

### ANALISIS SENTIMEN OPINI PENGGUNA TWITTER PADA APLIKASI BIBIT MENGUNAKAN MULTINOMIAL NAÏVE BAYES

Zelin Gaa Ngilo<sup>a</sup>, Nuryuliani<sup>b</sup>

<sup>a</sup>Fakultas Ilmu Komputer dan Teknologi Informasi, [zelin.ngilo@gmail.com](mailto:zelin.ngilo@gmail.com), Universitas Gunadarma

<sup>b</sup>Fakultas Teknologi Industri, [nryulia@staff.gunadarma.ac.id](mailto:nryulia@staff.gunadarma.ac.id), Universitas Gunadarma

#### ABSTRACT

The COVID-19 pandemic has increased the interest number of capital market investors in Indonesia. One of the factor in Indonesian's investment interest is the emergence of fintech in the investment sector. One of the fintech companies in mutual fund investment is "Bibit". To find out user opinions on the Bibit application, a sentiment analysis was carried out on Twitter's users. This study aims to analyze the sentiments of twitter users' opinions on the Bibit application using a combination of Lexicon-Based and Multinomial Naïve Bayes methods. The training data used were 2211 tweets and the validation data was 553 tweets. In the model training process, the training accuracy level is 91.50% and the validation accuracy rate is 85.35%. Model testing was carried out using 39 new tweet data and obtained an accuracy rate of 88%. Sentiment analysis using this method is visualized in the form of pie charts, graphs, and wordclouds. Based on the results of visualization of Twitter social media user sentiment towards the seedling application, it tends to be positive with a percentage of 52% positive and 48% negative.

**Keywords:** sentiment analysis, bibit, lexicon based, multinomial naïve bayes, twitter

#### ABSTRAK

Pandemi COVID-19 meningkatkan minat masyarakat Indonesia terhadap dunia investasi. Munculnya fintech di bidang investasi menjadi faktor pendukung meningkatnya minat investasi masyarakat Indonesia. Salah satu perusahaan fintech yang bergerak di investasi reksadana adalah "Bibit". Untuk mengetahui seberapa baik kualitas aplikasi Bibit di masyarakat maka dilakukan analisis sentiment opini, khususnya pengguna aplikasi Twitter. Penelitian ini bertujuan untuk melakukan analisis sentiment opini pengguna twitter terhadap aplikasi bibit menggunakan metode multinomial naïve bayes. Data latih yang digunakan sebanyak 2211 tweet dan data validasi sebanyak 553 tweet. Pada proses pelatihan model diperoleh tingkat akurasi pelatihan sebesar 91.50% dan tingkat akurasi validasi sebesar 85.35%. Pengujian model dilakukan menggunakan 39 data tweet baru. Hasil pengujian menunjukkan tingkat akurasi yang diperoleh sebesar 88%. Hasil dari analisis sentiment menggunakan kombinasi metode ini divisualisasikan dalam bentuk pie chart, grafik, dan wordcloud. Berdasarkan hasil visualisasi sentiment pengguna sosial media Twitter terhadap aplikasi bibit cenderung positif dengan persentase 52% positif dan 48% negatif.

**Kata Kunci:** analisis sentiment, bibit, lexicon based, multinomial naïve bayes, twitter

#### 1. PENDAHULUAN

Pandemi COVID-19 meningkatkan jumlah investor pasar modal di Indonesia. Merujuk data dari Otoritas Jasa Keuangan (OJK) per akhir Februari 2021, jumlah investor pasar modal meningkat sebesar 16,35 % dari jumlah investor tahun 2020. Industri reksa dana juga mengalami peningkatan yang signifikan. Data Kustodian Sentra Efek Indonesia (KSEI) juga menyebutkan bahwa jumlah investor reksadana di Indonesia meningkat 3,83 juta investor, sehingga peningkatan yang terjadi sebesar 20,50% dari data tahun 2020 sebesar 4,51 juta investor [1].

Peningkatan minat investasi masyarakat Indonesia ini dikarenakan kesadaran akan pentingnya menabung dana darurat, juga dikarenakan investasi semakin mudah dijangkau oleh masyarakat berpenghasilan menengah ke bawah. Perkembangan teknologi informasi yang begitu pesat juga mengambil

andil dalam meningkatnya investor di Indonesia. Hal ini dibuktikan dengan semakin berkembangnya financial technology (fintech). Beberapa contoh platform fintech yang ada di Indonesia adalah OVO, GoPay, Dana, LinkAja, Bibit dan masih banyak yang lainnya. Bibit merupakan perusahaan fintech yang berfokus pada investasi reksa dana [2].

Seirama dengan meningkatnya pengguna internet di Indonesia, maka angka pengguna media sosialnya pun meningkat. Berdasarkan laporan We Are Social, pengguna media sosial di Indonesia setara dengan 61,8 persn dari total populasi pada Januari 2021 [3]. Melalui media sosial para penggunanya bisa mengekspresikan kesehariannya, opini pribadinya maupun membagikan aktifitas serta momen dalam hidupnya. Salah satu media sosial yang paling banyak digunakan di Indonesia adalah Twitter. Untuk mengetahui seberapa baik kualitas aplikasi Bibit di masyarakat maka dilakukan analisis sentiment opini, khususnya pengguna aplikasi Twitter.

Ada banyak metode penelitian yang telah dilakukan mengenai analisis sentimen seperti Naïve Bayes [4,5] dan Algoritma KNN [6]. Penelitian ini bertujuan untuk melakukan analisis sentimen opini pengguna twitter terhadap aplikasi bibit menggunakan metode multinomial naïve bayes. Hasil penelitian ini diharapkan dapat bermanfaat untuk evaluasi dan membantu melakukan riset opini pengguna twitter mengenai aplikasi bibit yang mengandung sentimen positif atau negatif.

## 2. TINJAUAN PUSTAKA

### 2.1. Financial Technology

Financial technology atau yang lebih dikenal dengan fintech merupakan salah satu perkembangan teknologi di bidang keuangan. Fintech merupakan implementasi dan pemanfaatan teknologi untuk peningkatan layanan jasa perbankan dan keuangan yang umumnya dilakukan oleh perusahaan rintisan (startup) yang memanfaatkan teknologi software, internet, dan komunikasi [7]. Bentuk dasar Fintech antara lain:

- Pembayaran (digital wallets, Peer to Peer, payments),
- Investasi (equity crowdfunding, Peer to Peer Lending),
- Pembiayaan (crowdfunding, micro-loans, credit facilities),
- Asuransi (risk management), Lintasproses (big data analysis, predicitive modeling),
- Infrastruktur (security).

### 2.2. Bibit

Bibit merupakan salah satu fintech yang ada di Indonesia. Aplikasi bibit merupakan aplikasi investasi reksadana yang membantu investor pemula dalam memulai investasi. Aplikasi ini sangat cocok digunakan investor pemula karena memiliki teknologi robo advisor. Dengan teknologi robo advisor diharapkan calon investor dapat merancang portofolio investasi yang optimal [8]. Teknologi robo advisor akan mengategorikan calon investor berdasarkan umur, profil risiko, dan tujuan hidup secara otomatis. Terdapat empat jenis reksadana yang dapat dibeli melalui bibit, diantaranya [9]: Reksadana Pasar Uang, Reksadana Obligasi, Reksadana Saham, Reksadana Syariah.

### 2.3. Twitter

Dikutip dari situs resmi Twitter, Twitter merupakan suatu layanan bagi teman, keluarga, dan teman sekerja untuk berkomunikasi dan tetap terhubung melalui pertukaran pesan yang cepat dan sering. Pengguna dapat memposting tweet, yang dapat berisi foto, video, tautan, dan teks [10]. Melalui fitur tweet, pengguna dapat mengutarakan opini pribadinya, mendiskusikan isu-isu terkini, dan mengekspresikan banyak hal lainnya.

Di Indonesia, Twitter termasuk kedalam media sosial yang populer. Hal ini dibuktikan dengan data dari Statista per akhir Januari 2022, total pengguna Twitter di Indonesia mencapai 18,45 juta [11].

### 2.4. Multinomial Naïve Bayes

Multinomial Naïve Bayes termasuk kedalam metode supervised learning. Seperti metode supervised learning lainnya proses multinomial naïve bayes dibagi menjadi dua yaitu pelatihan dan validasi. Untuk pelatihan, probabilitas setiap kata dalam kelas dihitung dengan menggunakan rumus berikut [12]:

$$P(t|c) = \frac{T_{ct}}{\sum_{t' \in V} T_{ct'}} \quad (2.1)$$

Dimana  $T_{ct}$  adalah jumlah kemunculan kata  $t$  dalam dokumen pelatihan kelas  $c$ , dan  $\sum_{t' \in V} T_{ct'}$  adalah jumlah total atribut di kelas  $c$ . Atributnya adalah jumlah total kata dalam kelas  $c$  dan jumlah kata

dalam kosakata. Namun dalam rumus di atas, masalah muncul ketika  $T_{ct}$  sama dengan nol, atau ada kata-kata yang tidak muncul pada data latih. Untuk menghilangkan nilai nol, digunakan *laplace smoothing* untuk menambahkan 1 ke setiap persamaan :

$$P(t|c) = \frac{T_{ct} + 1}{\sum_{t' \in V} (T_{ct'} + 1)} = \frac{T_{ct} + 1}{(\sum_{t' \in V} T_{ct'}) + B'} \quad (2.2)$$

Rumus Maximum a Posteriori (MAP) digunakan untuk menghindari arus bawah dalam proses pengujian untuk memutuskan kelas terbaik untuk sebuah dokumen :

$$C_{map} = \operatorname{argmax}_{c \in C} [\log P(c) + \sum_{1 \leq k \leq n_r} \log P(t|c)] \quad (2.3)$$

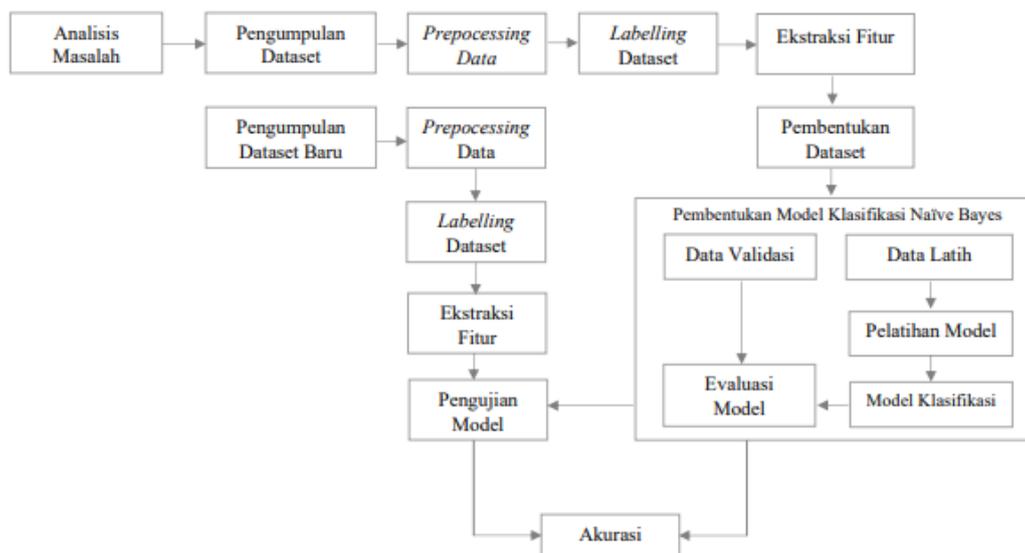
Sebelumnya  $P_{(c)}$  dihitung menggunakan rumus berikut:

$$P(c) = \frac{N_c}{N} \quad (2.4)$$

Dimana  $N_c$  adalah total dokumen di kelas  $c$  dan  $N$  adalah total dokumen dalam kumpulan data.

### 3. METODOLOGI PENELITIAN

Metode yang digunakan pada penelitian analisis sentimen opini pengguna twitter terhadap aplikasi bibit menggunakan *Multinomial Naïve Bayes*, dapat dilihat pada Gambar 1.



Gambar 1. Metode Penelitian

#### 3.1. Analisis Masalah

Twitter merupakan media sosial yang bagus untuk mengumpulkan berbagai opini masyarakat mengenai sesuatu hal. Pengguna aplikasi twitter dapat mengutarakan opini pribadinya, mendiskusikan isu-isu terkini, mengeluh, dan mengekspresikan banyak hal lainnya melalui fitur tweet. *Tweets* yang diungkapkan di akun pengguna twitter terkait aplikasi bibit tersebut mengandung sentimen yang dapat diklasifikasikan menjadi kategori positif dan negatif. Pada penelitian kali ini akan dilakukan analisis sentimen terhadap opini pengguna twitter mengenai aplikasi bibit menggunakan kombinasi *Lexicon Based* dan *Multinomial Naïve Bayes* yang menghasilkan sentimen positif atau negatif dan nilai akurasinya.

#### 3.2. Pengumpulan Dataset

Dataset yang digunakan dalam penelitian ini adalah data tweet masyarakat Indonesia mengenai aplikasi bibit pada media sosial twitter. Dataset dikumpulkan menggunakan teknik *scraping* menggunakan salah satu *library* python yaitu *twint*. *Scraping* data twitter dilakukan dengan pengambilan tweets berdasarkan beberapa kata kunci yang berhubungan dengan aplikasi bibit seperti bibitid, aplikasi bibit, reksadana bibit,

robo bibit, bibi reksadana, dan investasi bibit. *Tweet* yang diambil merupakan tweet dalam jangka waktu tertentu yaitu 1 Januari 2021 sampai 1 Mei 2021.

### 3.3. Preprocessing Data

*Preprocessing* dilakukan untuk mengubah data tweet yang belum terstruktur menjadi tweet yang lebih terstruktur sesuai dengan kebutuhan dalam proses analisis sentimen. *Preprocessing* data terdiri dari lima tahap yaitu:

- a. *Case Folding*  
Tahapan ini digunakan untuk mengubah data tweet menjadi huruf kecil (*lowercase*) agar konsisten.
- b. *Filtering*  
Proses ini akan menghapus delimiter, emoticon, mention, hashtag, link, tanda baca, angka, spasi, spasi ganda, dan *duplicate tweet* yang terdapat pada tweet tersebut
- c. *Tokenize*  
*Tokenize* adalah proses pemisahan teks menjadi potongan-potongan yang disebut sebagai token. Tweet hasil filtering akan diubah menjadi token untuk dianalisis.
- d. *Normalize*  
Tahapan *normalize* atau proses normalisasi digunakan untuk menyeragamkan tweet yang memiliki makna sama namun penulisannya berbeda, bisa diakibatkan kesalahan penulisan, penyingkatan kata, ataupun “bahasa gaul”.
- e. *Stopword Removal*  
Tahap *stopword removal* bertujuan untuk menghilangkan kata yang dianggap tidak penting.

### 3.4. Pelabelan Data

Tahap selanjutnya adalah pelabelan dataset yang telah melalui tahapan *preprocessing*. Proses *labelling* otomatis ini menggunakan metode *Lexicon Based*. Pada penelitian ini menggunakan kamus kata opini (*lexicon*) dari penelitian sebelumnya yaitu InSet Lexicon [13]. Metode ini akan mengklasifikasikan tweet menjadi tiga kategori yaitu positif, netral, dan negatif. Setelah proses *labelling* berhasil dilakukan, data tweet dengan hasil klasifikasi netral akan dihapus, karena pada proses selanjutnya hanya membutuhkan tweet dengan klasifikasi positif dan negatif.

### 3.5. Ekstraksi Fitur

Ekstraksi fitur dengan TF-IDF (Term Frequency - Inverse Document Frequency) adalah salah satu proses dari teknik ekstraksi fitur dengan proses memberikan nilai pada masing-masing kata yang ada pada data latih. Pembobotan kata bertujuan untuk memberikan bobot pada fitur kata berdasarkan frekuensi kemunculan kata [14].

### 3.6. Pembentukan Model Klasifikasi

Proses *modelling* dimulai pembagian dataset menjadi data latih dan data validasi. Dataset akan dibagi dengan perbandingan 80% untuk data latih dan 20% untuk data validasi. Data latih ini kemudian digunakan untuk melatih model *Multinomial Naïve Bayes* agar mendapatkan model yang dapat mengklasifikasikan sentimen suatu *tweets*. Langkah selanjutnya adalah melakukan evaluasi model menggunakan data validasi. Data validasi menggunakan 20% dari dataset yang diambil secara acak.

### 3.7. Pengujian Model

Pengujian model dilakukan dengan data *tweets* baru yang diambil dari Twitter. Data yang digunakan merupakan data *tweets* baru yang tidak terlibat dalam proses pelatihan dan validasi. Sebelum melakukan pengujian terhadap *tweets* baru, dilakukan tahap *preprocessing* terlebih dahulu. Hasil tahapan *preprocessing* akan dilabeli secara otomatis menggunakan metode *Lexicon Based* untuk dijadikan target pengujian model.

## 4. HASIL DAN PEMBAHASAN

### 4.1. Hasil Preprocessing

Tahapan *preprocessing* terdiri dari beberapa proses, yaitu *case folding*, *filtering*, *tokenize*, *normalize*, dan *stopword removal*. Hasil *preprocessing* dapat dilihat pada Tabel 1.

Tabel 1. Tahapan Preprocessing Data

| <i>Preprocessing</i> | Hasil   |
|----------------------|---|
| <i>Raw Tweet</i>     | Kapok bayar @bibitid pake @gopayindonesia. Pembayaran Gopay gak masuk ke Bibit, kejadian udah lebih dari 24 jam dan dananya belum kembali juga. Tiket |

|                         |  |
|-------------------------|--|
|                         | laporan masalah juga gak menyelesaikan masalah sama sekali. Kepastian uangnya balik pun gak ada kabarnya. <a href="https://t.co/KZHFMbr5vm">https://t.co/KZHFMbr5vm</a>  |
| <b>Case Folding</b>     | kapok bayar @bibitid pake @gopayindonesia. pembayaran gopay gak masuk ke bibit, kejadian udah lebih dari 24 jam dan dananya belum kembali juga. tiket laporan masalah juga gak menyelesaikan masalah sama sekali. kepastian uangnya balik pun gak ada kabarnya. <a href="https://t.co/kzhfmbr5vm">https://t.co/kzhfmbr5vm</a>                    |
| <b>Filtering</b>        | kapok bayar pake pembayaran gopay gak masuk ke bibit kejadian udah lebih dari jam dan dananya belum kembali juga tiket laporan masalah juga gak menyelesaikan masalah sama sekali kepastian uangnya balik pun gak ada kabarnya   |
| <b>Tokenize</b>         | ['kapok', 'bayar', 'pake', 'pembayaran', 'gopay', 'gak', 'masuk', 'ke', 'bibit', 'kejadian', 'udah', 'lebih', 'dari', 'jam', 'dan', 'dananya', 'belum', 'kembali', 'juga', 'tiket', 'laporan', 'masalah', 'juga', 'gak', 'menyelesaikan', 'masalah', 'sama', 'sekali', 'kepastian', 'uangnya', 'balik', 'pun', 'gak', 'ada', 'kabarnya']         |
| <b>Normalize</b>        | ['kapok', 'bayar', 'pakai', 'pembayaran', 'gopay', 'tidak', 'masuk', 'ke', 'bibit', 'kejadian', 'sudah', 'lebih', 'dari', 'jam', 'dan', 'dananya', 'belum', 'kembali', 'juga', 'tiket', 'laporan', 'masalah', 'juga', 'tidak', 'menyelesaikan', 'masalah', 'sama', 'sekali', 'kepastian', 'uangnya', 'balik', 'pun', 'tidak', 'ada', 'kabarnya'] |
| <b>Stopword Removal</b> | ['kapok', 'bayar', 'pakai', 'pembayaran', 'gopay', 'masuk', 'bibit', 'kejadian', 'jam', 'dananya', 'tiket', 'laporan', 'menyelesaikan', 'kepastian', 'uangnya', 'kabarnya']  |
| <b>Tweet Bersih</b>     | kapok bayar pakai pembayaran gopay masuk bibit kejadian jam dananya tiket laporan menyelesaikan kepastian uangnya kabarnya   |

#### 4.2. Hasil Pelabelan Dataset

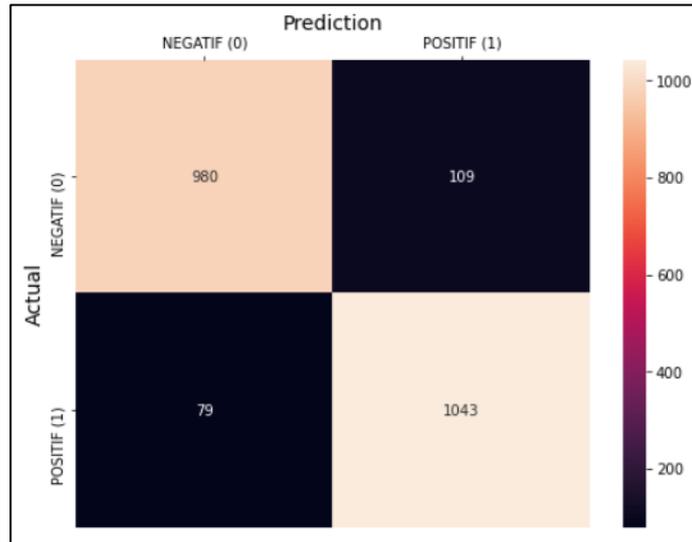
Pelabelan dataset ini dilakukan secara otomatis menggunakan metode *Lexicon Based*. Metode ini menggunakan kamus kata positif dan negatif dari penelitian sebelumnya yaitu InSet Lexicon. Dataset akan di klasifikasikan menjadi tiga kategori yaitu positif, negatif, dan netral. Namun hasil pelabelan dengan kategori netral akan dihapus, sehingga hanya tersisa 2.764 data tweet. Hasil pelabelan dataset dapat dilihat pada Tabel 2.

Tabel 2 Cuplikan Hasil Labelling Dataset

| Tweet Hasil Preprocessing  | Korpus Positif  | Korpus Negatif          | Score | Label Sentimen |
|--|---|-------------------------|-------|----------------|
| kapok bayar pakai pembayaran gopay masuk bibit kejadian jam dananya tiket laporan menyelesaikan kepastian uangnya kabarnya | kapok (-3), bayar (-3), masuk (-3), kejadian (-2), laporan (-5) | bayar (1), kejadian (3) | -12   | NEGATIF        |

#### 4.3. Hasil Pelatihan Model

Tingkat keakuratan dari hasil pelatihan model multinomial naïve bayes didapat dari perhitungan confusion matrix. Confusion matrix pelatihan model dapat dilihat pada Gambar 2.

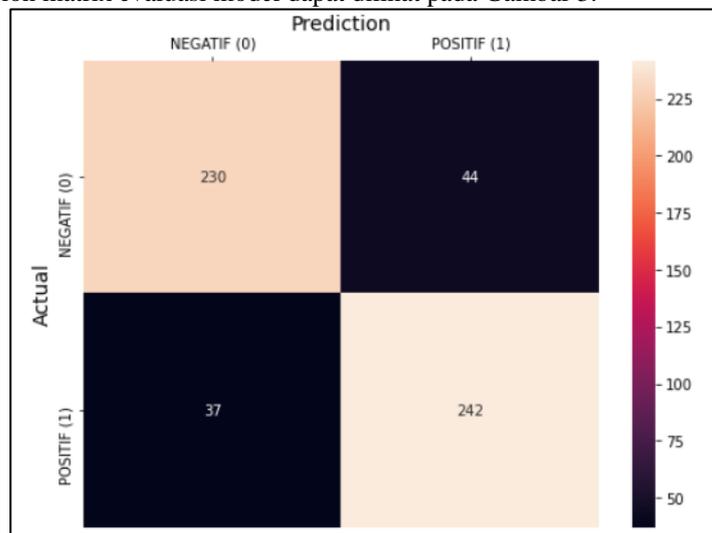


Gambar 2 Confusion Matrix Data Latih

Pada Gambar 2 terdapat 1.059 data tweet negatif, yang terdiri dari 980 tweet terprediksi benar dengan sentimen negatif (TN) dan 79 data tweet terprediksi salah dengan sentimen positif (FN) sedangkan dari 1.152 data tweet positif dihasilkan 1.043 data tweet terprediksi benar dengan sentimen positif (TP) dan 109 data tweet terprediksi salah dengan sentimen negatif (FP). Tingkat keakuratan pelatihan model diukur dengan perhitungan untuk melihat ketepatan model dalam mengklasifikasi tweets ke dalam sentimen positif dan negatif. Berdasarkan 2212 tweets data latih, didapat nilai akurasi sebesar 91.50% yang dihitung berdasarkan banyaknya tweets yang terprediksi dengan benar dibandingkan dengan seluruh tweets yang digunakan.

#### 4.4. Hasil Evaluasi Model

Evaluasi model dilakukan terhadap 20% data validasi yaitu sebanyak 553 data tweet yang diambil secara acak. Confusion matrix evaluasi model dapat dilihat pada Gambar 3.



Gambar 3 Confusion Matrix Data Validasi

Pada Gambar terdapat 267 data tweet negatif yang terdiri dari 230 tweet terprediksi benar dengan sentimen negatif (TN) dan 37 data tweet terprediksi salah dengan sentimen positif (FN) sedangkan dari 286 data tweet positif dihasilkan 242 data tweet terprediksi benar dengan sentimen positif (TP) dan 44 data tweet terprediksi salah dengan sentimen negatif (FP). Tingkat keakuratan evaluasi model diukur dengan perhitungan untuk melihat ketepatan model dalam mengklasifikasi tweets ke dalam sentimen positif dan negatif. Berdasarkan 553 tweets data validasi, didapat nilai akurasi sebesar 85.35% yang dihitung

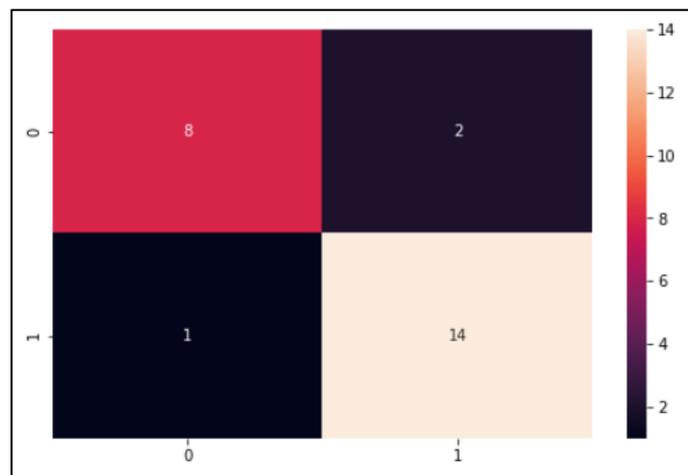
berdasarkan banyaknya tweets yang terprediksi dengan benar dibandingkan dengan seluruh tweets yang digunakan.

#### 4.5. Hasil Pengujian Model

Pengujian model dilakukan terhadap dataset baru yang diambil dari media sosial Twitter sebanyak 39 data tweet. Dataset ini merupakan data yang tidak terlibat dalam proses pelatihan dan evaluasi. Sebelum dataset diuji menggunakan model yang telah dibuat sebelumnya, dataset harus melalui tahapan *preprocessing* data dan pelabelan data menggunakan *Lexicon Based*. Setelah dilakukan pelabelan data yang tersisa adalah 25 tweet. Data tweet pengujian yang telah diklasifikasi oleh model dibandingkan dengan hasil pelabelan menggunakan *Lexicon Based* sebagai target. Cuplikan hasil pengujian dapat dilihat pada Tabel 3 dan confusion matrix pengujian model dapat dilihat pada Gambar 4.

Tabel 3 Cuplikan Hasil Pengujian

| Tweet Hasil Preprocessing  | Target  | Prediksi |
|--|---------|----------|
| rada risih saldo rekening uncontrollable semenjak pakai bibit nagih mindahin reksadana pasar uang rutin topupbulan diamankan aja manajer investasi kelola continuous profit hehe makasih<br>rada risih saldo rekening uncontrollable semenjak pakai bibit nagih mindahin reksadana pasar uang rutin topupbulan diamankan aja manajer investasi kelola continuous profit hehe makasih | POSITIF | POSITIF  |
| gue kayak gini aplikasi oke kadang buka situs chrome kacau gini ganti hape sih   | NEGATIF | NEGATIF  |
| gapernah promonya  | NEGATIF | POSITIF  |



Gambar 4 Confusion Matrix Pengujian Model

Berdasarkan Gambar 4 terdapat 9 data tweet negatif dihasilkan 8 tweet terprediksi benar dengan sentimen negatif (TN) dan 1 data tweet terprediksi salah dengan sentimen positif (FN) sedangkan dari 16 data tweet positif dihasilkan 14 data tweet terprediksi benar dengan sentimen positif (TP) dan 2 data tweet terprediksi salah dengan sentimen negatif (FP). Tingkat keakuratan evaluasi model diukur dengan perhitungan untuk melihat ketepatan model dalam mengklasifikasi tweet ke dalam sentimen positif dan negatif. Dari 25 data pengujian, model dapat memprediksi nilai akurasi sebesar 88%.

#### 5. KESIMPULAN

Pada penelitian ini telah berhasil dilakukan analisis sentimen opini pengguna twitter terhadap aplikasi bibit. Pendekatan yang digunakan untuk analisis sentimen adalah Hybrid Approach menggunakan kombinasi metode Lexicon Based dan Multinomial Naïve Bayes. Proses pelabelan dataset menggunakan metode lexicon based dengan corpus InSet Lexicon. Hasil pelabelan data akan digunakan dalam pembentukan model klasifikasi Multinomial Naïve Bayes. Dalam pembentukan model klasifikasi digunakan proporsi data latih sebanyak 80% dan data validasi sebanyak 20%. Data latih yang digunakan sebanyak 2211 tweet dan data

validasi sebanyak 553 tweet. Pada proses pelatihan model diperoleh tingkat akurasi pelatihan sebesar 91.50% dan tingkat akurasi validasi sebesar 85.35%.

Pengujian model klasifikasi dilakukan menggunakan 39 data tweet baru yang tidak terlibat dalam proses pelatihan dan validasi model. Hasil pengujian menunjukkan tingkat akurasi yang diperoleh sebesar 88% dengan nilai rata-rata recall sebesar 86.65%, nilai rata-rata precision sebesar 88.15% dan nilai rata-rata F1 sebesar 87.25%. Hasil dari analisis sentimen menggunakan kombinasi dua metode ini divisualisasikan dalam bentuk pie chart, grafik, dan wordcloud. Berdasarkan hasil visualisasi sentimen pengguna sosial media Twitter terhadap aplikasi bibit cenderung positif dengan persentase 52% positif dan 48% negatif.

Pada penelitian selanjutnya analisis sentimen opini pengguna twitter terhadap aplikasi bibit dapat dikembangkan dengan menggunakan pendekatan lain seperti Convolutional Neural Network (CNN), Support Vector Machine, atau mengkombinasikan metode Lexicon Based dengan algoritma lainnya. Selain itu pada penelitian selanjutnya dapat menggunakan data yang bersumber dari sosial media lain seperti Facebook dan Instagram.

#### DAFTAR PUSTAKA

- [1]. KSEI Indonesia Central Securities Depository, “Statistik Pasar Modal Indonesia,” Feb. 2021. Accessed: Apr. 10, 2021. [Online]. Available: [https://www.ksei.co.id/files/Statistik\\_Publik\\_Februari\\_2021.pdf](https://www.ksei.co.id/files/Statistik_Publik_Februari_2021.pdf)
- [2]. Bibit, “Gimana Sistem Kerja Fintech?,” Mar. 26, 2020. <https://artikel.bibit.id/keuangan1/gimana-sistem-kerja-fintech> (accessed Apr. 20, 2021).
- [3]. Kompas.com, “Riset Ungkap Lebih dari Separuh Penduduk Indonesia ‘Melek’ Media Sosial,” Feb. 24, 2021. Accessed: Mar. 15, 2021. [Online]. Available: <https://tekno.kompas.com/read/2021/02/24/08050027/riset-ungkap-lebih-dari-separuh-penduduk-indonesia-melek-media-sosial>
- [4]. E A Lisangan, A Gormantara, R Y Carolus, “Implementasi Naïve Bayes pada Analisis Sentimen Opini Masyarakat di Twitter terhadap Kondisi New Normal di Indonesia”, *Konsultasi: Konvergensi Teknologi dan Sistem Informasi*, Vol.2 No.1, April 2022.
- [5]. Aluisius Dwiki Adhi Putra, Safitri Juanita, “Analisis Sentimen pada Ulasan Pengguna Aplikasi Bibit dan Bareksa dengan Algoritma KNN”, *Jurnal Teknik Informatika dan Sistem Informasi*, Vol. 8, No. 2 Juni 2021.
- [6]. M. Saraswati and D. Rimirasih, “ANALISIS SENTIMEN TERHADAP PELAYANAN KRL COMMUTERLINE BERDASARKAN DATA TWITTER MENGGUNAKAN ALGORITMA BERNOULLI NAIVE BAYES,” *Jurnal Ilmiah Informatika Komputer*, vol. 25, no. 3, pp. 225–238, 2020, doi: 10.35760/ik.2020.v25i3.3256.
- [7]. N. Iman, “Financial Technology dan Lembaga Keuangan,” Yogyakarta, Nov. 2016. Accessed: Sep. 06, 2022. [Online]. Available: <http://nofieiman.com/wp-content/images/financial-technology-lembaga-keuangan.pdf>
- [8]. Bibit, “Apa Itu Robo Advisor,” Jul. 11, 2022. <https://faq.bibit.id/id/article/apa-itu-robo-advisor-lop0vt/> (accessed Jul. 14, 2022).
- [9]. Bibit, “Mengenai Bibit Investasi Reksadana Dan Cara Kerjanya,” Dec. 26, 2020. <https://artikel.bibit.id/investasi1/mengenai-bibit-investasi-reksadana-dan-cara-kerjanya> (accessed Mar. 10, 2021).
- [10]. Twitter, “Pertanyaan Umum Pengguna Baru.” <https://help.twitter.com/id/resources/new-user-faq> (accessed Jun. 01, 2021).
- [11]. Databoks, “Pengguna Twitter Indonesia Masuk Daftar Terbanyak di Dunia, Urutan Berapa?,” Mar. 23, 2022. <https://databoks.katadata.co.id/datapublish/2022/03/23/pengguna-twitter-indonesia-masuk-daftar-terbanyak-di-dunia-urutan-berapa> (accessed Jan. 25, 2022).
- [12]. G. Isabelle, W. Maharani, and I. Asror, “Analysis on Opinion Mining Using Combining Lexicon-Based Method and Multinomial Naive Bayes,” *Atlantis Highlights in Engineering*, vol. 2, pp. 214–219, 2019.
- [13]. F. Koto and G. Rahmanytyas, “InSet (Indonesia Sentiment Lexicon),” Dec. 2017. <https://github.com/fajri91/InSet> (accessed Jul. 06, 2021).
- [14]. N. Fitriyah, B. Warsito, D. Asih, and I. Maruddani, “ANALISIS SENTIMEN GOJEK PADA MEDIA SOSIAL TWITTER DENGAN KLASIFIKASI SUPPORT VECTOR MACHINE (SVM),” *JURNAL GAUSSIAN*, vol. 9, no. 3, pp. 376–390, 2020, Accessed: Jan. 07, 2022. [Online]. Available: <https://ejournal3.undip.ac.id/index.php/gaussian/>