

**SENTIMEN ANALISIS KOTA TEGAL BERBASIS ASPEK  
MENGUNAKAN ALGORITMA *NAÏVE BAYES***



**SKRIPSI**

**Diajukan Sebagai Salah Satu Syarat Untuk Menyelesaikan Studi Pada  
Program Studi Teknik Informatika**

**Oleh :**

**NOVITA FITRIA PUTRI**

**19090130**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**POLITEKNIK HARAPAN BERSAMA**

**TEGAL**

**2023**

**SENTIMEN ANALISIS KOTA TEGAL BERBASIS ASPEK  
MENGUNAKAN ALGORITMA *NAÏVE BAYES***

JUDUL



**SKRIPSI**

**Diajukan Sebagai Salah Satu Syarat Untuk Menyelesaikan Studi Pada  
Program Studi Teknik Informatika**

**Oleh :**

**NOVITA FITRIA PUTRI**

**19090130**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**POLITEKNIK HARAPAN BERSAMA**

**TEGAL**

**2023**



## HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama : Novita Fitria Putri

NIM : 19090130

adalah mahasiswa Program Studi Teknik Informatika Politeknik Harapan bersama.

Dengan ini saya menyatakan bahwa laporan Skripsi yang berjudul:

**“Sentimen Analisis Kota Tegal Berbasis Aspek Menggunakan  
Algoritma *Naïve Bayes*”**

merupakan hasil pemikiran sendiri secara orisinil yang saya susun secara mandiri dengan tidak melanggar kode etik hak karya cipta. Apabila dikemudian hari Laporan Skripsi ini terbukti melanggar kode etik karya cipta, maka saya bersedia untuk melakukan penelitian baru dan menyusun laporannya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenar-benarnya.

Tegal, 25 Agustus 2023

Yang membuat pernyataan,



Novita Fitria Putri

NIM. 19090130

## HALAMAN REKOMENDASI


Pembimbing Tugas Akhir memberikan rekomendasi kepada :

Nama : NOVITA FITRIA PUTRI  
NIM : 19090130  
Program Studi : Sarjana Terapan Teknik Informatika  
Judul Tugas Akhir : SENTIMEN ANALISIS KOTA TEGAL BERBASIS ASPEK  
MENGUNAKAN ALGORITMA *NAÏVE BAYES*


Untuk mengikuti Ujian Tugas Akhir karena telah memenuhi persyaratan yang telah ditentukan.

Pembimbing I,

Tegal, Juli 2023  
Pembimbing II,



Muhammad Fikri Hidayattullah, S.T., M.Kom.  
NIPY. 09.016.307



Dwi Intan Afidah, S.T., M.Kom.  
NIPY. 11.020.470

## HALAMAN PENGESAHAN

Nama : Novita Fitria Putri  
NIM : 19090130  
Program Studi : Sarjana Terapan Teknik Informatika  
Judul Skripsi : Sentimen Analisis Kota Tegal Berbasis Aspek  
Menggunakan Algoritma *Naïve Bayes*

dinyatakan lulus Ujian Skripsi pada Program Sptudi Sarjana Terapan Teknik Informatika Politeknik Harapan Bersama.




Tegal, Agustus 2023

Dewan Penguji

Nama

1. Ketua : Dyah Apriliani, S.T., M.Kom.
2. Anggota I : Dairoh, M.Sc.
3. Anggota II : Dwi Intan Af'idah, S.T., M.Kom.

Tanda Tangan

1. 
2. 
3. 

Mengetahui,

Ketua Program Studi Sarjana Terapan Teknik Informatika



Slamet Wiyono, S.Pd., M.Eng.

NIPY. 08.015.222

## ABSTRAK

Kurangnya ketentuan yang jelas mengenai penggunaan *e-mail* institusi oleh masyarakat Kota Tegal dalam menyampaikan informasi, saran, dan pengaduan kepada pemerintah. Sebagai gantinya, masyarakat cenderung menggunakan media sosial seperti *Twitter* untuk menyampaikan pendapat dan keluhan mereka, namun pemerintah mengalami kesulitan dalam menanggapi secara efektif. Penelitian ini bertujuan untuk mengembangkan sebuah aplikasi yang dapat mengklasifikasikan ulasan masyarakat Kota Tegal berdasarkan beberapa aspek seperti wisata/hiburan, pendidikan, fasilitas/layanan publik, dan kuliner. Melalui analisis sentimen berbasis aspek, aplikasi ini dapat memberikan informasi kepada pemerintah mengenai pendapat dan keluhan masyarakat secara spesifik. Penelitian ini menggunakan metode analisis sentimen dengan menggunakan algoritma Naïve Bayes untuk mengklasifikasikan ulasan masyarakat sebagai sentimen positif atau negatif. Pengujian dilakukan dengan menggunakan data ulasan masyarakat Kota Tegal yang dikumpulkan dari *Twitter*. Hasil penelitian menunjukkan bahwa aplikasi yang dikembangkan dapat mengklasifikasikan ulasan masyarakat Kota Tegal ke dalam beberapa aspek dengan tingkat akurasi pada data training diatas 90% dan tingkat akurasi pada data testing diatas 75% pada setiap kategori. Diharapkan aplikasi ini dapat membantu pemerintah dalam meningkatkan kualitas Kota Tegal dengan menanggapi pendapat dan keluhan masyarakat secara lebih efektif dan spesifik.

**Kata kunci :** Analisis Sentimen, Klasifikasi Aspek, *Naïve Bayes*, Kota Tegal.

## KATA PENGANTAR

Dengan memanjatkan puji syukur kehadirat Allah SWT, Tuhan Yang Maha Pengasih dan Maha Penyayang yang telah melimpahkan segala rahmat, hidayah dan inayah-Nya hingga terselesaikannya laporan Skripsi dengan judul

**“SENTIMEN ANALISIS KOTA TEGAL BERBASIS ASPEK  
MENGUNAKAN ALGORITMA *NAÏVE BAYES*”.**

Skripsi merupakan suatu kewajiban yang harus dilaksanakan untuk memenuhi salah satu syarat kelulusan dalam mencapai derajat Sarjana Sain Terapan pada Program Studi Sarjana Terapan Teknik Informatika Politeknik Harapan Bersama. Selama melaksanakan penelitian dan kemudian tersusun dalam laporan Tugas Akhir ini, banyak pihak yang telah memberikan bantuan, dukungan dan bimbingan.

Pada kesempatan ini, tak lupa penulis mengucapkan terimakasih yang sebesar-besarnya kepada :

1. Bapak Agung Hendarto, S.E., M.A. selaku Direktur Politeknik Harapan Bersama Tegal.
2. Slamet Wiyono, S.Pd., M.Eng. selaku Ketua Program Studi Sarjana Terapan Teknik Informatika Politeknik Harapan Bersama Tegal
3. Muhammad Fikri Hidayattullah, S.T., M.Kom. selaku Dosen Pembimbing I
4. Dwi Intan Af'idah, S.T., M.Kom. selaku Dosen Pembimbing II
5. Semua pihak yang telah mendukung, membantu serta mendoakan penyelesaian laporan Tugas Akhir ini.

Semoga laporan skripsi ini dapat memberikan sumbangan untuk pengembangan ilmu pengesatuan dan teknologi

Tegal, Agustus 2023  
Penulis

Novita Fitria Putri

## DAFTAR ISI

JUDUL .....	i
HALAMAN PERNYATAAN .....	ii
HALAMAN REKOMENDASI .....	ii
HALAMAN PENGESAHAN.....	iii
ABSTRAK .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR .....	ix
DAFTAR TABEL.....	x
DAFTAR LAMPIRAN.....	xi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Tujuan dan Manfaat.....	4
1.3. Tinjauan Pustaka .....	5
1.4. Data Penelitian.....	10
BAB II PRODUK .....	13
2.1. Perancangan.....	13
2.2. Perancangan Aplikasi .....	13
2.2.1. Desain Antarmuka.....	33
2.2.2. Implementasi Sistem .....	38
2.3. Pengujian.....	42
2.3.1. Pengujian Model .....	42
2.3.2. Pengujian Implementasi Model.....	45
2.3.3. Pengujian Aplikasi .....	46
2.3.4. Hasil Analisis Sentimen Perbulan pada Setiap Aspek Kategori .....	69
2.4. Kesimpulan dan Saran.....	71
BAB III HKI .....	73
1.1. Proses.....	73

1.2. Identitas HKI .....	73
DAFTAR PUSTAKA .....	75
LAMPIRAN .....	1

## DAFTAR GAMBAR

Gambar 1. 1 Dataset.....	11
Gambar 1. 2 Data Hasil Implementasi .....	12
Gambar 2. 1 Alur Pancangan Produk.....	14
Gambar 2. 2 TF-IDF .....	20
Gambar 2. 3 Confusion Matrix .....	27
Gambar 2. 4 Alur Implementasi Model .....	30
Gambar 2. 5 Alur Pembuatan Website.....	32
Gambar 2. 6 Menu Home.....	33
Gambar 2. 7 Menu Dashboard .....	34
Gambar 2. 8 Pilih Kategori .....	35
Gambar 2. 9 Tanpa Filter .....	36
Gambar 2. 10 Fitur Tanggal.....	37
Gambar 2. 11 Implementasi Menu Home .....	38
Gambar 2. 12 Implementasi Menu Dashboard .....	39
Gambar 2. 13 Implementasi Pilih Fitur.....	39
Gambar 2. 14 Tanpa Filter .....	40
Gambar 2. 15 Tanpa Filter .....	40
Gambar 2. 16 Implementasi Fitur Tanggal .....	41
Gambar 2. 17 Implementasi Model.....	46



## DAFTAR TABEL

Tabel 2. 1 Identifikasi Aspek .....	15
Tabel 2. 2 Pelabelan Kelas .....	15
Tabel 2. 3 Pengujian Model Aspek Training .....	42
Tabel 2. 4 Pengujian Model Aspek Testing .....	42
Tabel 2. 5 Hasil pengujian model sentimen aspek training .....	43
Tabel 2. 6 Hasil pengujian model sentimen aspek testing .....	43
Tabel 2. 7 Hasil Pengujian Klasifikasi Aspek.....	44
Tabel 2. 8 Hasil Pengujian Sentimen .....	44
Tabel 2. 9 Testing Black Box.....	47
Tabel 2. 10 Tabel Variabel Penilaian.....	50
Tabel 2. 11 Tabel Pernyataan Positif .....	51
Tabel 2. 12 Tabel Pernyataan Negatif .....	51
Tabel 2. 13 Tabel Kriteria Penilaian .....	52
Tabel 2. 14 Tabel Parameter Bobot Nilai Positif .....	53
Tabel 2. 15 Hasil Pengujian Usability Pernyataan Positif No. 1 .....	54
Tabel 2. 16 Hasil Pengujian Usability Pernyataan Positif No. 2 .....	55
Tabel 2. 17 Hasil Pengujian Usability Pernyataan Positif No. 3 .....	57
Tabel 2. 18 Hasil Pengujian Usability Pernyataan Positif No. 4 .....	58
Tabel 2. 19 Hasil Pengujian Usability Pernyataan Positif No. 5 .....	60
Tabel 2. 20 Hasil Pengujian Usability Pernyataan Negatif No. 1 .....	61
Tabel 2. 21 Hasil Pengujian Usability Pernyataan Negatif No. 2 .....	63
Tabel 2. 22 Hasil Pengujian Usability Pernyataan Negatif No. 3 .....	64
Tabel 2. 23 Hasil Pengujian Usability Pernyataan Negatif No. 4 .....	65
Tabel 2. 24 Hasil Pengujian Usability Pernyataan Negatif No. 5 .....	67
Tabel 2. 25 Hasil Sentimen Analisis Setiap Aspek.....	69

## **DAFTAR LAMPIRAN**

Lampiran 1 Surat Kesediaan Pembimbing.....	A-1
Lampiran 2 Surat Pernyataan HKI.....	B-1
Lampiran 3 Surat Pengalihan HKI.....	C-1
Lampiran 4 Manual Book .....	D-1
Lampiran 5 Dokumen Teknikal .....	E-1
Lampiran 6 Sertifikat HKI yang Terbit.....	F-1
Lampiran 7 Lembar Bimbingan .....	G-1

## **BAB I**

### **PENDAHULUAN**

#### **1.1. Latar Belakang**

Kota Tegal adalah salah satu kota di Jawa Tengah yang memiliki banyak potensi yang menarik untuk dikembangkan guna meningkatkan kualitas dari Kota Tegal. Sejalan dengan hal tersebut, Walikota Tegal telah membuat keputusan untuk menyediakan alamat *e-mail* resmi pemerintah kota sebagai sarana untuk memperkuat hubungan antara masyarakat dan instansi pemerintah. Tujuan dari keputusan tersebut adalah untuk memberikan pelayanan yang efektif dan terbuka kepada masyarakat dengan memanfaatkan teknologi informasi. Namun, hingga saat ini, masih belum terdapat ketentuan yang jelas mengenai standar operasional penggunaan *e-mail* oleh masyarakat ketika membutuhkan informasi, memberikan saran atau pendapat, serta menyampaikan pengaduan kepada pemerintah [1], masyarakat Kota Tegal lebih sering menyampaikan pendapat dan keluhannya melalui media sosial yang ada, salah satunya adalah *twitter*. Namun karena banyaknya opini masyarakat di *twitter* membuat pemerintah kesulitan untuk memberikan tanggapan tentang pendapat atau keluhan yang ada. Karena pengembangan kualitas Kota Tegal memberikan peranan penting dalam banyak hal, yaitu meningkatkan ekonomi dan meningkatkan kualitas dari fasilitas dan juga pelayanan publik, maka diperlukan klasifikasi tentang ulasan dari masyarakat yang diklasifikasikan kedalam beberapa aspek sehingga pemerintah bisa meninjau lebih dalam tentang keluhan yang ada dan lebih spesifik.

Media sosial merupakan teknologi yang memiliki peran penting untuk tanggap terhadap beberapa situasi, terlebih lagi karena teknologi sendiri adalah perangkat yang memungkinkan untuk bertukar informasi secara terbuka dan relatif cepat. Media sosial memudahkan pemerintah untuk menyebarkan informasi mengenai pemerintahan, serta tentu saja mengumpulkan umpan balik dari opini publik mengenai kebijakan daerah [2]. Karena banyaknya masyarakat Kota Tegal yang memanfaatkan media sosial sebagai sarana untuk menyalurkan pendapat dan juga keluhan seputar Kota Tegal, maka komentar masyarakat tersebut memberikan manfaat kepada pemerintah untuk dijadikan sebagai alat bantu untuk mengambil keputusan dalam meningkatkan kualitas Kota Tegal. Ulasan dari *twitter* ini merupakan ulasan masyarakat tentang Kota Tegal yang bersifat positif atau negatif.

Analisis sentimen berbasis aspek adalah tugas yang bertujuan untuk mengidentifikasi sentimen atau polaritas opini terhadap berbagai aspek yang telah disepakati secara jelas atau terdapat kategori aspek yang tersirat dalam teks bahasa alami yang dibuat oleh pengguna [3]. Untuk melakukan analisis sentimen, dapat digunakan metode klasifikasi, Salah satu metode yang umum digunakan adalah *Naïve Bayes*. Pada suatu penelitian *naive bayes* terbukti memiliki hasil akurasi, presisi, *recall*, dan *F1-Score* yang lebih baik dari pada metode *machine learning* SVM [4]. Tantangan dalam analisis sentimen sering kali muncul ketika terdapat ketidakseimbangan data antara kelas-kelas yang ada. Penggunaan model dengan data yang tidak seimbang cenderung menghasilkan akurasi prediksi yang rendah [5].

*Naïve bayes* merupakan model klasifikasi yang paling sederhana dan paling sering digunakan, *Naïve bayes* merupakan *machine learning* yang memanfaatkan perhitungan probabilitas suatu kelas berdasarkan persebaran kata di dalam dokumen [6]. Tujuan dari pengujian algoritma ini adalah agar dapat otomatis mengklasifikasikan ulasan sebagai sentiment yang bernilai positif atau bernilai negatif. Sentimen positif mengindikasikan adanya pandangan yang positif terhadap suatu konteks, sedangkan sentimen negatif mencerminkan pandangan negatif terhadap suatu konteks. Sentimen netral menyiratkan sikap yang netral, tidak memihak pada sisi positif atau negatif. Dalam penelitian ini, algoritma yang digunakan untuk analisis sentimen adalah *Naïve Bayes* [7].

Kesimpulan dari latar belakang ini adalah bahwa Kota Tegal memiliki potensi yang menarik untuk dikembangkan guna meningkatkan kualitasnya, dan pemerintah kota telah menyediakan alamat email resmi sebagai sarana untuk memperkuat hubungan dengan masyarakat melalui teknologi informasi. Namun, masih belum terdapat ketentuan yang jelas mengenai standar operasional penggunaan email oleh masyarakat, sehingga mereka lebih sering menyampaikan pendapat dan keluhan melalui media sosial, terutama Twitter. Ini menyebabkan kesulitan bagi pemerintah dalam memberikan tanggapan secara spesifik terhadap keluhan dan pendapat masyarakat. Oleh karena itu, diperlukan klasifikasi ulasan dari masyarakat ke dalam beberapa aspek agar pemerintah dapat meninjau lebih dalam tentang keluhan yang ada. Analisis sentimen berbasis aspek menggunakan metode klasifikasi *Naïve Bayes* dapat menjadi solusi untuk mengidentifikasi sentimen atau polaritas opini terhadap berbagai

aspek yang ada dalam teks bahasa alami. Hal ini akan membantu pemerintah dalam mengambil keputusan yang lebih tepat dalam meningkatkan kualitas Kota Tegal berdasarkan respon masyarakat yang diungkapkan melalui media sosial salah satunya adalah *twitter*.

## **1.2. Tujuan dan Manfaat**

### **1.2.1. Tujuan**

Tujuan penelitian ini adalah untuk membuat pemodelan sentimen analisis berbasis aspek terhadap tanggapan masyarakat mengenai beberapa aspek di Kota Tegal menggunakan algoritma naïve bayes dan diimplementasikan ke *platform* berbasis *website* secara *real time*

### **1.2.2. Manfaat**

Adapun manfaat yang didapat dari penelitian ini , antara lain:

#### **1. Bagi Mahasiswa**

- Menambah pengetahuan dan peningkatan pemahaman mengenai penggunaan dan penerapan Algoritma *Naïve Bayes*
- Penerapan pengetahuan yang diperoleh selama masa studi di Program Studi Sarjana Terapan Teknik Informatika Politeknik Harapan Bersama.

#### **2. Bagi Masyarakat**

- Memperoleh informasi tentang respon masyarakat terhadap aspek-aspek di Kota Tegal, seperti Aspek Wisata/Hiburan, Aspek Pendidikan, Aspek Fasilitas dan Publik, dan juga Aspek Kuliner

- Meningkatkan kualitas Kota Tegal secara keseluruhan dengan memperhatikan preferensi dan kebutuhan masyarakat

### 3. Bagi Politeknik Harapan Bersama

- Sebagai sumber pengetahuan dan referensi mengenai Algoritma *Naïve Bayes*
- Menambahkan kepustakaan di Politeknik Harapan Bersama Tegal

## 1.3. Tinjauan Pustaka

Terdapat penelitian terdahulu yang melakukan sentimen analisis berbasis aspek, pada penelitian ini berisi ulasan tentang kuliner di Kota Bandung pada situs TripAdvisor. Penelitian ini menggunakan teknik klasifikasi data mining dengan algoritma *Naïve Bayes* dilanjutkan dengan pelabelan multi aspek. Selain itu analisis dilakukan pada bobot ulasan berdasarkan jumlah kontribusi pemberi ulasan di *website* TripAdvisor. Hasil pengujian menunjukkan bahwa penggunaan seluruh kombinasi metode tersebut dalam proses klasifikasi sentiment dan menghasilkan nilai akurasi sebesar 98,67%. Dapat disimpulkan pada penelitian ini adalah kombinasi metode tersebut dapat menghasilkan akurasi yang lebih tinggi dari penelitian sebelumnya yang tingkat akurasinya sebesar 88.78% [8].

Pada penelitian lainnya yang bertujuan untuk melihat bagaimana respon pelanggan atau masyarakat umum terhadap perusahaan penyedia layanan internet dan komunikasi sehingga dapat dijadikan patokan untuk evaluasi terhadap layanan yang diberikan kepada pelanggan. Dari hasil penelitian tersebut dapat disimpulkan dengan

melihat nilai *f1-score* yang didapatkan bernilai tinggi yaitu 91.62% sehingga dapat dijadikan patokan untuk perusahaan tersebut agar dapat menjadikan evaluasi pelayanan dan kualitas [10].

Riset yang pernah dilakukan yang bertujuan menganalisis setiap teks untuk mengidentifikasi berbagai aspek kemudian menentukan tingkatan sentimen yaitu positif, negatif, dan netral menggunakan TF-IDF untuk pembobotan kata dan *Naïve Bayes* sebagai metode untuk mengklasifikasi sentimen. Dari hasil pengujian yang diperoleh nilai *F1-Score* sebesar 62,81% untuk data yang di terjemahkan ke dalam Bahasa Inggris kemudian ke dalam Bahasa Indonesia dan tidak menggunakan *stopword removal* [11].

Pada penelitian sebelumnya yang bertujuan untuk membantu perusahaan mengetahui respon masyarakat terhadap merk dagangnya melalui *review* yang ada dengan menggunakan metode *Naïve Bayes* untuk dapat memprediksi kelas sentimen pada ulasan produk online sesuai dengan sistem yang disiapkan. Nilai akurasi terendah yang dihasilkan sebesar 52,66% pada pengujian 5 kelas menggunakan dataset 80% latih dan 20% data uji. Sedangkan pada pengujian 3 kelas dengan 90% data latih dan 10% data uji memiliki nilai akurasi tertinggi sebesar 77.8% [12].

Pada penelitian sebelumnya yang bertujuan untuk mengembangkan pendukung keputusan menggunakan metode text mining dan sentiment analysis dalam menilai opini pengunjung wisata terhadap tempat wisata di Tegal dan sekitarnya. Metode klasifikasi *Naïve Bayes* dan *Decision Tree* diterapkan, dan hasil pengujian menunjukkan bahwa *Naïve Bayes* memiliki tingkat akurasi 73,33%, lebih unggul



daripada Decision Tree. Meskipun begitu, penelitian ini masih membutuhkan optimasi lebih lanjut dengan menggunakan metode feature weight atau feature selection guna meningkatkan akurasi model sentiment analysis untuk pengambilan keputusan yang lebih tepat bagi pengelolaan tempat wisata di wilayah tersebut [13]

Dari beberapa penelitian yang telah dilakukan terdapat banyak faktor yang dapat mempengaruhi tingkat akurasi yang dihasilkan pada setiap algoritma yang digunakan. Pada beberapa penelitian yang menggunakan algoritma *Naïve Bayes* menghasilkan nilai akurasi yang cukup tinggi karena pengambilan dataset dan *preprocessing* data yang dilakukan dengan membagi kedalam beberapa kelas. Dan juga dari berbagai penelitian sebelumnya telah dilakukan untuk menganalisis sentimen dalam berbagai konteks, seperti ulasan kuliner, ulasan gadget, tanggapan pelanggan terhadap layanan, dan opini pengunjung wisata. Dalam upaya mengidentifikasi sentimen, penelitian-penelitian tersebut telah menggunakan berbagai metode, seperti *Naïve Bayes*, Support Vector Machine (SVM), K-NN, dan Decision Tree, serta teknik text mining dan sentiment analysis.

Table 1 GAP Penelitian

No	Judul	Hasil	Perbedaan
1.	Analisis Sentimen Multi-Aspek Berbasis Konversi Ikon Emosi Dengan Algoritme Naïve Bayes Untuk Ulasan Wisata	Analisis sentimen untuk ulasan wisata kuliner menggunakan Naïve Bayes pada web Tripadvisor	Analisis sentimen Kota Tegal berbasis aspek dengan hasil klasifikasi data secara realtime

	Kuliner Pada Web Tripadvisor		
2.	Aspect Based Sentimen Analysis Opini Publik Pada Instagram dengan Convolutional Neural Network	Sentimen analisis opini publik terhadap operator telekomunikasi pada instagram dengan CNN yang bertujuan untuk meningkatkan kualitas layanan	Analisis sentimen Kota Tegal berbasis aspek pada twitter menggunakan pembobotan TF-IDF dan Naïve Bayes
3.	Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes	Analisis sentimen pada website Female Daily yang membahas tentang produk kecantikan	Analisis sentimen Kota Tegal berbasis aspek pada twitter dan di implementasikan ke dalam website secara realtime
4.	Aspect-Based Sentiment Analysis pada Ulasan E-Commerce dengan Metode Support Vector Machine untuk Mendapatkan Informasi Sentimen dari Beberapa Aspek	Sentimen analisis pada ulasan di E-commerce seperti platform Tokopedia dan Shopee dengan metode SVM	Analisis sentimen Kota Tegal berbasis Aspek di twitter menggunakan Naïve Bayes dan di implementasikan ke dalam website secara realtime
5.	Analisis Sentimen Penilaian Tempat Tujuan Wisata Kota Tegal Berbasis Text Mining	Pembuatan model sistem untuk memberikan sebuah informasi pada	Analisis sentimen Kota Tegal kedalam beberapa Aspek seperti Wisata/

		wisatawan dan pengelola tempat wisata.	hiburan, pendidikan, kuliner, dan fasilitas/ layanan publik
--	--	--	---

Hasil penelitian ini menunjukkan bahwa penggunaan kombinasi metode seperti Naïve Bayes dengan pelabelan berdasarkan aspek dapat menghasilkan akurasi yang lebih tinggi dalam klasifikasi sentimen. Teknik pengambilan bobot ulasan berdasarkan kontribusi pemberi ulasan juga memberikan kontribusi dalam meningkatkan hasil akurasi. Kesimpulannya, pada penelitian-penelitian sebelumnya dengan penelitian ini menunjukkan beragam pendekatan dan teknik yang dapat digunakan untuk menganalisis sentimen dari berbagai data ulasan atau tanggapan. Pemilihan metode dan pengaturan parameter yang tepat sangat berpengaruh terhadap akurasi hasil analisis sentimen. Meskipun beberapa penelitian mencapai hasil akurasi yang tinggi, ada ruang untuk peningkatan dan optimasi lebih lanjut dalam metode dan teknik yang digunakan.

## 1.4. Data Penelitian

### 1.4.1. Dataset

Pada penelitian ini, data yang digunakan berasal dari *Twitter* dan berupa kumpulan teks ulasan dari masyarakat yang membahas tentang Kota Tegal dalam bahasa Indonesia. Data yang diperoleh dari *Twitter* tersebut berkaitan dengan beberapa aspek Kota Tegal, yaitu Aspek Wisata/Hiburan, Aspek Fasilitas/Layanan Publik, dan Aspek Kuliner.

Untuk mengumpulkan data tersebut, beberapa parameter yang diperlukan adalah kata kunci "Kota Tegal" dan parameter lain seperti rentang tanggal. Data diambil menggunakan metode *crawling Twitter* dari tanggal 1 Januari 2022 hingga 31 Desember 2022. Dengan jumlah total seluruh data yang berhasil diambil dengan metode *crawling* adalah 3154 data selama tahun 2022. Setelah itu, data disimpan dalam *file .CSV* untuk digunakan pada tahapan berikutnya.

Pengumpulan data dilakukan dengan cara *crawling data twitter* dengan *Twint*, setelah data diambil dan disimpan kedalam *file .CSV*, maka selanjutnya data tersebut akan dilabeli secara manual dengan label angka 1 yaitu positif, angka 0 yaitu negatif, dan jika data tersebut tidak termasuk kedalam kategori maka akan dilabeli dengan angka -1. Dengan jumlah data pada aspek Wisata/Hiburan sebanyak 130 data positif dan 97 data negatif, pada aspek Pendidikan sebanyak 45 data positif dan 38 data negatif, pada aspek Fasilitas/Layanan Publik sebanyak 263 jumlah data positif dan 243 jumlah data negatif, dan pada aspek Kuliner terdapat 185 jumlah data positif dan 156

jumlah data negatif. Setelah data dilabeli dengan angka maka data tersebut siap digunakan untuk kebutuhan selanjutnya.

A	B	C	D	E	I
1	tweet				
2	Wisuda Universitas BSI Kampus Kabupaten Banyumas ke-16 & Kampus Kota Tegal ke-4 (06/12/2021) @Java Heritage Hotel Punwokerto <a href="https://t.co/MSS0lvvQgW">https://t.co/MSS0lvvQgW</a>	-1	1	-1	-1
3	(03/12/2021) Mengikuti kegiatan bersih pantai yang diselenggarakan Lanal Tegal di Pantai Alam Indah Kota Tegal. Kegiatan ini diikuti oleh berbagai instansi terkait di k	-1	-1	1	-1
4	@infopntegal Selamat malam , mohon di tindak lanjuti di daerah saya dari jam 17.30 WIB mati listrik sampai sekarang belum menyala Alamat di Jl. Dcripto mangunkusu	-1	-1	0	-1
5	[TF] Info banjir daerah tegal kota dong lurr, kayong pgn seblak kie, makasil	-1	-1	-1	1
6	Pemasangan portal di jalan masuk Alun-alun Kota Tegal diprotes pedagang. Wali Kota Tegal Deddy Yon Suriyono membeberkan alasan dari penutupan jalan tersebut. h	-1	-1	0	-1
7	Dalam rangka memperingati Hari Juang Kartika (HJK) TNI-AD tahun 2021 Kodim 0734/Kota Yogyakarta gelar kegiatan bhakti sosial donor darah bertempat di Jl. Tegal Gen	-1	-1	1	-1
8	Perusahaan di Kota Tegal wajib membayar upah sesuai Upah Minimum Rp 2.005.930,52 utk pekerja dengan masa kerja kurang dari 12 bulan; dan wajib menyusun Strukt	-1	-1	-1	-1
9	PPKM Level 3 Batal, Akhirnya Tegal Nyalakan Lagi Lampu-lampu Jalan di Kota #JawaTengah #KotaTegal <a href="https://t.co/kU3VzqvATP">https://t.co/kU3VzqvATP</a>	-1	-1	1	-1
10	Askrindo memberikan pembinaan dan bantuan Tanggung Jawab Sosial dan Lingkungan berupa sarana produksi kepada beberapa Kelompok seperti UMKM Seafood Olah	-1	-1	1	-1
11	Pembukaan Pelatihan Jabatan Fungsional Perawat Angkatan II (9 - 16 Desember 2021) dan Pelatihan Jabatan Fungsional Bidan (9 - 15 Desember 2021) BKPPD Kota Tegal	-1	1	-1	-1
12	Live Report Kasatlintas Polres Tegal Kota melaporkan pantauan Kamseltintarlintas dari Simpang Pramesthi Kota Tegal. Situasi Arus Lalu Lintas dari arah Timur maupun	-1	-1	1	-1
13	Aplikasi itu bertujuan memudahkan dan mendekatkan pelayanan PN Tegal dengan masyarakat, khususnya warga Kota Tegal. <a href="https://t.co/aTyuprYMCV">https://t.co/aTyuprYMCV</a>	-1	-1	1	-1
14	Waktunya kulineran sate kambing muda di Sate Tegal Marem lagi kuy! Satenya terbuat dari daging kambing muda pilihan yang didatangkan langsung dari Kota Tegal pa	-1	-1	-1	1
15	Ditentang Warga, Wali Kota Tegal Ngotot Tutup Jalan #protes #masyarakat #penutupan #jalan #tegal #alun #alun #protes <a href="https://t.co/UrOrYvTPR">https://t.co/UrOrYvTPR</a>	-1	-1	0	-1
16	Warga Kota Tegal Sekarang Bisa Mencetak Adminduk Mandiri <a href="https://t.co/2aQF86XjST">https://t.co/2aQF86XjST</a>	-1	-1	1	-1
17	Wali Kota Tegal Dedy Yon Supriyono mengungkapkan alasan kebijakan menutup kawasan Alun-alun dan Jalan Pancasila dengan portal. <a href="https://t.co/8tmIML3Z3a">https://t.co/8tmIML3Z3a</a>	-1	-1	1	-1
18	Sejak dipadamkan mulai 1 Desember 2021, lampu penerangan jalan umum (PJU) di Kota Tegal akhirnya kembali dinyalakan bertahap mulai Rabu (8/12/2021) malam. #pi	-1	-1	1	-1
19	Meski diprotes warga dan pedagang, pemortalan akses jalan menuju kawasan Alun-alun dan Jalan Pancasila Kota Tegal tetap diberlakukan dan tetap melarang kendara	-1	-1	0	-1
20	Jalan Dr. Soetomo Kota Tegal. Pemadaman Lampu Jalan oleh @PemkotTegal Mohon Pak @ganjarpranowo kasih solusi <a href="https://t.co/q5p6oRQpIM">https://t.co/q5p6oRQpIM</a>	-1	-1	0	-1

Gambar 1. 1 Dataset

#### 1.4.2. Data Hasil Implementasi Model

Berbeda dengan dataset sebelumnya yang digunakan untuk melatih model agar mendapatkan akurasi yang tinggi dalam mengklasifikasikan ulasan masyarakat Kota Tegal ke dalam beberapa aspek, data hasil implementasi model ini diperoleh dari proses *crawling* data *twitter* menggunakan *library*. Data yang masuk ke dalam *database* ini sebelumnya sudah diberikan label positif dan negatif sesuai dengan kategorinya secara otomatis. Jumlah data positif pada aspek Wisata/Hiburan sebanyak 558 dan jumlah data negatif sebanyak 6, untuk aspek Pendidikan sebanyak 261 Jumlah data positif, dan 27 Jumlah data negatif, untuk aspek Fasilitas/Layanan Publik sebanyak 1557 data positif, dan 27 data negatif, aspek Kuliner sebanyak 457 data positif, dan 4 data negatif. Data tersebut diambil secara realtime selama kurun waktu 1 Januari 2020 sampai 7 Agustus 2023, dan data tersebut akan terus bertambah.

Baru	tanggal	1	tweet	wisata_hiburan	pendidikan	fasilitas_layanan_publik	kuliner
dashboard_kotategal_db	2020-01-01		Ratusan hektar bawang merah di Kota Tegal terendam...	-1	-1	-1	-1
Baru	2020-01-01		Malam tahun baru menjadi kelabu, semalam di Kota T...	-1	-1	-1	-1
hasil_sentiment	2020-01-01		(Tegal, 01/01/2020) Izin melaporkan dari Terminal T...	-1	-1	-1	-1
information_schema	2020-01-02		Ini alun alun kota Tegal. Jadi bukan salah pak Ani...	1	-1	-1	1
klasifikasi_opini_publik	2020-01-02		@NyirRusmini2 @Faujiipn_ @yusuf_dumdum Lampu merah ...	-1	-1	1	-1
mysql	2020-01-02		@Faujiipn_ @NyirRusmini2 @yusuf_dumdum RSUD Kardina...	-1	-1	1	-1
performance_schema	2020-01-02		@ganjarpranowo @pembkabKEBUMEN @pembkab_klaten @Disk...	-1	-1	-1	-1
phpmyadmin	2020-01-02		Kantor Dinas Sosial (Dinsos) kota Tegal kebakaran...	-1	-1	-1	-1
test	2020-01-02		@ichi_fm Siang kak Ichi, disistem memang ditulisk...	-1	-1	-1	-1
	2020-01-02		Ratusan ekor ikan tambak yang berada di Jalan Mata...	-1	-1	0	-1
	2020-01-03		@infotegal Kayong enak es sagwan e mas Edi min. JL...	-1	-1	-1	1
	2020-01-03		#Pemberdayaan TEGAL Relawan inspirasi Rumah Zaka...	-1	-1	-1	-1
	2020-01-03		@CNNIndonesia "Pak bantu warga sendiri aja dulu no...	-1	-1	-1	-1
	2020-01-03		kunjungi markas PMI Kota Tegal yg terkena dampak p...	-1	-1	-1	-1
	2020-01-06		Pasang Ledeng Tanpa Uang Muka	-1	-1	-1	-1
	2020-01-06		Dalam rangka HUT ke...	-1	-1	-1	-1
	2020-01-06		Ditengah Gebyar HUT ke 44 PDAM Kota Tegal, Wakil W...	-1	-1	-1	-1
	2020-01-06		"Intinya kami dari nelayan siap bahwasanya Natuna ...	-1	-1	1	-1
	2020-01-06		#TerimaKasihFPI	-1	-1	-1	-1
	2020-01-07		Tegal Bukan Hanya Top Urusan Wart...	-1	-1	-1	-1
	2020-01-07		Muter Kab Pekalongan-Pemalang, kab Tegal, kota Teg...	-1	-1	-1	-1
	2020-01-07		@ganjarpranowo Kota Tegal, dari jam 15.00 - 21.30 ...	-1	-1	-1	-1
	2020-01-07		Pengumuman Gangguan Aliran Air	-1	-1	1	-1
	2020-01-07		Diumumkan kepada p...	-1	-1	1	-1
	Konsol						

Gambar 1. 2 Data Hasil Implementasi

## **BAB II**

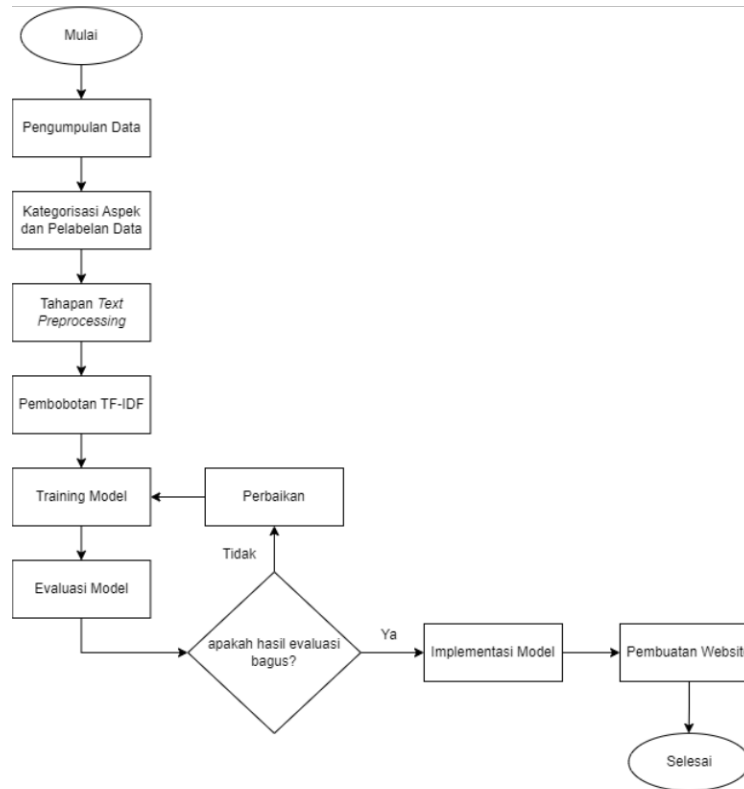
### **PRODUK**

#### **2.1. Perancangan**

Dalam pengembangan aplikasi ini, dibuat perancangan antarmuka pengguna yang responsif agar pengguna dapat dengan mudah menggunakan aplikasi. Aplikasi ini dapat menampilkan hasil klasifikasi sentimen masyarakat Kota Tegal yang sudah diklasifikasikan kedalam beberapa aspek secara otomatis. Aplikasi ini juga terdapat fitur tanggal yang dapat mempermudah pengguna untuk membaca data lebih rinci, dan juga fitur *WordCloud* untuk mengetahui trending topik yang sedang banyak diperbincangkan disetiap aspek.

#### **2.2. Perancangan Aplikasi**

Hasil dari penelitian ini menghasilkan suatu model dari *Naïve Bayes* dan *website* yang menampilkan dashboard klasifikasi ulasan masyarakat tentang Kota Tegal berbasis Aspek dari implementasi model yang telah dibuat sebelumnya berbasis *Website* secara *Realtime* dengan fitur-fitur tambahan seperti filter tanggal dan fitur *wordcloud*, serta fitur simpan data untuk memudahkan pengguna membaca data dan dapat mengelola data tersebut untuk kepentingan lain. Alur Perancangan Produk dapat dilihat pada Gambar 2.1.



Gambar 2. 1 Alur Pancangan Produk

**a. Pengumpulan Data**

Pada penelitian ini menggunakan data ulasan yang di dapat dari *twitter* menggunakan metode *Crawling*. Data ulasan yang digunakan adalah data yang membahas tentang Kota Tegal yang berbahasa Indonesia, yang nantinya akan di kategorikan kedalam beberapa aspek yaitu “Aspek Wisata/Hiburan”, “Aspek Fasilitas/Layanan Publik”, “Aspek Pendidikan”, dan “Aspek Kuliner”. Pengumpulan data ulasan masyarakat dari *twitter* dilakukan pada rentang waktu bulan Januari tahun 2022 sampai Bulan



Desember tahun 2022. Dengan total data dari proses *crawling* tersebut adalah sebanyak 3154 data.

Selanjutnya data yang sudah terkumpul akan dilabeli secara manual berdasarkan aspek yang ada. Dari setiap aspek terdapat beberapa kata yang menunjukkan bahwa ulasan tersebut membahas tentang aspek yang dituju. berdasarkan kata dijelaskan pada Tabel 2. 1.

Tabel 2. 1 Identifikasi Aspek

Aspek kategori	Kata
Wisata/Hiburan	Tempat, pemandangan, wisata
Fasilitas/layanan Publik	Akses, jalan, kota, fasilitas
Pendidikan	Sekolah, guru, siswa
Kuliner	Makanan, rasa, khas

Tahap selanjutnya yaitu pelabelan data dengan diklasifikasikan menjadi 2 kelas, yaitu kelas positif diberi nilai “1” dan kelas negatif diberi nilai “0”. Jika terdapat ulasan yang tidak memenuhi kategori dalam sebuah aspek maka akan diberi nilai “-1”. Contoh pelabelan dapat dilihat pada Tabel 2. 2.

Tabel 2. 2 Pelabelan Kelas

Ulasan	Wisata/Hiburan	Fasilitas/layanan Publik	Pendidikan	Kuliner
Menikmati pemandangan tegal dari atas, bisa jadi pengalaman dan sensasi tersendiri saat naik bianglala di rita park tegal. Kalau pas cuaca bagus bisa melihat gunung slamet di... bisa juga lihat lampu2 kota.	1	1	-1	-1

Berdasarkan identifikasi aspek yang telah dilakukan pada salah satu ulasan pada tabel 2 masuk kedalam 2 aspek, yaitu aspek wisata/hiburan karena didalam ulasan terdapat kata ‘wisata dan pemandangan’, dan aspek fasilitas/layanan publik karena didalam ulasan terdapat kata ‘lampu kota’. Untuk aspek Pendidikan dan kuliner bernilai “-1” dikarenakan ulasan tersebut tidak membahas tentang Pendidikan ataupun kuliner didalamnya, sehingga bisa dikategorikan sebagai *non-aspek*.

**b. *Text Preprocessing***

Pada tahap ini yaitu *Preprocessing Text* memiliki tujuan menghilangkan atau mengurangi data yang tidak perlu sehingga meningkatkan kualitas data dan algoritma dapat bekerja lebih efektif dan dapat memudahkan peneliti untuk menggunakan data di kemudian hari.

Berikut tahapan dari *preprocessing* antara lain :

1. *Casefolding*

*Casefolding* merupakan proses untuk mengubah semua huruf dalam text menjadi huruf kecil atau huruf besar. *Casefolding* yang dilakukan pada penelitian ini adalah mengubah semua huruf pada text menjadi huruf kecil. *Casefolding* dilakukan untuk menghindari kesalahan dalam pengolahan data karena banyaknya perbedaan besar kecilnya huruf.

## 2. *Filtering*

*Filtering* merupakan proses penyaringan data yang tidak relevan, penghapusan duplikat data yang muncul lebih dari satu kali dalam kumpulan data. *Filtering* dilakukan untuk memperoleh data yang lebih terstruktur dan relevan dengan penelitian yang sedang dilakukan

## 3. Tokenisasi

Tokenisasi merupakan proses untuk memecah text menjadi dokumen yang lebih kecil seperti kata, atau variable yang relevan. Pada proses ini juga dapat menghilangkan karakter yang tidak dibutuhkan seperti tanda baca, simbol, tagar, angka, mention.

## 4. *Stemming*

*Stemming* merupakan proses mengganti istilah tidak baku di dalam text menjadi bentuk yang lebih baku, dengan tujuan memudahkan pengolahan teks dan analisis yang dilakukan, serta mengurangi kemungkinan interpretasi atau ambiguitas pada text tertentu yang mengandung kata slang.

## 5. *Stopword*

*Stopword* merupakan proses pembuangan kata yang dianggap tidak penting. Kata tersebut sering muncul di dalam text namun dianggap tidak penting karna tidak memiliki makna dan tidak berpengaruh dalam proses analisis data.

**c. Pembobotan *TF-IDF***

Pretraining menggunakan metode *Term Frequency-Invers Document Frequency* (TF-IDF) menghasilkan representasi vektor untuk setiap kata. Proses pembobotan kata menggunakan *Term Frequency-Invers Document Frequency* (TF-IDF), pembobotan kata ini berfungsi untuk mengubah data berupa teks menjadi angka. *Machine learning* akan berjalan dengan baik dan optimal apabila ketika data yang diolah adalah data numerik atau angka. Oleh karena itu pada penelitian ini harus mengubah data berupa kumpulan kata menjadi data numerik. Ada banyak metode dalam proses pembobotan, termasuk *BoF (Bag of Words)*, *N-gram*, *Word2Vec*, dan TF-IDF [14]. TF-IDF adalah sebuah Teknik yang digunakan dalam pengolahan kata untuk memberikan bobot atau nilai pada kata didalam dokumen. Tujuan dari pembobotan tersebut adalah untuk mengidentifikasi kata yang paling penting atau paling banyak muncul dalam sebuah dokumen [15].

Pada metode ini ada dua faktor utama dalam pembobotan tersebut, *term frequency* (TF) adalah jumlah kemunculan *term* dalam suatu dokumen, semakin banyak jumlah kemunculan *term* didalam dokumen maka semakin besar bobot yang diberikan. Sedangkan *Invers Document Frequency* (IDF) digunakan untuk menghitung tingkat kepentingan sebuah *term* dalam seluruh kumpulan dokumen. IDF memberikan bobot yang lebih besar pada istilah atau *term* yang jarang muncul karena dianggap lebih penting daripada istilah yang sering muncul dalam dokumen. Semakin

sedikit jumlah dokumen yang mengandung *term* atau istilah tersebut, maka semakin besar nilai IDF yang diberikan [16]. Rumus TF dapat dituliskan dengan :

$$TF(t_a, d_b) = f(t_a, d_b)$$

$f(t_a, d_b)$  merupakan jumlah kemunculan *term* a pada dokumen b. semakin besar jumlah kemunculan *term* a pada sebuah dokumen b maka semakin besar nilai yang diberikan. Selanjutnya semakin banyak suatu *term* atau istilah muncul dibanyak dokumen maka nilai atau bobot yang diberikan akan semakin kecil [17]. IDF dapat dihitung dengan rumus sebagai berikut :

$$IDF = \log \frac{n}{df(t)}$$

TF-IDF dapat dirumuskan sebagai berikut :

$$TF-IDF(t_a, d_b) = TF(t_a, d_b) * IDF(t_a, d_b)$$

```

vektor_tfidf

array([[0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.],
       ...,
       [0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.]])

```

Gambar 2. 2 TF-IDF

#### d. Training Model

##### 1. Pembagian Data

Pada tahap ini, pembagian data dilakukan dengan membagi sebanyak 80% data *training* dan 20% data *testing*. Selanjutnya dilakukan proses klasifikasi pada data *training* menggunakan algoritma *Naïve Bayes*. Hasil klasifikasi dari data *training* tersebut diterapkan pada data *testing*. Setelah proses klasifikasi selesai, dilakukan evaluasi performa untuk mengukur kinerja model yang digunakan. Pada data *training* dan *testing* yang sudah di bagi menjadi 80:20 dimana data tersebut sudah diberi label sebelumnya. Pada model algoritma ini, data *training* digunakan sebagai panduan untuk membuat model klasifikasi, sementara data *testing* digunakan untuk menguji model tersebut dan melihat sejauh mana performa dan akurasi model dalam memprediksi label atau kelas dari contoh-contoh baru [18].

Pada pembagian rasio data 80:20 umum digunakan pada *machine learning* karena memiliki beberapa alasan yang penting, yaitu membantu mencegah *overfitting*, yaitu kondisi di mana model terlalu beradaptasi dengan data training dan kesulitan dalam menggeneralisasi pada data baru. Dengan memisahkan 80% data sebagai data *training* dan 20% sebagai data *testing*, model diuji pada data yang belum pernah dilihat sebelumnya, sehingga dapat menilai kemampuannya dalam melakukan prediksi yang lebih baik dan menghindari *overfitting*. Selain itu pembagian rasio data 80:20 juga bisa mendapatkan evaluasi yang akurat, Dengan menggunakan data *testing* yang terpisah, kita dapat mengukur seberapa baik model kita dalam memprediksi data baru. Rasio 80:20 memberikan ukuran yang cukup besar untuk data testing sehingga evaluasi model dapat dilakukan dengan lebih akurat [19]. Selanjutnya setiap kata akan dihitung jumlah probabilitas munculnya dalam suatu kelas tanpa mempertimbangkan fitur-fitur spesifik dari contoh. Estimasi probabilitas priori dilakukan dengan menghitung jumlah kemunculan setiap kelas dalam data latih. Probabilitas fitur adalah probabilitas kemunculan setiap fitur dalam setiap kelas. Pada pembuatan model ini menghasilkan 8 model dengan 4 model untuk klasifikasi aspek dan 4 model untuk klasifikasi sentimen pada aspek.

## 2. Model Aspek dan Model Sentimen Aspek

### A. Model Aspek

Pada model ini dataset dibagi menjadi data training dan data testing menggunakan fungsi `train_test_split` dari library `sklearn`. Kemudian, pada tahap ini dilakukan penyeimbangan data dengan *Random Oversampler* pada proses *training* agar data tersebut seimbang antara kelas minoritas dan kelas mayoritas. Tahap *oversampling* ini hanya dilakukan pada data *training* dan tidak dilakukan pada data *testing* karena untuk menghindari evaluasi model yang tidak realistis ketika diuji dengan data testing. Dengan melakukan *oversampling* pada data *training*, kita dapat meningkatkan kualitas dan keakuratan model pembelajaran mesin dalam mengenali dan memprediksi kelas minoritas. Setelah itu, model klasifikasi *MultinomialNB* dilatih pada data training yang telah di *oversampling* dan diuji menggunakan data training yang sama. Kemudian dilihat akurasi dan presisi pada model. Dalam melakukan *oversampling* pada data training, model klasifikasi dapat belajar dengan lebih baik untuk mengenali dan memprediksi kelas minoritas. Namun, penting untuk tetap melakukan evaluasi model pada data testing yang tidak diubah untuk mendapatkan hasil yang objektif dan dapat diandalkan



## B. Model Sentimen Aspek

Pada model data sentimen pada setiap aspek dan vektor TF-IDF dari dataset diambil dan dimasukkan ke dalam objek `READ_DATA`. Kemudian, data sentimen yang memiliki nilai -1 dihapus kemudian data training dan data testing dibuat dengan menggunakan fungsi *train\_test\_split* dari library *sklearn*. Selanjutnya, model klasifikasi *Multinomial Naive Bayes* dilatih pada data *training* dan diuji pada data *testing*. Kemudian dilihat hasil akurasi dan presisi pada model sentimen aspek yang ditampilkan pada tabel 2.5 dan tabel 2.6. Dalam melakukan klasifikasi sentimen pada setiap aspek, *oversampling* tidak digunakan. Namun, *oversampling* dapat digunakan untuk menangani ketidakseimbangan kelas pada dataset yang tidak seimbang. *Oversampling* dapat dilakukan dengan menggunakan teknik seperti *Random Oversampling* atau *Synthetic Minority Oversampling Technique (SMOTE)*.

### 3. *Naïve Bayes Classifier*

*Naive Bayes* adalah sebuah pendekatan yang berdasarkan pada teorema Bayes, yang menggabungkan pengetahuan yang didapat sebelumnya dengan informasi baru. Dengan demikian, algoritma ini merupakan salah satu metode klasifikasi yang simpel namun memiliki tingkat akurasi yang tinggi. Prinsip dasar di balik *Naive Bayes* adalah

mengklasifikasikan suatu data ke dalam kelas tertentu berdasarkan probabilitas yang terkait dengan kelas tersebut. Untuk melakukan ini, algoritma *Naive Bayes* menghitung probabilitas kelas berdasarkan data pelatihan yang diberikan. Secara khusus, ada dua tahap utama dalam penggunaan *Naive Bayes* untuk klasifikasi:

A. Tahap Pelatihan:

- **Menghitung Probabilitas Kelas**

Pada tahap ini, algoritma menghitung probabilitas setiap kelas berdasarkan jumlah data pelatihan yang termasuk dalam kelas tersebut. Misalnya, jika ada dua kelas A dan B, maka algoritma akan menghitung  $P(A)$  dan  $P(B)$  berdasarkan jumlah data yang berada di kelas A dan B.

- **Menghitung Probabilitas Fitur**

Algoritma juga menghitung probabilitas munculnya setiap fitur dalam kelas tertentu. Ini melibatkan menghitung berapa kali suatu fitur muncul dalam data pelatihan dari kelas tertentu dibagi dengan total data dalam kelas tersebut. Misalnya, untuk fitur  $X$  dalam kelas  $A$ , algoritma menghitung  $P(X|A)$ .

## B. Tahap Klasifikasi:

Setelah tahap pelatihan selesai, model *Naive Bayes* dapat digunakan untuk mengklasifikasikan data baru ke dalam salah satu kelas berdasarkan probabilitas yang telah dihitung sebelumnya. Ini melibatkan menghitung probabilitas kelas untuk data baru berdasarkan fitur-fiturnya dan memilih kelas dengan probabilitas tertinggi sebagai kelas prediksi untuk data tersebut.

Keuntungan dari algoritma *Naive Bayes* termasuk kemudahan implementasi, kemampuan bekerja dengan set data besar, dan performa yang baik dalam banyak kasus, terutama ketika asumsi independensi antar fitur tidak berdampak signifikan pada hasil klasifikasi. Namun, jika fitur-fitur dalam dataset sangat saling terkait, asumsi naif ini dapat menyebabkan performa model menurun. Oleh karena itu, perlu dilakukan analisis sebelumnya untuk memastikan bahwa asumsi ini sesuai dengan data yang akan diolah.

## 4. Pelatihan

Pada proses pelatihan model ini, yang pertama dilakukan sebelum proses *training* adalah pelabelan pada dataset dan juga harus melalui proses *preprocessing* untuk dijadikan input dari penggunaan metode *Naive Bayes* yang nantinya akan disimpan sebagai model data dengan

ekstensi .pickle. Pada proses pelatihan model sistem akan membaca data *tweet* yang akan dimasukan selanjutnya sistem akan mengklasifikasikan kedalam kelas atau aspek yang ada. Jika kata yang dimasukan tidak memenuhi kriteria aspek tertentu maka sistem akan melanjutkan proses klasifikasi pada aspek atau kategori yang lain, jika kata yang dimasukan terdeteksi bukan aspek maka akan dihapus secara otomatis oleh sistem. Setelah itu sistem akan melakukan ekstraksi fitur dari kata dan menyimpan data model tersebut. Selanjutnya, data model akan disimpan untuk digunakan pada tahap selanjutnya.

**e. Evaluasi Model**

Tahap terakhir pada penelitian ini adalah evaluasi, apakah algoritma yang digunakan dalam penelitian ini merupakan model klasifikasi yang baik sehingga mendapatkan nilai akurasi yang tinggi pada analisis sentimen ulasan terhadap Kota Tegal. Pada tahap evaluasi dilakukan menggunakan *confusion matrix*, *confusion matrix* merupakan alat untuk mengukur kualitas dari proses klasifikasi dokumen terhadap satu kelas atau lebih. Contoh *confusion matrix* 2x2 dapat dilihat pada Gambar 2.4.

		Actual Class	
		Yes	No
Predited Class	Yes	TP	FN
	No	FP	TN

Gambar 2. 3 Confusion Matrix

Terdapat 4 istilah nilai klasifikasi pada *confusion matrix* yang nantinya akan digunakan untuk menghitung nilai akurasi, presisi, recall, dan *f1-score*, nilai tersebut yaitu *True Positive* (TP), *False Negative* (FN), *False Positive* (FP), dan *True Negative* (TN). *True Positive* (TP) dan *True Negative* (TN) mengindikasikan hasil klasifikasi yang benar. *True Positive* (TP) adalah ketika hasil klasifikasi secara tepat memprediksi kepositifan suatu kondisi yang memang benar positif. *True Negative* (TN) adalah ketika hasil klasifikasi secara tepat memprediksi kenegatifan suatu kondisi yang memang benar negatif. *False Positive* (FP) terjadi ketika hasil prediksi secara keliru memprediksi kepositifan suatu kondisi yang sebenarnya negatif. Sedangkan *False Negative* (FN) terjadi ketika hasil prediksi secara keliru memprediksi kenegatifan suatu kondisi yang sebenarnya positif.

Dengan menggunakan nilai dalam confusion matrix tersebut, kita dapat menghitung akurasi, presisi, recall, dan f1-score menggunakan persamaan berikut.

$$\text{akurasi} = \frac{TN + TP}{(TN + FN + TP + FP)}$$

$$\text{precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}$$

$$\text{recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}$$

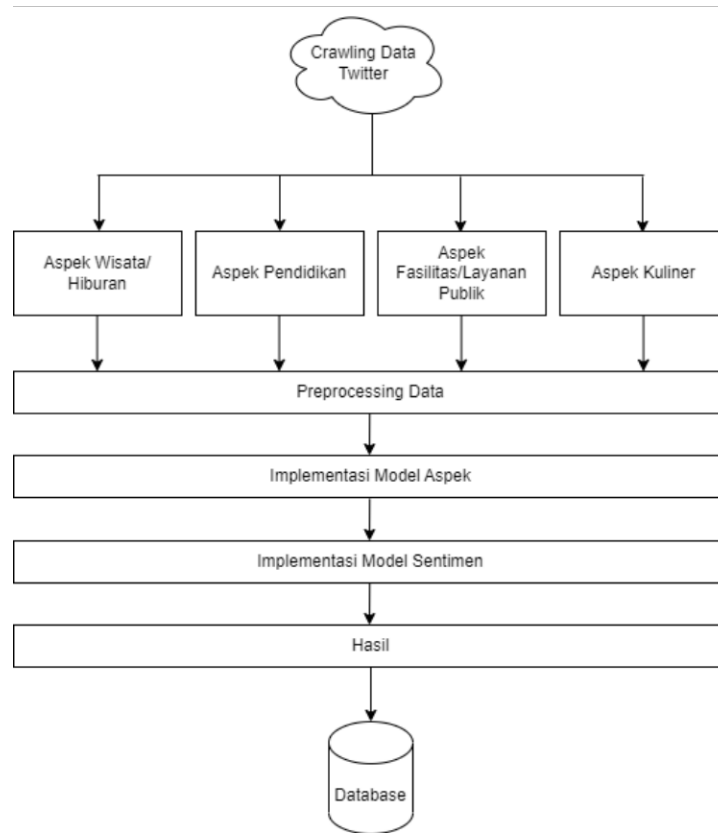
$$f1 - \text{score} = 2 \times \frac{(\text{precision} * \text{recall})}{(\text{precision} + \text{recall})}$$

Akurasi adalah proporsi persentase dari sentimen yang teridentifikasi dengan benar dari total keseluruhan. Untuk menghitung akurasi, kita membagi jumlah sentimen yang teridentifikasi dengan benar dengan total jumlah data yang diuji. *Precision* adalah rasio antara jumlah data yang relevan ditemukan dibandingkan dengan jumlah data yang ditemukan secara keseluruhan. Untuk menghitung *precision*, kita membagi jumlah data yang benar positif dengan jumlah data yang benar positif dan data yang salah positif. Jumlah data yang salah positif diperoleh dari total nilai positif yang tidak benar positif dalam setiap kelas yang sesuai. *Recall*

adalah rasio antara jumlah materi yang relevan ditemukan dibandingkan dengan jumlah total materi yang sebenarnya relevan. Nilai *f1-score* diperoleh dengan mengalikan hasil perkalian *precision* dan *recall*, kemudian membaginya dengan hasil penjumlahan *precision* dan *recall*, dan terakhir mengalikannya dengan dua [20].

**f. Implementasi Model**

Pada tahap implementasi model ini bertujuan untuk menggunakan model yang sudah dibuat sebelumnya. Untuk data baru yang akan melalui proses *crawling twitter* akan dilakukan klasifikasi aspek dan setelah itu akan dilakukan klasifikasi sentimen dan memasukan hasil klasifikasi tersebut kedalam *database*. Gambar 2.5 adalah alur dari implementasi model.



Gambar 2. 4 Alur Implementasi Model

#### 1. Proses *Crawling* Data Pada *Twitter*

Pada proses *crawling* data *twitter* ini metode yang digunakan untuk melakukan *crawling* adalah dengan menggunakan *Tweet Harvest*. Metode ini memungkinkan pengguna untuk mengumpulkan *tweet* berdasarkan kata kunci dan tanggal tertentu dari hasil pencarian *Twitter* [21]. Ada beberapa parameter yang digunakan dalam proses *crawling* ini, yaitu penggunaan kata kunci “Kota Tegal”, dan parameter lain berupa tanggal. Dari jumlah data baru yang berhasil dikumpulkan dengan proses *crawling* secara *realtime* sebanyak 5004 data.



## 2. Preprocessing Data

Setelah proses *crawling* data dari *twitter* menggunakan *Library Tweet Harvest* selesai, proses selanjutnya adalah tahapan preprocessing dari data yang sudah di *crawling* sama seperti tahapan 2.2.2. *Text Preprocessing*.

## 3. Implementasi Model Aspek

Setelah tahapan *preprocessing* data selesai, selanjutnya adalah tahapan implementasi model aspek yang sudah dibuat sebelumnya dengan menggunakan data baru yang telah melewati proses *crawling* dan melewati proses *text preprocessing*. Data tweet akan diuji dengan model aspek pertama, jika data tersebut merupakan aspek tersebut maka data akan diuji lagi dengan model sentimen. Tetapi, jika data tersebut merupakan bukan aspek maka data tersebut akan diuji dengan model aspek yang kedua. Jika data tersebut merupakan aspek yang kedua maka data tersebut akan diuji lagi dengan model sentimen. Jika data tersebut bukan aspek maka data tersebut akan diuji dengan model aspek yang ketiga, begitupun untuk model aspek yang keempat. Data tweet tersebut memiliki kemungkinan untuk memiliki beberapa aspek ketika diuji dengan model aspek.

## 4. Implementasi Model Sentimen

Setelah melewati implementasi model aspek, data yang merupakan aspek akan diuji lagi dengan model sentimen. Data tersebut

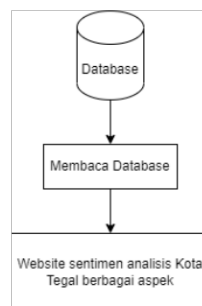
akan dilabeli dengan label angka 1 yaitu sentimen positif dan label angka 0 yaitu sentimen negatif. Setelah diuji dengan model aspek dan model sentimen yang sudah dibuat sebelumnya, selanjutnya data tersebut akan dimasukkan kedalam database.

#### 5. Memasukkan Hasil Implementasi Model Kedalam Database

Tahap selanjutnya adalah memasukan hasil implementasi model kedalam *database*. Dimulai dari proses *crawling* data, *preprocessing* data, implementasi model, dan memasukkan hasil implementasi model tersebut kedalam *database* dilakukan secara *realtime* dan otomatis menggunakan *Task Scheduler* setiap pukul 00.30 WIB.

#### g. Pembuatan Website

Tahap selanjutnya setelah melakukan implementasi model adalah proses pembuatan *website* menggunakan data hasil implementasi model sebelumnya yang sudah tersimpan didalam *database* dan akan ditampilkan kedalam *website* menggunakan *Streamlit*. Berikut merupakan alur dalam pembuatan *website* pada Gambar 2.6



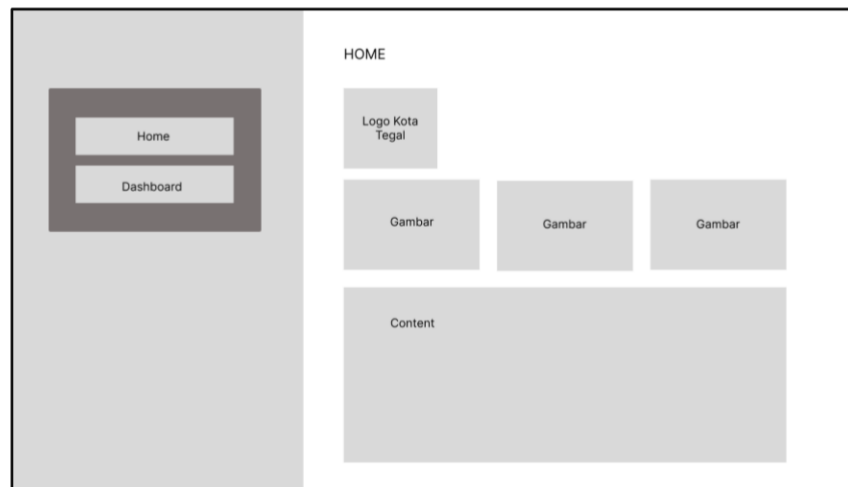
Gambar 2. 5 Alur Pembuatan *Website*

### 2.2.1. Desain Antarmuka

Desain antarmuka merupakan bentuk dasar dari sebuah rancangan aplikasi ketika membuat aplikasi pertama kali. Desain antarmuka ini bertujuan untuk memudahkan *programmer* dalam memahami *flow* aplikasi yang dibuat. Berikut merupakan desain antarmuka aplikasi :

#### 1. Menu Home

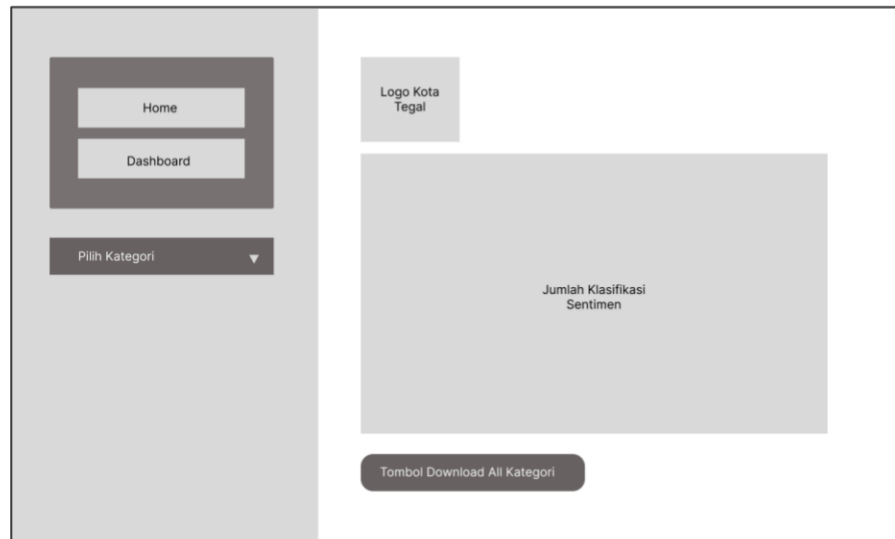
Menu Home merupakan tampilan awal ketika pengguna memasuki aplikasi. Terdapat beberapa komponen seperti Logo Kota Tegal dan Judul aplikasi. Dibagian samping kiri terdapat tampilan tombol untuk Menu *Home* dan Menu *Dashboard*. Terdapat juga keterangan singkat tentang aplikasi ini. Pada gambar 2.7 merupakan tampilan desain antarmuka untuk Menu Home.



Gambar 2. 6 Menu Home

## 2. Menu Dashboard

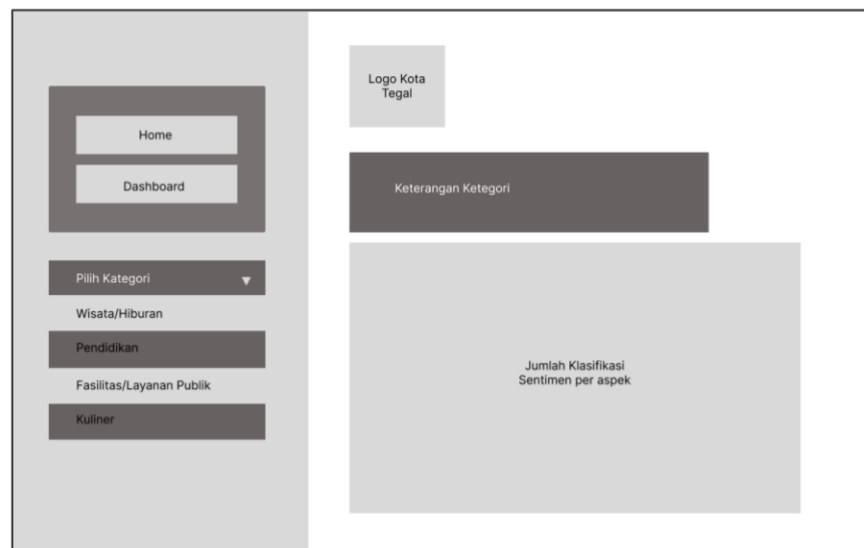
Pada menu *dahsboard* ini terdapat tombol untuk memilih kategori aspek yang diinginkan, selanjutnya terdapat jumlah klasifikasi sentimen pada aspek wisata/hiburan, pendidikan, fasilitas/layanan publik, dan kuliner berupa sentimen positif dan sentimen negatif dari masyarakat Kota Tegal dalam bentuk diagram seperti pada Gambar 2.8 dibawah ini. Pada halaman dashboard ini juga terdapat fitur berupa tombol untuk *download* data untuk semua kategori.



Gambar 2. 7 Menu Dashboard

### 3. Fitur Pilih Kategori

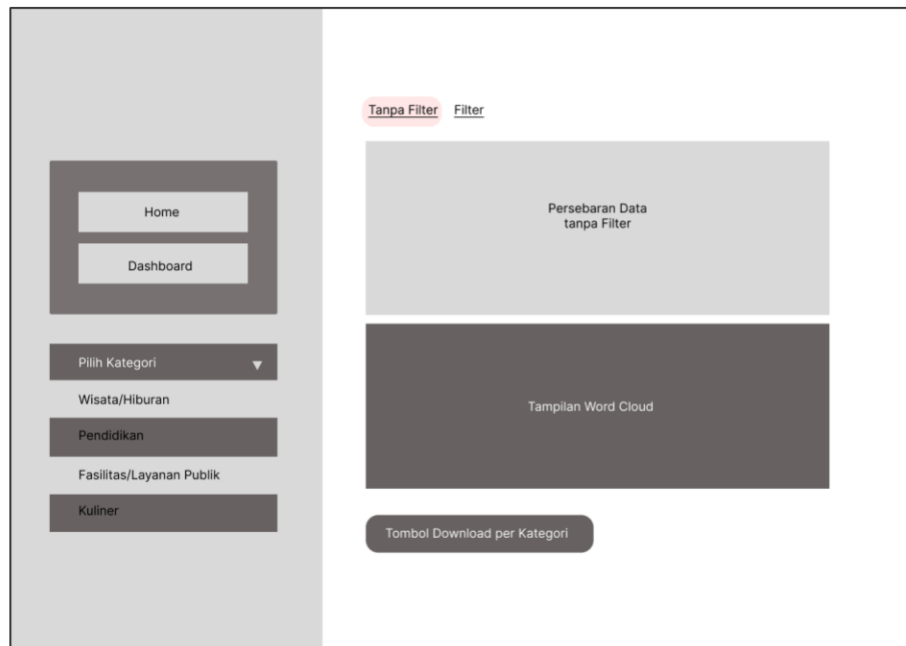
Fitur ini merupakan fitur untuk pengguna dapat memilih kategori yang akan dilihat sesuai dengan aspek yang dipilih. Sesuai dengan Gambar 2.8 dibawah ini terdapat tombol untuk memilih kategori mana yang akan ditampilkan, dan juga penjelasan singkat tentang masing-masing kategori aspek. Pada halaman ini juga terdapat hasil sentimen positif dan negatif dari setiap aspek yang ditampilkan dalam sebuah diagram yang interaktif.



Gambar 2. 8 Pilih Kategori

#### 4. Tanpa Filter

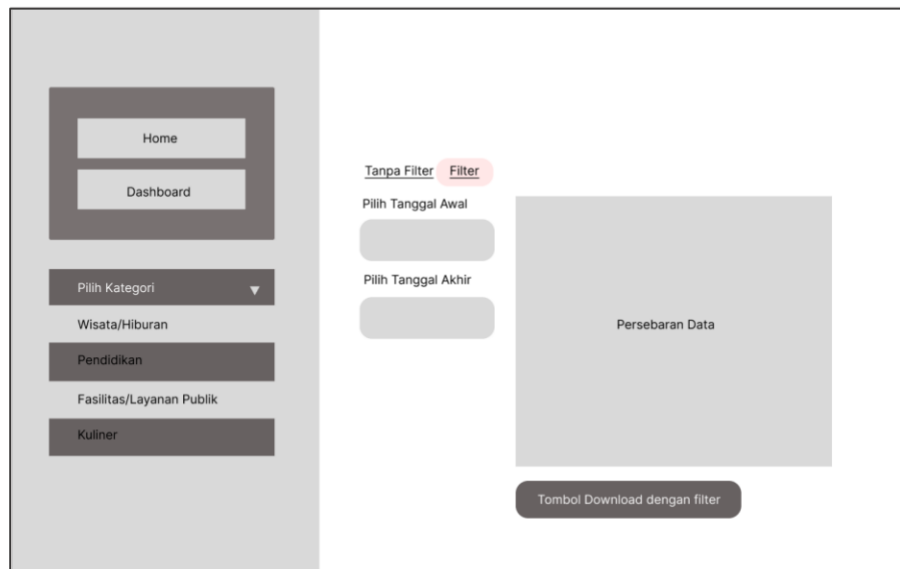
Pada halaman ini merupakan halaman persebaran data tanpa filter tanggal. Pada halaman ini terdapat persebaran data secara lebih rinci dengan diagram plot, pada halaman ini juga terdapat fitur berupa tombol untuk download data untuk setiap kategori aspek. Seperti pada gambar 2.9 ini. Dan terdapat fitur *wordcloud* di halaman ini



Gambar 2. 9 Tanpa Filter

## 5. Fitur Tanggal

Fitur tanggal ini merupakan fitur yang dapat membantu pengguna membaca data agar lebih rinci sesuai dengan rentang waktu yang diinginkan. Pengguna dapat memasukkan tanggal awal dan tanggal akhir maka akan muncul persebaran data sesuai dengan rentang waktu yang inputkan. Pada fitur ini juga dilengkapi dengan fitur untuk download data ketika rentang waktu tanggal diinputkan.

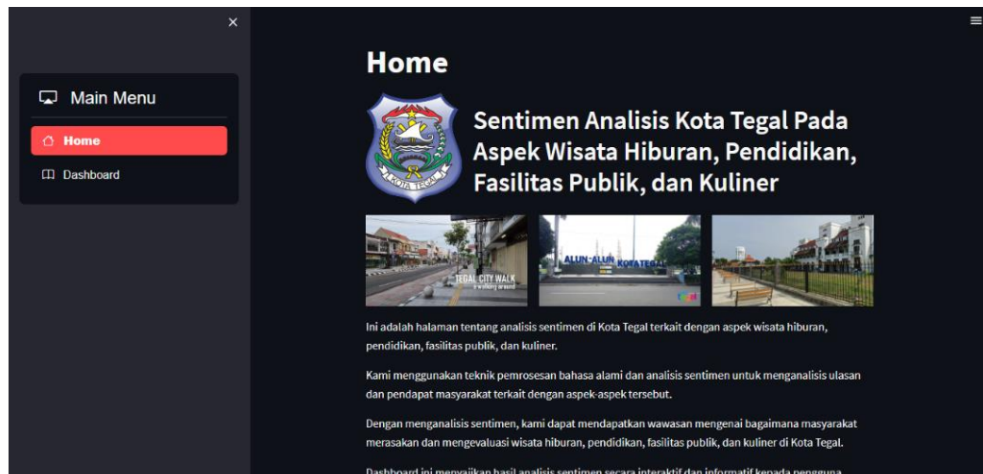


Gambar 2. 10 Fitur Tanggal

### 2.2.2. Implementasi Sistem

#### 1. Menu *Home*

Aplikasi ini akan menampilkan *dashboard* visualisasi hasil klasifikasi sentimen analisis Kota Tegal dalam beberapa aspek yang diambil dari data yang ada didalam *database*. Pada *website* ini terdapat dua menu yaitu menu “Home” dan menu “Dashboard”, seperti pada Gambar 2.12 merupakan tampilan dari menu *Home*.

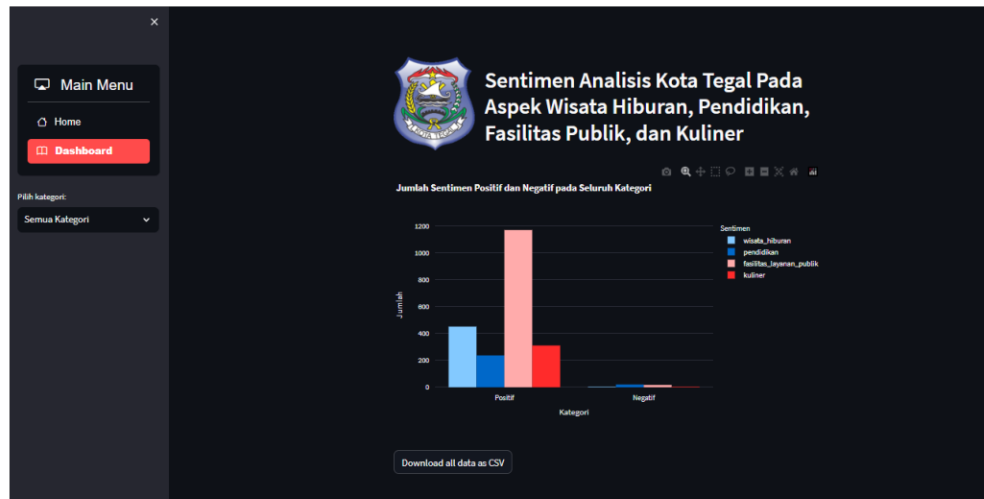


Gambar 2. 11 Implementasi Menu Home

#### 2. Menu Dashboard

Dan pada Gambar 2.12 merupakan tampilan dari menu *Dashboard* yang dapat menampilkan hasil klasifikasi dari ulasan masyarakat Kota Tegal yang sudah dikalsifikasikan kedalam beberapa aspek.





Gambar 2. 12 Implementasi Menu *Dashboard*

### 3. Fitur Pilih Kategori

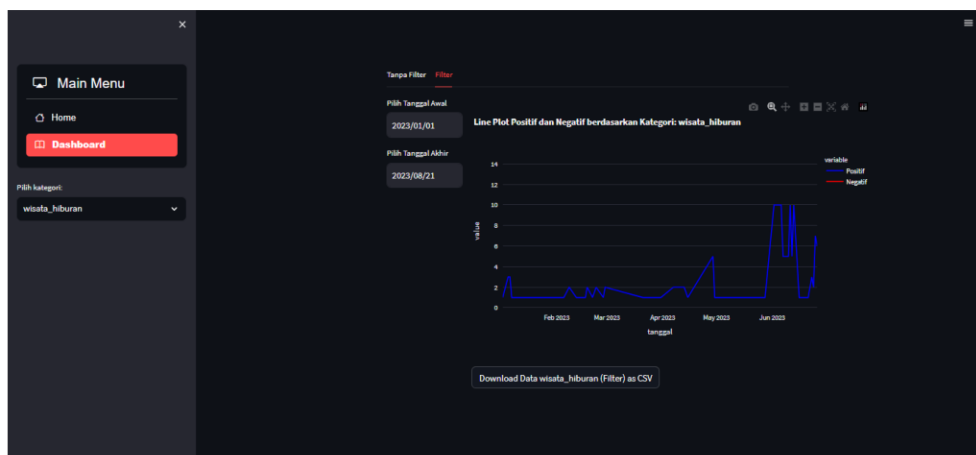
Pada halaman *dashboard* terdapat *button* untuk memilih kategori/aspek yang dapat memunculkan hasil kalsifikasi data. Pada aplikasi ini terdapat tombol untuk aspek wisata/hiburan, pendidikan, fasilitas/layanan publik, dan juga kuliner.



Gambar 2. 13 Implementasi Pilih Fitur



diinginkan seperti pada Gambar 2.16 , pengguna cukup memasukan tanggal awal dan tanggal akhir maka akan muncul visualisasi data dengan rentang waktu yang sudah diinputkan. Fitur tanggal ini berguna untuk memudahkan pengguna membaca data lebih rinci. Pada halaman ini juga terdapat fitur untuk download data perkategori dengan filter tanggal.



Gambar 2. 16 Implementasi Fitur Tanggal

## 2.3. Pengujian

### 2.3.1. Pengujian Model

Pada tahap ini merupakan tahap pengujian model dengan data *training* dan data *testing*. Data training digunakan untuk melatih model machine learning, sedangkan data testing digunakan untuk mengevaluasi kinerja model pada data yang belum pernah dilihat sebelumnya. Untuk melakukan pengujian, kita dapat membandingkan hasil prediksi model dengan nilai sebenarnya pada data *testing* menggunakan *Convolution Matrix* yang sesuai, seperti *Accuracy*, *Precision*, *Recall*, atau *F1-score*. Berikut tabel untuk pengujian model dengan data training pada setiap aspek.

Tabel 2. 3 Pengujian Model Aspek Training

Aspek/Kategori	<i>Accuracy</i>
Wisata/Hiburan	96%
Pendidikan	99%
Fasilitas/Layanan Publik	94%
Kuliner	95%

Tabel 2. 4 Pengujian Model Aspek Testing

Aspek/Kategori	<i>Accuracy</i>
Wisata/Hiburan	86%
Pendidikan	95%
Fasilitas/Layanan Publik	74%
Kuliner	79%

Pengujian selanjutnya yaitu pengujian pada model Sentimen. Dengan hasil akurasi dari pengujian data *training* pada tabel 2. 5. Dan hasil pengujian pada data *testing* pada tabel 2.6.

Tabel 2. 5 Hasil pengujian model sentimen aspek training

Aspek/Kategori	<i>Accuracy</i>
Wisata/Hiburan	99%
Pendidikan	96%
Fasilitas/Layanan Publik	100%
Kuliner	99%

Tabel 2. 6 Hasil pengujian model sentimen aspek testing

Aspek/Kategori	<i>Accuracy</i>
Wisata/Hiburan	91%
Pendidikan	86%
Fasilitas/Layanan Publik	82%
Kuliner	98%

Nilai akurasi pada pengujian data testing mengalami penurunan, hal tersebut disebabkan oleh beberapa faktor salah satunya yaitu overfitting dan kurangnya variasi pada data training, hal ini menyebabkan nilai akurasi pada data testing lebih kecil, namun untuk model tersebut nilai akurasi masih bagus dengan rata-rata diatas 70% untuk pengujian dengan data testing, dan rata-rata diatas 90% untuk pengujian dengan data testing.

Berikut pada tabel 2.7 dan pada tabel 2.8 merupakan hasil akurasi, *Precision*, *Recall*, dan *F1-score*.

Tabel 2. 7 Hasil Pengujian Klasifikasi Aspek

<b>Kategori</b>	<b><i>Precision</i> (%)</b>	<b><i>Recall</i> (%)</b>	<b><i>f1-score</i> (%)</b>
Wisata/Hiburan	94	86	89
Fasilitas/layanan	85	74	77
Pendidikan	97	95	96
Kuliner	93	79	83
Rata-Rata	92,25	83,5	86,25

Tabel 2. 8 Hasil Pengujian Sentimen

<b>Kategori</b>	<b><i>Precision</i> (%)</b>	<b><i>Recall</i> (%)</b>	<b><i>f1-score</i> (%)</b>
Wisata/Hiburan	98	98	98
Fasilitas/layanan	87	86	86
Pendidikan	89	82	83
Kuliner	91	91	91
Rata-Rata	91,25	89,25	89,5

Dari tabel dapat disimpulkan bahwa aspek pendidikan mencapai kinerja tertinggi, dengan rata-rata *f1-score* sebesar 86,25%, melebihi kinerja aspek lainnya. Hal ini dapat terjadi karena distribusi data pada aspek pendidikan lebih merata, dan proses identifikasi kata yang digunakan untuk melabeli aspek pendidikan lebih spesifik daripada ketiga aspek lainnya. Aspek fasilitas/layanan publik menjadi kelas mayor karena banyak ulasan yang tidak membahas secara khusus tentang fasilitas/layanan publik, sehingga ulasan tersebut diberi label netral. Hal ini menyebabkan ketidakseimbangan data, yang mengakibatkan model tidak mampu melakukan prediksi dengan akurat. Walaupun distribusi data pada aspek wisata/hiburan dan kuliner hampir

seimbang, namun performanya masih kurang memuaskan. Faktor yang menyebabkan hal tersebut adalah ketidak konsistenan dalam proses pelabelan, di mana identifikasi kata untuk aspek wisata/hiburan dan kuliner kurang spesifik karena ulasan mengenai produk juga diberi label berdasarkan opini secara umum.

Terdapat perbedaan yang signifikan antara nilai akurasi, *precision*, *recall*, dan *f1-score* pada setiap aspek. Perbedaan tersebut muncul karena akurasi hanya mempertimbangkan hasil prediksi yang benar (*True Positive* dan *True Negative*). Selain itu, nilai *recall* lebih tinggi daripada *precision*, yang menunjukkan bahwa model lebih baik dalam mengidentifikasi data positif daripada keseluruhan data yang sebenarnya positif, serta lebih berhati-hati dalam menghindari kesalahan *False Positive*. Karena ketidakseimbangan data, nilai *precision* rendah menunjukkan bahwa model sulit dalam memprediksi data yang sebenarnya positif. Sebagai akibatnya, model cenderung memprediksi kelas yang memiliki jumlah data yang lebih banyak. Karena nilai *precision* dan *recall* yang rendah, hal ini mengakibatkan nilai *f1-score* yang rendah.

### **2.3.2. Pengujian Implementasi Model**

Tahap selanjutnya adalah implementasi model yang telah dibuat selanjutnya, dengan langkah awal yaitu *crawling* data secara *realtime* dengan *twitter* kemudian data tersebut melalui proses *preprocessing* dan kemudian di *testing* menggunakan model yang telah dibuat sebelumnya, setelah itu hasilnya akan dimasukkan kedalam database.

Gambar 2.18 adalah tampilan dari data hasil yang sudah melalui tahap implementasi model yang telah dimasukkan kedalam database.

tanggal	tweet	wisata_hiburan	pendidikan	fasilitas_layanan_publik	kuliner
2020-01-01	Ratusan hektar bawang merah di Kota Tegal terendam...	-1	-1	-1	-1
2020-01-01	Malam tahun baru menjadi kelabu, semalam di Kota T...	-1	-1	-1	-1
2020-01-01	(Tegal,01/01/2020) Izin melaporkan dari Terminal T...	-1	-1	-1	-1
2020-01-02	Ini alun alun kota Tegal. Jadi bukan salah pak Ani...	1	-1	-1	1
2020-01-02	@NyiRusmini2 @Faujijpn_ @yusuf_dumdum Lampu merah ...	-1	-1	1	-1
2020-01-02	@Faujijpn_ @NyiRusmini2 @yusuf_dumdum RSUD Kardina...	-1	-1	1	-1
2020-01-02	@ganjarpranowo @pembkabKEBUMEN @pembkab_klaten @Disk...	-1	-1	-1	-1
2020-01-02	Kantor Dinas Sosial (Dinsos) kota Tegal kebanjiran...	-1	-1	-1	-1
2020-01-02	@ichi_fm Siang kak Ichi, disistem memang ditulisk...	-1	-1	-1	-1
2020-01-02	Ratusan ekor ikan tambak yang berada di Jalan Mata...	-1	-1	0	-1
2020-01-03	@infotegal Kayong enak es sagwan e mas Edi min. Jl...	-1	-1	-1	1
2020-01-03	#Pemberdayaan TEGAL. Relawan inspirasi Rumah Zaka...	-1	-1	-1	-1
2020-01-03	@CNNIndonesia "Pak bantu warga sendiri aja dulu no...	-1	-1	-1	-1
2020-01-03	kunjungi markas PMI Kota Tegal yg terkena dampak p...	-1	-1	-1	-1

Gambar 2. 17 Implementasi Model

### 2.3.3. Pengujian Aplikasi





Sebelum meluncurkan aplikasi, penting untuk melakukan pengujian menyeluruh guna memastikan aplikasi berfungsi dengan baik, bebas dari bug atau kesalahan, dan memberikan pengalaman pengguna yang memuaskan. Proses pengujian aplikasi ini mencakup berbagai jenis metode, seperti pengujian *Black box* yang bertujuan mengevaluasi kinerja dan fungsionalitas aplikasi, serta pengujian *Usability* yang fokus pada pengalaman pengguna saat menggunakan aplikasi. Dalam pelaksanaan pengujian *Usability*, beberapa hal penting harus diperhatikan, termasuk merancang pertanyaan yang relevan dan mudah dimengerti oleh peserta, serta melibatkan beragam pengguna untuk mengevaluasi interaksi mereka dengan aplikasi. Hasil dari pengujian *Black box* dan *Usability* ini sangat berharga bagi pengembangan aplikasi karena membantu meningkatkan kualitas aplikasi sebelum dipublikasikan.



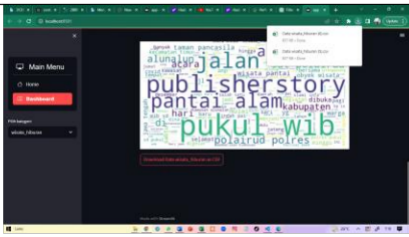





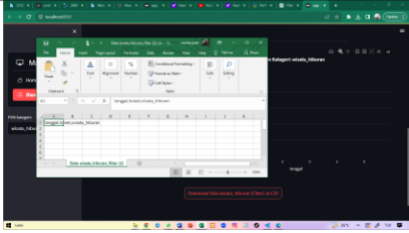
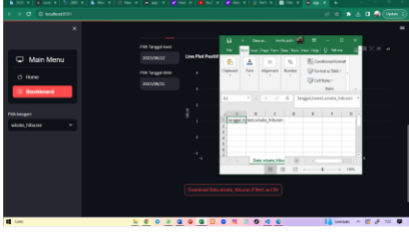
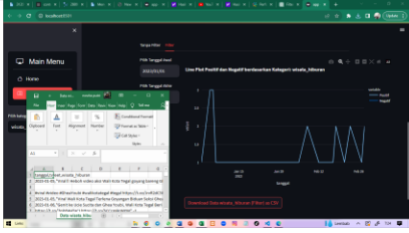
### A. Pengujian Aplikasi Menggunakan Metode *Black Box*


Pengujian dilakukan dengan menguji fungsionalitas aplikasi tanpa memeriksa atau mengetahui secara rinci struktur atau mekanisme kerja aplikasi tersebut. Berikut adalah hasil dari pengujian aplikasi menggunakan metode black box.

Tabel 2. 9 Testing Black Box

No	Object yang diuji	Skenario Pengujian	Hasil yang diharapkan	Hasil	Status
1.	Halaman Home	Klik menu home	Menampilkan tampilan home		Valid
2.	Halaman Dashboard	Klik menu dashboard	Menampilkan tampilan dashboard		Valid
3.	Download Data	Klik tombol untuk download data	Pengguna dapat mengunduh seluruh data dari aplikasi		Valid
4.	Fitur Pilih Kategori	Klik menu pilih kategori	Menampilkan pilihan beberapa kategori		Valid

5.	Fitur Pilih Kategori	Klik kategori wisata/hiburan	Menampilkan jumlah klasifikasi sentimen		Valid
6.	Fitur Tanpa Filter	Klik tombol tanpa filter	Menampilkan persebaran data tanpa filter		Valid
7.	Download Data Kategori	Klik tombol download data untuk setiap kategori	Pengguna dapat mendownload data sesuai kategori yang di pilih		Valid
8.	Filter Tanggal	Klik tombol filter	Menampilkan persebaran data dengan filter tanggal		Valid
9.	Filter Tanggal	Memasukkan awal dan tanggal akhir	Menampilkan persebaran data tanggal yang diinputkan		Valid
10	Filter Tanggal	Memasukkan tanggal sesudah	Aplikasi tidak akan menampilkan data apapun		Valid

11	Download Data per Tanggal	Klik tombol download data tanpa memasukk an tanggal awal dan tanggal akhir	Pengguna dapat mendownload tanggal dengan value data kosong		Valid
12	Download Data per Tanggal	Klik tombol download dengan memasukk an tanggal sesudah	Pengguna dapat mendownload tanggal dengan value data kosong		Valid
13	Download Data Per Tanggal	Klik tombol download data dengan memasukk an tanggal awal dan tanggal akhir	Pengguna dapat mendownload data sesuai dengan tanggal yang sudah dimasukkan		Valid

14	Fitur Wordcloud	Gerakan kursor dibawah pada setiap halaman kategori	Menampilkan tampilan wordcloud sesuai kategori yang dipilih		Valid
----	-----------------	---	---	--	-------

## B. Pengujian Aplikasi Menggunakan *Usability*

*Usability testing* adalah metode evaluasi yang digunakan untuk mengukur sejauh mana kemudahan dan efektivitas pengguna dalam berinteraksi dengan suatu produk, aplikasi, atau layanan. Hasil dari pengujian ini membantu mengidentifikasi masalah *usability*, seperti antarmuka yang membingungkan atau fungsi yang sulit digunakan, yang nantinya dapat diperbaiki untuk meningkatkan pengalaman pengguna secara keseluruhan.

Tabel variabel penilaian aplikasi menggunakan usability terdapat pada tabel 2.10

Tabel 2. 10 Tabel Variabel Penilaian

No	Variabel Penilaian
1.	Manfaat atau kegunaan Aplikasi
2.	Kemudahan bagi pengguna
3.	Tata Letak Fitur
4.	Interface / Tampilan Website
5.	Penggunaan Bahasa pada Aplikasi

Tabel 2. 11 Tabel Pernyataan Positif

No	Pernyataan Positif
1.	Saya merasa terbantu dengan adanya aplikasi ini karena memudahkan saya dalam mendapatkan informasi sentimen tentang tweet dari masyarakat terhadap Aspek Wisata/Hiburan, Kuliner, Pendidikan, dan Fasilitas/Layanan Publik di Kota Tegal.
2.	Saya tidak mengalami kesulitan dalam menemukan informasi yang saya cari berupa jumlah data, persebaran data, dan juga fitur-fitur lain
3.	Tata letak dan desain halaman ini membuat saya merasa nyaman dan mudah dipahami
4.	Tampilan dan tata letak pada aplikasi ini menarik
5.	Penggunaan bahasa pada aplikasi ini mudah dipahami

Tabel 2. 12 Tabel Pernyataan Negatif

No	Pernyataan Negatif
1.	Saya tidak merasa terbantu dengan adanya aplikasi ini karena tidak memudahkan saya dalam mendapatkan informasi sentimen tentang tweet dari masyarakat terhadap Aspek Wisata/Hiburan, Kuliner, Pendidikan, dan Fasilitas/Layanan Publik di Kota Tegal.
2.	Saya mengalami kesulitan dalam menemukan informasi yang saya cari berupa jumlah data, persebaran data, dan juga fitur-fitur lain
3.	Tata letak dan desain halaman ini membuat saya merasa tidak nyaman dan tidak mudah dipahami
4.	Tampilan dan tata letak pada aplikasi ini tidak menarik
5.	Penggunaan bahasa pada aplikasi ini tidak mudah dipahami

Data yang diperoleh dari penilaian responden terhadap aplikasi Sentimen Analisis Kota Tegal ini akan diolah menggunakan skala *Likert*. Pernyataan positif akan diberi skor 5, 4, 3, 2, dan 1, sementara pernyataan negatif akan diberi skor 1, 2, 3, 4, dan 5.

Tabel 2. 13 Tabel Kriteria Penilaian

No	Kriteria Penilaian	Sangat Setuju	Setuju	Cukup Setuju	Kurang Setuju	Tidak Setuju
1.	Pernyataan Positif	5	4	3	2	1
2.	Pernyataan Negatif	1	2	3	4	5

Setelah menentukan nilai skor, berikut merupakan rumus untuk hasil perhitungan penilaian responden dihitung menggunakan skala likert :

$$TXPn$$

Dimana : T = Total Jumlah Responden yang memilih

Pn = Pilihan angka skor likert.

a. Rumus *Index %*

Untuk menentukan nilai yang dihasilkan dari pernyataan, dapat menggunakan rumus *index %*. Rumus *index %* adalah sebagai berikut:

$$Index \% = \frac{Total\ skor}{y} \times 100$$

Sebelum melakukan perhitungan, langkah pertama adalah mengetahui interval atau jarak antara nilai terendah hingga tertinggi, serta interpretasi persen untuk mengetahui penilaian dengan metode mencari interval skor persen.

Rumus untuk menghitung interval skor persen adalah:

$$I = \frac{100}{\text{Jumlah skor(Likert)}}$$

Setelah mengetahui interval skor persen, tabel 2.14 dapat digunakan untuk melihat interpretasi skor berdasarkan interval, dengan rentang nilai dari 0% hingga 100%.

Tabel 2. 14 Tabel Parameter Bobot Nilai Positif

<b>Bobot Nilai</b>	<b>Keterangan</b>
<b>0% - 20%</b>	Tidak Setuju
<b>21% - 40%</b>	Kurang Setuju
<b>41% - 60%</b>	Cukup Setuju
<b>61% - 80%</b>	Setuju
<b>81% - 100%</b>	Sangat Setuju

Kemudian nilai responden dihitung dengan perhitungan sebagai berikut :

1. Saya merasa terbantu dengan adanya aplikasi ini karena memudahkan saya dalam mendapatkan informasi sentimen tentang tweet dari masyarakat terhadap Aspek Wisata/Hiburan, Kuliner, Pendidikan, dan Fasilitas/Layanan Publik di Kota Tegal

Tabel 2. 15 Hasil Pengujian Usability Pernyataan Positif No. 1

<b>Responden</b>	<b>Sangat Setuju</b>	<b>Setuju</b>	<b>Cukup Setuju</b>	<b>Kurang Setuju</b>	<b>Tidak Setuju</b>
1	0	1	0	0	0
2	0	1	0	0	0
3	1	0	0	0	0
4	1	0	0	0	0
5	0	1	0	0	0
6	1	0	0	0	0
7	1	0	0	0	0
8	0	1	0	0	0
9	1	0	0	0	0
10	1	0	0	0	0
11	0	1	0	0	0
12	0	1	0	0	0
13	1	0	0	0	0
14	1	0	0	0	0
15	0	1	0	0	0
16	0	1	0	0	0
17	0	1	0	0	0
18	0	1	0	0	0
19	0	1	0	0	0
20	0	1	0	0	0
<b>Jumlah (<i>T</i>)</b>	<b>8</b>	<b>12</b>	<b>0</b>	<b>0</b>	<b>0</b>
<b>Nilai (<i>Pn</i>)</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>2</b>	<b>1</b>
<b>Rumus : <i>T x Pn</i></b>	<b>40</b>	<b>48</b>	<b>0</b>	<b>0</b>	<b>0</b>
<b>Total Skor</b>	<b><math>40 + 48 + 0 + 0 + 0 = 88</math></b>				



Berdasarkan total skor yang diperoleh pada pernyataan positif No. 1 sebesar 88, maka nilai interpretasi responden terhadap aplikasi pada pernyataan ini dapat dihitung menggunakan rumus *Index %*. Dengan menggantikan nilai dalam rumus tersebut, didapatkan:

$$Index \% = \frac{Total\ skor}{y} \times 100 = \frac{88}{100} \times 100 = 88\%$$

Hasil nilai 88% ini masuk ke dalam kategori "Sangat Setuju" berdasarkan interpretasi nilai pada tabel parameter bobot nilai.

2. Saya tidak mengalami kesulitan dalam menemukan informasi yang saya cari berupa jumlah data, persebaran data, dan juga fitur-fitur lain

Tabel 2. 16 Hasil Pengujian Usability Pernyataan Positif No. 2

Responden	Sangat Setuju	Setuju	Cukup Setuju	Kurang Setuju	Tidak Setuju
1	0	1	0	0	0
2	0	1	0	0	0
3	0	1	0	0	0
4	1	0	0	0	0
5	1	0	0	0	0
6	0	1	0	0	0
7	0	1	0	0	0
8	0	1	0	0	0
9	0	1	0	0	0
10	0	0	1	0	0
11	0	1	0	0	0
12	0	0	1	0	0
13	1	0	0	0	0
14	1	0	0	0	0
15	0	1	0	0	0
16	0	1	0	0	0

17	0	1	0	0	0
18	0	1	0	0	0
19	1	0	0	0	0
20	0	1	0	0	0
<b>Jumlah (T)</b>	<b>5</b>	<b>13</b>	<b>2</b>	<b>0</b>	<b>0</b>
<b>Nilai (Pn)</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>2</b>	<b>1</b>
<b>Rumus : <math>T \times Pn</math></b>	<b>25</b>	<b>52</b>	<b>6</b>	<b>0</b>	<b>0</b>
<b>Total Skor</b>	<b><math>25 + 52 + 6 + 0 + 0 = 83</math></b>				

Berdasarkan total skor yang diperoleh pada pernyataan positif No. 2 sebesar 83, maka nilai interpretasi responden terhadap aplikasi pada pernyataan ini dapat dihitung menggunakan rumus *Index %*. Dengan menggantikan nilai dalam rumus tersebut, didapatkan:

$$Index \% = \frac{Total\ skor}{y} \times 100 = \frac{83}{100} \times 100 = 83\%$$

Hasil nilai 83% ini masuk ke dalam kategori "Sangat Setuju" berdasarkan interpretasi nilai pada tabel parameter bobot nilai.

3. Tata letak dan desain halaman ini membuat saya merasa nyaman dan mudah dipahami

Tabel 2. 17 Hasil Pengujian Usability Pernyataan Positif No. 3

<b>Responden</b>	<b>Sangat Setuju</b>	<b>Setuju</b>	<b>Cukup Setuju</b>	<b>Kurang Setuju</b>	<b>Tidak Setuju</b>
1	1	0	0	0	0
2	0	1	0	0	0
3	0	1	0	0	0
4	0	1	0	0	0
5	0	1	0	0	0
6	0	1	0	0	0
7	1	0	0	0	0
8	0	1	0	0	0
9	0	1	0	0	0
10	0	0	1	0	0
11	0	1	0	0	0
12	0	1	0	0	0
13	1	0	0	0	0
14	0	1	0	0	0
15	0	1	0	0	0
16	0	1	0	0	0
17	0	1	0	0	0
18	0	1	0	0	0
19	0	1	0	0	0
20	0	1	0	0	0
<b>Jumlah (I)</b>	<b>3</b>	<b>16</b>	<b>1</b>	<b>0</b>	<b>0</b>
<b>Nilai (Pn)</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>2</b>	<b>1</b>
<b>Rumus : <i>T x Pn</i></b>	<b>15</b>	<b>64</b>	<b>3</b>	<b>0</b>	<b>0</b>
<b>Total Skor</b>	<b><math>15 + 64 + 3 + 0 + 0 = 82</math></b>				

Berdasarkan total skor yang diperoleh pada pernyataan positif No. 3 sebesar 82, maka nilai interpretasi responden terhadap aplikasi pada pernyataan ini dapat dihitung menggunakan rumus *Index %*. Dengan menggantikan nilai dalam rumus tersebut, didapatkan:

$$Index \% = \frac{Total\ skor}{y} \times 100 = \frac{82}{100} \times 100 = 82\%$$

Hasil nilai 82% ini masuk ke dalam kategori "Sangat Setuju" berdasarkan interpretasi nilai pada tabel parameter bobot nilai.

#### 4. Tampilan dan tata letak pada aplikasi ini menarik

Tabel 2. 18 Hasil Pengujian Usability Pernyataan Positif No. 4

<b>Responden</b>	<b>Sangat Setuju</b>	<b>Setuju</b>	<b>Cukup Setuju</b>	<b>Kurang Setuju</b>	<b>Tidak Setuju</b>
1	1	0	0	0	0
2	0	1	0	0	0
3	1	0	0	0	0
4	1	0	0	0	0
5	0	0	1	0	0
6	0	0	1	0	0
7	1	0	0	0	0
8	0	1	0	0	0
9	0	1	0	0	0
10	0	0	1	0	0
11	0	1	0	0	0
12	0	1	0	0	0
13	1	0	0	0	0
14	0	1	0	0	0
15	0	0	1	0	0
16	0	1	0	0	0
17	0	0	1	0	0

18	0	1	0	0	0
19	0	0	1	0	0
20	1	0	0	0	0
<b>Jumlah (T)</b>	<b>6</b>	<b>8</b>	<b>6</b>	<b>0</b>	<b>0</b>
<b>Nilai (Pn)</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>2</b>	<b>1</b>
<b>Rumus : T x Pn</b>	<b>30</b>	<b>32</b>	<b>18</b>	<b>0</b>	<b>0</b>
<b>Total Skor</b>	<b>30 + 32 + 18 + 0 + 0 = 80</b>				

Berdasarkan total skor yang diperoleh pada pernyataan positif No. 4 sebesar 80, maka nilai interpretasi responden terhadap aplikasi pada pernyataan ini dapat dihitung menggunakan rumus *Index %*. Dengan menggantikan nilai dalam rumus tersebut, didapatkan:

$$Index \% = \frac{Total\ skor}{y} \times 100 = \frac{80}{100} \times 100 = 80\%$$

Hasil nilai 80% ini masuk ke dalam kategori "Setuju" berdasarkan interpretasi nilai pada tabel parameter bobot nilai.

5. Penggunaan bahasa pada aplikasi ini mudah dipahami

Tabel 2. 19 Hasil Pengujian Usability Pernyataan Positif No. 5

Responden	Sangat Setuju	Setuju	Cukup Setuju	Kurang Setuju	Tidak Setuju
1	0	1	0	0	0
2	0	1	0	0	0
3	0	0	1	0	0
4	1	0	0	0	0
5	0	1	0	0	0
6	0	1	0	0	0
7	1	0	0	0	0
8	0	1	0	0	0
9	1	0	0	0	0
10	0	1	0	0	0
11	0	0	1	0	0
12	0	1	0	0	0
13	1	0	0	0	0
14	0	1	0	0	0
15	0	1	0	0	0
16	0	1	0	0	0
17	0	1	0	0	0
18	0	1	0	0	0
19	0	1	0	0	0
20	1	0	0	0	0
<b>Jumlah (<i>T</i>)</b>	<b>5</b>	<b>13</b>	<b>2</b>	<b>0</b>	<b>0</b>
<b>Nilai (<i>Pn</i>)</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>2</b>	<b>1</b>
<b>Rumus : <i>T x Pn</i></b>	<b>25</b>	<b>52</b>	<b>6</b>	<b>0</b>	<b>0</b>
<b>Total Skor</b>	<b><math>25 + 52 + 6 + 0 + 0 = 83</math></b>				

Berdasarkan total skor yang diperoleh pada pernyataan positif No. 5 sebesar 83, maka nilai interpretasi responden terhadap aplikasi pada pernyataan ini dapat dihitung menggunakan rumus *Index %*. Dengan menggantikan nilai dalam rumus tersebut, didapatkan:

$$Index \% = \frac{Total\ skor}{y} \times 100 = \frac{83}{100} \times 100 = 83\%$$

Hasil nilai 83% ini masuk ke dalam kategori "Sangat Setuju" berdasarkan interpretasi nilai pada tabel parameter bobot nilai.

6. Saya tidak merasa terbantu dengan adanya aplikasi ini karena tidak memudahkan saya dalam mendapatkan informasi sentimen tentang tweet dari masyarakat terhadap Aspek Wisata/Hiburan, Kuliner, Pendidikan, dan Fasilitas/Layanan Publik di Kota Tegal.

Tabel 2. 20 Hasil Pengujian Usability Pernyataan Negatif No. 1

Responden	Sangat Setuju	Setuju	Cukup Setuju	Kurang Setuju	Tidak Setuju
1	0	0	0	1	0
2	0	0	0	1	0
3	0	0	0	1	0
4	0	0	0	0	1
5	0	0	0	1	0
6	0	0	0	1	0
7	0	0	0	0	1
8	0	0	0	0	1
9	0	0	0	1	0
10	0	0	0	1	0
11	0	0	0	1	0

12	0	0	0	1	0
13	0	0	0	0	1
14	0	0	0	1	0
15	0	0	0	1	0
16	0	0	0	1	0
17	0	0	1	0	0
18	0	0	0	0	1
19	0	0	0	1	0
20	0	0	0	0	1
<b>Jumlah (T)</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>13</b>	<b>6</b>
<b>Nilai (Pn)</b>	<b>-2</b>	<b>-1</b>	<b>0</b>	<b>1</b>	<b>2</b>
<b>Rumus : T x Pn</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>13</b>	<b>12</b>
<b>Total Skor</b>	<b>0 + 0 + 0 + 13 + 12 = 25</b>				

Berdasarkan total skor yang diperoleh pada pernyataan negatif No. 1 sebesar 25, maka nilai interpretasi responden terhadap aplikasi pada pernyataan ini dapat dihitung menggunakan rumus *Index %*. Dengan menggantikan nilai dalam rumus tersebut, didapatkan:

$$Index \% = \frac{Total\ skor}{y} \times 100 = \frac{25}{100} \times 100 = 25\%$$

Hasil nilai 25% ini masuk ke dalam kategori "Kurang Setuju" berdasarkan interpretasi nilai pada tabel parameter bobot nilai.

7. Saya mengalami kesulitan dalam menemukan informasi yang saya cari berupa jumlah data, persebaran data, dan juga fitur-fitur lain



Tabel 2. 21 Hasil Pengujian Usability Pernyataan Negatif No. 2

Responden	Sangat Setuju	Setuju	Cukup Setuju	Kurang Setuju	Tidak Setuju
1	0	0	0	0	1
2	0	0	0	1	0
3	0	0	0	0	1
4	0	0	0	1	0
5	0	0	0	1	0
6	0	0	0	1	0
7	0	0	0	0	1
8	0	0	0	1	0
9	0	0	0	1	0
10	0	0	0	1	0
11	0	0	0	1	0
12	0	0	0	0	1
13	0	0	0	0	1
14	0	0	0	1	0
15	0	0	0	0	1
16	0	0	0	1	0
17	0	0	1	0	0
18	0	0	0	1	0
19	0	0	0	1	0
20	0	0	0	0	1
<b>Jumlah (T)</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>12</b>	<b>7</b>
<b>Nilai (Pn)</b>	<b>-2</b>	<b>-1</b>	<b>0</b>	<b>1</b>	<b>2</b>
<b>Rumus : <math>T \times Pn</math></b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>12</b>	<b>7</b>
<b>Total Skor</b>	<b><math>0 + 0 + 0 + 12 + 7 = 19</math></b>				

Berdasarkan total skor yang diperoleh pada pernyataan negatif No. 2 sebesar 19, maka nilai interpretasi responden terhadap aplikasi pada pernyataan

ini dapat dihitung menggunakan rumus *Index %*. Dengan menggantikan nilai dalam rumus tersebut, didapatkan:

$$Index \% = \frac{Total\ skor}{y} \times 100 = \frac{19}{100} \times 100 = 19\%$$

Hasil nilai 19% ini masuk ke dalam kategori "Tidak Setuju" berdasarkan interpretasi nilai pada tabel parameter bobot nilai.

8. Tata letak dan desain halaman ini membuat saya merasa tidak nyaman dan tidak mudah dipahami

Tabel 2. 22 Hasil Pengujian Usability Pernyataan Negatif No. 3

<b>Responden</b>	<b>Sangat Setuju</b>	<b>Setuju</b>	<b>Cukup Setuju</b>	<b>Kurang Setuju</b>	<b>Tidak Setuju</b>
1	0	0	0	0	1
2	0	0	0	1	0
3	0	0	0	0	1
4	0	0	0	0	1
5	0	0	0	1	0
6	0	0	0	1	0
7	0	0	0	0	1
8	0	0	0	1	0
9	0	0	0	1	0
10	0	0	0	1	0
11	0	0	0	0	1
12	0	0	0	0	1
13	0	0	0	0	1
14	0	0	0	1	0
15	0	0	1	0	0
16	0	0	0	1	0
17	0	0	0	1	0
18	0	0	0	0	1
19	0	0	0	1	0

20	0	0	0	1	0
<b>Jumlah (T)</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>11</b>	<b>8</b>
<b>Nilai (Pn)</b>	<b>-2</b>	<b>-1</b>	<b>0</b>	<b>1</b>	<b>2</b>
<b>Rumus : T x Pn</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>11</b>	<b>16</b>
<b>Total Skor</b>	<b>0 + 0 + 0 + 11 + 16 = 27</b>				

Berdasarkan total skor yang diperoleh pada pernyataan negatif No. 3 sebesar 27, maka nilai interpretasi responden terhadap aplikasi pada pernyataan ini dapat dihitung menggunakan rumus *Index %*. Dengan menggantikan nilai dalam rumus tersebut, didapatkan:

$$Index \% = \frac{Total\ skor}{y} \times 100 = \frac{27}{100} \times 100 = 27\%$$

Hasil nilai 27% ini masuk ke dalam kategori "Kurang Setuju" berdasarkan interpretasi nilai pada tabel parameter bobot nilai.

9. Tampilan dan tata letak pada aplikasi ini tidak menarik

Tabel 2. 23 Hasil Pengujian Usability Pernyataan Negatif No. 4

<b>Responden</b>	<b>Sangat Setuju</b>	<b>Setuju</b>	<b>Cukup Setuju</b>	<b>Kurang Setuju</b>	<b>Tidak Setuju</b>
1	0	0	0	0	1
2	0	0	0	0	1
3	0	0	0	1	0
4	0	0	0	1	0
5	0	0	0	0	1
6	0	0	0	1	0
7	0	0	0	1	0
8	0	0	0	0	1

9	0	0	0	1	0
10	0	0	0	1	0
11	0	0	0	1	0
12	0	0	0	0	1
13	0	0	0	0	1
14	0	0	0	0	1
15	0	0	1	0	0
16	0	0	0	1	0
17	0	0	0	1	0
18	0	0	0	1	0
19	0	0	1	0	0
20	0	0	0	0	1
<b>Jumlah (T)</b>	<b>0</b>	<b>0</b>	<b>2</b>	<b>9</b>	<b>8</b>
<b>Nilai (Pn)</b>	<b>-2</b>	<b>-1</b>	<b>0</b>	<b>1</b>	<b>2</b>
<b>Rumus : T x Pn</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>9</b>	<b>16</b>
<b>Total Skor</b>	<b>0 + 0 + 0 + 9 + 16 = 25</b>				

Berdasarkan total skor yang diperoleh pada pernyataan negatif No. 4 sebesar 25, maka nilai interpretasi responden terhadap aplikasi pada pernyataan ini dapat dihitung menggunakan rumus *Index %*. Dengan menggantikan nilai dalam rumus tersebut, didapatkan:

$$Index \% = \frac{Total\ skor}{y} \times 100 = \frac{25}{100} \times 100 = 25\%$$

Hasil nilai 25% ini masuk ke dalam kategori "Kurang Setuju" berdasarkan interpretasi nilai pada tabel parameter bobot nilai.

10. Penggunaan bahasa pada aplikasi ini tidak mudah dipahami

Tabel 2. 24 Hasil Pengujian Usability Pernyataan Negatif No. 5

<b>Responden</b>	<b>Sangat Setuju</b>	<b>Setuju</b>	<b>Cukup Setuju</b>	<b>Kurang Setuju</b>	<b>Tidak Setuju</b>
1	0	0	0	0	1
2	0	0	0	0	1
3	0	0	0	0	1
4	0	0	0	0	1
5	0	0	0	1	0
6	0	0	0	0	1
7	0	0	0	0	1
8	0	0	0	1	0
9	0	0	0	1	0
10	0	0	0	0	1
11	0	0	0	0	1
12	0	0	0	1	0
13	0	0	0	0	1
14	0	0	0	0	1
15	0	0	0	0	1
16	0	0	0	1	0
17	0	0	0	0	1
18	0	0	0	0	1
19	0	0	0	1	0
20	0	0	0	1	0
<b>Jumlah (T)</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>7</b>	<b>13</b>
<b>Nilai (Pn)</b>	<b>-2</b>	<b>-1</b>	<b>0</b>	<b>1</b>	<b>2</b>
<b>Rumus : <i>T x Pn</i></b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>7</b>	<b>26</b>
<b>Total Skor</b>	<b><math>0 + 0 + 0 + 7 + 26 = 33</math></b>				

Berdasarkan total skor yang diperoleh pada pernyataan negatif No. 5 sebesar 33, maka nilai interpretasi responden terhadap aplikasi pada pernyataan ini dapat dihitung menggunakan rumus *Index %*. Dengan menggantikan nilai dalam rumus tersebut, didapatkan:

$$Index \% = \frac{Total\ skor}{y} \times 100 = \frac{33}{100} \times 100 = 33\%$$

Hasil nilai 33% ini masuk ke dalam kategori "Kurang Setuju" berdasarkan interpretasi nilai pada tabel parameter bobot nilai.

Evaluasi pengujian *usability* diukur dengan menggunakan persamaan berdasarkan data yang didapat dari tabel pernyataan positif dan negatif. Secara keseluruhan, hasil perhitungan pengujian *usability* dijabarkan seperti berikut ini:

$$Usability\ Pernyataan\ Positif\ (\%) = \frac{88+83+82+80+83}{5} \times 100\%$$

$$Usability\ (\%) = 83,2\% \text{ (Sangat Setuju)}$$

$$Usability\ Pernyataan\ Negatif\ (\%) = \frac{25+19+27+25+33}{5} \times 100\%$$

$$Usability\ (\%) = 25,8\% \text{ (Kurang Setuju)}$$

Demikian perencanaan dan rancangan untuk aplikasi Sentimen Analisis Kota Tegal.

### 2.3.4. Hasil Analisis Sentimen Perbulan pada Setiap Aspek Kategori

Tabel 2. 25 Hasil Sentimen Analisis Setiap Aspek

Tahun	Bulan	Aspek							
		Wisata/ Hiburan		Pendidikan		Fasilitas/ Layanan Publik		Kuliner	
		P	N	P	N	P	N	P	N
2020	Januari	10	0	5	0	19	1	6	0
	Februari	6	0	0	1	7	0	4	0
	Maret	18	0	112	0	27	0	7	0
	April	3	0	1	0	8	0	0	0
	Mei	16	0	7	0	30	0	7	0
	Juni	3	0	9	0	17	0	2	0
	Juli	3	0	2	0	1	0	0	0
	Agustus	12	0	2	0	7	0	2	0
	September	5	0	6	3	25	0	0	0
	Oktober	2	0	2	0	11	0	0	0
	November	4	0	3	0	16	2	2	0
	Desember	26	0	1	0	41	0	11	0
2021	Januari	21	1	5	3	35	0	5	0
	Februari	6	0	9	0	38	0	3	0
	Maret	9	0	6	0	26	1	4	1
	April	12	0	9	0	24	0	6	0
	Mei	13	0	2	0	38	0	11	0
	Juni	9	0	6	0	20	0	5	0
	Juli	3	0	5	0	44	0	2	0
	Agustus	2	0	8	1	30	0	6	0

	<b>September</b>	7	1	6	0	27	0	3	0
	<b>Oktober</b>	6	0	3	0	26	0	5	0
	<b>November</b>	6	0	2	0	27	1	1	0
	<b>Desember</b>	12	0	16	1	18	2	10	0
<b>2022</b>	<b>Januari</b>	9	0	7	0	40	0	3	0
	<b>Februari</b>	9	0	8	0	34	0	3	0
	<b>Maret</b>	11	0	3	0	22	0	6	0
	<b>April</b>	5	0	1	0	19	0	11	0
	<b>Mei</b>	14	0	22	1	33	1	5	0
	<b>Juni</b>	11	0	2	0	25	0	3	0
	<b>Juli</b>	11	0	6	0	30	1	8	0
	<b>Agustus</b>	2	0	0	0	2	0	0	0
	<b>September</b>	7	0	5	1	21	1	4	0
	<b>Oktober</b>	6	0	6	1	20	0	6	0
	<b>November</b>	7	0	0	2	18	0	10	0
	<b>Desember</b>	22	1	6	0	22	0	11	0
<b>2023</b>	<b>Januari</b>	11	0	4	1	31	1	8	0
	<b>Februari</b>	14	0	3	0	35	1	6	0
	<b>Maret</b>	5	0	15	0	25	0	6	0
	<b>April</b>	11	0	2	0	21	0	14	0
	<b>Mei</b>	8	0	6	3	11	0	4	0
	<b>Juni</b>	94	0	43	2	289	6	137	0
	<b>Juli</b>	86	2	16	7	216	5	77	0
	<b>Agustus</b>	1	1	0	0	81	4	33	3



Keterangan :

- Selama beberapa tahun, sentimen positif umumnya lebih banyak daripada sentimen negatif dalam berbagai aspek.
- Ada variasi dalam penyebaran sentimen dari tahun ke tahun dan dari bulan ke bulan yang berbeda, karena dipengaruhi oleh peristiwa atau tren khusus yang terjadi selama periode tersebut.
- Pada tahun 2023, terlihat peningkatan yang signifikan dalam aktivitas terkait sentimen, terutama pada bulan Juni dan Juli.

## **2.4. Kesimpulan dan Saran**

### **2.4.1. Kesimpulan**

Pada pengembangan aplikasi ini, dapat disimpulkan beberapa poin penting sebagai berikut:

1. Aplikasi ini mampu melakukan klasifikasi ulasan masyarakat Kota Tegal secara otomatis berdasarkan beberapa aspek utama seperti wisata/hiburan, pendidikan, fasilitas/layanan publik, dan kuliner.
2. Aplikasi ini menawarkan fitur *real-time*, sehingga pengguna dapat melihat hasil klasifikasi ulasan secara langsung tanpa adanya keterlambatan. Dan fitur lain seperti fitur filter tanggal yang memungkinkan pengguna untuk mengatur rentang waktu tertentu, sehingga mereka dapat membaca ulasan dengan lebih rinci dan spesifik sesuai dengan periode yang diinginkan.

3. Hasil pengujian menunjukkan bahwa aplikasi ini memiliki tingkat akurasi rata-rata diatas 80% dalam mengklasifikasikan ulasan masyarakat kedalam aspek-aspek yang telah ditentukan sebelumnya.
4. Pada tahun 2023, terlihat peningkatan yang signifikan dalam aktivitas terkait sentimen, terutama pada bulan Juni dan Juli

Secara keseluruhan, aplikasi ini memberikan kontribusi penting dalam membantu pemerintah Kota Tegal untuk mengumpulkan, mengklasifikasikan, dan menganalisis ulasan masyarakat secara efektif.

#### **2.4.2. Saran**

Berdasarkan penelitian yang telah dilakukan, aplikasi sentimen analisis Kota Tegal Berbasis Aspek ini dapat dikembangkan sebagai berikut :

1. Karena *twitter* memberlakukan kebijakan baru dengan membatasi hak akses *crawling* untuk pengguna *twitter*. Maka perlu dilakukan pengambilan data dengan membuat opsi platform lain seperti website, instagram, facebook, dan lain-lain.
2. Menggunakan metode machine learning yang berbeda untuk membandingkan nilai akurasi setiap model Seperti KNN ataupun SVM.
3. Menggunakan framework website lain yang dapat menampilkan dashboard yang berbeda dan lebih interaktif seperti flask, djanggo, dan lain-lain.

## **BAB III**

### **HKI**

#### **1.1. Proses**

Dalam proses pendaftaran Hak Kekayaan Intelektual di Pusat Pelatihan dan Pengabdian Masyarakat (P3M) Politeknik Harapan Bersama, ada beberapa dokumen yang diperlukan. Dokumen-dokumen ini digunakan untuk melakukan pendaftaran Hak Kekayaan Intelektual, termasuk Manuskrip Buku, Surat Pernyataan Pengajuan HKI, Surat Pengalihan HKI, Dokumen Teknis, dan KTP pemohon. Setelah proses pendaftaran Hak Kekayaan Intelektual (HKI) selesai, Anda akan menerima sertifikat HKI yang telah terdaftar dalam pangkalan data Kekayaan Intelektual Indonesia. Informasi mengenai HKI yang telah didaftarkan dapat diakses melalui situs web pangkalan data Kekayaan Intelektual Indonesia di (<https://pdki-indonesia.dgip.go.id/>).

#### **1.2. Identitas HKI**

Identitas HKI “Aplikasi Sentimen Analisis Kota Tegal Berbasis Aspek Menggunakan Naïve Bayes” dapat dilihat dibawah ini :

Nomor : EC00202355151

Tanggal : 13 Juli 2023  
dikeluarkan

Nama Pencipta : . Novita Fitria Putri

. Muhammad Fikri Hidayattullah, S.T., M.Kom.

. Dwi Intan Af'idah, S.T., M.Kom.

Nama : Pusat Penelitian dan Pengabdian Masyarakat (P3M)  
Pemegang Hak Politeknik Harapan Bersama Tegal  
Cipta  
Jenis Ciptaan : Program Komputer  
Judul Ciptaan : Aplikasi Sentimen Analisis Kota Tegal Berbasis Aspek  
Menggunakan Algoritma *Naïve Bayes*  
URL bukti : <https://e-hakcipta.dgip.go.id/index.php/c?code=YTIkOWJkM2YzZmY0MTZmM2MwY2E1Y2FlMmJkOTZmNTIK>

## DAFTAR PUSTAKA

- [1] T. rifda Khaerani and M. V. Sarai, “Analisis Pengembangan Electronic Government Di Kota Tegal,” vol. 8, 2020.
- [2] M. W. Graham, E. J. Avery, and S. Park, “The role of social media in local government crisis communications,” *Public Relat Rev*, vol. 41, no. 3, pp. 386–394, Sep. 2015, doi: 10.1016/j.pubrev.2015.02.001.
- [3] A. Zhao and Y. Yu, “Knowledge-enabled BERT for aspect-based sentiment analysis,” *Knowl Based Syst*, vol. 227, Sep. 2021, doi: 10.1016/j.knosys.2021.107220.
- [4] M. I. Fikri, T. S. Sabrila, Y. Azhar, and U. M. Malang, “Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter,” 2020.
- [5] S. K. Dirjen, P. Riset, D. Pengembangan, R. Dikti, and O. Heranova, “Terakreditasi SINTA Peringkat 2 Synthetic Minority Oversampling Technique pada Averaged One Dependence Estimators untuk Klasifikasi Credit Scoring,” *masa berlaku mulai*, vol. 1, no. 3, pp. 443–450, 2017.
- [6] W. Yulita *et al.*, “Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier,” *JDMSI*, vol. 2, no. 2, pp. 1–9, 2021.
- [7] V. A. Fitri, R. Andreswari, and M. A. Hasibuan, “Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve

- Bayes, decision tree, and random forest algorithm,” in *Procedia Computer Science*, Elsevier B.V., 2019, pp. 765–772. doi: 10.1016/j.procs.2019.11.181.
- [8] S. A. Azzahra and A. Wibowo, “Analisis Sentimen Multi-Aspek Berbasis Konversi Ikon Emosi Dengan Algoritme Naïve Bayes Untuk Ulasan Wisata Kuliner Pada Web Tripadvisor,” vol. 7, no. 4, 2020, doi: 10.25126/jtiik.202071907.
- [9] J. W. Iskandar and Y. Nataliani, “Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 6, pp. 1120–1126, Dec. 2021, doi: 10.29207/resti.v5i6.3588.
- [10] Eko Budi S and Aryo N, “Aspect Based Sentimen Analysis Opini Publik Pada Instagram dengan Convolutional Neural Network”.
- [11] C. H. Yutika, A. Adiwijaya, and S. Al Faraby, “Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 2, p. 422, Apr. 2021, doi: 10.30865/mib.v5i2.2845.
- [12] H. Gunawan Sulistio and A. Handojo, “Aspect-Based Sentiment Analysis pada Ulasan E-Commerce dengan Metode Support Vector Machine untuk Mendapatkan Informasi Sentimen dari Beberapa Aspek.”
- [13] O. Somantri, “JEPIN (Jurnal Edukasi dan Penelitian Informatika) Analisis Sentimen Penilaian Tempat Tujuan Wisata Kota Tegal Berbasis Text Mining”, [Online]. Available: [www.google.com/maps](http://www.google.com/maps)

- [14] A. Kadhim, "An Evaluation of Preprocessing Techniques for Text Classification Pattern Recognition View project Improvement text classification using log(TF-IDF) with K-NN Algorithm View project," 2018. [Online]. Available: <https://sites.google.com/site/ijcsis/>
- [15] R. Wati, S. Ernawati, and H. Rachmi, "Jurnal Manajemen Informatika (JAMIKA) Pembobotan TF-IDF Menggunakan Naïve Bayes Pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH TF-IDF Weighting Using Naïve Bayes on Public Sentiment on The Issue of Rising BIPIH," 2023, doi: 10.34010/jamika.v13i1.9424.
- [16] I. Mathilda Yulietha and S. Al Faraby, "Klasifikasi Sentimen Review Film Menggunakan Algoritma Support Vector Machine Sentiment Classification Of Movie Reviews Using Algorithm Support Vector Machine," 2017. [Online]. Available: <https://www.cs.cornell.edu/people/pablo/movie-review-data/>
- [17] C. H. Yutika, A. Adiwijaya, and S. Al Faraby, "Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 2, p. 422, Apr. 2021, doi: 10.30865/mib.v5i2.2845.
- [18] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional," vol. 15, no. 1, 2021.
- [19] R. Adinugroho, "Perbandingan Rasio Split Data Training Dan Data Testing Menggunakan Metode Lstm Dalam Memprediksi Harga Indeks Saham Asia."

- [20] R. Fajar, S. Program, P. Rekayasa, N. Lunak, and R. Bengkalis, “Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter,” vol. 3, no. 1.
- [21] “tweet-harvest.” <https://socket.dev/npm/package/tweet-harvest> (accessed Jul. 29, 2023).



## LAMPIRAN

### Lampiran 1 Surat Kesediaan Pembimbing

#### SURAT KESEPAKATAN BIMBINGAN SKRIPSI

Kami yang bertanda tangan di bawah ini :

Pihak Pertama

Nama : Novita Fitria Putri  
NIM : 19090130  
Program Studi : Sarjana Terapan Teknik Informatika

Pihak Kedua

Nama : Muhammad Fikri Hidayattullah, S.T.,M.Kom.  
Status : Dosen  
NIDN : 0623108801  
Jabatan Fungsional : Lektor  
Pangkat/Golongan : Penata-III/c

Pada hari Jumat tanggal 10 Maret 2023 telah terjadi sebuah kesepakatan bahwa Pihak Kedua bersedia menjadi Pembimbing I Skripsi Pihak Pertama dengan syarat Pihak Pertama wajib menyelesaikan skripsi selambat-lambatnya satu semester dan bersedia tidak menggunakan jasa joki dalam pembuatan skripsi tersebut, serta melakukan bimbingan Skripsi minimal 8 kali kepada Pihak Kedua. Adapun waktu dan tempat pelaksanaan disepakati antar pihak. Demikian kesepakatan ini dibuat dengan penuh kesadaran guna kelancaran penyelesaian Skripsi.

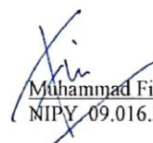
Tegal, 10 Maret 2023

Pihak Pertama



Novita Fitria Putri

Pihak Kedua



Muhammad Fikri Hidayattullah, S.T.,M.Kom  
NIPY 09.016.307

Mengetahui

Ketua Program Studi Sarjana Terapan Teknik Informatika



Slamet Wiyono, S.Pd., M.Eng.  
NIPY 08.015.222

### SURAT KESEPAKATAN BIMBINGAN SKRIPSI

Kami yang bertanda tangan di bawah ini :

Pihak Pertama

Nama : Novita Fitria Putri  
NIM : 19090130  
Program Studi : Sarjana Terapan Teknik Informatika

Pihak Kedua

Nama : Dwi Intan Af'idah, M. Kom.  
Status : Dosen  
NIDN : 0620089303  
Jabatan Fungsional : Asisten Ahli  
Pangkat/Golongan : Penata Muda Tk.I/III

Pada hari Selasa tanggal 7 Maret 2023 telah terjadi sebuah kesepakatan bahwa Pihak Kedua bersedia menjadi Pembimbing II Skripsi Pihak Pertama dengan syarat Pihak Pertama wajib melakukan bimbingan Skripsi minimal 8 kali kepada Pihak Kedua dengan syarat ada progress setiap dua minggu. Adapun waktu dan tempat pelaksanaan disepakati antar pihak.

Demikian kesepakatan ini dibuat dengan penuh kesadaran guna kelancaran penyelesaian Skripsi.

Tegal, 7 Maret 2023

Pihak Pertama

  
Novita Fitria Putri

Pihak Kedua

  
Dwi Intan Af'idah, M. Kom.

Mengetahui  
Ketua Program Studi Sarjana Terapan Teknik Informatika



Slamet Wiyono, S.Pd., M.Eng.  
NIPY. 08.015.222

## Lampiran 2 Surat Pernyataan HKI

### SURAT PERNYATAAN

Yang bertanda tangan di bawah ini, pemegang hak cipta:

1. N a m a : Novita Fitria Putri  
Kewarganegaraan : Indonesia  
Alamat : Griya Santika Blok N.15, Dukuhturi, Kab. Tegal
2. N a m a : Muhammad Fikri Hidayattullah, S.T., M.Kom.  
Kewarganegaraan : Indonesia  
Alamat : Jl. Glatik No. 68 Randugunting, Kota Tegal
3. N a m a : Dwi Intan Af'idah, S.T., M.Kom.  
Kewarganegaraan : Indonesia  
Alamat : Desa Grinting RT003/RW001, Kecamatan Bulakamba, Kabupaten Brebes, 52253

Dengan ini menyatakan bahwa:

1. Karya Cipta yang saya mohonkan:  
Berupa : Program Komputer  
Berjudul : Aplikasi Sentimen Analisis Kota Tegal Berbasis Aspek Menggunakan Algoritma *Naïve Bayes*
  - Tidak meniru dan tidak sama secara esensial dengan Karya Cipta milik pihak lain atau obyek kekayaan intelektual lainnya sebagaimana dimaksud dalam Pasal 68 ayat (2);
  - Bukan merupakan Ekspresi Budaya Tradisional sebagaimana dimaksud dalam Pasal 38;
  - Bukan merupakan Ciptaan yang tidak diketahui penciptanya sebagaimana dimaksud dalam Pasal 39;
  - Bukan merupakan hasil karya yang tidak dilindungi Hak Cipta sebagaimana dimaksud dalam Pasal 41 dan 42;
  - Bukan merupakan Ciptaan seni lukis yang berupa logo atau tanda pembeda yang digunakan sebagai merek dalam perdagangan barang/jasa atau digunakan sebagai lambang organisasi, badan usaha, atau badan hukum sebagaimana dimaksud dalam Pasal 65 dan;
  - Bukan merupakan Ciptaan yang melanggar norma agama, norma susila, ketertiban umum, pertahanan dan keamanan negara atau melanggar peraturan perundang-undangan sebagaimana dimaksud dalam Pasal 74 ayat (1) huruf d Undang-Undang Nomor 28 Tahun 2014 tentang Hak Cipta.
2. Sebagai pemohon mempunyai kewajiban untuk menyimpan asli contoh ciptaan yang dimohonkan dan harus memberikan apabila dibutuhkan untuk kepentingan penyelesaian sengketa perdata maupun pidana sesuai dengan ketentuan perundang-undangan.
3. Karya Cipta yang saya mohonkan pada Angka 1 tersebut di atas tidak pernah dan tidak sedang dalam sengketa pidana dan/atau perdata di Pengadilan.
4. Dalam hal ketentuan sebagaimana dimaksud dalam Angka 1 dan Angka 3 tersebut di atas saya / kami langgar, maka saya / kami bersedia secara sukarela bahwa:

- a. permohonan karya cipta yang saya ajukan dianggap ditarik kembali; atau
- b. Karya Cipta yang telah terdaftar dalam Daftar Umum Ciptaan Direktorat Hak Cipta, Direktorat Jenderal Hak Kekayaan Intelektual, Kementerian Hukum Dan Hak Asasi Manusia R.I dihapuskan sesuai dengan ketentuan perundang-undangan yang berlaku.
- c. Dalam hal kepemilikan Hak Cipta yang dimohonkan secara elektronik sedang dalam perkara dan/atau sedang dalam gugatan di Pengadilan maka status kepemilikan surat pencatatan elektronik tersebut ditangguhkan menunggu putusan Pengadilan yang berkekuatan hukum tetap.

Demikian Surat pernyataan ini saya/kami buat dengan sebenarnya dan untuk dipergunakan sebagaimana mestinya.

Tegal, 7 Juli 2023



10000  
SPESIAL KEM. HUKUM  
METERA  
TEMPEL  
776E1A0524900565

( Novita Fitria Putri )

Pemegang Hak Cipta\*



( Muhammad Fikri Hidayattullah, S.T., M.Kom.)

Pemegang Hak Cipta\*



( Dwi Intan Af'idah, S.T., M.Kom.)

Pemegang Hak Cipta\*

### Lampiran 3 Surat Pengalihan HKI

#### SURAT PENGALIHAN HAK CIPTA

Yang bertanda tangan di bawah ini :

1. N a m a : Novita Fitria Putri  
Kewarganegaraan : Indonesia  
Alamat : Griya Santika Blok N.15, Dukuhturi, Kab. Tegal
2. N a m a : Muhammad Fikri Hidayattullah, S.T., M.Kom.  
Kewarganegaraan : Indonesia  
Alamat : Jl. Glatik No. 68 Randugunting, Kota Tegal
3. N a m a : Dwi Intan Af'idah, S.T., M.Kom.  
Kewarganegaraan : Indonesia  
Alamat : Desa Grinting RT003/RW001, Kecamatan Bulakamba, Kabupaten Brebes, 52253

Adalah **Pihak I** selaku pencipta, dengan ini menyerahkan karya ciptaan saya kepada :

N a m a : Pusat Penelitian dan Pengabdian Masyarakat (P3M)  
Politeknik Harapan Bersama  
Alamat : Jl. Mataram no. 9 Pesurungan Lor Kota Tegal

Adalah **Pihak II** selaku Pemegang Hak Cipta berupa Program Komputer dengan judul "**Aplikasi Sentimen Analisis Kota Tegal Berbasis Aspek Menggunakan Algoritma Naïve Bayes**" untuk didaftarkan di Direktorat Hak Cipta dan Desain Industri, Direktorat Jenderal Kekayaan Intelektual, Kementerian Hukum dan Hak Asasi Manusia Republik Indonesia. Demikianlah surat pengalihan hak ini kami buat, agar dapat dipergunakan sebagaimana mestinya.

Tegal, Juli 2023  
Pencipta

Pemegang Hak Cipta Pencipta  
Ketua P3M

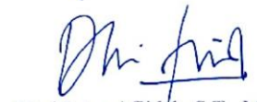


(Dr. Aldi Budi Riyanta, S.Si, M.T.)

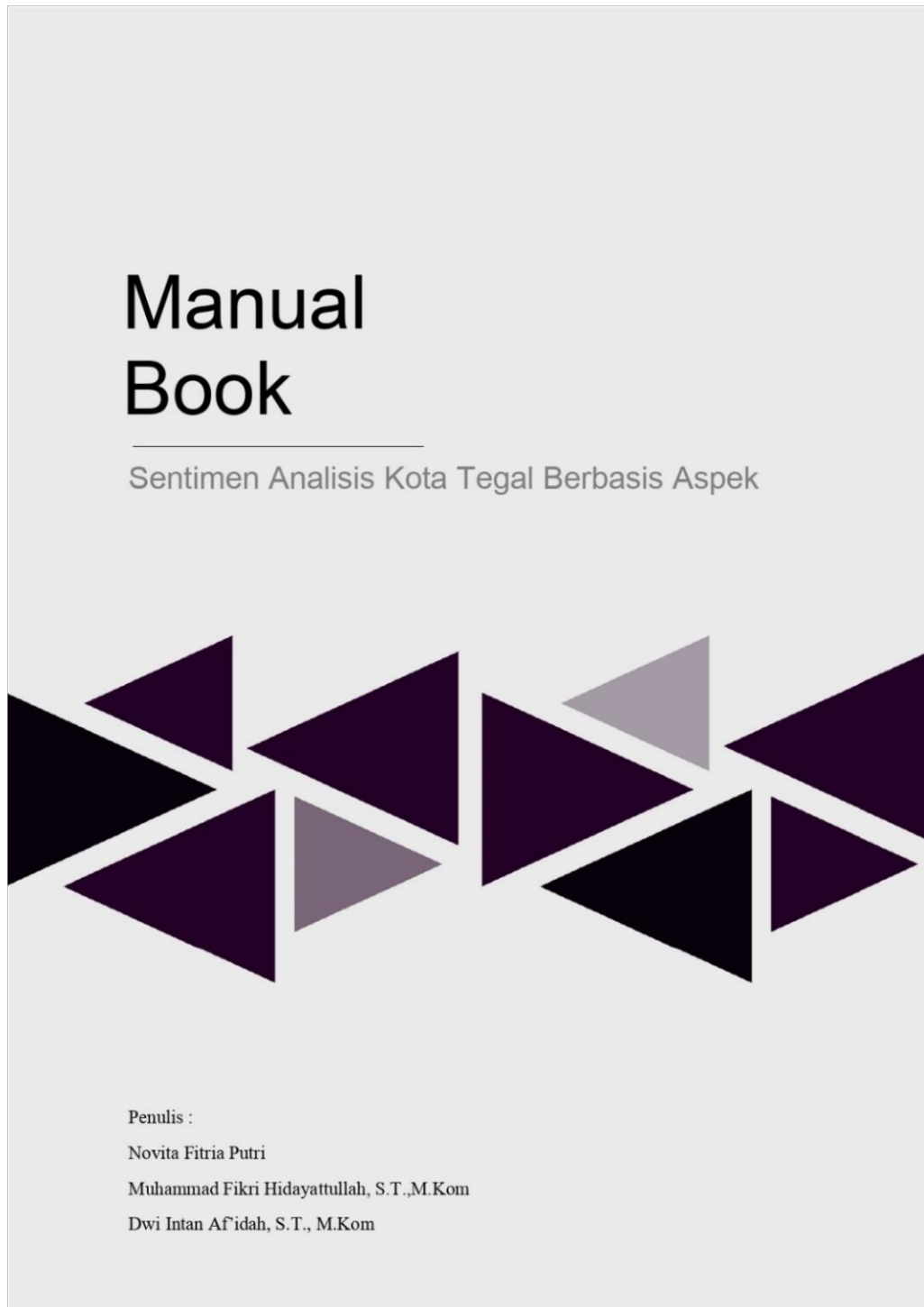


(Novita Fitria Putri)

(Muhammad Fikri Hidayattullah, S.T., M.Kom.)



(Dwi Intan Af'idah, S.T., M.Kom.)



## **1. PENDAHULUAN**

### **1.1. Tujuan Pembuatan Dokumen**

Dokumen user manual Website - "Sentimen Analisis Kota Tegal Berbasis Aspek" ini dibuat untuk tujuan sebagai berikut :

1. Menggambarkan dan menjelaskan penggunaan website "Sentimen Analisis Kota Tegal Berbasis Aspek".
2. Sebagai panduan penggunaan "Sentimen Analisis Kota Tegal Berbasis Aspek"

### **1.2. Deskripsi Umum Sistem**

#### **1.2.1. Deskripsi Umum Aplikasi**

Deskripsi umum aplikasi ini meliputi deskripsi pada *dashboard* "Sentimen Analisis Kota Tegal Berbasis Aspek". Fungsi utama dashboard ini adalah menampilkan informasi dari ulasan masyarakat tentang Kota Tegal pada beberapa aspek dengan tampilan yang informatif. *Dashboard* Sentimen Analisis Kota Tegal ini dirancang untuk membantu pemerintah Kota Tegal dalam mengklasifikasikan opini atau ulasan masyarakat Kota Tegal di *Twitter* secara cepat dan efektif. Dimana opini atau ulasan masyarakat Kota Tegal ini di klasifikasikan kedalam beberapa aspek atau kategori ulasan seperti; Aspek Wisata/Hiburan, Aspek Pendidikan, Aspek Fasilitas/Layanan Publik, dan Aspek Kuliner. Pada Aplikasi Dashboard ini sudah dilengkapi dengan fitur untuk menentukan filter tanggal untuk memudahkan pengguna membaca data serta tampilan dashboard yang interaktif dengan fitur persebaran data yang rinci

#### **1.2.2. Deskripsi Umum Kebutuhan Aplikasi**

Deskripsi umum kebutuhan aplikasi yang akan diimplementasikan meliputi semua informasi yang bersifat teknik yang menjadi acuan dalam pengembangan aplikasi.

### **1.3. Deskripsi Dokumen**

Deskripsi ini dibuat untuk memberikan panduan penggunaan website “Sentimen Analisis Kota Tegal Berbasis Aspek”. Dokumen ini berisikan informasi sebagai berikut :

#### **1. BAB I**

Berisikan informasi umum yang merupakan bagian pendahuluan, yang meliputi tujuan pembuatan dokumen, deskripsi umum sistem, dan deskripsi dokumen.

#### **2. BAB II**

Berisikan perangkat yang dibutuhkan untuk penggunaan *website* “Sentimen Analisis Kota Tegal Berbasis Aspek”. Meliputi perangkat lunak dan perangkat keras.

#### **3. BAB III**

Berisikan *user manual website* “Sentimen Analisis Kota Tegal Berbasis Aspek”.

## **2. PERANGKAT YANG DIBUTUHKAN**

### **2.1. Perangkat Lunak**

Perangkat lunak yang digunakan :

1. Sistem operasi *Windows*
2. *Tools* penjelajah (*Web Browser*) seperti ; *Firefox*, *Google Chrome*, *Microsoft Edge*.

#### **a. Perangkat Keras**

Perangkat keras yang dibutuhkan :

1. Laptop/PC (Personal Komputer)
2. Mouse, Keyboard, dan Monitor sebagai peralatan antarmuka



### **3. MENU DAN CARA PENGGUNAAN**

#### **a. Struktur Menu**

Adapun struktur menu pada website “Sentimen Analisis Kota Tegal Berbasis Aspek” adalah sebagai berikut :

- Home
- Dashboard
  - Halaman Utama
  - Pilihan Kategori Aspek
  - Pilihan Tanpa Filter
  - Pilihan dengan Filter

#### **b. Cara Penggunaan**

Pada bagian ini akan dijelaskan mengenai menu dan cara penggunaan pada dashboard Sentimen Analisis Kota Tegal Berbasis Aspek.

##### **1. Cara Membuka**

Bukalah website “Sentimen Analisis Kota Tegal Berbasis Aspek” dengan menggunakan Web Browser dengan alaman URL yang telah disediakan.

Maka secara otomatis akan diarahkan kedalam Halaman “**Home**”, seperti pada gambar dibawah ini :



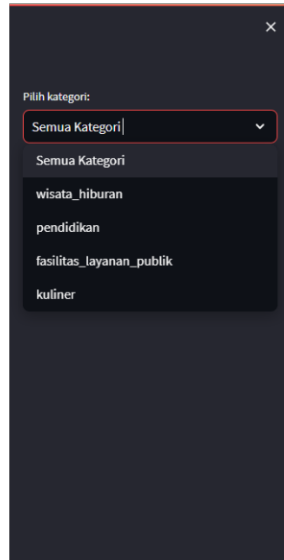
Gambar 1. Home

- Pada halaman terdapat pilihan **“Dashboard”**
- Klik** tombol **“Dashboard”** tersebut maka akan muncul halaman utama seperti gambar berikut :



Gambar 2. Halaman Dashboard

- Kemudian pada halaman utama terdapat fitur berisi pilihan kategori aspek yang dapat dipilih oleh pengguna seperti pada gambar berikut :



Gambar 3. Pilihan Kategori

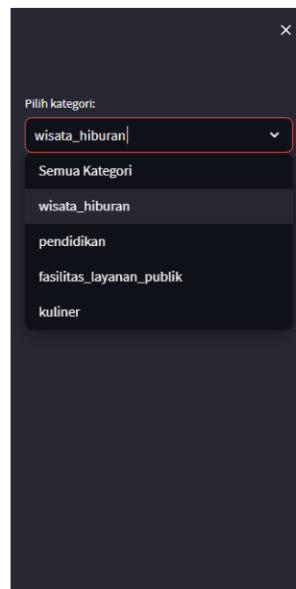
- d. Tekan pilihan “**semua kategori**”, kemudian akan menampilkan tampilan jumlah sentimen positif dan sentimen negatif dari semua kategori aspek seperti pada gambar berikut :



Gambar 4. Jumlah Sentimen Semua Kategori

## 2. Jumlah Sentimen Setiap Kategori

- a. Tekan salah satu kategori aspek “**wisata/hiburan**” seperti pada gambar berikut :



Gambar 5. Pilihan Kategori Aspek Wisata/Hiburan

- b. Setelah menekan pilihan **wisata/hiburan**, maka akan muncul tampilan jumlah sentimen positif dan sentimen negatif untuk kategori wisata/hiburan seperti pada gambar berikut :



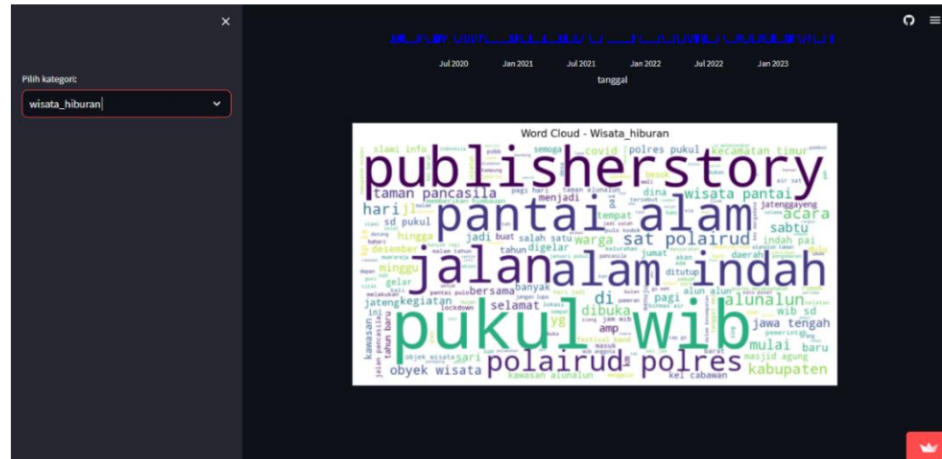
Gambar 6. Jumlah Sentimen Aspek Wsata/Hiburan

- c. Tekan tombol **“Tanpa Filter”** yang ada pada halaman dashboard aspek wisata/hiburan.
- d. Setelah menekan tombol **Tanpa Filter** maka akan muncul tampilan persebaran data sentimen positif dan sentimen negatif pada kategori aspek wisata/hiburan seperti pada gambar berikut :

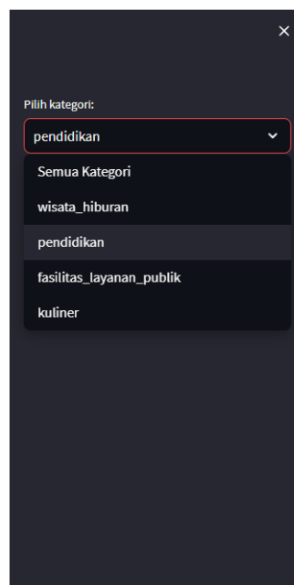


Gambar 7. Aspek Wisata/Hiburan Tanpa Filter

- e. Kemudian dashboard akan menampilkan **Wordcloud** pada aspek wisata/hiburan seperti pada gambar berikut :



f. Tekan salah satu kategori aspek **pendidikan** seperti pada gambar berikut :

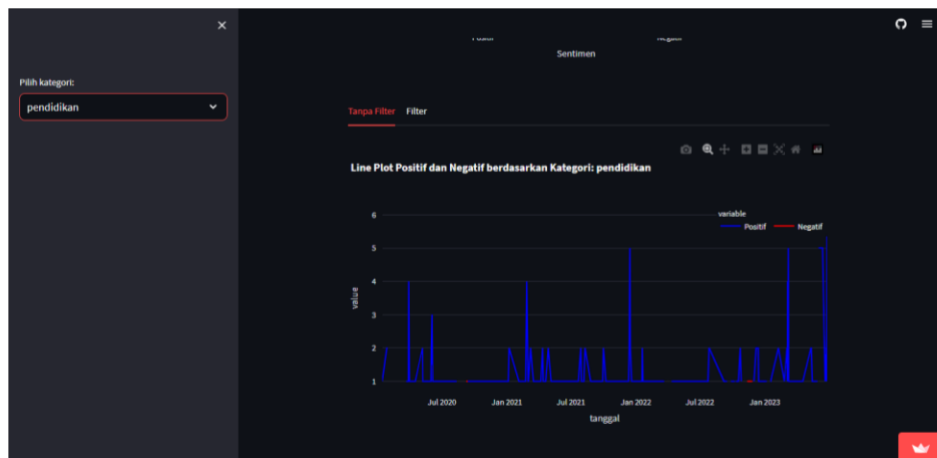


g. Setelah menekan pilihan “**pendidikan**”, maka akan muncul tampilan jumlah sentimen positif dan sentimen negatif untuk kategori pendidikan seperti pada gambar berikut:



Gambar 10. Jumlah Sentimen Aspek Pendidikan

- h. Tekan tombol **“Tanpa Filter”** yang ada pada halaman dashboard aspek pendidikan.
- i. Setelah menekan tombol **Tanpa Filter** maka akan muncul tampilan persebaran data sentimen positif dan sentimen negatif pada kategori aspek pendidikan seperti pada gambar berikut :



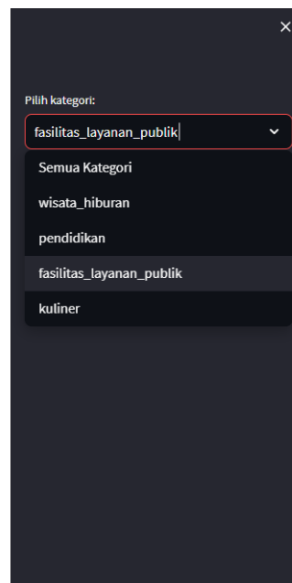
Gambar 11. Aspek Pendidikan Tanpa Filter

- j. Kemudian dashboard akan menampilkan **Wordcloud** pada aspek pendidikan seperti pada gambar berikut :



Gambar 12. WordCloud Aspek Pendidikan

- k. Tekan salah satu kategori aspek **“fasilitas/layanan publik”** seperti pada gambar berikut :



Gambar 13. Pilihan Aspek Fasilitas/Layanan Publik

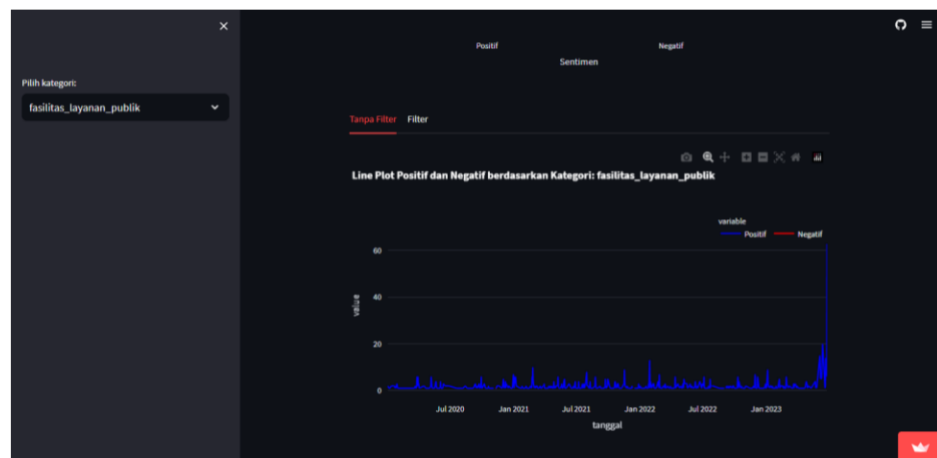
- l. Setelah menekan pilihan **fasilitas/layanan publik**, maka akan muncul tampilan jumlah sentimen positif dan sentimen negatif untuk kategori fasilitas/layanan publik seperti pada gambar berikut :





Gambar 14. Jumlah Sentimen Aspek Fasilitas/Layanan Publik

- m. Tekan tombol **Tanpa Filter** yang ada pada halaman dashboard aspek fasilitas/layanan publik.
- n. Setelah menekan tombol **Tanpa Filter** maka akan muncul tampilan persebaran data sentimen positif dan sentimen negatif pada kategori aspek fasilitas/layanan publik seperti pada gambar berikut :

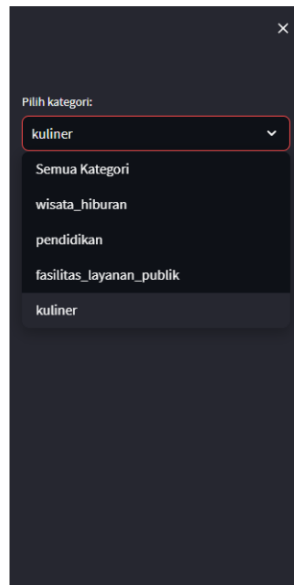


Gambar 15. Aspek Fasilitas/Layanan Publik Tanpa Filter

- o. Kemudian dashboard akan menampilkan **Wordcloud** pada aspek fasilitas/layanan publik seperti pada gambar berikut :



- p. Tekan salah satu kategori aspek “**kuliner**” seperti pada gambar berikut :

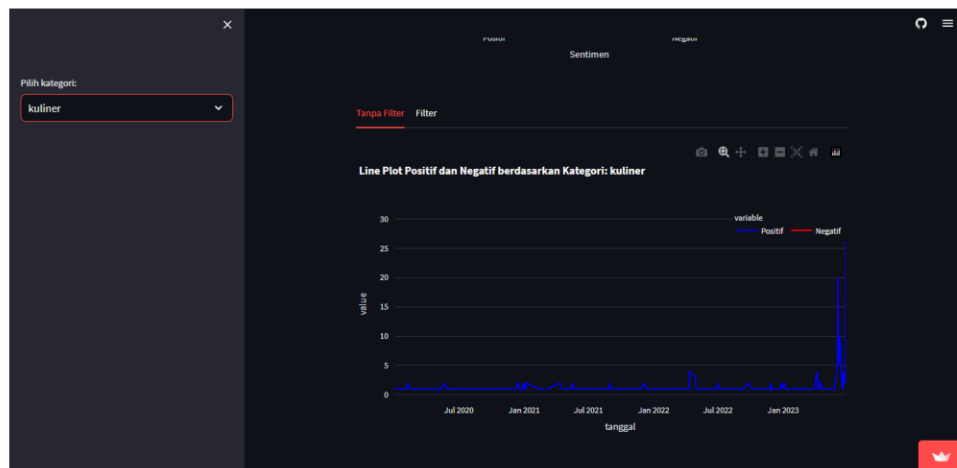


- q. Setelah menekan pilihan **kuliner**, maka akan muncul tampilan jumlah sentimen positif dan sentimen negatif untuk kategori kuliner seperti pada gambar berikut :



Gambar 18. Jumlah Sentimen Aspek Kuliner

- r. Tekan tombol **“Tanpa Filter”** yang ada pada halaman dashboard aspek fasilitas/layanan publik.
- s. Setelah menekan tombol **Tanpa Filter** maka akan muncul tampilan persebaran data sentimen positif dan sentimen negatif pada kategori aspek kuliner seperti pada gambar berikut :

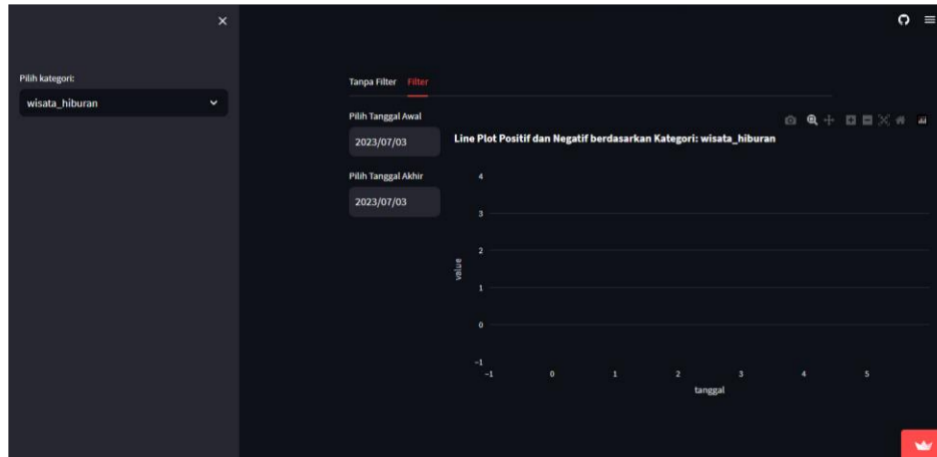


Gambar 19. Aspek Kuliner Tanpa Filter

- [illegible]

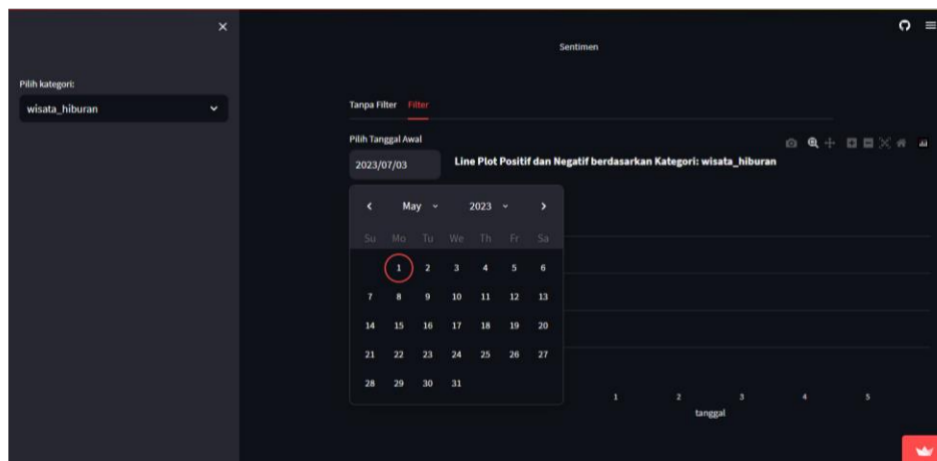
### 3. Persebaran Data Dengan Filter

- D-15



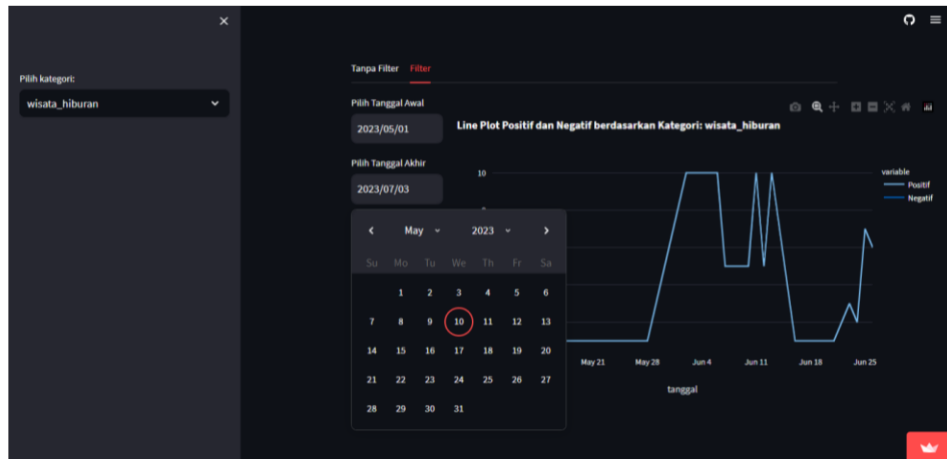
Gambar 21. Pilihan Filter Aspek Wisata/Hiburan

- d. Pilih **Tanggal Awal** pada dashboard seperti pada gambar berikut :



Gambar 22. Pilih Tanggal Awal Aspek Wisata/Hiburan

- e. Pilih **Tanggal Akhir** pada dashboard seperti pada gambar berikut :



Gambar 23. Pilih Tanggal Akhir Aspek Wisata/Hiburan

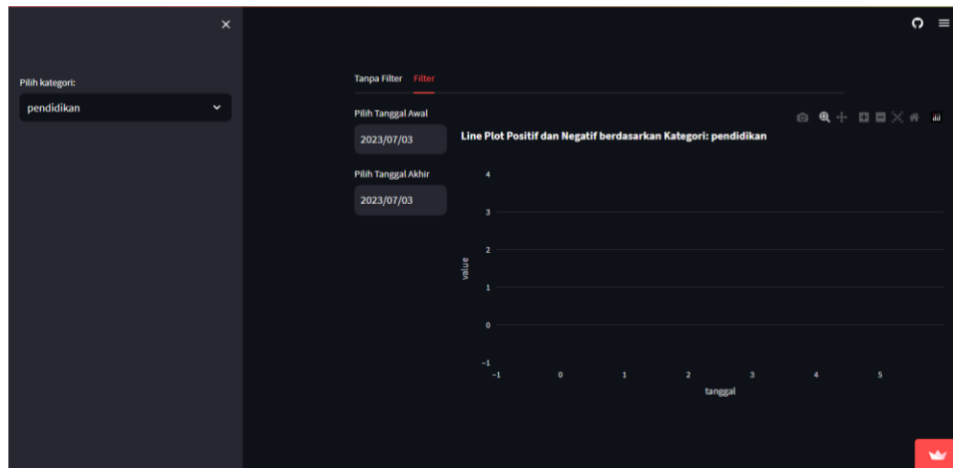
- f. Setelah menentukan **Tanggal Awal** dan **Tanggal Akhir** maka pada tampilan dashboard akan menampilkan hasil persebaran data pada aspek wisata/hiburan seperti pada gambar berikut :



Gambar 24. Aspek Wisata/Hiburan dengan Filter

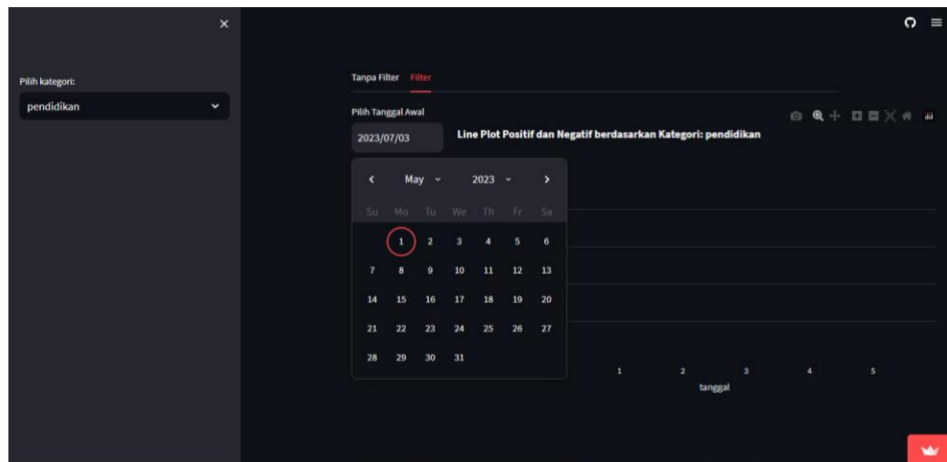
- g. Setelah menampilkan jumlah sentimen positif, sentimen negatif, dan juga persebaran data pada aspek pendidikan, pengguna bisa melihat persebaran data pada aspek pendidikan dengan filter tanggal yang sudah disediakan.

- h. Tekan tombol **“Filter”** yang ada pada halaman aspek pendidikan.
- i. Setelah menekan tombol **Filter** maka akan muncul tampilan seperti pada gambar berikut:



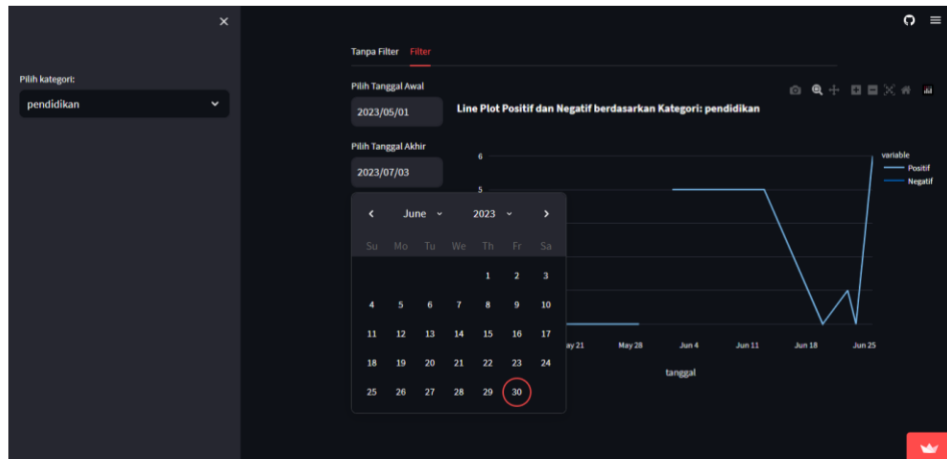
Gambar 25. Pilihan Filter Aspek Pendidikan

- j. Pilih **Tanggal Awal** pada dashboard seperti pada gambar berikut :



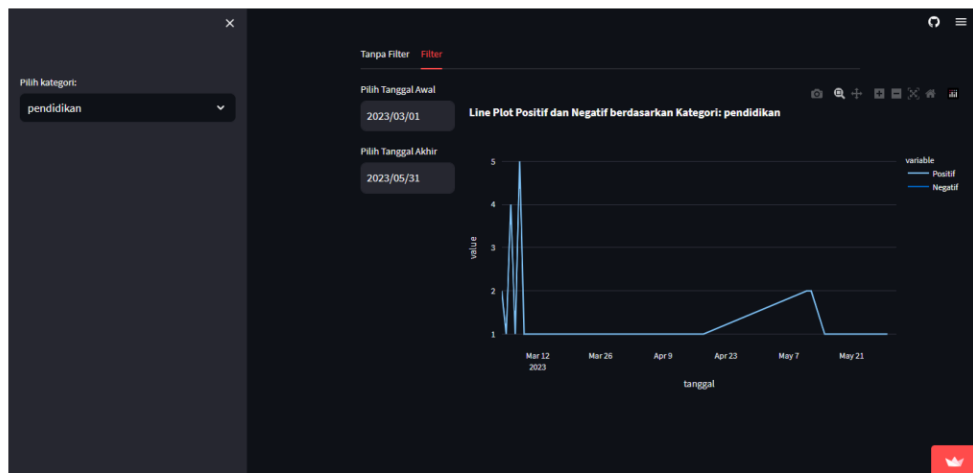
Gambar 26. Pilih Tanggal Awal Aspek Pendidikan

k. Pilih **Tanggal Akhir** pada dashboard seperti pada gambar berikut :



Gambar 27. Pilih Tanggal Akhir Aspek Pendidikan

l. Setelah menentukan **Tanggal Awal** dan **Tanggal Akhir** maka pada tampilan dashboard akan menampilkan hasil persebaran data pada aspek pendidikan seperti pada gambar berikut :



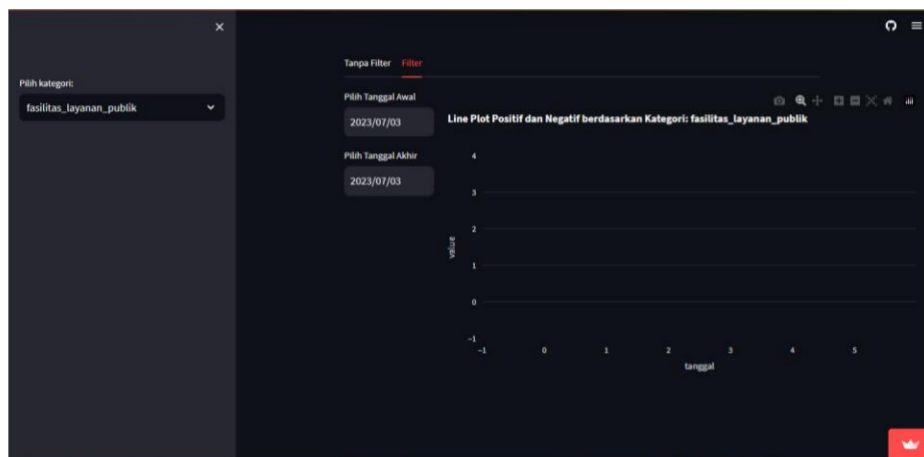
Gambar 28. Aspek Pendidikan Dengan Filter

m. Setelah menampilkan jumlah sentimen positif, sentimen negatif, dan juga persebaran data pada aspek fasilitas/layanan publik, pengguna bisa melihat



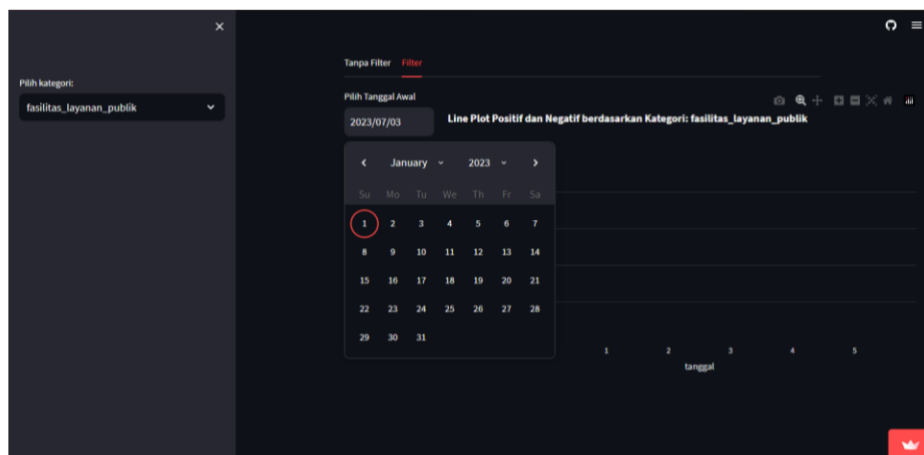
persebaran data pada aspek fasilitas/layanan publik dengan filter tanggal yang sudah disediakan.

- n. Tekan tombol **“Filter”** yang ada pada halaman aspek fasilitas/layanan publik.
- o. Setelah menekan tombol **Filter** maka akan muncul tampilan seperti pada gambar berikut:



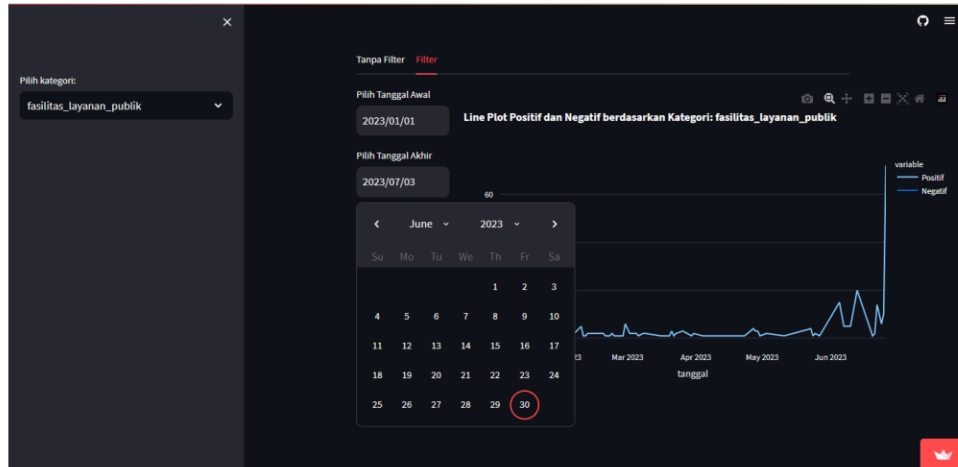
Gambar 29. Pilihan Filter Aspek Fasilitas/Layanan Publik

- p. Pilih **Tanggal Awal** pada dashboard seperti pada gambar berikut :



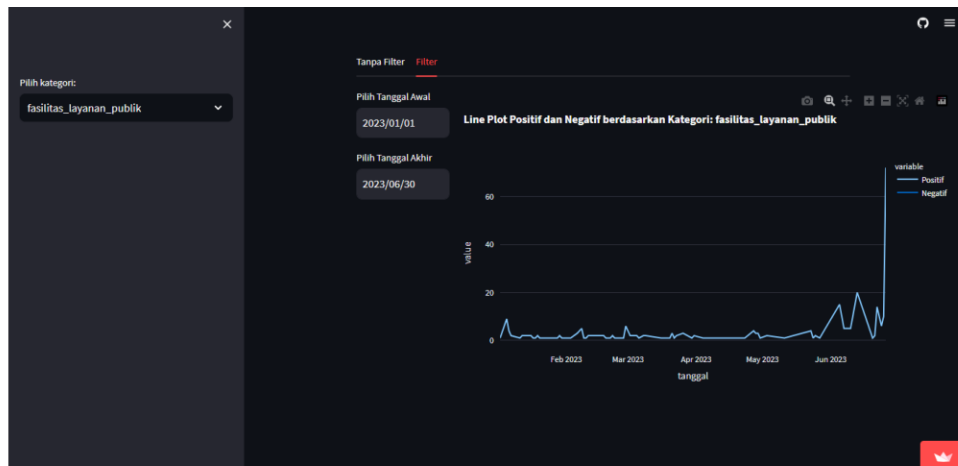
Gambar 30. Pilih Tanggal Awal Aspek Fasilitas/Layanan Publik

- q. Pilih **Tanggal Akhir** pada dashboard seperti pada gambar berikut :



Gambar 31. Pilih Tanggal Akhir Aspek Fasilitas/Layanan Publik

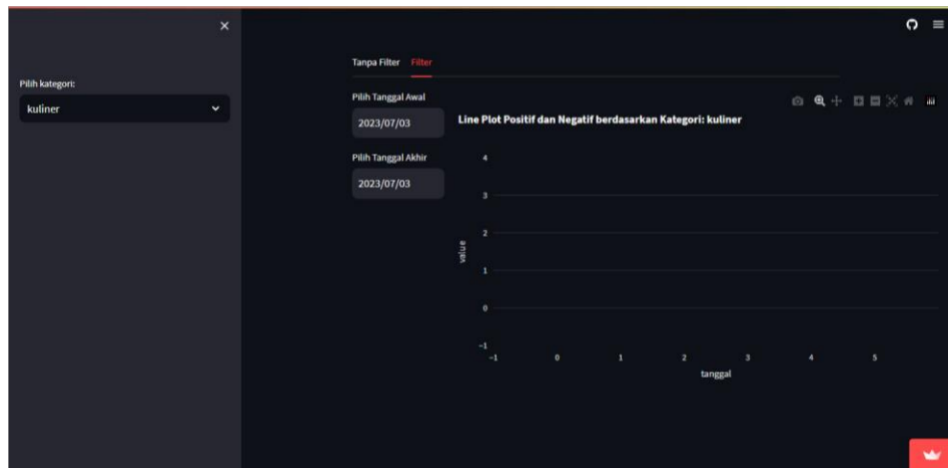
- r. Setelah menentukan **Tanggal Awal** dan **Tanggal Akhir** maka pada tampilan dashboard akan menampilkan hasil persebaran data pada aspek fasilitas/layanan publik seperti pada gambar berikut :



Gambar 32. Aspek Fasilitas/Layanan Publik dengan Filter

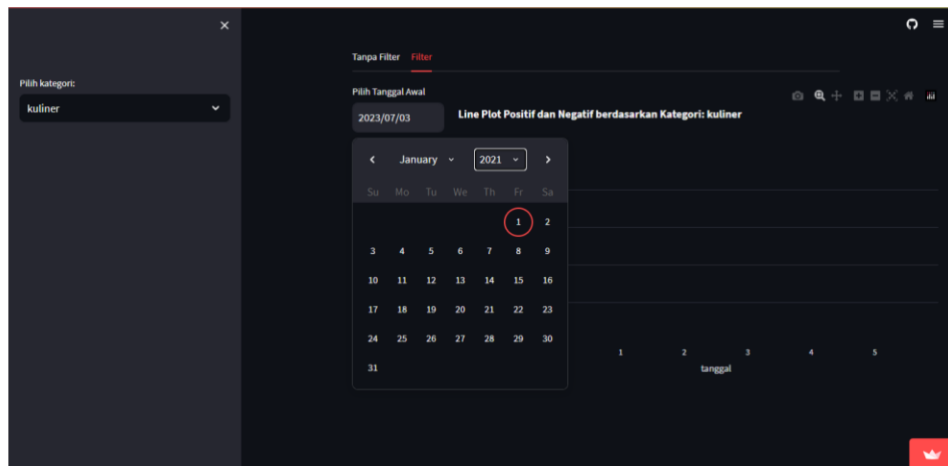
- s. Setelah menampilkan jumlah sentimen positif, sentimen negatif, dan juga persebaran data pada aspek kuliner, pengguna bisa melihat persebaran data pada aspek kuliner dengan filter tanggal yang sudah disediakan.

- t. Tekan tombol **“Filter”** yang ada pada halaman aspek kuliner.
- u. Setelah menekan tombol **Filter** maka akan muncul tampilan seperti pada gambar berikut:



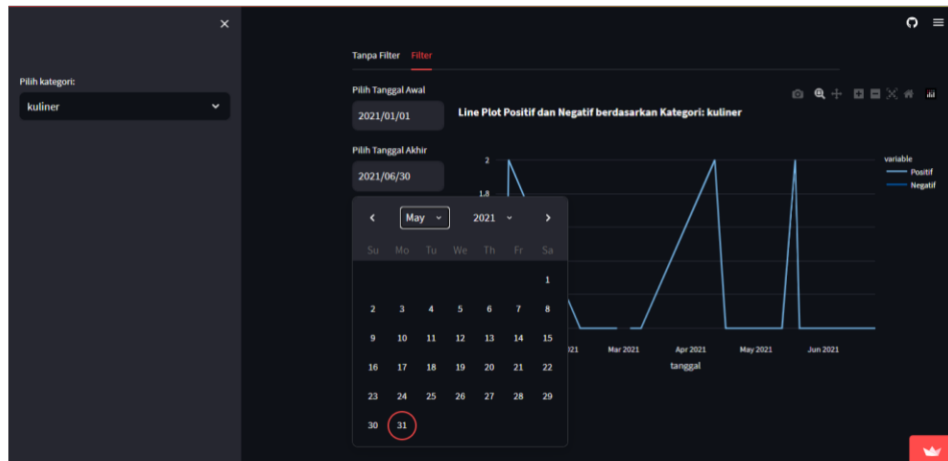
Gambar 33. Pilihan Filter Aspek Kuliner

- v. Pilih **Tanggal Awal** pada dashboard seperti pada gambar berikut :



Gambar 34. Pilih Tanggal Awal Aspek Kuliner

w. Pilih **Tanggal Akhir** pada dashboard seperti pada gambar berikut :



Gambar 35. Pilih Tanggal Akhir Aspek Kuliner

x. Setelah menentukan **Tanggal Awal** dan **Tanggal Akhir** maka pada tampilan dashboard akan menampilkan hasil persebaran data pada aspek kuliner seperti pada gambar berikut :



Gambar 36. Aspek Kuliner dengan Filte

## Lampiran 5 Dokumen Teknikal



## **Profil**

Dashboard Sentimen Analisis Kota Tegal ini dirancang untuk membantu pemerintah Kota Tegal dalam mengklasifikasikan opini atau ulasan masyarakat Kota Tegal di *Twitter* secara cepat dan efektif. Dimana opini atau ulasan masyarakat Kota Tegal ini di klasifikasikan kedalam beberapa aspek atau kategori ulasan seperti; Aspek Wisata/Hiburan, Aspek Pendidikan, Aspek Fasilitas/Layanan Publik, dan Aspek Kuliner. Pada Aplikasi Dashboard ini sudah dilengkapi dengan fitur untuk menentukan filter tanggal untuk memudahkan pengguna membaca data serta tampilan dashboard yang interaktif dengan fitur persebaran data yang rinci.

## **Latar Belakang**

Kota Tegal adalah salah satu kota di Jawa Tengah yang memiliki banyak potensi yang menarik untuk dikembangkan guna meningkatkan kualitas dari Kota Tegal itu sendiri. Sejalan dengan hal tersebut, Walikota Tegal telah membuat keputusan untuk menyediakan alamat *E-mail* resmi pemerintah kota sebagai sarana untuk memperkuat hubungan antara masyarakat dan instansi pemerintah. Tujuan dari keputusan tersebut adalah untuk memberikan pelayanan yang efektif dan terbuka kepada masyarakat dengan memanfaatkan teknologi informasi. Namun, hingga saat ini, masih belum terdapat ketentuan yang jelas mengenai standar operasional penggunaan *E-mail* oleh masyarakat ketika membutuhkan informasi, memberikan saran atau pendapat, serta menyampaikan pengaduan kepada pemerintah, masyarakat Kota Tegal lebih sering menyampaikan pendapat dan keluhannya melalui media sosial yang ada, salah satunya adalah *twitter*.

Media sosial merupakan teknologi yang memiliki peran penting untuk tanggap terhadap beberapa situasi, terlebih lagi karena teknologi sendiri adalah perangkat yang memungkinkan untuk bertukar informasi secara terbuka dan relatif cepat. Media sosial memudahkan pemerintah untuk menyebarkan informasi mengenai pemerintahan, serta tentu saja mengumpulkan umpan balik dari opini publik mengenai kebijakan daerah. Karena banyaknya masyarakat Kota Tegal yang memanfaatkan media sosial sebagai sarana untuk menyalurkan pendapat dan juga keluhan seputar Kota Tegal, maka

komentar masyarakat tersebut memberikan manfaat kepada pemerintah untuk dijadikan sebagai alat bantu untuk mengambil keputusan dalam meningkatkan kualitas Kota Tegal. Ulasan dari twitter ini merupakan ulasan masyarakat tentang Kota Tegal yang bersifat positif atau negatif.

Namun karena banyaknya opini masyarakat di *twitter* membuat pemerintah kesulitan untuk memberikan tanggapan tentang pendapat atau keluhan yang ada. Karena pengembangan kualitas Kota Tegal memberikan peranan penting dalam banyak hal, yaitu meningkatkan ekonomi dan meningkatkan kualitas dari fasilitas dan juga pelayanan publik, maka diperlukan klasifikasi tentang ulasan dari masyarakat yang diklasifikasikan kedalam beberapa aspek sehingga pemerintah bisa meninjau lebih dalam tentang keluhan yang ada dan lebih spesifik.

### **Manfaat**

Adapun manfaat yang dari aplikasi TourIn diuraikan sebagai berikut:

1. Pengelola Kota Tegal dapat memperoleh informasi tentang respon masyarakat terhadap aspek-aspek di Kota Tegal, seperti wisata, pendidikan, fasilitas publik, dan kuliner.
2. Aplikasi ini diharapkan dapat meningkatkan kualitas Kota Tegal secara keseluruhan dengan memperhatikan preferensi dan kebutuhan masyarakat

### **Spesifikasi Teknis**

Spesifikasi teknis meliputi :

- Source Code

Berikut uraian spesifikasi yang digunakan untuk membangun aplikasi :

- **Perangkat Keras**

1. Komputer/Laptop dengan spesifikasi processor Intel Pentium Gold 5405U, 256GB SSD, 4GB RAM


- **Perangkat Lunak**

1. Sistem operasi Windows 10
2. Visual Studio Code
3. Xampp
4. Python

Berikut uraian spesifikasi modul

1. Source Code

- **Crawling Dataset Twitter**



```
1 import twint
2 import nest_asyncio
3 nest_asyncio.apply()
4
5 def column_names():
6     return twint.output.panda.Tweets_df.columns
7
8 def twint_to_pd(columns):
9     return twint.output.panda.Tweets_df[columns]
10
11 c = twint.Config()
12 c.Since = "2022-01-01"
13 c.Until = "2022-12-10"
14 c.Search = "Kota Tegal"
15 c.Pandas = True
16
17 twint.run.Search(c)
18
19 hasil_scraping = twint_to_pd(['date', 'username', 'tweet'])
20 hasil_scraping.tail()
21
22 hasil_scraping.to_csv("datasetkota_tegal.csv", index=False)
```



	A	B	C	D	E	F
1	tweet	wisata/hil pendidik fasilitas/ku kuliner				
2	Widusa Universitas BSI Kampus Kabupaten Banyuwangi ke-16 & Kampu Kota Tegal ke-4 (06/12/2021) @ Java Heritage	-1	1	-1	-1	
3	(03/12/2021) Mengikuti kegiatan bersih pantai yang diselenggarakan Lanal Tegal di Pantai Alam Indah Kota Tegal. Kegiatan	-1	-1	1	-1	
4	@infopneltgalel Selamat malam ,mohon di tindak lanjuti di daerah saya dari jam 17.30 WIB mulai listrik sampai sekarang bi	-1	-1	0	-1	
5	[TF] Info banjir daerah tegal kota dong lurr, kayong pgn sebak me, makasii	-1	-1	-1	-1	
6	Pemasangan portal di jalan masuk Alun-alun Kota Tegal diprotes pedagang. Wali Kota Tegal Deddy Yon Suryiono membe	-1	-1	0	-1	
7	Dalam rangka memperingati Hari Juang Kartika (HKJ) TNI-AD tahun 2021 Kodim 0734/Kota Yogyakarta gelar kegiatan bhal	-1	-1	-1	-1	
8	Perusahaan di Kota Tegal wajib membayar upah sesuai Upah Minimum Rp 2.005.930,52 utk pekerja dengan masa kerja ki	-1	-1	-1	-1	
9	PPKM Level 3 Bakti, Akhirnya Tegal Nyalakan Jal Lampu-lampu Jalan di Kota #JawaTengah #KotaTegal https://t.co/ku3U	-1	-1	1	-1	
10	Askrindo memberikan pembinaan dan bantuan Tanggung Jawab Sosial dan Lingkungan berupa sarana produksi kepada t	-1	-1	-1	-1	
11	Pembukaan Pelatihan Jabatan Fungsional Perawat Angkatan II (9 - 16 Desember 2021) dan Pelatihan Jabatan Fungsional	-1	1	-1	-1	
12	Live Report Kasatlantas Polres Tegal Kota melaporkan pantauan Kamseltinclarlantas dari Simpang Pramesthi Kota Tegal.	-1	-1	-1	-1	
13	Aplikasi itu bertujuan memudahkan dan mendekatkan pelayanan PN Tegal dengan masyarakat, khususnya warga Kota T	-1	-1	-1	-1	
14	Waktunya kulineran sate kambing muda di Sate Tegal Marem Jal kuyy! Satenya terbuat dari daging kambing muda pilih	-1	-1	-1	-1	
15	Ditentang Warga, Wali Kota Tegal Ngotot Tutup Jalan #protes #masyarakat #penutupan #jalan #tegal #alun #alun #prote	-1	-1	0	-1	
16	Warga Kota Tegal Sekarang Bisa Mencetak Adminiduk Mandiri https://t.co/2QF6BX6t	-1	-1	1	-1	
17	Wali Kota Tegal Deddy Yon Suryiono mengungkapkan alasan kebijakan menutup kawasan Alun-alun dan Jalan Pancasila	-1	-1	-1	-1	
18	Sejak dipadamkan mulai 1 Desember 2021, lampu penerangan jalan umum (PJU) di Kota Tegal akhirnya kembali dinyalakan	-1	-1	-1	-1	
19	Meski diprotes warga dan pedagang, pemortalan akses jalan menuju kawasan Alun-alun dan Jalan Pancasila Kota Tegal t	-1	-1	0	-1	
20	Jalan Dr. Soetomo Kota Tegal. Pemadamn Lampu Jalan oleh @PemkoTegal Mohon Pak @ganjarpranowo kasih solusi i	-1	-1	0	-1	

Kode di atas adalah sebuah skrip Python yang menggunakan library Twint untuk mengumpulkan data tweet dari Twitter berdasarkan kriteria tertentu, kemudian menyimpan hasilnya dalam format CSV.

- Preprocessing Model Aspek

F-5

- Hasil Preprocessing Casefolding

```
... 0  wisuda universitas bsi kampus kabupaten banyum...
    1  (03/12/2021) mengikuti kegiatan bersih pantai...
    2  @infoplntegal selamat malam , mohon di tindak ...
    3  [tf] info banjir daerah tegal kota dong lurr, ...
    4  pemasangan portal di jalan masuk alun-alun kot...
    Name: tweet, dtype: object
```

- Hasil Preprocessing Tokenisasi

```
... [['wisuda',
      'universitas',
      'bsi',
      'kampus',
      'kabupaten',
      'banyumas',
      'ke',
      'amp',
      'kampus',
      'kota',
      'tegal',
      'ke',
      'java',
      'heritage',
      'hotel',
      'purwokerto'],
     ['mengikuti',
      'kegiatan',
```

Penjelasan :

Kode di atas merupakan bagian dari proses pra-pemrosesan data pada teks tweet yang telah diambil dan disimpan dalam variabel `data_aspek`. Langkah-langkah yang dilakukan dalam proses tersebut mencakup penyimpanan konten tweet dalam variabel `data_content`, kemudian dilanjutkan dengan proses casefolding untuk mengubah semua huruf menjadi huruf kecil dalam variabel `data_casefolding`. Tahap selanjutnya adalah filtering yang melibatkan penghapusan URL, menggantikan "(cont)" dengan spasi, menghapus tanda baca, menghapus tagger (kata

yang diawali dengan "#"), dan menghapus angka dari teks tweet. Semua tahap ini menghasilkan teks tweet yang lebih bersih dan siap untuk proses lebih lanjut.

#### - Preprocessing 2

```
26 # #tokenize
27 tknzs = TweetTokenizer()
28 data_tokenize = [tknzs.tokenize(tweet) for tweet in data_filtering]
29 data_tokenize
30
31 #slang word
32 path_datalang = open("kamus kata baku-clear (1).csv")
33 datalang = pd.read_csv(path_datalang, encoding = 'utf-8', header=None, sep=";")
34
35 def replaceSlang(word):
36     if word in list(datalang[0]):
37         indexslang = list(datalang[0]).index(word)
38         return datalang[1][indexslang]
39     else:
40         return word
41
42 data_formal = []
43 for data in data_tokenize:
44     data_clean = [replaceSlang(word) for word in data]
45     data_formal.append(data_clean)
46 len_data_formal = len(data_formal)
47
48 nltk.download('stopwords')
49 default_stop_words = nltk.corpus.stopwords.words('indonesian')
50 stopwords = set(default_stop_words)
51
52 def removeStopWords(line, stopwords):
53     words = []
54     for word in line:
55         word=str(word)
56         word = word.strip()
57         if word not in stopwords and word != "" and word != "&":
58             words.append(word)
59
60     return words
61 reviews = [removeStopWords(line,stopwords) for line in data_formal]
```

#### - Hasil Prerprocessing Slangword



Selanjutnya, kata-kata slang dalam teks tweet digantikan dengan kata-kata formal atau baku. Setelah menggantikan kata slang, langkah berikutnya adalah menghapus stop words. Stop words merupakan kata-kata umum yang biasanya tidak memberikan kontribusi signifikan dalam pemahaman makna teks dan sering dihilangkan dalam proses pra-pemrosesan teks. Setelah proses tersebut selesai, teks tweet telah mengalami tokenisasi, penggantian kata slang menjadi kata baku, dan penghapusan stop words.

#### - Modeling Aspek Wisata/Hiburan

```

1 # pembuatan vector kata
2 vectorizer = TfidfVectorizer()
3 reviews2 = [" ".join(x) for x in reviews]
4 vektor_tfidf = vectorizer.fit_transform(reviews2)
5 vektor_tfidf = vektor_tfidf.toarray()
6 vektor_tfidf.shape
7
8 #modeling aspek Wisata/Hiburan
9
10 #Panggil Data target aspek (y)
11 label_aspek_1 = data_aspek['wisata/hiburan']
12 label_aspek_1.shape
13 X_train, X_test, y_train, y_test = train_test_split(vektor_tfidf, label_aspek_1, test_size=0.3, shuffle=True)
14 nama = "Model_Aspek_Wisata_Hiburan"
15 y_train = np.array(y_train)
16 y_test = np.array(y_test)
17
18 # contoh menggunakan RandomOverSampler
19 from imblearn.over_sampling import RandomOverSampler
20 ros = RandomOverSampler(random_state=0)
21 X_resampled, y_resampled = ros.fit_resample(X_train, y_train)
22 from collections import Counter
23 print(sorted(Counter(y_resampled).items()))
24
25 from sklearn import metrics
26 clf = MultinomialNB()
27 clf.fit(X_resampled, y_resampled)
28 y_pred = clf.predict(X_resampled)
29 print("Akurasi : " + str(clf.score(X_resampled, y_resampled)))
30
31 print("Precision : " + str(metrics.precision_score(y_resampled, y_pred, labels=[-1, 0, 1], average='macro')))
32
33 # save model
34 # model_nvb = 'tfidf_sentimen_integritas_nvb.sav'
35 pickle.dump(clf, open('model/tfidf_{}.nvb.pkl'.format(nama), 'wb'))
36
37 #pengujian
38 model = pickle.load(open('model/tfidf_'+str(nama)+'_nvb.pkl', 'rb'))
39 pred_y_ = model.predict(X_test)
40 label_test = y_test
41 # result = model.score(X_test, y_test)
42 print('\n'+ '-'*45)
43 print('confusion matrix : ' + str(nama))
44 print('-'*45, '\n')
45 print(classification_report(label_test, pred_y_, labels=[-1, 0, 1]))

```

- Confusion Matrix Aspek Wisata/Hiburan

```

...
-----
confusion matrix : Model_Aspek_Wisata_Hiburan
-----

```

	precision	recall	f1-score	support
-1	0.99	0.86	0.92	880
0	0.41	0.97	0.57	29
1	0.24	0.68	0.35	38
accuracy			0.86	947
macro avg	0.54	0.84	0.61	947
weighted avg	0.94	0.86	0.89	947

Penjelasan :

Kode di atas merupakan proses pembuatan model klasifikasi untuk memprediksi kelas target berdasarkan teks tweet yang telah diolah menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Tahap awal adalah melakukan pengolahan teks dengan langkah-langkah seperti case folding, filtering untuk menghapus URL, tanda baca, dan tagger, serta penghilangan kata numerik. Selanjutnya, dilakukan tokenisasi menggunakan TweetTokenizer untuk membagi setiap tweet menjadi token-token kata. Kemudian, dilakukan penanganan kata slang dengan menggantikan kata slang menjadi kata baku sesuai dengan kamus kata baku yang telah disiapkan. Setelah itu, dilakukan pembersihan data dari kata-kata yang tidak relevan dengan menghapus stop words berbahasa Indonesia. Hasil dari proses ini adalah data teks yang telah diolah dan siap digunakan dalam pembuatan model. Langkah berikutnya adalah pembuatan vector kata menggunakan metode TF-IDF. Data teks yang telah diolah sebelumnya dikonversi menjadi bentuk vektor kata, di mana masing-masing kata diwakili oleh nilai TF-IDF-nya.

## - Modeling Aspek Pendidikan

```

1 # pembuatan vector kata
2 vectorizer = TfidfVectorizer()
3 reviews2 = [" ".join(x) for x in reviews]
4 vektor_tfidf = vectorizer.fit_transform(reviews2)
5 vektor_tfidf = vektor_tfidf.toarray()
6 vektor_tfidf.shape
7
8 #modeling aspek Pendidikan
9
10 #Panggil Data target aspek (y)
11 label_aspek_2 = data_aspek['pendidikan']
12 X_train, X_test, y_train, y_test = train_test_split(vektor_tfidf, label_aspek_2, test_size=0.3, shuffle=True)
13 nama = "Model_Aspek_Wisata_Pendidikan"
14 y_train = np.array(y_train)
15 y_test = np.array(y_test)
16 # contoh menggunakan RandomOverSampler
17 from imblearn.over_sampling import RandomOverSampler
18 ros = RandomOverSampler(random_state=0)
19 X_resampled, y_resampled = ros.fit_resample(X_train, y_train)
20 from collections import Counter
21 print(sorted(Counter(y_resampled).items()))
22 from sklearn import metrics
23 clf = MultinomialNB()
24 clf.fit(X_resampled, y_resampled)
25 y_pred = clf.predict(X_test)
26 print("Akurasi : " + str(clf.score(X_test, y_test)))
27 print("Precision : " + str(metrics.precision_score(y_test, y_pred, labels=[-1, 0, 1], average='macro'))))
28
29 # save model
30 model_nvb = 'tfidf_sentimen_integritas_nvb.sav'
31 pickle.dump(clf, open('model/tfidf_{}_nvb.pkl'.format(nama), 'wb'))
32 model = pickle.load(open('model/tfidf_{}_nvb.pkl'.format(nama), 'rb'))
33 pred_y_ = model.predict(X_test)
34 label_test = y_test
35 # result = model.score(X_test, y_test)
36 print('\n' + '-' * 45)
37 print('confusion matrix : ' + str(model))
38 print('-' * 45 + '\n')
39 print(classification_report(label_test, pred_y_, labels=[-1, 0, 1]))

```

## - Hasil Akurasi Aspek Pendidikan

```

... [(-1, 2153), (0, 2153), (1, 2153)]
Akurasi : 0.9908654590493885
Precision : 0.9909989025663376

```

- Confusion Matrix Aspek Pendidikan

```

...
-----
confusion matrix : Model_Aspek_Wisata_Pendidikan
-----

```

	precision	recall	f1-score	support
-1	0.99	0.96	0.98	918
0	0.50	0.86	0.63	14
1	0.26	0.60	0.37	15
accuracy			0.95	947
macro avg	0.59	0.81	0.66	947
weighted avg	0.97	0.95	0.96	947

Penjelasan :

Kode di atas merupakan proses pembuatan model klasifikasi kedua, kali ini untuk memprediksi kelas target yang berhubungan dengan aspek "pendidikan" dari teks tweet yang telah diolah menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Pertama-tama, data aspek "pendidikan" diambil sebagai label target, dan kemudian dilakukan pemisahan data menjadi data latih dan data uji dengan proporsi 70% untuk data latih dan 30% untuk data uji. Kemudian, seperti pada proses sebelumnya, data latih yang memiliki ketidakseimbangan jumlah sampel pada kelas target diatasi dengan menggunakan metode Random Oversampling.



- Modeling Aspek Fasilitas/Layanan Publik

```

1 # pembuatan vector kata
2 vectorizer = TfidfVectorizer()
3 reviews2 = [" ".join(x) for x in reviews]
4 vektor_tfidf = vectorizer.fit_transform(reviews2)
5 vektor_tfidf = vektor_tfidf.toarray()
6 vektor_tfidf.shape
7
8 #modeling aspek Pendidikan
9
10 #Panggil Data target aspek (y)
11 label_aspek_3 = data_aspek['fasilitas/layanan publik']
12 X_train, X_test, y_train, y_test = train_test_split(vektor_tfidf, label_aspek_3, test_size=0.3)
13 nama = "Model_Aspek_Wisata_Fasilitas_Layanan_Publik"
14 y_train = np.array(y_train)
15 y_test = np.array(y_test)
16
17 # contoh menggunakan RandomOversampler
18 from imblearn.over_sampling import RandomOverSampler
19 ros = RandomOverSampler(random_state=0)
20 X_resampled, y_resampled = ros.fit_resample(X_train, y_train)
21 from collections import Counter
22 print(sorted(Counter(y_resampled).items()))
23 from sklearn import metrics
24 clf = MultinomialNB()
25 clf.fit(X_resampled, y_resampled)
26 y_pred = clf.predict(X_resampled)
27 print("Akurasi : "+str(clf.score(X_resampled, y_resampled)))
28
29 print("Precision : "+str(metrics.precision_score(y_resampled, y_pred, labels=[-1,0,1], average='macro'))))
30
31 # save model
32 # model_nvb = 'tfidf_sentimen_integritas_nvb.sav'
33 pickle.dump(clf, open('model/tfidf_{}.nvb.pkl'.format(nama), 'wb'))
34 model = pickle.load(open('model/tfidf_'+str(nama)+'_nvb.pkl', 'rb'))
35 pred_y_ = model.predict(X_test)
36 label_test = y_test
37 # result = model.score(X_test, y_test)
38 print('\n'+ '-'*45)
39 print('confusion matrix : ' + str(nama))
40 print('-'*45, '\n')
41 print(classification_report(label_test, pred_y_, labels=[-1, 0, 1]))
42 ## Aspek Fasilitas dan Layanan Publik

```

- Hasil Akurasi Aspek Fasilitas/Layanan Publik

```

... [(-1, 1867), (0, 1867), (1, 1867)]
Akurasi : 0.943045884663453
Precision : 0.9477187725186903

```

- Confusion Matrix Aspek Fasilitas/Layanan Publik

```

...
-----
confusion matrix : Model_Aspek_Wisata_Fasilitas_Layanan_Publik
-----

```

	precision	recall	f1-score	support
-1	0.96	0.73	0.83	781
0	0.48	0.87	0.62	77
1	0.29	0.67	0.40	89
accuracy			0.74	947
macro avg	0.57	0.76	0.62	947
weighted avg	0.85	0.74	0.77	947

Penjelasan :

Kode di atas merupakan proses pembuatan model klasifikasi ketiga, yang bertujuan untuk memprediksi kelas target yang berhubungan dengan aspek "fasilitas/layanan publik" dari teks tweet yang telah diolah menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Pertama-tama, data aspek "fasilitas/layanan publik" diambil sebagai label target, dan kemudian dilakukan pemisahan data menjadi data latih dan data uji dengan proporsi 70% untuk data latih dan 30% untuk data uji. Selanjutnya, model klasifikasi Naive Bayes dilatih dengan data latih yang telah di-resampling. Setelah model terlatih, dilakukan evaluasi performa model dengan menggunakan data uji. Hasil prediksi model dibandingkan dengan label sebenarnya untuk mengukur akurasi dan precision model. Terakhir, model yang telah dilatih dan dievaluasi disimpan dalam file menggunakan modul pickle, sehingga dapat digunakan kembali untuk memprediksi kelas target pada data teks baru dengan aspek "fasilitas/layanan publik".

## - Modeling Aspek Kuliner

```

1 # pembuatan vector kata
2 vectorizer = TfidfVectorizer()
3 reviews2 = [" ".join(r) for r in reviews]
4 vektor_tfidf = vectorizer.fit_transform(reviews2)
5 vektor_tfidf = vektor_tfidf.toarray()
6 vektor_tfidf.shape
7
8 #modeling aspek fasilitas/layanan publik
9
10 #Panggil Data target aspek (y)
11 label_aspek_4 = data_aspek['kuliner']
12 X_train, X_test, y_train, y_test = train_test_split(vektor_tfidf, label_aspek_4, test_size=0.3)
13 nama = "Model_Aspek_Wisata_Kuliner"
14 y_train = np.array(y_train)
15 y_test = np.array(y_test)
16 # contoh menggunakan RandomOversampler
17 from imblearn.over_sampling import RandomOverSampler
18 ros = RandomOverSampler(random_state=0)
19 X_resampled, y_resampled = ros.fit_resample(X_train, y_train)
20 from collections import Counter
21 print(sorted(Counter(y_resampled).items()))
22 from sklearn import metrics
23 clf = MultinomialNB()
24 clf.fit(X_resampled, y_resampled)
25 y_pred = clf.predict(X_resampled)
26 print("Akurasi : "+str(clf.score(X_resampled, y_resampled)))
27
28 print("Precision : "+str(metrics.precision_score(y_resampled, y_pred, labels=[-1,0,1], average='macro'))))
29
30 # save model
31 model_nvb = 'tfidf_sentimen_integritas_nvb.sav'
32 pickle.dump(clf, open('model/tfidf_'+nama+'.pkl', 'wb'))
33 model = pickle.load(open('model/tfidf_'+nama+'.pkl', 'rb'))
34 pred_y_ = model.predict(X_test)
35 label_test = y_test
36 # result = model.score(X_test, y_test)
37 print('\n'+'-'*45)
38 print('confusion matrix : ' + str(nama))
39 print('-'*45+'\n')
40 print(classification_report(label_test, pred_y_, labels=[-1, 0, 1]))
41

```

## - Hasil Akurasi Aspek Kuliner

```

... [(-1, 1964), (0, 1964), (1, 1964)]
Akurasi : 0.9589273591310251
Precision : 0.9615307439254952

```

- Confusion Matrix Aspek Kuliner

```

...
-----
confusion matrix : Model_Aspek_Wisata_Kuliner
-----

```

	precision	recall	f1-score	support
-1	1.00	0.77	0.87	849
0	0.45	0.98	0.62	45
1	0.26	0.92	0.40	53
accuracy			0.79	947
macro avg	0.57	0.89	0.63	947
weighted avg	0.93	0.79	0.83	947

Penjelasan :

Kode tersebut adalah langkah-langkah untuk membuat model klasifikasi keempat, yang bertujuan untuk memprediksi kelas target yang berhubungan dengan aspek "kuliner" dari teks tweet yang telah diolah menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Pertama-tama, data aspek "kuliner" diambil sebagai label target, dan kemudian dilakukan pemisahan data menjadi data latih dan data uji dengan proporsi 70% untuk data latih dan 30% untuk data uji. Kemudian, dilakukan Random Oversampling pada data latih untuk menyeimbangkan jumlah sampel pada kelas target. Metode Random Oversampling ini akan menambahkan data duplikat pada kelas yang jumlahnya kurang sehingga jumlah sampel pada setiap kelas menjadi seimbang. Selanjutnya, model klasifikasi Naive Bayes dilatih dengan data latih yang telah di-resampling. Setelah model terlatih, dilakukan evaluasi performa model dengan menggunakan data uji.

- Preprocessing 1 Model Sentimen

Penjelasan :

F-17

spasi kosong. Data setelah proses filtering disimpan dalam bentuk Pandas Series dengan nama "data\_filtering". Selanjutnya, dilakukan tokenisasi menggunakan TweetTokenizer dari library nltk. Tokenisasi adalah proses pemisahan teks menjadi token atau kata-kata. Setiap tweet dalam dataset dipecah menjadi token-token menggunakan tokenisasi, dan hasilnya disimpan dalam variabel "data\_tokenize" dalam bentuk daftar.

#### - Preprocessing 2 Model Sentimen

```
36 #slang word
37 path_dataslang = open("kamus kata baku-clear (1).csv")
38 dataslang = pd.read_csv(path_dataslang, encoding = 'utf-8', header=None, sep=";")
39
40 def replaceSlang(word):
41     if word in list(dataslang[0]):
42         indexslang = list(dataslang[0]).index(word)
43         return dataslang[1][indexslang]
44     else:
45         return word
46
47 data_formal = []
48 for data in data_tokenize:
49     data_clean = [replaceSlang(word) for word in data]
50     data_formal.append(data_clean)
51 len_data_formal = len(data_formal)
52 # print(data_formal)
53
54 #Stopword
55 nltk.download('stopwords')
56 default_stop_words = nltk.corpus.stopwords.words('indonesian')
57 stopwords = set(default_stop_words)
58
59 def removeStopWords(line, stopwords):
60     words = []
61     for word in line:
62         word=str(word)
63         word = word.strip()
64         if word not in stopwords and word != "" and word != "&":
65             words.append(word)
66
67     return words
68 reviews = [removeStopWords(line,stopwords) for line in data_formal]
```

#### Penjelasan :

Pada kode diatas dilakukan pembacaan kamus kata baku atau kamus slang word dari file "kamus kata baku-clear (1).csv". Kamus ini berisi pasangan kata slang dan kata formal yang sesuai. File tersebut dibaca menggunakan Pandas, dan hasilnya disimpan dalam DataFrame "dataslang". Selanjutnya, dibuat fungsi "replaceSlang" yang akan digunakan untuk menggantikan kata slang dengan kata formal berdasarkan kamus yang telah dibaca. Fungsi ini menerima input kata, memeriksa

apakah kata tersebut ada dalam kamus slang word, dan mengembalikan kata formal jika ditemukan, atau mengembalikan kata asli jika tidak ditemukan. Selanjutnya, dilakukan proses penggantian kata slang dalam data tokenisasi. Setiap kata dalam setiap tweet diubah menggunakan fungsi "replaceSlang", dan hasilnya disimpan dalam list "data\_formal". Setelah itu, dilakukan proses penghapusan stopwords atau kata-kata umum yang tidak memberikan kontribusi signifikan dalam analisis teks. Dalam kode tersebut, stopwords bahasa Indonesia dari library NLTK diunduh dan disimpan dalam variabel "default\_stop\_words". Kemudian, stopwords tersebut disimpan dalam set "stopwords".

## - Modeling Sentimen Aspek Wisata/Hiburan

```

1 # pembuatan vector kata
2 vectorizer = TfidfVectorizer()
3 reviews2 = [" ".join(x) for x in reviews]
4 vektor_tfidf = vectorizer.fit_transform(reviews2)
5 vektor_tfidf = vektor_tfidf.toarray()
6 vektor_tfidf.shape
7
8 class READ_DATA:
9     aspek = ''
10     vektor_tfidf = ''
11     data_sentimen_wisata_hiburan = ''
12     data_sentimen_pendidikan = ''
13     data_sentimen_fasilitas_layanan publik = ''
14     data_sentimen_kuliner = ''
15     # data_sentimen_kontinuitas = ''
16     def __init__(self, data_sentimen, vektor_tfidf):
17         self.data_sentimen_wisata_hiburan = data_sentimen['wisata_hiburan']
18         self.data_sentimen_pendidikan = data_sentimen['pendidikan']
19         self.data_sentimen_fasilitas_layanan publik = data_sentimen['fasilitas_layanan publik']
20         self.data_sentimen_kuliner = data_sentimen['kuliner']
21         self.vektor_tfidf = vektor_tfidf
22
23     def get_data_sentimen(self, aspek):
24         data_sentimen = getattr(self, f'data_sentimen_{aspek}')
25         data_sentimen_list = data_sentimen.to_list()
26         # print(type(data_sentimen_list[0]))
27
28         # filter hapus data yang tidak memiliki sentimen {-1}
29         x_, y_ = [], []
30         for i, j in zip(vektor_tfidf, data_sentimen_list):
31             if j == -1:
32                 x_.append(i)
33                 y_.append(j)
34         return x_, y_
35
36 # Ubah kategori
37 read_data = READ_DATA(data_sentimen, vektor_tfidf)
38 x_content_sentimen_NVB, y_label_sentimen_NVB = read_data.get_data_sentimen('wisata_hiburan')
39 from sklearn.utils.multiclass import unique_labels
40 unique_labels(y_label_sentimen_NVB)
41 X_train, X_test, y_train, y_test = train_test_split(x_content_sentimen_NVB, y_label_sentimen_NVB, test_size=0.2, shuffle=True)
42 nama = "Model_Sentimen_Wisata_Hiburan"
43 unique_labels(y_test)
44
45 from sklearn import metrics
46 clf = MultinomialNB()
47 clf.fit(X_train, y_train)
48 y_pred = clf.predict(X_test)
49 print("Akurasi : " + str(clf.score(X_train, y_train)))
50
51 print("Precision : " + str(metrics.precision_score(y_train, y_pred, labels=[0,1], average='macro')))
52
53 # save model
54 # model_nvb = 'vektor_tfidf_sentimen_integritas_nvb.sav'
55 pickle.dump(clf, open('model/vektor_tfidf_{}.nbv.pkl'.format(nama), 'wb'))
56 model = pickle.load(open('model/vektor_tfidf_{}nbv.pkl'.format(nama), 'rb'))
57 pred_y_ = model.predict(X_test)
58 label_test = y_test
59
60 # result = model.score(X_test, y_test)
61 print('\n' + '-' * 45)
62 print('confusion matrix : ' + str(nama))
63 print('-' * 45 + '\n')
64 print(classification_report(label_test, pred_y_, labels=[0, 1]))

```

### Penjelasan :

Proses di atas merupakan tahapan pembuatan model analisis sentimen menggunakan metode Multinomial Naive Bayes dan vektor TF-IDF pada data teks terkait wisata hiburan. Pertama, data teks yang sudah dipreproses dilakukan vektorisasi menggunakan metode TF-IDF. Hal ini dilakukan untuk mengubah teks menjadi vektor kata berdasarkan bobot kemunculan kata-kata dalam dokumen, sehingga teks dapat diolah dalam bentuk numerik. Kemudian, dibuat kelas READ\_DATA yang berfungsi



untuk mengelola data sentimen berdasarkan aspek wisata tertentu. Kelas ini memisahkan data teks dan label sentimen, dan juga melakukan filter untuk menghapus data yang tidak memiliki sentimen atau yang bernilai -1. Selanjutnya, data teks dan label sentimen tersebut dibagi menjadi data latih (X\_train dan y\_train) dan data uji (X\_test dan y\_test) dengan proporsi 80:20 menggunakan fungsi train\_test\_split. Model klasifikasi Multinomial Naive Bayes kemudian dilatih dengan menggunakan data latih (X\_train dan y\_train). Akurasi dan precision model dihitung untuk mengevaluasi performanya pada data latih. Hasil model yang sudah dilatih kemudian disimpan dalam bentuk file pkl menggunakan pickle untuk digunakan kembali di lain waktu. Setelah itu, model di-load kembali untuk melakukan prediksi pada data uji (X\_test) dan hasilnya dievaluasi menggunakan classification\_report yang memberikan informasi berupa precision, recall, dan f1-score untuk setiap kelas sentimen.

#### - Modeling Sentimen Aspek Pendidikan

```
1 # Ubah kategori
2 read_data = READ_DATA(data_sentimen, vektor_tfidf)
3 x_content_sentimen_NVb, y_label_sentimen_NVb = read_data.get_data_sentimen('pendidikan')
4
5 from sklearn.utils.multiclass import unique_labels
6 unique_labels(y_label_sentimen_NVb)
7
8 X_train, X_test, y_train, y_test = train_test_split(x_content_sentimen_NVb, y_label_sentimen_NVb, test_size=0.2)
9 nama = "Model_Sentimen_Pendidikan"
10 # y_train = np.array(y_train)
11 # y_test = np.array(y_test)
12
13 unique_labels(y_test)
14
15 from sklearn import metrics
16 clf = MultinomialNB()
17 clf.fit(X_train, y_train)
18 y_pred = clf.predict(X_test)
19 print("Akurasi : " + str(clf.score(X_train, y_train)))
20 print("Precision : " + str(metrics.precision_score(y_train, y_pred, labels=[0,1], average='macro'))))
21
22 # save model
23 # model_nv = 'vektor_tfidf_sentimen_integritas_nv.sav'
24 pickle.dump(clf, open('model/vektor_tfidf_{}_nv.pkl'.format(nama), 'wb'))
25
26 model = pickle.load(open('model/vektor_tfidf_{}_nv.pkl'.format(nama), 'rb'))
27 pred_y_ = model.predict(X_test)
28 label_test = y_test
29 # result = model.score(X_test, y_test)
30 print("\n" + "-" * 45)
31 print("confusion matrix : " + str(nama))
32 print("-" * 45 + "\n")
33 print(classification_report(label_test, pred_y_, labels=[0, 1]))
```

Penjelasan :

Proses selanjutnya adalah mengubah kategori dan melakukan analisis sentimen terhadap data teks terkait pendidikan. Langkah-langkahnya mirip dengan analisis sebelumnya. Pertama, data teks yang sudah dipreproses menggunakan vektorisasi TF-IDF dipisahkan berdasarkan aspek pendidikan, dan label sentimen untuk aspek ini diekstraksi. Selanjutnya, data dibagi menjadi data latih (`X_train` dan `y_train`) dan data uji (`X_test` dan `y_test`) dengan perbandingan 80:20 menggunakan fungsi `train_test_split`. Model klasifikasi Multinomial Naive Bayes kemudian dilatih dengan menggunakan data latih (`X_train` dan `y_train`). Akurasi dan precision model dihitung untuk mengevaluasi performanya pada data latih. Setelah itu, model yang sudah dilatih disimpan dalam bentuk file `pkl` menggunakan `pickle`. Model tersebut di-load kembali untuk melakukan prediksi pada data uji (`X_test`) dan hasilnya dievaluasi menggunakan `classification_report` yang memberikan informasi berupa precision, recall, dan f1-score untuk setiap kelas sentimen.

## - Modeling Sentimen Aspek Fasilitas/Layanan Publik

```
1 # Ubah kategori
2 read_data = READ_DATA(data_sentimen, vektor_tfidf)
3 x_content_sentimen_NVB, y_label_sentimen_NVB = read_data.get_data_sentimen('fasilitas_layanan_publik')
4
5 from sklearn.utils.multiclass import unique_labels
6 unique_labels(y_label_sentimen_NVB)
7
8 X_train, X_test, y_train, y_test = train_test_split(x_content_sentimen_NVB, y_label_sentimen_NVB, test_size=0.2)
9 nama = "Model_Sentimen_fasilitas_layanan_publik"
10 # y_train = np.array(y_train)
11 # y_test = np.array(y_test)
12
13 unique_labels(y_test)
14
15 from sklearn import metrics
16 clf = MultinomialNB()
17 clf.fit(X_train, y_train)
18 y_pred = clf.predict(X_test)
19 print("Akurasi : "+str(clf.score(X_train, y_train)))
20 print("Precision : "+str(metrics.precision_score(y_train, y_pred, labels=[0,1], average='macro'))))
21
22 # save model
23 # model_nvb = 'vektor_tfidf_sentimen_integritas_nvb.sav'
24 pickle.dump(clf, open('model/vektor_tfidf_'+nama+'.pkl', 'wb'))
25
26 model = pickle.load(open('model/vektor_tfidf_'+str(nama)+'_nvb.pkl', 'rb'))
27 pred_y_ = model.predict(X_test)
28 label_test = y_test
29 # result = model.score(X_test, y_test)
30 print("\n"+str(nama))
31 print('confusion matrix : ' + str(nama))
32 print("-"+45+"\n")
33 print(classification_report(label_test, pred_y_, labels=[0, 1]))
```

### Penjelasan :

Selanjutnya, kita akan mengubah kategori dan melakukan analisis sentimen terhadap data teks yang terkait dengan fasilitas layanan publik. Prosesnya hampir mirip dengan analisis sebelumnya. Pertama-tama, data teks yang sudah dipreproses menggunakan vektorisasi TF-IDF akan dipisahkan berdasarkan aspek fasilitas layanan publik, dan label sentimen untuk aspek ini akan diekstraksi. Kemudian, data akan dibagi menjadi data latih (X\_train dan y\_train) dan data uji (X\_test dan y\_test) dengan perbandingan 80:20 menggunakan fungsi train\_test\_split. Selanjutnya, kita akan menggunakan model klasifikasi Multinomial Naive Bayes untuk melatih data latih (X\_train dan y\_train). Setelah model dilatih, kita akan menghitung akurasi dan precision dari model tersebut pada data latih. Model yang sudah dilatih akan disimpan dalam format file pkl menggunakan pickle. Selanjutnya, model yang telah disimpan akan di-load

kembali untuk melakukan prediksi pada data uji ( $X_{test}$ ) dan hasilnya akan dievaluasi menggunakan `classification_report`. Laporan ini akan memberikan informasi berupa precision, recall, dan f1-score untuk setiap kelas sentimen.

#### - Modeling Sentimen Aspek Kuliner

```
1 # Ubah kategori
2 read_data = READ_DATA(data_sentimen, vektor_tfidf)
3 x_content_sentimen_NVb, y_label_sentimen_NVb = read_data.get_data_sentimen('kuliner')
4
5 from sklearn.utils.multiclass import unique_labels
6 unique_labels(y_label_sentimen_NVb)
7
8 X_train, X_test, y_train, y_test = train_test_split(x_content_sentimen_NVb, y_label_sentimen_NVb, test_size=0.2)
9 nama = "Model_Sentimen_Kuliner"
10 # y_train = np.array(y_train)
11 # y_test = np.array(y_test)
12
13 unique_labels(y_test)
14
15 from sklearn import metrics
16 clf = MultinomialNB()
17 clf.fit(X_train, y_train)
18 y_pred = clf.predict(X_test)
19 print("Akurasi : " + str(clf.score(X_train, y_train)))
20 print("Precision : " + str(metrics.precision_score(y_train, y_pred, labels=[0,1], average='macro'))))
21
22 # save model
23 # model_nv = 'vektor_tfidf_sentimen_integritas_nv.sav'
24 pickle.dump(clf, open('model/vektor_tfidf_{0}_nv.pkl'.format(nama), 'wb'))
25
26 model = pickle.load(open('model/vektor_tfidf_{0}_nv.pkl'.format(nama), 'rb'))
27 pred_y_ = model.predict(X_test)
28 label_test = y_test
29 # result = model.score(X_test, y_test)
30 print('\n' + '-' * 45)
31 print('confusion matrix : ' + str(nama))
32 print('-' * 45 + '\n')
33 print(classification_report(label_test, pred_y_, labels=[0, 1]))
```

Penjelasan :

Selanjutnya, kita akan melakukan analisis sentimen terhadap data teks yang berkaitan dengan aspek kuliner. Langkah-langkahnya hampir mirip dengan analisis sebelumnya. Pertama-tama, data teks yang sudah dipreproses menggunakan vektorisasi TF-IDF akan dipisahkan berdasarkan aspek kuliner, dan label sentimen untuk aspek ini akan diekstraksi. Selanjutnya, data akan dibagi menjadi data latih ( $X_{train}$  dan  $y_{train}$ ) dan data uji ( $X_{test}$  dan  $y_{test}$ ) dengan perbandingan 80:20 menggunakan fungsi `train_test_split`. Kemudian, kita akan menggunakan model klasifikasi Multinomial Naive Bayes untuk melatih data latih ( $X_{train}$  dan  $y_{train}$ ). Setelah model dilatih, kita akan menghitung akurasi

dan precision dari model tersebut pada data latih. Model yang sudah dilatih akan disimpan dalam format file pkl menggunakan pickle. Selanjutnya, model yang telah disimpan akan di-load kembali untuk melakukan prediksi pada data uji (X\_test) dan hasilnya akan dievaluasi menggunakan classification\_report. Laporan ini akan memberikan informasi berupa precision, recall, dan f1-score untuk setiap kelas sentimen.

#### Implementasi Model secara realtime

##### - Crawling Model Realtime

```
1  ## Crawling
2  today = date.today()
3  yesterday = today - timedelta(days = -1)
4  today=str(today)
5  yesterday = str(yesterday)
6  kata_kunci="Kota Tegal"
7
8  twitter_search = f"(kata_kunci) lang:id until:(today) since:(yesterday)"
9
10 # Tentukan nama file dengan format "kueri pencarian_tanggal saat ini.json"
11 filename = "Kota_Tegal.json"
12 USING_TOP_SEARCH = True
13
14 smscraps_params = "--jsonl"
15 twitter_search_params = ""
16
17 if USING_TOP_SEARCH:
18     twitter_search_params += "--top"
19
20 os.system(f"cmd /c smscraps {smscraps_params} twitter-search {twitter_search_params} \"{twitter_search}\" > {filename}")
21 tweets_df = pd.read_json(filename, lines=True)
22 if len(tweets_df) > 0:
23     tweets_df.rename(columns={"date": "tanggal", "username": "username", "rawContent": "tweet"})
24     tweets_df_fix=tweets_df.rename(["tanggal", "username", "tweet"])
25     tweets_df_fix.to_csv(f"Kota_Tegal.csv", index=False)
26
27 data_crawling_kota_tegal = pd.read_csv("Kota_Tegal.csv")
28 df = data_crawling_kota_tegal[["tanggal", "tweet"]]
```

#### Penjelasan :

Kode ini merupakan kode untuk melakukan crawling data dari platform Twitter. Pada crawling ini, akan mencari tweet-tweet yang berisi kata kunci "Kota Tegal" dalam bahasa Indonesia (lang:id). Pencarian akan dilakukan hingga tanggal saat ini (today) dan dimulai sejak kemarin (yesterday). Data tweet hasil crawling akan disimpan dalam format JSON dengan nama file "Kota\_Tegal.json". Setelah data berhasil diambil, maka akan dibaca kembali sebagai Dataframe menggunakan pandas dan hanya kolom "tanggal" dan "tweet" yang dipertahankan. Proses crawling ini memiliki manfaat untuk memperoleh informasi yang berkaitan dengan Kota Tegal dari

perspektif pengguna Twitter. Data tersebut dapat berupa ulasan, komentar, atau pemikiran masyarakat mengenai Kota Tegal yang diungkapkan dalam bentuk tweet. Data yang berhasil di-crawling tersebut nantinya dapat diolah lebih lanjut untuk analisis sentimen, analisis tren, atau mendapatkan wawasan penting tentang topik atau kejadian yang sedang populer atau menjadi perhatian di wilayah tersebut.

—

```

1 # Preprocessing
2 #menyimpan tweet. (tipe data series pandas)
3 data_content = df['tweet']
4 # casefolding
5 data_casefolding = data_content.str.lower()
6 #filtering
7
8 #url
9 filtering_url = [re.sub('(?:https?://www\d{0,3}[.]?[a-z0-9\~]+.[a-z]{2,4})?(?:[^\s()<>+@!|\'{}[\]`~.,:;\\"''"]*)', '', tweet) for tweet in data_casefolding]
10 #cont
11 filtering_cont = [re.sub('[^a-zA-Z]', '', tweet) for tweet in filtering_url]
12 #punctuation
13 filtering_punctuation = [re.sub('[!"#$%&'()*+,-./:;=:@\{\}~-'"[]|\s]', ' ', tweet) for tweet in filtering_cont]
14 #hapus stagger
15 filtering_tagger = [re.sub(r'#([\w\s])+', ' ', tweet) for tweet in filtering_punctuation]
16 #numeric
17 filtering_numeric = [re.sub(r'(\d+)', ' ', tweet) for tweet in filtering_tagger]
18
19 # filterng RT , @ dan #
20 # fungsi_clean_rt = lambda x: re.compile('\n').sub('', re.compile('rt @').sub('@ ', x, count=1).strip())
21 # clean = [fungsi_clean_rt for tweet in filtering_numeric]
22
23 data_filtering = pd.Series(filtering_numeric)
24 # tokenizer
25 tknzr = TweetTokenizer()
26 data_tokenize = [tknzr.tokenize(tweet) for tweet in data_filtering]
27 data_tokenized = []
28 # slang word
29 path_datalang = open("kamus kata baku-clear (1).csv")
30 datalang = pd.read_csv(path_datalang, encoding='utf-8', header=None, sep=";")
31
32 def replaceSlang(word):
33     if word in list(datalang[0]):
34         indexslang = list(datalang[0]).index(word)
35         return datalang[1][indexslang]
36     else:
37         return word
38
39 data_formal = []
40 for data in data_tokenize:
41     data_clean = [replaceSlang(word) for word in data]
42     data_formal.append(data_clean)
43 len_data_formal = len(data_formal)
44 # print(data_formal)
45 # len_data_formal
46 nltk.download('stopwords')
47 default_stop_words = nltk.corpus.stopwords.words('indonesian')
48 stopwords = set(default_stop_words)
49
50 def removeStopWords(line, stopwords):
51     words = {}
52     for word in line:
53         word=str(word)
54         word = word.strip()
55         if word not in stopwords and word != "" and word != " ":
56             words.append(word)
57
58     return words
59
60 reviews = [removeStopWords(line, stopwords) for line in data_formal]
61
62 # Specify the file path of the pickle file
63 file_path = 'model2\data_train.pickle'

```

Penjelasan :

Proses pemrosesan teks yang dilakukan di atas bertujuan untuk membersihkan dan mengolah data tweet yang telah di-crawling sebelum dilakukan analisis lebih lanjut. Pertama, teks tweet diubah menjadi huruf kecil untuk menghindari perbedaan huruf kapital.

Selanjutnya, dilakukan filtering untuk menghapus URL, kata "cont", tanda baca, hashtag, dan angka agar teks menjadi lebih bersih dan dapat dianalisis dengan lebih baik. Setelah proses filtering, teks tweet dibagi menjadi kata-kata individual melalui tokenisasi menggunakan TweetTokenizer dari nltk. Selanjutnya, kata-kata slang dalam tweet digantikan dengan kata formal berdasarkan kamus kata baku yang telah disediakan dalam file CSV. Langkah selanjutnya adalah menghapus stopwords atau kata-kata yang sering muncul namun tidak memberikan kontribusi signifikan dalam analisis teks. Hal ini bertujuan untuk menyaring kata-kata yang umum digunakan dalam bahasa Indonesia dan tidak relevan dalam konteks analisis tweet. Hasil akhir dari pemrosesan teks ini adalah teks tweet yang telah dibersihkan dari elemen-elemen yang tidak diperlukan dan siap untuk dijadikan bahan analisis lebih lanjut. Data tweet yang telah diolah disimpan dalam variabel reviews dalam bentuk daftar tweet yang telah dibersihkan dan diproses.

- Implementasi Model Wisata/Hiburan Secara Realtime

```
1 # pembuatan vector kata
2 vectorizer = TfidfVectorizer()
3 train_vector = vectorizer.fit_transform(data_train)
4 reviews2 = [" ".join(r) for r in reviews]
5
6 ## Implementasi Aspek Hiburan
7 model_aspek = pickle.load(open('model2/tfidf_Model_Aspek_Wisata_Hiburan.nvb.pkl', 'rb'))
8 model_sentiment = pickle.load(open('model2/vektor_tfidf_Model_Sentimen_Wisata_Hiburan.nvb.pkl', 'rb'))
9
10 result = []
11
12 for test in reviews2:
13     test_data = [str(test)]
14     test_vector = vectorizer.transform(test_data)
15     pred = model_aspek.predict(test_vector)
16     if pred == 1:
17         result.append(-1)
18     else:
19         pred = model_sentiment.predict(test_vector)
20         result.append(pred[0])
21
22
23 from sklearn.utils.multiclass import unique_labels
24 unique_labels(result)
25
26 df['wisata_hiburan'] = result
```



Penjelasan :

Selanjutnya, kami mengimplementasikan model untuk aspek "Wisata Hiburan". Model aspek ini telah dilatih sebelumnya menggunakan algoritma Multinomial Naive Bayes (NB) dengan menggunakan pendekatan TF-IDF untuk ekstraksi fitur. Kami juga telah melatih model sentimen terkait aspek "Wisata Hiburan" dengan menggunakan pendekatan yang sama. Kami memuat kedua model tersebut dari file yang telah disimpan sebelumnya. Selanjutnya, kami melakukan prediksi terhadap setiap tweet yang telah diolah sebelumnya. Jika model aspek memprediksi bahwa tweet bukan tentang aspek "Wisata Hiburan" (nilai prediksi bukan 1), maka hasil sentimen diatur sebagai -1 (tidak relevan). Namun, jika model aspek memprediksi bahwa tweet berkaitan dengan aspek "Wisata Hiburan" (nilai prediksi 1), maka kami menggunakan model sentimen untuk memprediksi sentimen dari tweet tersebut. Hasil prediksi sentimen kemudian disimpan dalam sebuah list yang bernama "result". Setelah itu, kolom "wisata\_hiburan" ditambahkan ke dataframe "df" yang berisi hasil crawling tweet. Kolom ini berisi hasil prediksi sentimen dari setiap tweet terkait aspek "Wisata Hiburan".

#### - Implementasi Model Pendidikan Secara Realtime

```
28 ## Implementasi Aspek Pendidikan
29 model_aspek = pickle.load(open('model2/tfidf_Model_Aspek_Wisata_Pendidikan.nvb.pkl', 'rb'))
30 model_sentiment = pickle.load(open('model2/vektor_tfidf_Model_Sentimen_Pendidikan.nvb.pkl', 'rb'))
31
32 result = []
33
34 for test in reviews2:
35     test_data = [str(test)]
36     test_vector = vectorizer.transform(test_data)
37     pred = model_aspek.predict(test_vector)
38     if pred == 1:
39         result.append(-1)
40     else:
41         pred = model_sentiment.predict(test_vector)
42         result.append(pred[0])
43
44 from sklearn.utils.multiclass import unique_labels
45 unique_labels(result)
46
47 df['pendidikan'] = result
48
```

Penjelasan :

Selanjutnya, kami mengimplementasikan model untuk aspek "Pendidikan". Seperti sebelumnya, kami memuat model aspek "Pendidikan" yang telah dilatih sebelumnya menggunakan algoritma Multinomial Naive Bayes (NB) dengan pendekatan TF-IDF. Kami juga memuat model sentimen terkait aspek "Pendidikan" yang menggunakan pendekatan yang sama. Kami melakukan prediksi terhadap setiap tweet yang telah diolah sebelumnya menggunakan kedua model tersebut. Pertama, kami menggunakan model aspek untuk memprediksi apakah tweet tersebut berkaitan dengan aspek "Pendidikan" atau tidak. Jika prediksi aspek bukan 1, artinya tweet tersebut tidak relevan dengan aspek "Pendidikan", dan nilai sentimen diatur sebagai -1. Namun, jika prediksi aspek adalah 1, kami menggunakan model sentimen untuk memprediksi sentimen tweet tersebut. Hasil prediksi sentimen disimpan dalam list bernama "result". Kemudian, kolom "pendidikan" ditambahkan ke dataframe "df" yang berisi hasil crawling tweet. Kolom ini berisi hasil prediksi sentimen dari setiap tweet terkait aspek "Pendidikan".

- Implementasi Model Fasilitas/Layanan Publik Secara Realtime

```
49 ## Implementasi Aspek Fasilitas dan Layanan Publik
50 model_aspek = pickle.load(open('model2/tfidf_Model_Aspek_Wisata_Fasilitas_Layanan_Publik.nvb.pkl', 'rb'))
51 model_sentiment = pickle.load(open('model2/vektor_tfidf_Model_Sentimen_Fasilitas_Layanan_Publik.nvb.pkl', 'rb'))
52
53 result = []
54
55 for test in reviews2:
56     test_data = [str(test)]
57     test_vector = vectorizer.transform(test_data)
58     pred = model_aspek.predict(test_vector)
59     if pred != 1:
60         result.append(-1)
61     else:
62         pred = model_sentiment.predict(test_vector)
63         result.append(pred[0])
64
65 from sklearn.utils.multiclass import unique_labels
66 unique_labels(result)
67
68 df['fasilitas_layanan_publik'] = result
```

Penjelasan :

Berikutnya, kami mengimplementasikan model untuk aspek "Fasilitas dan Layanan Publik". Seperti sebelumnya, kami memuat model aspek "Fasilitas dan Layanan Publik" yang telah dilatih sebelumnya menggunakan algoritma Multinomial Naive Bayes (NB) dengan pendekatan TF-IDF. Kami juga memuat model sentimen terkait aspek "Fasilitas dan Layanan Publik" yang menggunakan pendekatan yang sama. Seperti sebelumnya, kami melakukan prediksi terhadap setiap tweet yang telah diolah sebelumnya menggunakan kedua model tersebut. Pertama, kami menggunakan model aspek untuk memprediksi apakah tweet tersebut berkaitan dengan aspek "Fasilitas dan Layanan Publik" atau tidak. Jika prediksi aspek bukan 1, artinya tweet tersebut tidak relevan dengan aspek "Fasilitas dan Layanan Publik", dan nilai sentimen diatur sebagai -1. Namun, jika prediksi aspek adalah 1, kami menggunakan model sentimen untuk memprediksi sentimen tweet tersebut. Hasil prediksi sentimen disimpan dalam list bernama "result". Kemudian, kolom "fasilitas\_layanan\_publik" ditambahkan ke dataframe "df" yang berisi hasil crawling tweet. Kolom ini berisi hasil prediksi sentimen dari setiap tweet terkait aspek "Fasilitas dan Layanan Publik".

## - Implementasi Model Kuliner Secara Realtime

```

70 ## Implementasi Aspek Kuliner
71 model_aspek = pickle.load(open('model2/tfidf_Model_Aspek_Wisata_Kuliner_nvb.pkl', 'rb'))
72 model_sentiment = pickle.load(open('model2/vektor_tfidf_Model_Sentimen_Kuliner_nvb.pkl', 'rb'))
73
74 result = []
75
76 for test in reviews2:
77     test_data = [str(test)]
78     test_vector = vectorizer.transform(test_data)
79     pred = model_aspek.predict(test_vector)
80     if pred != 1:
81         result.append(-1)
82     else:
83         pred = model_sentiment.predict(test_vector)
84         result.append(pred[0])
85
86 from sklearn.utils.multiclass import unique_labels
87 unique_labels(result)
88
89 df['kuliner'] = result
90
91 ## Store ke DB
92 df=df.dropna()
93 df=df.drop_duplicates()
94 df.head()
95 import pymysql
96 def insert_to_list_rekomen_db(xp2):
97     # Connect to the database
98     connection = pymysql.connect(host='127.0.0.1',
99                                 port=3306,
100                                user='root',
101                                # password='f#ur8J3N',
102                                database='dashboard_kotategal_db')
103
104     # create cursor
105     cursor=connection.cursor()
106
107     cols = ','.join([str(i) for i in xp2.columns.tolist()])
108     for i,row in xp2.iterrows():
109         sql = "INSERT INTO 'hasil_sentiment' (" + cols + ") VALUES (" + "%s,"*(len(row)-1) + "%s)"
110         cursor.execute(sql, tuple(row))
111
112     # the connection is not autocommitted by default, so we must commit to save our changes
113     connection.commit()
114 insert_to_list_rekomen_db(df)

```

### Penjelasan :

Selanjutnya, kami mengimplementasikan model untuk aspek "Kuliner". Kami memuat model aspek "Kuliner" yang telah dilatih sebelumnya menggunakan algoritma Multinomial Naive Bayes (NB) dengan pendekatan TF-IDF. Kami juga memuat model sentimen terkait aspek "Kuliner" yang menggunakan pendekatan yang sama. Seperti sebelumnya, kami melakukan prediksi terhadap setiap tweet yang telah diolah sebelumnya menggunakan kedua model tersebut. Pertama, kami menggunakan model aspek untuk memprediksi apakah tweet tersebut berkaitan dengan aspek "Kuliner" atau tidak. Jika prediksi aspek bukan 1, artinya tweet tersebut tidak relevan dengan aspek "Kuliner", dan nilai sentimen diatur sebagai -1. Namun, jika prediksi aspek adalah 1, kami menggunakan model sentimen untuk memprediksi sentimen tweet tersebut. Hasil prediksi sentimen disimpan dalam list bernama "result". Kemudian, kolom "kuliner" ditambahkan ke dataframe "df" yang berisi

hasil crawling tweet. Kolom ini berisi hasil prediksi sentimen dari setiap tweet terkait aspek "Kuliner". Selanjutnya, dataframe "df" yang telah diperbarui dengan kolom "kuliner" akan disimpan ke dalam database menggunakan fungsi "insert\_to\_list\_rekomen\_db". Dataframe tersebut akan disimpan dalam tabel "hasil\_sentiment" di database "dashboard\_kotategal\_db".

- **Membuat tampilan Website dashboard**

- Membuat Fungsi untuk membuat tampilan Bar Chart jumlah Sentiment Analisis seluruh aspek

```
1 def plot_sentiment_chart(dataset):
2     # Menghitung jumlah sentimen 1 dan 0 pada setiap kategori
3     sentiment_counts = dataset.iloc[:, 2:].apply(pd.value_counts)
4
5     # Membentuk DataFrame baru untuk chart
6     sentiment_df = pd.DataFrame(sentiment_counts.stack()).reset_index()
7     sentiment_df.columns = ['Kategori', 'Sentimen', 'Jumlah']
8
9     # Menghapus kategori dengan nilai -1
10    sentiment_df = sentiment_df[sentiment_df['Kategori'] != -1]
11
12    # Mengubah label kategori
13    sentiment_df['Kategori'] = sentiment_df['Kategori'].map({1: 'Positif', 0: 'Negatif'})
14
15    # Membuat chart menggunakan Plotly
16    fig = px.bar(sentiment_df, x='Kategori', y='Jumlah', color='Sentimen',
17                barmode='group', title='Jumlah Sentimen Positif dan Negatif pada Seluruh Kategori')
18    fig.update_layout(xaxis_title='Kategori', yaxis_title='Jumlah')
19
20    # Menampilkan chart menggunakan Streamlit
21    st.plotly_chart(fig)
```

Penjelasan :

Fungsi **plot\_sentiment\_chart** di atas digunakan untuk menghasilkan grafik yang menampilkan jumlah sentimen positif dan negatif pada setiap kategori dalam dataset. Pertama, fungsi ini menghitung jumlah sentimen positif dan negatif menggunakan metode `pd.value_counts()`. Kemudian, hasil perhitungan tersebut digunakan untuk membentuk DataFrame baru yang akan digunakan sebagai basis

untuk grafik. Setelah itu, label kategori pada DataFrame diubah dari nilai numerik menjadi label yang lebih deskriptif menggunakan metode `map()`. Selanjutnya, baris dengan kategori yang memiliki nilai -1 dihapus agar grafik hanya mencakup kategori yang valid. Selanjutnya, menggunakan library Plotly, fungsi `px.bar()` digunakan untuk membuat grafik batang. Grafik ini akan menampilkan jumlah sentimen positif dan negatif pada seluruh kategori, dengan grup bar yang membedakan antara sentimen positif dan negatif. Setelah mengatur judul dan label sumbu menggunakan metode `update_layout()`, grafik tersebut ditampilkan menggunakan Streamlit dengan menggunakan `st.plotly_chart()`.

- Membuat Fungsi untuk membuat tampilan Bar Chart jumlah Sentiment Analisis dengan filter kategori Aspek

```
10 def plot_sentiment_chart_by_category(dataset, category):
11     """
12     Fungsi untuk membuat chart menggunakan Plotly yang menampilkan jumlah sentimen 1 dan 0 pada kategori yang spesifik.
13
14     Args:
15         dataset (pandas.DataFrame): DataFrame yang berisi dataset.
16         category (str): Nama kolom kategori yang ingin ditampilkan.
17     """
18
19     # Menghitung jumlah sentimen 1 dan 0 pada kategori yang spesifik
20     sentiment_counts = dataset[category].value_counts()
21
22     # Membentuk DataFrame baru untuk chart
23     sentiment_df = pd.DataFrame(sentiment_counts).reset_index()
24     sentiment_df.columns = ['Sentimen', 'Jumlah']
25     sentiment_df['Sentimen'] = sentiment_df['Sentimen'].map({1: 'Positif', 0: 'Negatif'})
26
27     # Membuat chart menggunakan Plotly
28     fig = px.bar(sentiment_df, x='Sentimen', y='Jumlah',
29                 color='Sentimen', title=f'Jumlah Sentimen Positif dan Negatif pada Kategori {category}')
30     fig.update_layout(xaxis_title='Sentimen', yaxis_title='Jumlah')
31
32     # Menampilkan chart menggunakan Streamlit
33     st.plotly_chart(fig)
```

Penjelasan :

Fungsi `plot_sentiment_chart_by_category` yang disediakan digunakan untuk menghasilkan grafik yang menampilkan jumlah sentimen positif dan negatif pada kategori yang spesifik dalam dataset. Fungsi ini melakukan beberapa langkah untuk menghasilkan grafik

yang informatif. Pertama, fungsi menghitung jumlah sentimen positif (nilai 1) dan negatif (nilai 0) pada kategori yang spesifik dalam dataset menggunakan metode `value_counts()`. Hasil perhitungan ini kemudian digunakan untuk membentuk DataFrame baru, `sentiment_df`, yang akan digunakan sebagai basis untuk grafik. Selanjutnya, fungsi mengubah label sentimen dari nilai numerik menjadi label yang lebih deskriptif menggunakan metode `map()`. Hal ini dilakukan untuk memperjelas interpretasi grafik. Setelah itu, menggunakan library Plotly, fungsi `px.bar()` digunakan untuk membuat grafik batang. Grafik ini akan menampilkan jumlah sentimen positif dan negatif pada kategori yang spesifik, dengan warna batang yang membedakan sentimen. Setelah mengatur judul dan label sumbu menggunakan metode `update_layout()`, grafik tersebut ditampilkan menggunakan Streamlit dengan menggunakan `st.plotly_chart()`

- Membuat Fungsi untuk membuat tampilan Line Chart grafik Sentiment Analisis seluruh aspek

```

21 def create_line_plot(data, category):
22     # Filter data berdasarkan kategori dan negatif count = 0
23     category_data = data[(data[category] == -1)]
24
25     # Menghitung jumlah tweet positif dan negatif per tanggal
26     positive_counts = category_data[category_data[category] == 1].groupby('tanggal').size()
27     negative_counts = category_data[category_data[category] == 0].groupby('tanggal').size()
28
29     # Menggabungkan kedua seri menjadi satu dataframe
30     plot_data = pd.DataFrame({'Positif': positive_counts, 'Negatif': negative_counts}).reset_index()
31
32     # Membuat line plot menggunakan Plotly
33     fig = px.line(plot_data, x='tanggal', y=['Positif', 'Negatif'],
34                 title='Line Plot Positif dan Negatif berdasarkan Kategori: ' + category)
35
36     # Mengubah warna label
37     fig.update_traces(line=dict(color='blue'), selector=dict(name='Positif'))
38     fig.update_traces(line=dict(color='red'), selector=dict(name='Negatif'))
39
40     fig.update_layout(legend=dict(orientation="h",
41                                 yanchor="top", # Mengatur anchor ke atas
42                                 y=1.0, # Mengatur posisi y ke 1 (paling atas)
43                                 xanchor="right",
44                                 x=1)) # Mengatur posisi x ke 1 (paling kanan)
45
46     # Menampilkan plot menggunakan Streamlit
47     st.plotly_chart(fig)

```

Penjelasan :

Fungsi `create_line_plot` yang telah disediakan bertujuan untuk membuat line plot yang menampilkan jumlah tweet positif dan negatif per tanggal pada kategori yang spesifik dalam dataset. Langkah-langkah yang dilakukan oleh fungsi tersebut adalah sebagai berikut: Pertama, data difilter berdasarkan kategori yang spesifik dan negatif count yang tidak sama dengan 0. Langkah ini dilakukan dengan menggunakan operator perbandingan `!=`. Kemudian, dilakukan penghitungan jumlah tweet positif dan negatif per tanggal. Hal ini dilakukan dengan mengelompokkan data berdasarkan tanggal menggunakan metode `groupby()`, dan kemudian menghitung jumlah baris menggunakan metode `size()`. Selanjutnya, kedua seri tersebut digabungkan menjadi satu DataFrame baru yang disebut `plot_data`. DataFrame ini akan menjadi dasar untuk pembuatan line plot. Menggunakan library Plotly, fungsi `px.line()` digunakan untuk membuat line plot. Line plot tersebut akan menampilkan jumlah tweet positif dan negatif per tanggal, dengan warna garis yang membedakan antara positif dan negatif. Selanjutnya, label garis positif diubah menjadi warna biru dan label garis negatif diubah menjadi warna merah menggunakan metode `update_traces()`.



- Membuat Fungsi untuk membuat tampilan Line Chart grafik Sentiment Analisis seluruh aspek dengan filter Tanggal awal dan akhir

```

21 def create_line_plot_with_date(data, category, start_date=None, end_date=None):
22     data['tanggal'] = pd.to_datetime(data['tanggal'])
23     # Filter data berdasarkan kategori dan tanggal
24     category_data = data[data[category] != -1]
25
26     if start_date is not None:
27         start_date = datetime.combine(start_date, datetime.min.time())
28         category_data = category_data[category_data['tanggal'] >= start_date]
29     if end_date is not None:
30         end_date = datetime.combine(end_date, datetime.min.time())
31         category_data = category_data[category_data['tanggal'] <= end_date]
32
33     # Menghitung jumlah tweet positif dan negatif per tanggal
34     positive_counts = category_data[category_data[category] == 1].groupby('tanggal').size()
35     negative_counts = category_data[category_data[category] == 0].groupby('tanggal').size()
36
37     # Menggabungkan kedua seri menjadi satu dataframe
38     plot_data = pd.DataFrame({'Positif': positive_counts, 'Negatif': negative_counts}).reset_index()
39
40     # Membuat line plot menggunakan Plotly
41     fig = px.line(plot_data, x='tanggal', y=['Positif', 'Negatif'],
42                 title='Line Plot Positif dan Negatif berdasarkan Kategori: ' + category)
43
44     # Menampilkan plot menggunakan Streamlit
45     st.plotly_chart(fig)

```

#### Penjelasan :

Fungsi `create_line_plot_with_date` yang telah diberikan berfungsi untuk membuat line plot yang menampilkan jumlah tweet positif dan negatif per tanggal pada kategori yang spesifik dalam dataset. Fungsi ini juga memberikan fleksibilitas untuk memfilter data berdasarkan rentang tanggal yang ditentukan. Langkah-langkah yang dilakukan oleh fungsi tersebut adalah sebagai berikut: Pertama, kolom "tanggal" pada dataframe data dikonversi menjadi tipe data datetime menggunakan metode `pd.to_datetime()`. Selanjutnya, data difilter berdasarkan kategori yang spesifik dan rentang tanggal yang ditentukan. Jika tanggal awal (`start_date`) atau tanggal akhir (`end_date`) diberikan, maka data akan difilter berdasarkan rentang tanggal tersebut. Filter dilakukan dengan membandingkan kolom "tanggal" dengan tanggal awal dan akhir yang telah dikonversi menjadi tipe datetime. Kemudian, dilakukan penghitungan jumlah tweet positif dan negatif per tanggal pada data yang telah difilter. Hal ini dilakukan dengan

mengelompokkan data berdasarkan tanggal menggunakan metode `groupby()` dan menghitung jumlah baris menggunakan metode `size()`. Setelah itu, kedua seri tersebut digabungkan menjadi satu DataFrame baru yang disebut `plot_data`. DataFrame ini akan digunakan sebagai dasar untuk membuat line plot. Menggunakan library Plotly, fungsi `px.line()` digunakan untuk membuat line plot. Line plot tersebut akan menampilkan jumlah tweet positif dan negatif per tanggal, dengan garis yang membedakan antara positif dan negatif. selanjutnya, grafik yang dihasilkan ditampilkan menggunakan Streamlit dengan menggunakan `st.plotly_chart()`.

- Membuat Fungsi untuk membuat tampilan wordcloud Sentiment Analisis seluruh aspek

```
3 def preprocess_text(text):
4     # Remove URLs
5     text = re.sub(r'http\S+', '', text)
6
7     # Remove usernames
8     text = re.sub(r'@\s+', '', text)
9
10    # Remove special characters and numbers
11    text = re.sub(r'[^\s]', '', text)
12
13    # Convert text to lowercase
14    text = text.lower()
15
16    # Remove stop words
17    factory = StopWordRemoverFactory()
18    stop_words = factory.create_stop_word_remover()
19    text = stop_words.remove(text)
20
21    # Remove specific phrase
22    text = re.sub(r'tegal', '', text)
23    text = re.sub(r'kota', '', text)
24
25    return text
26
27 def create_wordcloud(data, category):
28     # Filter dataset by category
29     filtered_data = data[data[category] == 1]
30
31     # Concatenate all tweets in the selected category
32     text = ' '.join(filtered_data['tweet'])
33
34     # Preprocess the text
35     text = preprocess_text(text)
36
37     # Generate word cloud
38     wordcloud = WordCloud(width=800, height=400, background_color='white').generate(text)
39
40     # Display the word cloud
41     plt.figure(figsize=(10, 5))
42     plt.imshow(wordcloud, interpolation='bilinear')
43     plt.axis('off')
44     plt.title('Word Cloud - {}'.format(category.capitalize()))
45     st.pyplot(plt)
```

Penjelasan :

Fungsi **create\_wordcloud** menggunakan data yang difilter berdasarkan kategori yang spesifik. Selanjutnya, semua teks pada kategori tersebut digabungkan menjadi satu teks yang kemudian diproses menggunakan fungsi **preprocess\_text**. Setelah proses pra-pemrosesan selesai, word cloud dihasilkan dengan menggunakan pustaka **WordCloud**. Word cloud ini akan menampilkan kata-kata yang paling sering muncul dalam teks yang telah diproses. Grafik word cloud

kemudian ditampilkan menggunakan `plt.imshow()` dan `st.pyplot(plt)` untuk ditampilkan melalui Streamlit. Dengan menggunakan fungsi `create_wordcloud`, Anda dapat membuat visualisasi word cloud yang memvisualisasikan kata-kata yang paling sering muncul dalam teks pada kategori yang spesifik dalam dataset. Word cloud memberikan gambaran visual tentang kata-kata yang paling relevan atau umum terkait dengan aspek wisata hiburan, pendidikan, fasilitas publik, dan kuliner di Kota Tegal.

## - Membuat Fungsi main

```

1  def main():
2      with st.sidebar:
3          selected = option_menu(
4              menu_title="Main Menu",
5              options=["Home", "Dashboard"],
6              icons=["house", "book"],
7              menu_icon="cast",
8              default_index=0,
9          )
10     if selected == "Home":
11         st.title(f" {selected}")
12         #judul dan Logo Kota Tegal
13         col1, col2 = st.columns((2, 8))
14
15         with col1:
16             # st.header("")
17             image = Image.open("Logo-Kota-Tegal.png")
18             # resized_image = image.resize((150, 150))
19             st.image(image, caption="", use_column_width=True)
20
21         with col2:
22             st.header("Sentimen Analisis Kota Tegal Pada Aspek Wisata Hiburan, Pendidikan, Fasilitas Publik, dan Kuliner")
23             st.write('Ini adalah halaman tentang analisis sentimen di Kota Tegal terkait dengan aspek wisata hiburan,
24                 pendidikan, fasilitas publik, dan kuliner.')
25             st.write('Kami menggunakan teknik pemrosesan bahasa alami dan analisis sentimen untuk menganalisis ulasan dan
26                 pendapat masyarakat terkait dengan aspek-aspek tersebut.')
27             st.write('Dengan menganalisis sentimen, kami dapat mendapatkan wawasan mengenai bagaimana masyarakat merasakan dan
28                 mengevaluasi wisata hiburan, pendidikan, fasilitas publik, dan kuliner di Kota Tegal.')
29             st.write('Melalui halaman ini, kami bertujuan untuk menyajikan hasil analisis sentimen secara interaktif dan
30                 informatif kepada pengguna.')
31             st.write('Selamat menikmati dan semoga informasi yang disajikan bermanfaat!')
32
33     else:
34         #judul dan Logo Kota Tegal
35         col1, col2 = st.columns((2, 8))
36
37         with col1:
38             # st.header("")
39             image = Image.open("Logo-Kota-Tegal.png")
40             # resized_image = image.resize((150, 150))
41             st.image(image, caption="", use_column_width=True)
42
43         with col2:
44             st.header("Sentimen Analisis Kota Tegal Pada Aspek Wisata Hiburan, Pendidikan, Fasilitas Publik, dan Kuliner")
45
46             # Membaca data dari dataset Twitter Sentiment Analysis
47             data = pd.read_csv("data-26 Juni 2023.csv", sep=",")
48
49             # Membuat select box untuk memilih kategori
50             categories = ['wisata_hiburan', 'pendidikan', 'fasilitas_layanan_publik', 'kuliner']
51             category = st.sidebar.selectbox("Pilih kategori:", ['Semua Kategori']+categories)
52
53             if category == 'Semua Kategori' :
54                 #Bar Plot
55                 plot_sentiment_chart(data)
56
57             else:
58                 plot_sentiment_chart_by_category(data, category)
59                 tanpa_filter, filter = st.tabs(["Tanpa Filter", "Filter"])
60                 with tanpa_filter:
61                     create_line_plot(data, category)
62                     create_wordcloud(data, category)
63
64                 with filter:
65                     col1, col2 = st.columns((2, 8))
66                     with col1:
67                         start_date = st.date_input("Pilih Tanggal Awal", value=None)
68                         end_date = st.date_input("Pilih Tanggal Akhir", value=None)
69                     with col2:
70                         create_line_plot_with_date(data, category, start_date, end_date)
71
72     main()

```

## Penjelasan :

Fungsi main merupakan bagian utama dari program yang mengatur tampilan dan logika pada halaman utama ("Home") dan halaman dashboard pada aplikasi Sentimen Analisis Kota Tegal. Pada halaman "Home", tampilan terdiri dari judul aplikasi "Sentimen Analisis Kota Tegal Pada Aspek Wisata Hiburan, Pendidikan, Fasilitas

Publik, dan Kuliner" dan logo Kota Tegal. Diikuti dengan beberapa paragraf yang menjelaskan tentang analisis sentimen di Kota Tegal terkait dengan aspek wisata hiburan, pendidikan, fasilitas publik, dan kuliner. Penjelasan tersebut menggambarkan penggunaan teknik pemrosesan bahasa alami dan analisis sentimen untuk menganalisis ulasan dan pendapat masyarakat terkait dengan aspek-aspek tersebut. Tujuan utama dari halaman ini adalah menyajikan hasil analisis sentimen secara interaktif dan informatif kepada pengguna. Pada halaman "Dashboard", terdapat pemilihan kategori menggunakan select box pada sidebar. Pengguna dapat memilih untuk melihat data dan visualisasi sentimen untuk semua kategori atau memilih salah satu kategori spesifik. Jika memilih "Semua Kategori", maka akan ditampilkan bar plot yang menampilkan jumlah sentimen positif dan negatif pada setiap kategori. Jika pengguna memilih kategori spesifik, maka akan ditampilkan grafik bar plot yang menampilkan jumlah sentimen positif dan negatif pada kategori tersebut. Selain itu, terdapat tab "Tanpa Filter" dan "Filter". Pada tab "Tanpa Filter", ditampilkan line plot yang menunjukkan tren jumlah tweet positif dan negatif per tanggal pada kategori yang dipilih, serta word cloud yang memvisualisasikan kata-kata yang paling sering muncul pada kategori tersebut. Pada tab "Filter", pengguna dapat memilih tanggal awal dan tanggal akhir untuk melakukan filter pada line plot dan menampilkan tren sentimen pada rentang tanggal yang dipilih. Secara keseluruhan, fungsi main mengatur tampilan dan interaksi dengan pengguna pada halaman "Home" dan halaman dashboard. Halaman ini menyajikan informasi tentang analisis sentimen di Kota Tegal dan menyediakan visualisasi yang berguna untuk memahami sentimen pada aspek wisata hiburan, pendidikan, fasilitas publik, dan kuliner di Kota Tegal.

## Lampiran 6 Sertifikat HKI yang Terbit

 REPUBLIK INDONESIA KEMENTERIAN HUKUM DAN HAK ASASI MANUSIA	
<b>SURAT PENCATATAN CIPTAAN</b>	
Dalam rangka perlindungan ciptaan di bidang ilmu pengetahuan, seni dan sastra berdasarkan Undang-Undang Nomor 28 Tahun 2014 tentang Hak Cipta, dengan ini menerangkan:	
Nomor dan tanggal permohonan	: EC00202355151, 13 Juli 2023
<b>Pencipta</b>	
Nama	: Novita Fitria Putri, Muhammad Fikri Hidayattullah dkk
Alamat	: Griya Santika Blok N.15, RT 02/RW 04, Kelurahan Pengabean, Kecamatan Dukuhuri, Tegal, Jawa Tengah, 52192
Kewarganegaraan	: Indonesia
<b>Pemegang Hak Cipta</b>	
Nama	: Pusat Penelitian dan Pengabdian Masyarakat (P3M) Politeknik Harapan Bersama
Alamat	: Jalan Mataram No. 9, Pesurungan Lor, Kecamatan Margadana, Tegal, JAWA TENGAH 52142
Kewarganegaraan	: Indonesia
Jenis Ciptaan	: Program Komputer
Judul Ciptaan	: Aplikasi Sentimen Analisis Kota Tegal Berbasis Aspek Menggunakan Algoritma Naive Bayes
Tanggal dan tempat diumumkan untuk pertama kali di wilayah Indonesia atau di luar wilayah Indonesia	: 13 Juli 2023, di Tegal
Jangka waktu perlindungan	: Berlaku selama 50 (lima puluh) tahun sejak Ciptaan tersebut pertama kali dilakukan Pengumuman
Nomor pencatatan	: 000488086
adalah benar berdasarkan keterangan yang diberikan oleh Pemohon. Surat Pencatatan Hak Cipta atau produk Hak terkait ini sesuai dengan Pasal 72 Undang-Undang Nomor 28 Tahun 2014 tentang Hak Cipta	
	a.n. MENTERI HUKUM DAN HAK ASASI MANUSIA Direktur Hak Cipta dan Desain Industri
	 Anggoro Dasananto NIP. 196412081991031002
Disclaimer: Dalam hal pemohon memberikan keterangan tidak sesuai dengan surat pernyataan, Menteri berwenang untuk mencabut surat pencatatan permohonan.	

**LAMPIRAN PENCIPTA**

No	Nama	Alamat
1	Novita Fitria Putri	Griya Santika Blok N.15, RT 02/RW 04, Kelurahan Pengabean, Kecamatan Dukuhturi
2	Muhammad Fikri Hidayattullah	Jl. Glatik No. 68 Randugunting, Kecamatan Tegal Selatan, Kota Tegal, Jawa Tengah, 52131
3	Dwi Intan Af'idah	Desa Grinting RT003/RW001, Kecamatan Bulakamba, Kabupaten Brebes, Jawa Tengah, 52253





Lampiran 7 Lembar Bimbingan



**SARJANA TERAPAN TEKNIK INFORMATIKA**  
**POLITEKNIK HARAPAN BERSAMA**

**LEMBAR BIMBINGAN SKRIPSI**

**Nama** : Novita Fitria Putri

**NIM** : 19090130

**No. Ponsel** : 0822-2304-5336

**Judul Skripsi** : Sentimen Analisis Kota Tegal Berbasis Aspek Menggunakan  
Algoritma Naïve Bayes


**Dosen Pembimbing I** : Muhammad Fikri Hidayattullah, S.T.,M.Kom

No	Tanggal	Pemeriksaan	Perbaikan Yang Perlu Dilakukan	Paraf Pembimbing
1.	15/3 2023	Konsep	Judul "Aspek" mohon dicek sean mendalam!	
2.	29/5 2023	Model	Kuesi sudah terapan modeling!	
3.	3/6	Aplikasi	- Integrasi ke mode real time! - mess Kamis, 5 Juli 2023	
4.	9/7	Dokumen HKI	- Manual book diberi pendahuluan. - Dokumen terminal diberi pengelasan dan Capture tampilan!	

5.	7/7	laporan	Bur I	Fi
6.	17/7	laporan	- Bur I - II Revisi : - Tera feli's - Ramegn U1 - Keleng kapu	Fi
7.	25/7	laporan	lingkungan lokasi	Fi
8.	27/7	laporan	Ace	Fi

Tegal, 25 Juli .....2023

Dosen Pembimbing I

  
Muhammad Fikri Hidayattullah, S.T., M.Kom  
NIPY.09.016.307



SARJANA TERAPAN TEKNIK INFORMATIKA

POLITEKNIK HARAPAN BERSAMA

LEMBAR BIMBINGAN SKRIPSI

Nama : Novita Fitria Putri

NIM : 19090130

No. Ponsel : 0822-2304-5336

Judul Skripsi : Sentimen Analisis Kota Tegal Berbasis Aspek Menggunakan  
Algoritma Naïve Bayes

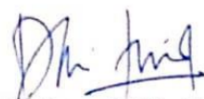
Dosen Pembimbing II : Dwi Intan Afidah, M. Kom

No	Tanggal	Pemeriksaan	Perbaikan Yang Perlu Dilakukan	Paraf Pembimbing
1.	27/3-2023	Konsep Aplikasi	Gunakan metode SMOTE / ADASYN untuk balancing data	Dwi Intan Afidah
2.	14/3-2023	Produk	Tambahkan tampilan agar informatif, Persiapkan artikel ilmiah	Dwi Intan Afidah
3.	22/5-2023	Produk	Perbaiki tampilan dari aplikasi agar informatif Selesaikan semua aplikasi agar berjalan dg baik	Dwi Intan Afidah
4.	7/7-2023	Produk	Sudah selesai Lanjutkan pengajuan hak cipta	Dwi Intan Afidah
5.	15/7 2023	Laporan	Jelaskan data penelitian. Perbaiki Bab 2 disub bab 2.2 Perbaiki font daftar pustaka	Dwi Intan Afidah

6.	16/7 2023	Laporan	<sup>kesimpulan</sup> Perbaiki <del>kesimpulan</del> menjadi Point - Point . Perbaiki <del>kesimpulan</del> Saran agar lebih Spesifik .	Dwi Intan
7.	18/7 2023	Aplikasi	Pastikan aplikasi sudah di hosting	Dwi Intan
8.	18/7 2023	Laporan dan aplikasi	Acc	Dwi Intan

Tegal, 18 Juli .....2023

Dosen Pembimbing II



Dwi Intan Afidah, M. Kom

NIPY. 11.020.470