KLASIFIKASI KOMENTAR PADA PEMBELAJARAN E-LEARNING MENGGUNAKAN ANALISIS SENTIMEN DENGAN METODE K-NEAREST NEIGHBOR

SKRIPSI

Untuk memenuhi sebagian persyaratan   
memperoleh gelar Sarjana Komputer

Disusun oleh:

Hanif Prasetyo Maulidina

NIM: 145150207111134



PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS BRAWIJAYA

MALANG

2018

MALANG

PENGESAHAN

JUDUL SKRIPSI

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan

memperoleh gelar Sarjana Komputer

Disusun Oleh :

Nama Mahasiswa

NIM: 123456789

Skripsi ini telah diuji dan dinyatakan lulus pada

2 Januari 2015

Telah diperiksa dan disetujui oleh:

|  |  |
| --- | --- |
| Dosen Pembimbing I  Nama Dosen Pembimbing I  NIK: 123456789  /\*jika terdapat NIK saja\*/ | Dosen Pembimbing II  Nama Dosen Pembimbing II  NIK: -  /\*jika tidak terdapat NIP, NIK, atau keduanya\*/ |

Mengetahui

Ketua Jurusan **NamaJurusan**

Contoh: Ketua Jurusan **Informatika**

Nama Ketua Jurusan

NIP: 123456789

/\*jika terdapat NIP\*/

PERNYATAAN ORISINALITAS

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disitasi dalam naskah ini dan disebutkan dalam daftar pustaka.

Apabila ternyata didalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur plagiasi, saya bersedia skripsi ini digugurkan dan gelar akademik yang telah saya peroleh (sarjana) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU No. 20 Tahun 2003, Pasal 25 ayat 2 dan Pasal 70).

Malang, 1 Januari 2015



­

Nama Mahasiswa

NIM: 123456789

KATA PENGANTAR

Bagian ini memuat pernyataan resmi untuk menyampaikan rasa terima kasih penulis kepada berbagai pihak yang telah membantu penyelesaian skripsi ini. Nama-nama penerima ucapan terima kasih sebaiknya dituliskan lengkap, termasuk gelar akademik, dan pihak-pihak yang tidak terkait dihindari untuk dituliskan. Bahasa yang digunakan seharusnya mengikuti kaidah bahasa Indonesia yang baku. Kata pengantar boleh diakhiri dengan paragraf yang menyatakan bahwa penulis menerima kritik dan saran untuk pengembangan penelitian selanjutnya. Terakhir, kata pengantar ditutup dengan mencantumkan kota dan tanggal penulisan kata pengantar, lalu diikuti dengan kata “Penulis”.

Malang, 1 Januari 2015

Penulis

email@domain.com

ABSTRAK

Bagian ini diisi dengan abstrak dalam Bahasa Indonesia. Abstrak adalah uraian singkat (umumnya 200-300 kata) yang merupakan intisari dari sebuah skripsi. Abstrak membantu pembaca untuk mendapatkan gambaran secara cepat dan akurat tentang isi dari sebuah skripsi. Melalui abstrak, pembaca juga dapat menentukan apakah akan membaca skripsi lebih lanjut. Oleh karena itu, abstrak sebaiknya memberikan gambaran yang padat tetapi tetap jelas dan akurat tentang (1) apa dan mengapa penelitian dikerjakan: sedikit latar belakang, pertanyaan atau masalah penelitan, dan/atau tujuan penelitian; (2) bagaimana penelitian dikerjakan: rancangan penelitian dan metodologi/metode dasar yang digunakan dalam penelitian; (3) hasil penting yang diperoleh: temuan utama, karakteristik artefak, atau hasil evaluasi artefak yang dibangun; (4) hasil pembahasan dan kesimpulan: hasil dari analisis dan pembahasan temuan atau evaluasi artefak yang dibangun, yang dikaitkan dengan pertanyaan/tujuan penelitian.

Yang harus dihindari dalam sebuah abstrak diantaranya (1) penjelasan latar belakang yang terlalu panjang; (2) sitasi ke literatur lainnya; (3) kalimat yang tidak lengkap; (3) singkatan, jargon, atau istilah yang membingungkan pembaca, kecuali telah dijelaskan dengan baik; (4) gambar atau tabel; (5) angka-angka yang terlalu banyak.

Di akhir abstrak ditampilkan beberapa kata kunci (normalnya 5-7) untuk membantu pembaca memposisikan isi skripsi dengan area studi dan masalah penelitian. Kata kunci, beserta judul, nama penulis, dan abstrak biasanya dimasukkan dalam basis data perpustakaan. Kata kunci juga dapat diindeks dalam basis data sehingga dapat digunakan untuk proses pencarian tulisan ilmiah yang relevan. Oleh karena itu pemilihan kata kunci yang sesuai dengan area penelitian dan masalah penelitian cukup penting. Pemilihan kata kunci juga bisa didapatkan dari referensi yang dirujuk.

Kata kunci: abstrak, skripsi, intisari, kata kunci, artefak

ABSTRACT

The absract of your skripsi in English is written here.

DAFTAR ISI

[PENGESAHAN ii](#_Toc428800273)

[PERNYATAAN ORISINALITAS iii](#_Toc428800274)

[KATA PENGANTAR iv](#_Toc428800275)

[ABSTRAK v](#_Toc428800276)

[ABSTRACT vi](#_Toc428800277)

[DAFTAR ISI vii](#_Toc428800278)

[DAFTAR TABEL ix](#_Toc428800279)

[DAFTAR GAMBAR x](#_Toc428800280)

[DAFTAR LAMPIRAN xi](#_Toc428800281)

[BAB 1 PENDAHULUAN 1](#_Toc428800282)

[1.1 Latar belakang 1](#_Toc428800283)

[1.2 Rumusan masalah 1](#_Toc428800284)

[1.3 Tujuan 4](#_Toc428800285)

[1.4 Manfaat 5](#_Toc428800286)

[1.5 Batasan masalah 6](#_Toc428800287)

[1.6 Sistematika pembahasan 6](#_Toc428800288)

[BAB 2 LANDASAN KEPUSTAKAAN 7](#_Toc428800289)

[2.1 Sub bab dua satu 7](#_Toc428800290)

[2.1.1 Seksi dua satu satu 7](#_Toc428800291)

[2.1.2 Seksi dua satu dua 7](#_Toc428800292)

[2.2 Sub bab dua dua 8](#_Toc428800293)

[2.2.1 Seksi dua dua satu tentang persamaan 8](#_Toc428800294)

[2.2.2 Seksi dua dua dua tentang tabel 8](#_Toc428800295)

[2.2.3 Gambar 10](#_Toc428800296)

[2.2.4 Lambang, satuan, dan singkatan 11](#_Toc428800297)

[2.2.5 Seksi dua dua satu tentang sitasi tabel dan gambar 12](#_Toc428800298)

[2.2.6 Seksi dua dua dua 13](#_Toc428800299)

[BAB 3 METODOLOGI 14](#_Toc428800300)

[3.1 Sub bab tiga satu 14](#_Toc428800301)

[3.1.1 Seksi tiga satu satu 14](#_Toc428800302)

[3.1.2 Seksi tiga satu dua 15](#_Toc428800303)

[3.2 Sub bab tiga dua 15](#_Toc428800304)

[BAB 4 HASIL 16](#_Toc428800305)

[4.1 Sub bab dua satu 16](#_Toc428800306)

[4.2 Sub bab dua dua 16](#_Toc428800307)

[4.2.1 Seksi empat dua satu 16](#_Toc428800308)

[4.2.2 Seksi empat dua dua 17](#_Toc428800309)

[4.3 Sub bab empat tiga 17](#_Toc428800310)

[4.3.1 Contoh struktur penelitian implementatif pengembangan 17](#_Toc428800311)

[4.3.2 Contoh struktur penelitian non-implementatif 18](#_Toc428800312)

[BAB 5 PEMBAHASAN 19](#_Toc428800313)

[5.1 Sub bab lima satu 19](#_Toc428800314)

[5.1.1 Seksi lima satu satu 19](#_Toc428800315)

[5.1.2 Seksi lima satu dua 19](#_Toc428800316)

[5.2 Seksi lima dua 20](#_Toc428800317)

[5.2.1 Seksi lima dua satu 20](#_Toc428800318)

[5.2.2 Seksi lima dua dua 20](#_Toc428800319)

[5.3 Sub bab lima tiga 20](#_Toc428800320)

[5.3.1 Contoh struktur penelitian implementatif pembangunan 20](#_Toc428800321)

[5.3.2 Contoh struktur penelitian non-implementatif eksperimental 22](#_Toc428800322)

[BAB 6 Penutup 23](#_Toc428800323)

[6.1 Kesimpulan 23](#_Toc428800324)

[6.2 Saran 23](#_Toc428800325)

[DAFTAR PUSTAKA 24](#_Toc428800326)

[LAMPIRAN A PERSYARATAN FISIK DAN TATA LETAK 27](#_Toc428800327)

[LAMPIRAN B PENGGUNAAN BAHASA 29](#_Toc428800328)

DAFTAR TABEL

[Tabel ‎2.1 Pembentukan bilangan random untuk Indeks Masa Tubuh (IMT) 8](#_Toc428800329)

[Tabel ‎2.2 Contoh tabel 2 9](#_Toc428800330)

DAFTAR GAMBAR

[Gambar ‎2.1 Pengaruh nilai K terhadap akurasi 11](#_Toc428800331)

DAFTAR LAMPIRAN

[LAMPIRAN A PERSYARATAN FISIK DAN TATA LETAK 27](#_Toc428800332)

[A.1 Kertas 27](#_Toc428800333)

[A.2 Margin 27](#_Toc428800334)

[A.3 Jenis dan ukuran huruf 27](#_Toc428800335)

[A.4 Spasi 27](#_Toc428800336)

[A.5 Kepala bab, sub bab, dan seksi 27](#_Toc428800337)

[A.6 Nomor halaman 28](#_Toc428800338)

[LAMPIRAN B PENGGUNAAN BAHASA 29](#_Toc428800339)

# PENDAHULUAN

## Latar belakang

Pembelajaran merupakan proses atau interaksi antara peserta didik dengan pendidik, dengan tujuan untuk menyampaikan atau menyalurkan ilmu pengetahuan dan saling bertukar informasi pada lingkungan belajar mengajar. Proses belajar mengajar sendiri terdapat seorang pendidik, atau biasa disebut guru, maupun peserta didik atau yang biasa disebut sebagai siswa. Pada proses pembelajaran, pendidik mengajar peserta didik agar mencapai objektif yang ditentukan (aspek kognitif), juga dapat mempengaruhi perubahan sikap (aspek afektif), serta aspek keterampilan (aspek psikomotor).

Dewasa ini, proses belajar mengajar dapat dilakukan dengan berbagai macam proses dan media. Menurut Syah (1998) proses belajar mengajar harus melibatkan peserta didik secara totalitas, dalam arti melibatkan pikiran, penglihatan, pendengaran, dan psikomotor. Dalam proses mengajar seorang guru harus mengajak siswa untuk mendengarkan, menyajikan media yang dapat dilihat, memberi kesmpatan untuk menulis dan mengajukan pertanyaan atau tanggapan sehingga terjadi dialog kreatif yang menunjukan proses belajar mengajar yang interaktif.

Untuk saat ini, metode pembelajaran interaktif adalah metode yang sedang berkembang dengan pesat, dimana metode ini meminimalisir tatap muka antara pendidik dan peserta didik. Menurut Faire & cosgrove dalam *Harlem* (1992), Model pembelajaran interaktif sering dikenal dengan nama pendekatan pertanyaan anak. Model ini dirancang agar siswa akan bertanya dan menemukan jawaban mereka sendiri. Pembelajaran interaktif tersebut dapat menggunakan berbagai media. Media yang digunakan bisa berupa media interaktif online, seperti *website*, maupun sebuah aplikasi.

Pada setiap pelajaran seorang pengajar akan membutuhkan *feedback* berupa komentar dari peserta didik. Pada pembelajaran interaktif sendiri, *feedback* yang diberikan adalah berupa text komentar, yang dituliskan pada akhir setiap pembelajaran. Komentar tu sendiri terdapat berbagai macam, yang dapat dibagi menjadi komentar positif dan negatif. Untuk itu, dibutuhkan suatu metode yang dapat menyaring antara komentar atau *feedback* baik yang positif maupun yang negatif.

Ada beberapa metode yang dapat digunakan untuk mengklasifikasikan dokumen berupa text, agar dapat dilakukan analisis sentimen. Diantaranya adalah metode *Support Vector Machine* (SVM), *Naive Bayes,*dan *K-Nearest Neighbor* (K-NN). Untuk penelitian ini, metode yang digunakan adalah metode K-NN. *K-Nearest Neighbor*  atau KNN merupakan metode yang melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang kemiripannya paling dekat dengan objek tersebut. (Hardiyanto & Rahutomo, 2016).

Penelitian sebelumnya yang dilakukan oleh Viswanath Bijalwan, Vinay Kumar, Pinki Kumar, dan Jordan Pasqua pada tahun 2014 yang berjudul *K-NN based Machine Learning Approach for Text and Document Mining*  menunjukkan bahwa metode *K-Nearest Neighbor* (K-NN) menghasilkan nilai akurasi tertinggi ketika dibandingkan dengan *Naive Bayes*  dan *Term Graph* dengan rata-rata K-NN sebesar 98,95%, *Naive Bayes* 62,66%, dan *Term Graph* 98,72% , sehingga menurut mereka metode K-NN lebih cocok digunakan untuk melakukan *mining*  pada dokumen text.

Penelitian yang dilakukan oleh Nurjanah,W., yang berjudul *Analisis* *Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter Menggunakan Metode K-Nearest Neighbor (KNN) dan pembobotan jumlah Retweet menggunakan Metode KNN untuk mengklasifikasi komentar positif maupun negatif pada media sosial Twitter*. Penelitian tersebut menghasilkan keluaran output berupa klasifikasi apakah dokumen bernilai positif atau negatif, dengan tingkat akurasi mencapai 80.83%, dimana hasil tersebut menunjukkan tingkat akurasi yang tinggi.

Penelitian lain dilakukan oleh Mentari,N., dengan judul *Analisis Sentimen Kurikulum 2013 Pada Sosial Media Twitter Menggunakan Metode K-Nearest Neighbor dan Feature Selection Query Expansion Ranking* bertujuan untuk analisis sentimen opini masyarakat tentang kurikulum 2013. Metode yang digunakan dalam penelitian ini adalah *K-Nearest Neighbor* , dan dari penelitian tersebut didapatkan hasil akurasi terbaik sebesar 96,36% pada presentase *feature selection* yang digunakan sebesar 50%.

Dari kedua penelitian tersebut, dapat diasumsikan jika metode K-NN memiliki tingkat akurasi yang tinggi sehingga pada penelitian ini akan menggunakan Algoritma K-NN untuk pengklasifikasian komentar pada pembelajaran dengan menggunakan analisis sentimen*.*

## Rumusan masalah

Berdasarkan latar belakang diatas maka penulis merumuskan masalah yang akan dibahas adalah :

1. Bagaimana implementasi analisis sentimen dengan metode *K-Nearest Neighbor* pada klasifikasi komentar pada pembelajaran
2. Bagaimana kinerja dari metode analisis sentimen dengan metode *K-Nearest Neighbor* pada pengklasifikasian komentar pembelajaran

## Tujuan

Berdasarkan rumusan masalah diatas, maka tujuan dari penelitian ini adalah

1.implementasi analisis sentimen untuk klasifikasi komentar pada pembelajaran

2. mengetahui kinerja dari metode analisis sentimen dengan menggunakan metode *K-Nearest Neighbor*  pada pengklasifikasian komentar pembelajaran.

## Manfaat

Adapun manfaat yang diharapkan dalam penelitian ini sebagai berikut :

* Bagi penulis : - Mempraktikkan ilmu yang didapat selama perkuliahan khususunya dibidang sistem komputasi cerdas
* Membantu pengajar dalam menerima *feedback*  dari sistem yang telah dikembangkan
* Bagi pembaca : Sebagai bahan dan acuan dalam mengembangkan sistem yang memiliki keterkaitan dengan penelitian.

## Batasan masalah

Batasan masalah pada penelitian ini meliputi :

1.komentar yang digunakan berasal dari web pembelajaran online.

2.penerapan dari metode ini sebatas hingga klasifikasi text dokumen.

## Sistematika pembahasan

Sistematika penyusunan laporan ditunjukkan untuk memberikan gambaran dan uraian dari laporan skripsi secara garis besar yang meliputi beberapa bab sebagai berikut.

**BAB I : Pendahuluan**

Menguraikan latar belakang, identifikasi masalah, rumusan masalah, batasan masalah, tujuan dan manfaat skripsi serta sistematika penyusunan laporan skripsi.

**BAB II : Tinjauan Pustaka**

Menguraikan teori-teori yang menjadi referensi dan sebagai pendukung dalam pelaksanaan skripsi.

**BAB III : Metode Penelitian**

Menguraikan tentang metode dan langkah kerja yang dilakukan dalam proses perancangan dan implementasi pada pelaksanaan skripsi dan perancangan sistem yang menjadi objek studi kasus skripsi.

**BAB IV : Perancangan**

Menguraikan proses komputasi cerdas dalam pembuatan sistem kecerdasan buatan

**BAB V : Implementasi**

Menguraikan proses implementasi dari dasar teori yang telah dipelajari sesuai hasil analisis dan perancangan sistem.

**BAB VI : Pengujian dan Analisis**

Menguraikan proses pengujian, hasil pengujian, serta kendala yang dihadapi selama pengujian sistem.

**BAB VII : Kesimpulan dan Saran**

Menguraikan kesimpulan yang diperoleh dari pembuatan dan pengujian sistem kecerdasan yang dikembangkan serta saran-saran untuk pengembangan lebih lanjut.

# LANDASAN KEPUSTAKAAN

Pada bab ini akan dibahas mengenai studi kepustakaan yang digunakan pada penelitian ini. Bab ini akan berisi penilitian penelitian yang dilakukan terdahulu dan metode terkait dalam penelitian ini.

## Kajian pustaka

Beberapa penelitian telah dilakukan yang berkaitan dengan analisis sentimen*.* Penelitian yang dilakukan oleh Nurul Dyah Mentari dengan judul *Analisis Sentimen Kurikulum 2013 Pada Sosial Media Twitter Menggunakan Metode K-Nearest Neighbor dan Feature Selection Query Expansion Ranking* bertujuan untuk analisis sentiment opini masyarakat tentang kurikulum 2013. Metode yang digunakan dalam penelitian ini adalah *K-Nearest Neighbor* , dan dari penelitian tersebut didapatkan hasil akurasi terbaik sebesar 96,36% pada presentase *feature selection* yang digunakan sebesar 50%.

Penelitian lain dilakukan oleh Viswanath Bijalwan, Vinay Kumar, Pinki Kumar, dan Jordan Pasqua pada tahun 2014 yang berjudul *K-NN based Machine Learning Approach for Text and Document Mining*  menunjukkan bahwa metode *K-Nearest Neighbor* (K-NN) menghasilkan nilai akurasi tertinggi ketika dibandingkan dengan *Naive Bayes*  dan *Term Graph* dengan rata-rata K-NN sebesar 98,95%, *Naive Bayes* 62,66%, dan *Term Graph* 98,72% , sehingga menurut mereka metode K-NN lebih cocok digunakan untuk melakukan *mining*  pada dokumen text.

Penelitian selanjutnya oleh Nurjanah,W., yang berjudul *Analisis* *Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter Menggunakan Metode K-Nearest Neighbor (KNN) dan pembobotan jumlah Retweet menggunakan Metode KNN untuk mengklasifikasi komentar positif maupun negatif pada media sosial Twitter*. Penelitian tersebut menghasilkan keluaran output berupa klasifikasi apakah dokumen bernilai positif atau negatif, dengan tingkat akurasi mencapai 80.83%, dimana hasil tersebut menunjukkan tingkat akurasi yang tinggi

## Pembelajaran

Pembelajaran adalah suatu sistem, yang bertujuan untuk membantu proses belajar siswa, yang berisi serangkaian peristiwa yang dirancang, disusun sedemikian rupa untuk mempengaruhi dan mendukung terjadinya proses belajar siswa yang bersifat internal. (Gagne dan Brigs, 1979). Pembelajaran merupakan gabungan dari dua aktivitas, yaitu aktivitas mengajar dan aktivitas belajar. Istilah pembelajaran sama dengan *instruction* atau pengajaran. Pengajaran mempunyai arti cara mengajar. Kegiatan belajar mengajar merupakan satu kesatuan dari dua kegiatan yang searah.

Dalam proses belajar mengajar akan terjadi interaksi antara peserta didik dan pendidik. Peserta didik anak didik adalah salah satu komponen manusiawi yang menempati posisi sentral dalam proses belajar-mengajar, sedang pendidik adalah salah satu komponen manusiawi dalam proses belajar-mengajar, yang ikut berperan dalam usaha pembentukan sumber daya manusia yang potensial dibidang pembangunan. (Slameto, 2003). Dalam proses pembelajaran, terdapat 2 metode umum yang diterapkan, yaitu pembelajaran konvensional dan interaktif.

### Pembelajaran Konvensional

 Pembelajaran ini merupakan jenis pembelajaran yang paling banyak digunakan di Indonesia di mana seorang tenaga pendidik menyampaikan materi pembelajaran kepada para peserta didiknya secara langsung. Di sini memiliki ciri utama yakni tenaga pendidik memiliki peranan yang lebih dominan dalam kegiatan

### Pembelajaran Interaktif

Perbedaan utama antara pembelajaran interaktif ini dengan pembelajaran ceramah adalah terletak pada penekanan interaksi yang lebih ditekankan antara tenaga pendidik kepada para peserta didiknya. Di sini tenaga pendidik tetap menyampaikan materi pembelajaran seperti pada jenis pembelajaran ceramah, tetapi ada kalanya memberikan pertanyaan atau memberikan kesempatan kepada para peserta didik untuk menyampaikan pendapatnya. Sehingga diharapkan para peserta didik akan lebih aktif selama kegiatan pembelajaran berlangsung. Anderson (1976) mengelompokkan media yang dapat digunakan pada pembelajaran menjadi 10 golongan berikut :

|  |  |  |
| --- | --- | --- |
| **No** | **Golongan Media** | **Contoh dalam Pembelajaran** |
| I | Audio | Kaset audio, siaran radio, CD, telepon |
| II | Cetak | Buku pelajaran, modul, brosur, leaflet, gambar |
| III | Audio-cetak | Kaset audio yang dilengkapi bahan tertulis |
| IV | Proyeksi visual diam | Overhead transparansi (OHT), Film bingkai (slide) |
| V | Proyeksi Audio visual diam | Film bingkai (slide) bersuara |
| VI | Visual gerak | Film bisu |
| VII |  | Audio Visual gerak, film gerak bersuara, video/VCD, televisi |
| VIII | Obyek fisik | Benda nyata, model, specimen |
| IX | Manusia dan lingkungan | Guru, Pustakawan, Laboran |
| X | Komputer | CAI (Pembelajaran berbantuan komputer), CBI (Pembelajaran berbasis komputer).[7] |

Tabel 2.1 Daftar pengelompokan media pembelajaran

Salah satu metode pembelajarn interaktif adalah menggunakan basis *e-learning*. E-learning adalah suatu sistem atau konsep pendidikan yang memanfaatkan teknologi informasi dalam proses belajar mengajar. Karakteristik E-learning menurut Nursalam (2008:135) adalah:

1. Memanfaatkan jasa teknologi elektronik.
2. Memanfaatkan keunggulan komputer (digital media dan komputer networks)
3. Menggunakan bahan ajar yang bersifat mandiri (self learning materials) kemudian disimpan di komputer, sehingga dapat diakses oleh doesen dan mahasiswa kapan saja dan dimana saja.
4. Memanfaatkan jadwal pembelajaran, kurikulum, hasil kemajuan belajar, dan hal-hal yang berkaitan dengan administrasi pendidikan dapat dilihat setiap saat di komputer.

## Analisis Sentimen

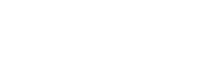
*Analisis sentimen* atau biasa disebut opinion mining merupakan salah satu cabang penelitian Text Mining. Opinion mining adalah riset komputasional dari opini, sentimen dan emosi yang diekspresikan secara tekstual. Jika diberikan suatu set dokumen teks yang berisi opini mengenai suatu objek, maka opinion mining bertujuan untuk mengekstrak atribut dan komponen dari objek yang telah dikomentasi pada setiap dokumen dan untuk menentukan apakah komentar tersebut bermakna positif atau negative . Analisis sentimen dapat dibedakan berdasarkan sumber datanya, beberapa level yang sering digunakan dalam penelitian Analisis sentimen adalah Analisis sentimen pada level dokumen dan Analisis sentimen pada level kalimat. Berdasarkan level sumber datanya Analisis sentimen terbagi menjadi 2 kelompok besar yaitu : Coarse-grained Analisis sentimen dan fined-grained Analisis sentimen Pada Analisis sentimen Coarse-grained, Analisis sentimen yang dilakukan adalah pada level dokumen. Secara garis besar fokus utama dari Analisis sentimen jenis ini adalah menganggap seluruh isi dokumen sebagai sebuah sentiment positif atau sentiment negatif. Fined-grained Analisis sentimen adalah Analisis sentimen pada level kalimat. Fokus utama fined-greined Analisis sentimen adalah menentukan sentimen pada setiap kalimat.

Dalam sebuah penelitian berjudul berjudul “*Affective-feature-based Analisis sentimen using SVM Classifier*” dilakukan oleh Luo et al., sentimen analisis bertujuan menemukan orientasi emosi dari ulasan pengguna secara otomatis. *Analisis sentimen* merupakan versi lebih kompleks dari *topic mining*, *analisis sentimen*  merupakan klasifikasi biner dengan pembagian klasifikasiannya menjadi dua jenis kelas yang pertama adalah positif dan yang kedua adalah negatif (Luo et al., 2016). Gambar 2.1 menunjukkan diagram alir yang diambil dari penelitian Luo et al.

SESI LATIH SESI UJI



|  |  |
| --- | --- |
| Data Latih  *Text* *pre-processing* | |
| Pembangkitan Fitur (Kata Unik) |  |
|  | |
| Pembobotan Fitur (Kata Unik) |  |
|  | |
| Penyeleksian Fitur (Kata Unik) |  |
| Representasi teks  Metode Klasifikasi | |



|  |  |
| --- | --- |
| Data Uji  *Text* *pre-processing* | |
| Pembangkitan Fitur (Kata Unik) |  |
|  | |
| Pembobotan Fitur (Kata Unik) |  |
|  | |
| Penyeleksian Fitur (Kata Unik) |  |
| Representasi teks  Terklasifikasi  Hasil Klasifikasi | |



Gambar 2.1 Diagram Alir proses analisis sentimen

Sumber : [Luo et al., 2016]

## *Text Mining*

Text mining dapat didefinisikan secara luas sebagai proses pengetahuan intensif dimana pengguna berinteraksi dengan koleksi dokumen dari waktu ke waktu dengan menggunakan separangkat alat analisis. Text mining berusaha mengekstrak informasi yang berguna dari sumber data melalui identifikasi dan eksplorasi pola yang menarik. Text mining cenderung mengarah pada bidang penelitian data mining. Oleh karena itu, tidak mengherankan bahwa text mining dan data mining berada pada tingkat arsitektur yang sama. Penambangan teks dapat dianggap sebagai proses dua tahap yang diawali dengan penerapan struktur terhadap sumber data teks dan dilanjutkan dengan ekstraksi informasi dan pengetahuan yang relevan dari data teks terstruktur ini dengan menggunakan teknik dan alat yang sama dengan penambangan data .

## *K-Nearest Neighbor*

*K-Nearest Neighbor*  atau K-NN adalah sebuah algoritma untuk melakukan klasifikas terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Kasus khusus di mana klasifikasi diprediksikan berdasarkan data pembelajaran yang paling dekat (dengan kata lain, k = 1) disebut algoritma nearest neighbor. Algoritma nearest neighbor berdasarkan pada proses pembelajaran menggunakan analogi / learning by analogi. Training sampelnya dideskripsikan dalam bentuk atribut numerik n-dimensi. Tiap sampel mewakili sebuah titik pada ruang n-dimensi. Dengan cara ini, semua training sampel disimpan pada pola ruang n-dimensi. Ketika diberikan “unknown” sampel, K-NN classifier mencari pola ruang K training sampel yang paling dekat “unknown” sampel tersebut. K training sampel ini adalah k nearest neighbor dari unknown sampel. Unknown sampel ditetapkan dengan class yang paling umum diantara k nearest neighborsnya. Ketika k = 1, unknown sampel ditetapkan dengan class dari training sampel yang paling dekat dengan pola ruangnya. Dalam pengklasifikasian teks, semakin besar nilai *cosine similarity* akan semakin dekat tingkat kemiripan antara data uji dan data latihnya dan sebaliknya jika nilai *cosine similarity*-nya semakin kecil maka akan semakin jauh tingkat kemiripan antara data uji dan data latihnya (Luhulima, 2015). Perhitungan *cosine similarity* adalah dengan persamaan :



*A* merupakan data uji, dan *B* merupakan data latih. *Ai* dan *Bi* merupakan bobot nilai yang diberikan untuk setiap *term* yang ada.

## Sistem Temu Kembali Informasi

Sistem Temu Kembali Informasi atau *Information Retrieval System* adalah sebuah media layanan bagi pengguna untuk memperoleh informasi yang diperlukan bagi pengguna. Dalam ODLIS, dijelaskan bahwa Temu Kembali Informasi adalah proses, metode, dan prosedur yang digunakan untuk menyeleksi informasi yang relevan yang tersimpan di database. Model sistem temu kembali informasi yaitu meliputi representasi dokume maupun *query*, fungsi pencarian (*retrieval function*) dan notasi kesesuaian (*relevance notation*) dokumen terhadap *query*. Salah satu model sistem temu kembali informasi yang paling awal digunakan adalah model *boolean*. Model Boolean mempresentasikan dokumen sebagai suatu himpunan kata-kunci. Sedangkan *query* direpresentasikan sebagai ekspresi *boolean*. *Query* dalam ekspresi *boolean* merupakan kumpulan kata kunci yang saling dihubungkan melalui operator *Boolean* seperti AND,OR dan NOT serta menggunakan tanda kurung untuk menentukan scope operator. Hasil pencarian dokumen dari model boolean adalah himpunan dokumen yang relevan.

QUERY

TEXT OPERATION

QUERY INFORMATION

Doc. 1

Doc. 2

Doc. 3

……

RANKING

TEXT OPERATION

INDEXING

COLLECTION INDEX

RANKED DOCUMENT

DOCUMENT COLLECTION

Gambar 2.2.2 Model Sistem Temu Kembali Informasi

***Text Operations*** (operasi terhadap teks) yang meliputi pemilihan kata- kata dalam query maupun dokumen dalam pentransformasian dokumen atau *query* menjadi *terms index* (indeks dari kata-kata).

***Query Formulation*** (formulasi terhadap query) yang memberi bobot pada indeks kata-kata query.

***Ranking***, mencari dokumen-dokumen yang relevan terhadap *query* dan mengurutkan dokumen tersebut berdasarkan kesesuaiannya dengan *query*.

***Indexing***, membangun data indeks dari koleksi dokumen. Dilakukan terlebih dahulu sebelum pencarian dokumen, sistem temu balik informasi menerima query dari pengguna, kemudian melakukan perangkingan terhadap pada koleksi berdasarkan kesesuaiannya dengan query. Hasil perangkingan yang diberikan kepada pengguna merupakan dokumen yang sistem, relevan dengan query, namun relevansi dokumen terhadap suatu query merupakan penilaian pengguna yang subjektif dan dipengaruhi banyak facto

## *Pre Processing*

### *Tokenizing*

Parsing / Tokenizing Teks adalah data unstructured yang harus dirubah dahulu menjadi terstuktur sebelum dianalis lebih lanjut. Teks dimasukan kedalam aplikasi yang disimpan kedalam array 1 dimensi. Kata – kata dalam kalimat dibagi berdasarkan spasi.

Manajemen pengetahuan adalah sebuah konsep baru di dunia bisnis

**Manajemen**

**Pengetahuan**

**Adalah**

**Sebuah**

**Konsep**

**Baru**

**Di**

**Dunia**

**bisnis**

Gambar 2.3 Contoh Tokenizing

### *Stopword Removal*

Stopword removal Setelah proses tokenizing setiap kata menjadi berdiri sendiri / tidak terikat dengan kata yang lain. Akibat dari pemisahlan kata tersebut, akan ada kata yang tidak memiliki arti yang relevan untuk menentukan ciri dari dokumen yang di tokenizing seperti “ini, itu, adalah, dan, atau” dan bayak lagi kata – kata sejenis. Kata – kata yang tidak memiliki arti yang relevan tersebut disebut stop word. Kumpulan dari stop word disebut stop list dan proses untuk menghapus stop word dalam dokumen disebut stopword removal.

**Manajemen**

**Pengetahuan**

**Adalah**

**Sebuah**

**Konsep**

**Baru**

**Di**

**Dunia**

**bisnis**

**Manajemen**

**Pengetahuan**

**Konsep**

**Baru**

**Dunia**

**bisnis**

Gambar 2.4 Contoh *Stopword Removal*

### *Stemming*

Stemming Stemming adalah proses pemetaan dan penguraian berbagai bentuk (variants) dari suatu kata menjadi bentuk kata dasarnya. Stemming Bahasa Indonesia telah banyak dikembangkan oleh para peniliti, salah satunya adalah *library* Sastrawi *stemming* yang ada pada Github. Penelitian yang dilakukan oleh Agusta (2009), melakukan perbandingan kinerja algoritme Nazief dan Adriani dengan algoritme Porter. Hasilnya menyatakan algoritme Nazief dan Adriani memiliki keakuratan yang lebih tinggi. Gambar 2.3 mencontohkan proses *Stemming.*

**Pengetahuan**

**bacaan**

**dipermasalahkan**

**kesayangan**

**tahu**

**baca**

**masalah**

**sayang**

Gambar 2.5 Contoh Proses  *Stemming*

## *Term Weighting* (TF-IDF)

Pembobotan kata (*term*) bertujuan untuk memberikan bobot pada setiap kata (*term*) yang terdapat pada dokumen teks yang akan diproses.

### *Term Frequency* (TF)

*Term Frequency* merupakan frekuensi kemunculan kata pada suatu dokumen teks. *Term Frequency* (*tft,d*) didefinisikan jumlah kemunculan term t pada dokumen d. Persamaan dari *Term Frequency* (*tft,d*) ditunjukkan pada Persamaan berikut :

𝑇𝐹 = 1 + log(𝐷) (2.1)

Keterangan :

*D* adalah kemunculan *term* pada dokumen

### *Document Frequency* (DF)

*Document Frequency* merupakan kata-kata yang banyak terdapat pada dokumen, kata tersebut tidak informatif, seperti kata dan, di, atau, bisa, merupakan.

### *Invers Document Frequency* (IDF)

*Invers Document Frequency* merupakan frekuensi kemunculan term pada keseluruhan dokumen teks. *Term* yang jarang muncul pada keseluruhan dokumen teks memiliki nilai *Invers Document Frequency* lebih besar dibandingkan dengan *term* yang sering muncul. Persamaan dari *Invers Document Frequency* (IDF) ditunjukkan pada persamaan berikut :

𝑵

𝑰𝑫𝑭 = 𝐥𝐨𝐠𝟏𝟎(𝑫𝑭) (2.2)

Dengan :

 Adalah jumlah dokumen text.

 Adalah jumlah dokumen yang mengandung *term t*

### *Term Frequency – Invers Document Frequency* (TF-IDF)

Nilai tf-idf dari sebuah kata merupakan kombinasi dari nilai tf dan nilai idf dalam perhitungan bobot. Persamaan dari TF-IDF ditunjukkan pada persamaan berikut :

𝑻𝒆𝒓𝒎 𝑾𝒆𝒊𝒈𝒉𝒕𝒊𝒏𝒈 = 𝑻𝑭 ∗ 𝑰𝑫𝑭 (2.3)

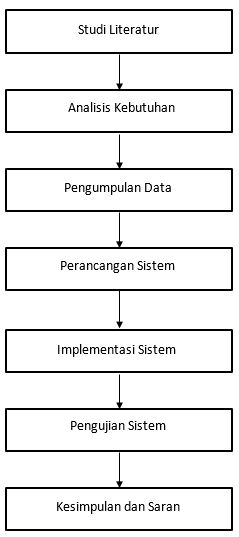
Keterangan:

*TF* = Nilai hasil *Term* *Frequency*

*IDF* = Nilai hasil *Inverse* *Document* *Frequency*

# METODOLOGI

Dalam bab tiga ini berisi tentang alur yang dilakukan dalam penelitian yang terdiri dari studi literatur, analisis kebutuhan, pengumpulan data, perancangan, implementasi, pengujian dan penarikan kesimpulan serta saran untuk penelitian selanjutnya. Langkah-langkah dalam penelitian ini akan dijelaskan pada diagram alir pada gambar 3.1



Gambar 3.1 Diagram alir metodologi penelitian

## Studi Literatur

Pada studi literatur tahapan utamanya adalah mencari sumber atau bahan referensi yang berisi tentang penjabaran teori yang digunakan dan beberapa penelitian yang telah dilakukan sebagai penunjang penelitian ini. Berikut merupakan teori yang digunakan dalam penelitian ini, antara lain :

*Analisis sentimen*

*K-Nearest Neighbor*

Pembelajaran

Sumber atau referensi yang digunakan bagi penelitian ini antara lain berasal dari buku, penelitian sebelumnya, jurnal, skripsi, serta konten yang terdapat di internet

## Analisisis Kebutuhan Sistem

Analisis kebutuhan sistem merupakan sebuah tahapan dalam menganalisis apa saja yang dibutuhkan dalam membangun sistem. Kebutuhan dibagi menjadi 2 yaitu kebutuhan fungsional dan non fungsional.

### Kebutuhan Fungsional

Sistem mampu menampilkan data latih

Sistem mampu melkakukan klasifikasi komentar dengan analisis sentimen menggunakan metode K-NN

Sistem mampu menghasilkan keluaran berupa data hasil klasifikasi analisis sentimen

### Kebutuhan Non Fungsional

1. Kebutuhan *software*:

* Sistem operasi Windows 10 64-bit.
* Microsoft Word 2013 sebagai editor dokumentasi penulisan.
* Microsoft Excell 2013 sebagai editor manualisasi sistem.
* NetBeans IDE 8.1 sebagai pengembangan sistem.
* MySQL sebagai manajemen database sistem.

1. Kebutuhan *hardware*:

* Intel® Core™i3 sebagai prosessor
* RAM 4 GB

## Perancangan Sistem

Tahap perancangan sistem merupakan tahap yang menjelaskan alur kerja sistem yang akan dibangun serta menjadi acuan dalam melakukan tahap implementasi. Sistem yang dibangun bertujuan untuk menklasifikasikan kkomentar pada pembelajaran menggunakan analisis sentimen dengan metode K-NN. Gambar 3.3

Data dimasukkan

Preprocessing text

Klasifikasi dengan metode K-NN

Analisis sentimen

selesai

Gambar 3.2 Diagram proses *Analisis sentimen*

## Implementasi

Tahap implementasi merupakan tahap untuk membangun sistem sesuai dengan tahap perancangan kedalam Bahasa pemrograman Java dengan editor pemrograman NetBeans IDE 8.1. Implementasi pembangunan sistem membutuhkan masukan berupa data komentar dari sistem pembelajaran yang telah disediakan sebelumnya, untuk dilakukan klasikasi dengan *Analisis sentimen* menggunakan metode *K-Nearest Neighbor*

## Pengujian Sistem

Tahap pengujian dilakukan pada hasil klasifikasi komentar pada pembelajaran dengan *analisis sentimen*  menggunakan metode *K-Nearest Neighbor* dengan cara menguji nilai dari beberapa parameter yang digunakan pada algoritma tersebut. Pada tahap pengujian juga dilakukan analisis terhadap kinerja sistem dengan cara menganalisis pengaruh beberapa parameter yang digunakan dari kualitas solusi yang dihasilkan.

## Kesimpulan dan Saran

Pengambilan kesimpulan dilakukan dengan cara menjawab pertanyaan pada rumusan masalah berdasarkan hasil dari pengujian yang dilakukan. Pemberian saran dilakukan dengan cara mengamati kekurangan sistem dan penelitian kemudian memberikan catatan perbaikan yang seharusnya digunakan untuk keperluan penelitian selanjutnya.

# PERANCANGAN

Pada tahap ini merupakan tahap yang menentukan hasil dari sistem yang akan dibuat. Dimana analisa yaitu tahap kegiatan dalam penentuan tindakan atau keputusan yang akan menentukan dari rancangan sistem. Sementara tahap perancangan menjadi tahap kegiatan menentukan rincian sistem yang akan dibuat.

## Deskripsi umum sistem

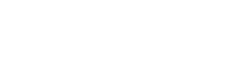
Pembangunan sistem dengan pengimplementasian metode *K-Nearest* *Neighbor* untuk mengklasifikasikan data berupa dokumen komentar mengenai pembelajaran bahasa inggris. Dalam satu dokumen terdapat beberapa jenis *term* yang berbeda dan unik, jenis *term* tersebut yang akan menjadi fitur untuk digunakan dalam proses klasifikasi. Semua *term* atau fitur dihitung bobotnya menggunakan rumus TF-IDF terlebih dahulu sebelum memasuki proses penyeleksian fitur dan pengklasifikasian pada saat pelatihan maupun pengujian.

Metode *K-Nearest* *Neighbor* bekerja dengan menghitung kemiripan antara dokumen uji terhadap setiap dokumen latih. Nilai kemiripan dihitung menggunakan *Cosine* *Similarity*, selanjutnya akan dipilih *k* tetangga terdekat. Dokumen uji akan ditentukan kelasnya berdasarkan kelas mayoritas dari *k* tetangga terdekatnya. Keluaran yang dihasilkan sistem berupa kelas “opini positif” atau “opini negatif”. Evaluasi sistem menggunakan akurasi, yang dapat dihitung dengan membandingkan hasil dari sistem dengan hasil dari pakar menggunakan Persamaan 2.8.

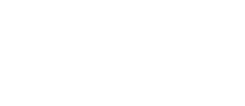
## Perancangan Proses

Ada empat tahapan utama yang dilakukan sistem yaitu, *text* *pre-processing*, melakukan perhitungan *term* *weighting* (TF-IDF), dan klasifikasi dengan *K-Nearest* *Neighbor*. Untuk memperdalam detail proses tersebut masih dibagi lagi menjadi beberapa tahapan. Dalam Gambar 4.1 menunjukkan gambaran proses jalannya sistem secara keseluruhan.

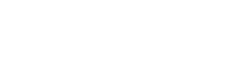
*Start*



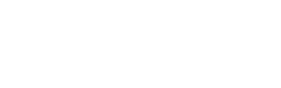
|  |  |  |
| --- | --- | --- |
|  | *Text* *pre-processing* |  |



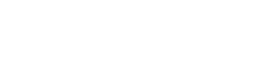
|  |  |  |
| --- | --- | --- |
|  | *Feature* *Selection* *Query* *Expansion* *Ranking* |  |



|  |  |  |
| --- | --- | --- |
|  | *Term* *Weighting* (TF-IDF) |  |



|  |  |  |
| --- | --- | --- |
|  | Klasifikasi *K-Nearest* *Neighbor* |  |



Masukkan: dokumen *tweets*

Keluaran: Hasil Analisis Sentimen

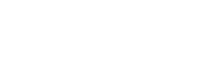
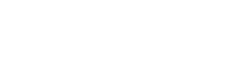
*End*

**Gambar** **4.1** **Diagram** **alir** **sistem** **4.2.1** ***Text*** ***pre-processing***

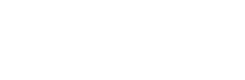
Tahapan*textpre-processing*meliputi tokenisasi,yang kedua *filtering*, danyang ketiga *stemming*. *Text* *pre-processing* ini dilakukan pada semua data yang diperoleh dengan tujuan untuk memperoleh data terstruktur agar sistem lebih mudah saat proses pengklasifikasiannya menggunakan metode *K-Nearest* *Neighbor*. Tahapan proses *text* *pre-processing* data dapat dilihat pada Gambar 4.2.

Masukkan pada sistem berupa data yang diperoleh dari Twitter dengan kata kunci Kurikulum 2013/Kurtilas/K-13. Kemudian data melalui proses tokenisasi. Setelah data tertokenisasi, lalu dilakukanlah proses *filtering* dan dilanjutkan dengan proses *stemming*. Semua tahapan *text* *pre-processing* telah selesai dilakukan, akan terdapat keluaran berupa kata dasar. Tahapan proses tersebut dijelaskan lebih rinci lagi pada penjelasan subproses berikutnya.

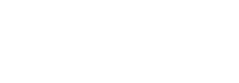
*Start*



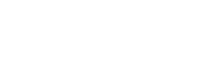
|  |  |  |
| --- | --- | --- |
|  | Tokenisasi |  |



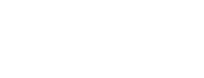
|  |  |  |
| --- | --- | --- |
|  | *Filtering* |  |



|  |  |  |
| --- | --- | --- |
|  | *Stemming* |  |



Masukkan: dokumen *tweets*



|  |  |  |
| --- | --- | --- |
|  | *Text* *pre-processing* |  |

Keluaran: kata dasar

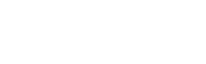
*End*

**Gambar** **4.2** **Diagram** **alir** ***text*** ***pre-processing*** **4.2.1.1** **Tokenisasi**

Tokenisasi merupakan subproses pertama di dalam tahapan *text* *pre-processing* dan fungsinya untuk melakukan perubahan pada setiap huruf besar menjadi huruf kecil serta menguraikan sebuah kalimat pada dokumen ke dalam token-token (terpisah menjadi kata demi kata) dengan pemisahnya menggunakan spasi. Selain itu tokenisasi memiliki fungsi untuk menghilangkan tanda baca pada kalimat. Kemudian menghilangkan beberapa karakter seperti *hashtag* (#), *emoticon*, tanda hubung (-) dan karakter angka. Subproses tokenisasi ini akan digambarkan pada pada Gambar 4.3.



*Start*

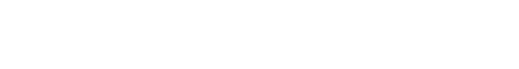
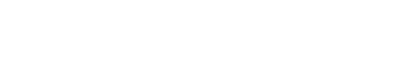
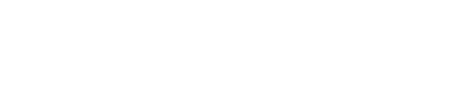


|  |  |  |
| --- | --- | --- |
|  | Tokenisasi |  |

Masukkan: Dokumen *tweets*

A

A



for i = 1 to n dokumen

Dokumen ke-i

Ganti karakter, tanda baca, tanda hubung dengan spasi jika bukan merupakan huruf, kata atau spasi

Ubah semua huruf besar menjadi huruf kecil

Pisahkan setiap kata yang ada pada dokumen (token)

*End* i

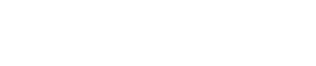
Kata hasil tokenisasi

*Return*

**Gambar** **4.3** **Diagram** **alir** **proses** **tokenisasi** **4.2.1.2** ***Filtering***

Tahapan proses *filtering* ini, kata penghubung, pengulangan kata, serta kata yang tidak berpengaruh dalam kalimat pada dokumen latih maupun uji akan dihilangkan. Daftar kata yang akan dihilangkan terdapat dalam kamus *stopword* yang sudah disediakan oleh peneliti yang nantinya disimpan dalam sistem. Berikut diagram alir tahapan dari proses *filtering* dapat dilihat pada Gambar 4.4.

*Start*



Masukkan: kata hasil tokenisasi



|  |  |  |
| --- | --- | --- |
|  | *Filtering* |  |

for i = 1 to n kata

A B C

A B C



kata ke-i

kata = stopword

Hapus kata

i

Kata hasil *filtering*

*End*

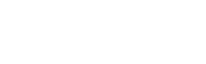
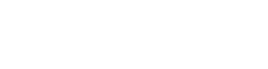
*Return*

**Gambar** **4.4** **Diagram** **alir** **proses** ***filtering***

**4.2.1.3** ***Stemming***

Tahapan proses selanjutnya adalah *stemming*, tujuannya untuk mengganti setiap kata yang ada pada dokumen hasil *filtering* menjadi kata dasarnya masing-masing. Algoritme *library* Sastrawi *stemming* yang ada pada github digunakan pada tahap ini. Penelitian yang dilakukan oleh Agusta (2009), melakukan perbandingan kinerja algoritme Nazief dan Adriani dengan algoritme Porter. Hasilnya menyatakan algoritme Nazief dan Adriani memiliki keakuratan yang lebih tinggi. Algoritme ini penerapannya seperti algoritme Nazief dan Adriani yang ditingkatkan algoritme *Confix* *Stripping* lalu peningkatan dilakukan lagi dengan algoritme Enhanced *Confix* *Stripping* dan terakhir peningkatan dilakukan kembali dengan algoritme *Modified* *Enhanced* *Confix* *Stripping*. Berikut gambaran diagram alir daritahapanproses*stemming*pada *library*Sastrawi dapatdilihatpadaGambar 4.5, Gambar 4.6, Gambar 4.7.

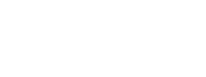
*Start*



|  |  |  |
| --- | --- | --- |
|  | *Stemming* |  |

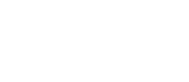
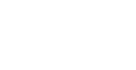
Kata hasil *filtering*

A

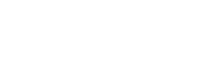


|  |  |  |
| --- | --- | --- |
|  | *Stemming* kata jamak |  |

A



|  |  |  |
| --- | --- | --- |
|  | *Stemming* kata jamak |  |



|  |  |  |
| --- | --- | --- |
|  | *Stemming* kata tunggal |  |



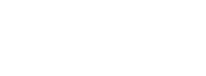
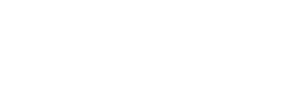
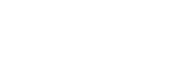
Kata jamak?

Kata hasil *stemming*

*End*

**Gambar** **4.5** **Diagram** **alir** **proses** ***stemming***

*Start*



|  |  |  |
| --- | --- | --- |
|  | *Stemming* kata tunggal |  |



ada pola “-“ ?

Ya

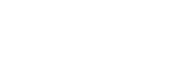
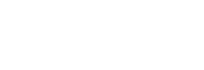
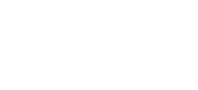
Ada akhiran ku,mu, nya ?

Tidak

Pisah akhiran dengan “-“

A B

A B



|  |  |
| --- | --- |
| Hasil = kataAwal | |
|  |  |



Ada di kamus dan hasil stem == kata ?

Tambah awalan “me”

Stem 1 == Stem 2 ?

Tidak Ya

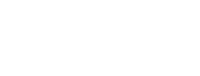
Hasil = hasilStem

Kata hasil *stemming*

*Return*

**Gambar** **4.6** **Diagram** **alir** ***stemming*** **kata** **jamak** **pada** ***Library*** **Sastrawi**

*Start*



|  |  |  |
| --- | --- | --- |
|  | *Stemming* Kata Tunggal |  |

Kata ada di kamus ?

Ya

Tidak

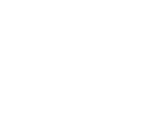
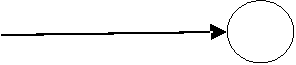
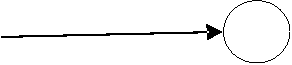
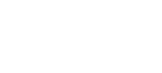
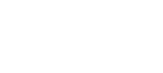
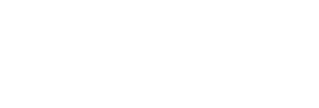
Kata <= 3 ? Ya

Tidak

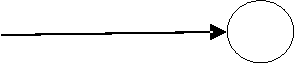
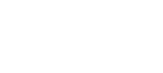
Hapus akhiraran lah, kah, tah, pun, ku, mu, nya, I, kan, an,

B A

A B



|  |  |  |
| --- | --- | --- |
|  | *Confix* *Stripping* |  |



Hapus awalan di, ke, se

Cek dan hapus awal dengan 56 aturan ambiguitas sastrawi

Kata ada di kamus ?

Ya

Tidak

Ya Terdapat pola be-lah, be-an, me-i, pe-i, ter-i ?

Tidak

Hapus akhiran

Kata ada di kamus ?

Ya

Tidak

Hapus awalan

Kata ada di kamus ?

**4.2.2** ***Term*** ***weighting*** **(TF-IDF)**

Tahapan proses *term* *weighting* berfungsi memberikan nilai bobot dari kata (*term*) dengan menggunakan metode perkalian nilai *term* *frequency* (TF) dan *inverse* *document* *frequency* (IDF). Untuk nilai *term* *frequency* (TF) diperoleh dengan Persamaan logarithmic TF, dan untuk nilai *inverse* *document* *frequency* (IDF) diperoleh dengan menghitung berapa banyak dokumen mengandung sebuah *term* tertentu (DF). Diagram alir proses perhitungan TF-IDF terdapat pada Gambar 4.9.

*Start*



|  |  |  |
| --- | --- | --- |
|  | Hitung DF |  |



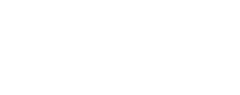
|  |  |  |
| --- | --- | --- |
|  | Hitung IDF |  |



|  |  |  |
| --- | --- | --- |
|  | Hitung TF-IDF |  |



|  |  |  |
| --- | --- | --- |
|  | Hitung TF |  |



|  |  |  |
| --- | --- | --- |
|  | *Term* *Weighting* (TF-IDF) |  |

Masukkan: Data kata dasar Dokumen Latih dan Uji

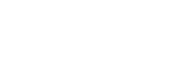
Hasil *term* *weighting*

*Return*

**Gambar** **4.9** **Diagram** **alir** **proses** ***term*** ***weighting*** Gambar4.9diatas*termweighting*memiliki susunanatasbeberapasub-proses.

Sub-proses ini menjadi satu kesatuan untuk menghitung *term* *weighting* (TF-IDF), berikut adalah susunannya sub-proses hitung *term* *frequency* (TF), hitung *document* *frequency* (DF), hitung *inverse* *document* *frequency* (IDF), dan paling akhir hitung TF-IDF. Untuk sub-sub proses lebih detail diagram alir prosesnya pada Gambar 4.10, Gambar 4.11, Gambar 4.12, dan Gambar 4.13.

*Start*



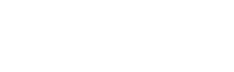
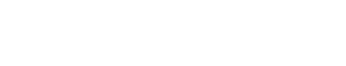
|  |  |  |
| --- | --- | --- |
|  | Hitung TF |  |

Masukkan: Data kata dasar Dokumen Latih dan Uji

A

28

A



i = 1 to sizeKategori

j = 1 to n dokumen

k = 1 to m *term* tokenDokumen Latih

kata == token[k] ?

Ya

Tidak

tfKata++

k

j

i

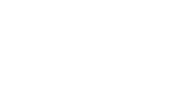
Hasil *term* *frequency*

*Return*

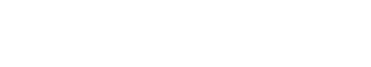
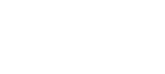
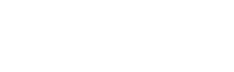
**Gambar** **4.10** **Diagram** **alir** **proses** **perhitungan** **TF**

Perhitungan *term* *frequency* (TF), memuat perhitungan banyaknya kemunculan *term* tertentu di setiap dokumen yang ada pada kategori positif dan negatif. Perhitungan *term* *frequency* (TF) telah selesai dilakukan, kemudian melakukan proses perhitungan *document* *frequency* dari *term* tersebut dihitung kemunculan pada seluruh dokumen pada seluruh kategori. Berikut diagram alir proses perhitungan *document* *frequency* dapat dilihat pada Gambar 4.11.

*Start*



|  |  |  |
| --- | --- | --- |
|  | Hitung DF |  |



Masukkan: Data kata dasar Dokumen Latih dan Uji

i = 1 to sizeKategori

j = 1 to n dokumen

tokenDokumen[j] terdapat kata ?

Ya

Tidak

dfKata++

j

dfKata[i]<-dfKata

dfKata = 0

i

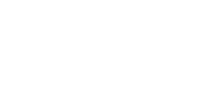
Hasil *document* *frequency*

*Return*

**Gambar** **4.11** **Diagram** **alir** **proses** **perhitungan** **DF**

Pada proses selanjutnya setelah perhitungan *document* *frequency* (DF), kemudian melakukan proses perhitungan *inverse* *document* *frequency* (IDF) dengan tujuan untuk digunakannya dalam pembobotan nilai *term* menggunakan metode TF-IDF. Berikut diagram alir proses perhitungan *document* *frequency* dapat dilihat pada Gambar 4.12.

*Start*



|  |  |  |
| --- | --- | --- |
|  | Hitung IDF |  |



Masukkan: Data kata dasar Dokumen Latih dan Uji

i = 1 to sizeKategori

idf = 10log\* (sizeDokumenLatih/dfKata[i])

i

Hasil IDF

*Return*

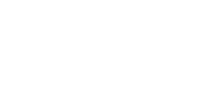
**Gambar** **4.12** **Diagram** **alir** **proses** **perhitungan** **IDF**

Nilai *term* *frequency* (TF) dan *inverse* *document* *frequency* (IDF) diketahui, lalu menghitung nilai bobot TF-IDF setiap kata atau *term* pada seluruh dokumen. Berikut diagram alir proses perhitungan *document* *frequency* dapat dilihat pada Gambar 4.13.



*Start*

Masukkan: Data kata dasar Dokumen Latih dan Uji, tfKata, idfKata



|  |  |  |
| --- | --- | --- |
|  | Hitung TF-IDF |  |

i = 1 to n kata

wKata = tfKata\*idfKata

i

Hasil wKata

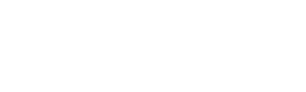
*Return*

**Gambar** **4.13** **Diagram** **alir** **proses** **perhitungan** **TF-IDF**

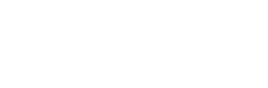
**4.2.3** **Klasifikasi** ***K-Nearest*** ***Neighbor***

Pada proses klasifikasi dokumen uji, dokumen uji akan dihitung kemiripannya dengan dokumen latih setiap kategorinya. Dihitung menggunakan metode *Cosine* *Similarity*. Setelah diketahuinya nilai *Cosine* *Similarity*, kemudian dilakukan pengklasifikasian menggunakan metode *K-Nearest* *Neighbor* untuk menentukan kelas dari tiap data uji. Hasil dari klasifikasi ini akan diamati dan dianalisa untuk dapat dilakukan perhitungan nilai akurasi pada tahap selanjutnya. Untuk alir dari proses klasifikasi dokumen uji dapat dilihat pada 4.19 berikut.

*Start*



|  |  |  |
| --- | --- | --- |
|  | Klasifikasi *K-Nearest* *Neighbor* |  |



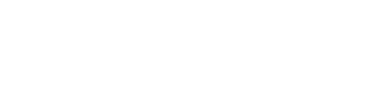
Menentukan nilai *k*

Masukkan: Data kata dasar Dokumen Latih dan Uji, wKata

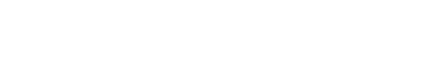
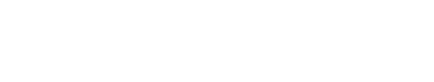
A



A



|  |  |  |
| --- | --- | --- |
|  | Hitung *Cosine* *similarity* antar dokumen |  |



Urutkan hasil *Cosine* *similarity*

Mengambil *k* tetangga terdekat

Tentukan kelas

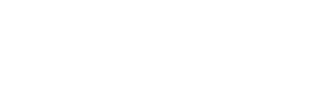
Hasil Klasifikasi Analisis Sentimen

*End*

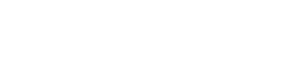
**Gambar** **4.19** **Diagram** **alir** **proses** **klasifikasi** ***K-Nearest*** ***Neighbor***

Sebelum dokumen-dokumen diklasifikasikan, dilakukan proses perhitungan *Cosine* *Similarity* terlebih dahulu untuk mengetahui nilai derajat kesamaan atau kemiripan antar dokumen uji dengan dokumen latih dari tiap kategorinya. *Cosine* *Similarity* dihitung setelah normalisasi hasil pembobotan *term* dilakukan. Berikut Gambar 4.20 merupakan gambaran proses perhitungan derajat kesamaan antar tiap dokumen dengan menggunakan *Cosine* *Similarity*.

*Start*



|  |  |  |
| --- | --- | --- |
|  | Hitung *Cosine* *Similarity* antar Dokumen |  |



|  |  |  |
| --- | --- | --- |
|  | Normalisasi Bobot |  |

Masukkan: Data kata dasar Dokumen Latih dan Uji, wKata

A

## Penghitungan Manual

Pada penelitian ini, manualisasi dimulai dari pengumpulan dataset akan diproses digunakan untuk mendapatkan hasil akhir berupa sentimen bernilai positif atau negatif. *Dataset* dengan jumlah 7 dokumen, terdiri 6 dokumen sebagai data latih dan 1 dokumen sebagai data uji. Pada data latih terdapat 3 data yang merupakan opini positif dan 3 data merupakan opini negatif.

### *Dataset*

*Dataset* yang digunakan terbagi menjadi dua, yaitu yang pertama data latih dan kemudian data uji. Pada data latih terdapat 6 dokumen yang digunakan. Sedangkan pada dokumen uji, terdapat 1 dokumen.

Tabel 4.1 Data Latih

|  |  |  |  |
| --- | --- | --- | --- |
| No | Label | Dokumen | Kategori |
| 1 | Dok1 | setelah membaca semua bacaan yang tertera dalam tes kali ini, saya merasa ingin sekali mengetahui banyak hal tentang bacaan-baccan yang belum saya ketahui sebelumnya sehingga memberi sebuah wawasan, ilmu ataupun gagasa-gagasan penting dalam sebuah bacaan tersebut sehingga bisa meningkatkan kemampuan membaca saya dalam proses pembelajaran bahasa inggris. | Positif |
| 2 | Dok2 | Sebenarnya saya pernah melakukan tugas semacam ini ketika saya masih di S1. Ini bukan hal baru buat saya. Saya perhatikan bahwa tingkat kompleksitas dari teks 1 hingga 3 meningkat. Ini baik. Dalam kesempatan lain, pengalaman belajar ini dapat ditingkatkan. | Postif |
| 3 | Dok3 | Ini adalah sebuah platform baru untuk saya dan tentu saja ini sangat bermanfaat. | Positif |
| 4 | Dok4 | Saya senang menjawab pertanyaan namun saya tidak suka membaca secara online, membuat mata saya sakit. | Negatif |
| 5 | Dok5 | Topik dari bacaan sebelumnya kurang menarik bagi saya sehingga saya kurang berminat membaca sampai detail. Soalnya juga agak sulit. Saya tidak menemukan kenikmatan membaca disini. | Negatif |
| 6 | Dok6 | Ketika saya berlatih kemampuan membaca dengan membaca beberapa teks dan beberapa soal yang topiknya tidak sesuai dengan pendidikan yang sedang saya tekuni, saya hanya merasakan satu manfaat saja, yakni mengasah skill saya dalam membaca. Akan tetapi jikalau topik yang dibahas dalam bahan bacaan sesuai dengan proses akademik yang sedang saya tekuni, saya yakin bahwa saya akan mendapatkan 2 manfaat, pertama, hal tersebut dapat mengasah skill membaca saya, dan yang kedua, saya mampu menambah ilmu yang sudah saya tekuni selama ini dalam proses akademik ini. | Negatif |
| 7 | Dok7 | Saya kurang menyukai jenis tes ini karena tema bacaan yang disajikan kurang menarik dan hanya menggunakan satu jenis pertanyaan saja. Layout atau desain web perlu dibuat lebih interkatif. | Negatif |

Tabel 4.2 Data uji

|  |  |  |  |
| --- | --- | --- | --- |
| No | Label | Dokumen | Kategori |
| 1 | Dok1 | kutipannya cukup menarik. Saya tidak begitu suka membaca, tetapi dua bagian terakhir cukup menarik untuk membuat saya ingin melakukan latihan ini. | ? |
|  |  |  |  |

**4.3.2** ***Text*** ***pre-processing***

Pada tahapan *text* *pre-processing* ini terdapat tiga tahapan, yaitu tokenisasi, *filtering*, dan *stemming*. *Pre-processing* menjelaskan bagaimana alur dari proses mengolahdata dokumen *tweets*yangnantinya akandiklasifikasikanmenggunakan metode *K-Nearest* *Neighbor*. Berikut hasil tokenisasi dari data latih dan data uji dapat dilihat pada Tabel 4.3 dan Tabel 4.4.

**Tabel** **4.3** **Hasil** **tokenisasi** **data** **latih**

|  |  |  |  |
| --- | --- | --- | --- |
| No | Label | Dokumen | Kategori |
| 1 | Dok1 | setelah membaca semua bacaan yang tertera dalam tes kali ini saya merasa ingin sekali mengetahui banyak hal tentang bacaan bacaan yang belum saya ketahui sebelumnya sehingga memberi sebuah wawasan ilmu ataupun gagasan gagasan penting dalam sebuah bacaan tersebut sehingga bisa meningkatkan kemampuan membaca saya dalam proses pembelajaran bahasa inggris | Positif |
| 2 | Dok2 | Sebenarnya saya pernah melakukan tugas semacam ini ketika saya masih di Strata satu Ini bukan hal baru buat saya Saya perhatikan bahwa tingkat kompleksitas dari teks hingga meningkat Ini baik Dalam kesempatan lain pengalaman belajar ini dapat ditingkatkan | Postif |
| 3 | Dok3 | Ini adalah sebuah platform baru untuk saya dan tentu saja ini sangat bermanfaat | Positif |
| 4 | Dok4 | Saya senang menjawab pertanyaan namun saya tidak suka membaca secara online membuat mata saya sakit | Negatif |
| 5 | Dok5 | Topik dari bacaan sebelumnya kurang menarik bagi saya sehingga saya kurang berminat membaca sampai detail Soalnya juga agak sulit Saya tidak menemukan kenikmatan membaca disini | Negatif |
| 6 | Dok6 | Ketika saya berlatih kemampuan membaca dengan membaca beberapa teks dan beberapa soal yang topiknya tidak sesuai dengan pendidikan yang sedang saya tekuni saya hanya merasakan satu manfaat saja yakni mengasah skill saya dalam membaca Akan tetapi jikalau topik yang dibahas dalam bahan bacaan sesuai dengan proses akademik yang sedang saya tekuni saya yakin bahwa saya akan mendapatkan manfaat pertama hal tersebut dapat mengasah kemampuan membaca saya dan yang kedua saya mampu menambah ilmu yang sudah saya tekuni selama ini dalam proses akademik ini | Negatif |
|  |  |  |  |

**Tabel** **4.4** **Hasil** **tokenisasi** **data** **uji**

|  |  |  |  |
| --- | --- | --- | --- |
| No | Label | Dokumen | Kategori |
| 1 | Dok1 | kutipannya cukup menarik saya tidak begitu suka membaca, tetapi dua bagian terakhir cukup menarik untuk membuat saya ingin melakukan latihan ini | ? |
|  |  |  |  |

Tahapan proses selanjutnya adalah *filtering*, berfungsi untuk menyimpan *term* yang penting dan menghapus *term* yang tidak penting. Berikut Tabel 4.5 dan Tabel 4.6 menampilkan hasil dari tahapan proses *filtering*.

**Tabel** **4.5** **Hasil** ***filtering*** **data** **latih**

|  |  |  |  |
| --- | --- | --- | --- |
| No | Label | Dokumen | Kategori |
| 1 | Dok1 | membaca semua bacaan tertera dalam tes kali bacaan bacaan wawasan ilmu gagasan gagasan penting bacaan meningkatkan kemampuan membaca proses pembelajaran bahasa inggris | Positif |
| 2 | Dok2 | tugas di strata satu baru perhatikan tingkat kompleksitas teks meningkat kesempatan pengalaman belajar ditingkatkan | Postif |
| 3 | Dok3 | platform bermanfaat | Positif |
| 4 | Dok4 | senang menjawab pertanyaan namun saya tidak suka membaca secara online membuat mata sakit | Negatif |
| 5 | Dok5 | Topik bacaan menarik berminat membaca detail sulit menemukan kenikmatan membaca | Negatif |
| 6 | Dok6 | berlatih membaca membaca teks soal topiknya sesuai pendidikan tekuni merasakan manfaat mengasah skill membaca topik dibahas bahan bacaan sesuai proses akademik tekuni manfaat mengasah kemampuan membaca menambah ilmu tekuni proses akademik | Negatif |
|  |  |  |  |

**Tabel** **4.6** **Hasil** ***filtering*** **data** **uji**

|  |  |  |  |
| --- | --- | --- | --- |
| No | Label | Dokumen | Kategori |
| 1 | Dok1 | kutipannya menarik suka membaca menarik latihan | ? |
|  |  |  |  |

**Tabel** **4.7** **Hasil** ***stemming*** **data** **latih**

|  |  |  |  |
| --- | --- | --- | --- |
| No | Label | Dokumen | Kategori |
| 1 | Dok1 | baca baca tertera tes kali baca baca wawasan ilmu gagasan gagasan penting baca tingkat baca proses belajar bahasa inggris | Positif |
| 2 | Dok2 | tugas strata hati tingkat kompleks teks tingkat sempat alam belajar tingkat | Postif |
| 3 | Dok3 | platform manfaat | Positif |
| 4 | Dok4 | senang jawab baca online mata sakit | Negatif |
| 5 | Dok5 | topik baca tarik minat baca detail sulit temu nikmat baca | Negatif |
| 6 | Dok6 | latih baca baca teks soal topik suai didik tekun rasa manfaat asah skill baca topik bahas bahan baca suai proses akademik tekun manfaat asah baca tambah ilmu tekun proses akademik | Negatif |
|  |  |  |  |

**Tabel** **4.8** **Hasil** **s*temming*** **data** **uji**

|  |  |  |  |
| --- | --- | --- | --- |
| No | Label | Dokumen | Kategori |
| 1 | Dok1 | Kutip tarik suka baca tarik latih | ? |
|  |  |  |  |

**4.3.3** ***Term*** ***weighting*** **(TF-IDF)**

Metode yang digunakan untuk menghitung bobot dari sebuah *term* (*term* *weighting*) adalah *Term* *Frequency* *–* *Inverse* *Document* *Frequency* (TF-IDF). Sebelumnya kita harus melakukan perhitungan *Term* *Frequency* (TF), *Document* *Frequency* (DF), lalu *Inverse* *Document* *Frequency* (IDF) untuk mengetahui nilainya dan dimasukkan pada perhitungan TF-IDF. Berikut pada Tabel 4.9 merupakan contoh perhitungan *termweighting*dari fitur data uji yang telahterpilihuntuknilai *Term* *Frequency* (TF), *Document* *Frequency* (DF), serta *Inverse* *Document* *Frequency* (IDF).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Kata | query | doc 1 | doc 2 | doc 3 |
| **column21** | **column22** | **column23** | **column24** | **column25** |
| baca | -0,60205999 | -1,07055373 | 0 | 0 |
| tertera | 0 | 0,602059991 | 0 | 0 |
| tes | 0 | 0,602059991 | 0 | 0 |
| kali | 0 | 0,602059991 | 0 | 0 |
| wawasan | 0 | 0,602059991 | 0 | 0 |
| ilmu | 0 | 0,301029996 | 0 | 0 |
| gagasan | 0 | 0,391649054 | 0 | 0 |
| penting | 0 | 0,602059991 | 0 | 0 |
| tingkat | 0 | 0 | 0 | 0 |
| proses | 0 | 0,124938737 | 0 | 0 |
| belajar | 0 | 0,301029996 | 0,301029996 | 0 |
| bahasa | 0 | 0,602059991 | 0 | 0 |
| inggris | 0 | 0,602059991 | 0 | 0 |
| tugas | 0 | 0 | 0,602059991 | 0 |
| strata | 0 | 0 | 0,602059991 | 0 |
| hati | 0 | 0 | 0,602059991 | 0 |
| kompleks | 0 | 0 | 0,602059991 | 0 |
| teks | 0 | 0 | 0,301029996 | 0 |
| sempat | 0 | 0 | 0,602059991 | 0 |
| alam | 0 | 0 | 0,602059991 | 0 |
| platform | 0 | 0 | 0 | 0,602059991 |
| manfaat | 0 | 0 | 0 | 0,602059991 |
| senang | 0 | 0 | 0 | 0 |
| jawab | 0 | 0 | 0 | 0 |
| online | 0 | 0 | 0 | 0 |
| mata | 0 | 0 | 0 | 0 |
| sakit | 0 | 0 | 0 | 0 |
| topik | 0 | 0 | 0 | 0 |
| tarik | 0,162549044 | 0 | 0 | 0 |
| minat | 0 | 0 | 0 | 0 |
| detail | 0 | 0 | 0 | 0 |
| sulit | 0 | 0 | 0 | 0 |
| temu | 0 | 0 | 0 | 0 |
| nikmat | 0 | 0 | 0 | 0 |
| latih | 0,602059991 | 0 | 0 | 0 |
| soal | 0 | 0 | 0 | 0 |
| suai | 0 | 0 | 0 | 0 |
| didik | 0 | 0 | 0 | 0 |
| tekun | 0 | 0 | 0 | 0 |
| rasa | 0 | 0 | 0 | 0 |
| asah | 0 | 0 | 0 | 0 |
| skill | 0 | 0 | 0 | 0 |
| bahas | 0 | 0 | 0 | 0 |
| bahan | 0 | 0 | 0 | 0 |
| akademik | 0 | 0 | 0 | 0 |
| tambah | 0 | 0 | 0 | 0 |
| kutip | 0,602059991 | 0 | 0 | 0 |
| suka | 0,602059991 | 0 | 0 | 0 |

**Tabel** **4.9** **Contoh** **hasil** **perhitungan** ***term*** ***weighting*** **(TF-IDF)**

|  |  |  |  |
| --- | --- | --- | --- |
| doc4 | doc 5 | doc 6 |  |
| **column26** | **column27** | **column28** | **column29** |
| -0,60205999 | -0,88931561 | -1,02288187 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0,162549044 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0,301029996 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |
| 0,602059991 | 0 | 0 |  |
| 0,602059991 | 0 | 0 |  |
| 0,602059991 | 0 | 0 |  |
| 0,602059991 | 0 | 0 |  |
| 0,602059991 | 0 | 0 |  |
| 0 | 0,124938737 | 0,162549044 |  |
| 0 | 0,124938737 | 0 |  |
| 0 | 0,602059991 | 0 |  |
| 0 | 0,602059991 | 0 |  |
| 0 | 0,602059991 | 0 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0,301029996 |  |
| 0 | 0 | 0,602059991 |  |
| 0 | 0 | 0,602059991 |  |
| 0 | 0 | 0,391649054 |  |
| 0 | 0 | 0,602059991 |  |
| 0 | 0 | 0,184549663 |  |
| 0 | 0 | 0,602059991 |  |
| 0 | 0 | 0,391649054 |  |
| 0 | 0 | 0,602059991 |  |
| 0 | 0 | 0,602059991 |  |
| 0 | 0 | 0,602059991 |  |
| 0 | 0 | 0,391649054 |  |
| 0 | 0 | 0,602059991 |  |
| 0 | 0 | 0 |  |
| 0 | 0 | 0 |  |

**Tabel** **4.10** **Contoh** **hasil** **perhitungan** ***term*** ***weighting*** **(TF-IDF) (Lanjutan)**

**4.3.5** **Normalisasi** **bobot**

Normalisasi bobot diperlukan pada saat proses perhitungan derajat kesamaan antara data uji dengan data latih dengan menggunakan perhitungan *Cosine* *Similarity*. Proses ini menggunakan nilai *term* *weighting* (TF-IDF) yang telah dihitung sebelumnya untuk dinormalisasi nilainya. Tabel 4.11 menunjukkan hasil proses normalisasi bobot *term* terpilih dari dokumen data uji.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 1kali 7 | 2 kali 7 | 3 kali 7 | 4 kali 7 | 5 kali 7 | 6 kali 7 |  |
| **column43** | **column48** | **column52** | **column55** | **column57** | **column58** |  |
| 0,379254284 | 1,199134488 | 0 | 0 | 0,379254284 | 0,827490067 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0,000412442 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0,008211814 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0,000412442 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0,13138902 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0 | 0 | 0 | 0 | 0 | 0 |  |
| 0,510643303 | 1,19954693 | 0,008211814 | 0 | 0,379254284 | 0,82790251 | sum |
| 2,626966156 | 4,342227674 | 3,318636034 | 1,840848059 | 3,188442368 | 2,98762757 | sqrt |
| 0,194385185 | 0,276251505 | 0,002474454 | 0 | 0,11894657 | 0,277110346 | cosin |

**Tabel** **4.11** **Contoh** **hasil** **perhitungan** ***cosine similarity***

**4.3.6** **Klasifikasi** ***K-Nearest*** ***Neighbor***

Nilai *k* (jumlah tetangga terdekat) pada *K-Nearest* *Neighbor* dapat mempengaruhi hasil klasifikasi. Berikut Tabel 4.12 menunjukkan hasil perhitungan KNN dari dokumen pada data uji.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| perhitungan KNN | | | | | |
| Doc1 | Doc2 | Doc3 | Doc4 | Doc5 | Doc6 |
| 0,19438519 | 0,27625151 | 0,00247445 | 0 | 0,11894667 | 0,27711035 |
| positif | positif | positif | negatif | negatif | negatif |
|  |  |  |  |  |  |
| Doc6 | Doc2 | Doc1 | Doc5 | Doc3 | Doc4 |
| 0,27711035 | 0,27625151 | 0,19438519 | 0,11894667 | 0,00247445 | 0 |
| negatif | positif | positif | negatif | positif | negatif |

**Tabel** **4.12** **Contoh** **hasil** **perhitungan** *K-Neaarest Neighbor*

Untuk pengklasifikasian hasil dari *Cosine* *Similarity* pada Tabel 4.111, diurutkan nilainya dari terbesar hingga terkecil. Pada Tabel 4.12 kelas yang paling dominan dengan nilai *k* = 3, dapat disimpulkan bahwa dokumen uji tersebut termasuk bernilai positif.

# PEMBAHASAN

Pembahasan berfungsi untuk menerjemahkan makna dari hasil yang diperoleh untuk menjawab pertanyaan atau masalah penelitian. Fungsi lainnya adalah untuk menjelaskan pemahaman baru yang didapatkan dari hasil penelitian, yang diharapkan berguna dalam pengembangan keilmuan. Dalam penelitian tingkat lanjut, fungsi pembahasan yang kedua ini sangat penting karena dapat menunjukkan kontribusi penulis terhadap pengembangan keilmuan. Akan tetapi, dalam penelitian tingkat skripsi, fungsi yang kedua ini dapat diterapkan secara terbatas karena pendidikan S1 tidak dituntut untuk pengembangan keilmuan secara substansial, tetapi cukup terhadap pemahaman personal dalam implementasi konsep atau teori.

## Sub bab lima satu

Dalam menjawab masalah penelitian, penulis diminta untuk melakukan evaluasi kritis terhadap hasil yang diperoleh. Tergantung dari fokus penelitian, beberapa contoh pertanyaan kritis yang dapat dijawab adalah:

* Seberapa jauh tujuan penelitian telah tercapai?
* Apakah aplikasi atau sistem yang dibangun sesuai dengan tujuannya?
* Apakah metode atau praktik perancangan dan implementasi yang baik telah dijalankan?
* Apakah teknologi implementasi yang tepat telah dipilih? Dan sebagainya.

### Seksi lima satu satu

Dalam menjelaskan pemahaman baru yang didapatkan, penulis dapat mengubungkan hasil penelitian dengan pengetahuan teoritik atau penelitian sebelumnya yang telah dibahas. Kaitan antara hasil penelitian dan pengetahuan teoritik misalnya berupa:

* pendapat tentang metode yang digunakan dari literatur, apakah dapat digunakan dengan baik secara langsung, dengan penyesuaian, atau dengan batasan tertentu;
* konfirmasi tentang batasan dari metodologi yang digunakan sehingga dapat berpengaruh pada hasil;
* penjelasan tentang informasi penting pada penelitian lainnya yang membantu penulis untuk menerjemahkan data penelitian penulis;
* penjelasan tentang kemungkinan hasil dari penelitian lainnya yang dapat dikombinasikan dengan penelitian penulis untuk memberikan pengetahuan baru; dan sebagainya.

### Seksi lima satu dua

Penulis dapat merefleksikan apa yang telah dipelajari selama melakukan penelitian, tetapi harus tetap terfokus dengan masalah penelitian ini dan tidak melebar ke masalah lainnya. Hal-hal yang berada di luar fokus peneltian tetapi penting dan menarik untuk diteliti dapat disarankan sebagai bahan penelitian berikutnya. Hal ini dapat dipertegas di bab Kesimpulan/ Penutup.

## Seksi lima dua

Hasil dan pembahasan dapat diletakkan dengan kemungkinan berikut:

1. Dipisahkan secara fisik ke dalam bab-bab yang berbeda
2. Dipisahkan secara fisik ke dalam dua atau lebih paragraf, seksi, atau sub bab yang berbeda tetapi dalam bab yang sama
3. Dileburkan menjadi satu dalam paragraf, dijelaskan secara naratif-deskriptif, terdistribusi ke satu atau lebih bab yang ada

### Seksi lima dua satu

Cara pertama atau kedua membantu pembaca yang ingin memisahkan observasi dan terjemahan dari observasi tersebut sehingga mereka dapat menilai kualitas dari masing-masing proses dengan lebih mudah. Kadang-kadang cara kedua lebih banyak dipilih daripada cara pertama jika data yang harus dipresentasikan yang cukup banyak dan laporan penelitian cukup panjang agar pembaca tidak perlu menunggu presentasi dari seluruh data selesai baru dapat membaca penerjemahannya. Cara pertama dan kedua ini banyak digunakan untuk penelitian yang bersifat kuantitatif, baik itu deskriptif, eksplanatori, maupun implementatif.

### Seksi lima dua dua

Cara ketiga biasanya digunakan jika data, analisis, dan penafsirannya sulit dipisahkan. Pemisahannya terkadang justru membuat laporan penelitian sulit dibaca. Hal ini dapat berlaku pada tipe penelitian yang bersifat kualitatif, baik itu deskriptif ataupun analitik/eksplanatori.

Pada dasarnya peletakan dan jumlah bab untuk hasil dan pembahasan sebaiknya disesuaikan karakter penelitian masing-masing. Judul bab pun tidak harus secara eksplisit “Hasil” dan “Pembahasan” tetapi dapat digantikan dengan nama yang lebih deskpritif dan tematik.

## Sub bab lima tiga

Contoh struktur skripsi untuk implementatif pembangunan dan non-implementatif eksperimental dapat dilihat pada kedua seksi berikut.

### Contoh struktur penelitian implementatif pembangunan

Berikut ini adalah contoh bab-bab yang terdapat pada penelitian implementatif pembangunan sistem perangkat lunak.

Bab 1 Pendahuluan

Bab 2 Landasan kepustakaan

Bab 3 Metodologi

Bab 4 Persyaratan

Bab 5 Perancangan dan implementasi

Bab 6 Pengujian

Bab 7 Penutup

Bab 1 sampai Bab 3 memuat informasi yang sesuai dengan panduan sebelumnya. Isi dari bab-bab berikutnya:

* Bab 4 Persyaratan:
* Pernyataan masalah (problem statement), yang lebih elaboratif daripada yang di Pendahuluan.
* Identifikasi pemangku kepentingan (stakeholders) dan aktor (actors) sistem.
* Daftar terstruktur persyaratan/kebutuhan perangkat lunak, secara fungsional, data, dan non-fungsional
* Use cases, use case diagrams, dan use case specifications, dan sebagainya.
* Bab 5 Perancangan dan implementasi:
* Rancangan arsitektur: deskripsi struktur dan setiap komponen utama
* Representasi data dalam model data dan basis data
* Detil implementasi dari fungsi-fungsi utama yang menjadi fokus
* Bab 6 Pengujian dan evaluasi
* Strategi, rencana, kasus, dan data pengujian
* Ringkasan hasil pengujian perangkat lunak, termasuk data dan analisisnya (detilnya di Lampiran)
* Evaluasi hasil proyek secara keseluruhan, misalkan
* Bab 7 Penutup
* Ringkasan dari capaian proyek
* Saran pengembangan lebih lanjut

Pada contoh struktur ini “hasil” tersebar di beberapa bab mulai Bab 4 Persyaratan sampai Bab 6, sedangkan “pembahasan” secara keseluruhan terhadap masalah penelitian terdapat di Bab 6. Yang dimaksud dengan pengujian dalam Bab 6 terfokus pada pengujian persyaratan perangkat lunak, sedangkan evaluasi berfungsi sebagai “pembahasan” secara keseluruhan, yaitu menentukan apakah “hasil” sudah menjawab masalah penelitian yang dirumuskan pada Bab 1.

Sebagai catatan, Bab 3 Metodologi umumnya menjelaskan model proses perangkat lunak yang digunakan. Jika strategi untuk setiap aktivitasnya (analisis persyaratan, perancangan, dan seterusnya) sudah dijelaskan di Bab 3 ini juga, maka bab-bab lainnya yang berhubungan dengan aktivitas-aktivitas ini masing-masing langsung dapat menjelaskan hasil pelaksanaan metodenya.

### Contoh struktur penelitian non-implementatif eksperimental

Berikut ini adalah contoh bab-bab yang terdapat pada penelitian implementatif pembangunan sistem perangkat lunak.

Bab 1 Pendahuluan

Bab 2 Landasan kepustakaan

Bab 3 Metodologi

Bab 4 Hasil

Bab 5 Pembahasan

Bab 6 Penutup

Isi dari setiap bab dapat menyesuaikan dengan panduan yang telah dijelaskan sebelumnya. Jika diperlukan, Bab 4 dapat digabungkan dengan Bab 5, menjadi Hasil dan Pembahasan.

Struktur dasar ini cukup universal sehingga dapat digunakan juga untuk tipe-tipe penelitian lainnya, khususnya jika belum ada struktur lain yang lebih tematik dan cocok untuk penelitian yang bersangkutan.

# Penutup

Bagian ini memuat kesimpulan dan saran terhadap skripsi. Kesimpulan dan saran disajikan secara terpisah, dengan penjelasan sebagai berikut:

## Kesimpulan

Kesimpulan merupakan pernyataan-pernyataan yang singkat, jelas, dan tepat tentang hasil penelitian yang diperoleh berdasarkan tujuannya. Bagian ini merupakan penegasan dari yang telah dijelaskan pada bagian Pembahasan dan tidak memuat informasi yang baru. Bagian ini juga mencerminkan jawaban dari rumusan masalah (pertanyaan penelitian).

## Saran

Saran berisi pernyataan-pernyataan yang ringkas dan jelas tentang masalah-masalah atau hal-hal yang dapat dilakukan untuk mengembangkan penelitian ini lebih lanjut. Saran itu dapat diarahkan pada aspek metode, instrumen, populasi/sampel, dan sebagainya.

DAFTAR PUSTAKA

Adobe Systems Incorporated, 2013. Adobe Air (3.5 beta). [program komputer] Adobe Labs. Tersedia di: <http://labs.adobe.com/technologies/  
flashruntimes/air/> [Diakses 1 Mei 2013]

Alif, A., 2013. *Komputasi cerdas untuk pemula*. Malang: ABC Press.

Angriawan, B., 2014. *Sistem pakar untuk penentuan kondisi tubuh ideal atlet sepakbola usia remaja*. S1. Universitas Malang Raya.

Berndtsson, M., Hansson, J., Olsson, B. & Lundell, B., 2008. *Thesis projects: a guide for students in Computer Science and Information Systems*. 2nd ed. London: Springer-Verlag London Limited.

Boots Group Plc., 2003. *Corporate social responsibility*. [online] Boots Group Plc. Tersedia di: <http://www.boots-plc.com/information/info.asp?id1=1a> [Diakses 1 April 2004]

Brata, K.C., 2012. *Rancang bangun aplikasi jejaring sosial kampus berbasis GPS pada ponsel cerdas Android*. S1. Universitas Brawijaya. Tersedia di <http://ptiik.ub.ac.id/skripsi> [Diakses 1 Agustus 2014]

British Standards Institution, 2011*. BS EN 594:2011 Timber structures. Test methods. Racking strength and stiffness of timbre frame wall panels*. British Standards Online [online] Tersedia melalui: Anglia Ruskin University Library <http://libweb.anglia.ac.uk> [Diakses 31 Augustus 2011]

Brodjonegoro, A., 2009a. *Dunia teknologi informasi bagi komunitas* *open source.* Bandung: Bandung Indah Press.

Brodjonegoro, A., 2009b. *Peran media sosial dalam pemasaran produk perangkat lunak.* Bandung: Bandung Indah Press.

Broughton, J.M., 2002a. The Brettow Woods proposal: a brief look. *Political Science Quarterly*, 42(6), p.564.

Broughton, J.M., 2002b. The Brettow Woods proposal: a brief look. *Political Science Quarterly*, [e-journal] 42(6). Tersedia melalui: Perpustakaan Universitas BX <http://perpustakaan.ubx.ac.id> [Diakses 1 Juli 2013]

Brown, J. 2005. Evaluating surveys of transparent governance. In: UNDESA (United Nations Department of Economic and Social Affairs), 2005. *6th Global forum for reinventing government: towards participatory dan transparent governance*. Seoul, Republic of Korea, 24-27 May 2005. New York: United Nations.

Cakraningrat, R., 2011. *Sistem pendukung keputusan untuk UMKM*. [e-book]. UBX Press. Tersedia melalui: Perpustakaan Universitas BX <http://perpustakaan.ubx.ac.id> [Diakses 1 Juli 2013]

Cookson, J. dan Church, S. eds., 2007. *Leisure and the tourist*. [e-book] Wallingford: ABS Publishers. Tersedia di: Google Books <http://booksgoogle.com> [Diakses 1 Juli 2009]

Cox, C., Brown, J.T. dan Tumpington, W.T., 2002. What health care assistants know about clean hands. *Nursing Today*, Spring Issue, pp.64-68.

Diponegoro, A., 2008. *The beauty of Indonesian oceans*. [electronic print] Tersedia di: <http://adiponegoro.com/store/product\_info.php?cPath=3&  
productss\_id=99> [Diakses 1 Januari 2011]

Esemka, 2012. *Esemka bisa*. [image online] Tersedia di: <http://www.esemka.co.id/esemkabisa.aspx> [Diakses 31 Januari 2011]

Goalie, D. 2008. Remote sensing technology for modern soccer. *Popular science and Technology*, [online] Tersedia di: <http://www.popsci.com/b012378/  
soccer.html> [Diakses 1 Juli 2009]

Haryanto, A. 2002. *Dua dunia*. [foto] (Koleksi pribadi Alan Haryanto)

Higher Education Act 2004. (c.8). London: HMSO

International Standards Office, 1998. *ISO 690 – 2 Information and documentation: Bibliographical references: Electronic documents*. Geneva: ISO.

Kartolo, R., 2010. *Wawancara pada Kabar Pagi*. Diwawancara oleh Sam Basman [televisi] TVRI Saluran 1, 17 Agustus 2010, 08:30.

Keene, E., ed., 1988. *Natural language*. Cambridge: University of Cambridge Press.

Kementerian Komunikasi dan Informatika, 2013. *Laporan Tahunan Layanan Informasi Publik Tahun 2012*. [pdf] Kementerian Komunikasi dan Informatika. Tersedia di: <http://publikasi.kominfo.go.id/bitstream/handle/  
54323613/976/laporan-dan-evaluasi-ppid-tahun-2012-ditambahkan-cover-untuk-online-ppid.pdf> [Diakses 1 Agustus 2014]

NHS Evidence, 2003. *National Library of Guidelines*. [online] Tersedia di: <http://www.library.nhs.uk/guidelinesfinder> [Diakses 1 Juli 2007]

Rahardjo, S. 2001. *Presiden Habibie*. [foto] (Jakarta, Koleksi Museum Presiden)

Richmod, J., 2005. *Customer expectations in the world of elctronic banking: a case study of the Bank of Britain*. PhD. Anglia Ruskin University.

Rumbaugh, J., Jacobson, I. & Booch, G., 2005. *The Unified Modeling Language reference manual*. 2nd ed. Boston: Addison-Wesley.

Samson, C., 1970. Problems of information studies in history. Dalam: S. Stone, ed. 1980. *Humanities information research*. Sheffield: CRUS. pp. 44-68.

Scottish Intercollegiate Guidelines, 2001. *Hypertension in the elderly*. (SIGN publication 20) [online] Edinburgh: SIGN (Diterbitkan 2001) Tersedia di:<http://www.sign.ac.uk/sign49.pdf> [Diakses 22 November 2004]

Silverman, D.F. dan Propp, K.K. eds., 1990. *The active interview*. BeverlyHills, CA: Sage.

Smith, J., 1975. A source of information. Dalam: W. Jones, ed. 2000. *One hundred and one ways to find information about health*. Oxford: Oxford University Press. Ch.2.

Sommerville, I., 2011. *Software engineering*. 9th ed. London: Addison-Wesley.

Sudirman, Z., 2011. *Pembahasan tentang sitasi dan perujukan*. [surat] (Komunikasi personal, 11 Juni 2011).

Tanenbaum, A.S., 1998. *Organisasi komputer terstruktur, jilid 1*. Diterjemahkan dari Bahasa Inggris oleh T.A.H Al-Hamdany. 2001. Jakarta: Salemba Teknika.

Thompson, A. dan Thomson, B., (in press) Innocent or guilty: a studi to ascertain the status of convicts in highly uncertain situations. *Journal of Crime Scene Investigation*. (Diterima untuk publikasi Januari 2002).

Undang-undang Republik Indonesia nomor 12 tahun 2012 tentang Pendidikan Tinggi. Jakarta: Kementerian Sekretariat Negara Republik Indonesia.

UNDESA (United Nations Department of Economic and Social Affairs), 2005. *6th Global forum for reinventing government: towards participatory dan transparent governance*. Seoul, Republic of Korea, 24-27 May 2005. New York: United Nations.

1. PERSYARATAN FISIK DAN TATA LETAK
   1. Kertas

Kertas yang digunakan adalah HVS 70 mg berukuran A4. Apabila terdapat gambar-gambar yang menggunakan kertas berukuran lebih besar dari A4, hendaknya dilipat sesuai dengan aturan yang berlaku. Pengetikan hanya dilakukan pada satu muka kertas, tidak bolak balik.

* 1. Margin

Batas pengetikan naskah adalah sebagai berikut :

* Margin kiri: 4 cm
* Margin atas: 3 cm
* Margin kanan: 3 cm
* Margin bawah: 3 cm
  1. Jenis dan ukuran huruf

Jenis huruf yang dipakai dalam skripsi adalah Calibri dengan ketentuan sebagai berikut:

* Judul bab pada level 1 berukuran 16 pt
* Judul sub bab/seksi pada level 2 berukuran 14 pt
* Judul sub bab/seksi pada level 3 berukuran 14 pt
* Judul sub bab/seksi pada level 4 berukuran 12 pt
* Badan teks berukuran 12 pt

Penggunaan jenis dan ukuran ini harus konsisten. Untuk memudahkan memelihara konsistensi sekaligus penyusunan struktur skripsi, fasilitas seperti *styles* dan *multilevel list* dalam program pengolah kata dapat digunakan. Sebuah *template* untuk skripsi ini telah disediakan untuk membantu mahasiswa. *Styles* dan *multilevel list* dalam template tersebut sudah dirancang untuk jenis dan ukuran huruf yang disyaratkan.

* 1. Spasi

Jarak standar antar baris dalam badan teks adalah satu spasi. Jarak antar paragraf, antara judul bab dan judul sub bab, antara judul sub bab dan badan teks, dan seterusnya, dapat dilihat pada masing-masing *style* yang digunakan dan tersedia dalam *template* untuk skripsi ini.

* 1. Kepala bab, sub bab, dan seksi

Kepala bab terdiri dari kata “BAB” yang diikuti dengan nomor bab dan judul dari bab tersebut, misalnya “BAB 1 PENDAHULUAN” . Kepala sub bab diawali dengan nomor sesuai tingkat hirarkinya dan diikuti dengan judul sub bab, misalnya “1.2 Rumusan masalah”. Penomoran sub bab disarankan tidak lebih dari 4 level (maksimal sub bab X.X.X.X). Kepala bab, sub bab, dan seksi tidak boleh mengandung *widow* atau *orphan* sehingga nampak menggantung atau terputus di bagian awal atau akhir sebuah halaman. *Widow* adalah sebuah paragraf dengan hanya satu baris pertama pada akhir halaman sedangkan sisanya berada pada halaman berikutnya. *Orphan* adalah baris terakhir dari satu paragraf yang tertulis pada awal suatu halaman sedangkan baris lainnya dari paragraf tersebut berada pada halaman sebelumnya.

* 1. Nomor halaman

Bagian awal skripsi menggunakan nomor halaman berupa angka Romawi kecil (i, ii, iii, iv, dan seterusnya) yang dimulai dari sampul dalam. Sedangkan bagian utama dan bagian akhir skripsi menggunakan nomor halaman berupa angka Arab (1, 2, 3, dan seterusnya) yang dimulai dari bab 1. Semua nomor halaman diletakkan di tengah bawah.

1. PENGGUNAAN BAHASA

Bahasa yang dipakai dalam skripsi adalah bahasa Bahasa Indonesia yang baku. Setiap kalimat harus memiliki subjek dan predikat, dan umumnya dilengkapi dengan objek, pelengkap, atau keterangan. Setiap paragraf biasanya terdiri dari beberapa kalimat. Penuturan isi dalam kalimat, paragraf, maupun antar paragraf harus menggunakan bahasa yang tepat dan menggambarkan alur logika yang runtut.

Penulisan bahasa asing yang sudah diserap dalam Bahasa Indonesia disesuaikan dengan kaidah Bahasa Indonesia. Sedapat mungkin dihindari penggunaan bahasa asing jika istilah dalam bahasa Indonesia sudah ada. Jika terpaksa menggunakan istilah dalam bahasa asing, maka penulisannya harus sesuai ejaan aslinya dan dicetak miring (*italic*), kecuali jika istilah tersebut adalah nama.

Sebagai referensi untuk penulisan Bahasa Indonesia yang baku, dokumen berikut dapat digunakan:

* Kamus Bahasa Indonesia, Tim Penyusun, Pusat Bahasa Departemen Pendidikan Nasional, Jakarta 2008
* Peraturan Menteri Pendidikan Nasional Republik Indonesia nomor 46 tahun 2009 tentang Pedoman Umum Ejaan Bahasa Indonesia yang Disempurnakan
* Kamus Besar Bahasa Indonesia dalam jaringan (KBBI daring): http://bahasa.kemdiknas.go.id/kbbi/index.php