

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP LAYANAN TRANSPORTASI BISKITA TRANSPAKUAN BOGOR MENGGUNAKAN METODE *ENSEMBLE*

NAUFAL NASHIF IMANUDDIN



**DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Analisis Sentimen Pengguna Twitter Terhadap Layanan Transportasi Biskita Transpakuan Bogor Menggunakan Metode *Ensemble*” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Januari 2024

Naufal Nashif Imanuddin
G14190025



ABSTRAK

NAUFAL NASHIF IMANUDDIN. Analisis Sentimen Pengguna Twitter Terhadap Layanan Transportasi Biskita Transpakuan Bogor Menggunakan Metode *Ensemble*. Dibimbing oleh YENNI ANGRAINI dan AKBAR RIZKI.

Bogor merupakan kota terpadat kelima di Indonesia pada tahun 2021. Tingginya jumlah angkutan umum dan kendaraan pribadi membuat pemerintah Kota Bogor berupaya meluncurkan angkutan umum yang lebih ramah pengguna dengan nama Biskita Transpakuan. Umpan balik dari masyarakat diperlukan untuk terus meningkatkan kualitas pelayanan dalam memenuhi kebutuhan pengguna. Umpan balik tersebut dapat berupa dukungan, kritik, dan saran. Analisis sentimen digunakan untuk mengekstrak informasi berharga salah satunya dari media sosial. Metode *Ensemble* dengan model dasar regresi logistik multinomial, random forest dan multinomial naïve bayes digunakan untuk melakukan klasifikasi. Tujuan penelitian ini untuk mengetahui performa dari model *Ensemble* dalam melakukan klasifikasi sentimen. Hasil penelitian menunjukkan bahwa model *Ensemble* dapat memanfaatkan kelebihan masing – masing model dasar dan mendapatkan Akurasi serta skor F1 tertinggi yaitu sebesar 78%.

Kata kunci: Analisis Sentimen, Biskita, Kemacetan, Metode *Ensemble*, Twitter

ABSTRACT

NAUFAL NASHIF IMANUDDIN. Sentiment Analysis of Twitter Users on Biskita Transpakuan Bogor Transportation Services Using the Ensemble Method. Supervised by YENNI ANGRAINI and AKBAR RIZKI.

Bogor is the fifth most populous city in Indonesia in 2021. The high number of public transportation and private vehicles has led the Bogor City government to launch a more user-friendly public transportation named Biskita Transpakuan. Feedback from the community is certainly needed to continue to improve service quality in meeting user needs. The feedback can be in the form of support, criticism, and suggestions. Sentiment analysis is used to extract valuable information from social media. Ensemble method with basic models of multinomial logistic regression, random forest and multinomial naïve bayes are used to perform classification. The purpose of this study is to determine the performance of the Ensemble model in performing sentiment classification. The results showed that the Ensemble model can utilize the advantages of each basic model and get the highest Akurasi and F1 score of 78%.

Keywords: Biskita, Congestion, Ensemble Method, Sentiment Analysis, Twitter



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2024
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP LAYANAN TRANSPORTASI BISKITA TRANSPAKUAN BOGOR MENGGUNAKAN METODE *ENSEMBLE*

NAUFAL NASHIF IMANUDDIN

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:
Rahma Anisa, S.Stat., M.Si.

Perpustakaan IPB University



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Judul Skripsi : Analisis Sentimen Pengguna Twitter Terhadap Layanan
Transportasi Biskita Transpakuan Bogor Menggunakan Metode
Ensemble

Nama : Naufal Nashif Imanuddin
NIM : G14190025

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:

Dr. Yenni Angraini, S.Si., M.Si.

Pembimbing 2:

Akbar Rizki, S.Stat., M.Si.

Diketahui oleh

Ketua Program Studi:

Dr. Bagus Sartono, S.Si., M.Si.
NIP. 19780411 200501 1002

Tanggal Ujian:
29 Desember 2023

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Februari 2023 sampai bulan Januari 2024 ini ialah analisis sentimen pengguna twitter, dengan judul “Analisis Sentimen Pengguna Twitter Terhadap Layanan Transportasi Biskita Transpakuan Bogor Menggunakan Metode *Ensemble*”.

Pada kesempatan ini, penulis ingin mengucapkan terima kasih kepada pihak-pihak yang telah berperan sejak penulis menempuh studi sampai proses penulisan karya ilmiah ini berakhir, diantaranya:

1. Bapak Purwono dan Ibu Amanatul Kirom selaku orang tua dan seluruh keluarga yang telah memberi doa dan dukungan
2. Dr. Yenni Angraini, S.Si., M.Si. dan Akbar Rizki, S.tat., M.Si., selaku Dosen pembimbing Skripsi yang telah memberikan arahan dan masukan dalam penulisan karya ilmiah ini.
3. Civitas akademika Departemen Statistika IPB yang telah memberikan ilmu dan telah membantu selama proses studi hingga menyelesaikan studi.
4. Seluruh teman-teman yang telah membantu yang tidak dapat penulis sebutkan satu per satu.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan. Penulis menyadari bahwa karya ilmiah ini masih jauh dari kesempurnaan, oleh karena itu penulis menyambut baik bila ada *feedback* berupa masukan yang membangun. Masukan seperti itu akan sangat berharga bagi penulis.

Bogor, Januari 2024

Naufal Nashif Imanuddin



DAFTAR ISI

DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiv
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	2
1.3 Manfaat	2
TINJAUAN PUSTAKA	3
2.1 Analisis Sentimen	3
2.2 <i>Term Frequency - Inverse Document Frequency (TF – IDF)</i>	3
2.3 <i>Synthetic Minority Over-sampling Technique (SMOTE)</i>	3
2.4 <i>Hyperparameter Tuning</i>	4
2.5 Multinomial Naïve Bayes	5
2.6 Regresi Logistik Multinomial	5
2.7 Random Forest	6
2.8 <i>Ensemble</i>	8
III METODE	10
3.1 Tahapan Penelitian	10
3.2 Data	10
3.3 Preproses Data	11
3.4 Eksplorasi Data	12
3.5 Pembuatan Model	13
3.6 Evaluasi Model	16
3.7 <i>Model Deployment</i>	17
IV HASIL DAN PEMBAHASAN	18
4.1 Preproses Data	18
4.2 Eksplorasi Data	19
4.3 Pembuatan Model	21
4.4 Evaluasi Model	23
4.5 <i>Model Deployment</i>	28
V SIMPULAN DAN SARAN	30
5.1 Simpulan	30
5.2 Saran	30
DAFTAR PUSTAKA	31
LAMPIRAN	34
RIWAYAT HIDUP	38

DAFTAR TABEL

1	Contoh Data Tweet	10
2	Aturan Pelabelan Tweet	11
3	Ilustrasi Hasil Ekstraksi Fitur	13
4	<i>Hyperparameter tuning Base Classifiers</i>	14
5	Karakteristik <i>Base Classifiers</i>	15
6	Ilustrasi <i>Weight Soft Voting</i>	16
7	Ilustrasi Matrik Konfusi Tiga Kelas	16
8	Contoh <i>Tweet Duplikat</i>	18
9	Hasil Pelabelan Data	18
10	Contoh Hasil Preproses Data	19
11	Hasil Tuning Parameter <i>Base Classifiers</i>	22
12	Akurasi <i>Base Classifiers</i> Pada Setiap <i>Fold</i>	23
13	Perbandingan Nilai <i>True Positif</i> Setiap Model	27
14	Hasil Evaluasi Model <i>Ensemble</i>	27

DAFTAR GAMBAR

1	Ilustrasi <i>Hard Voting</i> (da Silva <i>et al.</i> 2014).	9
2	Ilustrasi <i>Soft Voting</i> (da Silva <i>et al.</i> 2014).	9
3	Diagram Alir Penelitian	10
4	Ilustrasi <i>K-Fold Cross Validation</i> (Ren <i>et al.</i> 2019)	14
5	Diagram Alir Pembuatan Model <i>Ensemble</i>	15
6	Jumlah dan Presentase Sentimen	19
7	Lini Masa Sentimen	20
8	<i>Wordclouds</i> Sentimen (a) Negatif (b) Netral dan (c) Positif	20
9	10 Fitur Dengan Skor TF-IDF Tertinggi	21
10	<i>Oversampling</i> Menggunakan SMOTE	22
11	Matrik Konfusi Model Regresi Logistik Multinomial	23
12	Performa Model Regresi Logistik Multinomial pada setiap <i>Fold</i>	24
13	Matrik Konfusi Model Random Forest	24
14	Performa Model Random Forest pada setiap <i>Fold</i>	25
15	Matrik Konfusi Model Multinomial Naïve Bayes	25
16	Performa Model Multinomial Naive Bayes pada setiap <i>Fold</i>	26
17	Matrik Konfusi Model <i>Ensemble</i>	26
18	Performa Model <i>Ensemble</i> pada setiap <i>Fold</i>	27
19	Perbandingan Performa Akurasi <i>Classifiers</i>	28
20	Perbandingan Performa Skor F1 <i>Classifiers</i>	28



DAFTAR LAMPIRAN

1	Lampiran 1 Kamus Normalisasi Kata Gaul Tambahan Buatan Sendiri	35
2	Lampiran 2 <i>Screenshot</i> Tampilan Hasil <i>Deploying</i> Model	36
3	Lampiran 3 Hasil Prediksi Tweets Model <i>Ensemble</i>	37

I PENDAHULUAN

1.1 Latar Belakang

Kemacetan adalah salah satu permasalahan fundamental yang sering dihadapi oleh negara berkembang. Menurut Kamus Besar Bahasa Indonesia (KBBI) kemacetan merupakan keadaan tersendatnya atau bahkan terhentinya lalu lintas. Kemacetan sering terjadi di kota-kota besar di Indonesia seperti Jakarta, Surabaya, Bandung, tak terkecuali Kota Bogor. Analisis *Global Traffic Scorecard 2021* yang dilakukan INRIX menempatkan Kota Bogor dalam peringkat kelima kota termacet di Indonesia dan peringkat ke-821 kota termacet di dunia. Kemacetan mengakibatkan pengguna jalan merasakan stres, mengurangi jam kerja atau belajar, pemborosan bensin dan hilangnya pendapatan. Menurut (Soedomo *et al.* 2010) penyebab kemacetan lalu lintas di kota Bogor tidak hanya karena banyaknya angkutan kota, tapi juga karena banyaknya kendaraan pribadi yang digunakan oleh masyarakat. Oleh karena itu, pemerintah kota Bogor berupaya mengurangi jumlah angkot yang telah usang dengan cara mereduksi menjadi alat transportasi yang lebih ramah yaitu Biskita.

Biskita merupakan salah satu layanan angkutan umum di Kota Bogor yang memiliki pelayanan menyerupai Transjakarta namun, tidak menggunakan lajur khusus. Biskita memiliki kelebihan dibandingkan angkutan umum lainnya, seperti jarak rute, adanya pendingin udara (AC), CCTV, *passenger counting system*, *disability friendly* dan standar pelayanan minimum (SPM). Hal ini diharapkan mampu untuk menarik minat masyarakat beralih menggunakan angkutan umum. Oleh karena itu, persepsi masyarakat mengenai layanan Biskita merupakan hal yang penting untuk diperhatikan. Hal ini dikarenakan persepsi masyarakat terhadap pelayanan angkutan umum dapat mempengaruhi kinerja angkutan dan pengambilan kebijakan kedepannya. Angkutan umum yang berevolusi sesuai kebutuhan pengguna, dapat meningkatkan keinginan pengguna untuk menggunakan Biskita.

Persepsi masyarakat mengenai suatu topik dapat diamati pada platform Twitter. Twitter merupakan sebuah platform media sosial dan layanan *microblogging* untuk mengirim dan membaca pesan berbasis teks secara *real time*. Menurut penelitian dari (Loilatu *et al.* 2021) media sosial twitter memiliki peranan penting dalam media komunikasi transportasi publik. Disisi lain, berdasarkan laporan *We Are Social* dan *Hootsuite*, ada 24 juta pengguna twitter di Indonesia per Januari 2023. Hal ini berpotensi besar menjadikan twitter sebagai sumber data yang efisien untuk melihat persepsi masyarakat dengan cara melakukan analisis sentimen.

Analisis sentimen atau *opinion mining* merupakan suatu studi komputasi untuk memahami, mengekstrak, dan melakukan pengolahan data berbentuk teks untuk mendapatkan informasi sentimen dalam suatu kalimat terhadap suatu topik yang dapat bersifat positif, negatif, dan netral. Analisis sentimen melalui twitter umumnya banyak menggunakan metode klasifikasi seperti *Naïve Bayes Classifier*, *Maximum Entropy*, *Regresi Logistik*, dan *Support Vector Machine* (Wan dan Gao 2016). Akan tetapi, studi terkini menunjukkan bahwa metode *ensemble* dapat meningkatkan *robustness* dan performa dari analisis sentimen twitter (Kim *et al.* 2011). Metode *ensemble* merupakan metode yang mengkombinasikan hasil keputusan dari berbagai model klasifikasi dasar (*base classifier*) menjadi satu *classifier* sehingga hasil klasifikasi dapat menjadi lebih akurat.

Penelitian yang dilakukan (Kumari *et al.* 2021) menunjukkan bahwa analisis sentimen dengan metode *Ensemble* (kombinasi base classifier *Regresi Logistik*, *Multinomial Naïve Bayes*, dan *Random Forest* memiliki performa yang paling baik dengan akurasi 97,27% dibandingkan dengan performa saat menggunakan *single classifier Regresi Logistik, Multinomial Naïve Bayes, dan Random Forest* secara terpisah. Penelitian serupa juga dilakukan oleh (Wan dan Gao 2016) dimana metode *ensemble* meraih akurasi tertinggi yaitu sebesar 91,7% dibandingkan dengan metode yang menggunakan *single classifier*. Oleh karena itu, penelitian ini menggunakan metode *Ensemble* dalam analisis sentimen layanan moda transportasi Biskita.

1.2 Tujuan

Tujuan dari penelitian ini adalah mengidentifikasi performa penggunaan metode *Ensemble* dalam mengklasifikasikan ulasan dan opini publik pengguna twitter terhadap layanan transportasi Biskita Transpakuan Bogor.

1.3 Manfaat

Penelitian ini bermanfaat untuk melihat ulasan dan opini publik pengguna twitter serta memberikan gambaran mengenai performa algoritma *Ensemble* dalam melakukan analisis sentimen terhadap layanan Transportasi Biskita Transpakuan Bogor. Hasil dari penelitian ini diharapkan mampu memberikan suatu pengetahuan baru dan sebagai bahan evaluasi bagi pihak yang berkepentingan untuk terus meningkatkan mutu dan pelayanan yang lebih baik sehingga tetap mendapatkan kepercayaan dari para penggunanya.

II TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen adalah bidang studi yang melakukan analisis terhadap opini, emosi, dan sikap masyarakat yang berbentuk teks (Medhat *et al.* 2014). Analisis sentimen merupakan irisan dari bidang linguistik dan ilmu komputer yang digunakan untuk mengotomatisasi proses penentuan sentimen. Analisis sentimen merupakan salah satu cabang ilmu dari *text mining*, *natural language processing*, dan *artificial intelligence*. Penemuan informasi sentimen yang tersembunyi dari sejumlah besar data tekstual itu tergolong sulit. Namun, dengan adanya perkembangan teknologi memberikan kemampuan untuk mengekstraksi pendapat dan sentimen secara akurat. Tujuan utama analisis sentimen adalah mengartikan, mengekstrak, dan mengolah data teks secara otomatis untuk menghasilkan informasi yang bernilai.

2.2 Term Frequency - Inverse Document Frequency (TF – IDF)

TF-IDF adalah metode yang digunakan dalam pemrosesan teks dan pengambilan informasi untuk mengevaluasi sejauh mana sebuah kata kunci relevan terhadap sebuah dokumen dalam sebuah koleksi dokumen atau korpus. TF-IDF biasa digunakan sebagai proses pembobotan kata pada *information retrieval*. Model pembobotan TF-IDF mengasumsikan bahwa terms dengan frekuensi rendah memiliki bobot yang tinggi, sedangkan terms dengan frekuensi tinggi memiliki bobot yang rendah. Bobot ini diartikan sebagai pentingnya sebuah kata dalam suatu dokumen. Menurut (Ma *et al.* 2011) bobot dihitung dengan mengalikan antara nilai seberapa sering kata tersebut muncul dalam dokumen (*Term Frequency*) dengan nilai seberapa unik kata tersebut di dalam seluruh koleksi dokumen (*Inverse Document Frequency*).

$$w_{ij} = tf_{ij} \times idf_i \quad (1)$$

$$idf_i = \log \left(\frac{N}{df_i} \right) \quad (2)$$

dengan:

- i : 1,2, ..., p ; p = banyaknya kata
- j : 1,2, ..., N ;
- N : Banyaknya seluruh dokumen
- w_{ij} : Bobot dari kata i pada dokumen ke j
- tf_{ij} : Banyaknya kemunculan kata i pada dokumen ke j
- idf_i : Invers dokumen frekuensi
- df_i : Banyaknya dokumen yang mengandung kata i

2.3 Synthetic Minority Over-sampling Technique (SMOTE)

Synthetic Minority Oversampling Technique (SMOTE) merupakan salah satu pendekatan untuk mengatasi masalah ketidakseimbangan data. Keadaan ketidakseimbangan data terjadi ketika suatu kelas memiliki jumlah anggota yang

jauh lebih besar dibandingkan kelas lainnya, mengakibatkan klasifikasi cenderung memberikan akurasi yang tinggi untuk kelas mayoritas daripada kelas minoritas (He dan Garcia 2009). Permasalahan ini menjadi penting dalam analisis dan klasifikasi data, karena klasifikasi yang dilakukan pada data yang seimbang cenderung menghasilkan prediksi yang lebih akurat.

SMOTE menangani ketidakseimbangan data dengan menghasilkan data sintetis untuk kelas minoritas, sehingga proporsi data menjadi lebih seimbang. Data sintetis ini dihasilkan dengan memanfaatkan konsep tetangga terdekat. Pada data numerik, jarak kedekatannya diukur dengan jarak euclidean. Prosedur pembangkitan data sintesis untuk data numerik dimulai dengan menghitung perbedaan antara vektor utama dengan k tetangga terdekatnya. Selanjutnya, kalikan perbedaan dengan angka yang diacak antara nol dan satu. Terakhir, tambahkan perbedaan tersebut ke dalam nilai utama pada vektor utama asal sehingga diperoleh vektor utama yang baru (Chawla *et al.* 2002).

2.4 Hyperparameter Tuning

Hyperparameter adalah parameter yang menentukan arsitektur dan perilaku model, dan tidak dipelajari secara langsung dari data, namun ditentukan oleh pengguna sebelum model dilatih (James *et al.* 2013). *Hyperparameter tuning* adalah proses untuk menentukan kombinasi optimal dari *hyperparameter* pada model *machine learning* untuk meningkatkan performanya. Proses *hyperparameter tuning* biasanya melibatkan banyak percobaan yang dilakukan secara berulang dengan berbagai kombinasi nilai *hyperparameter*. Nantinya akan dipilih kombinasi yang bisa menghasilkan performa model terbaik. Menurut Pedregosa *et al* (2011) beberapa *hyperparameter* pada model *machine learning* yaitu sebagai berikut

1. *Hyperparameter* pada model Multinomial Naïve Bayes seperti *alpha* digunakan untuk mengontrol seberapa kuat estimasi model terhadap data pelatihan. *Alpha* dapat mempengaruhi tingkat *smoothing* pada distribusi peluang kelas.
2. Model Regresi Logistik Multinomial memiliki beberapa *hyperparameter* yang memengaruhi proses pelatihan. *Solver* digunakan untuk menentukan algoritma optimasi yang akan digunakan dalam proses pelatihan. *Hyperparameter penalty* memungkinkan pengguna untuk menyesuaikan jenis regularisasi yang diterapkan pada model, baik itu l1 atau l2. Sedangkan parameter C mengontrol kekuatan regularisasi, di mana nilai yang lebih tinggi mengindikasikan regularisasi yang lebih lemah. Secara umum, nilai C yang lebih kecil dapat memberikan akurasi yang lebih baik, tetapi dapat meningkatkan risiko *overfitting*. Nilai C yang lebih besar dapat mengurangi risiko *overfitting*, tetapi dapat menurunkan akurasi.
3. Pada model Random Forest, terdapat tiga *hyperparameter* utama yang memengaruhi konstruksi pohon keputusan, yaitu *min_samples_leaf*, *max_features*, dan *n_estimators*. *Min_samples_leaf* mengatur jumlah minimum sampel yang diperlukan untuk membentuk daun pohon, yang dapat mempengaruhi kompleksitas dan generalisasi model. *Max_features* mengontrol jumlah fitur yang dipertimbangkan untuk *split* pada setiap *node* pohon, sedangkan *n_estimators* menentukan jumlah pohon yang akan dibangun.

2.5 Multinomial Naïve Bayes

Multinomial Naive Bayes adalah metode klasifikasi secara statistik yang dapat memprediksi suatu data ke dalam kelas tertentu (Han *et al.* 2012). Metode ini merupakan varian dari algoritma Naive Bayes untuk klasifikasi dalam konteks data kategorik berasal dari teori probabilitas dan teorema Bayes. Metode Naïve Bayes disebut naive karena mengasumsikan independensi antara setiap peubah penjelas dalam setiap kelas respons. Algoritma ini bekerja dengan cara memprediksi probabilitas terjadinya suatu kejadian di masa depan berdasarkan kejadian sebelumnya (Yang 2018). Ini memanfaatkan peluang bersyarat untuk memprediksi kelas atau label suatu dokumen atau data berdasarkan fitur-fitur yang ada. Persamaan yang digunakan dalam klasifikasi Naïve Bayes didefinisikan pada Persamaan (3).

$$P(\text{label}|\text{fitur}) = \frac{P(\text{label}) \times P(\text{fitur}|\text{label})}{P(\text{fitur})} \quad (3)$$

Pada metode Naïve Bayes, setiap fitur dianggap independen sehingga memiliki peluang yang berbeda dan tidak berhubungan satu sama lain sehingga Persamaan (3) dijabarkan kembali menjadi Persamaan (4).

$$P(\text{label}|f_i) = \frac{P(\text{label}) \times P(f_1|\text{label}) \times P(f_2|\text{label}) \times \dots \times P(f_n|\text{label})}{P(\text{fitur})} \quad (4)$$

dengan:

$P(\text{label} \text{fitur})$: Probabilitas kondisional dari label diberikan fitur
$P(\text{label})$: Probabilitas keseluruhan dari label.
$P(\text{fitur} \text{label})$: Probabilitas kondisional dari fitur diberikan label.
$P(\text{fitur})$: Probabilitas keseluruhan dari fitur.
f_i	: Fitur ke – i.

2.6 Regresi Logistik Multinomial

Untuk melakukan analisis pada data kategorik, analisis regresi yang sesuai untuk digunakan adalah analisis regresi logistik. Regresi logistik dapat digunakan untuk menggambarkan hubungan peubah prediktor dan peubah respon yang memiliki dua atau lebih peubah kategori (Hosmer dan Lemeshow. 2000). Pada versi dasar Regresi Logistik, output dari metode ini merupakan bilangan *binary*. Akan tetapi, dapat juga berupa banyak kelas yang nantinya dikenal sebagai Regresi Logistik Multinomial (Cox 1958). Menurut Hosmer dan Lemeshow (2000) bentuk Persamaan regresi logistik biner adalah sebagai berikut:

$$\pi(x) = P(Y = j|x) = \frac{e^{g(x)}}{1 + e^{g(x)}} \quad (5)$$

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (6)$$

dengan:

$P(Y = j x)$: Peluang bersyarat dari peubah respon Y untuk kategori ke-j pada vektor x.
--------------	---

$\pi(x)$

: Persamaan regresi logistik untuk peubah respon Y untuk kategori ke- j dengan kovariat x

$g(x)$

: Logit pada peubah respon Y untuk kategori ke- j

x

: vektor dari observasi peubah prediktor

j

: indeks untuk kategori peubah respon ($j = 0, 1, 2$)

k

: indeks untuk peubah prediktor ($k = 0, 1, 2, \dots, p$)

i

: indeks untuk pengamatan ($i = 0, 1, 2, \dots, n$)

n

p

q

r

s

t

u

v

w

x

y

z

α

β

γ

δ

ϵ

ζ

η

θ

ι

κ

λ

μ

ν

ξ

\omicron

π

ρ

σ

τ

υ

ϕ

χ

ψ

ω

φ

ϑ

ϖ

$\var�$

$\var�$

$\var�$

$\var�$

$\var�$

$\var�$

$\var�$

Regresi Logistik Multinomial hampir sama dengan logistik biner. Regresi Logistik Multinomial melibatkan perbandingan antara dua level kategori, di mana satu level dijadikan referensi, sementara level lainnya dianggap sebagai peubah pembanding. Jika regresi logistik multinomial memiliki j kategori, persamaan yang terbentuk akan sebanyak $j-1$. Sebagai contoh, ketika $j = 3$, akan terbentuk dua persamaan logit seperti pada Persamaan 7 dan Persamaan 8.

$$g_1(x) = \ln \left[\frac{P(Y = 1|x)}{P(Y = 0|x)} \right] = \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 = \mathbf{X}\boldsymbol{\beta}_1 \quad (7)$$

dan

$$g_2(x) = \ln \left[\frac{P(Y = 2|x)}{P(Y = 0|x)} \right] = \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 = \mathbf{X}\boldsymbol{\beta}_2 \quad (8)$$

dengan:

$\boldsymbol{\beta}_j$: vektor koefisien kategori ke- j

\mathbf{X} : vektor yang berisi k peubah prediktor

Kemudian bentuk umum persamaan untuk setiap kategori terdapat pada Persamaan 9.

$$\pi_j(x) = \frac{e^{g_j(x)}}{\sum_{j=0}^2 e^{g_j(x)}} \quad (9)$$

Sehingga bentuk persamaan peluang untuk setiap kategori seperti pada Persamaan 10, 11 dan 12 berikut dimana $e^{g_0(x)} = 1$:

$$\pi_0(x) = \frac{1}{1 + e^{g_1(x)} + e^{g_2(x)}} \quad (10)$$

$$\pi_1(x) = \frac{e^{g_1(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}} \quad (11)$$

$$\pi_2(x) = \frac{e^{g_2(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}} \quad (12)$$

2.7 Random Forest

Sejak diperkenalkan pertama kali oleh Leo Breiman (2001), penggunaan random forest telah meningkat pesat dan terbukti memiliki beberapa keunggulan. Keunggulan tersebut melibatkan kemampuannya untuk menangani masalah *overfitting*, ketidakpekaan terhadap pencilan, dan kemampuan untuk menghasilkan

akurasi yang tinggi (Ali *et al.* 2012). Random forest juga efektif diterapkan pada data yang tidak seimbang dengan jumlah besar, memberikan kinerja yang baik dan waktu eksekusi yang cepat (More dan Rana 2017). Metode Random Forest juga dikenal sebagai bagian dari *ensemble learning*, metode ini membangun banyak pohon keputusan pada tahap pelatihan dan menggunakan metode *majority voting* pada tahap klasifikasi untuk menentukan keputusan akhir (Tan *et al.*, 2006).

Klasifikasi dengan random forest menerapkan metode *bagging* dan *random feature selection*. Selama tahap pelatihan, random forest menggunakan teknik umum yang dikenal sebagai *Bagging* terhadap setiap pohon keputusan. Teknik *Bagging* ini secara acak memilih sampel dari set pelatihan, dan proses *fitting* dilakukan pada sampel-sampel tersebut (Caie *et al.* 2020). Hal ini bertujuan untuk mengurangi korelasi antar pohon yang terbentuk (Hastie *et al.* 2008). Berikut adalah tahapan klasifikasi dalam menggunakan random forest pada gugus data yang terdiri atas a amatan dan p peubah penjelas (Sartono dan Syafitri 2010).

1. Tahapan *bootstrap*. *Bootstrap* adalah pengambilan contoh yang disertai dengan pengembalian. Pada tahap ini, ambil sebanyak n contoh acak dari data latih.
2. Tahapan *random feature subset*. Susun pohon berdasarkan data hasil *bootstrap* sebelumnya. Pada setiap proses pemisahan, pilih m peubah penjelas secara acak dimana $m < p$. Selanjutnya, lakukan pemisahan terbaik.
3. Ulangi langkah 1-2 sebanyak k kali sehingga diperoleh k buah pohon.
4. Lakukan penggabungan (agregasi) hasil prediksi k buah pohon dan gunakan majority vote (mengambil vote terbanyak) untuk menentukan hasil prediksi akhir.

Setiap proses pemisahan dalam pembentukan pohon perlu memperhatikan nilai entropi, kemudian dilakukan perhitungan *information gain*. Peubah dengan nilai *information gain* terbesar akan digunakan sebagai partisi atau pemisah terbaik. Andaikan dataset D dipartisi menjadi beberapa bagian ($D_1, D_2, \dots D_k$). Berikut adalah rumus untuk melakukan perhitungan entropi dan *information gain* yang dipelopori oleh Claude Shannon di dalam teori informasi (Gallager 2021).

$$Entropi(D) = - \sum_{i=1}^k p_i \log_2(p_i) \quad (13)$$

$$IG = Entropi(D) - \sum_{i=1}^k \frac{|D_i|}{|D|} \times Entropi(D_i) \quad (14)$$

dengan:

- D : Gugus data
 k : Jumlah partisi D
 p_i : Proporsi dari D_i terhadap D
 IG : *Information Gain*
 $|D_i|$: Jumlah observasi pada partisi ke- i
 $|D|$: Jumlah observasi pada D

Selain entropi terdapat ukuran lain yang dapat digunakan, yaitu gini. Setiap proses pemisahan akan memperhatikan nilai gini untuk kemudian dihitung nilai *gini decrease*. Berikut perhitungan rumus yang dilakukan.

$$Gini(D) = 1 - \sum_{i=1}^k (p_i^2) \quad (15)$$

$$GD = gini(D) - \sum_{i=1}^k \frac{|D_i|}{|D|} \times gini(D_i) \quad (16)$$

dengan:

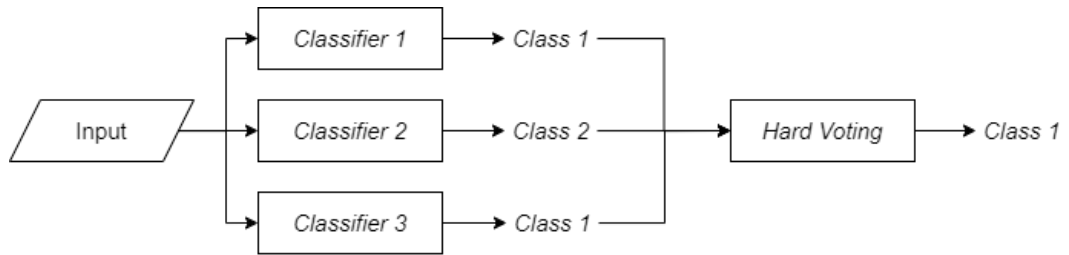
- D : Gugus data
- D : Jumlah partisi D
- D_i : Proporsi dari D_i terhadap D
- GD : *Gini decrease*
- $|D_i|$: Jumlah observasi pada partisi ke- i
- $|D|$: Jumlah observasi pada D

2.8 Ensemble

Metode *Ensemble* menggabungkan hasil keputusan dari berbagai model klasifikasi dasar menjadi satu classifier (Da Silva *et al.* 2014). Pendekatan ini meningkatkan *robustness* dan kinerja analisis sentimen. Dalam penelitiannya, (Dietterich 2000) menyatakan bahwa terdapat setidaknya tiga keuntungan dalam menggunakan sistem berbasis *Ensemble*:

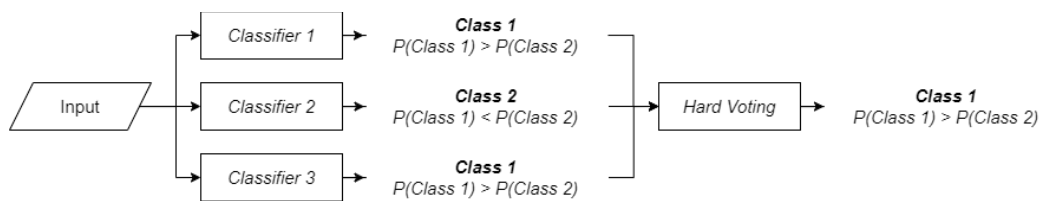
- a. Segi Statistik
Diasumsikan terdapat beberapa classifier dengan akurasi baik pada data latih. Jika hanya satu yang dipilih, kemungkinan kesalahan klasifikasi dapat terjadi pada data lain yang tidak dicakup oleh *classifier* tersebut. Dengan menggabungkan berbagai *classifier*, risiko kesalahan dapat dikurangi.
- b. Segi Komputasional
Ensemble memanfaatkan berbagai model untuk mencari solusi dari berbagai sudut, mengatasi keterbatasan pencarian lokal algoritma pembelajaran mesin. Dengan menggabungkan output model-model tersebut, *ensemble* dapat memberikan gambaran komprehensif yang mendekati solusi global. Hal ini tidak hanya meningkatkan stabilitas solusi tetapi juga memberikan aproksimasi yang lebih baik. Keunggulan ini terletak pada kemampuan *ensemble* untuk mengatasi variasi dan kompleksitas solusi lokal dari berbagai sumber.
- c. Segi Representasional
Ensemble mengatasi keterbatasan model dalam menggambarkan *decision boundary* yang kompleks dengan menggabungkan hasil dari berbagai *classifier*. Dengan cara ini, *Ensemble* dapat mengoptimalkan representasi *decision boundary*, meningkatkan kemampuan model untuk menangani keanekaragaman dan kompleksitas, serta meningkatkan performa klasifikasi secara keseluruhan.

Dalam membangun Model *Ensemble*, ada beberapa *hyperparameter* yang dapat diatur, seperti *voting* dan *weight*. Parameter *voting* memiliki dua opsi, yaitu metode Mayoritas (*Hard Voting*) dan Rata-rata (*Soft Voting*) (Pedregosa *et al.* 2011). Pendekatan *hard voting* akan menentukan hasil berdasarkan keputusan mayoritas dari semua *classifier*, sebagaimana dijelaskan dalam Gambar 1.



Gambar 1 Ilustrasi *Hard Voting* (da Silva *et al.* 2014).

Sedangkan pada *Soft Voting*, hasil akan dipilih berdasarkan rata-rata peluang kelas tersebut terpilih dari setiap *classifier*. Hasil yang akan dipilih yaitu merupakan rata - rata peluang terbesar. Ilustrasi dari *soft voting* dapat dilihat pada Gambar 2.

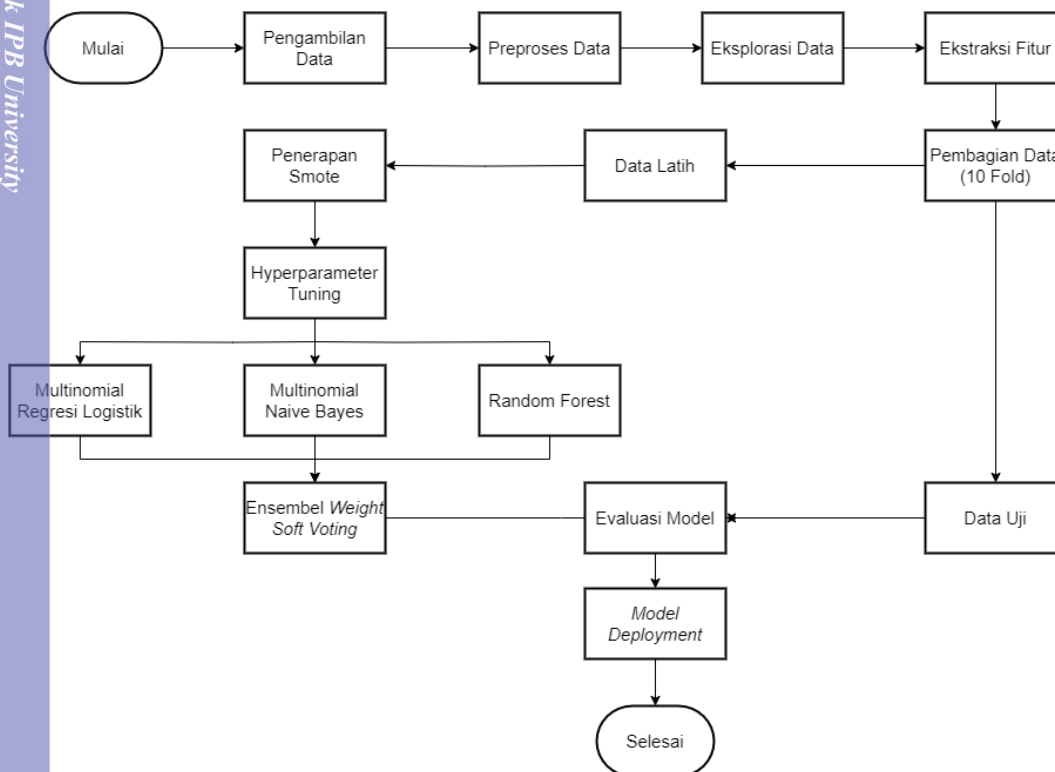


Gambar 2 Ilustrasi *Soft Voting* (da Silva *et al.* 2014).

III METODE

3.1 Tahapan Penelitian

Penelitian ini terdiri dari beberapa tahapan yaitu pengambilan data, praproses data, pelabelan data, eksplorasi data, ekstraksi fitur, pembagian data menjadi data latih dan uji, penerapan smote, *hyperparameter tuning*, pembuatan *base classifiers*, pembuatan model *ensemble*, kemudian evaluasi model. Diagram alir penelitian dapat dilihat pada Gambar 3.



Gambar 3 Diagram Alir Penelitian

3.2 Data

Penelitian ini menggunakan data sekunder dari media sosial Twitter sebanyak 4306 *tweets*. Data diambil dengan dua kata kunci yaitu “Biskita” sebanyak 3500 *tweets* dan “Transpakuan” sebanyak 806 *tweets* diambil dari tanggal 26 Juni 2022 sampai dengan 25 Juni 2023. Atribut yang diambil berupa ID tweet, isi tweet, waktu, dan username. Contoh data tweets dapat dilihat pada Tabel 1.

Tabel 1 Contoh Data Tweet

Date	Id	Tweet	Username
2023-06-14 11:49:30	1.66895E+18	najis banget gue nunggu biskita hampir 40 menit, lama banget dih	fre***vv
2023-03-08 04:45:02	1.63333E+18	@bogorfess_ Biskita enak bat k5 turun smp5 jalan deh 5 menit.	Ilminmit***in

3.3 Preproses Data

Praproses data adalah langkah-langkah yang dilakukan sebelum data disajikan ke model pembelajaran mesin atau sebelum dilakukan analisis statistik. Tujuan utama praproses data adalah meningkatkan kualitas data, mengurangi *noise*, dan memastikan data siap digunakan untuk tujuan analisis atau pembelajaran mesin.

Praproses data pada penelitian ini akan dilakukan meliputi proses – proses berikut.

3.3.1 Penghapusan Data Duplikat

Pengambilan data dilakukan menggunakan dua kata kunci yaitu ‘biskita’ dan ‘transpakuan’. Hal ini berpotensi menimbulkan duplikasi data ketika data yang sama terambil dua kali. Sehingga praproses data dimulai dengan tahapan pemilihan data untuk menghapus data duplikat.

3.3.2 Pelabelan Data

Pelabelan data dilakukan untuk menentukan kelas sentimen ke dalam kelas positif, negatif, netral atau spam. Pelabelan ini akan dilakukan secara manual agar meningkatkan akurasi dalam pengklasifikasian tweet (Anastasia dan Budi 2016). Pelabelan akan dilakukan oleh tiga orang dan dilakukan secara *online*, jika terdapat perbedaan label antara pelabel, maka akan dilakukan *checking* kembali pada *tweet* dengan label berbeda, lalu dilakukan persamaan persepsi untuk menentukan label final agar tingkat kesalahan bisa diminimalisir. Pelabelan akan dilakukan secara bersama – sama dengan aturan pelabelan dapat dilihat pada Tabel 2.

Tabel 2 Aturan Pelabelan Tweet

Label	Aturan
Negatif	Pernyataan berupa respon negatif terhadap topik “Pelayanan Transportasi Biskita Transpakuan Bogor” ditandai dengan adanya tolakan, kekecewaan, ketidakpuasan yang dapat disertai sindiran ataupun kritik.
Netral	Pernyataan netral yang tidak memuat informasi mengenai dukungan ataupun tolakan.
Positif	Pernyataan berupa respon positif terhadap topik “Pelayanan Transportasi Biskita Transpakuan Bogor” ditandai dengan adanya dukungan, ajakan, dan/atau pemberian informasi terkait manfaat positif sebagai upaya pencerdasan kepada masyarakat.
Spam	Kategori ini berisi tweet yang tidak relevan terhadap topik. Tweet ini akan dianggap sebagai spam.

3.3.3 Data Cleaning dan *Case Folding*

Pada proses data cleaning dilakukan menggunakan *Regular Expression* (*RegEx*) berupa penghapusan karakter non-ASCII, URL, mentions, tagar, simbol dan angka, spasi di awal dan di akhir kalimat, spasi ganda dan duplikasi tiga karakter beruntun atau lebih. Setelah dilakukan data cleaning

kemudian dilakukan *Case Folding* untuk menyeragamkan karakter pada data dengan mengkonversi menjadi karakter non kapital.

3.3.4 Normalisasi Kata

Normalisasi kata adalah melakukan penyeragaman terhadap format kata yang memiliki makna sama namun memiliki penulisan yang berbeda. Hal ini berguna karena tidak jarang ditemui terdapat kesalahan penulisan dan juga tidak digunakannya kata yang baku pada tweet. Pada proses ini akan menggunakan dua kamus. Kamus pertama yaitu *Colloquial Indonesian Lexicon* dari penelitian (Salsabila *et al.* 2019) dan kamus kedua yang peneliti sesuaikan dengan topik yang disusun secara manual dapat dilihat di Lampiran 1.

3.3.5 Stopwords Removal

Stopwords Removal merupakan proses untuk menghilangkan kata yang termasuk *stopwords*. *Stopwords* adalah kata-kata umum yang jumlahnya banyak tapi tidak memiliki kekuatan dan kepentingan prediksi dalam klasifikasi, seperti kata depan, kata ganti orang, dan kata hubung. Kata tersebut dinilai tidak relevan dan tidak memberikan kontribusi signifikan terhadap pemahaman konten teks, sehingga perlu untuk dihapus dari data (Amin 2012). Proses *stopword removal* ini dilakukan dengan bantuan library *Natural Language Toolkit* (NLTK) pada bahasa pemrograman python.

3.3.6 Stemming

Proses *stemming* merupakan proses penguraian kata berimbuhan menjadi bentuk kata dasarnya yang disebut "*stemma*" atau "akar kata". Proses ini dilakukan dengan menghapus semua imbuhan baik awalan, sisipan, dan akhiran. Proses *stemming* dilakukan dengan bantuan library *sastrawi* pada bahasa pemrograman python.

3.3.7 Tokenisasi

Tokenisasi adalah proses memecah teks kalimat menjadi unit-unit kecil yang disebut token. Token bisa berupa kata, frasa, atau karakter. Pilihan unit-token tergantung pada konteks dan kebutuhan penelitian. Namun pada penelitian ini, token akan berbentuk potongan kata. Tujuan utama dari tokenisasi adalah memudahkan analisis dan pemrosesan teks dengan mengubah teks kontinu menjadi serangkaian unit yang dapat dioperasikan.

3.4 Eksplorasi Data

Setelah dilakukan *preprocessing* pada data, data dapat dieksplorasi untuk membantu memahami dan menemukan informasi pada data. Langkah ini melibatkan pemahaman mendalam terhadap karakteristik dataset, termasuk distribusi sentimen, frekuensi kata kunci terkait transportasi, serta potensi aspek yang paling sering dibahas. Melalui eksplorasi data, dapat mengidentifikasi pola unik dalam data, seperti topik yang paling banyak dibicarakan, istilah yang sering muncul, dan variasi dalam sentimen pengguna. Eksplorasi data dapat dilakukan

dengan analisis sentimen secara umum, analisis sentimen berdasarkan waktu, dan visualisasi *wordclouds*.

3.5 Pembuatan Model

3.5.1 Ekstraksi Fitur

Ekstraksi fitur bertujuan untuk mengekstraksi tweet dari bentuk teks ke bentuk numerik sehingga dapat diproses sebagai input dari model klasifikasi (da Silva *et al.* 2014). Dalam penelitian ini ekstraksi fitur dilakukan menggunakan metode *Term Frequency – Inverse Document Frequency (TF-IDF)*. Metode TF-IDF memiliki beberapa *hyperparameter* yang bisa disesuaikan seperti *max_features*, dan *n_gram*. Ilustrasi dari hasil ekstraksi fitur dapat dilihat pada Tabel 3.

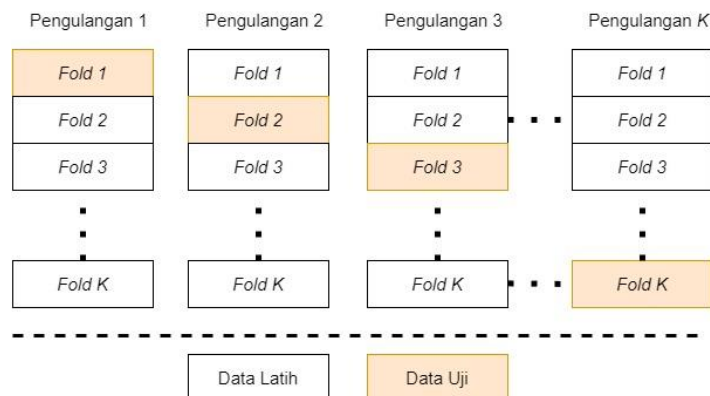
Tabel 3 Ilustrasi Hasil Ekstraksi Fitur

	<i>term₁</i>	<i>term₂</i>	...	<i>term_m</i>
<i>text₁</i>	<i>a₁₁</i>	<i>a₁₂</i>	...	<i>a_{1m}</i>
<i>text₂</i>	<i>a₂₁</i>	<i>a₂₂</i>	...	<i>a_{2m}</i>
...
<i>text_n</i>	<i>a_{n1}</i>	<i>a_{n2}</i>	...	<i>a_{nm}</i>

Pada Tabel 3, *n* dan *m* berturut-turut merepresentasikan jumlah *text* dan *terms*. Setiap *text* akan direpresentasikan dalam baris sebagai *text_i* = (*a_{i1}*, *a_{i2}*, *a_{i3}*, ..., *a_{im}*) dimana *a_{ij}* merupakan bobot dari *term_j* yang terdapat pada *text_i*. Bobot dari tiap *term* ini akan dihitung dengan metode *Term Frequency-Inverse Document Frequency (TF-IDF)* menggunakan modul *TfidfVectorizer* dari *library sklearn*.

3.5.2 Pembagian Data dan Penerapan SMOTE

Dalam membangun model perlu dilakukan evaluasi terhadap model. Sehingga dataset perlu dibagi menjadi data latih dan data uji. Data latih digunakan untuk melatih model dalam mempelajari pola dari data latih. Data uji digunakan untuk evaluasi terhadap performa model. Setiap proses pembagian data pada penelitian ini akan menggunakan metode *K-Fold Cross Validation*. Metode ini akan membagi data sebanyak *K* bagian (*fold*) dimana pada setiap bagian akan diterapkan validasi terhadap model dengan satu bagian sebagai data uji, dan sisanya sebagai data latih. Pada setiap iterasi akan dilakukan proses validasi pada *fold* yang berbeda sebanyak *K* kali. Ilustrasi pembagian data dapat dilihat pada Gambar 4.



Gambar 4 Ilustrasi *K-Fold Cross Validation* (Ren et al. 2019)

Pada penelitian ini nilai k yang digunakan adalah 10, sehingga *fold* yang terbentuk sebanyak 10 *fold*, dalam hal ini berarti data latih dan data uji akan terbentuk sebanyak 10 pasangan data latih dan data uji. Kemudian diterapkan SMOTE pada seluruh data latih untuk menangani ketidakseimbangan data.

3.5.3 Hyperparameter tuning

Setiap proses membangun model diperlukan *hyperparameter* yang dapat disesuaikan. Proses penyesuaian ini dapat dilakukan secara manual ataupun dengan bantuan modul. Proses penyesuaian dengan bantuan modul disebut *hyperparameter tuning*. Dalam penelitian ini, modul yang digunakan yaitu GridSearchCV pada bahasa pemrograman Python. Adapun *hyperparameter* yang akan dicari yaitu pada Tabel 4 berikut

Tabel 4 *Hyperparameter tuning Base Classifiers*

Model	Hyperparameter
Regresi Logistik Multinomial	Solver = ['liblinear', 'lbfgs'] Penalty = ('l1', 'l2') C = np.logspace(-4,4,20)
Random Forest	Min_samples_leaf = [1, 2, 4] Max_features = ['auto', 'sqrt'] N_estimators = [5, 20, 50, 100, 200, 400, 600, 800, 1000]
Multinomial Naïve Bayes	Alpha = [0,01; 0,1; 1]

3.5.4 Pembuatan Model *Ensemble*

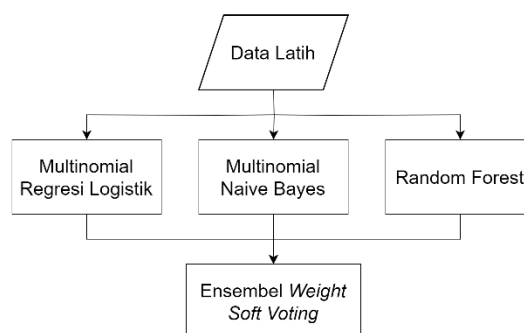
Pembuatan Model *Ensemble* dilakukan dengan teknik *Stacking*. Model akan dibangun dengan menggabungkan beberapa model yang disebut model dasar atau *base classifiers*. *Stacking* melibatkan penggabungan hasil prediksi dari model dasar dengan model meta yang memprediksi keluaran akhir berdasarkan hasil prediksi *base classifiers*. *Base classifiers* yang digunakan yaitu Regresi Logistik Multinomial (RLM), Random Forest (RF),

dan *Multinomial Naïve Bayes (MNB)*. Karakteristik masing-masing *base classifier* menurut (Pedregosa *et al.* 2011) dapat dilihat pada Tabel 5.

Tabel 5 Karakteristik *Base Classifiers*

No	Classifier	Kelebihan	Kekurangan
1	Regresi Logistik Multinomial	a. Mudah dan cepat untuk diimplementasikan b. Dapat memberikan informasi pada output berupa relevansi dan signifikansi fitur	a. Kurang cocok untuk data nonlinear b. sensitif terhadap overfitting
2	Random Forest	a. Relatif Akurat b. Memiliki performa yang baik dalam menangani berbagai kasus termasuk kasus pada data nonlinear.	a. Sulit diinterpretasikan b. Sensitif terhadap <i>hyperparameter</i>
3	<i>Multinomial Naïve Bayes</i>	a. Memiliki komputasi yang efisien b. <i>Robust</i> terhadap <i>noise</i> c. Mudah diimplementasikan	a. Atribut diasumsikan independen sehingga akan mengurangi akurasi jika diterapkan pada kasus yang memiliki atribut yang saling berkaitan

Model *Ensemble* diharapkan dapat memanfaatkan kelebihan dari masing – masing model dasar. Proses pembuatan model Ensemble dilakukan dengan membangun model dasar menggunakan *hyperparameter* terbaik yang dihasilkan dari proses *hyperparameter tuning*. Setelah model dasar dibangun, output dari masing-masing *classifier* tersebut akan menjadi input untuk membangun model *Ensemble* dengan ilustrasi pada Gambar 5 berikut:



Gambar 5 Diagram Alir Pembuatan Model *Ensemble*

Dalam membangun Model *Ensemble* dengan teknik *Stacking*, terdapat beberapa *hyperparameter* yang dapat disesuaikan seperti *voting* dan *weight*. Dalam penelitian ini model *Ensemble* akan menggunakan *hyperparameter soft voting* dan *weight*. Metode *weighted soft voting* akan mengembalikan nilai dari rata-rata penjumlahan peluang seluruh model dasar pada setiap kelas. Sebelumnya, nilai peluang masing-masing *classifier* akan dikalikan terhadap *weight* yang telah ditentukan. Kemudian kelas yang memiliki nilai peluang tertinggi akan dipilih menjadi output akhir klasifikasi (Pedregosa *et al.* 2011). *Weight* akan diberikan pada setiap *base classifiers* berdasarkan hasil akurasi dari masing – masing *base classifiers* dengan ilustrasi pada Tabel 6 berikut

Tabel 6 Ilustrasi *Weight Soft Voting*

Akurasi	Kelas		
	Negatif	Netral	Positif
Regresi Logistik Multinomial	$P_{11} \times w_1$	$P_{12} \times w_1$	$P_{13} \times w_1$
Random Forest	$P_{21} \times w_2$	$P_{22} \times w_2$	$P_{23} \times w_2$
Multinomial Naïve Bayes	$P_{31} \times w_3$	$P_{32} \times w_3$	$P_{33} \times w_3$
Rata - Rata	x_1	x_2	x_3

Perhitungan *weight classifier Ensemble* dengan *soft voting* ditunjukkan pada Tabel 6. P_{ij} merupakan peluang terprediksinya sebuah tweet masuk ke dalam kelas j oleh *classifier* i dan x_i merupakan rata-rata dari jumlah pada suatu kelas i . w_i merupakan *weight* yang ditentukan berdasarkan akurasi *classifier* tersebut. Kelas yang akan dipilih untuk menjadi label tweet merupakan nilai kelas dengan nilai x_i tertinggi untuk semua i .

3.6 Evaluasi Model

Model yang telah dibuat akan diuji dan dievaluasi dengan menggunakan matrik konfusi berdasarkan (Tharwat 2018). Dari matriks ini dapat evaluasi empat pengukuran yaitu akurasi, presisi, *recall*, dan skor F1. Ilustrasi matrik konfusi menurut Tharwat A (2018) dilihat pada Tabel 7

Tabel 7 Ilustrasi Matrik Konfusi Tiga Kelas

Prediksi	Aktual		
	$class_1$	$class_2$	$class_3$
$class_1$	TP ₁	E ₁₂	E ₁₃
$class_2$	E ₂₁	TP ₂	E ₂₃
$class_3$	E ₃₁	E ₃₂	TP ₃

Tabel 7 menunjukkan sembilan kemungkinan output dari model klasifikasi untuk tiga kelas 1, 2, dan 3. Ini mewakili elemen-elemen dari matrik konfusi 3×3. Pada Tabel 7, kolom-kolom mewakili kelas yang diprediksi, dan baris-baris

mewakili kelas yang sebenarnya. Terdapat sembilan kemungkinan di mana TP_1 adalah kasus di mana pengklasifikasi memprediksi sebagai kelas 1 dan sampel sebenarnya adalah kelas 1, dan E_{12} adalah jumlah prediksi kelas 1 yang keliru diklasifikasikan sebagai kelas 2. *False Positif (FP)* pada kelas ke - i merupakan jumlah seluruh prediksi yang sebenarnya bukan kelas ke - i tapi diklasifikasikan kedalam kelas ke - i , dalam hal ini berarti $FP_i = \sum E_i \dots$, sedangkan *False Negatif (FN)* pada kelas ke - i merupakan jumlah kelas ke - i yang sebenarnya tapi tidak terprediksi sebagai kelas ke - i , dalam hal ini berarti $FN_i = \sum E \dots_i$.

Menurut (Ballabio *et al.* 2018), akurasi didefinisikan sebagai akurasi keseluruhan menggunakan kelas yang diprediksi, sementara *recall*, dan presisi didefinisikan sebagai rata-rata statistik satu lawan semua atau *one versus all*. Rumus untuk menghitung akurasi, presisi, *recall*, dan skor F1 seperti berikut:

$$Accuracy = \frac{\sum_{i=1}^l TP_i}{n} \quad (17)$$

$$Precision = \frac{TP_i}{TP_i + FP_i} \quad (18)$$

$$Recall = \frac{TP_i}{TP_i + FN_i} \quad (19)$$

$$Skor F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (20)$$

3.7 Model Deployment

Model deployment merupakan tahapan implementasi model atau publikasi model yang telah dibangun agar bisa dimanfaatkan dan digunakan oleh orang lain. *Deploying* model dapat dilakukan dengan bahasa pemrograman Python dan *framework* Streamlit pada platform *huggingface.co*. Dalam hal ini model dapat memberikan label terhadap *tweet* dengan memprediksi apakah *tweet* tersebut mengandung sentimen negatif, netral atau positif.

IV HASIL DAN PEMBAHASAN

4.1 Preproses Data

Praproses data dimulai dengan menghapus tweet yang terduplikasi sebanyak 312 *tweets*. Contoh *tweets* duplikat dapat dilihat pada Tabel 8.

Tabel 8 Contoh *Tweet* Duplikat

Index	Date	Id Tweet	Tweet
408	2023-01-05 07:48:58+00:00	1610906186353872898	Tarif Biskita Transpakuan Diusulkan Rp5.500, Pemkot Kaji Tiket Terusan https://t.co/WcU10NgIRO
1670	2023-01-05 07:48:58+00:00	1610906186353872898	Tarif Biskita Transpakuan Diusulkan Rp5.500, Pemkot Kaji Tiket Terusan https://t.co/WcU10NgIRO

Setelah data bersih dari data duplikat, kemudian dilakukan pelabelan data untuk menentukan klasifikasi awal *tweets* kedalam empat kelas yaitu negatif, netral, positif dan spam. Pelabelan akan dilakukan oleh tiga orang, jika terdapat perbedaan label antara pelabel, maka akan dilakukan *checking* kembali pada *tweet* dengan label berbeda, lalu dilakukan persamaan persepsi untuk menentukan label final. Hal ini bertujuan agar tingkat kesalahan bisa diminimalisir. Pelabbelan dilakukan menggunakan *Spreadsheet*. Hasil final dapat dilihat pada Tabel 9.

Tabel 9 Hasil Pelabelan Data

No	Kelas	Jumlah
1	Negatif	907
2	Netral	1242
3	Positif	1318
4	Spam	627
Jumlah		4094

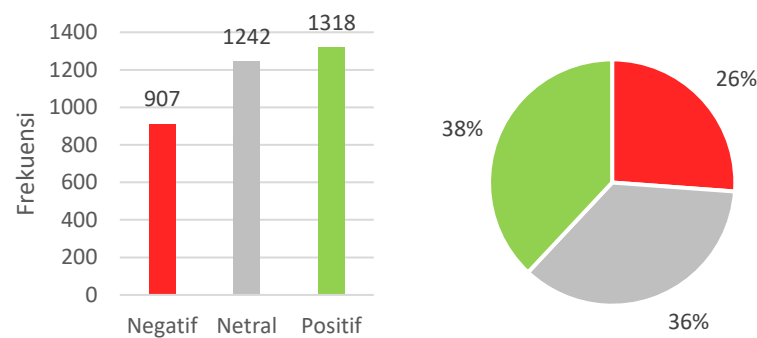
Tahapan selanjutnya yaitu dilakukan proses *data cleaning* dan *case folding*, normalisasi kata, penghapusan stopwords, *stemming*, dan tokenisasi. Pada proses normalisasi kata ditambahkan kamus kata gaul baru buatan sendiri, dapat dilihat pada Lampiran 1. Hasil preproses data dapat dilihat pada Tabel 10.

Tabel 10 Contoh Hasil Preproses Data

Tahapan	Tweet Setelah Preproses
Tweet Asli	@bogorfess_ Real, sabtu kemaren gw naek biskita cuma 2 orang, weekend yg biasanya penuh jaman gratisan dulu skrg bisa sepi gitu
Data Cleaning	real sabtu kemaren gw naek biskita cuma orang weekend yg biasanya penuh jaman gratisan dulu skrg bisa sepi gitu
Normalisasi Kata	real sabtu kemarin aku naik biskita cuma orang weekend yang biasanya penuh jaman gratisan dulu sekarang bisa sepi begitu
Stemming	real sabtu kemarin naik biskita orang weekend penuh jaman gratis sekarang bisa sepi
Tokenisasi	['real', 'sabtu', 'kemarin', 'naik', 'biskita', 'orang', 'weekend', 'penuh', 'jaman', 'gratis', 'sekarang', 'bisa', 'sepi']

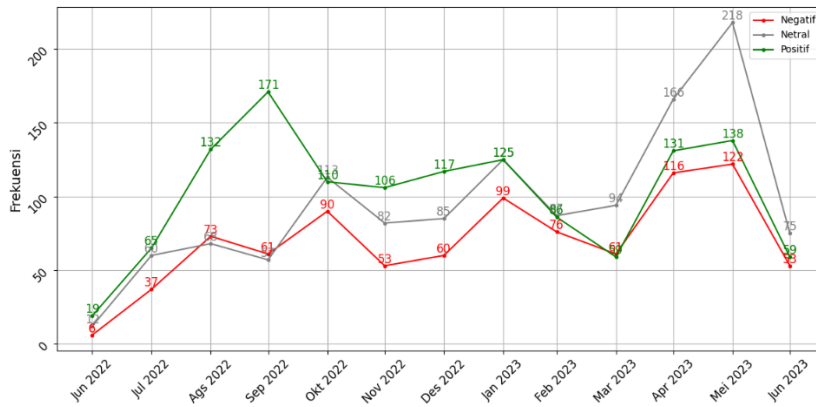
4.2 Eksplorasi Data

Pada tahap ini dilakukan eksplorasi data yaitu analisis sentimen secara umum, berdasarkan waktu, dan visualisasi. Secara umum, sentimen dari masing-masing kelas negatif, netral, dan positif berturut-turut memiliki data sebanyak 907, 1242, dan 1318 tweet. Tweet dengan sentimen positif memiliki persentase yang paling besar yaitu 38% diikuti oleh sentimen netral sebesar 36% dan sentimen negatif sebesar 26%. Visualisasi jumlah dan persentase tweet ini dapat dilihat pada Gambar 6.



Gambar 6 Jumlah dan Presentase Sentimen

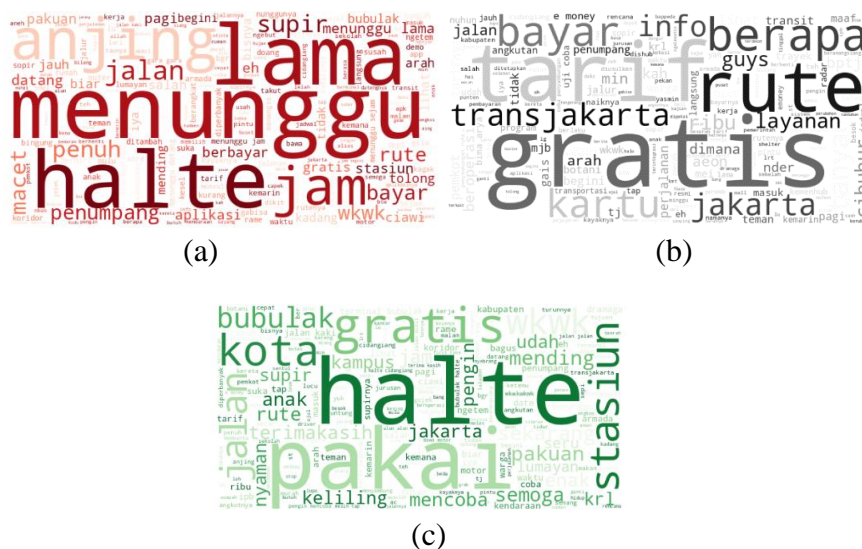
Setelah data dieksplorasi berdasarkan sentimen secara umum kemudian dilakukan eksplorasi data berdasarkan waktu untuk melihat tren dan perbandingan jumlah sentimen pada setiap bulannya sejak diambilnya tweet pada platform twitter. Hasil analisis sentimen berdasarkan waktu dapat dilihat pada Gambar 7.



Gambar 7 Lini Masa Sentimen

Sentimen positif tertinggi terjadi pada bulan september 2022 ini sesuai dengan adanya berita dengan judul *Biskita Transpakuan Mendapatkan Penghargaan Transportasi Ramah Anak* pada 2 September 2022 yang dimuat pada laman berita *klikterus*. Kemudian pada sentimen negatif titik tertinggi terjadi pada bulan Mei 2023, ini sesuai dengan adanya berita dengan judul *Biskita Transpakuan Bogor Berbayar Mulai 20 Mei, Tarif Rp 4000* pada 17 Mei 2023 yang dimuat pada laman berita *Detikcom*. Selain itu pada bulan maret hingga mei tahun 2023 terjadi tren kenaikan terhadap sentimen positif maupun sentimen negatif. Hal ini terjadi bersamaan dengan diberlakukannya tarif pada layanan Biskita Transpakuan Bogor. Sehingga peneliti menyimpulkan terdapat *pro* dan *kontra* terhadap kebijakan tarif pada layanan Biskita Transpakuan. Walaupun demikian, dalam rentang satu tahun sejak 26 juni 2022 sampai 25 juni 2023, sentimen positif selalu lebih tinggi pada setiap bulannya.

Setelah dilakukan eksplorasi data berdasarkan waktu dapat dilihat bahwa informasi sentimen terbatas hanya secara jumlah pada setiap bulannya. Untuk melihat informasi sentimen secara lebih detail kemudian dilakukan visualisasi *wordclouds* pada setiap kelas sentimen. Hasil visualisasi *wordclouds* dapat dilihat pada Gambar 8.

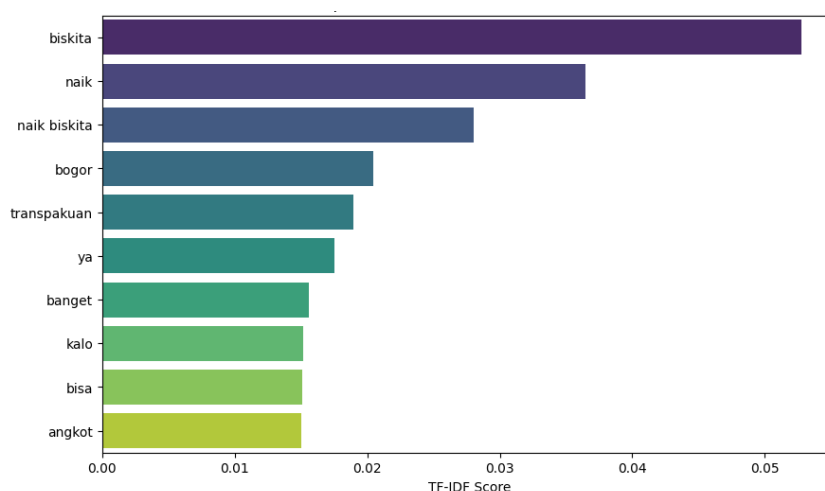


Gambar 8 *Wordclouds* Sentimen (a) Negatif (b) Netral dan (c) Positif

Hasil visualisasi *wordclouds* dapat diamati pada kelas negatif terdapat kata ‘menunggu’, ‘lama’, ‘anjing’, dan ‘penuh’ yang menunjukkan sentimen negatif berupa kekecewaan dan ketidakpuasan pengguna Biskita Transpakuan. Terdapat juga kata ‘supir’, ‘jam’, ‘jalan’, ‘berbayar’ dan ‘bubulak’ yang mengindikasikan bahwa kekecewaan pengguna ini terkait pengemudi biskita, waktu kedatangan, jalan pada rute biskita, kebijakan berbayar dan terkait halte Bubulak. Pada kelas netral terdapat kata ‘berapa’, ‘info’, dan ‘tanya’ yang menunjukkan sentimen netral berupa pertanyaan mengenai informasi Biskita Transpakuan. Terdapat juga kata ‘rute’, ‘tarif’ dan ‘kartu’ yang mengindikasikan bahwa pertanyaan dan informasi pada kelas netral terkait dengan informasi rute dan informasi pembayaran. Ini juga bisa memberikan *insight* bahwa Biskita Transpakuan masih kurang dalam memberikan layanan terkait informasi rute dan informasi pembayaran. Sedangkan pada kelas positif terdapat kata ‘pakai’, ‘mencoba’, ‘semoga’ dan ‘mending’ yang menunjukkan sentimen positif berupa antusiasme dan kepuasan pengguna terhadap layanan Biskita Transpakuan.

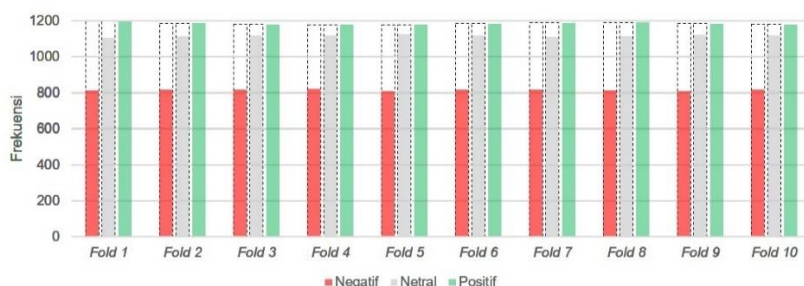
4.3 Pembuatan Model

Model *ensemble* dibangun menggunakan tiga *base classifiers* yaitu model Regresi Logistik Multinomial, *Multinomial Naïve Bayes* dan Random Forest. Tahapan pembuatan model dimulai dengan mengekstraksi *text* menjadi bentuk numerik. Pada tahap ini dilakukan ekstraksi fitur menggunakan *TfidfVectorizer* dari library *GridSearchCV* dengan *hyperparameter* yang telah dituning. Hasil tuning didapat *hyperparameter* terbaik yaitu $\text{max_features} = 22500$ dan $\text{ngram_range} = (1,2)$. Parameter max_features sebanyak 22500 berarti bahwa maksimal fitur yang dihasilkan berjumlah 22500 fitur. Fitur ini akan merepresentasikan nilai pada masing-masing sampel data yang berjumlah 3467 data. Sedangkan $\text{n_gram} = (1,2)$ berarti fitur akan dihasilkan dalam bentuk *unigram* (satu kata) dan *bigram* (dua kata). Hasil ekstraksi fitur akan berupa fitur dan skor TF-IDF. Skor TF-IDF akan merepresentasikan tingkat kepentingan suatu kata/fitur. Barchart 10 fitur dengan skor TF-IDF tertinggi dapat dilihat pada Gambar 9



Gambar 9 10 Fitur Dengan Skor TF-IDF Tertinggi

Tahapan selanjutnya yaitu pembagian data dan penerapan SMOTE. Pembagian data menggunakan *K-Fold Cross Validation* dengan *hyperparameter* $k = 10$ menghasilkan *fold* sebanyak 10 dengan data latih sebanyak 10 data berbeda. Smote diterapkan pada data latih pada setiap *fold*nya. Pada data latih kelas negatif, netral, dan positif berturut-turut memiliki proporsi masing-masing kelas yaitu sebesar 26%, 36% dan 38%. Terdapat kelas minoritas yaitu kelas negatif dengan presentase 26% dari keseluruhan data. Oleh karena itu, dilakukan proses *oversampling* data dengan algoritma SMOTE dapat dilihat pada Gambar 10.



Gambar 10 *Oversampling* Menggunakan SMOTE

Setelah dilakukan Smote pada data latih, kemudian data latih siap untuk digunakan untuk proses *training* model. Namun sebelum model dilatih, akan dilakukan proses *hyperparameter tuning* dengan menggunakan modul GridSearchCV untuk mencari *hyperparameter* terbaik berdasarkan data latih yang ada. Hasil *hyperparameter tuning* kemudian akan digunakan dalam membangun setiap model dasar. Hasil dari *Hyperparameter tuning* dapat dilihat pada Tabel 11.

Tabel 11 Hasil Tuning Parameter *Base Classifiers*

No	Classifier	Parameter Terbaik
1	Regresi Logistik Multinomial	$C = 11.28837$ penalty = l2 solver = liblinear
2	Random Forest	n_estimators = 200 min_samples_leaf = 1 max_features = log2
3	Multinomial Naïve Bayes	alpha = 0.1

Ensemble dibangun dengan menggunakan tiga *classifier* yang telah dilakukan *hyperparameter tuning* dan menggunakan metode *weighted soft voting*. *Weight* ditentukan berdasarkan akurasi masing-masing *classifier* pada setiap *fold*. Hasil akurasi dapat dilihat pada Tabel 12.

Tabel 12 Akurasi *Base Classifiers* pada setiap *Fold*

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
Regresi Logistik Multinomial	0,79	0,78	0,78	0,79	0,77	0,74	0,79	0,77	0,76	0,77
Random Forest	0,77	0,78	0,75	0,77	0,72	0,71	0,76	0,73	0,74	0,73
Multinomial Naive Bayes	0,82	0,75	0,72	0,75	0,74	0,70	0,74	0,73	0,75	0,72

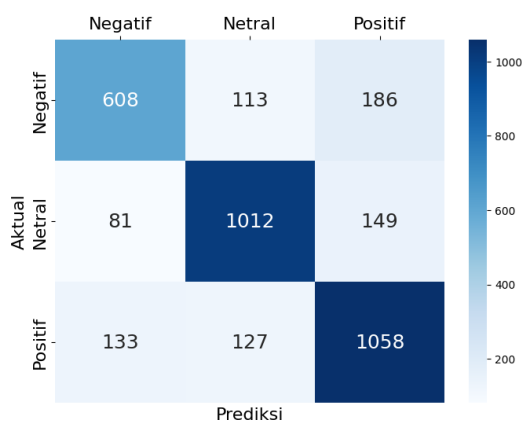
Weight pada pembuatan Model *Ensemble* akan dibuat berdasarkan akurasi setiap model pada setiap *fold*. Pada *fold* 1, *weight* yang diperoleh dari akurasi untuk *classifier* RLM yaitu sebesar 0,79; RF sebesar 0,77 dan MNB sebesar 0,82. Setelah *weight* untuk masing-masing *classifier* didapatkan, *classifier Ensemble* dibangun menggunakan keempat *base classifier* tersebut dengan *hyperparameter voting* = ‘soft’ dan *weight* = [0,79; 0,77; 0,82].

4.4 Evaluasi Model

Masing – masing model akan dievaluasi menggunakan Matrik konfusi dan akan dihitung akurasi dan skor-f1 pada setiap *fold* kemudian akan dihitung standar deviasi dari akurasi dan skor-f1.

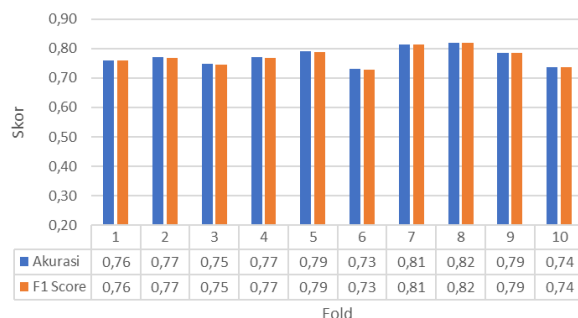
4.4.1. Regresi Logistik Multinomial

Regresi Logistik Multinomial memiliki nilai *true positif* (TP) kelas negatif, kelas netral dan kelas positif berturut – turut sebesar 608, 1012 dan 1058 prediksi benar. Nilai TP kelas positif pada *classifier* ini merupakan nilai *true positif* terbesar dari seluruh *base classifiers*. Sehingga model Regresi Logistik Multinomial merupakan model yang paling baik dalam memprediksi kelas positif. Namun *classifiers* ini memiliki nilai TP kelas negatif yang relatif kecil yaitu sebanyak 608.



Gambar 11 Matrik Konfusi Model Regresi Logistik Multinomial

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

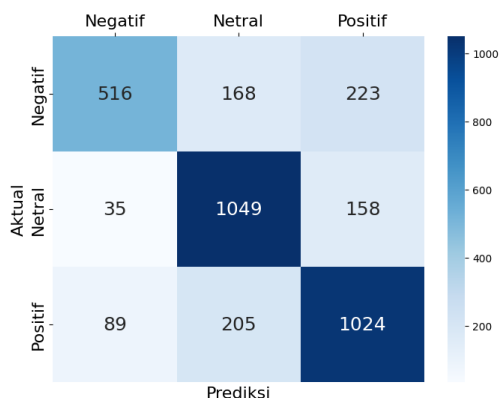


Gambar 12 Performa Model Regresi Logistik Multinomial pada setiap *Fold*

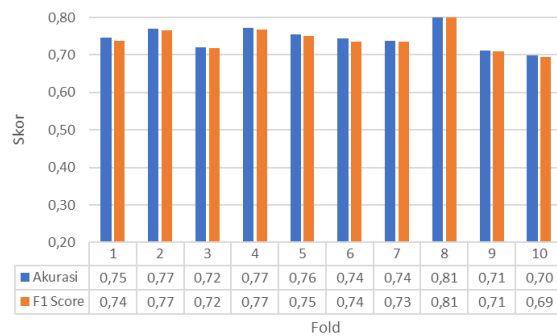
Dilihat dari performa setiap *fold* pada Gambar 12, *classifiers* ini memiliki standar deviasi akurasi dan skor F1 yang relatif tinggi dibanding base classifier lainnya yaitu sebesar 2,9% dan 3%, dengan akurasi rata-rata 77%. Standar deviasi yang tinggi membuktikan bahwa model Regresi Logistik Multinomial cukup sensitif terhadap *overfitting*.

4.4.2. Random Forest

Random Forest memiliki nilai *true positif* (TP) kelas negatif, kelas netral dan kelas positif berturut – turut sebesar 516, 1049 dan 1024 prediksi benar. Nilai TP kelas netral pada *classifier* ini merupakan nilai terbesar dari seluruh *base classifiers*. Sehingga model Random Forest merupakan model yang paling baik dalam memprediksi kelas netral. Namun *classifier* ini memiliki nilai TP kelas negatif terkecil yaitu sebanyak 516.



Gambar 13 Matrik Konfusi Model Random Forest

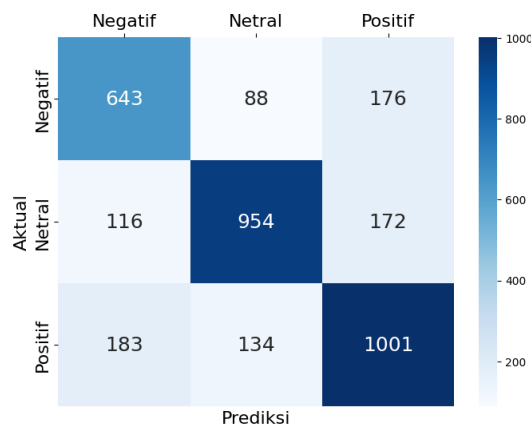


Gambar 14 Performa Model Random Forest pada setiap *Fold*

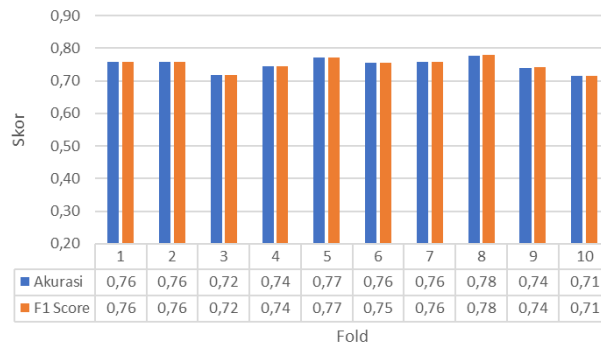
Dilihat dari performa setiap *fold* pada Gambar 14, *classifier* ini memiliki standar deviasi akurasi dan skor F1 paling tinggi yaitu masing – masing sebesar 3,3% dengan akurasi rata-rata 75%. Standar deviasi yang tinggi membuktikan bahwa model Random Forest kurang *robust* dan relatif sensitif terhadap *overfitting*.

4.4.3. Multinomial Naïve Bayes

Multinomial Naïve Bayes memiliki nilai *true positif* (TP) kelas negatif, kelas netral dan kelas positif berturut – turut sebesar 643, 954 dan 1001 prediksi benar. Nilai TP kelas positif pada *classifier* ini merupakan nilai *true positif* terbesar dari seluruh *base classifiers*. Sehingga model Multinomial Naïve Bayes merupakan model yang paling baik dalam memprediksi kelas negatif. Namun *classifier* ini memiliki nilai TP kelas netral dan kelas positif terkecil yaitu berturut – turut sebanyak 954 dan 1001.



Gambar 15 Matrik Konfusi Model Multinomial Naïve Bayes

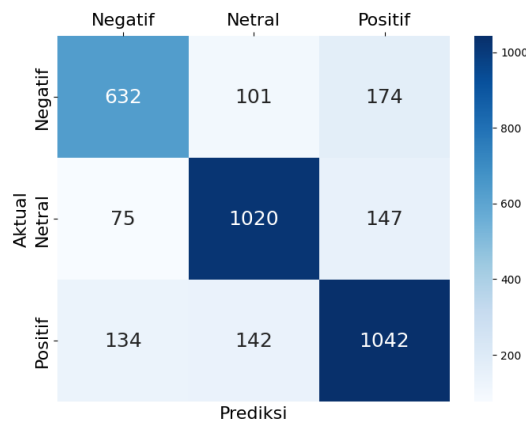


Gambar 16 Performa Model Multinomial Naive Bayes pada setiap *Fold*

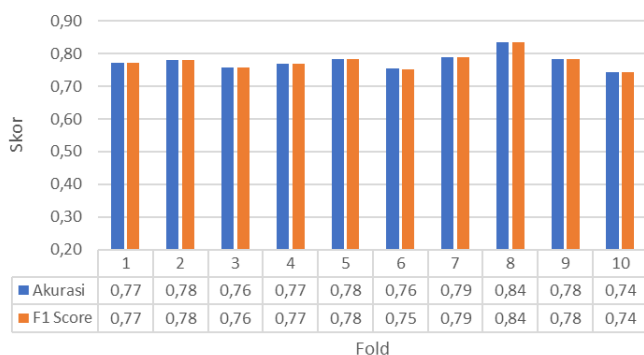
Dilihat dari performa setiap *fold* pada Gambar 16, *classifier* ini memiliki standar deviasi akurasi dan skor F1 paling rendah yaitu masing – masing sebesar 2,1% dengan akurasi rata-rata 75%. Standar deviasi yang rendah membuktikan bahwa model Multinomial Naïve Bayes merupakan model paling *robust* dibanding *classifiers* lainnya.

4.4.4. Ensemble

Model *Ensemble* memiliki nilai *true positif* (TP) kelas negatif, kelas netral dan kelas positif berturut – turut sebesar 632, 1020 dan 1042 prediksi benar. Nilai TP pada setiap kelas merupakan nilai *true positif* terbesar kedua dari seluruh *base classifiers*. Sehingga model *Ensemble* disimpulkan dapat memanfaatkan kelebihan dari masing – masing model dasar.



Gambar 17 Matrik Konfusi Model *Ensemble*

Gambar 18 Performa Model *Ensemble* pada setiap *Fold*

Adapun dilihat dari performa pada setiap *fold* dapat dilihat pada Gambar 18, model *Ensemble* memiliki nilai standar deviasi yang relatif rendah pada akurasi maupun skor F1 yaitu berturut – turut sebesar 2,5% dan 2,6% dengan akurasi rata-rata 78%. Hal ini membuktikan bahwa *Ensemble* cukup *robust* dalam menangani data yang bervariasi dan memiliki akurasi yang paling baik dibandingkan *classifier* lainnya. Dilihat dari nilai *true positif* pada setiap kelas dan pada setiap model, model *Ensemble* memberikan nilai *true positif* tertinggi ke – dua pada setiap kelas dapat dilihat pada Tabel 13.

Tabel 13 Perbandingan Nilai *True Positif* setiap Model

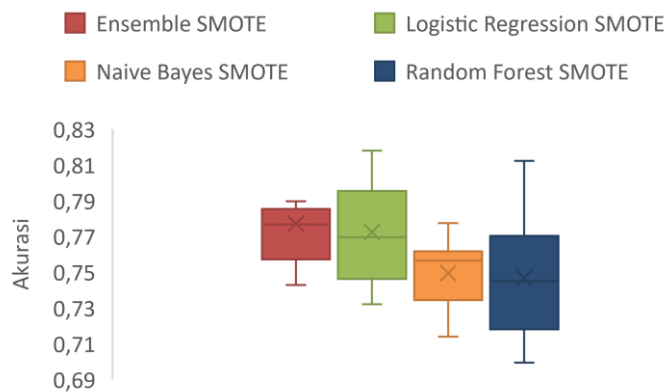
<i>True Positif</i> (TP)	Regresi Logistik Multinomial	Random Forest	Multinomial Naïve Bayes	<i>Ensemble</i>
Kelas Negatif	608	516	643	632
Kelas Netral	1012	1049	954	1020
Kelas Positif	1056	1024	1001	1042

Evaluasi lebih lanjut kemudian dilakukan dengan perhitungan akurasi, presisi, *recall*, dan Skor F1 berdasarkan hasil matrik konfusi yang telah dibuat. Hasil evaluasi dapat dilihat pada Tabel 14.

Tabel 14 Hasil Evaluasi Model *Ensemble*

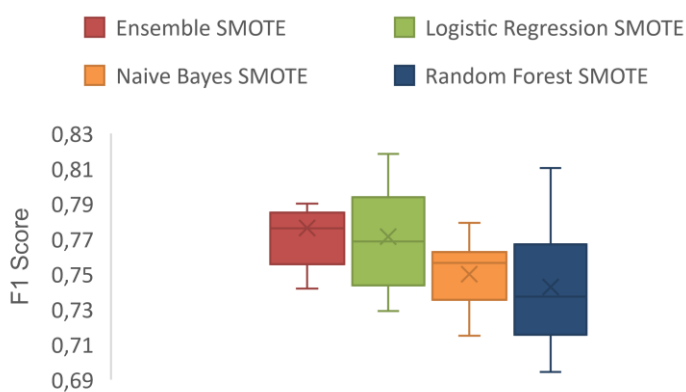
	Presisi	<i>Recall</i>	Skor F1	<i>Support</i>
Negatif	75%	69%	72%	907
Netral	81%	82%	81%	1242
Positif	76%	80%	78%	1318
Akurasi			78%	3467
<i>Macro Avg</i>	78%	77%	77%	3467
<i>Weighted Avg</i>	78%	78%	78%	3467

Model terbaik dapat dilihat berdasarkan skor Akurasi dan Skor F1 yang terbaik. Hasil perbandingan skor Akurasi dan Skor F1 Model *Ensemble* terhadap *base classifiers* dapat dilihat pada Gambar 19 dan Gambar 20.



Gambar 19 Perbandingan Performa Akurasi *Classifiers*

Model *Ensemble* memiliki rata – rata akurasi tertinggi dibandingkan model yang lain, yakni sebesar 78%. Jika dilihat dari standar deviasi akurasi, model *Ensemble* juga merupakan model yang cukup *robust* karena memiliki nilai standar deviasi yang kecil yakni sebesar 2,5%.



Gambar 20 Perbandingan Performa Skor F1 *Classifiers*

Model *Ensemble* memiliki rata – rata skor F1 tertinggi dibandingkan model yang lain, yakni sebesar 78%. Jika dilihat dari standar deviasi skor F1, model *Ensemble* juga merupakan model yang cukup *robust* karena memiliki nilai standar deviasi yang kecil yakni sebesar 2,6%.

4.5 Model Deployment

Tahapan selanjutnya adalah deploying model agar bisa dimanfaatkan dan digunakan orang lain. *Deploying* dapat dilakukan menggunakan bahasa Python dan *framework* Streamlit pada situs web online seperti *huggingface.co* yang relatif mudah dan tanpa biaya. Model *Ensemble* dapat diakses dengan membuka tautan <https://ipb.link/ensemble-model-stk25> dan tampilan hasil *deploying* dapat dilihat pada Lampiran 2. Orang lain dapat mencoba mengirimkannya *text* atau

file berisi *tweets* dan model akan memprediksi apakah teks/ *tweets* tersebut mengandung sentimen negatif, netral atau positif. Hasil dari prediksi beberapa *test case* dapat dilihat pada Lampiran 3.

@Hak cipta milik IPB University

IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

V SIMPULAN DAN SARAN

5.1 Simpulan

Hasil analisis sentimen dengan menggunakan *worldclouds* menunjukkan bahwa pada kelas negatif cenderung berisi kekecewaan pengguna terkait pengemudi biskita, waktu kedatangan, jalan pada rute biskita, kebijakan berbayar dan terkait halte Bubulak. Pada kelas netral cenderung berisi pertanyaan dan informasi. Informasi ini terkait dengan rute dan pembayaran., sedangkan kelas positif cenderung berisi antusiasme dan kepuasan pengguna terhadap layanan Biskita Transpakuan.

Model klasifikasi *Ensemble* dibuat menggunakan kombinasi *base classifier* Multinomial Naïve Bayes, *Random Forest*, dan Regresi Logistik yang telah dilakukan *hyperparameter tuning*. Model *Ensemble* dengan metode *weighted soft voting* mampu memanfaatkan kelebihan dari setiap *base classifiers* sehingga memperoleh *true positif* tertinggi ke - dua pada setiap kelas dan berhasil memprediksi sentimen dengan akurasi dan skor F1 terbaik sebesar 78%.

5.2 Saran

Saran untuk penelitian selanjutnya yaitu pada pelabelan dataset melibatkan pakar di bidang bahasa atau sentimen. Kontribusi pakar ini dapat meningkatkan konsistensi penilaian dan kualitas dataset yang digunakan dalam pemodelan. Kemudian, rentang waktu pengambilan data dapat diperluas kembali menyesuaikan peristiwa yang muncul terkait transportasi Biskita Transpakuan. Hal ini dapat berguna dalam menambah kekayaan dataset serta dapat menambah informasi yang dapat digali ketika melakukan analisis sentimen.

DAFTAR PUSTAKA

- Ali J, Khan R, Ahmad N, Maqsood I. 2012. Random forests and decision trees. *IJCSI Int. J. Comput. Sci. Issues*. 9(5):272–278.
- Aliyah Salsabila N, Ardhito Winatmoko Y, Akbar Septiandri A, Jamal A. 2019. Colloquial Indonesian Lexicon. *Proc. 2018 Int. Conf. Asian Lang. Process. IALP 2018*.:226–229. doi:10.1109/IALP.2018.8629151.
- Anastasia S, Budi I. 2017. Twitter sentiment analysis of online transportation service providers. *2016 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACISIS 2016*.:359–365. doi:10.1109/ICACISIS.2016.7872807.
- Ballabio D, Grisoni F, Todeschini R. 2018. Multivariate comparison of classification performance measures. *Chemom. Intell. Lab. Syst.* 174(December 2017):33–44. doi:10.1016/j.chemolab.2017.12.004.
- Caie PD, Dimitriou N, Arandjelović O. 2020. Precision medicine in digital pathology via image analysis and machine learning. *Artif. Intell. Deep Learn. Pathol.*(April 2023):149–173. doi:10.1016/B978-0-323-67538-3.00008-7.
- Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. 2002. SMOTE : Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*. 16 (2002):321–357.
- Cox DR. 1958. The Regression Analysis of Binary Sequences. *J. R. Stat. Soc. Ser. B*. 20(2):215–232. doi:10.1111/j.2517-6161.1958.tb00292.x.
- Dietterich TG. 2000. Ensemble methods in machine learning. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 1857 LNCS:1–15. doi:10.1007/3-540-45014-9_1.
- Gallager R. 2021. A Retrospective on His Life, Work, and Impact. *Ieee Trans. Inf. Theory*. 47(7):2685–2688.
- Han J, Kamber M, Pei J. 2012. Data Mining: Concepts and Techniques. 3rd editio. Massachusetts (MA): Morgan Kaufmann Publishers.
- Hastie T, Tibshirani R, Friedman J. 2008. The Elements of Statistical Learning. Ed ke-2. New York (NY): Springer.
- He H, Garcia EA. 2009. Learning from imbalanced data. *IEEE Trans. Knowl. Data Eng.* 21(9):1263–1284. doi:10.1109/TKDE.2008.239.
- Hosmer, D. W dan Lemeshow, S. 2000. *Applied Regresi Logistik Second Edition*. John Willey & Son, New York.
- Hosmer, D.W dan Lemeshow, S. 2013. *Applied Regresi Logistik Third Edition*. John Willey & Son, New York.
- Kim Hyunjoong, Kim Hyeuk, Moon H, Ahn H. 2011. A weight-adjusted voting algorithm for ensembles of classifiers. *J. Korean Stat. Soc.* 40(4):437–449. doi:10.1016/j.jkss.2011.03.002.
- Kumari S, Kumar D, Mittal M. 2021. An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier. *Int. J. Cogn.*

Comput. Eng. 2(November 2020):40–46. doi:10.1016/j.ijcce.2021.01.001.

Loilatu MJ, Irawan B, Salahudin S, Sihidi IT. 2021. Analysis of Twitter's Function as a Media communication of Public Transportation. *J. Komun.* 13(1):54. doi:10.24912/jk.v13i1.8707.

Ma Z, Jing F, Liang C, Hu X, Shi Y. 2011. An improved approach to terms weighting in text classification. *2011 Int. Conf. Comput. Manag. CAMAN 2011*.(January 2013). doi:10.1109/CAMAN.2011.5778755.

Medhat W, Hassan A, Korashy H. 2014. Sentiment analysis algorithms and applications: A survey. *Ain Shams Eng. J.* 5(4):1093–1113. doi:10.1016/j.asej.2014.04.011.

More AS, Rana DP. 2017. Review of random forest classification techniques to resolve data imbalance. *Proc. - 1st Int. Conf. Intell. Syst. Inf. Manag. ICISIM 2017*. 2017-Janua:72–78. doi:10.1109/ICISIM.2017.8122151.

Nurrun Muchammad Shiddieqy H, Paulus Insap S, Wing Wahyu W. 2016. Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter. *Semin. Nas. Teknol. Inf. dan Komun.* 7(2):57–64.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, *et al.* 2011. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 12(May 2014):2825–2830.

Ren Q, Li M, Han S. 2019. Tectonic discrimination of olivine in basalt using data mining techniques based on major elements: a comparative study from multiple perspectives. *Big Earth Data.* 3(1):8–25. doi:10.1080/20964471.2019.1572452.

Sartono B, Syafitri UD. 2010. Metode Pohon Gabungan: Solusi Pilihan untuk Mengatasi Kelemahan Pohon Regresi dan Klasifikasi Tunggal. *Forum Stat. dan Komputasi.* 15(1):1–7.

Saputra SA, Rosiyadi D, Gata W, Husain SM. 2019. Google play e-wallet sentiment analysis using naive bayes algorithm based on particle swarm optimization. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi).* 3(3):377–382.

Da Silva NFF, Hruschka Eduardo R., Hruschka Estevam R. 2014. Tweet sentiment analysis with classifier ensembles. *Decis. Support Syst.* 66:170–179. doi:10.1016/j.dss.2014.07.003.

Soedomo RP, Mulyono S, Falatehan AF. 2010. Pemilillian Kebijakan Mengatasi Kemacetan Lalu Lintas Di Kota Bogor (Studi Penerapan Analiti.Cal Hirarchi Process (AHP)). *War. Penelit. Perhub.* 22(2):134–149. doi:10.25104/warlit.v22i2.1031.

Tan PN, Steinbach M, Kumar V. 2006. Introduction to Data Mining. Boston: Pearson Addison Wesley.

Tharwat A. 2018. Classification assessment methods. *Appl. Comput. Informatics.* 17(1):168–192. doi:10.1016/j.aci.2018.08.003.

Wan Y, Gao Q. 2016. An Ensemble Sentiment Classification System of Twitter Data for Airline Services Analysis. *Proc. - 15th IEEE Int. Conf. Data Min.*

Work. *ICDMW 2015*.:1318–1325. doi:10.1109/ICDMW.2015.7.

Yang FJ. 2018. An implementation of naive bayes classifier. *Proc. - 2018 Int. Conf. Comput. Sci. Comput. Intell. CSCI 2018*.:301–306. doi:10.1109/CSCI46756.2018.00065.

@Hak cipta milik IPB University

IPB University



@Hak cipta milik IPB University

LAMPIRAN

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 1 Kamus Normalisasi Kata Gaul Tambahan Buatan Sendiri

No	Kata Sebelum	Kata Sesudah	No	Kata Sebelum	Kata Sesudah
1	kurangg	kurang	41	cpek	lelah
2	udaa	udah	42	anjrit	anjing
3	mnurut	menurut	43	anjig	anjing
4	anyinh	anjing	44	anjigg	anjing
5	seputat	seputar	45	anjingg	anjing
6	ijo	hijau	46	bukann	bukan
7	dmma	dimana	47	skrgg	sekarang
8	anjrot	anjing	48	makasihh	terimakasih
9	ajgg	anjing	49	asu	anjing
10	keboen	kebun	50	moga	semoga
11	aseekk	asik	51	cok	jancok
12	bliau	beliau	52	cokk	jancok
13	aseek	asik	53	cook	jancok
14	berpaa	berapa	54	cookk	jancok
15	berpa	berapa	55	amgkot	angkot
16	bggtt	banget	56	gua	aku
17	cntoh	contoh	57	gweh	aku
18	anzink	anjing	58	guah	aku
19	jrg	jarang	59	gw	aku
20	msi	masih	60	gwah	aku
21	anjirt	anjing	61	gue	aku
22	kesampeian	kesampaian	62	wkwkwk	wkwk
23	dtgnya	datangnya	63	dah	udah
24	dtg	datang	64	tkt	takut
25	dngin	dingin	65	umumm	umum
26	ktub	kutub	66	umuum	umum
27	brngkt	berangkat	67	yah	yah
28	antra	antara	68	drtd	daritadi
29	pinuh	penuh	69	drtdi	daritadi
30	anjink	anjing	70	ges	gais
31	anjir	anjing	71	gays	gais
32	ajg	anjing	72	geys	gais
33	smpet	sempat	73	trans pakuan	transpakuan
34	sempet	sempat	74	anjr	Anjir
35	makai	memakai	75	anjer	anjing
36	bgst	bangsat	76	anjer	anjing
37	anjg	anjing	77	njir	anjing
38	cpk	lelah	78	anjr	anjing
39	capek	lelah	79	trans pakuan	transpakuan
40	capk	lelah	80	gblk	goblok

@Hak cipta milik IPB University

IPB University

Lampiran 2 *Screenshot* Tampilan Hasil *Deploying* Model

Spaces

naufalnashif / sentiment-analysis-ensemble-model

like 1

Running

@Hakcipta milik IPB University

IPB University

Settings :

General Settings :

Pilih Model Sentimen:

Ensemble

Pilih metode input:

Teks Manual

Unggah Berkas XLSX

Unggah berkas XLSX

Drag and drop file here

Limit 200MB per file • XLSX

Browse files

test_tweets.xlsx
20.2KB

Pastikan berkas XLSX Anda memiliki kolom yang bernama **Text** (Maks.500 data).

Jika terdapat kolom type **datetime**, ganti nama kolom menjadi **Date**

Preference Settings:

Pilih Warna Wordclouds :

Blues

Tekan "Analysis" kembali jika tampilan menghilang

Analysis

Statistics and Data Science

Sentiment Analysis : Based on Tweets Biskita Transpakuan Bogor 2022-2023

Documentation Results Creator More

Results:

Wordclouds:

Bar Chart :

Hasil Analisis Sentimen

	e	label	emotikon
0	pang biskita turun setengahnya	Negatif	😞
1	ar ngelunjak banget ya bangsat menunggu datang cash pasan gabi	Positif	😄
2	ibales seabad ngolesin krim geliga kebanyakan rencana potong ram	Netral	😐
3	us subsidi ditjen hubdat memiliki koridor dilayani armada bus arm	Positif	😄
4	ok ditemuin trans pakuan duduk sebelahn pas ambil hp hp merk j	Positif	😄
5	ny high school heartbreaks duduk kursi transpakuan lihat motor ny	Positif	😄
6	hnya bogor naik biskita biaya alias gratis	Netral	😐
7	uk yuk buka trayek depok cilodong apa-apa	Netral	😐
8	kita tarif angkot naik	Positif	😄
9	gsi dibangun jalur trem pr nya biskita difokuskan membantu kota j	Netral	😐

Unduh CSV

Thank you for trying the demo!

Best regards, Naufal Nashif | © 2023

Lampiran 3 Hasil Prediksi *Tweets* Model *Ensemble*

N o	Tweet	Label Aktual	Prediksi LRM	Prediksi MNB	Prediksi RF	Prediksi Ensemble
1	@senjipr Betul, saya kerjaannya banyak bagian di biskita 🙏	Netral	Netral	Positif	Positif	Netral
2	@ipb_menfess naik biskita aja nder, lgsg turun di halte kebon raya nah dri situ udh jalan kaki bntar ada botani	Positif	Positif	Positif	Positif	Positif
3	sumpah gue di sb bakal naik biskita aja REAL karena enak bgt njir haltenya depan perumahan gue trs turunnya lgsg depan fakultas 🤗🤗🤗🤗 rehat nyetir cuy	Positif	Positif	Positif	Positif	Positif
4	[bgr] guys kalo dari whiz prime hotel Pajajaran, mau ke cipaku naik biskita nya yg mana sama darimana ya?	Netral	Netral	Netral	Netral	Netral
5	Supir 03/02 gak tau diri emang, kan halte jembatan emang buat yg mau naik biskita. Bisa bisanya supir biskita dimarahin njir, kagak tau diri emang supirnya dah. Trs supir biskita dibawa emosi gara gara supir 03, ini bawa nya ngebut parah 😭	Negatif	Negatif	Negatif	Negatif	Negatif
6	Biskita lama nih anjir datengnya, terminal uda penuh ama orang	Negatif	Negatif	Negatif	Negatif	Negatif
7	[bgr] transpakuan yang ke belanova masih ada ga yaaaa gaiss	Netral	Netral	Netral	Netral	Netral
8	Jadi tadi gue naik biskita trayek K1. Gue naik sama sepupu 2 orang, tapi yang punya emoney cuma gue aja. Sebelumnya gue udah pernah beberapa kali naik biskita dan pernah dikasih tau temen, 1 kartu maksimal boleh 3 orang. Karena pikir gue boleh2 aja, yaudah naik.	Positif	Positif	Positif	Positif	Positif
9	naik Biskita, gratis. dari deket St Bogor ke Cidangiang. K2. bismillah 15.00 https://t.co/QdahNwmGYq	Positif	Netral	Netral	Positif	Netral
10	biskita tuh masih gratis apa bayar si	Netral	Netral	Netral	Netral	Netral

@Hak cipta milik IPB University

IPB University

RIWAYAT HIDUP

Penulis dilahirkan di Cilacap pada tanggal 4 Mei 2001 sebagai anak pertama dari pasangan bapak Purwono dan ibu Amanatul Kirom. Pendidikan sekolah menengah atas (SMA) ditempuh di sekolah SMA Negeri 1 Kroya, dan lulus pada tahun 2019. Pada tahun 2019, penulis diterima sebagai mahasiswa program sarjana (S-1) di Program Studi Statistika dan Sains Data, Fakultas Matematika dan Ilmu Pengetahuan Alam di IPB.

Selama mengikuti program S-1, penulis aktif menjadi staff Media dan Branding di Gamma Sigma Beta (GSB), Anggota UKM Forces Angkatan 17, dan Kepala Divisi Media dan Branding Kompetisi Statistika Junior 2022. Penulis juga pernah mengikuti magang di PT Xquisite Analitika Indonesia, serta Magang dan studi Independen Bersertifikat di Ruangguru. Selain itu penulis juga pernah mengikuti lomba Infografis dan terpilih sebagai finalis di Universitas Syiah Kuala tahun 2020.