

**ANALISIS SENTIMEN MAHASISWA TERHADAP MATA KULIAH
DI FAKULTAS ILMU KOMPUTER UNIVERSITAS ICHSAN
GORONTALO MENGGUNAKAN PEMBOBOTAN TF.IDF**

Oleh

HIZRAWATI YAHYA

T3119106

SKRIPSI

**Untuk memenuhi salah satu syarat ujian
guna memperoleh gelar Sarjana**



**PROGRAM SARJANA
TEKNIK INFORMATIKA
UNIVERSITAS ICHSAN GORONTALO
GORONTALO
2023**

PERSETUJUAN SKRIPSI

**ANALISIS SENTIMEN MAHASISWA TERHADAP MATA KULIAH
DI FAKULTAS ILMU KOMPUTER UNIVERSITAS ICHSAN
GORONTALO MENGGUNAKAN PEMBOBOTAN TF.IDF**

Oleh

HIZRAWATI YAHYA

T3119106

SKRIPSI

Untuk memenuhi salah satu syarat ujian
guna memperoleh gelar Sarjana
Program Studi Teknik Informatika,
ini telah disetujui oleh Tim Pembimbing
Gorontalo,

Pembimbing I

Ksc



Zohrahayaty, S.Kom, M.Kom

NIDN : 0912117702

Pembimbing II



Mohamad Efendi Lasulika, M.Kom

NIDN : 0929048902

PENGESAHAN SKRIPSI

**ANALISIS SENTIMEN MAHASISWA TERHADAP MATA KULIAH
DI FAKULTAS ILMU KOMPUTER UNIVERSITAS ICHSAN
GORONTALO MENGGUNAKAN PEMBOBOTAN TF.IDF**

Oleh

HIZRAWATI YAHYA

T3119106

Diperiksa oleh Panitia Ujian Strata Satu (S1)

Universitas Ichsan Gorontalo

1. Ketua Penguji

HADITSAH ANNUR, M.Kom

2. Anggota

ANDI BODE, M.Kom

3. Anggota

YULIANTY LASENA, M.Kom

4. Anggota

ZOHRAHAYATY, S.Kom, M.Kom

5. Anggota

MOHAMAD EFENDI LASULIKA, M.Kom

Mengetahui

Dekan Fakultas Ilmu Komputer



Ketua Program Studi



PERNYATAAN SKRIPSI

Dengan ini saya menyatakan bahwa :

1. Karya tulis (Skripsi) saya ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana) baik di Universitas Ichsan Gorontalo maupun diperguruan tinggi lainnya.
2. Karya tulis (Skripsi) saya ini adalah murni gagasan, rumusan, dan penelitian saya sendiri, tanpa bantuan pihak lain, kecuali arahan dai Tim Pembimbing.
3. Dalam karya tulis (Skripsi) saya ini tidak terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali secara tertulis dicantumkan sebagai acuan/sitasi dalam naskah dan dicantumkan pula dalam daftar pustaka.
4. Peryataan ini saya buat dengan sesungguhnya dan apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, saya bersedia menerima sanksi akademi berupa pencabutan gelar yang telah diperoleh karena karya tulis ini, serta sanksi lainnya sesuai dengan norma-norma yang berlaku di Universitas Ichsan Gorontalo.

Kota, Mei 2023

Yang Membuat Pernyataan,

A red official stamp from the Indonesian Ministry of Education, Culture, and Higher Education (Kemendikbudristek) is visible. The stamp contains the text 'KEMENDIKBUDRISTEK', 'METRAKAI', 'TEMPEL', and a unique identification number '82AKX475175080'. A handwritten signature in black ink is written over the stamp.

Hizrawati Yahya

ABSTRACT

HIZRAWATI YAHYA. T3119106. THE ANALYSIS OF STUDENT SENTIMENT TOWARDS COURSES AT THE FACULTY OF COMPUTER SCIENCE OF UNIVERSITAS ICHSAN GORONTALO USING TF.IDF WEIGHTING

Student satisfaction with the courses taught by lecturers is a problem in a university. Universitas Ichsan Gorontalo implements online questionnaire filling as a student assessment to assess a lecturer's performance on the courses concerned. At Universitas Ichsan Gorontalo, especially the Faculty of Computer Science, filling out questionnaires is implemented. It is mandatory for all students in the Faculty of Computer Science as one of the requirements for accessing grades in the final exam and taking courses in the following semester. Student assessment during learning in each course is said to be important to improve the learning process. Therefore, this study aims to determine the level of student satisfaction with the course through lecturer performance. In this study, sentiment analysis is carried out using Term Frequency-Inverse Document Frequency (TF.IDF) weighting and looking at the level of similarity between documents (comments) using Cosine Similarity. The results of this study indicate the level of student satisfaction with courses through the results of lecturer performance. Of 1,010 data, there are 942 preprocessing results with 423 positive sentiments, 293 neutral sentiments, and 226 negative sentiments. The results of TF.IDF obtains an accuracy of 85.7%.

Keywords: courses, sentiment analysis, TF.IDF, Cosine Similarity



ABSTRAK

HIZRAWATI YAHYA. T3119106. ANALISIS SENTIMEN MAHASISWA TERHADAP MATA KULIAH DI FAKULTAS ILMU KOMPUTER UNIVERSITAS ICHSAN GORONTALO MENGGUNAKAN PEMBOBOTAN TF.IDF

Kepuasan mahasiswa terhadap mata kuliah yang diajarkan dosen ialah suatu hal dalam sebuah perguruan tinggi. Universitas Ichsan Gorontalo menjadi salah satu yang menerapkan pengisian kuesioner secara online sebagai penilaian mahasiswa untuk menilai suatu kinerja dosen terhadap mata kuliah yang bersangkutan. Di Universitas Ichsan Gorontalo khususnya Fakultas Ilmu Komputer telah menerapkan pengisian kuesioner, hal ini diterapkan wajib untuk seluruh mahasiswa di fakultas ilmu komputer sebagai salah satu syarat mengakses nilai di ujian akhir dan mengontrak mata kuliah di semester berikutnya. Penilaian mahasiswa selama pembelajaran di setiap mata kuliah dikatakan penting agar dapat meningkatkan proses pembelajaran. Oleh karena itu, penelitian ini bertujuan agar dapat mengetahui tingkat kepuasan mahasiswa terhadap mata kuliah melalui kinerja dosen. Pada penelitian ini dilakukan dengan analisis sentimen menggunakan pembobotan *Term Frequency-Inverse Document Frequency* (TF.IDF) dan melihat tingkat kemiripan antar dokumen (komentar) menggunakan *Cosine Similarity*. Hasil penelitian ini tujuannya untuk melihat tingkat kepuasan dari mahasiswa terhadap mata kuliah melalui hasil kinerja dosen dari 1.010 data terdapat 942 hasil *preprocessing* dengan terdapat 423 sentimen positif, 293 sentimen netral dan 226 sentimen negatif. Hasil dari TF.IDF yang didapatkan akurasi sebesar 85,7%.

Kata kunci: mata kuliah, analisis sentimen, TF.IDF, *Cosine Similarity*

KATA PENGANTAR

Alhamdulillah, penulis dapat menyelesaikan usulan penelitian proposal ini dengan judul “Analisis Sentimen Mahasiswa Terhadap Mata Kuliah Di Fakultas Ilmu Komputer Universitas Ichsan Gorontalo Menggunakan Pembobotan TF.IDF”, untuk memenuhi salah satu syarat penyusunan Skripsi Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.

Penulis menyadari sepenuhnya bahwa usulan penelitian ini tidak mungkin terwujud tanpa bantuan dan dorongan dari berbagai pihak, baik bantuan moril maupun materil. Untuk itu, dengan segala keikhlasan dan kerendahan hati, penulis mengucapkan banyak terima kasih dan penghargaan yang setinggi-tingginya kepada :

1. Bapak Muh. Ichsan Gaffar, SE.,M.Ak, selaku Keta Yayasan Pengembangan Ilmu Pengetahuan dan Teknologi (YPIPT) Ichsan Gorontalo
2. Bapak Dr. Abdul Gaffar La Tjokke, M.Si, selaku Rektor Universitas Ichsan Gorontalo
3. Bapak Irvan A Salihi, M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo
4. Bapak Sudirman Melangi, M.Kom selaku pembantu Dekan I Bidang Akademik Fakultas Ilmu Komputer Universitas Ichsan Gorontalo
5. Ibu Irma Surya Kumala Idris, M.Kom selaku Pembantu Dekan II Bidang Administrasi Umum dan Keuangan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo
6. Bapak Sudirman S. Panna, M.Kom, selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo
7. Ibu Zohrahayaty, M.Kom, selaku Pembimbing I
8. Bapak Mohamad Efendi Lasulika, M.Kom, selaku Pembimbing II
9. Bapak/Ibu Dosen Universitas Ichsan Gorontalo yang telah mendidik dan mengajarkan berbagai disiplin ilmu kepada penulis
10. Kedua Orang Tua saya yang tercinta, atas segala kasih sayang, jerih payah dan doa restunya dalam membesarkan dan mendidik penulis

11. Rekan-rekan seperjuangan Sitti Meautiah D.A Jan dan Putri Siti Salsabila Ibrahim, serta teman-teman dekat yang telah banyak memberikan batuan dan mendukung moril yang sangat besar kepada penulis
12. Kepada semua pihak yang ikut membantu dalam penyelesaian proposal/skripsi ini yang tak sempat penulis sebut satu-persatu

Semoga Allah SWT melimpahkan balasan atas jasa-jasa mereka kepada kami. Penulis menyadari sepenuhnya bahwa apa yang telah tercapai ini masih jauh dari kesempurnaan dan masih banyak terdapat kekurangan. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran konstruktif. Akhirnya penulis berharap semoga hasil yang telah dicapai ini dapat bermanfaat bagi kita semua, Amin.

Gorontalo, April 2023

Hizrawati Yahya

DAFTAR ISI

HALAMAN JUDUL	i
PERSETUJUAN SKRIPSI	ii
PENGESAHAN SKRIPSI	iii
PERNYATAAN SKRIPSI	iv
ABSTRAK	v
KATA PENGANTAR.....	vii
DAFTAR ISI	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Identifikasi Masalah	5
1.3. Rumusan Masalah	5
1.4. Tujuan Penelitian	5
1.5. Manfaat Penelitian	5
BAB II LANDASAN TEORI.....	7
2.1. Tinjauan Studi.....	7
2.2. Tinjauan Pustaka.....	9
2.2.1. Mata Kuliah	9
2.2.2. Data Mining.....	11
2.2.3. Text Mining.....	11
2.2.4. Analisis Sentimen	12
2.2.5. Preprocessing	12

2.2.6.	Algoritma TF.IDF	15
2.2.7.	<i>Cosine Similarity</i>	16
2.2.8.	<i>Confussion Matrix</i>	17
2.3.	Kerangka Pikir	19
BAB III METODE PENELITIAN		20
3.1.	Jenis, Metode, Subjek, Objek, Waktu, dan Lokasi Penelitian....	20
3.2.	<i>Dataset</i>	20
3.3.	Pemodelan	21
3.4.	<i>Preprocessing</i>	22
3.5.	Pembobotan TF.IDF.....	22
3.6.	<i>Cosine Similarity</i>	23
3.7.	Hasil Klasifikasi.....	23
3.8.	Evaluasi	23
BAB IV HASIL PENELITIAN		24
4.1.	Pengumpulan Data	24
4.2.	Model Usulan.....	25
4.2.1.	<i>Pra Processing</i>	26
4.2.2.	Pembobotan Dokumen (TF-IDF)	33
4.2.3.	Perhitungan <i>Cosine Similarity</i>	45
BAB V HASIL EVALUASI.....		51
5.1.	Evaluasi Model	51
5.1.1.	<i>Confusion Matriks</i>	53
5.2.	Uji Coba Model.....	54
5.2.1.	<i>Pra Processing</i>	55
5.2.2.	Pembobotan Dokumen (TF.IDF).....	57

5.2.3. <i>Cosine Similarity</i>	57
BAB VI PENUTUP	59
6.1. KESIMPULAN	59
6.2. SARAN	59
DAFTAR PUSTAKA	61
LAMPIRAN	63

DAFTAR GAMBAR

Gambar 2.1 Kerangka Pikir.....	19
Gambar 3.1 Pemodelan	21
Gambar 4.1 Model yang di Usulkan	25
Gambar 4.2 <i>Pra Precossing</i>	26
Gambar 4.3 Hasil <i>Case Folding</i>	27
Gambar 4.4 Hasil <i>Tokenizing</i>	28
Gambar 4.5 Hasil <i>Stopword/Number Removal</i>	29
Gambar 4.6 Hasil <i>Stemming</i>	30
Gambar 4.7 Tahapan TF.IDF	33
Gambar 4.8 Visualisasi TF.IDF.....	44
Gambar 4.9 Tahapan <i>Cosine Similarity</i>	45
Gambar 5.1 Model Visualisasi	54
Gambar 5.2 Visualisasi <i>Case Folding</i>	55
Gambar 5.3 Visualisasi <i>Tokenizing</i>	55
Gambar 5.4 Visualisasi <i>Stopword/Number Removal</i>	56
Gambar 5.5 Visualisasi <i>Stemming</i>	56
Gambar 5.6 Visualisasi Perhitungan TF.IDF	57
Gambar 5.7 Visualisasi <i>Cosine Similarity</i>	57

DAFTAR TABEL

Tabel 1 <i>State Of The Art</i>	7
Tabel 2 Sampel Data Kuesioner Mahasiswa	9
Tabel 3 Proses <i>Case Folding</i>	12
Tabel 4 Proses <i>Tokenizing</i>	13
Tabel 5 Proses <i>Filtering Stopword</i>	13
Tabel 6 Proses <i>Stemming</i>	15
Tabel 2.7 Visualisasi <i>Confusion Matrix</i>	18
Tabel 4.1 Hasil Pengumpulan Data	24
Tabel 4.2 Data Saran	26
Tabel 4.3 Hasil <i>Pra Processing</i>	30
Tabel 4.4 Hasil Pemilihan <i>Term</i>	32
Tabel 4.5 Menghitung Kemunculan Kata	34
Tabel 4.6 Perhitungan DF	35
Tabel 4.7 Perhitungan IDF	40
Tabel 4.8 Perhitungan TF.IDF	43
Tabel 4.9 Perhitungan Q dan D	46
Tabel 4.10 Perhitungan Akar Kuadrat <i>Cosine Similarity</i>	47
Tabel 4.11 Pengurutan Tingkat Kemiripan	49
Tabel 5.1 Data Validasi	51
Tabel 5.2 Hasil Pengujian Model	52
Tabel 5.3 <i>Confusion Matrix</i>	53
Tabel 5.4 Perhitungan <i>Precision</i>	53
Tabel 5.5 Perhitungan <i>Recall</i>	54

DAFTAR LAMPIRAN

Lampiran 1 : HASIL TURNITIN

Lampiran 2 : Listing Program

Lampiran 3 : Dataset

Lampiran 4 : RIWAYAT HIDUP PENELITI

Lampiran 5 : Surat Penelitian

Lampiran 6 : Surat Keterangan Penelitian

Lampiran 7 : Surat Bebas Plagiasi

Lampiran 8 : Surat Bebas Pustaka

BAB I

PENDAHULUAN

1.1. Latar Belakang

Institut seperti perguruan tinggi di Indonesia sudah memasuki era dimana saling bersaing baik dari kondisi fasilitas hingga kualitas sumber daya manusia yang berkompeten, maka hal tersebut merupakan salah satu peran dan tanggung jawab dari seorang dosen agar menciptakan mahasiswa yang kompeten [1]. Pembelajaran yang ada perguruan tinggi ditugaskan kepada dosen dengan tujuan memberikan lulusan yang berkompeten dan berkualitas [2]. Dosen menjadi salah satu yang berperan dalam kegiatan belajar mengajar serta berkaitan dengan mahasiswa. Dosen senantiasa diharuskan meningkatkan pengetahuan dan proses mengajar lebih efisien dan efektif agar ilmu yang diberikan dapat diterima dengan baik oleh mahasiswa [3]. Adanya mahasiswa dapat menilai dosen sehingga perguruan tinggi dapat menilai kinerja dosen melalui pengisian kotak saran mahasiswa terhadap dosen pada setiap mata kuliah. Hal ini menjadi bagian dari penilaian di perguruan tinggi untuk mengetahui evaluasi pembelajaran terhadap dosen melalui tingkat kepuasan mahasiswa ke dalam aspek penilaian.

Dengan adanya penilaian mahasiswa khususnya mahasiswa di Universitas Ichsan Gorontalo yang merupakan salah satu lembaga perguruan tinggi swasta menerapkan suatu penilaian dengan pengisian kuesioner untuk kotak saran [4], sebagai penilaian mata kuliah terhadap kinerja dosen dan fasilitas kampus. Dengan adanya pengisian kotak saran untuk tingkat kepuasan pihak kampus dapat mengumpulkan opini mahasiswa melalui pengisian kuesioner kotak saran [1]. Pada setiap akhir semester mahasiswa Universitas Ichsan Gorontalo melakukan pengisian kotak saran terhadap suatu penilaian baik dari proses belajar terhadap mata kuliah dan sarana prasarana kampus [4]. Kuesioner ini diisi mahasiswa secara online yang terdiri dari data tingkat kepuasan hingga disimpulkan melalui pengisian kotak saran terhadap mata kuliah dosen yang bersangkutan [3]. Di Universitas Ichsan Gorontalo khususnya di Fakultas Ilmu Komputer menerapkan pengisian kuesioner secara online kepada mahasiswa ialah menjadi syarat utama

untuk mengakses nilai di akhir semester serta syarat untuk mengontrak semester berikutnya.

Pengisian kuesioner untuk mahasiswa di fakultas ilmu komputer Universitas Ichsan Gorontalo umumnya berisikan tingkat kepuasan serta pengisian kotak saran baik positif maupun negatif yang menjadi alasan utama digunakan sebuah penyedia layanan [5]. Hal ini diadakan untuk menjadi bahan penelitian untuk mengetahui kepuasan mahasiswa melalui kotak saran di fakultas ilmu komputer terhadap mata kuliah sebagai penilaian untuk kinerja dosen yang menjadi salah satu sumber data untuk analisis sentimen hasil evaluasi dengan menggunakan data yang diperoleh.

Pada penelitian ini hanya menggunakan data hasil dari komentar pada kotak saran mahasiswa di fakultas ilmu komputer terdiri dari komentar-komentar yang secara acak. Tujuan peneliti ini menggunakan data pengisian kotak saran ialah melihat mata kuliah mana yang paling banyak disukai mahasiswa dalam proses pembelajaran apakah mahasiswa menerima dengan baik sehingga hal ini dapat dilihat dari suatu kinerja dosen dari mata kuliah tersebut. Untuk saat ini data kuesioner yang diisi oleh mahasiswa melalui kotak saran yang terkumpul belum seberapa maksimal dan efektif diketahui hasil dari tingkat kepuasan mahasiswa terhadap mata kuliah dosen yang secara acak terdiri dari komentar positif, negatif maupun ambigu yang bersifat ganda. Dengan data yang diperoleh dari hasil pengisian kotak saran mahasiswa maka peneliti menguji coba menggunakan metode perhitungan TF.IDF dengan teknik analisis sentimen untuk mengklasifikasi suatu data dari kotak saran dengan tujuan menilai serta meningkatkan proses pembelajaran agar lebih efektif.

Terdapat hal yang dilakukan untuk mengetahui hasil dengan cara mengklasifikasi secara manual tetapi hal tersebut bias saja memakan waktu. Klasifikasi merupakan proses dimana melakukan pengelompokan suatu kelas berbeda menjadi satu kelas yang sama dengan suatu kesamaan. Tugas dari klasifikasi berupa 1). Membentuk suatu model memori yang akan menyimpan prototipe, 2). Menggunakan model untuk klasifikasi yang telah dibentuk untuk sebuah objek data lain sehingga dapat mengetahui suatu kelas objek [1]. Terdapat

suatu solusi untuk mempercepat dalam mengetahui hasil secara otomatis ialah dengan melakukan evaluasi analisis sentimen agar mempermudah mengetahui hasil kotak saran pada tingkat kepuasan mahasiswa terhadap mata kuliah.

Analisis sentimen ialah fokus pada seseorang yang memberikan pendapat secara sentiment baik dari positif dan negative yang berkaitan dengan pengolahan kata tertentu dan menjadi salah satu proses memperoleh hasil sentimen di fakultas ilmu komputer yang diklasifikasi menggunakan metode TF.IDF (*Term Frequency-Inverse Document Frequency*) [6]. Analisis sentimen bertugas untuk mengelompokkan teks dalam dokumen baik dari kalimat sehingga dari tanggapan ini dapat diklasifikasikan dalam aspek positif ataupun negatif dengan metode TF.IDF (*Term Frequency-Inverse Document Frequency*) [1].

TF.IDF (*Term Frequency-Inverse Document Frequency*) digunakan untuk mengklasifikasi suatu tanggapan/opini yang bersifat positif ataupun negatif [1]. Keuntungan dari penggunaan metode TF.IDF (*Term Frequency-Inverse Document Frequency*) ialah metode klasifikasi sederhana yang digunakan dalam ringkasan sehingga mempermudah menghitung kesamaan kata yang sering digunakan. Metode TF.IDF (*Term Frequency-Inverse Document Frequency*) ini menghitung nilai *tf* dan *idf* antar kata yang sering muncul [7]. Sehingga hasil dari pembobotan kata menggunakan TF.IDF diperoleh dari hasil perkalian *tf* dan *idf*. Data yang akan digunakan dalam penelitian ini ialah data dari pengisian kotak saran mahasiswa di fakultas ilmu komputer dengan berbahasa Indonesia dalam pembobotan kata menggunakan metode TF.IDF (*Term Frequency-Inverse Document Frequency*) [8].

Berdasarkan penelitian sebelumnya, dengan judul Analisis Sentimen Pada Data Kuesioner Evaluasi Dosen Oleh Mahasiswa Menggunakan Algoritma *Support Vector Machine* [3] menghasilkan 1140 komentar positif, 84 negatif dan 395 netral dengan nilai akurasi 75%. Peneliti selanjutnya, [8] tentang Analisis Sentimen Penggunaan Twitter Terhadap Polemik Pesepakbolaan Indonesia Menggunakan Pembobotan TF.IDF dan *K-Nearest Neighbor* dari 2.000 data *tweet* menghasilkan 790 *tweet* positif dan 1.210 *tweet* negatif dengan penggunaan metode KNN dan TF.IDF menghasilkan nilai akurasi sebesar 79,99%. Dan

penelitian[5], tentang Analisis Sentimen Kepuasan Pelanggan Pada Penyedia Layanan Pengantaran Makanan dari 1.000 data *tweet* yang diperoleh menghasilkan 343 *tweet* positif, 176 *tweet* negatif dan 481 *tweet* netral. Kemudian memperoleh nilai akurasi sebesar 86%. Berdasarkan uraian diatas disimpulkan analisis sentimen dengan menggunakan metode klasifikasi TF.IDF (*Term Frequency-Inverse Document Frequency*) dapat dikatakan efektif karena dari hasil tingkat akurasi yang dihasilkan baik sehingga metode ini membantu untuk mengetahui dosen dimata kuliah yang paling banyak diminati mahasiswa secara otomatis dan efektif sehingga tidak diperlukan lagi proses perhitungan secara manual dengan adanya metode klasifikasi TF.IDF (*Term Frequency-Inverse Document Frequency*).

Pada penelitian ini, peneliti hanya berfokus untuk uji coba perhitungan menggunakan Algoritma TF.IDF terhadap analisis data pada teks serta mengetahui hasil tingkat akurasi dari algoritma TF.IDF. Tujuan dari menggunakan metode TF.IDF (*Term Frequency-Inverse Document Frequency*) agar dapat mengetahui seberapa akurat metode ini digunakan untuk analisis sentimen dalam mengklasifikasi. Sehingga hasil akhirnya dapat diketahui jumlah komentar positif dan negatif terhadap mata kuliah untuk kinerja dosen dari pengisian kotak saran oleh mahasiswa dan mengetahui nilai akurasi dari penggunaan metode TF.IDF. Jadi hasil dari analisis pengolahan bertujuan mengetahui kepuasan mahasiswa terhadap mata kuliah yang diajarkan oleh dosen sehingga dapat dievaluasi dan dapat meningkatkan cara pembelajaran sehingga bisa lebih efektif lagi pada proses pembelajaran melalui kinerja dosen.

Pada proses perhitungan dari algoritma TF.IDF untuk analisis data digunakan suatu tahapan berupa *preprocessing* untuk menghasilkan data yang secara teratur. Dalam *preprocessing* atau (*Natural Language Preprocessing*) ialah suatu proses dimana melakukan proses pelabelan secara otomatis dengan suatu teknik berupa *case folding*, *cleansing*, *tokenizing*, *filtering stopword*, hingga *stemming* dengan menggunakan *tools* pada python dengan *library streamlit* untuk memproses tahapan *preprocessing* dengan menggunakan *Natural Language Tool Kit* (NLTK) [9]. Pada penggunaan python untuk tahapan *preprocessing* NLTK

digunakan untuk mengolah data dalam jumlah yang banyak untuk menganalisis data.

Berdasarkan penjelasan diatas, peneliti tertarik melakukan penelitian dengan judul **“Analisis Sentimen Mahasiswa Terhadap Mata Kuliah di Fakultas Ilmu Komputer Universitas Ichsan Gorontalo Menggunakan Pembobotan TF.IDF”**

1.2. Identifikasi Masalah

Proses mengklasifikasi yang jika dilakukan secara manual akan memakan waktu dan tenaga, sehingga diperlukan suatu metode untuk mempermudah dan membantu proses mengklasifikasi hasil pengisian kotak saran terkait komentar terhadap tingkat kepuasan melalui data teks sehingga dapat memperoleh jumlah nilai positif dan negatif dan nilai akurasi.

1.3. Rumusan Masalah

Bagaimana hasil akurasi yang dihasilkan dari uji coba perhitungan Algoritma TF.IDF (*Term Frequency-Inverse Document Frequency*) dan *Cosine Similarity* pada Analisis Sentimen terhadap tingkat kepuasan mahasiswa melalui data pada kotak saran untuk mata kuliah?

1.4. Tujuan Penelitian

Memperoleh hasil akurasi yang baik dari uji coba perhitungan Algoritma TF.IDF *Cosine Similarity* dari kotak saran tingkat kepuasan mahasiswa terhadap mata kuliah yang paling disenangi terhadap kinerja dosen.

1.5. Manfaat Penelitian

1. Manfaat Teoritis : memberikan masukan bagi perkembangan ilmu pengetahuan dan teknologi, khususnya dibidang ilmu komputer yaitu berupa permutakhiran metode analisis sentimen serta uji coba metode TF.IDF (*Term Frequency-Inverse Document Frequency*)
2. Manfaat Praktis : sumbangan pemikiran, karya, bahan pertimbangan atau solusi untuk mengambil suatu keputusan dalam rangka mengambil suatu analisis yang akurat peneliti selanjutnya dapat memberikan suatu

kontribusi untuk pengembangan apabila terdapat kekurangan pada sistem dan metode ini pada penelitian selanjutnya.

BAB II

LANDASAN TEORI

2.1. Tinjauan Studi

Terdapat beberapa penelitian yang sebelumnya telah dilakukan terkait analisis sentimen dan penggunaan metode TF.IDF (*Term Frequency-Inverse Document Frequency*), yaitu sebagai berikut :

Tabel 1 *State Of The Art*

No.	PENELITI	JUDUL	TAHUN	METODE	HASIL
1	Jeremy Andrew Septian,Tresna Maulana Fahrudin, Aryo Nugroho [7]	Analisis Sentimen Pengguna Twitter Terhadap Polemik Pesepakbolan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor	2019	TF.IDF, K-Nearest Neighbor	Dari 2.000 data <i>tweet</i> yang diklasifikasi menghasilkan 790 <i>tweet</i> positif dan 1.210 <i>tweet</i> negatif. Dengan menggunakan metode KNN menghasilkan nilai akurasi sebesar 79,99%
2	Indah Nur Fitri Astuti, Irfan Darmawan, Dita	Analisis Sentimen Pada Data Kuesioner Evaluasi Dosen Oleh Mahasiswa	2020	Algoritma <i>Support Vector Machine</i> , TF.IDF	Dari data yang diklasifikasi sebanyak 2.465

No.	PENELITI	JUDUL	TAHUN	METODE	HASIL
	Pramesti [3]	(EDOM) Prodi Sistem Informasi Telkom University menggunakan Algoritma <i>Support Vector Machine</i>		(<i>Term Frequency-Inverse Document Frequency</i>),	dihasilkan terdapat 1.440 komentar positif, 84 komentar negatif, dan 395 komentar netral. Dari penggunaan Algoritma SVM dan TF.IDF menghasilkan akurasi sebesar 75%.
3	Andrew, Erick Dazki, Richardus Eko Indrajit [5]	Analisis Sentimen Kepuasan Pelanggan Pada Penyedia Layanan Pengantaran Makanan	2022	Algoritma TF.IDF, SGD (<i>Stochastic Gradient Descent</i>), Algoritma SVM (<i>Support Vector Machine</i>), K-Nearest Neighbor	Dari jumlah 1.000 data <i>tweet</i> yang diperoleh hasil terdapat 343 <i>tweet</i> positif, 481 <i>tweet</i> negatif dan 176 <i>tweet</i> netral. Kemudian diklasifikasikan dengan

No.	PENELITI	JUDUL	TAHUN	METODE	HASIL
					metode TF.IDF menghasilkan nilai akurasi sebesar 86% .

2.2. Tinjauan Pustaka

2.2.1. Mata Kuliah

Mata kuliah ialah sesuatu yang dipelajari oleh mahasiswa dan diajarkan oleh seorang dosen dalam perguruan tinggi. Dosen mengajarkan mata kuliah mencakup suatu materi pembelajaran, metode, penilaian yang mengandung unsur sikap, pengetahuan dan keterampilan yang menjadi salah satu aspek dalam satuan kredit semester (SKS) yang menjadi bahan penilaian dari dosen terhadap mahasiswa. Apabila mata kuliah tersebut disenangi oleh mahasiswa maka kinerja dosen dalam mata kuliah tersebut diterima dengan baik oleh mahasiswa [10] .

Kinerja ialah suatu penilaian pada seseorang terkait suatu pencapaian dalam pekerjaan [4]. Kinerja dosen dilihat dari bagaimana mahasiswa berinteraksi dengan dosen tersebut dalam proses pembelajaran mata kuliah dan mahasiswa dapat memberikan penilaian terhadap kinerja dosen pada mata kuliah [11]. Mahasiswa dapat menilai kinerja dosen pada mata kuliah tersebut melalui pemberian saran dalam pengisian kuesioner[12].

Tabel 2 Sampel Data Kuesioner Mahasiswa

NO.	DOSEN	MATA KULIAH	SARAN
1	AZWAR, M. Kom.	ALGORITMA DAN STRUKTUR DATA	insa Allah bisa menjadi dosen yg selalu baik hati #Lebih Semangat Lagi Untuk mengajarkan Ilmu kepada kami #TERBAIK #Terima kasih atas Ilmunya #Lebih ditingkatkan lagi dalam proses pembelajaran #Smoga kampus boalemo kembali seperti dlu Td akan bermasalah lagi. Dan jangan terlalu menekan mahasiswa #Baik jika praktek terus

			dilakukan dalam melakukan pembelajaran #Semoga kedepannya hubungan antar dosen dan mahasiswa tetap terjaga #Terima kasih #Semoga kedepannya lebih baik lagi #Lebih di perhatikan lagi mahasiswanya #Cara mengajar sudah baik #lebih semangat lagi untuk mengajar kami (mahasiswa). #tetap semangat mengajari kami untuk menuntut ilmu. #TERIMA KASIH #Menurut saya pemberian tugas lebih ditingkatkan #Semangat terus dalam mengajar #Tetap Semangat #Terima kasih #Pertahankan cara mengajarnya pak, saya suka cara mengajar bapak Terima kasih. #Kalau bisa membuat contoh latihan di kelas lebih sulit lagi. Supaya kami terbiasa dengan hal-hal yang sulit #Terima kasih #Lebih di perhatikan lagi ke aktifan mahasiswa #Jangan pernah bosan d
2	YASIN ARIL MUSTOFA, M.Kom	ALGORITMA DAN STRUKTUR DATA	Cara mengajar #No komen #Baik #Sangat baik #Baik #Baik #No komen #Tidak ada #Cara mengajar sangat baik #Bapak mengajar dan memberikan materinya sangat baik #Lebih baik lagi kedepannya amin #Lebih mengenal mahasiswa-nya #Lebih ditingkat cara mengajar yang lebih baik lagi #Saya harap Pak Aril dapat mengajar lebih baik lagi dan dapat mengembangkan mahasiswa yang baik #bapak paling baik #terima kasih #Tidak ada #gaga #Baik #Tidak ada #Bapak sangat baik
3	SUHARDI RUSTAM, M.Kom	ALGORITMA DAN STRUKTUR DATA	Tdk ada #Sudah cukup baik #Saya harap dengan hasil kuesioner ini dapat membuat dosen lebih semangat dan tanpa ada rasa bosan dalam meningkatkan pengetahuan mahasiswa #Semoga kedepannya tetap memberikan yang terbaik #Lebih di perbanyak praktek nya dibandingkan materi. #Proses belajar di tingkatkan #Dosen mengajar lebih menyenangkan #Jadilah dosen yang selalu terampil dan tepat waktu #Tidak Ada #Y #Saran saya mohon untuk jam mknnya masuk #Terima

			kasih atas ilmu nya pak #Selalu bimbing kami ya pak [] #Tetap dipertahankan #Lebih memberikan praktek #Bisa lebih ditingkatkan lagi pembelajarannya #Memuaskan #Perlu bimbingan coding #-#Tidak ada #sudah baik #Lebih di tingkatkan lagi cara mengajarnya#Lebih sering memberikan praktek , karna sebagian mahasiswa masih ada yang belum, bisa mengerjakan tugas, seperti coding .#Lebih detail memperluas wawasan mahasiswa #jangan terlalu cepat menjelaskan dalam pembelajaran #The best #Alhamdulillah baik #Lebih sering memberikan praktek #Saran saya sebaiknya mahasiswa lebih dikontro
--	--	--	---

Sumber : Siakun Universitas Ichsan Gorontalo, 2020/2021 Ganjil

2.2.2. Data Mining

Data mining ialah proses mengerjakan suatu teknik dalam menganalisis suatu pola yang diinginkan untuk membantu suatu pengambilan keputusan yang dikhususkan pada pencarian pola data yang diinginkan dan dapat dimengerti dalam pengambilan keputusan yang ingin dikenali [13]. *Data mining* juga dapat menganalisa data yang kemudian dipelajari agar teliti apabila menggunakan suatu perangkat keputusan lain[14] .

2.2.3. Text Mining

Text mining disebut juga *text* dari *Data Mining* yang secara umum mengacu pada suatu proses ekstraksi [7]. *Text mining* merupakan suatu proses agar menghasilkan informasi teks yang didapatkan dengan pola dan tren dalam pola statistik. *Text mining* termasuk pembobotan kata dengan tujuan member nilai bobot untuk *term* yang ada pada suatu dokumen[14]. *Text mining* juga termasuk dalam proses dimana melakukan pencarian suatu informasi dalam sebuah teks dengan jumlah besar.

2.2.4. Analisis Sentimen

Analisis sentimen ini adalah suatu bidang untuk menganalisis pendapat orang terhadap entitas seperti layanan, organisasi, peristiwa hingga suatu topik yang berisikan kritikan ataupun saran [8].

Analisis sentimen/*Opinion mining* ialah proses dimana mengolah dan mengekstrak data dengan otomatis agar memperoleh suatu informasi sentimen yang terdapat kalimat opini[15]. Analisis sentimen ialah menganalisis, mengolah, memahami serta mengekstrak data yang berisikan opini, emosi pada suatu objek baik dari komentar, saran dan ulasan terkait masalah yang akan diteliti dengan tujuan mengetahui opini tersebut bernilai positif, negatif ataupun netral [16].

2.2.5. Preprocessing

Preprocessing adalah tahapan awal mengkondisikan data agar selanjutnya diolah pada klasifikasi [17]. *Preprocessing* bertujuan mengolah data teks yang tidak terstruktur menjadi suatu data yang terstruktur (Andriyani), [4]. Proses dalam menentukan fitur-fitur yang mewakili setiap kata yang ada pada data teks dengan *text preprocessing* menggunakan *tools* yang ada pada python dengan *library streamlit* yang merupakan bagian untuk menampilkan suatu pemodelan. Pada *preprocessing* terdapat suatu library berupa NLTK (*Natural Language Tool Kit*) ialah suatu platform yang banyak digunakan untuk membangun analisis teks [18]. Terdapat beberapa tahapan yang dilakukan dalam *preprocessing* [19]:

- a) *Case Folding* ialah mengubah semua huruf menjadi sebuah huruf kecil (lowercase) dalam suatu kalimat. Tahapan case folding :

Tabel 3 Proses *Case Folding*

Sebelum <i>Case Folding</i>	Sesudah <i>Case Folding</i>
lebih semangat lagi untuk mengajar kami (mahasiswa).	lebih semangat lagi untuk mengajar kami (mahasiswa).
Saya harap Pak Aril dapat mengajar lebih baik lagi dan dapat mengembangkan	saya harap pak aril dapat mengajar lebih baik lagi dan dapat mengembangkan mahasiswa yang

mahasiswa yang baik	baik
jangan terlalu cepat menjelaskan dalam pembelajaran	jangan terlalu cepat menjelaskan dalam pembelajaran

- b) *Tokenizing* ialah tahap pemisahan kata pada suatu kalimat menjadi suatu kata biasa. Tahapan *tokenizing* :

Tabel 4 Proses *Tokenizing*

Sebelum <i>Tokenizing</i>	Sesudah <i>Tokenizing</i>
lebih semangat lagi mengajar kami mahasiswa	[lebih], [semangat], [lagi], [mengajar], [kami], [mahasiswa]
saya harap pak aril dapat mengajar lebih baik lagi dan dapat mengembangkan mahasiswa yang baik	[saya], [harap], [pak], [aril], [dapat], [mengajar], [lebih], [lebih], [baik], [lagi], [dan], [dapat], [mengembangkan], [mahasiswa], [yang], [baik]
jangan terlalu cepat menjelaskan dalam pembelajaran	[jangan], [terlalu], [cepat], [menjelaskan], [dalam], [pembelajaran]

- c) *Filtering Stopword*

Ialah tahap mengambil kata yang penting dari proses sebelumnya dan kemudian membuang kata yang dianggap tidak penting dengan tujuan mengurangi jumlah kata yang akan disimpan. Tahapan *filtering stopword* :

Tabel 5 Proses *Filtering Stopword*

Sebelum <i>Filtering Stopword</i>	Sesudah <i>Filtering Stopword</i>
lebih	
semangat	Semangat
lagi	
mengajar	Mengajar

Sebelum <i>Filtering Stopword</i>	Sesudah <i>Filtering Stopword</i>
kami	
mahasiswa	
saya	
harap	
pak	
aril	
dapat	
mengajar	
lebih	Lebih
baik	Baik
lagi	
dan	
dapat	
mengembangkan	Mengembangkan
mahasiswa	
yang	
baik	
jangan	Jangan
terlalu	
cepat	Cepat
menjelaskan	Menjelaskan
dalam	
pembelajaran	Pembelajaran

d) *Stemming*

ialah mengubah semua kata menjadi suatu kata dasar dengan menghilangkan kata yang berimbuhan. Tahapan *stemming* :

Tabel 6 Proses *Stemming*

Sebelum <i>Stemming</i>	Sesudah <i>Stemming</i>
semangat	Semangat
mengajar	Ajar
lebih	Lebih
baik	Baik
mengembangkan	Kembang
jangan	Jangan
cepat	Cepat
menjelaskan	Jelas
pembelajaran	Belajar

2.2.6. Algoritma TF.IDF

Dari data yang sudah melalui tahap *preprocessing* telah dalam bentuk numerik yang diubah dengan menggunakan metode TF.IDF (*Term Frequency-Inverse Document Frequency*) ini[14].

Apabila semakin besar nilai perhitungan bobot yang diperoleh, maka semakin besar tingkat kemiripan setiap kata terhadap dokumen. Jadi bobot kata yang paling besar ialah data3 yaitu “jangan terlalu cepat menjelaskan dalam pembelajaran.”

Algoritma TF.IDF ialah suatu metode yang termasuk dalam bidang *information retrieval* paling sederhana untuk menghitung bobot kemunculan kata dalam dokumen dan memperhatikan suatu kemunculan *term* pada kumpulan dokumen [19].

Metode TF.IDF (*Term Frequency-Inverse Document Frequency*) digunakan agar dapat menilai TF.IDF membutuhkan sebuah nilai idf_t yang diperoleh dengan Persamaan 1 :

$$idf_t = \log\left(\frac{n}{df_i}\right) \quad (1)$$

Persamaan 2 :

$$W_{dt} = TF_{dt} \times idf_t \quad (2)$$

Ket :

- N = Jumlah yang akan diuji dalam dokumen
 df_t = Jumlah *term* (t) yang terdapat dalam dokumen
 W = Bobot suatu *term* pada dokumen
 TF = Banyak *term* yang sering muncul pada dokumen
 d = dokumen ke-d
 t = kata kunci pada kata ke-t
 v = hasil dari TF.IDF dari setiap dokumen

Setelah semua kata yang diambil dari proses *preprocessing*, maka selanjutnya dilakukan perhitungan TF.IDF dengan menggunakan persamaan (1) dan (2) seperti :

Tabel 2.1 Perhitungan TF.IDF

<i>Term</i>	<i>Tf</i>			<i>df</i>	<i>Idf</i> $1 \div df$	<i>Tf \times Idf</i>		
	Data 1	Data 2	Data 3			Data 1	Data 2	Data 3
semangat	1	0	0	1	1	1	0	0
mengajar	1	0	0	1	1	1	0	0
lebih	1	1	0	2	0,5	0,5	0,5	0
baik	0	1	0	1	1	0	1	0
mengembangkan	0	1	0	1	1	0	1	0
jangan	0	0	1	1	1	0	1	1
cepat	0	0	1	1	1	0	0	1
menjelaskan	0	0	1	1	1	0	0	1
pembelajaran	0	0	1	1	1	0	0	1

2.2.7. Cosine Similarity

Cosine similarity merupakan suatu model vektor dari bobot *tf.idf* yang gunanya untuk menempatkan suatu nilai numerik pada dokumen agar dapat dihitung seberapa dekat nilai antar dokumen yang artinya menghitung tingkat kemiripan antar dokumen dengan menggunakan fungsi mengukur kemiripan

(*similarity measure*). Apabila hasil dari fungsi semakin besar maka, kemiripan dari dokumen tersebut semakin mirip [20].

Dalam proses pengukuran tersebut untuk mengetahui hasil dengan menggunakan rumus dibawah ini :

$$\text{Persamaan 3 : } \cos \alpha = \frac{A \cdot B}{|A||B|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

Ket :

A = Vektor A, yang akan dibandingkan kemiripan

B = Vektor B, yang akan dibandingkan kemiripan

A*B = dot product antara vektor A dan vektor B

|A| = panjang vektor A

|B| = panjang vektor B

|A||B| = cross product antara |A| dan |B|

(3)

2.2.8. Confussion Matrix

Confussion matrix adalah proses pengukuran untuk mengetahui hasil parameter dari metode dibawah ini[14]:

a) *Accuracy*

$$\text{akurasi} = \frac{\text{jumlah data yang dinilai benar}}{\text{jumlah seluruh data validasi}} \times 100\%$$

(4)

b) *Precision*

$$\text{precision} = \frac{\text{Precion Positif} + \text{Precision Netral} + \text{Precsion negatif}}{\text{jumlah seluruh data validasi}} \times 100\%$$

(5)

c) *Recall*

$$\text{Recall} = \frac{\text{RecallPositif} + \text{PrecRecallision Netral} + \text{Recallnegatif}}{\text{jumlah seluruh data validasi}} \times 100\%$$

(6)

Pada *confussion matrix* terdapat empat kombinasi yang digunakan untuk pengukuran dari berupa hasil proses klasifikasi berupa tabel berikut [19] :

Tabel 2.7 Visualisasi *Confusion Matrix*

#	PREDISKI	
	Positif	Negatif
REALITA	TF	TP
	FN	FP

Ket :

TN (True Negatif) ialah jumlah komentar negatif pada dokumen yang sudah benar dilabeli *classifier*

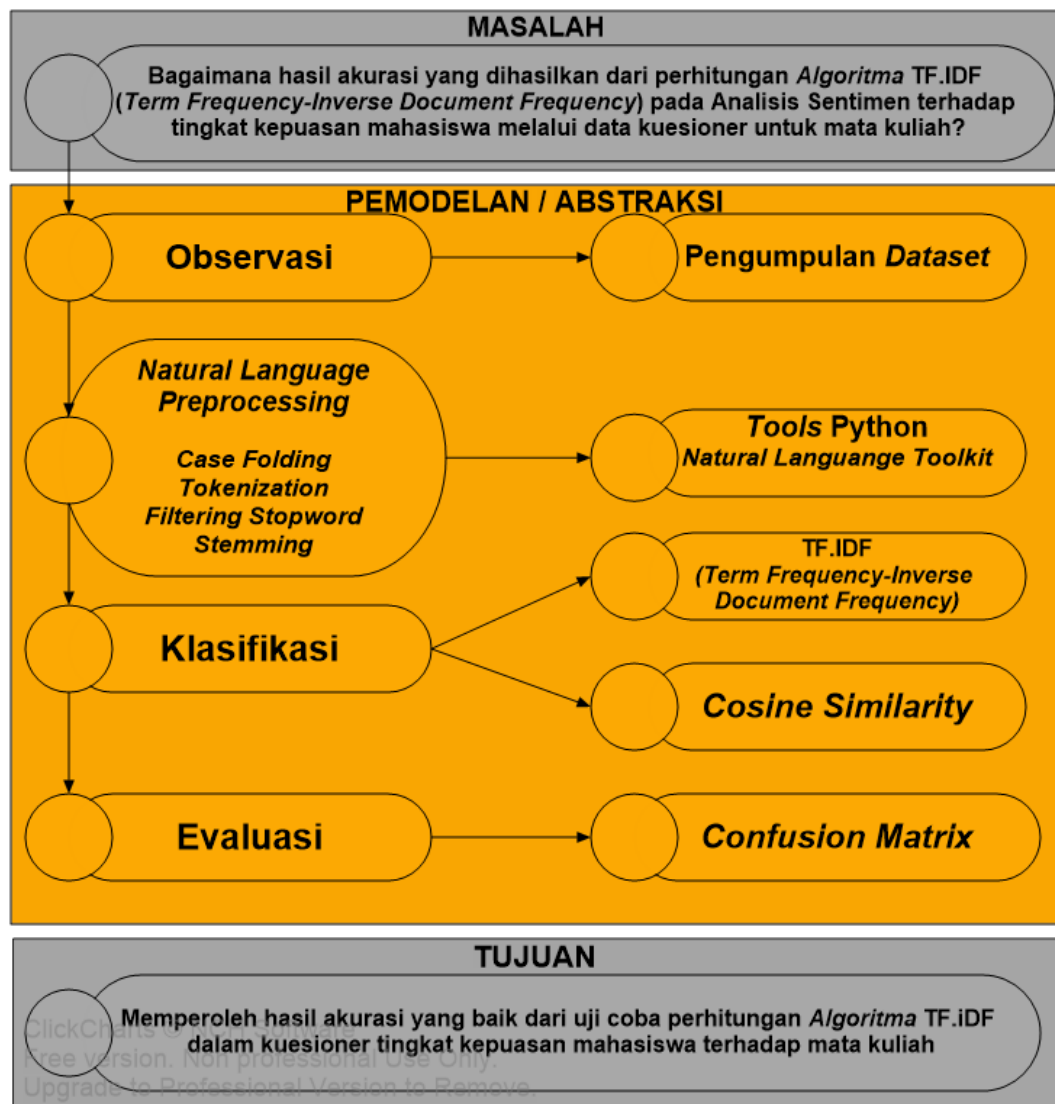
TP (True Positif) ialah jumlah komentar positif pada dokumen yang sudah benar dilabeli *classifier*

FN (False Negatif) ialah jumlah komentar negatif pada dokumen yang sudah benardilabeli *classifier*

FP (False Positif) ialah jumlah komentar positif pada dokumen yang sudah benar dilabeli *classifier*

Confussion matrix dilakukan setelah data diuji untuk evaluasi agar mengukur tingkat akurasi dengan menggunakan persamaan (4) [16].

2.3. Kerangka Pikir



Gambar 2.1 Kerangka Pikir

BAB III

METODE PENELITIAN

3.1. Jenis, Metode, Subjek, Objek, Waktu, dan Lokasi Penelitian

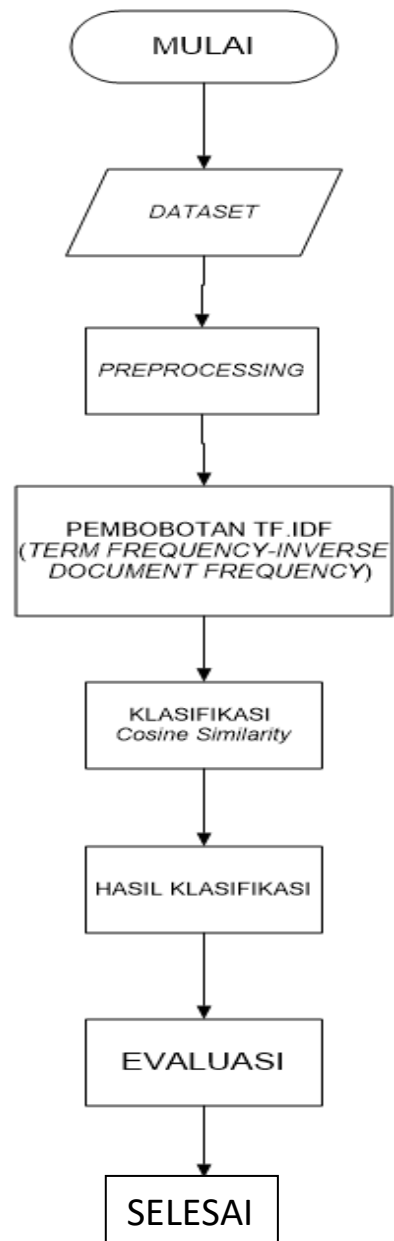
Dilihat pada penjelasan sebelumnya mengenai analisis sentimennya, pada penelitian ini ialah suatu penelitian yang berfokus pada analisis sentimen dalam menerapkan metode agar memberikan sebuah solusi dalam suatu permasalahan agar dapat mempermudah. Dilihat dari informasi yang diolah, penelitian ini ialah suatu penelitian yang bersifat kuantitatif ialah dengan menguji coba metode atau teori yang digunakan jika dilihat dari data yang akan diperoleh.

Dipandang dari penelitian ini menggunakan suatu penelitian dengan tujuan menemukan jawaban melalui percobaan penelitian kuantitatif. Subjek yang digunakan dari penelitian ini ialah analisis sentimen mahasiswa terhadap mata kuliah yang khususnya pada data kuesioner mahasiswa. Penelitian ini dimulai dari bulan Oktober sampai Februari dengan lokasi di Universitas Ichsan Gorontalo.

3.2. Dataset

Pada penelitian ini data yang akan digunakan ialah sebuah data hasil dari pengisian kotak saran mahasiswa pada Fakultas Ilmu Komputer yang diperoleh dari BAAK (Biro Administrasi Akademik dan Mahasiswa) ialah selaku yang menyediakan pengisian kuesioner kotak saran untuk penilaian mahasiswa terhadap kinerja dosen pada mata kuliah tersebut. Data yang diperoleh dari BAAK dalam bentuk file excel dengan total “1.011” data tanggapan mahasiswa di tahun 2020/2021 untuk semester ganjil. Data yang diolah ialah data yang terdapat dalam kotak saran/tanggapan, apabila terdapat tanggapan sama yang lebih dari satu, maka data yang diambil hanya satu tanggapan saja untuk dataset.

3.3. Pemodelan



Gambar 3.1 Pemodelan

3.4. *Preprocessing*

Selanjutnya hal yang dilakukan setelah mengumpulkan *dataset*, maka pada tahap selanjutnya ialah proses dimana melakukan tahap *preprocessing* data. Tahap *preprocessing* ini ialah tahap dimana mengerjakan sebelum melakukan proses mengklasifikasi. Di tahap ini melakukan suatu pengolahan data awal menjadi sebuah data yang akan diolah ke tahapan selanjutnya.

Pada tahapan teks *preprocessing* ini terdapat beberapa hal yang akan dilakukan seperti awalnya huruf kapital diubah menjadi huruf kecil (*Case Folding*), pemisahan kalimat menjadi suatu kata (*Tokenization*), kemudian memisahkan kata yang tidak penting (*Filtering Stopword*), tahapan terakhir ialah mengubah suatu kata dasar dengan menghilangkan kata yang berimbuhan (*Stemming*) dengan menggunakan *tools python*.

3.5. **Pembobotan TF.IDF**

TF.IDF (*Term Frequency-Inverse Document Frequency*) merupakan proses dimana melakukan pembobotan kata atau fitur. Maksud dari TF.IDF ialah TF (*Term Frequency*) ialah seberapa penting kata yang muncul dari tiap dokumen dalam sebuah dokumen. Sedangkan, IDF (*Inverse Document Frequency*) ialah suatu kata umum yang sering digunakan dalam dokumen. Tahap pembobotan TF.IDF ialah proses perhitungan bobot antar TF yang dikalikan dengan IDF untuk mengetahui seberapa sering muncul *term* setiap kata pada dokumen dengan proses perhitungan menggunakan rumus pada persamaan (1) sampai (2).

Untuk tahap ini ialah dimana melakukan suatu klasifikasi sentimen menggunakan metode (*Term Frequency-Inverse Document Frequency*) TF.IDF. Metode TF.IDF (*Term Frequency-Inverse Document Frequency*) ialah proses pemisahan data untuk klasifikan dengan menghilangkan nilai-nilai yang tidak penting dalam kata pada seluruh dataset [5].

Proses dari klasifikasi pembobotan TF.IDF dari setiap kata bertujuan untuk menghitung dari setiap kata. TF.IDF ialah performa yang dapat meningkatkan suatu model klasifikasi teks alasannya dari setiap kata dengan bobot nilai berbeda, maka nilai TF.IDF dengan nilai besar bisa memberikan yang sama besar pada model yang diuji pada setiap data. Dari jumlah data yang diperoleh dilakukan

suatu pelabelan secara manual antar data tanggapan ke dalam suatu kategori tanggapan positif maupun negatif. Selain itu dilakukan juga pelabelan secara otomatis dengan tahapan *preprocessing* menggunakan *tools* python yang dimana mengubah kata berimbuhan menjadi kata dasar yang termasuk pada proses *stemming*. Apabila seluruh data selesai dilabeli, selanjutnya yang dilakukan ialah proses klasifikasi akan dihitung.

3.6. *Cosine Similarity*

Untuk proses perhitungan *cosine similarity* dilakukan ialah dengan mengukur tingkat kemiripan antar dokumen dengan cara memasukkan nilai *term* yang dihasilkan dari klasifikasi pembobotan TF.IDF digunakan untuk mengukur jarak kemiripan kata antar dokumen. *Cosine similarity* digunakan untuk mendeteksi suatu kemiripan, apabila hasil yang diperoleh dari kedua buah objek besar, maka hasil fungsi *similarity* dikatakan mirip. Hasil yang diperoleh dari metode ini ialah seberapa persen kemiripan dari setiap data/dokumen yang ada.

3.7. Hasil Klasifikasi

Hasil dari klasifikasi pada penelitian ini adalah berupa seberapa akurat uji coba metode yang digunakan. Dari uji coba metode ini menghasilkan nilai yang diperoleh seperti kategori positif, negatif dan netral serta tingkat kemiripan antar dokumen untuk tanggapan dari kotak saran terhadap keputusan mahasiswa terhadap kinerja dosen pada mata kuliah.

3.8. Evaluasi

Pada tahap evaluasi ini dilakukan untuk kinerja dari metode yang digunakan dengan mengukur seberapa baik metode yang digunakan, sehingga dilakukan proses perhitungan menggunakan *confussion matrix*gunanya menguji apabila metode ini akurat dengan hasil klasifikasi yang diperoleh dari metode yang digunakan.

BAB IV

HASIL PENELITIAN

4.1. Pengumpulan Data

Data yang digunakan pada penelitian ini ialah suatu data berupa saran dari satu semester dalam tahun 2020/2021 sebanyak 1.010 data. Dari data saran yang diperoleh berbentuk file excel diseleksi secara manual terlebih dahulu untuk menghilangkan data saran yang bersifat kosong yang tidak sesuai. Selanjutnya data yang telah diseleksi secara manual disimpan dalam database yang diperoleh untuk diproses pada tahapan *preprocessing*. Setelah dilakukan seleksi secara manual maka diperoleh dari jumlah sebelumnya 1.010 data, menjadi 942 data bersih yang akan dilakukan tahapan *preprocessing* dengan membuang 68 data yang tidak memiliki arti atau bersifat kosong.

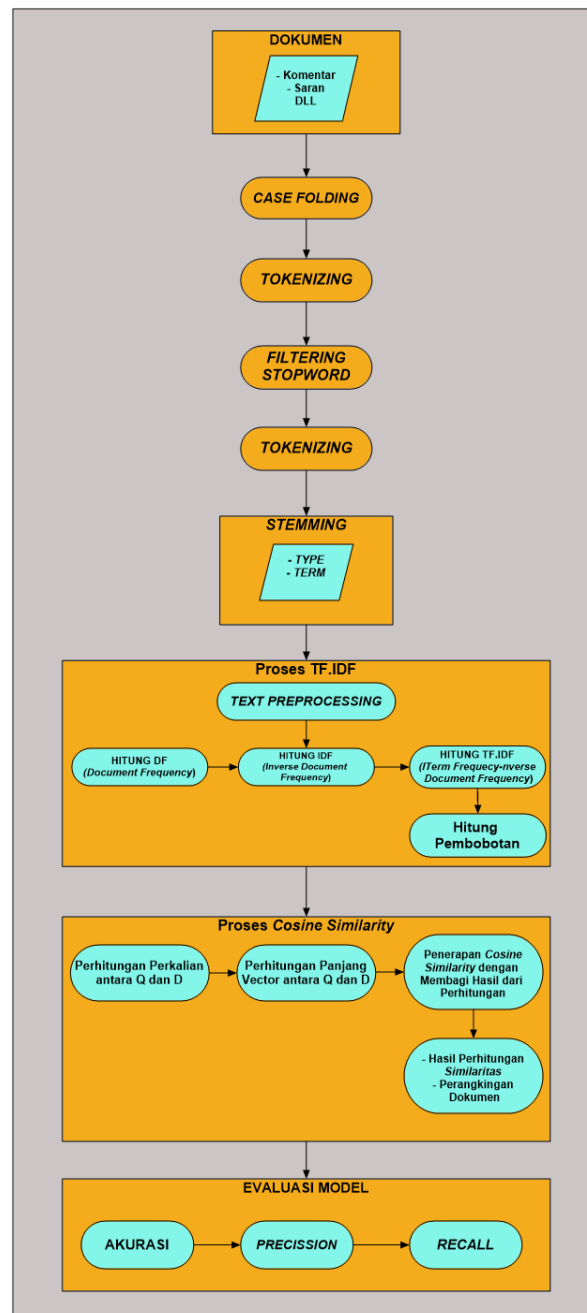
Berdasarkan dari hasil pengumpulan data, diperoleh data keseluruhan sebagai berikut :

Tabel 4.1 Hasil Pengumpulan Data

No.	Dosen	Mata Kuliah	Saran
1	IRVAN A SALIHI, M.Kom	ALGORITMA DAN STRUKTUR DATA	Kalau bisa cara mengajar lebih dipermudah..
2	IRVAN A SALIHI, M.Kom	ALGORITMA DAN STRUKTUR DATA	saran saya dosen harus lebih kocak sedikit wkwk
3	IRVAN A SALIHI, M.Kom	ALGORITMA DAN STRUKTUR DATA	Tidak ada
....			
1.008	SARLIS MOODUTO, M.Kom	KOMUNIKASI DATA	Alhamdulillah baik
1.009	SARLIS MOODUTO, M.Kom	KOMUNIKASI DATA	Lebih memberikan praktek
1.010	SARLIS MOODUTO, M.Kom	KOMUNIKASI DATA	Terima kasih atas ilmu nya pak 🙏

4.2. Model Usulan

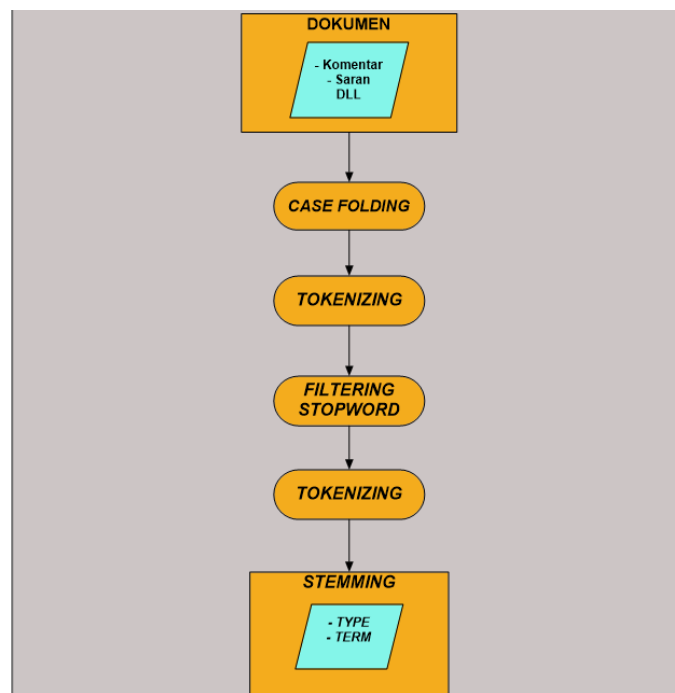
Dibawah ini ialah suatu tahapan atau model yang akan dilakukan untuk penelitian ini adapun *tools* atau alat bantu yang di gunakan adalah menggunakan bahasa pemrograman python dengan *library streamlit* untuk menampilkan hasil pemodelan dalam bentuk *framework* seperti :



Gambar 4.1 Model yang di Usulkan

4.2.1. *Pra Processing*

Tahapan *praprocessing* adalah tahapan dimana mengubah dokumen yang berupa saran menjadi suatu *term* adapun tahapannya adalah sebagai berikut :



Gambar 4.2 *Pra Precossing*

Dibawah ini ialah data saran yang telah dilabeli kelas secara manual seperti:

Tabel 4.2 Data Saran

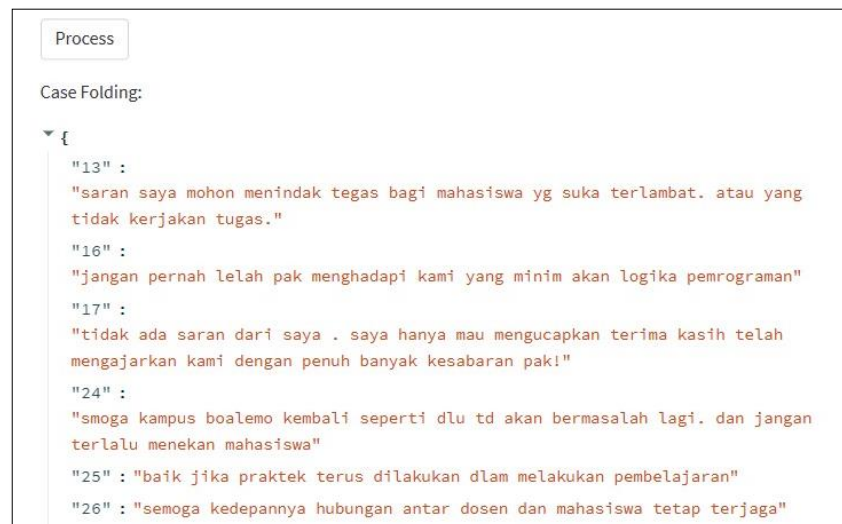
No.	Saran	Sentimen
1.	Kalau bisa cara mengajar lebih dipermudah..	Positif
2.	saran saya dosen harus lebih kocak sedikit wkwk	Netral
3.	Tidak ada	Netral
.....		
940.	jangan terlalu cepat dalam menjelaska materi	Negatif
941.	Alhamdulillah baik	Positif
942.	Lebih memberikan praktek	Negatif
943.	Terima kasih atas ilmu nya pak ☺	Positif

- a) *Case Folding* ialah pengubahan huruf besar menjadi huruf kecil dengan secara otomatis seperti :

Proses Case Folding

```
{1: 'kalau bisa cara mengajar lebih dipermudah..', 2: 'saran saya dosen
harus lebih kocak sedikit wkwk', 3: 'tidak ada', 4: 'tidak ada', 5: 'baik',
6: 'seketika di jelaskan mudah paham', 7: 'harus bisa lebih baik dari
biasanya', 8: 'terimakasih atas ilmunya', 9: 'penjelasan materi harus
pelan-pelan', 10: 'tetap di pertahankan sistem mengajarnya', 11: 'tidak
ada', 12: 'saran saya semoga mahasiswa bisa lebih paham', .....939:
'jangan terlalu cepat dalam menjelaska materi', 940: 'alhamdulillah baik',
941: 'lebih memberikan praktek', 942: 'terima kasih atas ilmu nya pak '}]
```

Hasil *case folding* menggunakan python dan *libabry* sastrwai :



Gambar 4.3 Hasil *Case Folding*

- b) *Tokenizing* ialah pemisahan kalimat menjadi suatu kata dengan secara otomatis seperti :

Proses Tokenizing

```
{1: ['kalau', 'bisa', 'cara', 'mengajar', 'lebih', 'dipermudah'], 2:
['saran', 'saya', 'dosen', 'harus', 'lebih', 'kocak', 'sedikit', 'wkwk'],
3: ['tidak', 'ada'], 4: ['tidak', 'ada'], 5: ['baik'], 6: ['seketika',
'di', 'jelaskan', 'mudah', 'paham'], 7: ['harus', 'bisa', 'lebih', 'baik',
'dari', 'biasanya'], 8: ['terimakasih', 'atas', 'ilmunya'], 9:
['penjelasan', 'materi', 'harus', 'pelan', 'pelan'], 10: ['tetap', 'di',
'pertahankan', 'sistem', 'mengajarnya'], 11: ['tidak', 'ada'], 12:
['saran', 'saya', 'semoga', 'mahasiswa', 'bisa', 'lebih',
'paham'], .....939: ['jangan', 'terlalu', 'cepat', 'dalam', 'menjelaska',
```

```
'materi'], 940: ['alhamdulillah', 'baik'], 941: ['lebih', 'memberikan',
'praktek'], 942: ['terima', 'kasih', 'atas', 'ilmu', 'nya', 'pak']}]}
```

Dibawah ini merupakan tampilan visualisasi dari *tokenizing* :



Gambar 4.4 Hasil *Tokenizing*

- c) *Number Removal* ialah dimana menghapus suatu angka yang ada dalam dokumen tetapi, tidak menghapus suatu angka yang menyatu dengan kata secara otomatis seperti :

Proses *Number Removal*

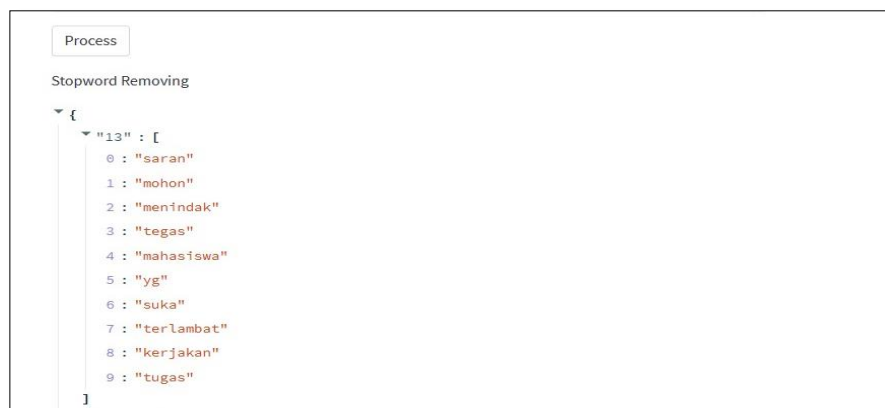
```
{1: ['kalau', 'bisa', 'cara', 'mengajar', 'lebih', 'dipermudah'], 2:
['saran', 'saya', 'dosen', 'harus', 'lebih', 'kocak', 'sedikit', 'wkwk'],
3: ['tidak', 'ada'], 4: ['tidak', 'ada'], 5: ['baik'], 6: ['seketika',
'di', 'jelaskan', 'mudah', 'paham'], 7: ['harus', 'bisa', 'lebih', 'baik',
'dari', 'biasanya'], 8: ['terimakasih', 'atas', 'ilmunya'], 9:
['penjelasan', 'materi', 'harus', 'pelan', 'pelan'], 10: ['tetap', 'di',
'pertahankan', 'sistem', 'mengajarnya'], 11: ['tidak', 'ada'], 12:
['saran', 'saya', 'semoga', 'mahasiswa', 'bisa', 'lebih', 'paham'], 13:
['saran', 'saya', 'mohon', 'menindak', 'tegas', 'bagi', 'mahasiswa', 'yg',
'suka', 'terlambat', 'atau', 'yang', 'tidak', 'kerjakan', 'tugas'], 14:
['tingkatkan', 'lagi', 'kualitas', 'dosen'], 15: ['tidak', 'ada', 'lagi',
'saran', 'karena', 'cara', 'mengajar', 'sudah', 'bagus'],
.....939: ['jangan', 'terlalu', 'cepat',
'dalam', 'menjelaska', 'materi'], 940: ['alhamdulillah', 'baik'], 941:
['lebih', 'memberikan', 'praktek'], 942: ['terima', 'kasih', 'atas',
'ilmu', 'nya', 'pak']}]}
```

- d) *Stopword Removal* ialah menyimpan suatu kata yang penting dan membuang kata dasar yang dianggap tidak penting oleh sistem secara otomatis seperti :

Proses Stopword Removal

```
{1: ['kalau', 'cara', 'mengajar', 'lebih', 'dipermudah'], 2: ['saran',
'dosen', 'lebih', 'kocak', 'sedikit', 'wkwk'], 3: [], 4: [], 5: ['baik'],
6: ['seketika', 'jelaskan', 'mudah', 'paham'], 7: ['lebih', 'baik',
'biasanya'], 8: ['terimakasih', 'atas', 'ilmunya'], 9: ['penjelasan',
'materi', 'pelan', 'pelan'], 10: ['tetap', 'pertahankan', 'sistem',
'mengajarnya'], 11: [], 12: ['saran', 'semoga', 'mahasiswa', 'lebih',
'paham'],.....935: ['sangat', 'baik'], 936:
['kehadiran', 'lebih', 'tingkatkan'], 937: ['memberikan', 'prakter',
'diskusi'], 938: ['utamakan', 'waktu', 'masuk', 'waktu', 'mengajar'], 939:
['jangan', 'terlalu', 'cepat', 'menjelaska', 'materi'], 940:
['alhamdulillah', 'baik'], 941: ['lebih', 'memberikan', 'praktek'], 942:
['terima', 'kasih', 'atas', 'ilmu', 'nya', 'pak']
```

Dibawah ini merupakan visualisasi dari *stopword/number removal* :



Gambar 4.5 Hasil *Stopword/Number Removal*

- e) *Stemming* ialah dimana menghilangkan kata yang berimbuhan dan menjadikan suatu kata dasar secara otomatis seperti :

Proses Stemming

```
{1:['kalau', 'cara', 'ajar', 'lebih', 'mudah'], 2: ['saran', 'dosen',
'lebih', 'kocak', 'sedikit', 'wkwk'], 3: [], 4: [], 5: ['baik'], 6:
['ketika', 'jelas', 'mudah', 'paham'], 7: ['lebih', 'baik', 'biasa'], 8:
['terimakasih', 'atas', 'ilmu'], 9: ['jelas', 'materi', 'pelan', 'pelan'],
10: ['tetap', 'tahan', 'sistem', 'ajar'], 11: [], 12: ['saran', 'moga',
'mahasiswa', 'lebih', 'paham'], 13: ['saran', 'mohon', 'tindak', 'tegas',
'mahasiswa', 'yg', 'suka', 'lambat', 'kerja', 'tugas'], 14: ['tingkat',
'kualitas', 'dosen'], 15: ['saran', 'cara', 'ajar', 'bagus'], 16:
['jangan', 'pernah', 'lelah', 'pak', 'hadap', 'minim', 'logika',
'pemrograman'], 17: .....938: ['utama', 'waktu', 'masuk',
'waktu', 'ajar'], 939: ['jangan', 'terlalu', 'cepat', 'menjelaska',
```

```
'materi'], 940: ['alhamdulillah', 'baik'], 941: ['lebih', 'beri',
'praktek'], 942: ['terima', 'kasih', 'atas', 'ilmu', 'nya', 'pak']}]
```

Dibawah ini merupakan visualisasi dari proses *stemming* :



Gambar 4.6 Hasil *Stemming*

Dari total data asli dari saran/komentar pengumpulan sebanyak 1.010 data. Dengan membuang data yang bernilai kosong secara manual sebanyak 68 data, sehingga total data yang digunakan pada pengujian ialah sebanyak 942 data. Kemudian dari data 942 dilakukan proses *preprocessing* dengan membuang data yang mengandung kata tidak penting atau kata dasar, sehingga menghasilkan data sebanyak 874.

Pada tabel dibawah ini terdapat 10 data yang dipilih dari jumlah data 942 data untuk dilakukan proses *preprocessing* dengan menggunakan suatu *query* seperti :

Tabel 4.3 Hasil *Pra Processing*

IDE	Hasil Case Folding	Hasil Tokenizing	Hasil Stopword Removal	Hasil Stemming
D1	'kalau bisacara mengajar lebih dipermudah..'	[kalau], [bisa], [cara], [mengajar], [lebih], [dipermudah]	['kalau', 'cara', 'mengajar', 'lebih', 'dipermudah']	['kalau', ' cara', 'ajar', 'lebih', 'mudah']
D2	'saran saya dosen harus	['saran', 'saya',	['saran', 'dosen',	['saran', ' dosen',

	lebih kocak sedikit wkwk'	'dosen', 'harus', 'lebih', 'kocak', 'sedikit', 'wkwk']	'lebih', 'kocak', 'sedikit', 'wkwk']	'lebih', 'kocak', 'sedikit', 'wkwk']
D3	'tidak ada'	['tidak', 'ada']	[]	[]
D4	'tidak ada'	['tidak', 'ada']	[]	[]
D5	'baik'	['baik']	['baik']	['baik']
D6	'seketika di jelaskan mudah paham'	['seketika', 'di', 'jelaskan', , 'mudah', 'paham' ']	['seketika', 'jelaskan', 'mudah', 'paham']	['ketika', 'jelas', 'mudah', 'paham']
D7	'harus bisa lebih baik dari biasanya'	['harus', 'bisa', 'lebih', 'baik', 'dari', 'biasanya']	['lebih', 'baik', 'biasanya']	['lebih', 'baik', 'biasa']
D8	'terimakasih atas ilmunya'	['terimakasih', 'atas', 'ilmunya']	['terimakasih', , 'atas', 'ilmunya']	['terimaka sih', 'atas', 'il mu']
D9	'penjelasan materi harus pelan-pelan'	['penjelasan', 'materi', 'harus', 'pelan', , 'pelan']	['penjelasan', , 'materi', 'pelan', 'pelan']	['jelas', ' materi', 'pelan', 'pelan']
D10	'tetap di pertahankan sistem mengajarnya'	['tetap', 'di', 'pertahankan', 'sistem', 'mengajarnya']	['tetap', 'pertahankan', , 'sistem', 'mengajarnya' ']	['tetap', ' tahan', 'sistem', 'ajar']
..	..	...	...	...
Q	'di pertahankan'	'di', 'pertahankan'	'pertahankan'	'tahan'

Dibawah ini ialah suatu tabel hasil *term* yang dihasilkan dari tahapan *preprocessing* seperti :

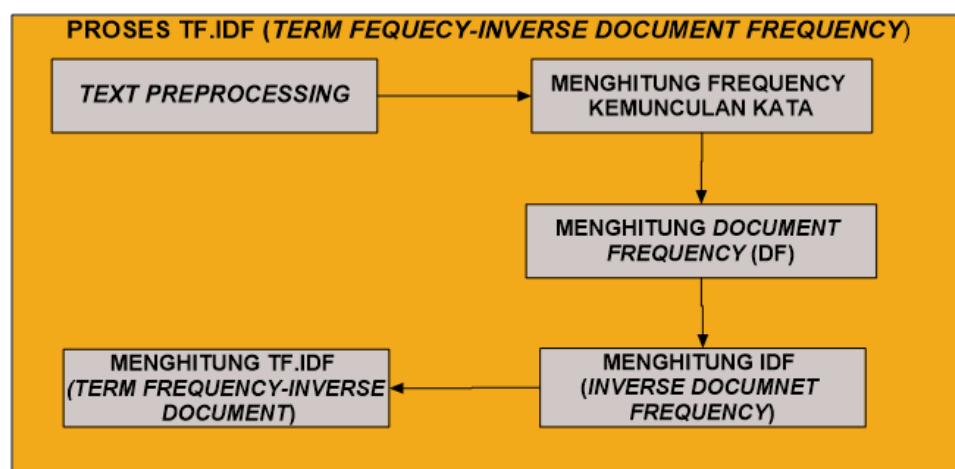
Tabel 4.4 Hasil Pemilihan *Term*

Type	Term	Term yang Diperoleh
kalau	kalau	kalau
bisa	bisa	
cara	cara	cara
ajar	ajar	ajar
lebih	lebih	lebih
mudah	mudah	mudah
saran	saran	saran
saya	saya	
dosen	dosen	dosen
harus	harus	
lebih		
kocak	kocak	kocak
sedikit	sedikit	sedikit
wkwk	wkwk	wkwk
tidak	tidak	
ada	ada	
tidak		
ada		
baik	baik	baik
ketika	ketika	ketika
di	di	
jelas	jelas	jelas
mudah		
paham	paham	paham
harus	harus	harus
bisa	bisa	
lebih		
baik		

dari	dari	
biasa	biasa	biasa
terimakasih	terimakasih	terimakasih
atas	atas	atas
ilmu	ilmu	ilmu
jelas		
materi	materi	materi
harus		
pelan	pelan	pelan
pelan		
tetap	tetap	tetap
di		
tahan	tahan	tahan
sistem	sistem	sistem
ajar		
tahan		
....		

4.2.2. Pembobotan Dokumen (TF-IDF)

Proses TF.IDF pada penelitian ini yang di implementasikan pada bahasa pemrograman python menggunakan library *sklearn* atau *scikit-learn*. Adapaun prosesnya adalah sebagai berikut:



Gambar 4.7 Tahapan TF.IDF

4.2.2.1 Menghitung *Frequency* Kemunculan Kata

Tahapan pertama yang dilakukan adalah menghitung seberapa banyak kata sering muncul atau digunakan mahasiswa pada setiap komentar/saran yang diberikan, dibawah ini adalah sebagian proses menghitung kemunculan kata pada TF (*Term Frequency*) :

Tabel 4.5 Menghitung Kemunculan Kata

Term	1	2	3	4	5	6	7	8	9	10	Q
kalau	1	0	...	...	0	0	0	0	0	0	0
cara	1	0	...	...	0	0	0	0	0	0	0
ajar	1	0	...	...	0	0	0	0	0	1	0
lebih	1	0	...	...	0	0	1	0	0	0	0
mudah	1	0	...	...	0	1	0	0	0	0	0
saran	0	1	...	...	0	0	0	0	0	0	0
dosen	0	1	...	...	0	0	0	0	0	0	0
kocak	0	1	...	...	0	0	0	0	0	0	0
sedikit	0	1	...	...	0	0	0	0	0	0	0
wkwk	0	1	...	...	0	0	0	0	0	0	0
baik	0	0	...	...	1	0	1	0	0	0	0
ketika	0	0	...	...	0	1	0	0	0	0	0
jelas	0	0	...	...	0	1	0	0	1	0	0
paham	0	0	...	...	0	1	0	0	0	0	0
biasa	0	0	...	...	0	1	0	0	0	0	0
terimakasih	0	0	...	...	0	0	0	1	0	0	0
atas	0	0	...	...	0	0	0	1	0	0	0
ilmu	0	0	...	...	0	0	0	1	0	0	0
materi	0	0	...	...	0	0	0	0	1	0	0
pelan	0	0	...	...	0	0	0	0	2	0	0
tetap	0	0	...	...	0	0	0	0	0	1	0
tahan	0	0	...	...	0	0	0	0	0	1	1
sistem	0	0	...	...	0	0	0	0	0	1	0
tahan	0	0	...	...	0	0	0	0	0	0	1

4.2.2.2 Pembobotan TF.IDF

Setelah menghitung kemunculan kata dari setiap dokumen, selanjutnya dilakukan proses perhitungan TF.IDF (*Term Frequency-Inverse Document Frequency*). Proses perhitungan TF.IDF ini dapat dilihat pada tabel berikut :

Tabel 4.6 Perhitungan DF

Term	Doc (d)	DF
kalau	D1	1
cara	D1	1
ajar	D1,D10	2
lebih	D1,D7	2
mudah	D1,D6	2
saran	D2	1
dosen	D2	1
kocak	D2	1
sedikit	D2	1
wkwk	D2	1
baik	D5,D7	2
ketika	D6	1
jelas	D6	2
paham	D6	1
biasa	D6	1
terimakasih	D8	1
atas	D8	1
ilmu	D8	1
materi	D9	1
pelan	D9	2
tetap	D10	1
tahan	D10,DQ	2
sistem	D10	1
...	...	...
tahan	DQ	1

Dari tabel diatas adalah sebagian dari kata yang diambil dari proses perhitungan DF (*Document Frequency*) yang dimana menghitung kata yang muncul dari setiap dokumen. Setelah dilakukan perhitungan DF, maka dilakukan proses perhitungan IDF secara manual dengan menggunakan rumus dibawah ini :

Term [kalau] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$\begin{aligned} idf_t &= \log \left(\frac{11}{1} \right) \\ &= 1,04139 \end{aligned}$$

Term [cara] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$\begin{aligned} idf_t &= \log \left(\frac{11}{1} \right) \\ &= 1,04139 \end{aligned}$$

Term [ajar] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$\begin{aligned} idf_t &= \log \left(\frac{11}{2} \right) \\ &= 0,74036 \end{aligned}$$

Term [lebih] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$\begin{aligned} idf_t &= \log \left(\frac{11}{2} \right) \\ &= 0,74036 \end{aligned}$$

Term [mudah] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{2} \right)$$

$$= 0,74036$$

Term [saran] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [dosen] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [kocak] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [sedikit] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [wkwk] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [baik] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{2} \right)$$

$$= 0,74036$$

Term [ketika] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [jelas] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{2} \right)$$

$$= 0,74036$$

Term [paham] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [biasa] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [terimakasih] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [atas] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [ilmu] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [materi] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [pelan] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{2} \right)$$

$$= 0,74036$$

Term [tetap] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Term [tahan] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{2} \right)$$

$$= 0,74036$$

Term [sistem] :

$$idf_t = \log \left(\frac{n}{df_t} \right)$$

$$idf_t = \log \left(\frac{11}{1} \right)$$

$$= 1,04139$$

Selanjutnya hasil dari perhitungan IDF (*Inverse Document Frequency*) dapat dilihat pada tabel dibawah ini :

Tabel 4.7 Perhitungan IDF

Term	DF	IDF
kalau	1	1,04139
Cara	1	1,04139
Ajar	2	0,74036
Lebih	2	0,74036
Mudah	2	0,74036
Saran	1	1,04139
Dosen	1	1,04139
Kocak	1	1,04139
Sedikit	1	1,04139
wkwk	1	1,04139
baik	2	0,74036
ketika	1	1,04139
jelas	2	0,74036
paham	1	1,04139
biasa	1	1,04139
terimakasih	1	1,04139
atas	1	104139

ilmu	1	1,04139
materi	1	1,04139
pelan	2	0,74036
tetap	1	1,04139
tahan	2	0,74036
sistem	1	1,04139
Q	1	1,04139

Setelah dilakukan proses perhitungan IDF (*Inverse Document Frequency*), maka dilakukan proses perhitungan selanjutnya adalah menghitung nilai TF.IDF dengan cara nilai TF $\times$ IDF sehingga menghasilkan nilai TF.IDF. Dibawah ini ialah sebagian proses perhitungan manual dengan seperti :

$$\text{Term [kalau]} : W_{dt} = TF_{dt} \times idf_t$$

$$W = 1 \times 1,041 = 1,041$$

$$\text{Term [cara]} : W_{dt} = TF_{dt} \times idf_t$$

$$W = 1 \times 1,041 = 1,041$$

$$\text{Term [ajar]} : W_{dt} = TF_{dt} \times idf_t$$

$$W = 2 \times 0,74 = 1,48$$

$$\text{Term [lebih]} : W_{dt} = TF_{dt} \times idf_t$$

$$W = 2 \times 0,74 = 1,48$$

$$\text{Term [mudah]} : W_{dt} = TF_{dt} \times idf_t$$

$$W = 2 \times 0,74 = 1,48$$

$$\text{Term [saran]} : W_{dt} = TF_{dt} \times idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [dosen]} : W_{dt} = TF_{dt} \times idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [kocak]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [sedikit]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [wkwk]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [baik]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 0,74 = 0$$

$$\text{Term [ketika]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [jelas]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 2 \times 0,74 = 1,48$$

$$\text{Term [paham]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [biasa]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [terimakasih]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [atas]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [ilmu]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [materi]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [pelan]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 2 \times 0,74 = 1,48$$

$$\text{Term [tetap]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

$$\text{Term [tahan]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 2 \times 0,74 = 1,48$$

$$\text{Term [sistem]} : W_{dt} = TF_{dt} X idf_t$$

$$W = 0 \times 1,041 = 0$$

Pada tabel dibawah ini dapat dilihat keseluruhan perhitungan dari setiap dokumen seperti :

Tabel 4.8 Perhitungan TF.IDF

D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Q
1,041	0	...	...	0	0	0	0	0	0	0
1,041	0	...	...	0	0	0	0	0	0	0
1,48	0	...	...	0	0	0	0	0	1,48	0
1,48	0	...	...	0	0	1,48	0	0	0	0
1,48	0	...	...	0	1,48	0	0	0	0	0
0	1,041	...	...	0	0	0	0	0	0	0
0	1,041	...	...	0	0	0	0	0	0	0
0	1,041	...	...	0	0	0	0	0	0	0

0	1,041	...	...	0	0	0	0	0	0	0
0	1,041	...	...	0	0	0	0	0	0	0
0	0	...	...	1,48	0	1,48	0	0	0	0
0	0	...	...	0	1,041	0	0	0	0	0
0	0	...	...	0	1,48	0	0	1,48	0	0
0	0	...	...	0	1,041	0	0	0	0	0
0	0	...	...	0	1,041	0	0	0	0	0
0	0	...	...	0	0	0	1,041	0	0	0
0	0	...	...	0	0	0	1,041	0	0	0
0	0	...	...	0	0	0	1,041	0	0	0
0	0	...	...	0	0	0	0	1,041	0	0
0	0	...	...	0	0	0	0	1,481	0	0
0	0	...	...	0	0	0	0	0	1,041	0
0	0	...	...	0	0	0	0	0	1,48	1,48
0	0	...	...	0	0	0	0	0	1,041	0
0	0	...	...	0	0	0	0	0	0	1,041

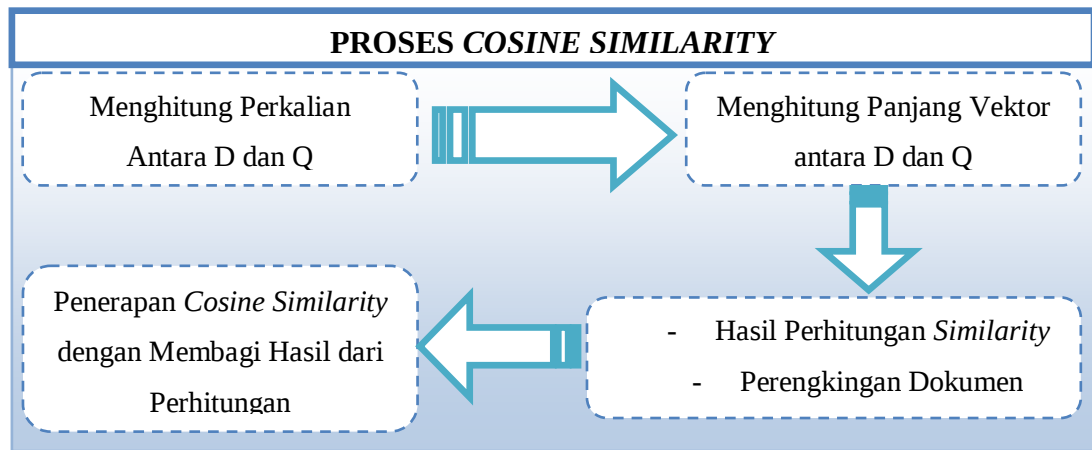
Berikut ini merupakan hasil TF-IDF dengan bahasa pemrograman python menggunakan *library sklearn* sebagai visualisasinya :

Process												
	490	13	587	16	17	24	25	26	38	39	55	63
490	1	0.0762	0.035	0	0	0.0314	0	0.0397	0	0	0.0381	0
13	0.0762	1	0.0316	0	0.068	0.0284	0	0.0359	0	0	0.0345	0
587	0.035	0.0316	1	0	0.0313	0.134	0.0324	0.0358	0	0.2833	0.0716	0.1265
16	0	0	0	1	0.0679	0.0759	0	0	0.2851	0	0.0807	0
17	0	0.068	0.0313	0.0679	1	0	0.0297	0	0	0	0.108	0.0299
24	0.0314	0.0284	0.134	0.0759	0	1	0	0.0321	0.1181	0	0.0309	0
25	0	0	0.0324	0	0.0297	0	1	0	0.097	0	0.1712	0.0905
26	0.0397	0.0359	0.0358	0	0	0.0321	0	1	0	0	0.039	0.1065
38	0	0	0	0.2851	0	0.1181	0.097	0	1	0	0.2226	0.0975
39	0	0	0.2833	0	0	0	0	0	0	1	0.0567	0.0497

Gambar 4.8 Visualisasi TF.IDF

4.2.3. Perhitungan *Cosine Similarity*

Setelah nilai TF.IDF diketahui maka, tahapan selanjutnya ialah proses perhitungan *cosine similarity* yang artinya menghitung kesamaan dalam membandingkan kemiripan antar data/dokumen.



Gambar 4.9 Tahapan *Cosine Similarity*

Dengan menggunakan rumus seperti dibawah ini :

$$Cos \propto \frac{A \cdot B}{|A||B|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

Ket :

A = Vektor A, yang akan dibandingkan kemiripan

B = Vektor B, yang akan dibandingkan kemiripan

A*B = dot product antara vektor A dan vektor B

|A| = panjang vektor A

|B|= panjang vektor B

|A||B| = cross product antara |A| dan |B|

Dibawah ini ialah suatu tahapan perhitungan secara manual untuk algoritma *cosine similarity* dengan menggunakan rumus diatas seperti :

- Mengalikan nilai Q yang ditemukan dari hasil TF.IDF sebelumnya dengan setiap dokumen yang ada.

- Selanjutnya dengan menghitung hasil perkalian masing-masing *query*, kemudian hasil jawaban dari perkalian untuk setiap *query* dijumlahkan.
- Kemudian menghitung hasil dari perkalian vektor tiap *query* dengan jawaban *query*
- Terakhir menghitung nilai *cosine similarity* antara nilai vektor D terhadap Q dengan menggunakan rumus diatas.

Tabel 4.9 Perhitungan Q dan D

<i>Term</i>	1	2	3	4	5	6	7	8	9	10
kalau	0	0	...	...	0	0	0	0	0	0
cara	0	0	...	...	0	0	0	0	0	0
ajar	0	0	...	...	0	0	0	0	0	0
lebih	0	0	...	...	0	0	0	0	0	0
mudah	0	0	...	...	0	0	0	0	0	0
saran	0	0	...	...	0	0	0	0	0	0
dosen	0	0	...	...	0	0	0	0	0	0
kocak	0	0	...	...	0	0	0	0	0	0
sedikit	0	0	...	...	0	0	0	0	0	0
wkwk	0	0	...	...	0	0	0	0	0	0
baik	0	0	...	...	0	0	0	0	0	0
ketika	0	0	...	...	0	0	0	0	0	0
jelas	0	0	...	...	0	0	0	0	0	0
paham	0	0	...	...	0	0	0	0	0	0
biasa	0	0	...	...	0	0	0	0	0	0
terimakasih	0	0	...	...	0	0	0	0	0	0
atas	0	0	...	...	0	0	0	0	0	0
ilmu	0	0	...	...	0	0	0	0	0	0
materi	0	0	...	...	0	0	0	0	0	0

pelan	0	0	...	...	0	0	0	0	0	0
tetap	0	0	...	...	0	0	0	0	0	0
tahan	0	0	...	...	0	0	0	0	0	2,190
sistem	0	0	...	..	0	0	0	0	0	0
Total D0.Di	0	0	...	...	0	0	0	0	0	2,190

Dibawah adalah suatu proses perhitungan yang dimana untuk mengetahui hasil dari perhitungan akar kuadrat seperti pada tabel berikut :

Tabel 4.10 Perhitungan Akar Kuadrat *Cosine Similarity*

D1	D2	D3	D4	D5	D6	D7	D8	D9	D10		Q
1,084	0	...	...	0	0	0	0	0	0	...	0
1,084	0	...	...	0	0	0	0	0	0	...	0
2,190	0	...	...	0	0	0	0	0	2,190	...	0
2,190	0	...	...	0	0	2,190	0	0	0	...	0
2,190	0	...	...	0	2,190	0	0	0	0	...	0
0	1,085	...	...	0	0	0	0	0	0	...	0
0	1,085	...	...	0	0	0	0	0	0	...	0
0	1,085	...	...	0	0	0	0	0	0	...	0
0	1,085	...	...	0	0	0	0	0	0	...	0
0	1,085	...	...	0	0	0	0	0	0	...	0
0	0	...	...	2,190	0	2,190	0	0	0	...	0
0	0	...	...	0	1,085	0	0	0	0	...	0
0	0	...	...	0	2,190	0	0	0	0	...	0
0	0	...	...	0	1,085	0	0	2,190	0	...	0
0	0	...	...	0	1,085	0	0	0	0	...	0
0	0	...	...	0	0	0	1,084	0	0	...	0
0	0	...	...	0	0	0	1,084	0	0	...	0
0	0	...	...	0	0	0	1,084	0	0	...	0
0	0	...	...	0	0	0	0	1,085	0	...	0
0	0	...	...	0	0	0	0	2,190	0	...	0

	0	0	...	...	0	0	0	0	0	1,084	...	0
	0	0	...	...	0	0	0	0	0	2,190	...	2,190
	0	0	...	...	0	0	0	0	0	1,084	...	0
Total Kuadrat	8,739	5,425	...	...	2,190	6,550	4,380	3,252	5,465	6,548	...	2,190
Akar Kuadrat	2,956	2,329	...	...	1,479	2,559	2,092	1,083	2,337	2,558	...	1,479

Dibawah ini ialah suatu proses perhitungan manual untuk *cosine similarity* antara Q dengan setiap dokumen untuk menentukan tingkat kemiripan seperti :

a) *Similarity* Q dengan D1 :

$$\text{Cosine similarity} = D1 / (\text{Akar Kuadrat Q} \times \text{Akar Kuadrat D1})$$

$$\text{Cosine similarity} = 0 / (1,479 \times 2,956) = 0$$

$$\text{Presentase} = 0\%$$

b) *Similarity* Q dengan D2 :

$$\text{Cosine similarity} = D2 / (\text{Akar Kuadrat Q} \times \text{Akar Kuadrat D2})$$

$$\text{Cosine similarity} = 0 / (1,479 \times 2,329) = 0$$

$$\text{Presentase} = 0\%$$

c) *Cosine similarity* = D3/(Akar Kuadrat Q × Akar Kuadrat D3)

$$\text{Cosine similarity} = -$$

$$\text{Presentase} = -$$

d) *Cosine similarity* = D4/(Akar Kuadrat Q × Akar Kuadrat D4)

$$\text{Cosine similarity} = -$$

$$\text{Presentase} = -$$

e) *Cosine similarity* = D5/(Akar Kuadrat Q × Akar Kuadrat D5)

$$\text{Cosine similarity} = 0 / (1,479 \times 1,479) = 0$$

$$\text{Presentase} = 0\%$$

f) *Cosine similarity* = D6/(Akar Kuadrat Q × Akar Kuadrat D6)

$$\text{Cosine similarity} = 0 / (1,479 \times 2,559) = 0$$

$$\text{Presentase} = 0\%$$

g) *Cosine similarity* = D7/(Akar Kuadrat Q × Akar Kuadrat D7)

$$\text{Cosine similarity} = 0 / (1,479 \times 2,093) = 0$$

$$\text{Presentase} = 0\%$$

$$\text{h) } \text{Cosine similarity} = D8/(\text{Akar Kuadrat } Q \times \text{Akar Kuadrat } D8)$$

$$\text{Cosine similarity} = 0/(1,479 \times 1,083) = 0$$

$$\text{Presentase} = 0\%$$

$$\text{i) } \text{Cosine similarity} = D9/(\text{Akar Kuadrat } Q \times \text{Akar Kuadrat } D9)$$

$$\text{Cosine similarity} = 0/(1,479 \times 2,337) = 0$$

$$\text{Presentase} = 0\%$$

$$\text{j) } \text{Cosine similarity} = D10/(\text{Akar Kuadrat } Q \times \text{Akar Kuadrat } D10)$$

$$\text{Cosine similarity} = 0/(1,479 \times 2,559) = 0$$

$$\text{Presentase} = 0\%$$

Setelah dilakukan proses perhitungan keseluruhan *cosine similarity* maka dapat dilihat pada tabel dibawah ini tingkat kemiripan dari dokumen yang ada antara lain :

Tabel 4.11 Pengurutan Tingkat Kemiripan

IDE	Komentar	Nilai Kemiripan
10	Tetap di pertahankan sistem mengajarnya	0%
1	Kalau bisa cara mengajarnya lebih dipermudah	0%
2	Sara saya dosen harus lebih kocak sedikit wkwk	0%
5	Baik	0%
6	Seketika dijelaskan mudah paham	0%
7	Harus bisa lebih dari biasanya	0%
8	Terimakasih atas ilmunya	0%
9	Penjelasan materi harus pelan-pelan	0%
3	Tidak ada	-
4	Tidak ada	-
..	..	...

Jadi, dapat dilihat dari proses perhitungan *cosine similarity* dari 10 komentar yang digunakan apabila dengan menggunakan *query* “Di Pertahanakan”, maka hasilnya tidak memiliki suatu kemiripan dengan *query* yang dimasukkan.

BAB V

HASIL EVALUASI

5.1. Evaluasi Model

Setelah Model Selesai dibangun maka langkah selanjutnya adalah melakukan evaluasi terhadap model tersebut, evaluasi dilakukan dengan menggunakan *Confusion matrix* sehingga didapatkan *akurasi precision* dan *recall* terhadap data validasi.

Adapun data validasi dalam penelitian ini terdiri dari 14 data yang di ambil dari dataset sebagai berikut :

Tabel 5.1 Data Validasi

No	Komentar	Sentimen
1	Lebih di perbaiki untuk penguasaan materinya	Negatif
2	Lebih memberikan tugas dan praktek	Negatif
3	Cara mengajar sudah baik	Netral
4	Ditingkatkan proses belajar	Negatif
5	Lebih ditingkatkan diskusi nya	Negatif
6	Kehadiran lebih di tingkatkan lagi	Negatif
7	Agar bisa memberikan prakter atau diskusi	Negatif
8	Utamakan waktu masuk dan waktu mengajar	Negatif
9	Tidak ada saran dari saya. Saya hanya mau mengucapkan TERIMA KASIH TELAH MENGAJARKAN KAMI DENGAN PENUH BANYAK KESABARAN PAK!	Positif
10	Sudah baik	Netral
11	Jangan pernah lelah pak menghadapi kami yang minim akan logika pemrograman	Positif
12	Semoga kedepannya hubungan antar dosen dan mahasiswa tetap terjaga	Positif
13	Semoga cara mengajarnya akan lebih baik dari Hari hari kemarin.	Positif
14	Jangan pernah bosan memberikan yang terbaik kepada kami sebagai mahasiswa	Positif

Selanjutnya data validasi tersebut di uji ke dalam model yang sudah di bangun dan membandingkan sentimen hasil uji dengan sentimen aktual pada data valiasi. Hasil pengujian dapat di lihat di tabel 5.2 sebagai berikut:

Tabel 5.2 Hasil Pengujian Model

No	Komentar	Sentimen	Hasil Uji Coba
1	Lebih di perbaiki untuk penguasaan materinya	Negatif	Negatif
2	Lebih memberikan tugas dan praktek	Negatif	Negatif
3	Cara mengajar sudah baik	Netral	Positif
4	Ditingkatkan proses belajar	Negatif	Negatif
5	Lebih ditingkatkan diskusi nya	Negatif	Negatif
6	Kehadiran lebih di tingkatkan lagi	Negatif	Negatif
7	Agar bisa memberikan prakter atau diskusi	Negatif	Negatif
8	Utamakan waktu masuk dan waktu mengajar	Negatif	Negatif
9	Tidak ada saran dari saya. Saya hanya mau mengucapkan TERIMA KASIH TELAH MENGAJARKAN KAMI DENGAN PENUH BANYAK KESABARAN PAK!	Positif	Positif
10	Sudah baik	Netral	Positif
11	Jangan pernah lelah pak menghadapi kami yang minim akan logika pemrograman	Positif	Positif
12	Semoga kedepannya hubungan antar dosen dan mahasiswa tetap terjaga	Positif	Positif
13	Semoga cara mengajarnya akan lebih baik dari Hari hari kemarin.	Positif	Positif
14	Jangan pernah bosan memberikan yang terbaik kepada kami sebagai mahasiswa	Positif	Positif

5.1.1. Confusion Matriks

Hasil Pengujian tersebut selanjutnya di masukkan ke dalam *confusion matriks* sebagai berikut :

Tabel 5.3 Confusion Matrix

		Sentimen aktual		
		Positif	Netral	Negatif
Sentimen uji coba	Positif	a=5	b=0	c=0
	Netral	d=2	e=0	f=0
	Negatif	g=0	h=0	i=7

Dari tabel *confusion matriks* akan didapatkan nilai Akurasi, *Precision* dan *Recall*

1. Akurasi

$$akurasi = \frac{jumlah\ data\ yang\ dibernilai\ benar}{jumlah\ seluruh\ data\ validasi} \times 100$$

$$\begin{aligned}
 akurasi &= \frac{a + e + i}{a + b + c + d + e + f + g + h + i} \times 100 \\
 &= \frac{5 + 0 + 7}{5 + 0 + 0 + 2 + 0 + 0 + 0 + 0 + 7} \times 100 \\
 &= \frac{12}{14} \times 100 \\
 &= 0,857 * 100 \\
 &= 85,7\%
 \end{aligned}$$

2. Precision

Tabel 5.4 Perhitungan Precision

	Positif	netral	Negatif
TAPI	5	0	7
FP	2+0	0+0	0+0
Precision	5/(2+0)=2,5	0/(0+0)=0	7/(0+0)=1

$$precision = \frac{Precision\ Positif + Precision\ Netral + Precision\ negatif}{jumlah\ seluruh\ data\ validasi} \times 100$$

$$\begin{aligned}
 precision &= \frac{2,5 + 0 + 1}{14} \times 100 \\
 &= \frac{3,5}{14} \times 100 \\
 &= 0,25 \times 100 \\
 &= 25\%
 \end{aligned}$$

3. Recall

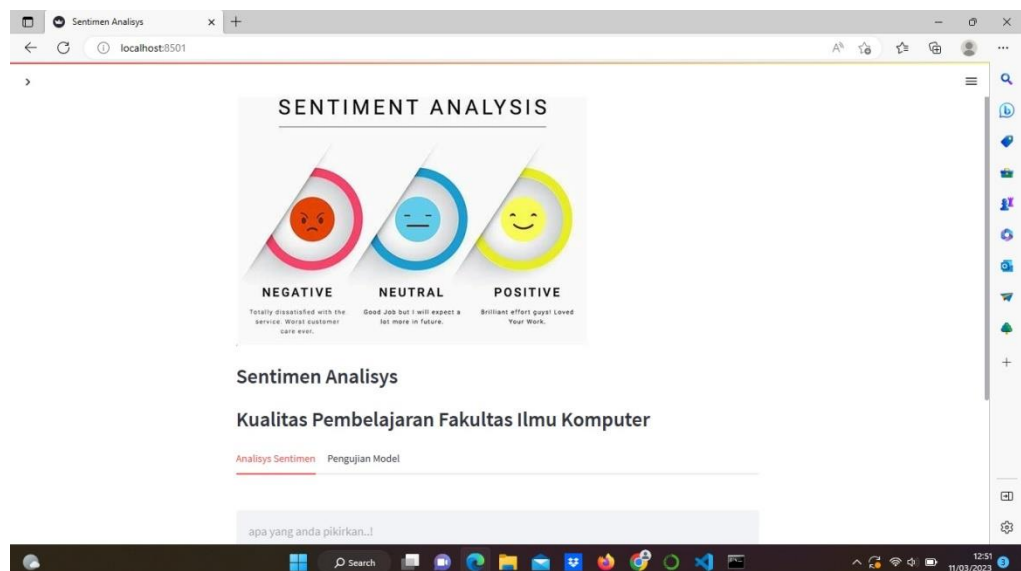
Tabel 5.5 Perhitungan Recall

	Positif	Netral	Negatif
TP	5	0	7
FP	0+0	2+0	0+0
Recall	$5/(0+0)=1$	$0/(2+0)=0$	$7/(0+0)=1$

$$Recall = \frac{RecallPositif + PrecRecallision Netral + Recallnegatif}{jumlah seluruh data validasi} \times 100$$

$$\begin{aligned}
 Recall &= \frac{1 + 0 + 1}{14} \times 100 \\
 &= \frac{2}{14} \times 100 \\
 &= 0,142 \times 100 \\
 &= 14,2\%
 \end{aligned}$$

5.2. Uji Coba Model

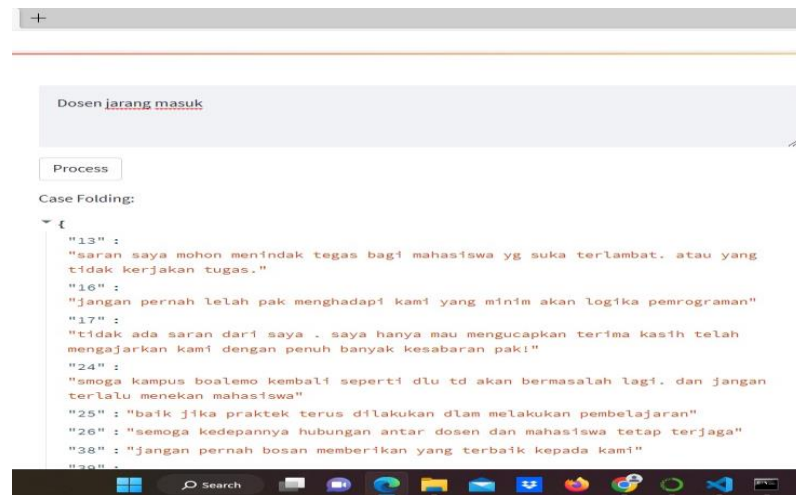


Gambar 5.1 Model Visualisasi

Tampilan diatas ialah sebuah visualisasi yang diakses melalui terminal pada pyhton untuk menampilkan antara komentar positif, negatif dan netral, proses *preprocessing*, hingga seberapa persen tingkat kemiripan komentar yang ada.

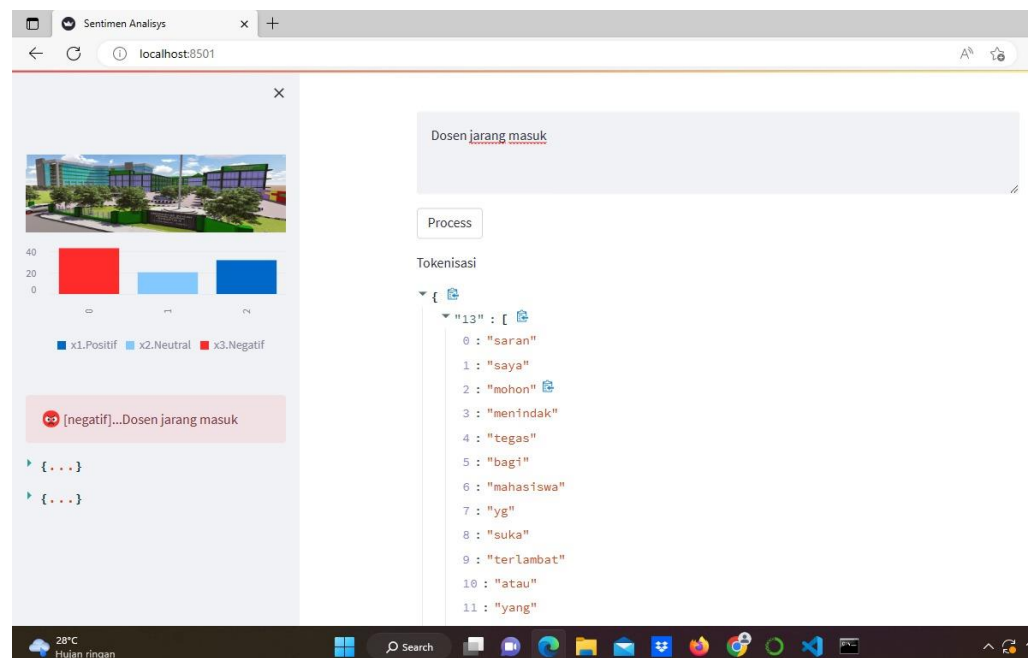
5.2.1. *Pra Processing*

- a) Dibawah ini merupakan tampilan dari visualisasi pada *preprocessing* untuk *case folding* yang dimana mengubah seluruh huruf kapital pada kalimat menjadi huruf kecil secara otomatis dengan menggunakan *library streamlit* :



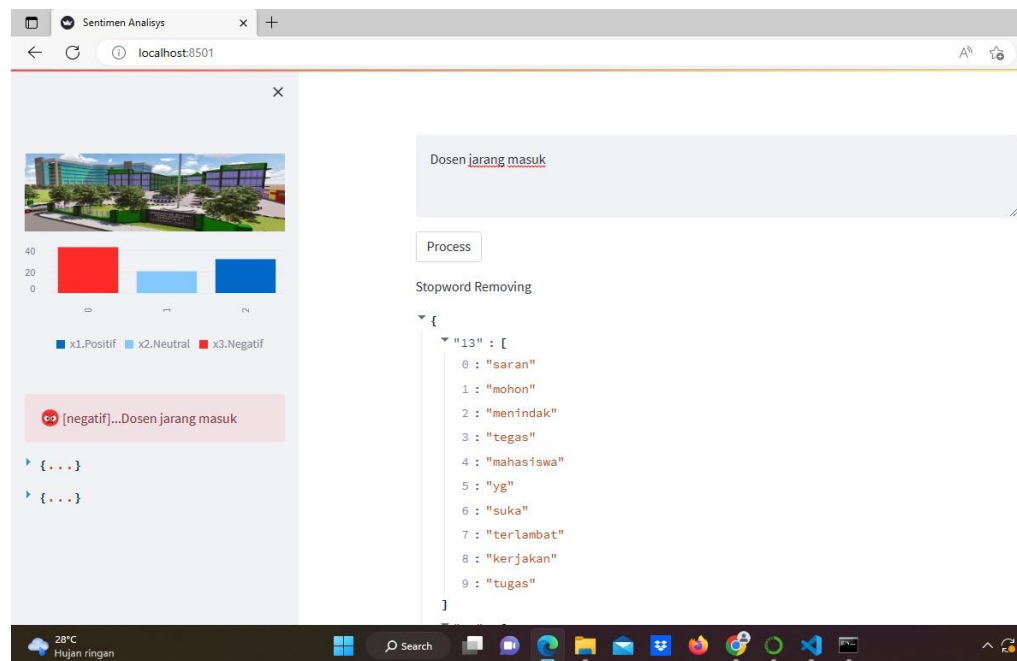
Gambar 5.2 Visualisasi *Case Folding*

- b) *Tokenizing* ialah pemisahan kalimat menjadi suatu kata setelah dilakukan proses *case folding* dengan secara otomatis untuk visualisasinya seperti :



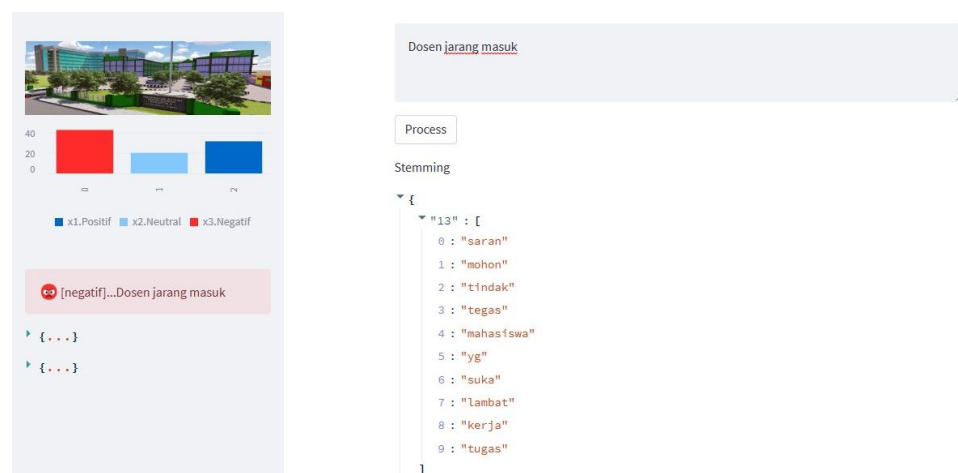
Gambar 5.3 Visualisasi *Tokenizing*

- c) *Number Removal* ialah dimana menghapus suatu angka yang ada dalam dokumen tetapi, tidak menghapus suatu angka yang menyatu dengan kata secara otomatis
- d) *Stopword Removal* ialah menyimpan suatu kata yang penting dan membuang kata dasar yang dianggap tidak penting oleh sistem secara otomatis untuk visualisasinya seperti :



Gambar 5.4 Visualisasi *Stopword/Number Removal*

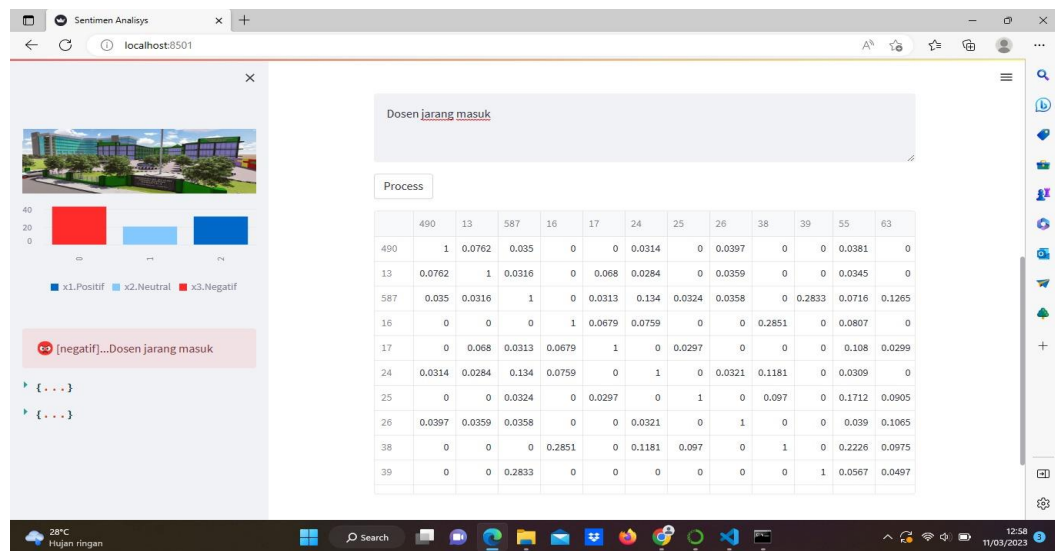
- e) *Stemming* ialah proses dimana menghilangkan kata yang berlebihan dan menjadikan suatu kata dasar secara otomatis untuk visualisasinya :



Gambar 5.5 Visualisasi *Stemming*

5.2.2. Pembobotan Dokumen (TF.IDF)

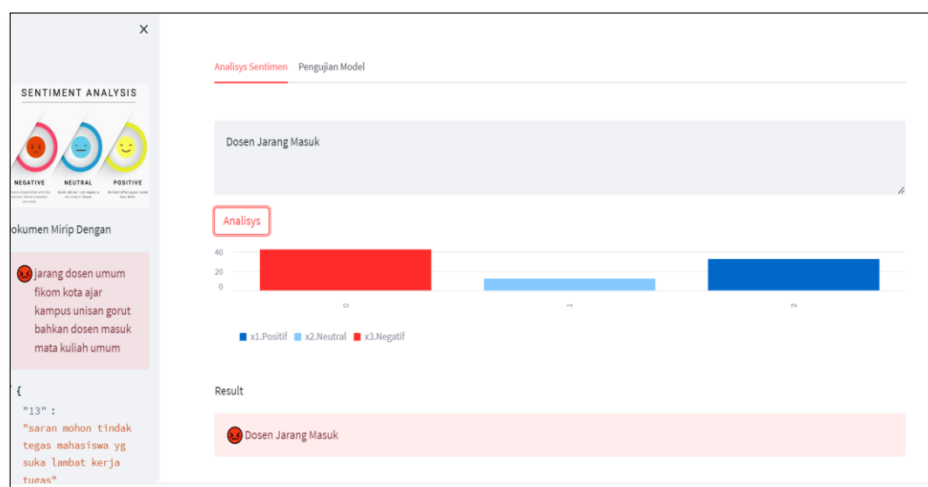
Proses perhitungan untuk pembobotan TF.IDF dilakukan setelah tahapan *preprocessing* secara otomatis menggunakan *tools* python untuk visualisasinya dalam menghitung nilai bobot keseluruhan kata yang ada pada setiap dokumen seperti :



Gambar 5.6 Visualisasi Perhitungan TF.IDF

5.2.3. Cosine Similarity

Proses *cosine similarity* didapatkan dari proses perhitungan setelah didapatkan hasil perhitungan TF.IDF, maka hasil dari TF.IDF dikalikan dengan nilai Q dengan tujuan untuk melihat tingkat kemiripan antar data/dokumen (komentar) pada suatu kelas antara positif, netal, maupun negatif seperti dibawah ini :



Gambar 5.7 Visualisasi Cosine Similarity

Dari gambar tersebut dapat dilihat jika komentar yang dimasukkan dapat dilihat komentar tersebut termasuk kedalam suatu komentar negatif. Dalam tampilan visualisasi ini terdapat tahapan-tahapan yang termasuk dalam memproses data melalui tahapan awal ialah *preprocessing* data dengan hasil akhir ialah berupa kata yang penting untuk selanjutnya dilanjutkan dalam penerapan metode dalam penelitian ini ialah metode TF.IDF dan *Cosine Similarity*.

BAB VI

PENUTUP

6.1. KESIMPULAN

Hasil dari penggunaan metode TF.IDF dan *Cosine Similarity* menunjukkan bahwa analisis sentimen terhadap data kuesioner kinerja dosen di Universitas Ichsan Gorontalo lebih dominan pada sentimen positif, yang dimana dari total keseluruhan ialah 1.010 data kemudian dilabeli secara manual, maka data yang diperoleh dari pelabelan tersebut sebanyak 942 data terdapat 423 data termasuk dalam sentimen positif, 226 data sentimen negatif dan 293 sentimen netral. Komentar yang bersifat negatif tersebut dapat diperbaiki dengan cara meningkatkan proses pembelajaran agar kinerja dosen tersebut lebih baik lagi khususnya di Fakultas Ilmu Komputer Universitas Ichsan Gorontalo. Setelah dilakukan proses klasifikasi menggunakan algoritma *Term Frequency-Inverse Document Frequency* (TF.IDF) dan *Cosine Similarity* akurasi yang dihasilkan baik. Hal ini dilihat dari hasil evaluasi dengan *confusion matrix* perhitungan secara manual tingkat akurasi sebesar 85,7%, *preccision* sebesar 25% dan *recall* sebesar 14,2%.

6.2. SARAN

Dari peneltian ini hanya menggunakan satu metode ialah metode TF.IDF, sedangkan metode TF.IDF hasil akhirnya apabila tidak digunakan metode pembanding maka hasil akhirnya terlalu sederhana dan hasil akurasi yang dihasilkan kurang baik. Jadi pada penelitian ini terdapat sebuah kekurangan ialah dari penggunaan metode, sehingga untuk menyempurnakan penggunaan metode pada penelitian ini ditambahkan suatu metode sebagai pembanding ialah metode *cosine similarity* dan untuk menghasilkan hasil akurasi yang baik. Metode *cosine similarity* digunakan pada penelitan ini fungsinya ialah melihat suatu tingkat kemiripan dari dokumen/data yang digunakan. Jadi pada penelitian ini digunakan metode TF.IDF dan *Cosine Similarity* sebagai metode untuk mengolah data komentar, kemudian dievaluasikan menggunakan *Confusion Matrix* untuk melihat kinerja dari metode yang digunakan. Pada penelitian ini metode TF.IDF dan

Confison Matrix sudah baik, dikarenakan hasil akurasi yang dihasilkan dari evaluasi metode baik. Untuk selanjutnya dari penggunaan metode dapat dikembangkan dengan menggunakan metode lain untuk menghasilkan hasil akurasi yang lebih baik lagi.

DAFTAR PUSTAKA

- [1] S. R. P. S. Jan, “Penerapan Algoritma Naive Bayes pada Analisis Sentimen terhadap Data Kuesioner Evaluasi Dosen Universitas Ichsan Gorontalo,” p. 22, 2022.
- [2] U. Kristen, D. Wacana, M. Estera, P. S. Informatika, and F. T. Informasi, “Hasil evaluasi pembelajaran dosen,” 2021.
- [3] I. Nur, F. Astuti, I. Darmawan, and D. Pramesti, “Analisis Sentimen Pada Data Kuesioner Evaluasi Dosen Oleh Mahasiswa (EDOM) Prodi Sistem Informasi Telkom University Menggunakan Algoritma Support Vector Machine Analisis Of Sentiment In The Data Evaluation Of Dosage Students (EDOM) Information Systems Tel.”
- [4] M. A. Muzakki, P. P. Sarjana, M. T. Informasi, B. A. Keuangan, A. Sentimen, and A. Rocchio, “Klasifikasi dan Analisa Sentimen Kuesioner Fasilitas dan Layanan untuk Universitas Qomaruddin Gresik,” vol. 5, pp. 68–76, 2020.
- [5] E. Dazki and R. E. Indrajit, “Analisis Sentimen Kepuasan Pelanggan Pada Penyedia Layanan Pengantaran Makanan,” vol. 9, no. 3, pp. 1699–1708, 2022.
- [6] J. A. Septian, T. M. Fahrudin, and A. Nugroho, “Journal of Intelligent Systems and Computation 43,” *J. Intell. Syst. Comput.*, pp. 43–49, 2019.
- [7] B. Harjito, K. N. Aini, and B. Murtiyasa, “Klasifikasi Dokumen berkonten Serangan jaringan menggunakan Multinomial Naive Bayes 2 . Text Mining , text processing and Term Frequency – Inverse,” pp. 112–118, 2018.
- [8] J. A. Septian, T. M. Fahrudin, and A. Nugroho, “Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF - IDF dan K - Nearest Neighbor,” pp. 43–49.
- [9] M. Yunus, “Basic Text Preprocessing Menggunakan NLTK.”
- [10] “campus.quipper.com.”
- [11] B. A. Sevsa, P. Studi, and T. Informatika, “Analisis Sentimen pada Indeks

- Kinerja Dosen Fakultas SAINTEK UIN Sunan Kalijaga Menggunakan Naive Bayes Classifier,” vol. 10, pp. 112–123, 2019.
- [12] R. Taufiq, “PENILAIAN KINERJA DOSEN DALAM BIDANG BELAJAR,” vol. 5, no. 1, pp. 77–85, 1979.
- [13] M. K. Hermawati, Fajar Astuti S.Kom, *Data Mining*. ANDI, 2013.
- [14] M. T. Informatika and U. Amikom, “ANALISIS PEMBOBOTAN KATA PADA KLASIFIKASI TEXT MINING,” vol. 3, no. 2, pp. 179–184, 2019.
- [15] D. A. E. Rozi Fahrur Imam, Pranomo Hadi Sholeh, “Implementasi Opinion Mining (Analisis Sentimen) Untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi,” *Implementasi Opin. Min. (Analisis Sentimen) Untuk Ekstraksi Data Opini Publik pada Perguru. Tinggi*, vol. Vol. 6, 2012.
- [16] J. S. Komputer, “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” vol. 5, no. November 2019, pp. 697–711, 2021.
- [17] Y. T. Pratama, F. A. Bachtiar, and N. Y. Setiawan, “PARIWISATA PANTAI MALANG SELATAN MENGGUNAKAN TF-IDF DAN SUPPORT VECTOR MACHINE SKRIPSI memperoleh gelar Sarjana Komputer Disusun oleh: Yoga Tika Pratama,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, pp. 6244–6252, 2018.
- [18] R. A. Sasmita and A. Z. Falani, “Pemanfaatan Algoritma TF/IDF Pada Sistem Informasi Ecomplaint Handling,” *J. Link*, vol. 27, no. 1, pp. 27–33, 2018.
- [19] S. Wulan, U. Vitandy, A. A. Supianto, and F. A. Bachtiar, “Analisis Sentimen Evaluasi Kinerja Dosen menggunakan Term Frequency- Inverse Document Frequency dan Naïve Bayes Classifier,” vol. 3, no. 6, 2019.
- [20] B. Herwijayanti, D. E. Ratnawati, and L. Muflikhah, “Klasifikasi Berita Online d engan menggunakan Pembobotan TF-IDF dan Cosine Similarity,” vol. 2, no. 1, pp. 306–312, 2018.

LAMPIRAN

Lampiran 1 : HASIL TURNITIN

 **Similarity Report ID: oia25211:33351233**

PAPER NAME	AUTHOR
SKRIPSI_T3119106_HIZRAWATI YAHYA.docx	T3119106-HIZRAWATI YAHYA hjayhya@gmail.com

WORD COUNT	CHARACTER COUNT
9753 Words	53058 Characters
PAGE COUNT	FILE SIZE
63 Pages	1.1MB
SUBMISSION DATE	REPORT DATE
Mar 31, 2023 12:38 PM GMT+8	Mar 31, 2023 12:39 PM GMT+8

● 13% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

• 13% Internet database

• 5% Publications database

• Crossref database

• Crossref Posted Content database

• 0% Submitted Works database

● Excluded from Similarity Report

• Cited material

• Small Matches (Less than 10 words)

Summary



Similarity Report ID: oia:25211:33351233

13% Overall Similarity

Top sources found in the following databases:

- 13% Internet database
- 5% Publications database
- Crossref database
- Crossref Posted Content database
- 0% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	andi.ddns.net	2%
	Internet	
2	123dok.com	1%
	Internet	
3	repository.ub.ac.id	1%
	Internet	
4	repository.nusaputra.ac.id	<1%
	Internet	
5	repository.ittelkom-pwt.ac.id	<1%
	Internet	
6	docplayer.info	<1%
	Internet	
7	mutu.unindra.ac.id	<1%
	Internet	
8	barim165.wordpress.com	<1%
	Internet	

[Sources overview](#)



Similarity Report ID: oia25211:33351233

9	repository.uin-suska.ac.id	<1%
	Internet	
10	publikasi.dinus.ac.id	<1%
	Internet	
11	jurnal.stts.edu	<1%
	Internet	
12	media.neliti.com	<1%
	Internet	
13	eprints.uns.ac.id	<1%
	Internet	
14	etheses.uin-malang.ac.id	<1%
	Internet	
15	seminar.laili.or.id	<1%
	Internet	
16	repository.mercubuana.ac.id	<1%
	Internet	
17	researchgate.net	<1%
	Internet	
18	ejurnal.stmik-budidarma.ac.id	<1%
	Internet	
19	jurnal.itg.ac.id	<1%
	Internet	
20	text-id.123dok.com	<1%
	Internet	

Sources overview

21	hariiniadalahhadiah.wordpress.com	<1%
	Internet	
22	coursehero.com	<1%
	Internet	
23	library.binus.ac.id	<1%
	Internet	
24	download.garuda.kemdikbud.go.id	<1%
	Internet	
25	Pratiwi Juniar Achmad Gani, Adithya Yudistira. "PENGEMBANGAN KA..."	<1%
	Crossref	
26	journal.unusida.ac.id	<1%
	Internet	
27	openlibrarypublications.telkomuniversity.ac.id	<1%
	Internet	
28	Danang Kastowo, Andy Saputra, Wachid Daga Suryono, Erna Setyowati...	<1%
	Crossref	
29	digilib.unhas.ac.id	<1%
	Internet	
30	ejournal.pelitaindonesia.ac.id	<1%
	Internet	
31	jtiik.ub.ac.id	<1%
	Internet	
32	kukamkon.blogspot.com	<1%
	Internet	

33	Dedy Kurniadi, Sam Farisa Chaerul Haviana, Andika Novianto. "Implem...	<1%
	Crossref	
34	Sri Mulyani, Sri Astuti Thamrin, Siswanto Siswanto. "ANALISIS SENTIM...	<1%
	Crossref	
35	doku.pub	<1%
	Internet	
36	ejournal.bsi.ac.id	<1%
	Internet	
37	es.scribd.com	<1%
	Internet	
38	jutif.if.unsoed.ac.id	<1%
	Internet	
39	ojs.unikom.ac.id	<1%
	Internet	
40	repository.pip-semarang.ac.id	<1%
	Internet	
41	repository.usd.ac.id	<1%
	Internet	

Lampiran 2 : Listing Program

```
from nltk.tokenize import RegexpTokenizer
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import euclidean_distances
from sklearn.metrics.pairwise import cosine_similarity
from Sastrawi.StopWordRemover.StopWordRemoverFactory import StopWordRemoverFactory
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
import pandas as pd
import streamlit as st
import plotly.express as px
from PIL import Image
import numpy as np
import time
import operator
import time

st.set_page_config(page_title='Sentimen Analisis')
image = Image.open('images/sentimen.jpg')
st.image(image)
st.subheader('Sentimen Analisis')
st.subheader('Kualitas Pembelajaran Fakultas Ilmu Komputer')

### --- LOAD DATAFRAME
excel_file = 'Data.xlsx'
sheet_name = 'Latih'

df = pd.read_excel(excel_file,
                   sheet_name=sheet_name,
                   header=0)

jumlahdata=df.No.count()
dataset={}
data={}
sentimen={}
for i in range(0,jumlahdata):
    dataset[df.No[i]]=df.Saran[i]
    sentimen[df.No[i]]=df.Kelas[i]
```

```

data[df.No[i]]="["+df.dosen[i]+"]+"df.Saran[i]

image = Image.open('images/unisan.png')
st.sidebar.image(image)
tab1, tab2= st.tabs(["Analisis Sentimen","Pengujian Model"])
with tab1:
    txt=st.text_area('',height=5,placeholder='apa yang anda pikirkan..!',)
    if st.button('Process'):
        dataset["q"]=txt

##Case Folding
for k in dataset.keys():

    hasil_case_folding = dataset[k].lower()
    dataset[k]= hasil_case_folding
st.write("Case Folding:",dataset)
# Tokenizing
tokenizer = RegexpTokenizer(r"\w+")
for k in dataset.keys():
    tokens = tokenizer.tokenize(dataset[k])
    dataset[k] = tokens
st.write("Tokenisasi",dataset)
# Number Removal
for k in dataset.keys():
    tokens = []
    for t in dataset[k]:
        if t.isnumeric() == False:
            tokens.append(t)
    dataset[k] = tokens
# Stopword Removal by Sastrawi
factory = StopWordRemoverFactory()
stopword_list = factory.get_stop_words()
for k in dataset.keys():
    tokens = []
    for t in dataset[k]:
        if t not in stopword_list:
            tokens.append(t)
    dataset[k] = tokens
st.write("Stopword Removing",dataset)

```

```

# Stemming by Sastrawi
factory = StemmerFactory()
stemmer = factory.create_stemmer()
for k in dataset.keys():
    tokens = []
    for t in dataset[k]:
        tokens.append(stemmer.stem(t))
    dataset[k] = tokens
# Mengembalikan Format Dataset Awal
st.write("Stemming",dataset)
for k in dataset.keys():
    dataset[k] = " ".join(dataset[k])
# Frekuensi Kemunculan Kata
tf = CountVectorizer()
term_doc_matrix = tf.fit_transform(dataset.values())
pd.DataFrame(term_doc_matrix.toarray(),
index=dataset.keys(), columns=tf.get_feature_names())

# Pembobotan TF.IDF
tfidf = TfidfVectorizer()
inverted_index = tfidf.fit_transform(dataset.values())
pd.DataFrame(inverted_index.toarray(),
index=dataset.keys(), columns=tfidf.get_feature_names())

# Perhitungan Cosine Similarity
cs = cosine_similarity(inverted_index, inverted_index)
df_cs = pd.DataFrame(cs, index=dataset.keys(),
columns=dataset.keys())
#df_cs

# Pengurutan Reranking Cosine Similarity
rank_cs = {}
for k in dataset.keys():
    if k != "q":
        rank_cs[k] = df_cs.at[k, 'q']

top_rank_cs = dict(sorted(rank_cs.items(),
key=operator.itemgetter(1), reverse=True))
tampil=pd.DataFrame(top_rank_cs.values(),
index=top_rank_cs.keys(),columns=["Cosine Similarity"])

```

```

positif=[]
negatif=[]
netral=[]
id_rank=[]
for j in top_rank_cs.keys():
    if (top_rank_cs[j] >= 1):
        top_rank_cs[j]=1.00
        nilai=top_rank_cs[j]*100
    else :
        nilai=top_rank_cs[j]*100
    id_rank.append(j)
    if (sentimen[j]=="negatif"):
        negatif.append(top_rank_cs[j])
    elif sentimen[j]=="positif":
        positif.append(top_rank_cs[j])
    elif sentimen[j]=="netral":
        netral.append(top_rank_cs[j])
    st.write(j,':',sentimen[j],nilai,'%')
x1=netral[0]*100
x2=negatif[0]*100
x3=positif[0]*100
nos=id_rank[0]
chart_data = pd.DataFrame(
    [[0,0,x2],[0,x1,0],[x3,0,0]],
    columns=["x1.Positif", "x2.Neutral", "x3.Negatif"])
st.sidebar.bar_chart(chart_data, width=100,height=150)

if (sentimen[nos]=="negatif"):
    st.sidebar.error("[ "+sentimen[nos]+" ]..." +txt,
icon="🚫")
elif sentimen[nos]=="positif":
    st.sidebar.info("[ "+sentimen[nos]+" ]..." +txt,
icon="✅")
elif sentimen[nos]=="netral":
    st.sidebar.success("[ "+sentimen[nos]+" ]..." +txt,
icon="👍")

```

Lampiran 3 : Dataset

Dosen	Mata Kuliah	Komentar
IRVAN A SALIHI, M.Kom	ALGORITMA DAN STRUKTUR DATA	Kalau bisa cara mengajar lebih dipermudah..
		saran saya dosen harus lebih kocak sedikit wkwk
		Tidak ada
		Tidak ada
		Baik
		Seketika di jelaskan mudah paham
		Harus bisa lebih baik dari biasanya
		Terimakasih atas ilmunya
		Penjelasan materi harus pelan-pelan
		Tetap di pertahankan sistem mengajarnya
		Tidak ada
		Saran saya semoga mahasiswa bisa lebih paham
		saran saya mohon menindak tegas bagi mahasiswa yg suka terlambat. atau yang tidak kerjakan tugas.
		Tingkatkan lagi kualitas dosen
		tidak ada lagi saran karena cara mengajar sudah bagus
		Jangan pernah lelah pak menghadapi kami yang minim akan logika pemrograman
		Tidak ada saran dari saya . Saya hanya mau mengucapkan TERIMA KASIH TELAH MENGAJARKAN KAMI DENGAN PENUH BANYAK KESABARAN PAK!
		Tidak ada saran
ZULFRIANTO Y LAMASIGI,M.Kom	PENGOLAHAN CITRA	Lebih ditingkatkan, maaf pak dosen.
		Mantap
		Baik skaliinii
		Baik
		Sanagt mengenagkan
		Tidak ada saran
		Tidak ada
		Tidak ada
		Menjelaskan materi harus pelan-pelan
		Mengajar harus lebih di perjelas lagi walaupun sudah ada buku di tiap2 mahasiswa agar mahasiswa bisa lebih cepat memahami.
		Tidak ada
		Saran saya semoga bisa lebih paham dengan cara mengajar bapak
		dipertahankan cara mendidik mahasiswanya
		Tingkatkan lagi kualitas dosen
		Tidak ada saran pak . Makasih
		berikan solusi yang pasti kepada mahasiswa yg mengalami kesulitan dalam codingnya yg bermasalah
ANDI BODE, M.Kom	JARINGAN KOMPUTER LANJUTAN	Semoga akan lebih baik dari hari hari kemarin
		semoga kedepanya lebih baik dari hari ini
		Kedepannya harus lebih baik lagi
		sangat baik
		Baik
		Dosen keren, tidak ada saran, mohon maaf Pak saya tidak masuk"

		Menurut saya Pak Bayu memberikan metode belajar yang sangat baik sehingga mahasiswa dapat mudah mengerti
		Terima kasih pak
		Tidak ada
		Sangat baik
		Tidak ada saran
		Tidak ada saran
		Baik
		saran, hanya berusaha untuk menjadi yang lebih baik lagi. terimah kasih
		Tidak ada saran
		Terima kasih atas pemberian ilmunya
		UntuK MK ini sangat bagus dan luar biasa banyak prakteknya jadi mahasiswa lebih cept mengerti
		No saran
		Permantap cara mengajar
		Semoga kedepannya pak mengajar di reguler B lagi
DOSEN UMUM	SISTEM DIGITAL	Tdk ada
		Sudah cukup baik
		tidak ada
		Semoga kedepannya tetap memberikan yang terbaik
		Dosen umum yang datang mengajar kurang
		Tingkatkan proses belajar
		Dosen yang sangat baik
		Diharapkan agar selalu memberikan pembelajaran dengan baik
		Tidak Ada
		Hak dan kewajiban jgn d lupakan
		Mohon bimbingannya
		Dipertahankan
		Lebih memberikan tugas dan praktek
		Sebaiknya memberikan diskusi setiap pertemuan
		Tidak ada
		sangat baik
		Lebih tingkatkan kehadiran
		Agar bisa memberikan diskusi.
REZQIWATI ISHAK, M.Kom	METODE NUMERIK	Semoga bisa menjadi pribadi yang lebih baik..
		Tidak ada
		Sudah baik
		Masih bisah di pahami
		tetap dalam proses yg seperti ini
		Tidak ada saran
		Luar biasa
		Pertahankan cara mengajar ibu.. Karena cara mengajar ibu sngat baik
		Tidak ada saran, pokoknya ibu kiki terbaik□
		Tidak ada
		Terbaik.
		Tetap pertahankan cara mengajarnya
		Tidak ada
		ibu mengajar dengan sangat baik sangat disiplin waktu saran saya lebih di tambah lagi disiplinnya

REZQIWATI ISHAK, M.Kom	METODE NUMERIK	dipertahankan cara mendidik mahasiswanya.
		Tingkatkan lagi kualitas dosen
		Tidak ada lagi saran karena cara menjelaskan sudah baik sehingga mahasiswa mudah paham apa yang di sampaikan dosen
		Sehat selalu ibu
		Tidak ada saran dari saya . Saya hanya mau mengucapkan TERIMA KASIH TELAH MENGAJARKAN KAMI DENGAN PENUH BANYAK KESABARAN BU!
		Perfect.
		Sangat baik
		semoga lebih semangat lagi cara mengajarnya
YULIANTY LASENA, M.Kom	METODE NUMERIK	Semoga bisa menjadi pribadi yang lebih baik..
		Tidak ada
		Ok
		Bagus
		lebih menarik lagi dalam proses pengajar agar kami mahasiswa bisa tertarik dalam mengikuti kegiatan pebelajaran
		Tidak ada saran
		Terimah kasih sudah mengajar di semester ini
		Mantap
		Mata kuliah yang paling menyenangkan dan paling di tggu
		Tidak ada
		semoga lebih terampil dalam mengajar
		Metode penjelasan materi harus diperjelas dan dikembangkan lagi.
		Cara mengajar boleh di perjelas lagi agar mahasiswa bisa lebih cepat memahami
		Tidak ada
		Saran saya semoga cara mengajar di MK ini akan lebih baik dan dapat di mengerti
		dipertahankan cara mengajar mahasiswanya
		Tingkatkan lagi kualitas dosen
		Diharapkan lagi cara mengajarnya lebih baik agar mahasiswa lebih mengerti lagi
		Lebih baik lagi
		Tidak ada saran dari saya . Saya hanya mau mengucapkan TERIMA KASIH TELAH MENGAJARKAN KAMI DENGAN PENUH BANYAK KESABARAN BU!
		Cara mengajar perlu ditingkatkan, karena hanya sebagian mahasiswa yg paham selebihnya mereka mencari tau sendiri materi yg sudah dijelaskan
dst		

Lampiran 4 :
RIWAYAT HIDUP PENELITI



Nama : Hizrawati Yahya
Nim : T3119106
Tempat,Tanggal Lahir: Gorontalo, 28 Juni 2001
Email : hjrayhya@gmail.com

Riwayat Pendidikan:

1. Tahun 2006, menyelesaikan Pendidikan TK di Taman Kanak-Kanak Ade Irma, Tamboo, Kec.Tilongkabila
2. Tahun 2012, menyelesaikan Pendidikan di Sekolah Dasar Negeri 03 Tilongkabila, Kec. Tilongkabila
3. Tahun 2016, menyelesaikan Pendidikan di Sekolah Menengah Pertama Negeri 1 Kabila, Kec. Kabila
4. Tahun 2019, menyelesaikan Pendidikan di Sekolah Menengah Atas Negeri 1 Kabila, Kec.Kabila
5. Tahun 2019, telah diterima menjadi Mahasiswa di Perguruan Tinggi Swasta Universitas Ichsan Gorontalo



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS ICHSAN GORONTALO
LEMBAGA PENELITIAN

Kampus Unisan Gorontalo LLS - Jln. Achmad Nadjamuddin No. 17 Kota Gorontalo
Telp: (0435) 8724465, 829975 E-Mail: lembaga penelitian@unisan.ac.id

Nomor : 4374/PIP/LEMLIT-UNISAN/GTO/X/2022

Lampiran : -

Hal : Permohonan Izin Penelitian

Kepada Yth,

Dekan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo

di,-

Tempat

Yang bertanda tangan di bawah ini :

Nama : Dr. Rahmisyari, ST.,SE.,MM

NIDN : 0929117202

Jabatan : Ketua Lembaga Penelitian

Meminta kesediaannya untuk memberikan izin pengambilan data dalam rangka penyusunan **Proposal /**

Skripsi, kepada :

Nama Mahasiswa : Hizrawati Yahya

NIM : T3119106

Fakultas : Fakultas Ilmu Komputer

Program Studi : Teknik Informatika

Lokasi Penelitian : FAKULTAS ILMU KOMPUTER UNIVERSITAS ICHSAN
GORONTALO

Judul Penelitian : ANALISIS SENTIMEN MAHASISWA TERHADAP MATA
KULIAH DI FAKULTAS ILMU KOMPUTER UNIVERSITAS
ICHSAN GORONTALO MENGGUNAKAN PEMBOBOTAN
TF-IDF

Atas kebijakan dan kerja samanya diucapkan banyak terima kasih.

Gorontalo, 31 Oktober 2022
Ketua

Dr. Rahmisyari, ST.,SE.,MM
NIDN 0929117202

+



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI

UNIVERSITAS ICHSAN GORONTALO

FAKULTAS ILMU KOMPUTER

SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001

Jl. Achmad Nadjamuddin No. 17 Telp (0435) 829975 Fax (0435) 829976 Gorontalo

SURAT KETERANGAN PENELITIAN

Nomor : I1/PIKOM-UIG/SKP/1/2023

Yang beranda tangan di bawah ini :

Nama : Irvan Abraham Salibi, M.Kom
N I D N : 0928028101
Jabatan : Dekan Fakultas Ilmu Komputer

Dengan ini Menerangkan bahwa :

Nama Mahasiswa : Hizrawati Yahya
N I M : T3119106
Program Studi : Teknik Informatika

Bahwa yang bersangkutan benar-benar telah melakukan penelitian tentang "Analisis Sentimen Mahasiswa Terhadap Mata Kuliah Di Fakultas Ilmu Komputer Universitas Ihsan Gorontalo Menggunakan Pembobotan TF.IDF(studi kasus Fakultas Ilmu Komputer) " Guna untuk menyelesaikan Studi pada Program Studi Teknik Informatika Fakultas Ilmu Komputer, dan bersangkutan telah menyelesaikan penelitian Tersebut pada TGL 10 Januari 2023 sesuai dengan waktu yang telah di tentukan.

Demikian Surat Keterangan ini dibuat dan digunakan untuk seperlunya.

Gorontalo, 10 Januari 2023
Dekan,

Irvan Abraham Salibi, M.Kom
NIDN. 0928028101



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS ICHSAN GORONTALO

FAKULTAS ILMU KOMPUTER

SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001

Jl. Achmad Najamuddin No. 17 Telp. (0435) 829975 Fax (0435) 829976 Gorontalo

SURAT REKOMENDASI BEBAS PLAGIASI

No. 120/FIKOM-UIG/R/IV/2023

Yang bertanda tangan di bawah ini :

Nama : Irvan Abraham Salihi, M.Kom
NIDN : 0928028101
Jabatan : Dekan Fakultas Ilmu Komputer

Dengan ini menerangkan bahwa :

Nama Mahasiswa : Hizrawati Yahya
NIM : T3119106
Program Studi : Teknik Informatika (S1)
Fakultas : Fakultas Ilmu Komputer
Judul Skripsi : Analisis Sentimen Mahasiswa Terhadap Mata Kuliah
Di Fakultas Ilmu Komputer Universitas Ichsan
Gorontalo Menggunakan Pembobotan TF.IDF

Sesuai hasil pengecekan tingkat kemiripan skripsi melalui aplikasi **Turnitin** untuk judul skripsi di atas diperoleh hasil *Similarity* sebesar **13%**, berdasarkan Peraturan Rektor No. 32 Tahun 2019 tentang Pendeteksian Plagiat pada Setiap Karya Ilmiah di Lingkungan Universitas Ichsan Gorontalo dan persyaratan pemberian surat rekomendasi verifikasi calon wisudawan dari LLDIKTI Wil. XVI, bahwa batas kemiripan skripsi maksimal 30%, untuk itu skripsi tersebut di atas dinyatakan **BEBAS PLAGIASI** dan layak untuk diujikan.

Demikian surat rekomendasi ini dibuat untuk digunakan sebagaimana mestinya.

Mengetahui
Dekan,

Irvan Abraham Salihi, M.Kom
NIDN: 0928028101

Gorontalo, 3 April 2023
Tim Verifikasi,


Zulfrianto W. Lamasigi, M.Kom
NIDN. 0914039101

Terlampir :
Hasil Pengecekan Turnitin



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS ICHSAN GORONTALO
FAKULTAS ILMU KOMPUTER
UPT. PERPUSTAKAAN FAKULTAS
SK. MENDIKNAS RI NO. 84/D/0/2001
Jl. Achmad Nadjamuddin No.17 Telp(0435) 829975 Fax. (0435) 829976 Gorontalo

SURAT KETERANGAN BEBAS PUSTAKA

No : 013/Perpustakaan-Fikom/III/2023

Perpustakaan Fakultas Ilmu Komputer (FIKOM) Universitas Ichsan Gorontalo dengan ini menerangkan bahwa :

Nama Anggota : Hizrawati Yahya
No. Induk : T3119106
No. Anggota : M202319

Terhitung mulai hari, tanggal : Senin, 13 Maret 2023, dinyatakan telah bebas pinjam buku dan koleksi perpustakaan lainnya.

Demikian keterangan ini di buat untuk di gunakan sebagaimana mestinya.



Gorontalo, 13 Maret 2023

**Mengetahui,
Kepala Perpustakaan**

Apriyanto Alhamad, M.Kom

NIDN : 0924048601

