

Analisis Ulasan Pengguna Aplikasi Diagnosa Tanaman Di Play Store Menggunakan Naïve Bayes

Hafiz Irsyad, Akhsani Taqwiyum

1Program Studi Informatika, Universitas Multi Data Palembang, Palembang, Indonesia
Email: 1hafizirsyad@mdp.ac.id , 2 akhsani.taqwiyum@mdp.ac.id

Abstrak– Mengenal penyakit tanaman butuh literasi yang sangat dalam sehingga para petani pemula merasa enggan untuk mendalami bidang pertanian. Pertanian 4.0 telah diterapkan di beberapa negara, sehingga para pemula untuk bertani tidak perlu khawatir lagi dengan adanya pertanian 4.0. Aplikasi diagnosa penyakit tanaman dapat diunduh pada Google Play Store dan banyak ulasan komentar dari para pengguna. Dengan begitu banyak ulasan dari komentar yang ada, sehingga menjadi sulit untuk diolah secara manual, walaupun ada rating. Pada umumnya Rating belum tentu sesuai dengan isi ulasan pengguna. Maka dari itu, diperlukan untuk mengolah hasil ulasan pengguna serta dapat melihat kecenderungan pengguna terhadap aplikasi tersebut. Metode yang digunakan adalah Naïve Bayes. Untuk pelabelan data memerlukan pakar Bahasa Indonesia dilabelkan secara manual berdasarkan keilmuan Bahasa Indonesia dan KBBI. Pelabelan adalah Positif, negatif dan netral. Dataset didapatkan sebanyak 252 ulasan. Dari rata-rata pengujian mendapatkan nilai akurasi 79%. Sementara untuk nilai Precision sentiment positif 100%, Sentimen netral 76% dan untuk sentiment negatif 77%. Untuk nilai Recall Sentimen positif bernilai 37%, Sentimen netral 100% dan untuk sentiment negatif 100%. Dan untuk F1-Score sendiri memiliki Sentimen positif bernilai 56%, Sentimen netral 85% dan untuk sentiment negatif 87%.

Kata Kunci: Pertanian 4.0, Google Play Store, Naïve Bayes

Abstract– Recognizing plant diseases requires very deep literacy so novice farmers feel reluctant to study agriculture. Agriculture 4.0 has been implemented in several countries, so beginners to farming don't need to worry anymore about agriculture 4.0. The plant disease diagnostic application can be downloaded on the Google Play Store and many reviews and comments from users. With so many reviews from existing comments, it becomes difficult to process them manually, even though there are ratings. In general, ratings are not necessarily in accordance with the contents of user reviews. Therefore, it is necessary to process the results of user reviews and be able to see user tendencies towards the application. The method used is Naïve Bayes. For data labeling, an Indonesian language expert is required who is labeled manually based on Indonesian knowledge and KBBI. Labeling is Positive, Negative and Neutral. The dataset obtained as many as 252 reviews. From the average test, it gets an accuracy value of 79%. Meanwhile, the Precision value is 100% positive sentiment, 76% for neutral sentiment and 77% for negative sentiment. The Recall value for positive sentiment is 37%, neutral sentiment is 100% and for negative sentiment is 100%. And the F1-Score itself has a positive sentiment of 56%, a neutral sentiment of 85% and a negative sentiment of 87%.

Keywords: Agriculture 4.0, Google Play Store, Naïve Bayes

1. PENDAHULUAN

Pertanian 4.0 adalah berasal dari “industri 4.0”, yang mana mengarahkan dan mengkhususkan berpadu bersama Teknologi Informasi. Ketika Pertanian 4.0 sudah digaungkan, secara bersamaan bermunculan inovasi-inovasi baru terkait dengan pertanian dan teknologi informasi. Salah satu inovasi dalam bidang pertanian yang mendukung system produksi tanaman untuk pelaku pertanian secara umum adalah mengetahui jenis-jenis penyakit terhadap tanaman dan menangani permasalahan terhadap penyakit tanaman.

Penyakit dan hama serangga adalah sebagian besar masalah yang dihadapi dalam pertanian. Tetapi banyak stresor abiotik lain yang dipelajari, seperti kekurangan nutrisi, toksisitas, faktor lingkungan seperti angin dan hujan es, dan masih banyak lagi. Kemampuan untuk membedakan antara cekaman biotik dan abiotik penting dan dapat menghemat aplikasi pestisida yang tidak diperlukan [1].

Sekarang dengan berkembangnya Pertanian 4.0 para petani tidak perlu khawatir dan kesulitan lagi untuk mendeteksi hama penyakit yang menyerang tanaman mereka. Cukup dengan foto tanaman yang terserang penyakit, kemudian mengunggahnya di aplikasi yang dapat di download melalui google play. Kemudian secara otomatis aplikasi tersebut akan membaca foto tersebut dan memberikan informasi mengenai hama atau penyakit yang menyerang tanaman tersebut serta memberikan solusi cara penanganannya [2].

Aplikasi Plantix dan Agrio adalah aplikasi yang membantu para petani untuk mengelola perlindungan tanaman dengan cara yang lebih baik. Pada aplikasi ini juga dapat melakukan identifikasi penyakit, sehingga dapat membantu para Petani atau Petani rumahan yang mencari solusi dari permasalahan terhadap tanaman.

Analisis sentimen adalah proses penggalian opini, tindakan, dan emosi dari dokumen teks untuk memperoleh informasi [3].

Analisis sentimen mempunyai banyak keuntungan. Salah satunya adalah untuk mengetahui apakah pengguna merespon positif aplikasi pendeteksi penyakit tanaman ini. Dapat digunakan untuk pengembangan Pertanian 4.0 masa depan. Karena manfaatnya, banyak penelitian yang menggunakan sentimen analisis, seperti perusahaan para penyedia

aplikasi pada Google Play Store yang berguna untuk mengetahui kualitas dari aplikasi dengan adanya data ulasan dari pengguna pada aplikasi yang digunakan [4].

Pada Penelitian yang dilakukan oleh Dita, Firda, Fita, Belasana & Muhadi (2020) mengenai Analisa Sentimen Terhadap Review Produk Kecantikan Menggunakan Metode Naive Bayes Classifier. Data berjumlah 220 ulasan dengan perbandingan rasio menjadi 110 ulasan dengan sentimen negatif dan 110 sentimen positif. Algoritma naive bayes classifier mendapatkan 110 data terklasifikasi negatif dan 85 data dinyatakan negatif seperti yang diharapkan. Juga 25 hasilnya negatif, Jadi jika ada 110 data positif, 92 akan diklasifikasikan menurut prediksi, dan 18 data diperkirakan negatif, tetapi hasilnya positif. Didapatkan hasil perhitungan manual yang dilakukan diperoleh akurasi yaitu 80,45% [5].

Pada Penelitian yang dilakukan oleh Rita Apriani & Dudih Gustian mengenai Analisis Sentimen Dengan Naive Bayes Terhadap Komentar Aplikasi Tokopedia. Data pelatihan yang dipilih secara acak hingga 1.000 komentar, kemudian 500 komentar positif dan 500 komentar negatif secara manual. Untuk data pengujian, 1.500 komentar terbaru dipilih dengan mengumpulkan data yang sama seperti pelatihan. Hasil pengujian untuk 1.500 data uji menunjukkan nilai akurasi sebesar 97,13%, sedangkan nilai recall kelas sebesar 95,49%. Di sisi lain, ketika nilai AUC adalah 0,980. Data komentar sentimen negatif diperoleh sebesar 63,53% dan positif 36,47% [6].

Berdasarkan uraian terkait penelitian terdahulu, algoritma Naive Bayes terbukti cukup baik dalam kasus analisis sentimen, sehingga pada penelitian ini akan digunakan juga algoritma Naive Bayes untuk mengetahui sentimen masyarakat terhadap sebuah masalah tentang banyaknya aplikasi investasi yang ada pada Google Play Store. Penelitian ini memberikan solusi untuk menentukan persentase komentar dan umpan balik positif atau negatif dari pengguna aplikasi berdasarkan umpan balik pengguna pada aplikasi identifikasi penyakit tanaman, dimana persentase dan feedback akan meningkatkan kepercayaan pengguna menggunakan aplikasi dan memberikan informasi untuk meningkatkan kualitas layanan, kualitas produk dan kualitas internal perusahaan. Data yang diambil dalam penelitian ini selama 4 (empat) bulan dimulai Agustus 2022 sampai November 2022. Dan data reviewnya pada aplikasi yang akan diklasifikasikan adalah teks bahasa Indonesia yang meliputi bahasa formal dan non-formal

2. METODOLOGI PENELITIAN

2.1 Text Mining

Penambangan teks merupakan penggalan informasi yang bermanfaat dari sejumlah sekumpulan dokumen dari suatu masa dengan mengidentifikasi ciri khas yang menarik [7].

2.2 Naive Bayes

Penerapan Naive Bayes ke analisis sentimen mempunyai dua tahap penting, yaitu tahap pelatihan dan pengujian. Naive Bayes merupakan sebuah metode pengklasifikasian dengan menggunakan probabilitas sederhana yang berakar pada Teorema Bayes dan memiliki asumsi ketidaktergantungan (independent) yang tinggi dari masing - masing kondisi atau kejadian. Probabilitas dihitung menggunakan persamaan (1) [8].

$$P(c|x) = \frac{P(x|c)}{P(x)} \cdot P(c) \dots (1)$$

Keterangan :

x :Data dengan class yang belum diketahui

c :Hipotesis data merupakan suatu class spesifik P(c|x) :Probabilitas hipotesis berdasar kondisi P(c) :Probabilitas hipotesis

P(x|c) :Probabilitas berdasarkan kondisi pada hipotesis P(x) :Probabilitas x

2.3 K-Fold Cross Validation

K-Fold Cross Validation adalah metode yang digunakan untuk memprediksi kesalahan dalam mengevaluasi sebuah model. Data dibagi menjadi himpunan bagian K berjumlah hampir sama. Di setiap pengulangan, salah satu himpunan bagian akan digunakan sebagai data training dan data testing .

Metode K-Fold Cross Validation membagi dataset menjadi K subset data (D1,D2,...,Dk) dengan jumlah data setiap subset dengan sama rata. Pada iterasi pertama, subset data D1 digunakan sebagai data uji, sedangkan sisa subset lainnya (D2,...,Dk) akan digunakan sebagai data latih dan seterusnya sampai iterasi ke-K. Untuk proses Proses cross validation dapat dilihat pada gambar 1 [9].



Gambar 1. Proses Cross Validation

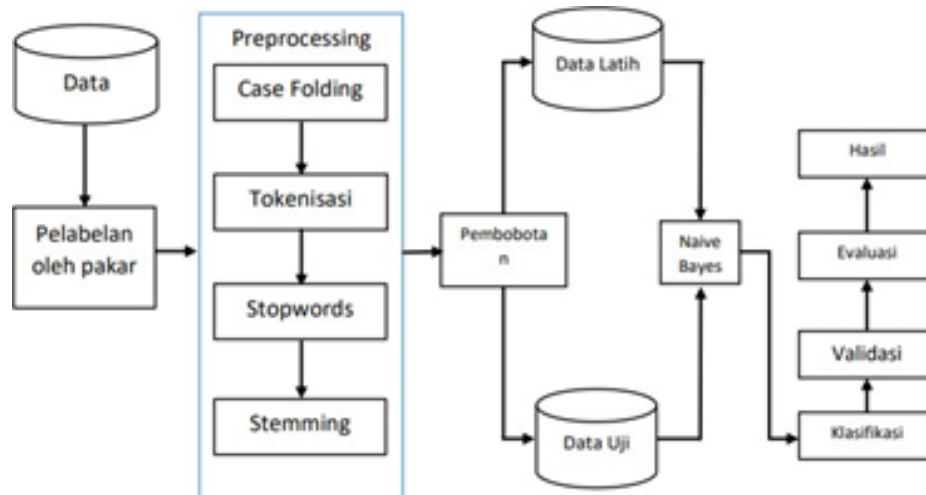
1. METODOLOGI PENELITIAN

3.1 Tahapan Penelitian



Gambar 2. Desain Penelitian

1. Identifikasi Masalah
 Pada tahap ini dilakukan identifikasi masalah mengenai analisis sentimen pada ulasan aplikasi Plantix dan Agrio di Google Play Store menggunakan metode Naive Bayes.
2. Studi Literatur
 Pada tahap ini dilakukan pembelajaran literatur berupa jurnal dan buku yang berkaitan dengan topik ini agar mendapatkan informasi dan data yang bisa digunakan untuk penelitian.
3. Pengumpulan DataSet
 Pada pengumpulan data ulasan dilakukan dengan mengambil data dari link Google Play Store aplikasi Plantix dan Agrio menggunakan tools Google Colab dengan bahasa pemrograman python. Proses pengambilan data diambil dari 1 Agustus 2022 sampai 17 November 2022 pada ulasan aplikasi Plantix dan Agrio. Dengan total data yang didapatkan sebanyak 252 yang akan digunakan dalam analisis sentimen menggunakan metode Naive Bayes.
4. Pelabelan dataset
 Pada tahap pelabelan dataset, pelabelan data menggunakan kuesioner yang diberikan kepada Ahli Bahasa Indonesia. Sehingga data diambil pada ulasan di Google Play Store dan akan dilabeli untuk membedakan sentimen positif, dengan lambang '1', sentiment netral dengan lambang '0', dan sentimen negatif dengan lambang '-1'.
5. Perancangan Sistem
 Pada tahap ini dilakukan perancangan sistem mulai dari pengumpulan data, data yang telah dikumpulkan kemudian diberi label positif, label netral dan label negative kemudian dilakukan preprocessing. Perbandingan untuk rasio data dibagi menjadi 70% data latih dan 30% data uji, kemudian dilakukan Pembobotan TF-IDF, dimana proses perubahan kata menjadi bentuk angka atau vector, sedangkan TF-IDF digunakan untuk menentukan frekuensi sebuah kata didalam sebuah dokumen dan pemberian bobot pada setiap kata pada setiap dokumen untuk mencari dan menghitung berapa kali kata itu muncul. Untuk klasifikasi menggunakan Algoritma Naive Bayes. Perancangan Sistem dapat dilihat pada Gambar 3.



Gambar 3. Perancangan Sistem

6. Evaluasi

Tahap Tahap ini melakukan evaluasi terhadap sistem yang telah dirancang. Tujuan dari tahap ini adalah memastikan kebenaran hasil yang telah didapat dan meminimalisir kesalahan. Pada tahap evaluasi ini digunakan metode Confusion Matrix yang akan menghitung recall, precision dan accuracy. Perhitungan Confusion Matrix dapat dilakukan dengan persamaan berikut ini [7].

$$\text{Recall} = TP / TP + FN \times 100\% \quad \text{Precision} = TP / TP + FP \times 100\%$$

$$\text{Accuracy} = TP + TN / TP + FN + FP + FN \times 100\%.$$

- Keterangan :
- TP : Jumlah data positif jenis sentimen yang terklasifikasi dengan benar oleh sistem.
 - TN : Jumlah data negatif jenis sentimen yang terklasifikasi dengan benar oleh sistem.
 - FN : Jumlah data negatif jenis sentimen yang terklasifikasi dengan salah oleh sistem.
 - FP : Jumlah data positif jenis sentimen yang terklasifikasi dengan salah oleh sistem

2. HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini berupa ulasan pengguna aplikasi diagnosa penyakit tanaman yang diambil melalui dua aplikasi di Google Play Store yaitu, Plantix dan Agrio. Proses pengumpulan data dilakukan dengan scraping. Dari hasil scraping terkumpul data mentah sebanyak Agrio 71 ulasan komentar terbaru dan Plantix 181 ulasan komentar terbaru.

	reviewId	userName	userImage	content	score
0	00c51986-083e-4328-921c-6166b659b2cf	Lisardo Sitorus	lh.googleusercontent.com/a/ALm5wu...	Aplikasi ini untuk saya tidak berguna, sementara...	1
1	adf630e2-0952-4dde-8d67-664904581b2a	Moh Rifan F M	lh.googleusercontent.com/a/ALm5wu...	Ga lengkap	1
2	0be19233-35de-41f5-9679-9eee83002af8	An To	lh.googleusercontent.com/a/ALm5wu...	Admin tolong ditambah untuk identifikasi tanam...	5
3	bd93ebf2-7e64-470b-be10-cfebbe6ccb1b	Seven Sinukaban	lh.googleusercontent.com/a/-ACNPE...	So recommended	3
4	54e9fc53-041d-404a-a4c9-4c06d77ac119	Rama Ctb	lh.googleusercontent.com/a/ALm5wu...	okkkkk	5

Gambar 4. Data Scraping

Sebelum diklasifikasikan, data tersebut harus melalui beberapa tahapan proses agar bisa diklasifikasikan dalam proses selanjutnya, berikut adalah tahapan prosesnya:

1. Pelabelan Data

Pada tahap pelabelan dataset, pelabelan data menggunakan kuesioner yang diberikan kepada pakar Bahasa Indonesia. Sehingga data diambil pada ulasan di Google Play Store dan akan dilabeli untuk membedakan

sentimen positif, dengan lambang ‘1’, sentiment netral dengan lambang ‘0’, dan sentimen negatif dengan lambang ‘-1’. Pelabelan ini diberikan kepada pakar dalam bentuk kuesioner. Hasil dari pelabelan dapat dilihat pada Tabel 1

Tabel 1. Pelabelan Data

Label	Review
-1	Aplikasi ini untuk saya tidak berguna, sementara...
0	Admin tolong ditambah untuk identifikasi tanam...
1	So recommended
-1	Ga lengkap
0	okkkkk

2. Pre Processing

Case Folding adalah tahapan pertama dari proses Preprocessing data. Dalam tahap ini dilakukan perubahan huruf pada data teks Google Play Store dari huruf kapital menjadi huruf kecil. Adapun hasil dari data yang telah di labeli oleh pakar, proses ini dapat dilihat pada Tabel 2.

Tabel 2. Case Folding

Sebelum Case Folding	Sesudah Case Folding
ingin Mempelajari dan Melindungi Tanaman dari serangan Hama..	ingin mempelajari dan melindungi tanaman dari serangan hama

Setelah data teks melalui tahapan Case Folding maka dilakukan tahapan tokenisasi. Tokenisasi dilakukan untuk memisahkan antar kata pada kalimat dan menghilangkan tanda baca, angka dan emoticon yang terdapat pada teks. Hasil yang didapat dari proses tokenisasi yang dimana data inputan yang digunakan untuk proses ini adalah data yang berasal dari proses Case Folding, hasilnya dapat dilihat pada Tabel 3.

Tabel 3. Tokenisasi

Sebelum	Sesudah
ingin Mempelajari dan Melindungi Tanaman dari serangan Hama..	‘ingin’, ‘mempelajari’, ‘dan’, ‘melindungi’, ‘tanaman’, ‘dari’, ‘serangan’, ‘hama’

Stopword adalah proses yang dilakukan setelah proses Case Folding dan Tokenisasi, digunakan untuk menghilangkan kata-kata yang sering muncul dan juga tidak dibutuhkan dalam proses analisis. Pada proses Stopword dilakukan dengan bantuan modul NLTK yang disediakan dalam bahasa pemrograman Python. Hasil yang didapat dalam tahap Stopword ini dapat dilihat pada Tabel 4.

Tabel 4. Stopword

Sebelum	Sesudah
ingin Mempelajari dan Melindungi Tanaman dari serangan Hama..	‘mempelajari’, ‘melindungi’, ‘tanaman’, ‘serangan’, ‘hama’

Tahapan terakhir pada Preprocessing adalah Stemming. Memiliki fungsi untuk mengembalikan atau mengubah kata yang mengandung imbuhan menjadi kata dasar. Pada penelitian ini, proses Stemming dilakukan dengan menggunakan modul Sastrawi yang disediakan dalam bahasa pemrograman Python. Hasil dari Stemming dapat dilihat pada Tabel 5.

Tabel 5. Stemming

Sebelum	Sesudah
'ingin', 'mempelajari', 'melindungi', 'tanaman', 'serangan', 'hama'	'pelajari', 'lindung', 'tanam', 'serang', 'hama'

3. Pembobotan

Pembobotan TF-IDF adalah tahap yang dilakukan untuk mengetahui hubungan antar kalimat pada sekumpulan dokumen teks. Pada penelitian ini, untuk proses Pembobotan TF-IDF dilakukan dengan modul Sklearn yang disediakan dalam bahasa pemrograman Python, dimana hasil yang diperoleh dari modul Sklearn. Hasil TF-IDF dapat dilihat pada Tabel 6.

Tabel 6. TF-IDF

	about	admin	ajar	akurat	amazing
0	0	0	0	0	0
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0,47	0	0
4	0	0	0	0	0
5	0	0	0	0	0
6	0	0	0	0	0
7	0	0,67	0	0	0
8	0	0	0	0	0
9	0	0	0	0	0
10	0	0	0	0	0
11	0	0	0	0	0
12	0	0	0	0	0
13	0	0	0	0	0
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	0	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	0	0	0
26	0	0	0	0	0
27	0	0	0	0	0
28	0	0	0	0	0
29	0	0	0	0	0
30	0	0	0	0	0
31	0	0	0	0	0
32	0	0	0	0	0
33	0	0	0	0	0
34	0	0	0	0	0
35	0	0	0	0	0
36	0	0	0	0	0
37	0	0	0	0	0
38	0	0	0	0	0
39	0	0	0	0	0
40	0	0	0	0	0
41	0	0	0	0	0
42	0	0	0	0	0
43	0	0	0	0	0
44	0	0	0	0	0
45	0	0	0	0	0
46	0	0	0	0	0
47	0	0	0	0	0
48	0	0	0	0	0
49	0	0	0	0	0
50	0	0	0	0	0
51	0	0	0	0	0
52	0	0	0	0	0
53	0	0	0	0	0
54	0	0	0	0	0
55	0	0	0	0	0
56	0	0	0	0	0
57	0	0	0	0	0
58	0	0	0	0	0
59	0	0	0	0	0
60	0	0	0	0	0
61	0	0	0	0	0
62	0	0	0	0	0
63	0	0	0	0	0
64	0	0	0	0	0
65	0	0	0	0	0
66	0	0	0	0	0
67	0	0	0	0	0
68	0	0	0	0	0
69	0	0	0	0	0
70	0	0	0	0	0
71	0	0	0	0	0
72	0	0	0	0	0
73	0	0	0	0	0
74	0	0	0	0	0
75	0	0	0	0	0
76	0	0	0	0	0
77	0	0	0	0	0
78	0	0	0	0	0
79	0	0	0	0	0
80	0	0	0	0	0
81	0	0	0	0	0
82	0	0	0	0	0
83	0	0	0	0	0
84	0	0	0	0	0
85	0	0	0	0	0
86	0	0	0	0	0
87	0	0	0	0	0
88	0	0	0	0	0
89	0	0	0	0	0
90	0	0	0	0	0
91	0	0	0	0	0
92	0	0	0	0	0
93	0	0	0	0	0
94	0	0	0	0	0
95	0	0	0	0	0
96	0	0	0	0	0
97	0	0	0	0	0
98	0	0	0	0	0
99	0	0	0	0	0
100	0	0	0	0	0

4. Visualisasi

Word Cloud merupakan modul yang disediakan oleh bahasa pemrograman Python yang dapat digunakan untuk memvisualisasikan teks. Teks dengan frekuensi yang berbeda-beda digambarkan dengan ukuran dan warna yang berbeda. Dari hasil pengumpulan data ulasan aplikasi Plantix dan Agrio pada Google Play Store, kata-kata yang muncul terkait dengan review ulasan tersebut ditampilkan pada Gambar 6.



Gambar 5. Data Scraping

Dari gambar di atas dapat diketahui dari 252 Ulasan yang digunakan pada penelitian dan menunjukkan kata yang paling banyak muncul adalah “tanam” sebanyak 75 kata, “bantu” sebanyak 25 kata, “tani” sebanyak 20 kata, dan “bagus” sebanyak 18 kata.

5. Naive bayes
 Pada tahap klasifikasi data, data yang dimaksud adalah data yang telah melewati proses Preprocessing dan Pembobotan TF-IDF dibagi menjadi 2 yaitu data latih dan data uji, dimana perbandingan rasio data sebesar 70% untuk data latih dan 30% untuk data uji.
6. Evaluasi
 Dalam tahapan validasi hasil yang didapatkan digunakan metode K-Fold Cross Validation dengan nilai $k = 5$. Adapun hasil dari pengujian akan mendapatkan nilai accuracy, precision, recall dan F1-Score. Hasil dari validasi tersebut dapat ditampilkan pada tabel 6.

Tabel 6. Hasil Validasi

Tabel 6. Hasil Validasi

Fold	Acc	Precision (%)			Recall (%)			F1-Score (%)		
		-	0	+	-	0	+	-	0	+
1	0.82	0.81	0.7	1.0	1.	1.0	0.	0.8	0.	0.6
			6	0	0	0	48	9	86	5
2	0.77	0.74	0.7	1.0	1.	1.0	0.	0.8	0.	0.4
			5	0	0	0	30	5	86	6
3	0.82	0.89	0.7	1.0	1.	1.0	0.	0.9	0.	0.4
			2	0	0	0	31	4	84	8
4	0.78	0.69	0.8	1.0	1.	1.0	0.	0.8	0.	0.5
			1	0	0	0	41	2	89	8
5	0.78	0.75	0.7	1.0	1.	1.0	0.	0.8	0.	0.6
			3	0	0	0	46	6	84	3
AVG	0,79	0,77	0,7	1	1	1	0,	0,8	0,	0,5
			6				37	7	85	6

Pada skenario pertama diuji dengan menggunakan 3 (tiga) label yaitu, positif, negatif dan netral. Data pada tabel 6 diatas dapat menunjukkan hasil yang dilakukan oleh algoritma Naive Bayes berjalan dengan baik, dimana memperoleh nilai sebesar 82%. Untuk masing-masing kelas sentimen pada tabel 6, algoritma naive bayes dapat mengetahui dan melakukan klasifikasi berdasarkan sentimen positif cukup baik, yang terdapat pada model sentimen positif pada bagian F1-Score sebesar 65%. Untuk pengenalan kelas yang kurang baik dapat dilihat pada Recall sebesar 48%, dan untuk membedakan dengan kelas lainnya sangat baik ditukukan pada nilai precisióń sebesar 100%. Untuk model sentimen negatif F1-Score memperoleh nilai sebesar 89% yang berarti sangat baik mengenali kelasnya. Dan untuk sentimen netral nilai yang paling rendah adalah 76%.

Pada skenario kedua memperoleh nilai akurasi kurang baik dari pada skenario pertama, yaitu 77%. Untuk sentimen positif memiliki nilai tertinggi pada precisióń sebesar 100%, yang artinya dapat mengenali kelasnya dengan baik. Untuk sentimen negatif nilai tertinggi terdapat pada recall sebesar 100% dan untuk sentimen netral memperoleh nilai terendah 75%.

Pada skenario ketiga memperoleh nilai akurasi sangat baik, yaitu 82%. Dimana masing-masing kelas dapat mengklasifikasi dengan baik. Untuk sentimen positif nilai tertinggi terdapat pada precisióń 100%. Sementara sentimen netral memperoleh nilai terendah pada 81%.

Pada skenario keempat hasil akurasi sebesar 78%. Untuk pengenalan sentimen positif tertinggi terdapat pada bagian precisióń sebesar 100%. Sementara hasil negatif paling tinggi adalah pada bagian 100% pada bagian recall. Dan untuk mengenali sentimen netral nilai terendah adalah 81% pada bagian precisióń.

Pada skenario kelima memperoleh akurasi nilai sebesar 78%. Untuk kelas tertinggi mengenali sentimen positif adalah precisióń dengan nilai 100%. Sementara untuk kelas sentimen negatif memperoleh nilai tertinggi pada recall sebesar 100%. Dan untuk sentimen netral memperoleh nilai recall sebesar 100%. Dari semua percobaan tersebut mendapatkan nilai rata-rata cukup baik yaitu 79%. Sementara itu rata-rata terendah terdapat pada Recall dengan nilai 37% terhadap sentimen positif. Untuk rata-rata kelas tertinggi pada sentimen negatif adalah 87 pada recall. Dan untuk rata-rata kelas tertinggi pada sentimen netral adalah recall sebesar 100%

4. KESIMPULAN

Rata-rata nilai akurasi dari plantix dan agrio dapat diperoleh sebesar 79%. dimana artinya algoritma Naive Bayes berjalan dengan baik melakukan klasifikasi terhadap masing-masing kelas sentimen positif, netral dan negatif. Faktor dalam memperoleh nilai akurasi rata-rata sebesar 79% adalah pada tahapan pelabelan yang sangat teliti. Yang mana pada pelabelan data dilakukan terhadap pakar bahasa indonesia sekaligus didampingi dengan KBBI (Kamus Besar Bahasa Indonesia). Maka dengan hasil pada penelitian ini para warganet yang awam tentang diagnose penyakit tanaman ini bisa menggunakan aplikasi Plantix dan Agrio sebagai sarana pendukung untuk memulai pertanian 4.0..

UCAPAN TERIMAKASIH

Terima kasih disampaikan kepada pihak-pihak yang telah mendukung terlaksananya penelitian ini.

REFERENCES

- [1] Agrio, "Aplikasi untuk mengidentifikasi penyakit dan hama tanaman," 17 11 2022. [Online]. Available: <https://agrio.app/Aplikasi-untuk-mengidentifikasi-penyakit-dan-hama-tanaman/>.
- [2] R. A.F, "Canggih, Aplikasi Plantix Pendeteksi Opt," 11 August 2021. [Online]. Available: <https://jagadtani.com/read/2295/canggih-aplikasi-plantix-pendeteksi-opt>. [Accessed 19 November 2022].
- [3] D. A. Kristiyanti, "ANALISIS SENTIMEN REVIEW PRODUK KOSMETIK MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN PARTICLE SWARM OPTIMIZATION SEBAGAI METODE SELEKSI FITUR," Seminar Nasional Inovasi dan Tren (SNIT) 2015, pp. 134-141, 2015.
- [4] A. Rahman and E. Utami, "Sentimen Analisis Terhadap Aplikasi pada Google Playstore Menggunakan Algoritma Naïve Bayes dan Algoritma Genetika," Jurnal Komtika (Komputasi dan Informatika), pp. 60-71, 2021.
- [5] D. N. Sari, F. Adelia, F. Rosdiana, B. B. Butar and M. Hariyanto, "NALISA SENTIMEN TERHADAP REVIEWPRODUK KECANTIKAN MENGGUNAKAN METODE NAIVE BAYESCLASSIFIER," JIKA (Jurnal Informatika) Universitas Muhammadiyah Tangerang , pp. 109-118, 2020.
- [6] R. Apriani and D. Gustian, "ANALISIS SENTIMEN DENGAN NAÏVE BAYES TERHADAP KOMENTAR APLIKASI TOKOPEDIA," Jurnal Rekayasa Teknologi Nusa Putra, pp. 54-62, 2019.
- [7] J. S. R. Feldman, The Text Mining Handbook, New York: Cambridge University Press, 2007.
- [8] S. D. H. P. G. D. N. C. Mahesh Kini M, "Text Mining Approach to Classify Technical," International Journal of Advanced Research in Computer and Communication Engineering, vol. 4, no. 7, pp. 386-391, 2015.



- [9] D. K. N. S. Linda Mardiana, "ANALISIS DISKRIMINAN DENGAN K FOLD CROSS VALIDATION UNTUK," Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster), vol. 11, no. 1, pp. 97-102, 2022.
- [10] F. Ratnawati, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter," JURNAL INOVTEK POLBENG - SERI INFORMATIKA, pp. 51-59, 2018.
- [11] A. Karim, S. Esabella, M. Hidayatullah, and T. Andriani, "Sistem Pendukung Keputusan Aplikasi Bantu Pembelajaran Matematika Menggunakan Metode EDAS," vol. 4, no. 3, 2022, doi: 10.47065/bits.v4i3.2494.
- [1] M. Bobbi, K. Nasution, S. Suryadi, and A. Karim, "Sistem Pendukung Keputusan Dalam Rekomendasi Kelayakan nasabah Penerima Kredit Menerapkan Metode MOORA dan MOOSRA," vol. 4, no. 3, pp. 1284–1292, 2022, doi: 10.47065/bits.v4i3.2610.