

**ANALISIS SENTIMEN OBJEK WISATA DI KABUPATEN
BANGGAI LAUT MENGGUNAKAN
METODE NAIVE BAYES**

Oleh

FATMAWATI

T3120013

SKRIPSI

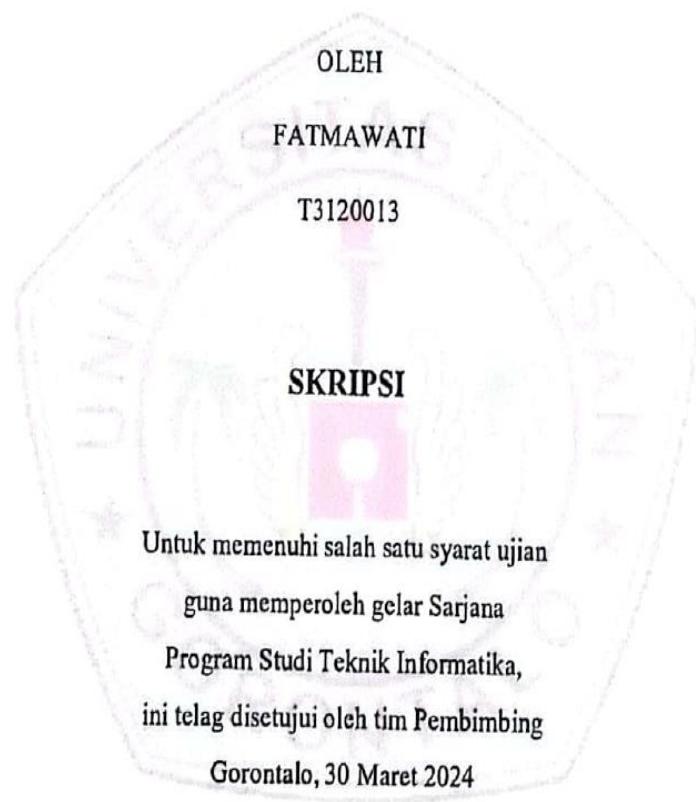
**Untuk memenuhi salah satu syarat ujian
guna memperoleh gelar sarjana**



**PROGRAM SARJANA
TEKNIK INFORMATIKA
UNIVERSITAS ICHSAN GORONTALO
GORONTALO
2024**

Persetujuan Skripsi

ANALISIS SENTIMEN OBJEK WISATA DI KABUPATEN BANGGAI LAUT MENGGUNAKAN METODE NAIVE BAYES



Pembimbing I

Husdi, M.Kom

NIDN. 0907108701

Pembimbing II

Kartika Chandra Pelangi, M.Kom

NIDN. 0916038304

PENGESAHAN SKRIPSI

ANALISIS SENTIMEN OBJEK WISATA DI KABUPATEN BANGGAI LAUT MENGGUNAKAN METODE NAIVE BAYES

OLEH

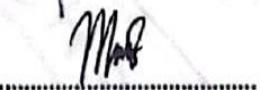
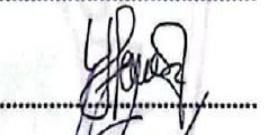
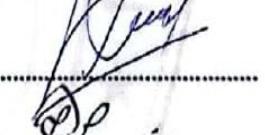
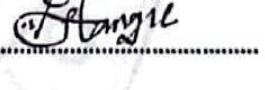
FATMAWATI

T3120013

Dipersembahkan oleh Panitia Ujian Strata Satu (S1)

Universitas Ichsan Gorontalo

1. Ketua Pengudi

Rezqiwati Ishak, M.Kom
2. Anggota

Muis Nanja, M.Kom
3. Anggota

Yulianty Lasena, M.Kom
4. Anggota

Husdi, M.Kom
5. Anggota

Kartika Chandra Pelangi, M.Kom

Mengetahui



PERYATAAN SKRIPSI

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

Gorontalo,



ABSTRACT

FATMAWATI. T3120013. THE SENTIMENT ANALYSIS OF TOURIST ATTRACTIONS IN BANGGAI LAUT REGENCY USING THE NAIVEBAYES METHOD

A tourist attraction is a place visited for its beauty where tourism activities are carried out and a place to have fun for a long time to achieve satisfaction, enjoy good service, and bring home beautiful memories of the tourist attractions. In Banggai Laut Regency, many tourist attractions are attractive for tourism activities. The number of visitors increases every year. The discussion about tourist attractions in Banggai Laut Regency is interesting to the local community. There are a lot of public comments about the tourist attractions in Banggai Laut. The number of opinions about tourist attractions in Banggai Laut makes it difficult to determine the sentiment of the comments manually. Therefore, sentiment analysis is required to classify the comments and whether or not they tend to be positive. In this case, this study employs the Naïve Bayes algorithm to classify the problems. Based on the sentiment analysis, it is proven that the Naïve Bayes method or algorithm can classify comments with good results. The accuracy generated in this sentiment analysis is 87%, with a division of training data of 90% and test data of 10%. The acquisition of these accuracy results indicates that the proposed algorithm has a Fairly Good diagnostic level.

Keywords: *sentiment analysis, tourist attraction, Naïve Bayes*



ABSTRAK

FATMAWATI. T3120013. ANALISIS SENTIMEN OBJEK WISATA DI KABUPATEN BANGGAI LAUT MENGGUNAKAN METODE NAIVE BAYES

Objek wisata adalah tempat yang dikunjungi karena keindahannya, tempat di mana aktivitas pariwisata dilakukan, tempat untuk bersenang-senang dalam waktu yang cukup lama demi mencapai kepuasan, menikmati pelayanan yang baik, dan membawa pulang kenangan indah dari tempat wisata tersebut. Di Kabupaten Banggai Laut, terdapat banyak objek wisata yang sangat menarik untuk aktivitas pariwisata, sehingga jumlah pengunjungnya meningkat setiap tahunnya. Pembahasan mengenai objek wisata Kabupaten Banggai Laut sangat menarik perhatian masyarakat setempat, sehingga banyak sekali komentar-komentar masyarakat mengenai objek wisata yang ada di Banggai Laut sendiri. Banyaknya opini mengenai objek wisata yang ada di Banggai Laut ini membuat adanya kesulitan dalam menentukan sentimen dari komentar-komentar secara manual. Oleh karena itu, diperlukan analisis sentimen yang dapat mengklasifikasi komentar tersebut apakah cenderung bersifat positif atau negatif. Dalam hal ini, penulis menggunakan algoritma *Naïve Bayes* untuk mengklasifikasi permasalahan tersebut. Berdasarkan analisis sentimen yang telah dilakukan, maka dapat diketahui bahwa metode atau algoritma *Naïve Bayes* ini mampu mengklasifikasi komentar dengan hasil yang baik. Adapun akurasi yang dihasilkan pada analisis sentimen ini adalah sebesar 87% dengan pembagian data latih sebesar 90% dan data uji sebesar 10%. Dengan perolehan hasil akurasi tersebut, maka dapat dikatakan bahwa algoritma yang diusulkan ini memiliki tingkat diagnostik yang cukup baik.

Kata kunci: analisis sentimen, objek wisata, *Naïve Bayes*



KATA PENGANTAR

Alhamdulillah, penulis dapat menyelesaikan usulan penelitian atau skripsi dengan judul: “**Analisis Sentimen Objek Wisata Di Kabupaten Banggai Laut Menggunakan Metode Naive Bayes**”, untuk memenuhi salah satu syarat penyusunan Skripsi Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.

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

1. Ibu Dr. Yuriko Abdulsamat, M.Si, selaku Ketua Yayasan Pengembangan Ilmu Pengetahuan dan Teknologi (YPIPT) Ichsan Gorontalo;
2. Bapak Dr. Abdul Gaffar La Tjokke, M.Si, selaku Rektor Universitas Ichsan Gorontalo;
3. Bapak Irvan Abraham Salihi S.Kom.,M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
4. Bapak Sudirman Melangi, M.Kom, selaku Pembantu Dekan I Bidang Akademik Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
5. Ibu Irma Surya Kumala Idris, M.Kom, selaku Pembantu Dekan II Bidang Administrasi Umum dan Keuangan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
6. Bapak Sudirman S. Panna, S.Kom, M.Kom, selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
7. Bapak Husdi, M.kom, selaku Pembimbing I dalam penelitian ini yang telah membimbing penulis selama penelitian;
8. Ibu Kartika Chandra Pelangi, M.Kom, selaku Pembimbing II dalam penelitian ini telah membimbing penulis selama penelitian;

9. Bapak dan Ibu Dosen Universitas Ichsan Gorontalo yang telah mendidik dan mengajarkan berbagai disiplin ilmu kepada penulis;
10. Kedua Orang Tua saya yang tercinta, atas segala kasih sayang, jerih payah dan doa restunya dalam membesar dan mendidik penulis;
11. Rekan-rekan seperjuangan yang telah banyak memberikan bantuan dan dukungan moril yang sangat besar kepada penulis;
12. Kepada semua pihak yang ikut membantu dalam penyelesaian proposal atau skripsi ini yang tak sempat penulis sebutkan satu-persatu.

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

Gorontalo,.../.../2024

Penulis

DAFTAR ISI

PERSETUJUAN SKRIPSI	Error! Bookmark not defined.
PENGESAHAN SKRIPSI	ii
PERYATAAN SKRIPSI.....	iii
ABSTRAK	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
BAB I PENDAHULUAN	1
1.1 Latang Belakang	1
1.2 Identifikasi Masala.....	5
1.3 Rumusan Masalah.....	5
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
BAB II LANDASAN TEORI	6
2.1 Tinjauan Studi	6
2.2 Tinjauan Pustaka.....	8
2.2.1 Objek Wisata	8
2.2.2 Analisis Sentimen	8
2.2.3 Text Mining.....	9
2.2.4 Preprocessing.....	9
2.2.5 Pembobotan TF-IDF.....	13
2.2.6 <i>Naïve Bayes</i>	15
2.3 Kerangka Pikir	19
BAB III METODE PENELITIAN	20
3.1 Jenis, Metode, Subjek, Objek, Waktu, dan Lokasi Penelitian	20
3.2 Tahap Penelitian	20
3.2.1 Pengumpulan Data.....	21
3.2.2 Pengembangan Model.....	22

3.2.3	Implementasi Model	24
3.2.4	Evaluasi Model.....	24
BAB IV HASIL PENELITIAN.....		25
4.1	Hasil Pengumpulan Data.....	25
4.2	Hasil Pemodelan	28
4.2.1	Hasil <i>Preprocessing</i>	29
4.2.2	Hasil Pembobotan TF.IDF	53
4.2.3	Klasifikasi <i>Naïve Bayes</i>	63
BAB V PEMBAHASAN		65
5.1	Pembahasan Model.....	65
5.1.1	<i>Preprocessing</i>	65
5.1.1	Pembobotan.....	66
5.1.2	Analisis	67
5.1.3	Evaluasi.....	68
5.1.4	Implementasi	69
BAB VI PENUTUP		71
6.1	Kesimpulan.....	71
6.2	Saran	71
DAFTAR PUSTAKA		72

DAFTAR GAMBAR

Gambar 3. 1 Tahap Penelitian	20
Gambar 3. 2 Gambar Pengembangan Model	22
Gambar 4. 1 Hasil Pengumpulan Data.....	27
Gambar 4. 2 Pemodelan	28
Gambar 4. 3 Flowchart Preprocessing	29
Gambar 4. 4 Flowchart TF-IDF.....	53
Gambar 4. 5 Flowchart Klasifikasi Naïve Bayes	63
Gambar 4. 6 Uji Akurasi	69
Gambar 5. 1 Preprocessing.....	65
Gambar 5. 2 Pembobotan.....	66
Gambar 5. 3 Analisis.....	67
Gambar 5. 4 Evaluasi	68
Gambar 5. 5 Hasil Implementasi 1	69
Gambar 5. 6 Hasil Implementasi 2	70

DAFTAR TABEL

Tabel 1. 1 Laporan data wisata Kabupaten Banggai Laut	2
Tabel 2. 1 Penelitian Terdahulu.....	6
Tabel 2. 2 Proses Case Folding	10
Tabel 2. 3 Proses Tokenizing	10
Tabel 2. 4 Proses Filtering.....	11
Tabel 2. 5 Proses Stemming	13
Tabel 2. 6 Perhitungan TF.IDF.....	15
Tabel 4. 1 Hasil Pengumpulan Data	25
Tabel 4. 2 Hasil Case Folding	29
Tabel 4. 3 Hasil dari Tokenizing	32
Tabel 4. 4 Hasil dari Stopword Removal	40
Tabel 4. 5 Hasil dari Stemming	47
Tabel 4. 6 Sampel Term dan Perhitungan manual TF	53
Tabel 4. 7 Sampel Perhitungan Manual TF Normalisasi	56
Tabel 4. 8 Sampel Perhitungan Manual DF dan IDF	58
Tabel 4. 9 Sampel Perhitungan Manual TF-IDF	61

BAB I

PENDAHULUAN

1.1 Latang Belakang

Pariwisata adalah kegiatan perjalanan wisata yang dapat dilakukan secara berulang-ulang atau mengelilingi suatu tempat, baik secara terencana maupun tanpa rencana, yang mampu memberikan pengalaman mendalam bagi para pelakunya. Pariwisata merupakan salah satu sektor yang memiliki potensi besar untuk menjadi penyumbang devisa terbesar di Indonesia. Pertumbuhan industri pariwisata Indonesia mencapai 7,4 persen per tahun, melebihi rata-rata pertumbuhan industri pariwisata dunia sebesar 5,8 persen. Pemerintah telah menghadirkan berbagai program untuk meningkatkan jumlah kunjungan wisatawan, baik yang berasal dari dalam negeri maupun dari luar negeri [1].

Banggai laut sebagai salah satu Kabupaten di Provinsi Sulawesi Tengah yang merupakan jalur pelayaran nusantara yang menghubungkan beberapa daerah kepulauan di Sulawesi Tengah serta didukung dengan kekayaan alam, pesona budaya daerah, keunikan sejarah. Berbagai upaya pengembangan dan promosi pariwisata daerah sebagai salah satu Daerah Tujuan wisata.

Salah satu upaya dinas pariwisata Kabupaten Banggai Laut untuk memperkenalkan objek-objek wisata yang ada di Kabupaten Banggai Laut dilakukan berbagai strategi marketing yang bersifat terpadu dan provisional dengan melibatkan seluruh stakeholder melalui pendekatan: *co-based marketing*, *society-based marketing* dan *IT-based marketing* dengan mengutamakan karakteristik, dan kearifan lokal yang berbudaya. Mempromosikan dan memasarkan objek dan daya tarik wisata (wisata alam, wisata budaya, wisata sejarah, wisata religi), maka perlu diciptakan sebuah media yang berbasis IT (*IT-based marketing*) dalam bentuk sebuah website yang menggambarkan berbagai informasi tentang Potensi Industri Pariwisata Kabupaten Banggai Laut.

Tabel 1. 1 Laporan data wisata Kabupaten Banggai Laut

JENIS WISATA	JUMLAH WISATA	NAMA-NAMA WISATA
Wisata Alam	6	1. Bukit Popisi 2. Paisu Batongan 3. Air Terjun Kokungan 4. Gua Tokubet 5. Padang Laya 6. Air Terjun
Wisata Budaya	6	1. Banggai Lalongo 2. Upacara Adat Sasampe 3. Upacara Adat Mabangun Tunggul Kamali Boneaka 4. Upacara Adat Mabangun Tunggul Kamali Banggai Lalongo 5. Upacara Adat Mabangun Tunggul Kamali Padang Laya 6. Air Terjun Matube
Wisata Religi	1	1. Makam Raja Mandar
		1. Spot Diving Tinakin Laut 2. Spot Diving Tolobundu

Wisata Bahari	20	3. Spot Diving Pulau Togong
		4. Spot Diving Pulau Bangkurung
		5. Spot Diving Pulau Bakakan
		6. Spot Diving Rudi Gaja
		7. Spot Diving Taman Laut Kima
		8. Pantai Pasir Putih
		9. Menara Pandang
		10. Pulau Tulungan
		11. Taman Laut Kima
		12. Pantai Binotik
		13. Pantai Kalapa Lima
		14. Pantai Monosan
		15. Pantai Lambangan Pauno
		16. Pantai Bontolan
		17. Pantai Oyama
		18. Pulau Bandang
		19. Spot Diving Batu Kapal 2
		20. Pantai Pompon

Sumber: Dinas Pariwisata Dan Kebudayaan Kabupaten Banggai Laut

Berdasarkan table 1.1 diatas menunjukan bahwa Kabupaten Banggai Laut Memiliki banyak tempat wisata antara lain wisata alam, wisata budaya, wisata religi, dan wisata bahari.

Sesuai dengan observasi atau wawancara yang peneliti lakukan di kantor dinas pariwisata dan kebudayaan kabupaten banggai laut, pemerintah Banggai Laut belum menyiapkan masyarakat yang sadarkan pariwisata dan menjalani kerjasama dengan pemanku kepentingan sehingga upaya meningkatkan daya saing destinasi pariwisata masih minim. Pemerintah masih acuh terhadap potensi pariwisata yang ada di Banggai Laut.

Karena sekarang pengunjung wisata yang ada di banggai laut meningkat maka pemerintah harus mempertahankan atau meningkatkan sarana dan prasarana maupun pelayanan maka dibutuhkan analisis untuk mengetahui tanggapan pengunjung tehadap layanan tersebut. Jadi analisis sentimen akan memberikan suatu pengambilan informasi berkualitas tinggi dan komentar-komentar untuk diklasifikasikan. Teknik klasifikasian sentiment bias dipakai dalam menentukan polaris pada setiap komentar tunggal dan kemudian disatukan menjadi ranting sentiment.

Untuk menganalisis sentimen pengunjung, peneliti menggunakan metode Bayes. Metode Bayes adalah metode yang diciptakan oleh Thomas Bayes yang digunakan untuk menghitung probabilitas suatu peristiwa berdasarkan informasi sebelumnya. Text mining adalah proses penambangan yang dilakukan oleh komputer untuk menggali informasi baru yang sebelumnya tidak diketahui [2].

Penelitian mengenai Analisis sentimen objek wisata menggunakan *naive bayes* untuk mengetahui komentar wisatawan pernah dilakukan sebelumnya oleh beberapa penelitian. Pada penelitian dengan judul “Analisis Sentimen Wisata Jawa Tengah Menggunakan *Naive Bayes*” [3]. Penelitian berikutnya dengan judul “Analisis Sentimen Objek Wisata Bali Di *Google Maps* Menggunakan Algoritma *Naive Bayes*” [4]. Penelitian berikutnya dengan judul “Sentimen Analysis Destinasi Wisata Kabupaten Bekasi Berdasarkan Oponi Masyarakat Menggunakan *Naive Bayes*” [5]. Kelebihan dari model Naive Bayes yang sederhana ini adalah kemampuannya bersaing dengan model algoritma lain. Selain itu, penerapannya juga sangat cepat karena

probabilitas dapat dihitung secara langsung. Algoritma ini sangat berguna saat Anda memerlukan pelatihan yang cepat.

Berdasarkan latar belakang diatas maka penulis tertarik melakukan penelitian yang berjudul **“Analisis Sentimen Objek Wisata Di Kabupaten Banggai Laut Menggunakan Metode Naive Bayes”**.

1.2 Identifikasi Masala

Berdasarkan penjelasan latar belakang diatas, permasalahan utama pada penelitian ini adalah Pemerintah banggai belum menyiapkan masyarakat yang sadarkan pariwisata.

1.3 Rumusan Masalah

Berdasarkan penjelasan latar belakang diatas, maka rumusan masalah untuk penelitian ini adalah bagaimana cara kinerja metode *naive bayes* dalam mengklasifikasikan tingkat sentimen pada objek wisata kabupaten banggai laut?

1.4 Tujuan Penelitian

Untuk mengetahui kinerja metode *naive bayes* dalam mengklasifikasikan tingkat sentimen pada objek wisata kabupaten banggai laut.

1.5 Manfaat Penelitian

1. Manfaat Teoritis

Memberikan masukan kepada Dinas Pariwisata Kabupaten Banggai Laut terhadap pengembangan ilmu pengetahuan dibidang text mining.

2. Manfaat Praktis

Sumbangan pemikiran, karya, bahan pertimbangan, atau solusi bagi Dinas Pariwisata Kabupaten Banggai Laut dalam meningkatkan daya saing destinasi pariwisata.

BAB II

LANDASAN TEORI

2.1 Tinjauan Studi

Tinjauan studi sangat berguna bagi peneliti dalam memberikan pedoman serta pegangan peneliti yang selanjutnya yang nantinya dengan adanya penelitian sebelumnya akan mempermudah peneliti dalam melakukan penelitian yang sesuai dengan topik pembahasan. Berikut terdapat beberapa jurnal yang relevan dengan penelitian ini yaitu:

Tabel 2. 1 Penelitian Terdahulu

No	Peneliti	Judul	Tahun	Metode	Hasil
1.	Khofifah Diah Indarwati, Herny Februariyanti[6]	Analisis Sentimen tehadap Kualiatas Pelayanan Aplikasi Gojek Menggunakan Metode Naïve Bayes Classifier	2023	Menggunakan metode Naïve Bayes	Hasil penelitian yang telah dilakukan mengenai analisis sentiment data ulasan pada pengguna aplikasi gojek memiliki dampak yang cukup besar dalam mencari transportasi umum

2.	Wawah Khofifah, Dhian Nur Rahayu, Arif Maulana Yusuf [7]	Analisis Sentimen Menggunakan Naïve Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps	2022	Metode yang digunakan Naïve Bayes	Dapat membantu pemerintah utnuk mengetahui review masyarakat menggunakan aplikasi RapidMiner.
3.	Yerik Afrianto Singgalen [8]	Analisis Sentimen dan Pemodelan Topik dalam Optimalisasi Pemasaran Destinasi Pariwisata Prioritas di Indonesia	2022	Metode Penelitian yang digunakan dalam kajian Literature ini ialah <i>Sistematic Literature Review</i> (SLR).	Hasil stud keputusan ini menunjukkan bahwa pendekatan analisis sentimen dan pemodelan topik bermanfaat bagi optimalisasi pemasaran destinasi pariwisata

					prioritas di Indonesia.
--	--	--	--	--	-------------------------

2.2 Tinjauan Pustaka

2.2.1 Objek Wisata

Objek wisata adalah segala sesuatu yang terdapat di daerah tujuan wisata yang menjadi daya tarik untuk mengundang orang-orang untuk datang berkunjung ke tempat tersebut. Pariwisata mencakup segala hal yang terkait dengan wisata, termasuk pengembangan atraksi pariwisata serta usaha-usaha terkait dalam bidang tersebut [1]. Tujuan utama pariwisata adalah untuk memberikan pengalaman yang menyenangkan dan menghibur, sekaligus untuk mengeksplorasi serta mempelajari keunikan tempat atau budaya yang berbeda. Selain itu, pariwisata juga dapat memberikan manfaat ekonomi, sosial, dan lingkungan bagi masyarakat setempat dan para pengunjung.

Berdasarkan definisi di atas, objek wisata adalah tempat yang dikunjungi karena keindahannya, tempat di mana aktivitas pariwisata dilakukan, tempat untuk bersenang-senang dalam waktu yang cukup lama demi mencapai kepuasan, menikmati pelayanan yang baik, dan membawa pulang kenangan indah dari tempat wisata tersebut. Di Kabupaten Banggai Laut, terdapat banyak objek wisata yang sangat menarik untuk aktivitas pariwisata, sehingga jumlah pengunjungnya meningkat setiap tahunnya.

2.2.2 Analisis Sentimen

Analisis Sentimen adalah bidang menarik yang berkembang pesat di era digital, karena saat ini masyarakat secara luas dapat menyampaikan pendapat dan pemikiran mereka melalui dunia maya dengan menggunakan bahasa teks. Namun, terkadang pembaca dapat memiliki kesalahpahaman. Analisis sentimen adalah proses untuk menentukan sentimen dan mengelompokkan polaritas teks dalam dokumen atau kalimat, sehingga dapat mengategorikannya sebagai sentimen positif, negatif, atau netral. Sentimen analisis merupakan bagian dari teks mining

dan data mining, di mana ia mengolah berbagai macam opini yang diberikan oleh masyarakat [9]. Dalam teks, sentimen dapat ditangkap pada berbagai tingkatan diantaranya pada tingkat dokumen, paragraph, kalimat, atau klausula.

Sentimen positif memberikan nilai yang baik, sementara sentimen negatif memberikan nilai yang buruk dalam konteks teks, sedangkan sentimen netral memberikan nilai yang tidak berpihak baik maupun buruk, dan biasanya kalimat dengan sentimen ini tidak mengandung opini di dalamnya. Penelitian di bidang analisis sentimen dari suatu data menjadi sangat penting dan dibutuhkan di era big data seperti saat ini. Analisis sentimen adalah tren terkini di bidang Natural Language Processing (NLP). Analisis sentimen sering menggunakan informasi dari postingan-postingan atau berita yang ada di media sosial.

2.2.3 Text Mining

Text mining adalah proses ekstraksi pola (informasi dan pengetahuan yang berguna) dari sejumlah data tak terstruktur. Proses ini bertujuan untuk mengidentifikasi pola-pola data, tren, dan ekstraksi pengetahuan yang potensial dari data teks. Text mining membantu dalam memperoleh informasi berkualitas tinggi dari teks dengan memperhatikan pola statistik dan tren [10].

Teks mining digunakan dalam teknik seperti kategorisasi, ekstraksi identitas, dan analisis sentimen untuk mengidentifikasi wawasan dan tren dalam volume data dasar yang tidak terstruktur atau semi-terstruktur. Objek dari teks mining ini biasanya berupa dokumen yang tidak terstruktur atau semi-terstruktur, dan proses informasi pada teks mining dapat menghasilkan analisis perasaan yang.

2.2.4 Preprocessing

Preprocessing dilakukan agar peneliti dapat mengolah data yang ada dengan baik, sehingga dapat menghindari gangguan terhadap data yang tidak konsisten. Tahap preprocessing adalah proses persiapan teks menjadi data yang akan diolah pada tahap berikutnya. Input awal dalam proses ini berupa dokumen utuh dengan tujuan agar hasil keluaran dari klasifikasi memiliki tingkat akurasi yang tinggi. [11] Berikut ini adalah tahapan dari preprocessing :

a) *Case Folding*

Case folding adalah salah satu bentuk *preprocessing* teks yang paling sederhana dan efektif, meskipun sering diabaikan. Hanya huruf dari a-z yang diterima, sedangkan karakter selain huruf akan dihilangkan dan dianggap sebagai *delimiter* (pemisah) [12]. Berikut adalah contoh proses *case folding* :

Tabel 2. 2 Proses *Case Folding*

Sebelum <i>Case Folding</i>	Sesudah <i>Case Folding</i>
Sektor pariwisata saat ini menjadi salah satu investasi yang penting di Banggai Laut	sektor pariwisata saat ini menjadi salah satu investasi yang penting di banggai laut
Objek wisata yang ada di Banggai Laut terbilang banyak.	objek wisata yang ada di banggai laut terbilang banyak.
Pemerintah lebih meningkatkan lagi pengembangan pariwisata yang ada di Banggai Laut.	pemerintah lebih meningkatkan lagi pengembangan pariwisata yang ada di banggai laut

b) *Tokenizing*

Tokenisasi adalah proses pemisahan teks menjadi potongan-potongan yang disebut token untuk kemudian dianalisis. Kata-kata, angka, simbol, tanda baca, dan entitas penting lainnya dapat dianggap sebagai token [12]. Contoh tahapan ini dapat dilihat pada table dibawah ini :

Tabel 2. 3 Proses *Tokenizing*

Sebelum <i>Tokenizing</i>	Sesudah <i>Tokenizing</i>
sektor pariwisata saat ini menjadi salah satu investasi yang penting di banggai laut	[sektor], [pariwisata], [saat], [ini], [menjadi], [salah], [satu], [investasi], [yang], [penting], [di], [banggai], [laut]

objek wisata yang ada di banggai laut terbilang banyak.	[objek], [wisata], [yang], [ada], [di], [banggai], [laut], [terbilang], [banyak]
pemerintah lebih meningkatkan lagi pengembangan pariwisata yang ada di banggai laut	[pemerintah], [lebih], [meningkatkan], [lagi], [pengembangan], [pariwisata], [yang], [ada], [di], [banggai], [laut]

c) *Filtering*

Filtering adalah tahap mengambil kata-kata penting dari hasil token dengan menggunakan algoritma *stoplist* (untuk mengeluarkan kata-kata kurang penting) atau *wordlist* (untuk menyimpan kata-kata penting) [12]. *Stopword* adalah kata-kata umum yang biasanya muncul dalam jumlah besar dan dianggap tidak memiliki makna. Contohnya dapat dilihat pada table di bawah ini.

Tabel 2. 4 Proses *Filtering*

Sebelum <i>Filtering Stopword</i>	Sesudah <i>Filtering Stopword</i>
sektor	Sektor
pariwisata	Pariwisata
saat	
ini	
menjadi	
salah	
satu	
investasi	Investasi
yang	

Sebelum Filtering Stopword	Sesudah Filtering Stopword
penting	Penting
di	
banggai	Banggai
laut	Laut
objek	Objek
wisata	
yang	
ada	
di	
banggai	
laut	
terbilang	
banyak	
pemerintah	Pemerintah
lebih	
meningkatkan	Meningkatkan
lagi	
pengembangan	Pengembangan
pariwisata	
yang	
ada	
di	
banggai	
laut	

d) *Stemming*

Stemming adalah proses menghilangkan infleksi kata hingga mencapai bentuk dasarnya, namun penting untuk diingat bahwa bentuk dasar tersebut tidak selalu sama dengan akar kata (*root word*) [12]. Contoh tahapan ini dapat dilihat pada gambar 3.

Tabel 2. 5 Proses *Stemming*

Sebelum <i>Stemming</i>	Sesudah <i>Stemming</i>
sektor	Sektor
pariwisata	Pariwisata
investasi	Investasi
penting	Penting
banggai	Banggai
laut	Laut
objek	Objek
meningkatkan	Tingkat
pengembangan	Kembang

2.2.5 Pembobotan TF-IDF

Term Frequency-Inverse Document Frequency (TF-IDF) adalah proses ekstraksi fitur yang memberikan bobot pada setiap kata yang terdapat dalam dokumen, biasanya dengan mengubah kata-kata tersebut menjadi nilai numerik. Algoritma TF-IDF memberikan bobot pada setiap kata kunci dalam setiap kategori untuk mengidentifikasi tingkat kemiripan antara kata kunci dan kategori yang tersedia. Langkah untuk melakukan TF-IDF adalah sebagai berikut :

1. Menentukan kata kunci
 - a. Menentukan nilai TF bias menggunakan binner, yaitu jika suatu kata terdapat dalam sebuah dokumen akan diberi nilai 1 (satu), jika tidak akan diberi nilai 0

- b. Menentukan nilai df dengan cara menghitung jumlah dokumen yang mendukung kata tertentu.
- c. Hitung nilai frekuensi dokumen terbalik (IDF) yang berguna untuk menentukan pentingnya kata-kata dalam dokumen. IDF dirumuskan sebagai berikut.

$$\text{IDF} = \log D/\text{df} \quad (1)$$

Keterangan :

Df = jumlah dokumen terhitung yang mengandung istilah

D = jumlah total dokumen

2.2.5.1 Contoh Penerapan TF-IDF

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

Persamaan 1 :

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

Persamaan 2 :

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

Ket :

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

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

Tabel 2. 6 Perhitungan TF.IDF

<i>Term</i>	<i>Tf</i>		<i>df</i>	<i>Idf</i> $1 \div df$	<i>Tf</i> \times <i>Idf</i>	
	Data 1	Data 2			Data 1	Data 2
Sektor	1	0	1	1	1	0
Pariwisata	1	0	1	1	1	0
Investasi	1	0	1	1	1	0
Penting	1	0	1	1	1	0
Banggai	1	1	2	0,5	0,5	0,5
Laut	1	1	2	0,5	0,5	0,5
Objek	0	1	1	1	0	1
Tingkat	0	0	0	0	0	0
Kembang	0	0	0	0	0	0

2.2.6 *Naïve Bayes*

Naïve Bayes adalah sebuah metode pengelompokan yang bersumber dari Teorema Bayes. Metode pengelompokan ini menggunakan probabilitas atau kemungkinan yang terjadi, dan klasifikasi pertama kali ditemukan oleh ilmuwan Inggris yang bernama Thomas Bayes. Ini berarti metode ini dapat memprediksi peluang suatu kejadian di masa depan dengan mempelajari kejadian di masa lalu, sehingga dikenal dengan sebutan Teorema Bayes. Karakteristik dari *Naïve Bayes*

adalah tingkat independensi yang kuat terhadap masing-masing kejadian atau kondisi.

Menurut pendapat Olson Delen pada tahun 2008, ia mengatakan bahwa dalam setiap pengambilan keputusan, *Naïve Bayes* menghitung peluang kemungkinan dengan syaratnya, kemudian mempertimbangkan vektor informasi objek. Keuntungan penggunaan metode *Naïve Bayes* adalah kemampuannya untuk melakukan pengklasifikasian dengan data pelatihan atau penelitian yang jumlahnya kecil untuk memperkirakan kriteria suatu operasi pengklasifikasian. Hal ini karena metode ini mengasumsikan bahwa setiap nilai adalah independen, sehingga perubahan acak pada nilai dalam suatu kelas sudah cukup untuk menentukan pengelompokan, tanpa memerlukan perhitungan terhadap seluruh matriks kovarians. Ada beberapa kegunaan dari *naïve bayes* adalah :

- Pengelompokan teks seperti teks akademis atau teks berita
- Naïve bayes adalah metode machine learning
- Menggunakan kemungkinan terjadi
- Membuat suatu analisis medis
- Menyaring atau mendeteksi spam
- Bahasa pemogamannya atau codenya bias dibingka sederhana

Adapun tahap melakukan klasifikasi menggunakan algoritma *Naïve Bayes Multinomial* sebagai berikut.

Menghitung probabilitas prior masing-masing kelas :

$$P(c) = \frac{N_c}{N} \quad (4)$$

Dimana:

N_c : Adalah jumlah c pada seluruh dokumen

N : jumlah seluruh dokumen

1. Menghitung probabilitas *term* ke-n data dokumen menggunakan persamaan sebagaimana berikut:

$$P(t_n | C) = \frac{w_{ct} + 1}{(\sum_{WF \in V} w_{ct}) + B}, \quad (5)$$

Dimana:

$P(t_n | C)$: Probabilitas *term* t_n berada di dokumen pada kelas c

W_{ct} : Bobot TF-IDF *term* t pada dokumen dengan kelas c

$\sum_w F_{\in V} W_{ct}$: Jumlah W dari keseluruhan *term* yang berada di kelas c

B' : Jumlah W (nilai idf tidak dikali tf) kata unik pada seluruh kelas

2. Menghitung apakah suatu dokumen masuk pada kelas yang mana sesuai dengan probabilitas yang diperoleh.

$$P(c|d) = P(c) \times P(t_1|c) \times P(t_2|c) \times \dots \times P(t_n|c) \quad (6)$$

Dimana:

$P(c|d)$: Probabilitas dokumen d berada di kelas c

$P(c)$: Probabilitas prior dari kelas c

t_n : Kata dokumen d ke-n

$P(t_n|c)$: Probabilitas kata ke-n dengan diketahui kelas c

3. Setelah itu akan ditentukan kelas dari data yang di proses tersebut dengan berpatokan pada probabilitas tertinggi.

2.2.6.1 Contoh Penerapan Naïve Bayes

Berikut ini merupakan contoh perhitungan manual dari *Multinomial Naïve Bayes*.

- a. Menghitung probabilitas prior masing-masing kelas menggunakan persamaan (4).

$$P_{positif} = \frac{4}{9} = 0,44 \qquad \qquad P_{negatif} = \frac{5}{9} = 0,55$$

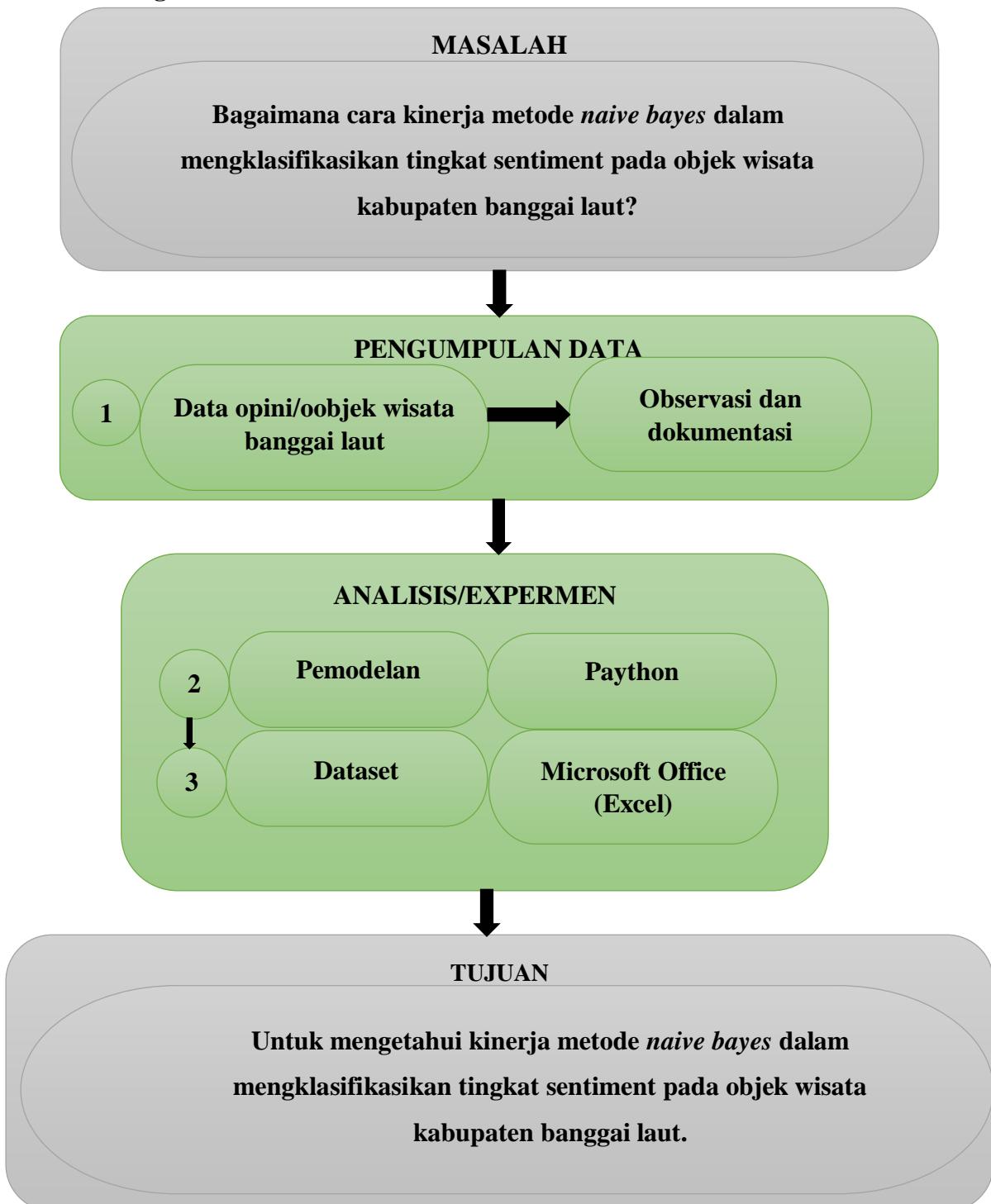
- b. Menghitung probabilitas *term* ke-n data dokumen menggunakan persamaan (5).
- c. Menghitung suatu probabilitas suatu dokumen masuk ke dalam suatu kelas tertentu menggunakan persamaan (6). Misalnya :

$$P_{(positif \mid D11)} = 0,44 \times 0,0237 \times 0,0237 \times 0,0237 \times 0,0464 \times 0,0464 \times \\ 0,0237 \times 0,0237 \times 0,0237 = 1,7404E - 13$$

$$P_{(negatif \mid D11)} = 0,55 \times 0,0164 \times 0,0164 \times 0,0164 \times 0,0164 \times 0,0164 \times \\ 0,0164 \times 0,0164 \times 0,0164 = 2,9866E - 15$$

- d. Menentukan kelas dokumen dengan memilih probabilitas tertinggi. Berdasarkan perhitungan yang dilakukan pada poin ke-3, dapat disimpulkan bahwa masuk ke dalam kelas positif.

2.3 Kerangka Pikir



BAB III

METODE PENELITIAN

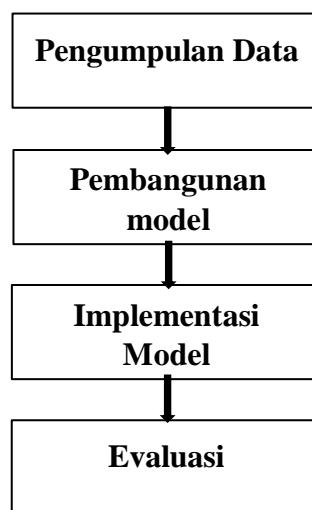
3.1 Jenis, Metode, Subjek, Objek, Waktu, dan Lokasi Penelitian

Dipandang dari tingkat penerapannya, maka penelitian ini memerlukan penelitian terapan. Berdasarkan dari jenis informasi yang diolah, maka penelitian ini merupakan penelitian kuantitatif ialah dengan menguji coba metode atau teori yang digunakan jika dilihat dari data yang diperoleh.

Dipandang dari penelitian ini menggunakan suatu penelitian dengan tujuan menemukan jawaban melalui percobaan penelitian kuantitatif. Subjek yang digunakan dari penelitian ini adalah analisis sentimen objek wisata yang ada di banggai laut yang khususnya pada data kuesioner pengunjung. Penelitian ini dimulai dari bulan Oktober sampai Februari yang berlokasi pada Dinas Pariwisata dan Kebudayaan Kabupaten Banggai Laut.

3.2 Tahap Penelitian

Berikut adalah tahap penelitian yang peneliti akan lakukan :



Gambar 3. 1 Tahap Penelitian

3.2.1 Pengumpulan Data

1. Observasi

Pengamatan atau observasi merupakan suatu kegiatan yang dilakukan pada suatu objek atau secara eksklusif atau lebih. Observasi langsung dilapangan, yaitu dengan megamati secara langsung ke lokasi pariwisata di kabupaten banggai laut.

2. Wawancara

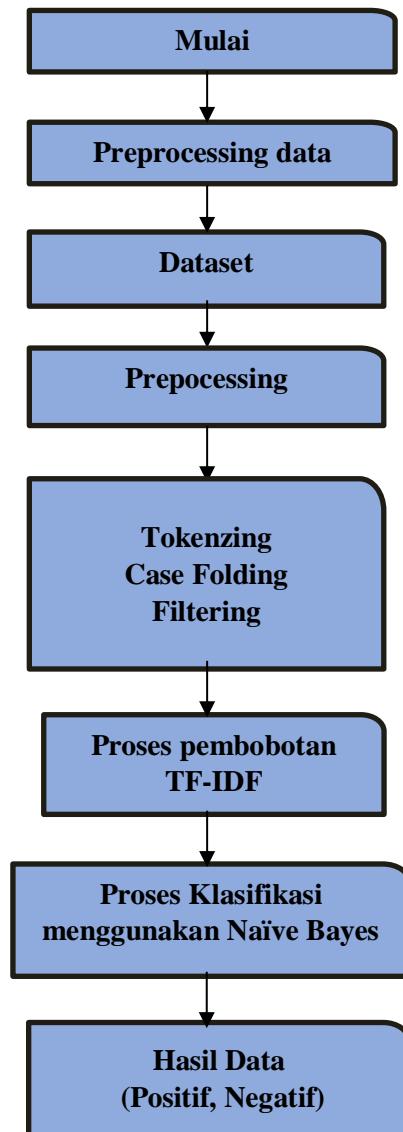
Wawancara merupakan suatu percakapan atau tahapan transfer informasi, metode ini digunakan untuk mengajukan beberapa pertanyaan kepada Karyawan Kantor Dinas Pariwisata Dan Kebudayaan Kabupaten Banggai Laut.

3. Dokumentasi

Pengarsipan atau dokumentasi adalah sebuah gambaran dari aktivitas atau kegiatan yang terstruktur secara sistematis pada saat melakukan pencarian, penyidikan, pengumpulan dan penyediaan dokumen untuk mendapatkan penjelasan pengetahuan, keterangan serta bukti validasi untuk diberikan pada pihak yang berkepentingan dan pada metode ini penulis mengumpulkan data dari jurnal atau literatur yang berkaitan serta beberapa penelitian terkait sebagai referensi sehingga dapat membantu penulis dalam menyelesaikan tugas akhir ini.

3.2.2 Pengembangan Model

Pengembangan model yang diusulkan dapat digambarkan menggunakan *flowchart* dokument yang ditunjukkan pada gambar 3.2 Berikut ini.



Gambar 3. 2 Gambar Pengembangan Model

a) *Preprocessing*

Tahap *preprocessing* ini adalah tahap yang dilakukan sebelum proses klasifikasi dimulai. Pada tahap ini, data awal diolah menjadi data yang akan digunakan dalam tahap selanjutnya. Dalam tahap *preprocessing* teks, beberapa tindakan yang dilakukan mencakup mengubah huruf kapital menjadi huruf kecil

(*Case Folding*), memisahkan kalimat menjadi kata-kata (*Tokenization*), menghapus kata-kata yang tidak relevan (*Filtering Stopword*), dan menghilangkan imbuhan dari kata-kata untuk mendapatkan kata dasar (*Stemming*) dengan menggunakan perangkat lunak *Python*.

b) **Pembobotan TF.IDF**

TF.IDF (*Term Frequency-Inverse Document Frequency*) adalah sebuah proses di mana kata atau fitur diberi bobot. TF.IDF memiliki maksud bahwa TF (*Term Frequency*) mengukur seberapa penting sebuah kata muncul dalam setiap dokumen dalam suatu koleksi dokumen. Sementara itu, IDF (*Inverse Document Frequency*) mengukur sejauh mana suatu kata umumnya digunakan dalam dokumen-dokumen tersebut. Proses pembobotan TF.IDF melibatkan perhitungan bobot antara TF yang dikalikan dengan IDF untuk menentukan seberapa sering kata tertentu muncul dalam sebuah dokumen, dengan menggunakan rumus dalam persamaan (1) hingga (2).

Proses klasifikasi berdasarkan pembobotan TF.IDF dari setiap kata bertujuan untuk menghitung nilai TF.IDF dari masing-masing kata. TF.IDF adalah metode yang dapat meningkatkan kinerja model klasifikasi teks karena memberikan bobot nilai yang berbeda pada setiap kata, sehingga nilai TF.IDF yang tinggi dapat memberikan dampak yang signifikan pada model yang diuji untuk setiap data. Setelah mendapatkan data yang cukup, dilakukan pelabelan manual terhadap respons data ke dalam kategori tanggapan positif maupun negatif. Selain itu, dilakukan pelabelan otomatis dengan tahap preprocessing menggunakan alat bahasa Python yang mengubah kata-kata berimbuhan menjadi kata dasar, yang merupakan bagian dari proses stemming. Setelah seluruh data selesai dilabeli, langkah selanjutnya adalah menghitung proses klasifikasi.

3.2.3 Implementasi Model

Hasil dari klasifikasi dalam penelitian ini adalah tingkat akurasi dari uji coba metode yang digunakan. Uji coba metode ini menghasilkan nilai yang mencakup kategori positif dan negatif serta tingkat kemiripan antar dokumen untuk mengevaluasi tanggapan dari kotak saran terhadap kepuasan pengunjung terhadap objek wisata yang ada di Banggai Laut.

3.2.4 Evaluasi Model

Pada tahap evaluasi ini, dilakukan untuk mengukur kinerja metode yang digunakan dengan menguji sejauh mana metode tersebut akurat dalam hasil klasifikasinya, dengan menggunakan proses perhitungan *Naïve Bayes*.

BAB IV

HASIL PENELITIAN

4.1 Hasil Pengumpulan Data

Data pada penelitian ini merupakan analisis sentimen mengenai komentar masyarakat terhadap objek wisata di Banggai Laut, terdiri dari 73 data. Data ini terbagi menjadi dua kelompok, yaitu 41 data positif dan 32 data negatif. Kemudian, data tersebut akan diproses dan dianalisis berdasarkan model yang telah dibangun oleh peneliti.

Komentar masyarakat terhadap objek wisata di Banggai Laut diambil secara online melalui *Google Form*, kemudian penulis melakukan pemberian sentimen negatif dan positif pada data komentar dengan ahli bahasa Indonesia agar data yang peneliti kumpulkan lebih akurat dalam pemberian sentimen, dan kemudian disalin ke dalam file Excel untuk dimasukkan ke dalam database. Data dari file Excel ini dipilih dan dilabeli secara manual terlebih dahulu untuk menghapus data yang kosong dan tidak sesuai untuk diproses. Selanjutnya, komentar yang ada dalam database tersebut akan dilakukan *preprocessing* dan disimpan dalam database data bersih.

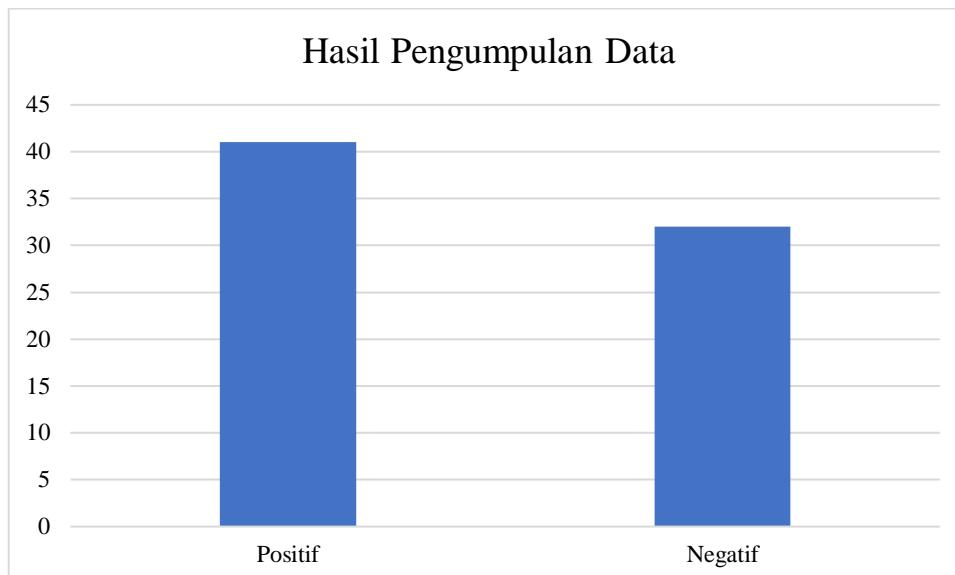
Berdasarkan hasil penggumpulan data, diperoleh data primer sebagai berikut :

Tabel 4. 1 Hasil Pengumpulan Data

No	Komentar	Sentimen
1	Menurut pendapat saya, objek wisata yang berada di kabupaten banggai laut sangatlah banyak, dan sangat bagus serta sangat indah.	Positif
2	Menurut saya objek wisata yang ada di Kabupaten Banggai Laut itu memiliki keindahan alam yang sangat indah dan bisa memanjakan mata bagi para pengujung.	Positif

3	Objek wisata yang ada di banggai laut sudah banyak peningkatan dari berbagai objek. Sehingga kedepannya diharapkan objek wisata ini dapat memberikan kontribusi yang besar bagi banggai laut.	Positif
4	Wisata yang ada di kabupaten banggai laut menurut saya sudah cukup berkembang, terutama dalam membantu pendapatan daerah dari wisata itu masing-masing	Positif
5	Menurut saya, objek wisata di kabupaten Banggai laut sangat bagus dan keren, di mana banyak sekali tempat wisata yang bisa mempunyai keragaman yang sangat bagus.	Positif
6	Beberapa wisata di kabupaten Banggai laut sangat indah dan harus tetap di kembangkan	Positif
7	objek wisata yang ada di banggai laut sangat menarik dari wisata pantai oyama, pantai lambangan pauna dan masih banyak lagi wisata yang ada di banggai laut yang menarik	Positif
8	objek wisata di Banggai Laut sangat luar biasa keindahannya. Mulai dari pantai yang amat luas memanjang, pasir yang putih dan halus serta air laut yang jernih.	Postif
9	Objek wisata di kabupaten Banggai Laut belum bisa di nikmati banyak orang dikarenakan kurangnya pemanfaatan peluang yang ada	Negatif
10	objek wisata yang ada di banggai laut belum banyak orang yang mengetahuinya belum lagi akses yang lumayan susah untuk sampai di tempat-tempat objek wisata tersebut	Negatif
11	Objek wisata di banggai laut patut dikunjungi oleh semua orang apalagi yang dari luar banggai laut karena keberagamannya dan masih asri dari alam	Positif
12	Kabupaten Banggai Laut memiliki banyak Objek wisata alam yang keindahannya sangat luar biasa dan layak untuk dikunjungi oleh wisatawan dalam dan luar negeri	Positif

13	Objek wisata di banggai laut sangat banyak namun masih kurang di publikasi	Negatif
14	Dengan adanya objek wisata yang ada di banggai laut membuat penambahan penduduk akibat dari pendatang baru yang datang dari luar daerah	Negatif
15	beberapa pengunjung mengeluh tentang kurangnya infrastruktur pendukung dan kurangnya fasilitas yang memadai di objek wisata yang ada di Banggai Laut.	Negatif
....
73	Kabupaten Banggai Laut memiliki peluang besar untuk menjadi destinasi unggulan yang menarik minat wisatawan lokal maupun internasional.	Positif

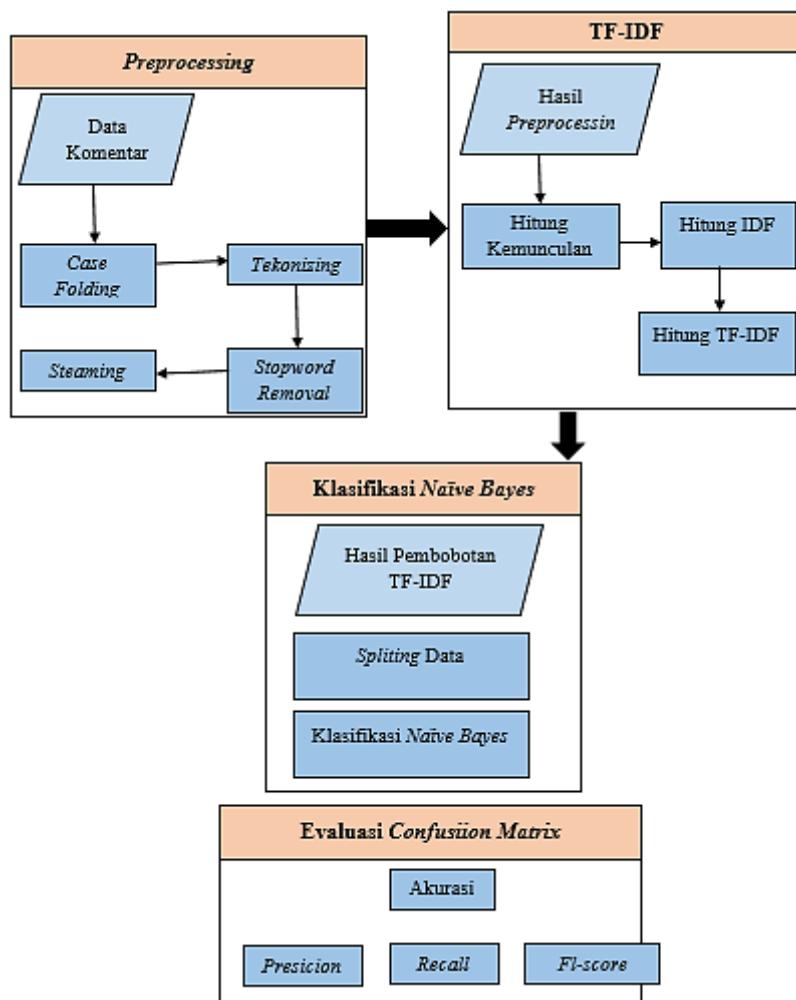


Gambar 4. 1 Hasil Pengumpulan Data

Ket : Dari 73 data di atas dapat diketahui jumlah data dibagi menjadi dua yaitu Negatif sebanyak 32 data dan Positif sebanyak 41 data.

4.2 Hasil Pemodelan

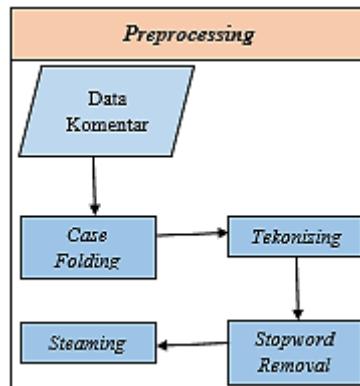
Pembangunan model dalam penelitian ini dilakukan secara terstruktur, dimulai dari tahap *preprocessing*, pembobotan TF-IDF, klasifikasi *Naïve Bayes*, dan diakhiri dengan evaluasi menggunakan matriks kebingungan (*confusion matrix*).



Gambar 4. 2 Pemodelan

4.2.1 Hasil *Preprocessing*

Tahap selanjutnya setelah dataset siap untuk diolah adalah melakukan tahap *preprocessing*. Pada tahap ini, proses dibagi menjadi beberapa tahapan, yaitu *Case Folding*, *Tokenizing*, *Stopword Removal*, dan *Stemming*.



Gambar 4. 3 Flowchart Preprocessing

Dalam tahap *preprocessing* yang dilakukan menggunakan Bahasa pemrograman *Python* pada *Streamlit*, diperoleh hasil sebagai berikut:

1) *Case Folding*

Pada tahap ini, seluruh komentar akan diubah menjadi huruf kecil secara keseluruhan. Berikut adalah hasil dari *Case Folding* untuk Komentar 1-10.

Tabel 4. 2 Hasil Case Folding

No	Komentar	Hasil <i>Case Folding</i>
1	Menurut pendapat saya, objek wisata yang berada di kabupaten banggai laut sangatlah banyak, dan sangat bagus serta sangat indah.	menurut pendapat saya, objek wisata yang berada di kabupaten banggai laut sangatlah banyak, dan sangat bagus serta sangat indah.
2	Menurut saya objek wisata yang ada di Kabupaten Banggai Laut	menurut saya objek wisata yang ada di kabupaten banggai laut itu memiliki keindahan alam yang sangat

	itu memiliki keindahan alam yang sangat indah dan bisa memanjakan mata bagi para pengujung.	indah dan bisa memanjakan mata bagi para pengujung.
3	Objek wisata yang ada di banggai laut sudah banyak peningkatan dari berbagai objek. Sehingga kedepannya diharapkan objek wisata ini dapat memberikan kontribusi yang besar bagi banggai laut.	objek wisata yang ada di banggai laut sudah banyak peningkatan dari berbagai objek. sehingga kedepannya diharapkan objek wisata ini dapat memberikan kontribusi yang besar bagi banggai laut.
4	Wisata yang ada di kabupaten banggai laut menurut saya sudah cukup berkembang, terutama dalam membantu pendapatan daerah dari wisata itu masing-masing	wisata yang ada di kabupaten banggai laut menurut saya sudah cukup berkembang, terutama dalam membantu pendapatan daerah dari wisata itu masing-masing
5	Menurut saya, objek wisata di kabupaten Banggai laut sangat bagus dan keren, di mana banyak sekali tempat wisata yang bisa mempunyai keragaman yang sangat bagus.	menurut saya, objek wisata di kabupaten banggai laut sangat bagus dan keren, di mana banyak sekali tempat wisata yang bisa mempunyai keragaman yang sangat bagus

6	Beberapa wisata di kabupaten Banggai laut sangat indah dan harus tetap di kembangkan	beberapa wisata di kabupaten banggai laut sangat indah dan harus tetap di kembangkan
7	objek wisata yang ada di banggai laut sangat menarik dari wisata pantai oyama, pantai lambangan pauna dan masih banyak lagi wisata yang ada di banggai laut yang menarik	objek wisata yang ada di banggai laut sangat menarik dari wisata pantai oyama, pantai lambangan pauna dan masih banyak lagi wisata yang ada di banggai laut yang menarik
8	objek wisata di Banggai Laut sangat luar biasa keindahannya. Mulai dari pantai yang amat luas memanjang, pasir yang putih dan halus serta air laut yang jernih.	objek wisata di banggai laut sangat luar biasa keindahannya. mulai dari pantai yang amat luas memanjang, pasir yang putih dan halus serta air laut yang jernih
9	Objek wisata di kabupaten Banggai Laut belum bisa di nikmati banyak orang dikarenakan kurangnya pemanfaatan peluang yang ada	objek wisata di kabupaten banggai laut belum bisa di nikmati banyak orang dikarenakan kurangnya pemanfaatan peluang yang ada
10	objek wisata yang ada di banggai laut belum banyak orang yang mengetahuinya belum	objek wisata yang ada di banggai laut belum banyak orang yang mengetahuinya belum lagi akses yang

	lagi akses yang lumayan susah untuk sampai di tempat-tempat objek wisata tersebut	lumayan susah untuk sampai di tempat-tempat objek wisata tersebut
--	-----------------------------------------------------------------------------------	-------------------------------------------------------------------

2) *Ekstrasi Fitur Ikon*

Menggunakan metode Naive Bayes untuk mengelola kalimat yang mengandung ikon memerlukan beberapa tahap pemrosesan dan ekstraksi fitur. Ikon dan emotikon harus diubah menjadi representasi teks atau fitur khusus yang dapat diolah oleh algoritma pembelajaran mesin. Dengan cara ini, model dapat memahami dan mengklasifikasikan teks dengan ikon secara efektif.

3) *Tokenizing*

Pada tahap tokenisasi ini, setiap kalimat dalam setiap komentar akan dipenggal menjadi kata per kata. Berikut adalah hasil proses tokenisasi dari komentar 1-10.

Tabel 4. 3 Hasil dari *Tokenizing*

No	Hasil <i>Case Folding</i>	Hasil <i>Tokenizing</i>
1	menurut pendapat saya, objek wisata yang berada di kabupaten banggai laut sangatlah banyak, dan sangat bagus serta sangat indah.	menurut pendapat saya objek wisata yang berada di kabupaten banggai laut

		sangatlah banyak dan sangat bagus serta sangat indah
2	menurut saya objek wisata yang ada di kabupaten banggai laut itu memiliki keindahan alam yang sangat indah dan bisa memanjakan mata bagi para pengujung.	menurut saya objek wisata yang ada di kabupaten banggai laut itu memiliki keindahan alam yang sangat indah dan bisa memanjakan mata bagi para

		pengujung
3	objek wisata yang ada di banggai laut sudah banyak peningkatan dari berbagai objek. sehingga kedepannya diharapkan objek wisata ini dapat memberikan kontribusi yang besar bagi banggai laut.	objek wisata yang ada di banggai laut sudah banyak peningkatan dari berbagai objek sehingga kedepannya diharapkan objek wisata ini dapat memberikan kontribusi yang besar bagi banggai laut
4	wisata yang ada di kabupaten banggai laut menurut saya	wisata yang

	sudah cukup berkembang, terutama dalam membantu pendapatan daerah dari wisata itu masing-masing	ada di kabupaten banggai laut menurut saya sudah cukup berkembang terutama dalam membantu pendapatan daerah dari wisata itu masing masing
5	menurut saya, objek wisata di kabupaten banggai laut sangat bagus dan keren, di mana banyak sekali tempat wisata yang bisa mempunyai keragaman yang sangat bagus	menurut saya objek wisata di kabupaten banggai laut sangat bagus dan

		<p>keren di mana banyak sekali tempat wisata yang bisa mempunyai keragaman yang sangat bagus</p>
6	beberapa wisata di kabupaten banggai laut sangat indah dan harus tetap di kembangkan	<p>beberapa wisata di kabupaten banggai laut sangat indah dan harus tetap di kembangkan</p>
7	objek wisata yang ada di banggai laut sangat menarik dari wisata pantai oyama,	<p>objek wisata yang</p>

	pantai lambangan pauna dan masih banyak lagi wisata yang ada di banggai laut yang menarik	ada di banggai laut sangat menarik dari wisata pantai oyama pantai lambangan pauna dan masih banyak lagi wisata yang ada di banggai laut yang menarik
8	objek wisata di banggai laut sangat luar biasa keindahannya. mulai dari pantai yang amat luas memanjang, pasir yang putih	objek wisata di banggai laut sangat

	dan halus serta air laut yang jernih	luar biasa keindahannya mulai dari pantai yang amat luas memanjang pasir yang putih dan halus serta air laut yang jernih
9	objek wisata di kabupaten banggai laut belum bisa di nikmati banyak orang dikarenakan kurangnya pemanfaatan peluang yang ada	objek wisata di kabupaten banggai laut belum bisa di nikmati banyak

		orang dikarenakan kurangnya pemanfaatan peluang yang ada
10	objek wisata yang ada di banggai laut belum banyak orang yang mengetahuinya belum lagi akses yang lumayan susah untuk sampai di tempat-tempat objek wisata tersebut	objek wisata yang ada di banggai laut belum banyak orang yang mengetahuinya belum lagi akses yang lumayan susah untuk sampai di tempat tempat objek

		wisata tersebut
--	--	--------------------

4) *Stopword Removal*

Pada tahap ini, dilakukan penghapusan kata-kata yang cenderung tidak memiliki arti. Berikut adalah hasil dari tahap Penghapusan Stopword dari Komentar 1-10.

Tabel 4. 4 Hasil dari *Stopword Removal*

No	Hasil Tokenizing	Hasil Stopword Removal
1	menurut pendapat saya objek wisata yang berada di kabupaten banggai laut sangatlah banyak dan sangat bagus serta sangat indah	pendapat objek wisata berada kabupaten banggai laut sangatlah banyak sangat bagus sangat indah
2	menurut saya	objek wisata

	objek wisata yang ada di kabupaten banggai laut itu memiliki keindahan alam yang sangat indah dan bisa memanjakan mata bagi para pengujung	kabupaten banggai laut memiliki keindahan alam sangat indah memanjakan mata pengujung
3	objek wisata yang ada di banggai laut sudah banyak	objek wisata banggai laut banyak peningkatan berbagai objek kedepannya

	<p>peningkatan dari berbagai objek sehingga kedepannya diharapkan objek wisata ini dapat memberikan kontribusi yang besar bagi banggai laut</p>	<p>diharapkan objek wisata memberikan kontribusi besar banggai laut</p>
4	<p>wisata yang ada di kabupaten banggai laut menurut saya sudah cukup berkembang terutama</p>	<p>wisata kabupaten banggai laut cukup berkembang terutama membantu pendapatan daerah wisata masing masing</p>

	dalam membantu pendapatan daerah dari wisata itu masing masing	
5	menurut saya objek wisata di kabupaten banggai laut sangat bagus keren mana banyak dan keren di mana banyak sekali tempat wisata yang bisa mempunyai keragaman	objek wisata kabupaten banggai laut sangat bagus keren mana banyak sekali tempat wisata mempunyai keragaman sangat bagus

	yang sangat bagus	
6	Beberapa wisata di kabupaten banggai laut sangat indah dan harus tetap di kembangkan	beberapa wisata kabupaten banggai laut sangat indah tetap kembangkan
7	objek wisata yang ada di banggai laut sangat menarik dari wisata pantai oyama pantai	objek wisata banggai laut sangat menarik wisata pantai oyama pantai lambangan pauna banyak wisata

	lambangan pauna dan masih banyak lagi wisata yang ada di banggai laut yang menarik	banggai laut menarik
8	objek wisata di banggai laut sangat luar biasa keindahannya mulai pantai dari luas memanjang yang amat pasir halus air laut	objek wisata banggai laut sangat luar biasa keindahannya mulai pantai luas memanjang pasir putih halus air laut

	yang putih dan halus serta air laut yang jernih	jernih
9	objek wisata di kabupaten banggai laut belum bisa di nikmati banyak orang dikarenakan kurangnya pemanfaatan peluang yang ada	objek wisata kabupaten banggai laut nikmati banyak orang dikarenakan kurangnya pemanfaatan peluang
10	objek wisata yang	objek wisata banggai

	ada di banggai laut belum banyak orang yang mengetahuinya belum lagi akses yang lumayan susah untuk sampai di tempat tempat objek wisata tersebut	laut banyak orang mengetahuinya akses lumayan susah tempat tempat objek wisata tersebut
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------

5) *Stemming*

Pada tahap ini, akan dilakukan penghapusan terhadap imbuhan dari setiap kata, baik yang terletak di depan maupun di belakang, untuk mengubah setiap kata menjadi kata dasarnya. Berikut ini adalah hasil dari tahapan Stemming.

Tabel 4. 5 Hasil dari *Stemming*

No	Hasil Stopword Removal	Hasil Stemming
1	pendapat objek wisata berada kabupaten banggai laut sangatlah banyak sangat bagus sangat indah	dapat objek wisata ada kabupaten bangga laut sangat banyak sangat bagus sangat indah
2	objek wisata kabupaten banggai laut memiliki keindahan alam sangat indah memanjakan mata pengujung	objek wisata kabupaten bangga laut milik indah alam sangat indah manja mata ujung
3	objek wisata banggai	objek wisata bangga

	laut banyak peningkatan berbagai objek kedepannya diharapkan objek wisata memberikan kontribusi besar banggai laut	laut banyak tingkat bagai objek depan harap objek wisata beri kontribusi besar bangga laut
4	wisata kabupaten banggai laut cukup berkembang terutama membantu pendapatan daerah wisata masing masing	wisata kabupaten bangga laut cukup kembang utama bantu dapat daerah wisata masing masing
5	objek wisata kabupaten	objek wisata kabupaten

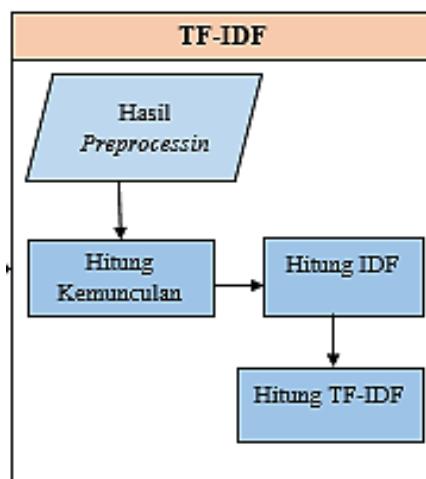
	banggai laut sangat bagus keren mana banyak sekali tempat wisata mempunyai keragaman sangat bagus	bangga laut sangat bagus keren mana banyak sekali tempat wisata punya ragam sangat bagus
6	beberapa wisata kabupaten banggai laut sangat indah tetap kembangkan	beberapa wisata kabupaten bangga laut sangat indah tetap kembang
7	objek wisata banggai laut sangat menarik wisata	objek wisata bangga laut sangat tarik wisata

	pantai oyama pantai lambangan pauna banyak wisata banggai laut menarik	pantai oyama pantai lambang pauna banyak wisata bangga laut tarik
8	objek wisata banggai laut sangat luar biasa keindahannya mulai pantai luas memanjang pasir putih halus air laut jernih	objek wisata bangga laut sangat luar biasa indah mulai pantai luas panjang pasir putih halus air laut jernih
9	objek	objek

	wisata kabupaten banggai laut nikmati banyak orang dikarenakan kurangnya pemanfaatan peluang	wisata kabupaten bangga laut nikmat banyak orang karena kurangnya manfaat peluang
10	objek wisata banggai laut banyak orang mengetahuinya akses lumayan susah tempat tempat objek wisata tersebut	objek wisata bangga laut banyak orang tahu akses lumayan susah tempat tempat objek wisata sebut

4.2.2 Hasil Pembobotan TF.IDF

Setelah semua tahapan preprocessing selesai dilakukan, langkah selanjutnya adalah melakukan ekstraksi fitur. Dalam hal ini, peneliti menggunakan metode pembobotan TF-IDF yang bertujuan untuk memberikan bobot pada setiap kalimat. Pada pembobotan TF-IDF, dilakukan penggabungan dua konsep perhitungan, yaitu frekuensi kata dan invers.



Gambar 4. 4 Flowchart TF-IDF

1) *Term Frequency (TF)*

Pada tahap ini, akan dilakukan perhitungan frekuensi kemunculan kata berdasarkan term yang diperoleh. Jadi, dalam tahap ini akan ditentukan berapa kali suatu kata muncul dalam sebuah Komentar. Sebagai contoh, kata atau term “objek” muncul sebanyak 2 kali dalam komentar pertama atau D1, dan muncul sebanyak 1 kali dalam Komentar 3 atau D3. Apabila term tersebut tidak muncul dalam komentar atau dokumen lainnya, maka akan ditulis sebagai 0. Hal yang sama juga berlaku untuk term-term dalam dokumen-dokumen lainnya.

Berikut merupakan perhitungan *Term Frequency* dari 10 dokumen, yaitu komentar 1 hingga komentar 10 (D1 sampai D10), sebagai contoh perhitungan manual dari TF-IDF.

Tabel 4. 6 Sampel Term dan Perhitungan manual TF

TREM	TF									
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
dapat	1	0	0	1	0	0	0	0	0	0
objek	1	1	3	0	1	0	1	1	1	2
wisata	1	1	2	2	2	1	3	1	1	2
ada	1	0	0	0	0	0	0	0	0	0
kabupaten	1	1	0	1	1	1	0	0	1	0
bangga	1	1	2	1	1	1	2	1	1	1
laut	1	1	2	1	1	1	2	2	1	1
sangat	3	1	0	0	2	1	1	1	0	0
banyak	1	0	1	0	1	0	1	0	1	1
bagus	1	0	0	0	2	0	0	0	0	0
indah	1	2	0	0	0	1	0	1	0	0
milik	0	1	0	0	0	0	0	0	0	0
alam	0	1	0	0	0	0	0	0	0	0
manja	0	1	0	0	0	0	0	0	0	0
mata	0	1	0	0	0	0	0	0	0	0
ujung	0	1	0	0	0	0	0	0	0	0
tingkat	0	0	1	0	0	0	0	0	0	0
bagai	0	0	1	0	0	0	0	0	0	0
depan	0	0	1	0	0	0	0	0	0	0
harapan	0	0	1	0	0	0	0	0	0	0
beri	0	0	1	0	0	0	0	0	0	0
kontribusi	0	0	1	0	0	0	0	0	0	0
besar	0	0	1	0	0	0	0	0	0	0
cukup	0	0	0	1	0	0	0	0	0	0
kembang	0	0	0	1	0	0	0	0	0	0
utama	0	0	0	1	0	0	0	0	0	0
bantu	0	0	0	1	0	0	0	0	0	0
daerah	0	0	0	1	0	0	0	0	0	0

tahu	0	0	0	0	0	0	0	0	0	1
akses	0	0	0	0	0	0	0	0	0	1
lumayan	0	0	0	0	0	0	0	0	0	1
susah	0	0	0	0	0	0	0	0	0	1
sebut	0	0	0	0	0	0	0	0	0	1
Panjang Dok	13	13	17	13	17	8	17	18	12	14

Setelah mendapatkan hasil dari *Term Frequency*, langkah selanjutnya adalah melakukan normalisasi pada *Term Frequency* tersebut. Normalisasi dilakukan karena panjang dokumen atau komentar tidak selalu sama, yang dapat memengaruhi nilai TF. Oleh karena itu, normalisasi dilakukan dengan membagi hasil TF sebelumnya dengan panjang dari masing-masing dokumen atau komentar. Berikut ini adalah hasil dari TF Normalisasi.

Tabel 4. 7 Sampel Perhitungan Manual TF Normalisasi

0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,077	0	0	0	0	0	0	0
0	0	0,077	0	0	0	0	0	0	0
0	0	0,077	0	0	0	0	0	0	0
0	0	0,077	0	0	0	0	0	0	0
0	0	0,077	0	0	0	0	0	0	0
0	0	0,077	0	0	0	0	0	0	0
0	0	0,154	0	0	0	0	0	0	0
0	0	0	0	0,059	0	0	0	0	0
0	0	0	0	0,059	0	0	0	0	0
0	0	0	0	0,059	0	0	0	0	0
0	0	0	0	0,059	0	0	0	0	0,143
0	0	0	0	0,059	0	0	0	0	0
0	0	0	0	0,059	0,143	0	0	0	0
0	0	0	0	0	0,143	0	0	0	0
0	0	0	0	0	0	0,118	0	0	0
0	0	0	0	0	0	0,118	0,056	0	0
0	0	0	0	0	0	0,059	0	0	0
0	0	0	0	0	0	0,059	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0
0	0	0	0	0	0	0,056	0	0	0,083
0	0	0	0	0	0	0,083	0	0	0

0	0	0	0	0	0	0	0	0,083	0
0	0	0	0	0	0	0	0	0,083	0
0	0	0	0	0	0	0	0	0,083	0
0	0	0	0	0	0	0	0	0,083	0
0	0	0	0	0	0	0	0	0,083	0
0	0	0	0	0	0	0	0	0,071	
0	0	0	0	0	0	0	0	0,071	
0	0	0	0	0	0	0	0	0,071	
0	0	0	0	0	0	0	0	0,071	
0	0	0	0	0	0	0	0	0,071	

2) *Inverse Document Frequensi (IDF)*

Pada tahap ini, dilakukan perhitungan IDF. Sebelum proses perhitungan IDF dilakukan, terlebih dahulu dilakukan perhitungan DF. DF merupakan jumlah kemunculan suatu term dalam dokumen. Sebagai contoh, jika term “objek” muncul pada 2 dokumen, maka nilai DF untuk term tersebut diisi dengan angka 2, tanpa memperhatikan frekuensi kemunculannya. IDF adalah hasil dari pembagian jumlah total dokumen dengan nilai DF untuk setiap term. Berikut adalah hasil perhitungan DF dan IDF.

Tabel 4. 8 Sampel Perhitungan Manual DF dan IDF

TREM	DF	IDF
dapat	2	0,699
objek	8	0,097
wisata	10	0
ada	1	1
kabupaten	6	0,222
bangga	10	0
laut	10	0
sangat	6	0,222

banyak	6	0,222
bagus	2	0,699
indah	4	0,398
milik	1	1
alam	1	1
manja	1	1
mata	1	1
ujung	1	1
tingkat	1	1
bagai	1	1
depan	1	1
harapan	1	1
beri	1	1
kontribusi	1	1
besar	1	1
cukup	1	1
kembang	1	1
utama	1	1
bantu	1	1
daerah	1	1
masing	1	1
keren	1	1
mana	1	1
sekali	1	1
tempat	2	0,699
punya	1	1
ragam	1	1
beberapa	1	1
tetap	1	1
tarik	1	1

pantai	2	0,699
oyama	1	1
lambang	1	1
pauna	1	1
luar	1	1
biasa	1	1
mulai	1	1
luas	1	1
panjang	1	1
pasir	1	1
putih	1	1
halus	1	1
air	1	1
jernih	1	1
nikmat	1	1
orang	1	1
karena	1	1
kurangnya	1	1
manfaat	1	1
peluang	1	1
tahu	1	1
akses	1	1
lumayan	1	1
susah	1	1
sebut	1	1

3) Perhitungan TF-IDF

Berikut adalah langkah terakhir dari pembobotan TF-IDF, yaitu perhitungan hasil TF-IDF yang diperoleh dari perkalian antara nilai TF

yang telah dinormalisasi dan nilai IDF. Berikut adalah hasil perhitungan manual TF-IDF menggunakan 10 dokumen.

Tabel 4. 9 Sampel Perhitungan Manual TF-IDF

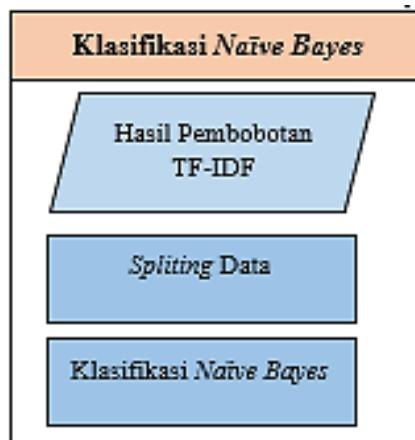
TF-IDF									
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
0,054	0	0	0,054	0	0	0	0	0	0
0,007	0,007	0,017	0	0,006	0	0,006	0,005	0,008	0,014
0	0	0	0	0	0	0	0	0	0
0,077	0	0	0	0	0	0	0	0	0
0,017	0,017	0	0,017	0,013	0,032	0	0	0,018	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0,051	0,017	0	0	0,026	0,032	0,013	0,012	0	0
0,017	0	0,013	0	0,013	0	0,013	0	0,018	0,016
0,054	0	0	0	0,082	0	0	0	0	0
0,031	0,061	0	0	0	0,057	0	0,022	0	0
0	0,077	0	0	0	0	0	0	0	0
0	0,077	0	0	0	0	0	0	0	0
0	0,077	0	0	0	0	0	0	0	0
0	0,077	0	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0	0	0	0	0	0	0
0	0	0,059	0,077	0	0	0	0	0	0
0	0	0	0,077	0	0	0	0	0	0
0	0	0	0,077	0	0	0	0	0	0
0	0	0	0,077	0	0	0	0	0	0
0	0	0	0,077	0	0	0	0	0	0
0	0	0	0,154	0	0	0	0	0	0

Ket: Penting untuk dicatat bahwa perhitungan TF-IDF ini merupakan kelanjutan dari perhitungan TF, dimulai dari langkah awal perhitungan TF

hingga penyelesaian perhitungan TF-IDF. Oleh karena itu, jika sebuah term tidak muncul pada dokumen tertentu, bobotnya akan ditetapkan sebagai "0", dan hal ini berlaku hingga pada hasil akhir dari perhitungan TF-IDF itu sendiri.

4.2.3 Klasifikasi *Naïve Bayes*

Setelah dokumen-dokumen tersebut diberi bobot menggunakan TF-IDF, langkah selanjutnya adalah melakukan proses klasifikasi. Dalam penelitian ini, penulis melakukan klasifikasi menggunakan algoritma Naïve Bayes Classifier dengan menggunakan Library Scikit-learn.



Gambar 4. 5 Flowchart Klasifikasi *Naïve Bayes*

Sebelum melakukan klasifikasi, dataset yang ada akan ditraining terlebih dahulu. Dalam konteks ini, penelitian akan melakukan proses pembagian data, yang disebut splitting data, dengan membagi dataset menjadi data training dan data uji. Dalam proses splitting data ini, peneliti menggunakan variabel pengatur (sliders) agar pembagian data dapat disesuaikan secara fleksibel, dimulai dari pembagian: 90%:10%, 80%:20%, 70%:30%, dan 60%:40%. Ketika lebih kecil pembagian maka nilai yang dihasilkan lebih kecil juga dalam klasifikasi ini peneliti mengambil kesimpulan pembagian 90%:10%.

Cara kerja *naive bayes* dalam melakukan proses klasifikasi dapat dilihat pada contoh komentar “beberapa wisata dikabupaten banggai laut sangat indah dan harus tetap di kembangkan”. Komentar ini kemudian akan dilakukan

preprocessing sehingga menjadi “beberapa wisata kabupaten bangga laut sangat indah tetap kembang”. Setelah itu menghitung probabilitas prior masing-masing seperti berikut.

$$P_{positif} = \frac{8}{10} = 0,80 \quad P_{negatif} = \frac{2}{10} = 0,020$$

Kemudian menghiung probabilitas *term* ke-n data dokumen menggunakan persamaan (5). Kemudian menghitung suatu probabilitas suatu dokumen masuk ke dalam suatu kelas tertentu menggunakan persamaan (6). Seperti berikut :

$$\begin{aligned} P_{(positif | D6)} &= P(\text{positif}) \times P(\text{beberapa}|\text{positif}) \times \dots p(\text{kembang}|\text{positif}) \\ &= 1,4904E - 7 \end{aligned}$$

$$\begin{aligned} P_{(negatif | D6)} &= P(\text{negatif}) \times P(\text{beberapa}|\text{negatif}) \times \dots p(\text{kembang}|\text{negatif}) \\ &= 0,0023 \end{aligned}$$

Dari dua perhitungan probabilitas diatas maka dapat diketahui probabilitas kemunculan dari kategori positif lebih besar dari kategori negatif. Oleh karena itu contoh tersebut diklasifikasikan ke kategori positif.

BAB V

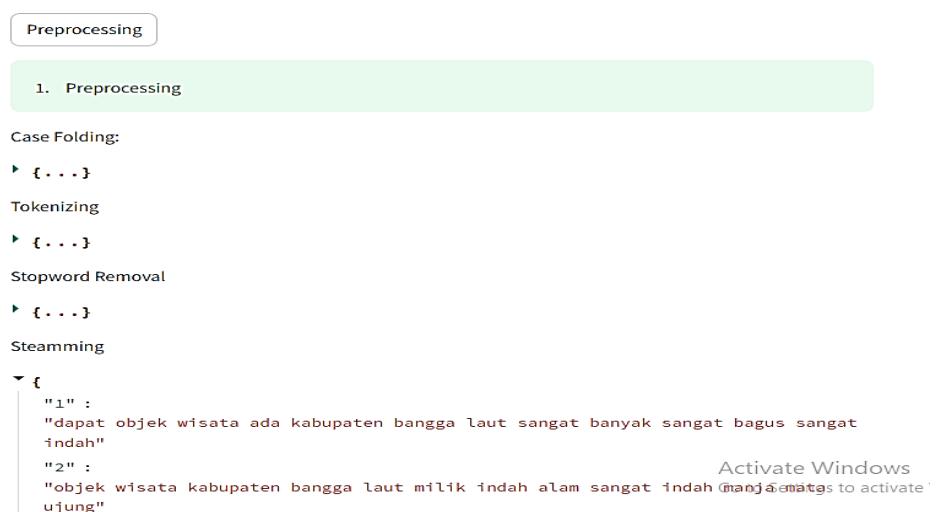
PEMBAHASAN

5.1 Pembahasan Model

Pada pengembangan model penelitian ini, kami memanfaatkan *Framework Streamlit* berbasis *Python* untuk melakukan klasifikasi dengan algoritma *Naïve Bayes*. Penggunaan Streamlit ini bersifat *open source* dan dirancang khusus untuk membangun aplikasi *web* interaktif dalam bidang sains data dan *machine learning*.

5.1.1 Preprocessing

Tahap preprocessing ini dilakukan sebelum proses klasifikasi dimulai. Pada tahap ini, data awal diolah menjadi bentuk yang akan digunakan dalam tahap selanjutnya. Dalam *preprocessing* teks, langkah-langkah yang dilakukan termasuk mengubah huruf kapital menjadi huruf kecil (*Case Folding*), memisahkan kalimat menjadi kata-kata (*Tokenization*), menghapus kata-kata yang tidak relevan (*Filtering Stopword*), dan menghilangkan imbuhan dari kata-kata untuk mendapatkan kata dasar (*Stemming*) menggunakan perangkat lunak *Python*. Berikut adalah tahapan *Preprocessing* dalam pemodelan:



Gambar 5. 1 Preprocessing

5.1.1 Pembobotan

Pada tahap pembobotan kata, penulis menggunakan metode TF-IDF ini dengan tujuan memberikan bobot pada setiap kata, sehingga hasil yang diperoleh dari pemodelan adalah sebagai berikut:

	ada	adalah	adanya	agar	air	akan	akibat	akomodasi	akses	alam	amat	apalagi	area	asri	atau	bagi	bagus	bagus ²	baik	bakau	balut	banggai	banyak	baru		
7	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	
8	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	
9	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
10	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
11	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	2	0	0
12	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
14	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
15	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
16	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
17	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0
19	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
21	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	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 5. 2 Pembobotan

5.1.2 Analisis

Peneliti menggunakan metode *Naïve Bayes* untuk melakukan klasifikasi dengan memanfaatkan pustaka *Scikit-learn* pada kerangka kerja Streamlit. Berikut adalah gambaran dari proses analisis yang dilakukan.

The screenshot shows a Streamlit application interface. At the top, there is a button labeled "Mulai Analisis". Below it, a green box contains the text "3. Naive Bayes". Underneath this box, the text "Kernel GaussianNB" is displayed. Following that, "Data Testing Sebenarnya" is shown, with a code snippet:

```
<NumpyExtensionArray> ['Negatif', 'Positif', 'Negatif', 'Positif', 'Negatif', 'Negatif', 'Negatif', 'Negatif'] Length: 8, dtype: object
```

Next is "Data Testing Yg Diprediksi", also with a code snippet:

```
<StringArray> ['Negatif', 'Negatif', 'Negatif', 'Positif', 'Negatif', 'Negatif', 'Negatif'] Length: 8, dtype: string
```

Below these, "Nilai Probabilitas" is listed, followed by the text "nilai akurasinya adalah 87.5". In the top right corner, there is a message: "Activate Wind" and "Go to Settings to a".

Gambar 5. 3 Analisis

5.1.3 Evaluasi

Metode penelitian dimulai dari pengumpulan data hingga melakukan perhitungan nilai probabilitas menggunakan *Naïve Bayes*. Peneliti berhasil mencapai nilai akurasi sebesar 87% dari 73 dataset. Data tersebut dibagi menjadi subset, di mana 10% digunakan sebagai data uji (*testing*) dan 90% sebagai data latih (*training*). Pada pembagian tersebut, diperoleh nilai akurasi sebesar 87%. Namun, ketika pembagian dilakukan dengan perbandingan 20% data uji dan 80% data latih, nilai akurasi menurun menjadi 80%. Begitu juga pada pembagian 30% data uji dan 70% data latih, nilai akurasi turun lagi menjadi 77%. Visualisasi pembagian data terbaik dapat dilihat pada framework streamlit dengan nilai akurasi tertinggi.

Mulai Evaluasi

3. Evaluasi

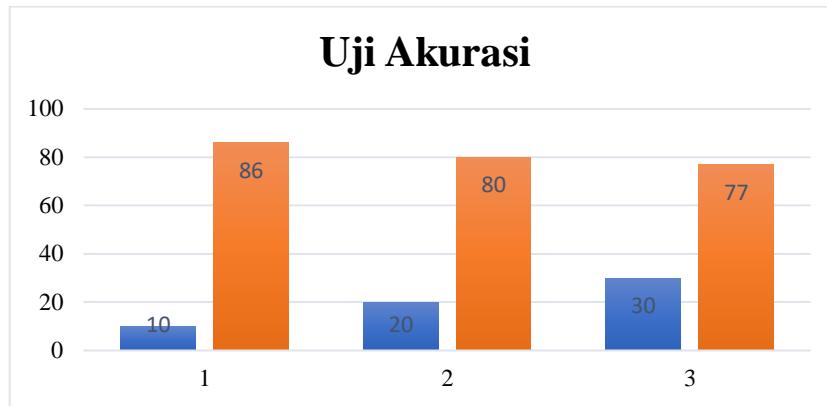
0	1
6	0
1	1

precision recall f1-score support

Negatif	0.86	1.00	0.92	6
Positif	1.00	0.50	0.67	2
accuracy			0.88	8

macro avg 0.93 0.75 0.79 8 weighted avg 0.89 0.88 0.86 8

Gambar 5. 4 Evaluasi



Gambar 4. 6 Uji Akurasi

Ket : Warna biru adalah nilai persen pembagian untuk mencari nilai Akurasi dan yang warna orens adalah hasil nilai akurasi.

5.1.4 Implementasi

Pada tahap implementasi model, dilakukan proses uji coba terhadap model yang telah dibangun. Implementasi model berfungsi untuk menguji komentar-komentar baru agar dapat dianalisis sebagai positif atau negatif.

Berikut merupakan visualisasi hasil dari implementasi model pada *streamlit* dengan menggunakan komentar pertama.

Komentar Masyarakat Pemodelan **Sentimen Analisis** Grafik

Komentar Masyarakat

Menurut pendapat saya, objek wisata yang berada di kabupaten banggai laut sangatlah banyak, dan sangat bagus serta sangat indah.

Analisis

[Positif] Menurut pendapat saya, objek wisata yang berada di kabupaten banggai laut sangatlah banyak, dan sangat bagus serta sangat indah.

Gambar 5. 5 Hasil Implementasi 1

Dari hasil pengujian pada implementasi model, terlihat bahwa komentar tersebut diklasifikasikan sebagai sentimen “Positif”.

Berikut adalah hasil visualisasi dari implementasi model pada streamlit dengan menggunakan komentar kedua.



Komentar Masyarakat Pemodelan **Sentimen Analisis** Grafik

Komentar Masyarakat

objek wisata yang ada di banggai laut belum banyak orang yang mengetahuinya belum lagi akses yang lumayan susah untuk sampai di tempat-tempat objek wisata tersebut

Analisis

[Negatif] objek wisata yang ada di banggai laut belum banyak orang yang mengetahuinya belum lagi akses yang lumayan susah untuk sampai di tempat-tempat objek wisata tersebut

Gambar 5. 6 Hasil Implementasi 2

Dari hasil pengujian pada implementasi model, terlihat bahwa komentar tersebut diklasifikasikan sebagai sentimen “Negatif”

BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan analisis sentiment yang telah dilakukan, dapat disimpulkan bahwa metode atau algoritma *Naïve Bayes* mampu menganalisis komentar dengan hasil yang baik terhadap objek wisata Kabupaten Banggai Laut. Algoritma ini efektif dalam mengidentifikasi kelas sentiment positif dan negatif terhadap komentar tersebut. Penggunaan algoritma ini juga menunjukkan tingkat akurasi yang tinggi, mencapai 87% dari total 73 dataset dengan pembagian 90% data latih dan 10% data uji. Dengan tingkat akurasi tersebut, dapat disimpulkan bahwa algoritma yang diusulkan memiliki tingkat diagnostik yang baik.

6.2 Saran

Pada penelitian selanjutnya, perlu dipertimbangkan untuk menganalisis komentar-komentar di media sosial lainnya. Selain itu, peneliti dapat menambahkan metode atau algoritma tambahan sebagai pembanding. Hal ini dikarenakan dalam penelitian ini, proses analisis yang dilakukan oleh algoritma hanya dilakukan term demi term atau berdasarkan banyaknya kata saja. Oleh karena itu, untuk penelitian selanjutnya, perlu dipertimbangkan penggunaan metode yang mampu melakukan pembelajaran lebih dari satu kata atau per kalimat. Selanjutnya, penelitian juga dapat dikembangkan dengan membangun sistem serupa.

DAFTAR PUSTAKA

- [1] S. Nurhayati and V. G. Ristanto, “SISTEM INFORMASI PARIWISATA PROVINSI PAPUA BERBASIS WEB,” 2017. [Online]. Available: www.merdeka.com,
- [2] F. V. Sari and A. Wibowo, “ANALISIS SENTIMEN PELANGGAN TOKO ONLINE JD.ID MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER BERBASIS KONVERSI IKON EMOSI,” *Jurnal SIMETRIS*, vol. 10, no. 2, 2019.
- [3] R. Ferryawan and F. Wahyu Wibowo, “Analisis Sentimen Wisata Jawa Tengah Menggunakan Naïve Bayes,” 2019.
- [4] A. Sentimen *et al.*, “Analisis Sentimen Objek Wisata Bali Di Google Maps Menggunakan Algoritma Naive Bayes,” 2022.
- [5] K. Keahlian Rekayasa *et al.*, “Terbit online pada laman web jurnal: <http://journal.ittelkom-pwt.ac.id/index.php/dinda> Journal of Dinda Sentiment Analysis Destinasi Wisata Kabupaten Bekasi Berdasarkan Opini Masyarakat Menggunakan Naive bayes,” *Data Institut Teknologi Telkom Purwokerto*, vol. 2, no. 2, pp. 64–74, 2022, [Online]. Available: <http://journal.ittelkom-pwt.ac.id/index.php/dinda>
- [6] K. Diah Indarwati and H. Februariyanti, “ANALISIS SENTIMEN TERHADAP KUALITAS PELAYANAN APLIKASI GO-JEK MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER.”
- [7] W. Khofifah, D. N. Rahayu, and A. M. Yusuf, “Analisis Sentimen Menggunakan Naive Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps,” *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol. 16, no. 4, pp. 28–38, Jan. 2022, doi: 10.35969/interkom.v16i4.192.
- [8] Y. A. Singgalen, “Analisis Sentimen dan Pemodelan Topik dalam Optimalisasi Pemasaran Destinasi Pariwisata Prioritas di Indonesia,” *Journal of Information*

Systems and Informatics, vol. 4, no. 1, [Online]. Available: <http://journal-isi.org/index.php/isi>

- [9] N. Aula, M. Ula, and L. Rosnita, “ANALISIS SENTIMEN REVIEW CUSTOMER TERHADAP PERUSAHAAN EKSPEDISI JNE, J&T EXPRESS DAN POS INDONESIA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM) ANALYSIS OF CUSTOMER REVIEW SENTIMENT TO JNE, J&T EXPRESS AND POS INDONESIA EXPEDITION COMPANIES USING SVM METHOD,” *Journal of Informatics and Computer Science*, vol. 9, no. 1, 2023.
- [10] D. Auliya Agustina, S. Subanti, E. Zukhronah, P. Studi Statistika, and U. Sebelas Maret, “Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Marketplace di Indonesia Menggunakan Algoritma Support Vector Machine,” 2020.
- [11] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, “Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19),” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 2, p. 406, Apr. 2021, doi: 10.30865/mib.v5i2.2835.
- [12] L. Aji Andika and P. Amalia Nur Azizah, “Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier,” 2019.

LAMPIRAN

```

from cgitb import text
from turtle import pd
import streamlit as st
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.metrics import confusion_matrix
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
from Sastrawi.StopWordRemover.StopWordRemoverFactory import
StopWordRemoverFactory
from nltk.tokenize import RegexpTokenizer
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
vectorizer = TfidfVectorizer(max_features=2500)
st.header('Analisi Sentimen Objek Wisata Di Kabupaten Banggai Laut',
divider='blue')
st.image("objek.png", width=700)
st.sidebar.image("logo.png", width=300)
st.sidebar.header('Sejarah Banggai Laut', divider='blue')
st.sidebar.text_area(
    "Kabupaten Banggai Laut adalah salah satu kabupaten yang berada di provinsi
    Sulawesi Tengah."
    "Perisai Bersegi Lima melambangkan kekokohan perdamaian persatuan dan
    kesatuan Banggai"
    "Laut bagian dari Negara Kesatuan Republik Indonesia berdasarkan Pancasila;"
    "Rumah Berpilar Emas, merupakan sebuah latar belakang sejarah kejayaan
    serta keemasan"
    "kerajaan Banggai yang selalu memiliki semangat membangun; Dua Tiang
    berpilar Empat"
)

```

"menggambarkan lembaga adat kerajaan banggai diperkuat 2 lembaga demokrasi yaitu basalo"

"sangkap dan komisi sangkap; Susunan Tangga Rumah, artinya tahapan menuju kemandirian"

"didalam membangun Kabupaten Banggai Laut; Bintang Kuning Emas melambangkan keagungan Ketuhanan Yang Maha Esa;"

"Pohon Kelapa, merupakan hasil potensi daratan atau kesuburan tanah;"

"Padi dan Kapas, melambangkan kemakmuran dan kesejahteraan sandang dan pangangan;"

"Laut Biru melambangkan wilayah kabupaten banggai laut sebagai daerah maritime;"

"Ikan Cardinal, merupakan endemik biota laut asli dari Kabupaten Banggai Laut;"

"Pita Merah dan Putih merupakan ciri khas bendera budaya Banggai yang bermakna ketegaran,"

"keperkasaan, kemegahan, serta kesucian didalam mengembang, menegakkan dan menyatukan"

"pendapat dan kemajemukan masyarakat banggai;"

"Untaian Mutiara, artinya keakraban dan keramahtamahan serta keindahan Banggai Laut;"

"Lahirnya Kabupaten Banggai Laut tertuang pada gambar dalam untaian 11 mutiara, 1 rumah"

"berpilar emas, 20 bunga padi, dan 13 bunga kapas, kesemuanya tersimpul angka 11 Januari 20"

)

Using object notation

```
add_selectbox = st.sidebar.selectbox(
    "How would you like to be contacted?",
    ("Email", "Home phone", "Mobile phone")
)
```

--- LOAD DATAFRAME

```

excel_file = 'Komentar.xlsx'
sheet_name = 'Sheet1'
df = pd.read_excel(excel_file,
                   sheet_name=sheet_name,
                   header=0)
jumlahdata=df.No.count()
dataset = {}
No = []
Komentar = []
Sentimen = []
# Using "with" notation
with st.sidebar:
    add_radio = st.radio(
        "Choose a shipping method",
        ("Standard (5-15 days)", "Express (2-5 days)")
    )
tab1, tab2, tab3 = st.tabs(["Komentar Masyarakat", "Pemodelan", "Sentimen Analisis"])
with tab1:
    st.header("Komentar Masyarakat")
    for i in range(0,jumlahdata):
        dataset[df.No[i]]=df.Komentar[i]
        if (df.Sentimen[i] == 'Negatif'):
            st.error(df.Komentar[i], icon=':(')
        if (df.Sentimen[i] == 'Positif'):
            st.info(df.Komentar[i], icon="(:")
        Sentimen.append(df.Sentimen[i])
        Komentar.append(df.Komentar[i])
with tab2:
    #st.write(dataset)
    dnb = pd.DataFrame (columns=['text','label'])

```

```

dnb ['Sentimen'] = Sentimen
dnb ['Komentar'] = Komentar
st.header("Pemodelan")
if st.button('Preprocessing'):
    st.success("1. Preprocessing")
# Case Folding
for k in dataset.keys():
    hasil_case_folding = dataset[k].lower()
    dataset[k] = hasil_case_folding
    st.write("Case Folding:", dataset)
# Tokenizing
tokenizer = RegexpTokenizer(r"\w+")
for k in dataset.keys():
    tokens = tokenizer.tokenize(dataset[k])
    dataset[k] = tokens
    st.write("Tokenizing", dataset)
# Number Removal
for k in dataset.keys():
    tokens = []
    for t in dataset[k]:
        if t.isnumeric() == False:
            tokens.append(t)
    dataset[k] = tokens
# Stopword Removal by Sastrawi
factory = StopWordRemoverFactory()
stopword_list = factory.get_stop_words()
for k in dataset.keys():
    tokens = []
    for t in dataset[k]:
        if t not in stopword_list:
            tokens.append(t)

```

```

dataset[k] = tokens
st.write("Stopword Removal", dataset)

# Stemming by Sastrawi
factory = StemmerFactory()
stemmer = factory.create_stemmer()
for k in dataset.keys():
    tokens = []
    for t in dataset[k]:
        tokens.append(stemmer.stem(t))
    dataset[k] = tokens
st.write("Steamming",dataset)

# Mengembalikan Format Dataset Awal
#for k in dataset.keys():
#    dataset[k] = " ".join(dataset[k])
# st.write("Steamming",dataset)

if st.button('Pembobotan'):
    st.success("2. Pembobotan Dokumen (TF-IDF)")

    #Frekuensi Kemunculan Kata
    tf = CountVectorizer()
    term_doc_matrix = tf.fit_transform(dataset.values())
    TAMPIL=pd.DataFrame(term_doc_matrix.toarray(), index=dataset.keys(),
    columns=tf.get_feature_names_out())
    st.write(TAMPIL)

    # Pembobotan TF-IDF
    vectorizer = TfidfVectorizer(max_features=2500)
    model_g = GaussianNB()
    v_data = vectorizer.fit_transform(dataset.values()).toarray()
    st.write(v_data)
    st.success("Split Data")
    t_size = st.radio(
        "Pilih pembagian splitting data",

```

```

(0.1,0.2,0.3))

st.write("size splitting :", t_size)

if st.button('Mulai Analisis'):

    st.success("3. Naive Bayes")

    st.write("Kernel GaussianNB")

    #vectorizer = TfidfVectorizer(max_features=2500)

    from sklearn.naive_bayes import GaussianNB

    model_g = GaussianNB()

    from sklearn.model_selection import train_test_split

    v_data = vectorizer.fit_transform(dataset.values()).toarray()

    X_train, X_test, y_train, y_test = train_test_split(v_data,
                                                       dnb['Sentimen'], test_size=t_size, random_state=0)

    nbmodel = model_g.fit(X_train, y_train)

    #nbmodel.class_count_

    from sklearn.metrics import accuracy_score

    y_preds = model_g.predict(X_test)

    predict = pd.array(y_preds)

    true_Sentimen = pd.array(y_test)

    st.write("Data Testing Sebenarnya")

    st.write(true_Sentimen)

    st.write("Data Testing Yg Diprediksi")

    st.write(predict)

    proba = nbmodel.predict_proba(X_test)

    st.write("Nilai Probabilitas")

    print(proba)

    st.write('nilai akurasinya adalah ',

            accuracy_score(y_test, y_preds)*100)

if st.button('Mulai Evaluasi'):

    st.success("3. Evaluasi")

    #vectorizer = TfidfVectorizer(max_features=2500)

    model_g = GaussianNB()

```

```

v_data = vectorizer.fit_transform(dataset.values()).toarray()
X_train, X_test, y_train, y_test = train_test_split(v_data,
dnb['Sentimen'], test_size=t_size, random_state=0)
model_g.fit(X_train, y_train)
y_preds = model_g.predict(X_test)
from sklearn.metrics import classification_report,confusion_matrix
st.write(confusion_matrix(y_test, y_preds))
st.write(classification_report(y_test, y_preds))

```

with tab3:

```

#st.write(dataset)
dnb = pd.DataFrame (columns=['text','label'])
dnb ['Sentimen'] = Sentimen
dnb ['Komentar'] = Komentar
txt = st.text_area("Komentar Masyarakat","")
from openpyxl import load_workbook
# Existing Excel file
existing_file = 'hasil.xlsx'
sheet_name2 = 'Sheet1'
df_hasil =pd.read_excel(existing_file,
sheet_name=sheet_name2,
header=0)
jumlahhasil=df_hasil.No.count()
datahasil = {}
No_hasil = []
Komentar_hasil = []
Sentimen_hasil = []
no_baru=jumlahhasil+1
if st.button('Analisis'):
    dataset["q"]=txt
    # Case Folding
    for k in dataset.keys():

```

```

    hasil_case_folding = dataset[k].lower()
    dataset[k] = hasil_case_folding
    # st.write("Case Folding:",dataset)

    # Tokenizing
    tokenizer = RegexpTokenizer(r"\w+")
    for k in dataset.keys():
        tokens = tokenizer.tokenize(dataset[k])
        dataset[k] = tokens
    # st.write("Tokenisasi",dataset)

    # Number Removal
    for k in dataset.keys():
        tokens = []
        for t in dataset[k]:
            if t.isnumeric() == False:
                tokens.append(t)
        dataset[k] = tokens

    # Stopword Removal by Sastrawi
    factory = StopWordRemoverFactory()
    stopword_list = factory.get_stop_words()
    for k in dataset.keys():
        tokens = []
        for t in dataset[k]:
            if t not in stopword_list:
                tokens.append(t)
        dataset[k] = tokens
    # st.write("Stopword Removing",dataset)

    # Stemming by Sastrawi
    factory = StemmerFactory()
    stemmer = factory.create_stemmer()
    for k in dataset.keys():
        tokens = []

```

```

for t in dataset[k]:
    tokens.append(stemmer.stem(t))
dataset[k] = tokens
# Mengembalikan Format Dataset Awal
# st.write("Stemming",dataset)
for k in dataset.keys():
    dataset[k] = " ".join(dataset[k])
# Frekuensi Kemunculan Kata
tf = CountVectorizer()
term_doc_matrix = tf.fit_transform(dataset.values())
pd.DataFrame(term_doc_matrix.toarray(), index=dataset.keys(),
columns=tf.get_feature_names_out())
# Pembobotan TF.IDF
vectorizer = TfidfVectorizer(max_features=2500)
model_g = GaussianNB()
v_data = vectorizer.fit_transform(dataset.values()).toarray()
# st.write(v_data)
X_train, X_test, y_train, y_test = train_test_split(v_data, dnb['Sentimen'],
test_size=0.3, random_state=0)
model_g.fit(X_train, y_train)
y_preds = model_g.predict(X_test)
tweet = txt
v_data = vectorizer.transform([tweet]).toarray()
y_preds = model_g.predict(v_data)
#st.write(y_preds)
# dengan asumsi bahwa 1 merupakan label positif
if y_preds == "Negatif":
    st.error('[Negatif] '+txt, icon=':( ')
    hasil="Negatif"
else:
    st.success('Positif] '+txt, icon='(:')

```

```
hasil="Positif"

# New data to append
new_data = [[no_baru, tweet, hasil]]

# Load existing workbook
wb = load_workbook(existing_file)

# Select the active sheet
ws = wb.active

# Append new data
for row in new_data:
    ws.append(row)

# Save the workbook
wb.save(existing_file)

for i in range(0,jumlahhasil):
    datahasil[df_hasil.No[i]]=df_hasil.Komentar[i]
    if (df_hasil.Sentimen[i] == 'Negatif '):
        st.error(df_hasil.Komentar[i], icon=':(')
    elif (df_hasil.Sentimen[i] == 'Positif'):
        st.info(df_hasil.Komentar[i], icon="(:")
    else:
        st.info(df_hasil.Komentar[i], icon="(:")

    #Sentimen_hasil.append(df.Senstimen[i])
    #Sentimen_hasil.append(df.Sentimen[i])
```



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

SURAT KETERANGAN BEBAS PUSTAKA

No : 003/Perpustakaan-Fikom/III/2024

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

Nama Anggota	:	Fatmawati
No. Induk	:	T3120013
No. Anggota	:	M20244

Terhitung mulai hari, tanggal : Rabu, 27 Maret 2024, dinyatakan telah bebas pinjam buku dan koleksi perpustakaan lainnya.

Demikian keterangan ini di buat untuk di pergunakan sebagaimana mestinya.



Gorontalo, 27 Maret 2024

**Mengetahui,
Kepala Perpustakaan**



Apriyanto Alhamad, M.Kom
NIDN : 0924048601



PEMERINTAH KABUPATEN BANGGAI LAUT
DINAS PARIWISATA DAN KEBUDAYAAN
Jln. BRAWIJAYA No. 01 Telp. (0462)....., Kp. 94891
BANGGAI

SURAT KETERANGAN TELAH MELAKUKAN PENELITIAN

Nomor : 556/ 82 /DISPARBUD/III/2024

Yang bertanda tangan di bawah ini :

Nama	:	LAODE KAIMUDIN, SE
N I P	:	19711020 200012 1 004
Pangkat /Gol. Ruang	:	Pembina Utama Muda, IV/c
Jabatan	:	Kepala Dinas Pariwisata dan Kebudayaan
Instansi	:	Dinas Pariwisata Dan kebudayaan Kab. Banggai Laut

Dengan ini menerangkan bahwa :

Nama	:	Fatmawati
Tempat/Tgl Lahir	:	Lalong, 03/03/2003
Jenis Kelamin	:	Perempuan
Pekerjaan	:	Mahasiswa
NIM	:	T3120013
Fakultas	:	Ilmu Komputer
Program Studi	:	Teknik Informatika
Asal Perguruan Tinggi : Universitas Ichsan Gorontalo		

Adalah benar telah melakukan Penelitian tentang **Analisis Sentimen Objek Wisata di Kabupaten Banggai Laut Menggunakan Metode Naive Bayes** di Dinas Pariwisata dan Kebudayaan Kabupaten Banggai Laut pada Tanggal 25 s/d 26 Maret 2024.

Demikian surat Keterangan ini dibuat untuk dapat digunakan dengan perlunya.

Banggai, 27 Maret 2024





**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
 UNIVERSITAS ICHSAN GORONTALO
 FAKULTAS ILMU KOMPUTER
 SK MENDIKNAS NOMOR 84/D/O/2001
 JL. Achmad Nadjamuddin No. 17 Telp. (0435) 829975 Fax (0435) 829976 Gorontalo**

Berita Acara Perbaikan/Revisi Ujian PROPOSAL

Pada hari ini, kamis 26 Oktober 2023, Pukul 14.00-15.00 Wita. Telah dilaksanakan Ujian PROPOSAL mahasiswa Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.

Nama : FATMAWATI
 Nim : T3120013
 Pembimbing I : Husdi, M.Kom
 Pembimbing II : Kartika Chandra Pelangi, M.Kom
 Judul PROPOSAL : Analisis Sentimen Objek Wisata Dikabupaten Banggai Laut Menggunakan Metode Naive Bayes

Oleh Komite Seminar sebagai berikut :

No	Komite Seminar	Status	Tanda Tangan
1	Rezqiwati Ishak, M.Kom	Ketua	
2	Muis Nanja, M.Kom	Anggota	
3	Yulianty Lasena, M.Kom	Anggota	
4	Husdi, M.Kom	Anggota	
5	Kartika Chandra Pelangi, M.Kom	Anggota	



**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS IHSAN GORONTALO
FAKULTAS ILMU KOMPUTER
SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001
Jl. Achmad Najamuddin No. 17 Telp. (0435) 829975 Fax (0435) 829976 Gorontalo**

SURAT REKOMENDASI BEBAS PLAGIASI
No. 74/FIKOM-UIG/R/III/2024

Yang bertanda tangan di bawah ini :

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

Dengan ini menerangkan bahwa :

Nama Mahasiswa	:	Falmawati
NIM	:	T3120013
Program Studi	:	Teknik Informatika (S1)
Fakultas	:	Fakultas Ilmu Komputer
Judul Skripsi	:	Analisis Sentimen Objek Wisata di Kabupaten Banggai Laut Menggunakan Metode Naive Bayes

Sesuai hasil pengecekan tingkat kemiripan skripsi melalui aplikasi **Turnitin** untuk judul skripsi di atas diperoleh hasil **Similarity** sebesar **15%**, berdasarkan Peraturan Rektor No. 32 Tahun 2019 tentang Pendekripsi Plagiat pada Setiap Karya Ilmiah di Lingkungan Universitas Ichsan Gorontalo dan persyaratan pemberian surat rekomendasi verifikasi calon wisudawan dari LLDIKTI Wil. XVI, bahwa batas kemiripan skripsi maksimal 30%, untuk itu skripsi tersebut di atas dinyatakan **BEBAS PLAGIASI** dan layak untuk diujangkan.

Demikian surat rekomendasi ini dibuat untuk digunakan sebagaimana mestinya.

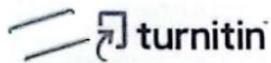


Irvan Abraham Salihi, M.Kom
NIDN. 0928028101

Gorontalo, 30 Maret 2024
Tim Verifikasi,

Zulfrianto Y. Lamasigi, M.Kom
NIDN. 0914089101

Terlampir :
Hasil Pengecekan Turnitin



PAPER NAME	AUTHOR
SKRIPSI_T3120013_FATMAWATI.pdf	FATMA WATI
WORD COUNT	CHARACTER COUNT
9641 Words	47055 Characters
PAGE COUNT	FILE SIZE
71 Pages	2.1MB
SUBMISSION DATE	REPORT DATE
Mar 28, 2024 3:23 PM GMT+8	Mar 28, 2024 3:23 PM GMT+8

● 15% Overall Similarity

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

- 14% Internet database
 - Crossref database
 - 0% Submitted Works database
 - 4% Publications database
 - Crossref Posted Content database

- Excluded from Similarity Report

- Bibliographic material
 - Cited material
 - Quoted material
 - Small Matches (Less than 10 words)

Summary

DAFTAR RIWAYAT HIDUP



Nama	: Fatmawati
Nim	: T3120013
Tempat,Tanggal Lahir	: Lalong, 03 Maret 2003
Alamat	: Jln. Kenangan
Agama	: Islam
Kewarganegaraan	: WNI
Email	: fatmawatialloe@gmail.com

Riwayat pendidikan dan pekerjaan :

1. Tahun 2014, menyelesaikan pendidikan di Sekolah Dasar Negeri Potil Pololoba.
2. Tahun 2017, menyelesaikan pendidikan di Sekolah Menengah Pertama Negeri 2 Banggai.
3. Tahun 2020, menyelesaikan pendidikan di Sekolah Menengah Kejuruan Negeri 1 Banggai.
4. Tahun 2020, diterima menjadi mahasiswa di Perguruan Tinggi Swasta Universitas Ichsan Gorontalo.