfungsi sebagai halaman yang menampilkan hasil dari penambangan data yang sudah dilakukan.	Terdapat pada gambar L.3.13. Desain <i>Interface</i> Halaman Data Hasil <i>Crawling</i> Lampiran 4. Desain.
Desain <i>Interface</i> Hasil Ekstraksi Data (<i>Case Folding</i>)	berfungsi sebagai halaman yang menampilkan hasil dari proses ekstraksi data (<i>case folding</i>).	Terdapat pada gambar L3.14. Desain <i>Interface</i> Data Hasil Ekstraksi (<i>Case Folding</i>) Lampiran 4. Desain.
Desain <i>Interface</i> Hasil Ekstraksi Data (<i>Tokenizing</i>)	berfungsi sebagai halaman yang menampilkan hasil dari proses ekstraksi data (<i>tokenizing</i>).	Terdapat pada gambar L.3.15. Desain <i>Interface</i> Data Hasil Ekstraksi (<i>Tokenizing</i>) Lampiran 4. Desain.
Desain <i>Interface</i> Hasil Ekstraksi Data (<i>Filtering</i>)	sebagai halaman yang menampilkan hasil dari proses ekstraksi data (<i>filtering</i>).	Terdapat pada gambar L.3.16. Desain <i>Interface</i> Data Hasil Ekstraksi (<i>Filtering</i>) Lampiran 4. Desain.
Desain <i>Interface</i> Hasil Ekstraksi Data (<i>Stemming</i>)	berfungsi sebagai halaman yang menampilkan hasil dari proses ekstraksi data (<i>stemming</i>).	Terdapat pada gambar L.3.17. Desain <i>Interface</i> Data Hasil Ekstraksi (<i>Stemming</i>) Lampiran 4. Desain.
Desain <i>Interface</i> Hasil Klasifikasi Data.	berfungsi sebagai halaman yang menampilkan hasil dari proses klasifikasi data.	Terdapat pada gambar L.3.18. Desain <i>Interface</i> Data Klasifikasi Lampiran 4. Desain.

3.2.3. Pengkodean

Tahapan selanjutnya yang dilakukan pada tahap pengembangan yaitu pengkodean. Tahap Pengkodean bertujuan untuk menjalankan proses pengkodean pada tahap yang dilakukan pada penelitian ini. Pengkodean yang dilakukan pada penelitian ini adalah adalah pengkodean Penambangan Data, Ekstraksi Data, dan Klasifikasi Data.

A. Penambangan Data

Pengkodean yang dilakukan pertama kali pada penelitian ini adalah melakukan penambangan data. Penambangan data digunakan untuk mengambil data yang akan digunakan pada penelitian ini. Untuk pengkodean yang digunakan pada penambangan terdapat pada gambar L.3.19. Kode Penambangan Data Lampiran 5. Pengkodean.

Tahapan selanjutnya yang dilakukan dalam pengkodean adalah melakukan ekstraksi data. Ekstraksi data bertujuan untuk mengolah hasil penambangan data yang formatnya tidak terstruktur hingga menjadi format kata dasar. Ekstraksi data yang dilakukan pada penelitian ini terdiri dari 4 (empat) tahap, yaitu *case folding*, *tokenizing*, *filtering*, dan *stemming*. Pada penelitian ini tahapan *tokenizing*, *filtering*, dan *stemming* menggunakan *plugin sastrawi*.

B. Case Folding

Tahapan pertama yang dilakukan dalam ekstraksi data adalah *case folding*. *Case folding* memiliki fungsi untuk merubah semua huruf menjadi huruf kecil. Untuk pengkodean yang digunakan pada ekstraksi data *case folding* terdapat pada gambar L.3.20. Kode Ekstraksi Data *Case Folding* Lampiran 5. Pengkodean.

C. *Tokenizing*

Tahapan kedua yang dilakukan dalam ekstraksi data adalah *tokenizing*. *Tokenizing* memiliki fungsi untuk memotong atau membagi kalimat menjadi kata-kata. Untuk pengkodean yang digunakan pada ekstraksi data *tokenizing* terdapat pada gambar L.3.21. Kode Ekstraksi Data *Tokenizing* Lampiran 5. Pengkodean.

D. *Filtering*

Tahapan ketiga yang dilakukan dalam ekstraksi data adalah *filtering*. *Filtering* memiliki fungsi untuk menghilangkan kata-kata yang tidak digunakan atau tidak penting. Untuk pengkodean yang digunakan pada ekstraksi data *filtering* terdapat pada gambar L.3.22. Kode Ekstraksi Data *Filtering* Lampiran 5.

Pengkodean.

E. *Stemming*

Tahapan keempat yang dilakukan dalam ekstraksi data adalah *stemming*. *Stemming* memiliki fungsi untuk mengembalikan kata imbuhan menjadi kata dasar. Untuk pengkodean yang digunakan pada ekstraksi data *stemming* terdapat pada gambar L.3.23. Kode Ekstraksi Data *Stemming* Lampiran 5. Pengkodean.

F. *Klasifikasi Data*

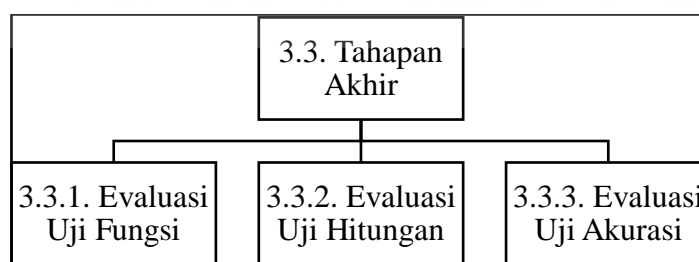
Tahapan selanjutnya yang dilakukan dalam pengkodean adalah melakukan klasifikasi data. Klasifikasi data pada penelitian ini memiliki fungsi untuk menentukan *class* pada data. Untuk pengkodean yang digunakan pada klasifikasi data terdapat pada gambar L.3.24. Kode Klasifikasi Data Lampiran 5. Pengkodean.

3.2.4. Pengujian

Tahapan selanjutnya yang dilakukan pada tahap pengembangan yaitu pengujian. Tahap pengujian bertujuan untuk melakukan uji coba pada fungsi yang digunakan pada penelitian ini. pengujian yang dilakukan pada penelitian ini adalah pengujian fungsi penambahan data, ekstraksi data, dan klasifikasi data. Sumber data yang digunakan pada penelitian ini menggunakan *mediakonsumen.com*. Pengujian terdapat pada tabel L3.15. Uji Coba Penambahan Data, tabel L3.16. Uji Coba Ekstraksi Data, tabel L3.17. Uji Coba Klasifikasi Data dan tabel L3.18. Uji Coba Visualisasi Data Lampiran 6. Pengujian.

3.3. Tahap Akhir

Bagian ini menjelaskan tentang Tahap Akhir untuk menyelesaikan penelitian ini. Tahap Akhir meliputi 3 (Tiga) Subbab, yaitu Evaluasi Uji Fungsi, Evaluasi Uji Hitungan, dan Evaluasi Uji Akurasi. Untuk gambaran tahap akhir terdapat pada gambar 3.25. Tahap Akhir.



Gambar 3.4. Tahap Akhir
(Sumber: Olahan Penulis)

3.3.1. Evaluasi Uji Fungsi

Tahapan selanjutnya yang dilakukan pada tahap akhir yaitu evaluasi uji fungsi. Evaluasi uji fungsi dilakukan dengan tujuan untuk mengetahui apakah hasil dari uji fungsi yang ada pada penelitian ini dapat diimplementasikan sesuai dengan

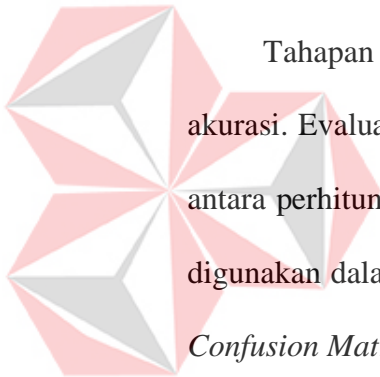
harapan. Evaluasi uji fungsi yang dilakukan pada penelitian ini berupa pengujian fungsi Penambangan Data, Ekstraksi Data, Klasifikasi Data, dan Visualisasi Data.

3.3.2. Evaluasi Uji Perhitungan

Tahapan selanjutnya yang dilakukan pada tahap akhir yaitu evaluasi uji hitungan. Evaluasi uji hitungan dilakukan dengan tujuan untuk mengetahui perbandingan antara perhitungan manual dan perhitungan dalam bentuk program. Metode yang digunakan dalam melakukan perhitungan manual dan perhitungan dalam bentuk program menggunakan metode *naïve bayes*.

3.3.3. Evaluasi Uji Akurasi

Tahapan selanjutnya yang dilakukan pada tahap akhir yaitu evaluasi uji akurasi. Evaluasi uji hitungan dilakukan dengan tujuan untuk mengetahui akurasi antara perhitungan manual dan perhitungan dalam bentuk program. Metode yang digunakan dalam melakukan uji akurasi pada penelitian ini menggunakan metode *Confusion Matrix*.



UNIVERSITAS
Dinamika

BAB IV

HASIL DAN PEMBAHASAN

Bab keempat yang ditulis pada penelitian ini adalah hasil dan pembahasan. Pada bab ini akan dibahas hasil dan implementasi dari penelitian berdasarkan tahapan-tahapan yang sudah dilakukan sesuai dengan metode yang digunakan. Hasil dan implementasi dibagi menjadi 3 (tiga) subbab, yaitu: Implementasi, Pengujian, dan Evaluasi.

4.1. Implementasi

Bagian ini menjelaskan tentang implementasi yang dilakukan pada penelitian ini. Implementasi berisi tampilan hasil rancangan aplikasi yang disesuaikan dengan Kebutuhan Fungsional dan Kebutuhan Non Fungsional. Berikut adalah implementasi pada tampilan hasil rancangan aplikasi.

4.1.1. Halaman Login

Pada Halaman ini *Digital Consumer* akan melakukan *login* untuk mengakses aplikasi dengan cara memasukkan *username* dan *password*. Setelah mengisi *username* dan *password*, *Digital Consumer* menekan tombol *login* agar terproses dan masuk ke dalam aplikasi *website*. Gambaran dari halaman *login* terdapat pada gambar L.4.1. Halaman *Login* Lampiran 7. Implementasi.

4.1.2. Halaman Utama

Setelah melakukan *login*, *Digital Consumer* dapat masuk ke dalam halaman utama dan melihat hasil visualisasi. Pada halaman utama menampilkan visualisasi dari hasil melakukan analisis sentimen sesuai dengan tahapan dan metode yang

digunakan. Halaman utama berisi data perbulan dalam bentuk grafik garis (*line chart*) dan kelas analisis sentimen positif atau negatif dalam bentuk grafik lingkaran (*pie chart*). Gambaran dari halaman utama terdapat pada gambar L.4.2. Halaman Utama Lampiran 7. Implementasi.

Selain itu, pada halaman utama terdapat tombol untuk melakukan penambangan (*crawling*) data. Ketika menekan tombol ‘perbarui data sekarang’ pada halaman utama, muncul halaman *pop up* (halaman *modal*) untuk melakukan penambangan data. Gambaran dari proses pengambilan data terdapat pada gambar L.4.3. Halaman *Pop Up* Penambangan Data Lampiran 7. Implementasi.

Setelah data berhasil diambil, muncul halaman *pop up* (halaman *modal*) pemberitahuan dimana data berhasil didapatkan dan pada halaman *pop up* dan terdapat tombol ‘mulai ekstraksi data’ yang digunakan untuk melakukan ekstraksi data. Gambaran dari pemberitahuan pengambilan data selesai dan proses ekstraksi data terdapat pada gambar L.4.4. Halaman *Pop Up* Ekstraksi Data Lampiran 7. Implementasi.

Setelah data berhasil diekstraksi, muncul halaman *pop up* (halaman *modal*) pemberitahuan dimana data berhasil diekstraksi dan pada halaman *pop up* juga terdapat tombol ‘lihat data’ untuk melihat data yang sudah melalui penambangan data dan ekstraksi data. Gambaran dari proses ekstraksi data terdapat pada gambar L.4.5. Halaman *Pop Up* Ekstraksi Data Lampiran 7. Implementasi.

4.1.3. Halaman Hasil Data *Crawling*

Halaman Hasil *Crawling* Data menampilkan hasil dari penambangan data. halaman ini berisi nama *author*, tanggal *posting*, *link*, sumber, imputan dan aksi. Pada saat menekan tombol aksi, halaman *pop up* detail dari hasil dari penambangan

data akan ditampilkan. Gambaran dari halaman hasil data *crawling* terdapat pada gambar L.4.6. Halaman Hasil Data *Crawling* dan gambar L.4.7. Halaman *Pop Up* Detail Hasil Data *Crawling* Lampiran 7. Implementasi.

4.1.4. Halaman Hasil Ekstraksi Data (*Case Folding*)

Halaman hasil ekstraksi data (*case folding*) menampilkan hasil dari *case folding* dimana data inputan diubah menjadi *lowercase*. Halaman hasil ekstraksi data (*case folding*) berisikan kode, hasil *case folding*, dan aksi. Pada saat menekan tombol aksi, maka akan keluar halaman *pop up* detail dari hasil dari ekstraksi data (*case folding*). Gambaran dari halaman hasil ekstraksi data (*case folding*) terdapat pada gambar L.4.8. Halaman Hasil Ekstraksi Data (*Case Folding*) dan gambar L.4.9. Halaman Hasil Ekstraksi Data (*Case Folding*) Lampiran 7. Implementasi.

4.1.5. Halaman Hasil Ekstraksi (*Tokenizing*)

Halaman hasil ekstraksi data (*tokenizing*) menampilkan hasil dari *tokenizing* dimana data inputan dipotong-potong. Halaman hasil ekstraksi data (*tokenizing*) berisikan kode, hasil *tokenizing*, dan aksi. Pada saat menekan tombol aksi, maka akan keluar halaman *pop up* detail dari hasil dari ekstraksi data (*tokenizing*). Gambaran dari halaman hasil ekstraksi data (*tokenizing*) terdapat pada gambar L.4.10. Halaman Hasil Ekstraksi Data (*Tokenizing*) dan gambar L.4.11. Halaman Detail Hasil Ekstraksi Data (*Tokenizing*) Lampiran 7. Implementasi.

4.1.6. Halaman Hasil Ekstraksi (*Filtering*)

Halaman hasil ekstraksi data (*filtering*) menampilkan hasil dari *filtering* dimana data inputan dihilangkan kata-kata imbuhan dsb. Halaman hasil ekstraksi data (*filtering*) berisikan kode, hasil *filtering*, dan aksi. Pada saat menekan tombol

aksi, maka akan keluar halaman *pop up* detail dari hasil dari ekstraksi data (*filtering*). Gambaran dari halaman hasil ekstraksi data (*filtering*) terdapat pada gambar L.4.12. Halaman Hasil Ekstraksi Data (*Filtering*) dan gambar L.4.13. Halaman Detail Hasil Ekstraksi Data (*Filtering*) Lampiran 7. Implementasi.

4.1.7. Halaman Hasil Ekstraksi (*Stemming*)

Halaman hasil ekstraksi data (*stemming*) menampilkan hasil dari *stemming* dimana data inputan dikembalikan ke kata dasar. Halaman hasil ekstraksi data (*stemming*) berisikan kode, hasil *stemming*, dan aksi. Pada saat menekan tombol aksi, maka akan keluar halaman *pop up* detail dari hasil dari ekstraksi data (*stemming*). Gambaran dari halaman hasil ekstraksi data (*stemming*) terdapat pada gambar L.4.14. Halaman Hasil Ekstraksi Data (*Stemming*) dan gambar L.4.15. Halaman Detail Hasil Ekstraksi Data (*Stemming*) Lampiran 7. Implementasi.

4.1.8. Halaman Hasil Klasifikasi Data

Halaman hasil klasifikasi data menampilkan hasil dari klasifikasi data. Halaman hasil klasifikasi berisikan tanggal *posting*, sumber, inputan, hasil klasifikasi, dan aksi. Pada halaman hasil klasifikasi data terdapat tombol ‘lakukan perhitungan data’ untuk mengubah halaman hasil klasifikasi data menjadi halaman perhitungan klasifikasi data. Gambaran dari halaman hasil klasifikasi terdapat pada gambar L.4.16. Halaman Hasil Klasifikasi Data Lampiran 7. Implementasi.

Setelah pada halaman hasil klasifikasi data menekan tombol ‘lakukan perhitungan data’, isi halaman berubah menjadi total seluruh data, data untuk *training*, dan data untuk *testing* serta terdapat tombol ‘mulai lakukan perhitungan’. Pada saat menekan tombol ‘mulai lakukan perhitungan’, sistem akan melakukan

perhitungan klasifikasi data pada data *training* dan data *testing* menggunakan metode perhitungan *naïve bayes*. Gambaran dari halaman perhitungan klasifikasi terdapat pada gambar L.4.17. Halaman Perhitungan Klasifikasi Data Lampiran 7. Implementasi.

Setelah perhitungan klasifikasi data selesai, akan muncul pemberitahuan bahwa perhitungan klasifikasi data selesai disertai dengan detail data yang dihitung dan hasil dari data *testing*. Gambaran dari halaman perhitungan klasifikasi data terdapat pada gambar L.4.18. Halaman Hasil Perhitungan Klasifikasi Data Lampiran 7. Implementasi.

Pada halaman hasil klasifikasi data terdapat tombol aksi yang berguna untuk menampilkan detail klasifikasi data. Pada saat menekan tombol aksi, maka akan keluar halaman *pop up* detail dari klasifikasi data. Gambaran dari halaman detail klasifikasi data terdapat pada gambar L.4.19. Halaman Detail Klasifikasi Data Lampiran 7. Implementasi.

4.2. Pengujian

Bagian ini menjelaskan tentang pengujian yang dilakukan pada penelitian. Pengujian yang dilakukan terdiri dari pengujian fungsi, pengujian perhitungan, dan pengujian akurasi. Pengujian menggunakan sumber yang sama, yaitu *mediakonsumen.com*, *twitter.com*, dan *detik.com*. Berikut adalah hasil pengujian yang ada pada penelitian ini.

4.2.1. Pengujian Fungsi

Bagian ini menjelaskan tentang pengujian fungsi yang digunakan pada penelitian ini. Pengujian fungsi yang dilakukan terdiri dari 4 (empat) pengujian, yaitu Pengujian Fungsi Mengambil Data/*Crawling* Data, Pengujian Fungsi Mengekstraksi Data, Pengujian Fungsi Mengklasifikasi Data, dan Pengujian Fungsi Memvisualisasikan Data.

A. Pengujian Fungsi *Crawling* Data

Pengujian yang dilakukan pada fungsi *crawling* sebanyak 3 (tiga) kali pengujian. Pengujian yang pertama pada fungsi *crawling* dengan sumber mediakonsumen.com, pengujian yang kedua pada fungsi *crawling* dengan sumber twitter.com, dan pengujian yang ketiga pada fungsi *crawling* dengan sumber detik.com. hasil pengujian fungsi *crawling* data terdapat pada tabel L.4.1. Pengujian Fungsi *Crawling* Data (mediakonsumen.com), tabel L.4.2. Pengujian Fungsi *Crawling* Data (twitter.com), dan tabel L.4.3. Pengujian Fungsi *Crawling* Data (detik.com) Lampiran 8. Pengujian Fungsi.

B. Pengujian Fungsi Ekstraksi Data

Pengujian yang dilakukan pada fungsi ekstraksi data sebanyak 3 (tiga) kali pengujian. Pengujian yang pertama pada fungsi ekstraksi data dengan sumber mediakonsumen.com, pengujian yang kedua pada fungsi ekstraksi data dengan sumber twitter.com, dan pengujian yang ketiga pada fungsi ekstraksi data dengan sumber detik.com. pengujian fungsi ekstraksi data terdapat pada tabel L.4.4. Pengujian Fungsi Ekstraksi Data (mediakonsumen.com), tabel L.4.5. Pengujian Fungsi Ekstraksi Data (twitter.com), dan tabel L.4.6. Pengujian Fungsi Ekstraksi Data (detik.com) Lampiran 8. Pengujian Fungsi.

C. Pengujian Fungsi Klasifikasi Data

Pengujian yang dilakukan pada fungsi klasifikasi data sebanyak 3 (tiga) kali pengujian. Pengujian yang pertama pada fungsi klasifikasi data dengan sumber mediakonsumen.com, pengujian yang kedua pada fungsi klasifikasi data dengan sumber twitter.com, dan pengujian yang ketiga pada fungsi klasifikasi data dengan sumber detik.com. pengujian fungsi klasifikasi terdapat pada tabel L.4.7. Pengujian Fungsi Klasifikasi Data (mediakonsumen.com), tabel L.4.8. Pengujian Fungsi Klasifikasi Data (twitter.com), dan tabel L.4.9. Pengujian Fungsi Klasifikasi Data (detik.com) Lampiran 8. Pengujian Fungsi.

D. Pengujian Fungsi Visualisasi Data

Pengujian yang dilakukan pada fungsi visualisasi data sebanyak 1 (tiga) kali pengujian. Pengujian pada fungsi visualisasi data dengan sumber data hasil dari klasifikasi data. Pengujian fungsi visualisasi data terdapat pada tabel L.4.10. Pengujian Fungsi Visualisasi Data Lampiran 8. Pengujian Fungsi.

4.2.2. Pengujian Perhitungan

Bagian ini menjelaskan tentang pengujian perhitungan yang digunakan pada penelitian ini. Pengujian perhitungan yang dilakukan terdiri dari 2 (dua) pengujian. Pengujian perhitungan yang dilakukan adalah pengujian perhitungan aplikasi dan pengujian perhitungan manual.

A. Pengujian Perhitungan Aplikasi

Bagian ini menjelaskan tentang pengujian perhitungan pada aplikasi. pada perhitungan dengan aplikasi terdapat tampilan tabel data master. Tabel data master berfungsi untuk menampilkan rekaman dokumen dan kata-kata positif atau negatif

apa saja yang muncul. Tabel data master terdapat pada gambar L.4.20. Tabel Data Master Lampiran 9. Pengujian Perhitungan.

Dalam melakukan perhitungan tabel data master dibagi menjadi 2 (dua) tabel, yaitu tabel data *training* positif dan tabel data *training* negatif. Setelah dibagi menjadi 2 tabel, dilakukan proses perhitungan data pada tabel data *training* positif dan tabel data *training* negatif. Tabel data *training* positif dan tabel data *training* negatif terdapat pada gambar L.4.21. Tabel Data *Training* Positif dan gambar L.4.22. Tabel Data *Training* Negatif Lampiran 9. Pengujian Perhitungan.

Kemudian, setelah melakukan proses perhitungan pada data *training* positif dan negatif, hasil perhitungan data *testing* ditampilkan pada halaman hasil perhitungan *testing*. Hasil perhitungan *testing* berisikan hasil *stemming*, nilai positif, nilai negatif, dan kesimpulan. Pada pengujian perhitungan aplikasi menggunakan sumber data mediakonsumen.com, twitter.com, dan detik.com. data *testing* yang digunakan berdasarkan *crawling* data pada tanggal 24 februari 2020. Dari *crawling* data yang dilakukan didapatkan 3 data baru. Hasil pengujian perhitungan *testing* terdapat pada gambar L.4.23. Hasil Perhitungan *Testing* mediakonsumen.com, L.4.24. Hasil Perhitungan *Testing* twitter.com, dan gambar L.4.25. Hasil Perhitungan *Testing* detik.com Lampiran 9. Pengujian Perhitungan.

B. Pengujian Perhitungan Manual

Bagian ini menjelaskan tentang pengujian perhitungan secara manual. Pada perhitungan secara manual dibuat terlebih dahulu tabel data *training* dan data *testing*. Tabel data *training* dan data *testing* terdapat pada tabel L.4.11. Tabel Data *Training* Dan Data *Testing* Lampiran 9. Pengujian Perhitungan. Setelah membuat data *training* dan data *testing*, dibuat tabel data master. tabel data master berfungsi

untuk menampilkan rekaman dokumen data dan kata-kata positif atau negatif apa saja yang muncul. Tabel data master terdapat pada tabel L.4.12. Tabel Data Master Lampiran 8. Pengujian Perhitungan.

Dalam melakukan perhitungan, tabel data master dibagi menjadi 2 (dua) tabel, yaitu tabel data *training* positif dan tabel data *training* negatif. Setelah dibagi menjadi 2 tabel, dilakukan proses perhitungan data pada tabel data *training* positif dan tabel data *training* negatif. Tabel data *training* positif dan tabel data *training* negatif terdapat pada tabel L.4.13. Tabel Data *Training* Positif dan tabel L.4.14. Tabel Data *Training* Negatif Lampiran 9. Pengujian Perhitungan.

Setelah data *training* positif dan tabel data *training* negatif dibuat, dilakukan perhitungan pada tiap data *training* positif dan data *training* negatif. Detail hasil perhitungan data *training* terdapat pada tabel L.4.15. Perhitungan Data *Training* Positif dan Negatif Lampiran 9. Pengujian Perhitungan. Setelah data *training* positif dan data *training* negatif dihitung, dilakukan perhitungan pada tiap data *testing*. data *testing* yang digunakan menggunakan sumber yang sama pada perhitungan aplikasi. sumber yang digunakan adalah mediakonsumen.com, twitter.com, dan detik.com. Berikut adalah proses perhitungan data *testing* secara manual.

Dokumen 41 = Menanti Nikmatnya Layanan IndiHome di Rumah Saya

stemming = 'rumah'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$P(\text{positif}) = \text{null}$

= null

= null

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$P(\text{negatif}) = \text{null}$

$= \text{null}$

$= \text{null}$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 41 tidak dapat diklasifikasi, karena nilai perhitungan dokumen uji pada kelas negatif dan kelas positif tidak memiliki nilai.

Dokumen 42 = Akhir Cerita Pahit Saya dengan IndiHome

stemming = 'akhir | cerita | pahit'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$P(\text{positif}) = \text{null}$

$= \text{null}$

$= \text{null}$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$P(\text{negatif}) = \text{null}$

$= \text{null}$

$= \text{null}$

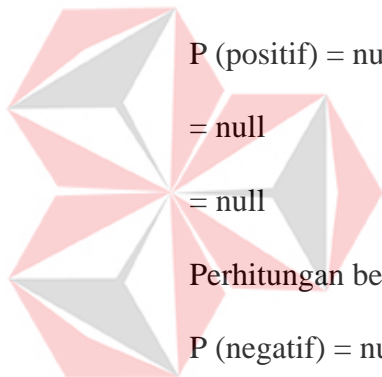
Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 42 tidak dapat diklasifikasi, karena nilai perhitungan dokumen uji pada kelas negatif dan kelas positif tidak memiliki nilai.

Dokumen 43 = Pemasangan Sambungan Baru IndiHome Mengecewakan, Janji

7×24 Jam Akan Diproses Tidak Dipenuhi

stemming = 'baru | janji | jam | tidak'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji



$$P(\text{positif}) = P(\text{tidak} | \text{positif})$$

$$= 0,018867925$$

$$= 0,018867925$$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$$P(\text{negatif}) = P(\text{tidak} | \text{negatif})$$

$$= 0,131147541$$

$$= 0,131147541$$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 43 ada dalam kelas negatif, karena nilai perhitungan dokumen uji pada kelas negatif lebih besar daripada nilai perhitungan dokumen uji pada kelas positif.

Dokumen 44 = Jangan. @IndiHome mahal, jelek, lemot dan gak bisa buka

@NetflixID . Maklum produknya @TelkomIndonesia. Sama kayak si

@Telkomsel . BUMN

tolol.<https://twitter.com/detikcom/status/1232271218625122304> ...

stemming = 'jangan | mahal | jelek | bisa | buka | maklum | sama | kayak | si | tolol |

status'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$$P(\text{positif}) = P(\text{mahal} | \text{positif}) + P(\text{jelek} | \text{positif}) + P(\text{bisa} | \text{positif})$$

$$= 0,018867925 + 0,018867925 + 0,018867925$$

$$= 0,056603775$$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$$P(\text{negatif}) = P(\text{mahal} | \text{negatif}) + P(\text{jelek} | \text{negatif}) + P(\text{bisa} | \text{mahal})$$

$$= 0,032786885 + 0,032786885 + 0,016393443$$

$$= 0,081967213$$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 44 ada dalam kelas negatif, karena nilai perhitungan dokumen uji pada kelas negatif lebih besar daripada nilai perhitungan dokumen uji pada kelas positif.

Dokumen 45 = @IndiHome min kmren sy di tlp dr telkomsel indihome dan ditawarkan tambahan paket khusus dibuka 17 chanel dgn tambahan hrg 20rb dan aktif mulai hari ini tp sy cek ko chanel nya blm terbuka min? @TelkomCare bisa dibantu ? Terlampir no indihome sy pic.twitter.com/Mxhg5fzP76

stemming = 'min | paket | khusus | aktif | mulai | hari | cek | min | bisa'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$$P(\text{positif}) = P(\text{bisa} | \text{positif})$$

$$= 0,018867925$$

$$= 0,018867925$$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$$P(\text{negatif}) = P(\text{bisa} | \text{negatif})$$

$$= 0,016393443$$

$$= 0,016393443$$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 45 ada dalam kelas positif, karena nilai perhitungan dokumen uji pada kelas positif lebih besar daripada nilai perhitungan dokumen uji pada kelas negatif.

Dokumen 46 = kenapa ni indihome tiba-tiba sinyalnya gak connect? dan use tv pun tidak bisa di akses @IndiHome pic.twitter.com/21bu9m2Igm

stemming = 'ni | pun | tidak | bisa | akses'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$$P(\text{positif}) = P(\text{tidak} | \text{positif}) + P(\text{bisa} | \text{positif})$$

$$= 0,018867925 + 0,018867925$$

$$= 0,03773585$$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$$P(\text{negatif}) = P(\text{tidak} | \text{negatif}) + P(\text{bisa} | \text{mahal})$$

$$= 0,131147541 + 0,016393443$$

$$= 0,147540984$$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 46 ada dalam kelas negatif, karena nilai perhitungan dokumen uji pada kelas negatif lebih besar daripada nilai perhitungan dokumen uji pada kelas positif.

Dokumen 47 = @IndiHome pagi, koneksi kenapa lagi ini ?

pic.twitter.com/1gq6aaZE8g

stemming = 'pagi | koneksi | lagi'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$$P(\text{positif}) = \text{null}$$

$$= \text{null}$$

$$= \text{null}$$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$$P(\text{negatif}) = \text{null}$$

$$= \text{null}$$

$$= \text{null}$$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 47 tidak dapat diklasifikasi, karena nilai perhitungan dokumen uji pada kelas negatif dan kelas positif tidak memiliki nilai.

Dokumen 48 = Telkom Rilis IndiHome Lite, Paket Internet Rp 199 Ribu

stemming = 'rilis | paket | ribu'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$$P(\text{positif}) = \text{null}$$

$$= \text{null}$$

$$= \text{null}$$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$$P(\text{negatif}) = \text{null}$$

$$= \text{null}$$

$$= \text{null}$$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 48 tidak dapat diklasifikasi, karena nilai perhitungan dokumen uji pada kelas negatif dan kelas positif tidak memiliki nilai.

Dokumen 49 = Telkom soal Blokir Netlix: Kami Tidak Diskriminasi

stemming = 'soal | blokir | tidak | diskriminasi'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$$P(\text{positif}) = P(\text{tidak} | \text{positif})$$

$$= 0,018867925$$

$$= 0,018867925$$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$$P(\text{negatif}) = P(\text{tidak} | \text{negatif})$$

$$= 0,131147541$$

$$= 0,131147541$$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 49 ada dalam kelas negatif, karena nilai perhitungan dokumen uji pada kelas negatif lebih besar daripada nilai perhitungan dokumen uji pada kelas positif.

Dokumen 50 = Paket Telepon IndiHome (Single Play) Lebih Mahal dari Paket

(Dual Play) Internetnya

stemming = 'paket | telepon | lebih | mahal | paket'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$$P(\text{positif}) = P(\text{mahal} | \text{positif})$$

$$= 0,018867925$$

$$= 0,018867925$$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$$P(\text{negatif}) = P(\text{mahal} | \text{negatif})$$

$$= 0,032786885$$

$$= 0,032786885$$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 50 ada dalam kelas

negatif, karena nilai perhitungan dokumen uji pada kelas negatif lebih besar

daripada nilai perhitungan dokumen uji pada kelas positif.

Dokumen 51 = Kecewa dengan Pelayanan Indihome

stemming = 'kecewa'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$$P(\text{positif}) = P(\text{kecewa} | \text{positif})$$

$$= 0,018867925$$

$$= 0,018867925$$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$$P(\text{negatif}) = P(\text{kecewa} | \text{negatif})$$

$$= 0,06557377$$

$$= 0,06557377$$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 51 ada dalam kelas

negatif, karena nilai perhitungan dokumen uji pada kelas negatif lebih besar

daripada nilai perhitungan dokumen uji pada kelas positif.

Dokumen 52 = Pagi, Kak Rachellyca. Perihal keluhannya silahkan dicoba restart modemnya terlebih dahulu. Tunggu 3-5 menit lalu coba browsing. Jika masih berkendala bisa informasikan nomor internet, atas nama dan nomor HP aktifnya via DM guna pengecekan lebih lanjut. (1/2)

stemming = 'pagi | kak | perihal | dahulu | tunggu | menit | coba | masih | bisa | nomor | nama | nomor | via | lebih | lanjut'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$$P(\text{positif}) = P(\text{bisa} | \text{positif})$$

$$= 0,018867925$$

$$= 0,018867925$$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$$P(\text{negatif}) = P(\text{bisa} | \text{negatif})$$

$$= 0,016393443$$

$$= 0,016393443$$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 52 ada dalam kelas

positif, karena nilai perhitungan dokumen uji pada kelas positif lebih besar

daripada nilai perhitungan dokumen uji pada kelas negatif.

Dokumen 53 = Terima kasih -Diar (2/2)

stemming = 'kasih'

Perhitungan berdasarkan hasil probabilitas positif dengan dokumen uji

$$P(\text{positif}) = P(\text{kasih} | \text{positif})$$

$$= 0,056603774$$

$$= 0,056603774$$

Perhitungan berdasarkan hasil probabilitas negatif dengan dokumen uji

$$P(\text{negatif}) = P(\text{kasih} | \text{negatif})$$

$$= 0,016393443$$

$$= 0,016393443$$

Dari hasil perhitungan di atas dapat diketahui bahwa dokumen 53 ada dalam kelas positif, karena nilai perhitungan dokumen uji pada kelas positif lebih besar daripada nilai perhitungan dokumen uji pada kelas negatif.

4.2.3. Pengujian Akurasi

Bagian ini menjelaskan tentang pengujian akurasi yang digunakan pada penelitian ini. Pengujian akurasi pada penelitian ini menggunakan metode *confusion matrix*. Pengujian akurasi menggunakan hasil dari perhitungan aplikasi dan perhitungan manual. Pengujian akurasi terdapat pada tabel L.4.16. tabel data pengujian akurasi data *testing* dan gambar L.4.26. *Confusion Matrix Data Testing* Lampiran 10. Pengujian Akurasi. Setelah dibuat tabel data pengujian akurasi data *testing*, dilakukan perhitungan akurasi dengan menggunakan rumus *confusion matrix*. Berikut adalah proses perhitungan akurasi.

Berdasarkan Rumus (3)

$$Akurasi = \frac{3 + 6}{3 + 6 + 1 + 3}$$

$$Akurasi = 0,69 \times 100\% = 69\%$$

4.3. Evaluasi

Bagian ini menjelaskan tentang evaluasi yang dilakukan pada penelitian. Evaluasi yang dilakukan terdiri dari Evaluasi Hasil Uji Fungsi, Evaluasi Uji Perhitungan, dan Evaluasi Uji Akurasi. Berikut adalah hasil pengujian yang ada pada penelitian.

4.3.1. Evaluasi Hasil Uji Fungsi

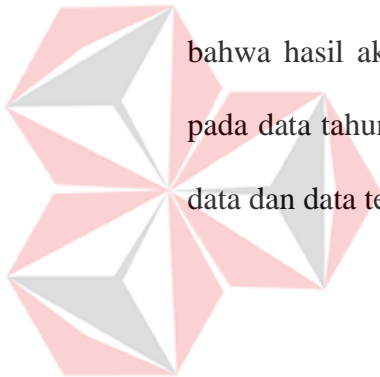
Bagian ini membahas mengenai evaluasi berdasarkan pada pengujian fungsi yang sudah dilakukan. Pengujian fungsi yang dilakukan berupa pengujian fungsi penambahan data, ekstraksi data, klasifikasi data, dan visualisasi data. Pengujian pada penambahan data dilakukan sebanyak 3 (tiga) kali, pengujian pada ekstraksi data dilakukan sebanyak 3 (tiga) kali, pengujian pada klasifikasi data dilakukan sebanyak 3 (tiga) kali, dan pengujian pada visualisasi data dilakukan sebanyak 1 (satu) kali. Dari 3 pengujian penambahan data, 3 pengujian berhasil dilakukan dengan tingkat keberhasilan 100%. Dari 3 pengujian ekstraksi data, 3 pengujian berhasil dilakukan dengan tingkat keberhasilan 100%. Dari 3 pengujian klasifikasi data, 3 pengujian berhasil dilakukan dengan tingkat keberhasilan 100%. Dari 1 pengujian visualisasi data, 1 pengujian berhasil dilakukan dengan tingkat keberhasilan 100%. Tingkat keberhasilan 100% berdasarkan pada jumlah pengujian yang dilakukan dibagi dengan jumlah keberhasilan pengujian.

4.3.2. Evaluasi Hasil Uji Hitungan

Bagian ini membahas mengenai evaluasi berdasarkan pada pengujian perhitungan yang sudah dilakukan. Pengujian perhitungan yang dilakukan berupa pengujian perhitungan aplikasi dan pengujian perhitungan manual. Berdasarkan pengujian perhitungan aplikasi dan pengujian perhitungan manual didapatkan hasil berupa nilai yang berbeda dengan nilai yang sama. Hasil tersebut terdapat pada tabel L.4.17. Data Perbandingan Hasil Uji Aplikasi Dan Uji Manual Lampiran 11. Evaluasi Hitungan.

4.3.3. Evaluasi Hasil Uji Akurasi

Bagian ini membahas mengenai evaluasi berdasarkan pada pengujian perhitungan akurasi yang sudah dilakukan. Pengujian perhitungan akurasi yang dilakukan berupa pengujian perhitungan menggunakan metode *confusion matrix*. Pengujian perhitungan akurasi menggunakan sumber data *mediakonsumen.com*, *twitter.com*, dan *detik.com*. Dari tabel data pengujian akurasi data *testing* didapatkan bahwa didapatkan *Actual Value / Predictive Value* berupa 3 *True Positive* (TP), 6 *True Negative* (TN), 1 *False Positive* (FP) dan 3 *False Negative* (FN). *Actual Value / Predictive Value* didapatkan dari hasil perhitungan aplikasi dan perhitungan manual. Berdasarkan perhitungan dengan uji akurasi didapatkan bahwa hasil akurasi perhitungan analisis sentimen yang didapatkan berdasarkan pada data tahun 2017 sampai dengan 2020 dari jumlah data training sebanyak 40 data dan data testing sebanyak 13 data didapatkan nilai sebesar 69%.



BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan dari hasil dan pembahasan yang ada pada penelitian ini, maka didapatkan kesimpulan sebagai berikut:

1. Analisis sentimen pada penelitian ini dapat menampilkan visualisasi dari hasil klasifikasi analisis sentimen dalam bentuk diagram lingkaran (*pie chart*).
2. Menghasilkan klasifikasi analisis sentimen yang dibuktikan dengan pengujian fungsi *Crawling* Data, Ekstraksi, Klasifikasi dan Visualisasi dengan nilai sebesar 100%, pengujian perhitungan dengan nilai yang berbeda dengan klasifikasi yang sama, dan pengujian akurasi pada data tahun 2017 sampai dengan 2020 dari jumlah data training sebanyak 40 data dan data testing sebanyak 13 data didapatkan nilai sebesar 69%.

5.2. Saran

Adapun saran-saran yang dapat menjadi masukan untuk penelitian-penelitian yang akan datang antara lain adalah:

1. Menambah data yang dilakukan pengujian untuk penelitian yang akan datang.
2. Mengembangkan dan menambah penggunaan data *training* dan data *testing* dari sumber yang berbeda untuk penelitian yang akan datang.
3. Mengembangkan algoritma atau metode lain untuk melakukan pengolahan data pada penelitian yang akan datang.
4. Meningkatkan akurasi dengan menambah data uji (*testing*).

DAFTAR PUSTAKA

- Baj-Rogowska, A. (2017). Sentiment Analysis of Facebook Posts: the Uber case. *The 8th IEEE International Conference on Intelligent Computing and Information Systems (ICICIS 2017)*, 1-5.
- Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook*. New York: Cambridge University Press.
- Gurusamy, V. (2014). RTRIC. *Preprocessing Techniques for Text Mining*, 1-6.
- Indonesia, PT. Telkom. (2019, April 29). *Apa Itu IndiHome?* Retrieved from IndiHome: <https://indihome.co.id/pusat-bantuan/kenali-indihome/apa-itu-indihome>
- Kadhim, A. I., Cheah, Y.-N., & Ahamed, N. H. (2014). 2014 4th International Conference on Artificial Intelligence with Applications in Engineering and Technology. *Text Document Preprocessing and Dimension Reduction Techniques for Text Document Clustering*, 1-5.
- Kosara, R. (2007). 2007 11th International Conference Information Visualization (IV '07). *Visualization Criticism - The Missing Link Between Information Visualization and Art*, 1-6.
- Langgeni, D. P. (2010). Seminar Nasional Informatika 2010 (semnasIF 2010) ISSN: 1979-2328. *CLUSTERING ARTIKEL BERBAHASA INDONESIA MENGGUNAKAN UNSUPERVISED FEATURE SELECTION*, 1-10.
- Liu, B. (2012). Sentiment Analysis and Opinion Mining. In B. Liu, *Sentiment Analysis and Opinion Mining*. Morgan & Claypool.
- Manning, C. D., Raghavan, P., & Schütze, H. (2009). *An Introduction To Information Retrieval*. Cambridge: Cambridge University Press.
- Saleh, A. (2015). Implementasi Metode Klasifikasi Naïve Bayes. *Citec Journal*, Vol. 2, No. 3, Mei 2015 – Juli 2015, 1-12.
- Saragih, M. H. (2017). 2017 International Conference on Sustainable Information Engineering and Technology (SIET). *Sentiment Analysis of Customer Engagement on Social Media in Transport Online*, 1-6.
- Shrivastava, V. (2018). ISSN : 2248-9622 Vol. 8, Issue 11 (Part -I). *A Methodical Study of Web Crawler*, 1-8.
- Srividhya, V. (2010). International Journal of Computer Science and Application Issue 2010 ISSN 0974-0767. *Evaluating Preprocessing Techniques in Text Categorization*, 1-3.

Sukanto, & Shalahuddin, M. (2013). *Rekayasa Perangkat Lunak Terstruktur Dan Berorientasi Objek*. Bandung: Informatika, 1-9.

Wahono, R. S. (2019, August 22). *Research Methodology*. Retrieved from romisatriawahono: <https://romisatriawahono.net>

Weiss, S. M., Indurkha, N., Zhang, T., & Damerou, F. J. (2005). *Text Mining : Predictive Methods for Analyzing Unstructured Information*. New York: Springer.



UNIVERSITAS
Dinamika