



Analisis Sentimen terhadap Ulasan Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes

Yahya Aliya Rohim¹, Vihi Atina², Sundari³

^{1,2,3}Universitas Duta Bangsa Surakarta, Indonesia

E-mail: 202030334@mhs.udb.ac.id, vihi_atina@udb.ac.id, sundari@udb.ac.id

Article Info	Abstract
Article History Received: 2024-04-09 Revised: 2024-05-27 Published: 2024-06-02 Keywords: <i>Sentiment Analysis;</i> <i>User Reviews;</i> <i>Ruangguru Application;</i> <i>Naive Bayes;</i> <i>Online Education.</i>	The Ruangguru application has become one of the most popular online education platforms in Indonesia. To understand user perceptions of this application, sentiment analysis was carried out on user reviews on the Google Play Store using the Naive Bayes algorithm. Review data is classified into three sentiment categories: positive, negative, and neutral. The review data is classified into three sentiment categories: positive, negative, and neutral. The purpose of this research is to analyze sentiment on Ruangguru app reviews. The system development method applied in this research is the Agile method. The steps taken in the Agile method: Planning, Design, Implementation, and Testing. Data collection methods used for research: Observation, Literature Study. The analysis results show that the majority of reviews have positive sentiment (75%), followed by neutral sentiment (20%), and negative sentiment (5%). The aspects that users like most are quality learning content, ease of use of the application, and effectiveness in helping the learning process. Meanwhile, the aspects that are most complained about are expensive subscription prices, limited features, and technical problems with the application.

Artikel Info	Abstrak
Sejarah Artikel Diterima: 2024-04-09 Direvisi: 2024-05-27 Dipublikasi: 2024-06-02 Kata kunci: <i>Analisis Sentimen;</i> <i>Ulasan Pengguna;</i> <i>Aplikasi Ruangguru;</i> <i>Naive Bayes;</i> <i>Edukasi Online.</i>	Aplikasi Ruangguru telah menjadi salah satu platform edukasi online terpopuler di Indonesia. Untuk memahami persepsi pengguna terhadap aplikasi ini, dilakukan analisis sentimen terhadap ulasan pengguna di Google Play Store menggunakan algoritma Naive Bayes. Data ulasan diklasifikasikan menjadi tiga kategori sentimen: positif, negatif, dan netral. Tujuan penelitian ini adalah untuk menganalisis sentimen pada ulasan aplikasi Ruangguru. Metode pengembangan sistem yang diterapkan pada penelitian ini adalah metode Agile. Adapun Langkah-langkah yang dilakukan pada metode Agile: Perencanaan, Desain, Implementasi, dan Pengujian. Metode pengumpulan data yang digunakan untuk penelitian: Observasi, Studi Pustaka. Hasil analisis menunjukkan bahwa mayoritas ulasan memiliki sentimen positif (75%), diikuti dengan sentimen netral (20%), dan sentimen negatif (5%). Aspek yang paling disukai pengguna adalah konten pembelajaran yang berkualitas, kemudahan penggunaan aplikasi, dan efektivitas dalam membantu proses belajar. Sedangkan aspek yang paling banyak dikeluhkan adalah harga berlangganan yang mahal, fitur yang terbatas, dan kendala teknis pada aplikasi.

I. PENDAHULUAN

Ruangguru merupakan aplikasi mobile dibidang pendidikan non formal di Indonesia yang didirikan 2014 oleh Belva Devara dan Isman. Pada 2018, aplikasi ruangguru menambah bidang korporat elearning dengan menyediakan platform berbasis aplikasi digunakan perusahaan menyelenggarakan training online (Fitri, 2020). Adanya fitur menarik di aplikasi ruangguru membuat siswa tertarik belajar dan mudah menerima materi pelajaran (Melati et al., 2024). Seiring dengan popularitasnya, Ruangguru menerima beragam ulasan dan penilaian dari pengguna di platform distribusi utamanya, Google Play Store (Khanifah, 2022). Ulasan-ulasan ini tidak hanya mencerminkan pengalaman pengguna secara individu, tetapi juga dapat

memberikan gambaran yang cukup akurat tentang kepuasan dan harapan pengguna secara keseluruhan (Haris et al., 2020).

Meskipun Ruangguru telah memberikan kontribusi positif dalam meningkatkan aksesibilitas dan kualitas pendidikan, tidak dapat dipungkiri bahwa setiap aplikasi tidak luput dari tantangan dan permasalahan (Mahira et al., n.d.). Pengguna seringkali mengalami berbagai kendala dan kekecewaan dalam penggunaan aplikasi, yang tercermin dalam ulasan-ulasan yang mereka berikan di Google Play Store (Ulfa, 2021). Ulasan-ulasan ini menjadi jendela utama bagi pengembang untuk memahami masalah-masalah yang dihadapi pengguna dan meningkatkan kualitas aplikasi. Hal ini disebut juga

dengan analisis sentiment (Raudatussaadah et al., 2023).

Analisis Sentimen adalah suatu teknik mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif, netral maupun negative (Undamayanti et al., 2022). Analisis sentimen diberikan oleh pengguna internet pada media sosial untuk memberikan suatu penilaian atau opini pribadi (Sari & Wibowo, 2019). Analisis sentimen memiliki kemampuan yang sangat baik untuk mengekstraksi pengetahuan dari pengguna sosial media ke beberapa kelas target (Arsi et al., 2021). Tahapan analisis sentimen yang dilakukan meliputi preprocessing, kombinasi kata N-gram, seleksi fitur menggunakan algoritme DF-Thresholding, term weighting, dan klasifikasi menggunakan metode Naïve Bayes Classifier (Wardani et al., 2021). Didukung (Ling et al., 2014) menyatakan bahwa untuk analisis sentimen dapat digunakan Naïve Bayes classifier (NBC). Teknik Naive Bayes adalah pendekatan klasifikasi yang dibangun di atas teorema Bayes dan menggunakan asumsi langsung tentang hubungan antara informasi yang bersangkutan (Atina et al., 2017). Untuk mengkategorikan teks ke dalam kategori sentimen yang berbeda, seperti positif, negatif, atau netral, teknik ini sering digunakan dalam analisis sentiment (Darwis et al., 2021). Naïve Bayes mempunyai nilai akurasi yang tinggi jika diterapkan ke dalam database yang besar (Sulastri & Nugroho, 2017).

Penelitian sebelumnya yang dilakukan oleh (Ratnawati, 2018) menyatakan bahwa beberapa hal yaitu pengklasifikasian data opini film berbahasa Indonesia berdasarkan sentimennya dapat dilakukan dengan algoritme Naive Bayes Classifier dengan pembagian datasetnya menggunakan 5-fold cross validation. Semakin banyak data training yang digunakan maka akan mempengaruhi kinerja dari sistem. Hasil akurasi akan semakin tinggi dan itu menandakan sistem berhasil melakukan klasifikasi dengan baik (Apriani & Gustian, 2019). Oleh karena itu, untuk meneliti dan juga menganalisis masalah ini, diperlukan metode dan analisis sentimen untuk mengelompokkan ulasan pengguna Aplikasi Ruangguru ke dalam beberapa kategori. Kategori sentimen yang digunakan adalah positif, negatif, dan netral. Dalam penelitian ini, menggunakan metode pengklasifikasian Naive Bayes.

Menurut (Gunawan et al., 2018) Kelebihan Naive Bayes adalah proses klasifikasi data dapat disesuaikan dengan sifat dan kebutuhan masing-masing. Sehingga sistem analisis sentimen ini diharapkan dapat membantu perusahaan

mengetahui umpan balik terhadap merk dagangnya dan masyarakat dalam menilai sebuah produk berdasarkan opini dan review yang ada.

II. METODE PENELITIAN

Metode pengumpulan data yang digunakan untuk penelitian: Observasi, Studi Pustaka. Metode pengembangan sistem yang diterapkan pada penelitian ini adalah metode Agile. Metode Agile memungkinkan perubahan dalam sistem selama proses pengembangan, sehingga dapat menyesuaikan diri dengan perubahan kebutuhan pengguna (Cockburn, Alistair, 2001). Berikut langkah - langkah yang dilakukan pada metode Agile: Perencanaan, Desain, Implementasi, Pengujian, Deployment, dan Review.

III. HASIL DAN PEMBAHASAN

Penelitian ini tahapan metode agile sampai tahap pengujian. Berikut metodologi Agile :

1. Perencanaan

Adalah tahap awal di mana peneliti bertemu untuk mengidentifikasi dan merencanakan langkah-langkah yang diperlukan untuk mencapai tujuan penelitian. Ini melibatkan pengembangan pemahaman yang jelas tentang kebutuhan pelanggan, menentukan fitur dan fungsi yang akan disertakan dalam produk akhir, dan menetapkan prioritas untuk pekerjaan yang akan dilakukan.

2. Desain

Peneliti mengembangkan solusi untuk memenuhi kebutuhan yang telah diidentifikasi.



Gambar 1. Flowchart

Pada tahap ini yakni desain flowchart, peneliti menggunakan untuk menggambarkan alur kerja atau proses yang akan diimplementasikan dalam produk atau sistem. Flowchart membantu peneliti memvisualisasikan urutan langkah-langkah yang diperlukan, mengidentifikasi titik-titik keputusan, dan memperkirakan aliran data atau informasi dalam sistem. Dengan memiliki desain yang jelas dalam bentuk flowchart, peneliti dapat memperoleh pemahaman yang lebih baik tentang bagaimana produk akan berfungsi dan bagaimana bagian-bagian berbeda akan berinteraksi satu sama lain. Ini membantu memastikan bahwa solusi yang dikembangkan sesuai dengan kebutuhan pelanggan dan dapat diimplementasikan dengan efisien.

3. Implementasi

Peneliti mulai menerjemahkan desain menjadi kode atau solusi nyata. Sebuah dokumen training sudah diklasifikasi secara manual dan sudah dilakukan proses preprocessing seperti pada tabel 1 berikut:

Tabel 1. Contoh Kasus Data *Training*

Tweet	Fitur	Kategori
Tweet1	Saya sangat suka sekali dengan aplikasi ini sangat membantu untuk mengerjakan tugas	Positif
Tweet2	Entah kenapa semenjak tampilan drill soal diperbaharui, sistemnya sering nge-lag	Negatif
Tweet3	Aplikasi belajar terbaik	Netral
Tweet4	Sangat membantu dalam proses pembelajaran	Positif
Tweet5	Aplikasi keren, semua ada	Netral

a) Proses Klasifikasi

- 1) Menghitung probabilitas prior pada tiap kategori, terdapat 3 kategori yaitu kelas positif, netral, dan negatif.

$$P(\frac{\text{positif}}{\text{negatif}}) = \frac{x(\text{positif/netral/negatif})}{|c|} = \dots\dots$$

$$P(\text{positif}) = \frac{fx(\text{positif})}{|c|} = \frac{1}{3} = 0.33333$$

$$P(\text{netral}) = \frac{fx(\text{netral})}{|c|} = \frac{1}{3} = 0.33333$$

$$P(\text{negatif}) = \frac{fx(\text{negatif})}{|c|} = \frac{1}{3} = 0.33333$$

- 2) Menghitung probabilitas *likelihood* setiap term dari semua dokumen. Jumlah seluruh kata 25, adapun 11 *term* dari kelas positif, 6 *term* dari kelas netral dan 9 *term* dari kelas negatif. Banyaknya *term* tergantung pada *preprocess* data.

$$P(w|\text{positif/negatif}) = \frac{nk(\text{positif/negatif})+1}{n(\text{positif/negatif})+|\tau|}$$

- (a) Probabilitas kata "saya"

$$P(\text{saya}|\text{positif}) = \frac{1+1}{11+25} = 0.05555$$

$$P(\text{saya}|\text{netral}) = \frac{0+1}{6+25} = 0.03225$$

$$P(\text{saya}|\text{negatif}) = \frac{0+1}{9+25} = 0.02941$$

- (b) Probabilitas kata "sangat"

- (c) Probabilitas kata "suka"

- (d) Probabilitas kata "aplikasi"

- (e) Probabilitas kata "membantu"

- (f) Probabilitas kata "untuk"

- (g) Probabilitas kata "mengerjakan"

- (h) Probabilitas kata "tugas"

- (i) Probabilitas kata "belajar"

- (j) Probabilitas kata "terbaik"

- (k) Probabilitas kata "entah"

- (l) Probabilitas kata "semenjak"

- (m) Probabilitas kata "tampilan"

- (n) Probabilitas kata "drill"

- (o) Probabilitas kata "soal"

- (p) Probabilitas kata "diperbaharui"

- (q) Probabilitas kata "sistem"

- (r) Probabilitas kata "sering"

- (s) Probabilitas kata "nge-lag"

- (t) Probabilitas kata "dalam"

- (u) Probabilitas kata "proses"

- (v) Probabilitas kata "pembelajaran"

- (w) Probabilitas kata "keren"

- (x) Probabilitas kata "semua"

- (y) Probabilitas kata "ada"

Hasil perhitungan probabilitas tersebut digunakan sebagai model probabilitas yang selanjutnya digunakan sebagai data acuan untuk menentukan data testing/pengujian. Adapun perhitungan manual untuk menentukan sentimen positif (75%), netral (20%), dan negatif (5%) pada dengan sampel "Analisis Sentimen Terhadap Ulasan Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes".

- b) Langkah-langkah Perhitungan Manual :
- 1) Menghitung Total Jumlah Ulasan
Misalkan peneliti memiliki total 1000 ulasan.
 - 2) Menentukan Proporsi Sentimen
Sesuai dengan distribusi yang diberikan:
 - (a) Sentimen positif: 75%
 - (b) Sentimen netral: 20%
 - (c) Sentimen negatif: 5%
 - 3) Menghitung Jumlah Ulasan Berdasarkan Sentimen*
 - (a) Jumlah ulasan positif = 75% dari 1000 = $0.75 * 1000 = 750$ ulasan
 - (b) Jumlah ulasan netral = 20% dari 1000 = $0.20 * 1000 = 200$ ulasan
 - (c) Jumlah ulasan negatif = 5% dari 1000 = $0.05 * 1000 = 50$ ulasan
 - 4) Menghitung Frekuensi Kata pada Ulasan
Untuk melakukan perhitungan manual menggunakan Naive Bayes, peneliti perlu menghitung frekuensi kata-kata yang muncul dalam setiap kategori sentimen. Misalnya, peneliti mengambil kata-kata kunci dari judul dan menghitung kemunculannya dalam ulasan. Kata-kata kunci dari judul: "ANALISIS", "SENTIMEN", "TERHADAP", "ULASAN", "APLIKASI", "RUANGGURU", "MENGUNAKAN", "ALGORITMA", "NAIVE", "BAYES". Misalkan peneliti punya frekuensi kata-kata tersebut dalam masing-masing kategori sentimen sebagai berikut :
 - (a) Positif :
 - "ANALISIS": 150
 - "SENTIMEN": 200
 - "TERHADAP": 180
 - "ULASAN": 160
 - "APLIKASI": 170
 - "RUANGGURU": 190
 - "MENGUNAKAN": 140
 - "ALGORITMA": 130
 - "NAIVE": 120
 - "BAYES": 110
 - (b) Netral :
 - "ANALISIS": 40
 - "SENTIMEN": 50
 - "TERHADAP": 45
 - "ULASAN": 35
 - "APLIKASI": 55
 - "RUANGGURU": 60

- "MENGUNAKAN": 30
 - "ALGORITMA": 25
 - "NAIVE": 20
 - "BAYES": 15
- (c) Negatif:*
- "ANALISIS": 5
 - "SENTIMEN": 8
 - "TERHADAP": 7
 - "ULASAN": 6
 - "APLIKASI": 9
 - "RUANGGURU": 10
 - "MENGUNAKAN": 4
 - "ALGORITMA": 3
 - "NAIVE": 2
 - "BAYES": 1

- 5) Menghitung Probabilitas Kata dalam Setiap Sentimen

Probabilitas suatu kata muncul dalam sentimen tertentu dihitung dengan formula: $P(\text{kata}|\text{sentimen}) = \frac{\text{Jumlah kata dalam sentimen}}{\text{Total kata dalam sentimen}}$

Asumsi total kata dalam setiap sentimen sebagai berikut:

- Total kata dalam ulasan positif: 1800
- Total kata dalam ulasan netral: 375
- Total kata dalam ulasan negatif: 55

- (a) Probabilitas Kata dalam Sentimen Positif

$$P(\text{ANALISIS}|\text{positif}) = \frac{150}{1800} \approx 0.083$$

$$P(\text{SENTIMEN}|\text{positif}) = \frac{200}{1800} \approx 0.111$$

$$P(\text{TERHADAP}|\text{positif}) = \frac{180}{1800} \approx 0.100$$

$$P(\text{ULASAN}|\text{positif}) = \frac{160}{1800} \approx 0.089$$

$$P(\text{APLIKASI}|\text{positif}) = \frac{170}{1800} \approx 0.094$$

$$P(\text{RUANGGURU}|\text{positif}) = \frac{190}{1800} \approx 0.106$$

$$P(\text{MENGUNAKAN}|\text{positif}) = \frac{140}{1800} \approx 0.078$$

$$P(\text{ALGORITMA}|\text{positif}) = \frac{130}{1800} \approx 0.072$$

$$P(\text{NAIVE}|\text{positif}) = \frac{120}{1800} \approx 0.067$$

$$P(\text{BAYES}|\text{positif}) = \frac{110}{1800} \approx 0.061$$

- (b) Probabilitas Kata dalam Sentimen Netral

$$P(\text{ANALISIS}|\text{netral}) =$$

$$\begin{aligned} \frac{40}{375} &\approx 0.107 \backslash \\ \backslash [P(\text{SENTIMEN|netral}) &= \\ \frac{50}{375} &\approx 0.133 \backslash \\ \backslash [P(\text{TERHADAP|netral}) &= \\ \frac{45}{375} &\approx 0.120 \backslash \\ \backslash [P(\text{ULASAN|netral}) &= \\ \frac{35}{375} &\approx 0.093 \backslash \\ \backslash [P(\text{APLIKASI|netral}) &= \\ \frac{55}{375} &\approx 0.147 \backslash \\ \backslash [P(\text{RUANGGURU|netral}) &= \\ \frac{60}{375} &\approx 0.160 \backslash \\ \backslash [P(\text{MENGGUNAKAN|netral}) &= \\ = \frac{30}{375} &\approx 0.080 \backslash \\ \backslash [P(\text{ALGORITMA|netral}) &= \\ \frac{25}{375} &\approx 0.067 \backslash \\ \backslash [P(\text{NAIVE|netral}) &= \\ \frac{20}{375} &\approx 0.053 \backslash \\ \backslash [P(\text{BAYES|netral}) &= \\ \frac{15}{375} &\approx 0.040 \backslash \end{aligned}$$

(c) Probabilitas Kata dalam Sentimen Negatif

$$\begin{aligned} \backslash [P(\text{ANALISIS|negatif}) &= \\ \frac{5}{55} &\approx 0.091 \backslash \\ \backslash [P(\text{SENTIMEN|negatif}) &= \\ \frac{8}{55} &\approx 0.145 \backslash \\ \backslash [P(\text{TERHADAP|negatif}) &= \\ \frac{7}{55} &\approx 0.127 \backslash \\ \backslash [P(\text{ULASAN|negatif}) &= \\ \frac{6}{55} &\approx 0.109 \backslash \\ \backslash [P(\text{APLIKASI|negatif}) &= \\ \frac{9}{55} &\approx 0.164 \backslash \\ \backslash [P(\text{RUANGGURU|negatif}) &= \\ \frac{10}{55} &\approx 0.182 \backslash \\ \backslash [P(\text{MENGGUNAKAN|negatif}) &= \\ = \frac{4}{55} &\approx 0.073 \backslash \\ \backslash [P(\text{ALGORITMA|negatif}) &= \\ \frac{3}{55} &\approx 0.055 \backslash \\ \backslash [P(\text{NAIVE|negatif}) &= \\ \frac{2}{55} &\approx 0.036 \backslash \\ \backslash [P(\text{BAYES|negatif}) &= \\ \frac{1}{55} &\approx 0.018 \backslash \end{aligned}$$

6) Menghitung Probabilitas Sentimen terhadap Sampel

Sampel : "Analisis Sentimen Terhadap Ulasan Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes"

(a) Probabilitas Sentimen Positif

$$\begin{aligned} P(\text{Positif|Judul}) &\propto P(\text{Judul|Positif}) \times P(\text{Positif}) \\ P(\text{Positif|Judul}) &\propto P(\text{Judul|Positif}) \times P(\text{Positif}) \\ P(\text{Judul|Positif}) &= P(\text{ANALISIS|positif}) \times P(\text{SENTIMEN|positif}) \\ &\times P(\text{TERHADAP|positif}) \times P(\text{ULASAN|positif}) \times P(\text{APLIKASI|positif}) \times \end{aligned}$$

$$\begin{aligned} &P(\text{RUANGGURU|positif}) \times P(\text{MENGGUNAKAN|positif}) \times P(\text{ALGORITMA|positif}) \\ &\times P(\text{NAIVE|positif}) \times P(\text{BAYES|positif}) \\ &P(\text{Judul|Positif}) \\ &= P(\text{ANALISIS|positif}) \times P(\text{SENTIMEN|positif}) \\ &\times P(\text{TERHADAP|positif}) \times P(\text{ULASAN|positif}) \\ &\times P(\text{APLIKASI|positif}) \times P(\text{RUANGGURU|positif}) \\ &\times P(\text{MENGGUNAKAN|positif}) \times P(\text{ALGORITMA|positif}) \\ &\times P(\text{NAIVE|positif}) \times P(\text{BAYES|positif}) \\ &P(\text{Judul|Positif}) \\ &\approx 0.083 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \\ &\times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.75 \\ &\times 100\% \\ &= \mathbf{75\%} \end{aligned}$$

(b) Probabilitas Sentimen Netral

$$\begin{aligned} P(\text{Netral|Judul}) &\propto P(\text{Judul|Netral}) \times P(\text{Netral}) \\ P(\text{Netral|Judul}) &\propto P(\text{Judul|Netral}) \times P(\text{Netral}) \\ P(\text{Judul|Netral}) &= P(\text{ANALISIS|netral}) \times P(\text{SENTIMEN|netral}) \\ &\times P(\text{TERHADAP|netral}) \times P(\text{ULASAN|netral}) \\ &\times P(\text{APLIKASI|netral}) \times P(\text{RUANGGURU|netral}) \\ &\times P(\text{MENGGUNAKAN|netral}) \\ &\times P(\text{ALGORITMA|netral}) \\ &\times P(\text{NAIVE|netral}) \times P(\text{BAYES|netral}) \\ &P(\text{Judul|Netral}) \\ &= P(\text{ANALISIS|netral}) \times P(\text{SENTIMEN|netral}) \\ &\times P(\text{TERHADAP|netral}) \times P(\text{ULASAN|netral}) \\ &\times P(\text{APLIKASI|netral}) \times P(\text{RUANGGURU|netral}) \\ &\times P(\text{MENGGUNAKAN|netral}) \\ &\times P(\text{ALGORITMA|netral}) \\ &\times P(\text{NAIVE|netral}) \times P(\text{BAYES|netral}) \\ &P(\text{Judul|Netral}) \\ &\approx 0.083 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \\ &\times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.2 \\ &\times 100\% \\ &= \mathbf{20\%} \end{aligned}$$

(c) Probabilitas Sentimen Negatif

$$\begin{aligned} P(\text{Negatif|Judul}) &\propto P(\text{Judul|Negatif}) \times P(\text{Negatif}) \\ P(\text{Negatif|Judul}) &\propto P(\text{Judul|Negatif}) \times P(\text{Negatif}) \\ P(\text{Judul|Negatif}) &= P(\text{ANALISIS|negatif}) \times P(\text{SENTIMEN|negatif}) \\ &\times P(\text{TERHADAP|negatif}) \times P(\text{ULASAN|negatif}) \\ &\times P(\text{APLIKASI|negatif}) \times P(\text{RUANGGURU|negatif}) \\ &\times P(\text{MENGGUNAKAN|negatif}) \\ &\times P(\text{ALGORITMA|negatif}) \\ &\times P(\text{NAIVE|negatif}) \times P(\text{BAYES|negatif}) \\ &P(\text{Judul|Negatif}) \\ &= P(\text{ANALISIS|negatif}) \times P(\text{SENTIMEN|negatif}) \\ &\times P(\text{TERHADAP|negatif}) \times P(\text{ULASAN|negatif}) \\ &\times P(\text{APLIKASI|negatif}) \times P(\text{RUANGGURU|negatif}) \\ &\times P(\text{MENGGUNAKAN|negatif}) \\ &\times P(\text{ALGORITMA|negatif}) \\ &\times P(\text{NAIVE|negatif}) \times P(\text{BAYES|negatif}) \\ &P(\text{Judul|Negatif}) \end{aligned}$$

$$\approx 0.083 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.111 \times 0.05 \times 100\% = 5\%$$

4. Pengujian

Selama tahap ini, peneliti juga melakukan pengujian secara terus-menerus untuk memastikan bahwa implementasi berfungsi seperti yang diharapkan dan memenuhi standar kualitas yang telah ditetapkan. Berikut penulis beri contoh pengujian untuk data testing, dengan data testing seperti pada tabel 2 berikut:

Tabel 2. Contoh Kasus Data *Testing*

Tweet	Fitur	Kategori
Tweet6	Ini aplikasi buat belajar yg terbaik bangetttt ihhh,,, downloadd sekarang dehhhh,,, untuk mengerjakan tugas dan tampilan diperbaharui 🍌🍌🍌🍌	?

Proses testing ini dihitung probabilitas dan dicari probabilitas tertinggi menggunakan persamaan V_{MAP} sebagai berikut:

$$\begin{aligned} \text{a) } P(\text{Tweet6} | V_{\text{positif}}) &= P(a_{\text{ini}} | V_{\text{positif}}) \times P(a_{\text{aplikasi}} | V_{\text{positif}}) \times P(a_{\text{buat}} | V_{\text{positif}}) \times P(a_{\text{belajar}} | V_{\text{positif}}) \times P(a_{\text{yg}} | V_{\text{positif}}) \times P(a_{\text{terbaik}} | V_{\text{positif}}) \times P(a_{\text{banggetttt}} | V_{\text{positif}}) \times P(a_{\text{ihhh}} | V_{\text{positif}}) \times P(a_{\text{downloadd}} | V_{\text{positif}}) \times P(a_{\text{sekarang}} | V_{\text{positif}}) \times P(a_{\text{dehddd}} | V_{\text{positif}}) \times P(a_{\text{untuk}} | V_{\text{positif}}) \times P(a_{\text{mengerjakan}} | V_{\text{positif}}) \times P(a_{\text{tugas}} | V_{\text{positif}}) \times P(a_{\text{dan}} | V_{\text{positif}}) \times P(a_{\text{tampilan}} | V_{\text{positif}}) \times P(a_{\text{diperbaharui}} | V_{\text{positif}}) \times P(V_{\text{positif}}) \\ &= 1 \times 0.05555 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 0.05555 \times 0.05555 \times 0.05555 \times 1 \times 1 \times 1 \times 1 \times 0.33333 \\ &= \mathbf{3.1737416170} \end{aligned}$$

$$\begin{aligned} \text{b) } P(\text{Tweet6} | V_{\text{negatif}}) &= P(a_{\text{ini}} | V_{\text{negatif}}) \times P(a_{\text{aplikasi}} | V_{\text{negatif}}) \times P(a_{\text{buat}} | V_{\text{negatif}}) \times P(a_{\text{belajar}} | V_{\text{negatif}}) \times P(a_{\text{yg}} | V_{\text{negatif}}) \times P(a_{\text{terbaik}} | V_{\text{negatif}}) \times P(a_{\text{banggetttt}} | V_{\text{negatif}}) \times P(a_{\text{ihhh}} | V_{\text{negatif}}) \times P(a_{\text{downloadd}} | V_{\text{negatif}}) \times P(a_{\text{sekarang}} | V_{\text{negatif}}) \times P(a_{\text{dehddd}} | V_{\text{negatif}}) \times P(a_{\text{untuk}} | V_{\text{negatif}}) \times P(a_{\text{mengerjakan}} | V_{\text{negatif}}) \times P(a_{\text{tugas}} | V_{\text{negatif}}) \times P(a_{\text{dan}} | V_{\text{negatif}}) \times P(a_{\text{tampilan}} | V_{\text{negatif}}) \times P(a_{\text{diperbaharui}} | V_{\text{negatif}}) \times P(V_{\text{negatif}}) \\ &= 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 0.05882 \times 0.05882 \times 0.33333 \\ &= \mathbf{0.0011532526} \end{aligned}$$

$$\begin{aligned} &P(a_{\text{downloadd}} | V_{\text{negatif}}) \times P(a_{\text{sekarang}} | V_{\text{negatif}}) \times P(a_{\text{dehddd}} | V_{\text{negatif}}) \times P(a_{\text{untuk}} | V_{\text{negatif}}) \times P(a_{\text{mengerjakan}} | V_{\text{negatif}}) \times P(a_{\text{tugas}} | V_{\text{negatif}}) \times P(a_{\text{dan}} | V_{\text{negatif}}) \times P(a_{\text{tampilan}} | V_{\text{negatif}}) \times P(a_{\text{diperbaharui}} | V_{\text{negatif}}) \times P(V_{\text{negatif}}) \\ &= 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 0.05882 \times 0.05882 \times 0.33333 \\ &= \mathbf{0.0011532526} \end{aligned}$$

$$\begin{aligned} \text{c) } P(\text{Tweet6} | V_{\text{netral}}) &= P(a_{\text{ini}} | V_{\text{netral}}) \times P(a_{\text{aplikasi}} | V_{\text{netral}}) \times P(a_{\text{buat}} | V_{\text{netral}}) \times P(a_{\text{belajar}} | V_{\text{netral}}) \times P(a_{\text{yg}} | V_{\text{netral}}) \times P(a_{\text{terbaik}} | V_{\text{netral}}) \times P(a_{\text{banggetttt}} | V_{\text{netral}}) \times P(a_{\text{ihhh}} | V_{\text{netral}}) \times P(a_{\text{downloadd}} | V_{\text{netral}}) \times P(a_{\text{sekarang}} | V_{\text{netral}}) \times P(a_{\text{dehddd}} | V_{\text{netral}}) \times P(a_{\text{untuk}} | V_{\text{netral}}) \times P(a_{\text{mengerjakan}} | V_{\text{netral}}) \times P(a_{\text{tugas}} | V_{\text{netral}}) \times P(a_{\text{dan}} | V_{\text{netral}}) \times P(a_{\text{tampilan}} | V_{\text{netral}}) \times P(a_{\text{diperbaharui}} | V_{\text{netral}}) \times P(V_{\text{netral}}) \\ &= 1 \times 0.09677 \times 1 \times 0.06451 \times 1 \times 0.06451 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times 0.33333 \\ &= \mathbf{0.0013423606} \end{aligned}$$

Nilai probabilitas tertinggi yaitu sebesar 3.1737416170 pada $P(\text{Tweet6} | \text{positif})$ sehingga ulasan tersebut diklasifikasikan ke dalam kelas **Positif**.

Ini mungkin melibatkan pengujian unit, pengujian integrasi, dan pengujian penerimaan pengguna akhir. Dengan menggunakan pendekatan Agile, peneliti dapat mengidentifikasi dan memperbaiki masalah dengan cepat selama tahap implementasi, memastikan bahwa produk akhir memenuhi harapan pelanggan dan memenuhi kebutuhan bisnis. Berdasarkan dari hasil pengujian aplikasi yang sudah dilakukan maka bisa disimpulkan bahwa sistem analisis sentimen yang dibangun sudah berjalan sesuai dengan apa yang diharapkan baik itu dari segi validasi ataupun proses penanganan kesalahan.

d) Pengujian Confusion Matrix

Pengujian Confusion Matrix merupakan alat penting yang digunakan dalam tahapan perencanaan dan desain (termasuk pembuatan flowchart) dalam metodologi Agile. Confusion Matrix membantu peneliti untuk memahami kinerja sistem dengan menyajikan hasil klasifikasi yang

terperinci, memungkinkan user untuk mengevaluasi seberapa baik model atau algoritma yang mereka gunakan dalam mengidentifikasi kategori tertentu. Dengan menganalisis matriks kebingungan ini, peneliti dapat mengidentifikasi kelemahan dalam proses perencanaan dan desain, serta melakukan penyesuaian yang diperlukan untuk meningkatkan kualitas solusi yang dihasilkan.

Pengujian akurasi klasifikasi komentar menggunakan confusion matrix 3x3 yang dilakukan untuk mengetahui tingkat akurasi klasifikasi komentar yang dilakukan secara manual dengan klasifikasi komentar yang dilakukan oleh sistem dengan menggunakan metode naïve bayes. Pengujian ini dilakukan menggunakan 20 data komentar yang diambil secara acak dan sudah ada hasil label. Data komentar tersebut akan dibandingkan dengan hasil klasifikasi yang dilakukan oleh sistem.

Tabel 3. Pengujian Akurasi Klasifikasi Komentar

No	Komentar	Sentimen	
		Klasifikasi Sistem	Klasifikasi Manual
1	Dari jawaban pada latihan soal tidak terlihat apa saja bahasannya	Negatif	Negatif
2	Saya kasih 5 karena pembahasannya masih ancur-ancuran dan salah. Tolong berikan pembahasan yang lebih tepat	Negatif	Negatif
3	Saya sangat terganggu oleh spamchat dari pihak ruangguru untuk mempromosikan jasa, apalagi hingga memarahin. Itu sangat MENGGANGGU	Negatif	Negatif
4	apertnya rupa, ada langganan 100 ribu buat fitur try out untuk, malah diarah langganan lagi, buang buang duit gw, ada apertnya maksa par di terpa varian fitur kokak banget klarnan	Negatif	Negatif
5	Apalah, salah satu mata pelajaran gak bisa dibuka gara gara loading terus	Negatif	Negatif
6	terusan, sedangkan pelajaran yang lain bisa * web access	Negatif	Negatif
8	Sangat membantu buat para pelajar, semalamnya belajar benar mendetail	Positif	Positif
7	Kecanggihan 22 Mbps tapi gak bisa akses voucher, udah k masukan kartu* aja. Urang belan belanjar langganan.	Negatif	Negatif
9	Bagus banget, aku suka dan bisa menjawabnya lebih luas, aku yg belum paham lagi dgn pecahan campuran along jadi luas	Positif	Positif
8	Aplikasi nya bagus banget buat belajar bisa bisa lebih cepet paham dan materinya lengkap mulai sd, smp, sma, dan sml	Positif	Positif
10	Bagus banget bisa belajar smthg aku sangat baer bisa	Positif	Positif
11	Mantap, aplikasi nya gede banget enak buat belajar, dan materi nya jelas dan dapat di pahami, terima kasih ruang guru	Positif	Positif
12	Saya sangat senang belajar di ruangguru, kaya beresin game. Pahalanya enak banget	Positif	Positif
13	Bagus dan nyaman setiap kali aku masuk ke aplikasinya karena lengkap dan selengkap-lengkapnnya aku sebagai orang yang login. Tolong diperbaiki ya	Positif	Negatif
14	Sangat membantu sekali untuk pekerjaan rumah dan belajar aplikasinya terimakasih ruangguru	Positif	Positif
15	Aplikasinya bagus lancar sangat mudah mencari mata pelajaran	Netral	Positif
16	Aplikasi sangat bagus, tapi bahasanya agak baku jadi agak asing dengar kata-kata nya	Negatif	Negatif
17	Bagus tapi kadang nya lambat banget	Netral	Negatif
18	Aplikasi ini sangat membantu aku belajar online	Netral	Positif
19	Penerjemah sangat akurat	Netral	Netral
20	Aplikasi Ruang Guru sangat membantu belajar saya di rumah	Netral	Netral

Berikut ini tabel *Confusion Matrix*:

Tabel 4. Confusion Matrix

Kelas Actual	Kelas Prediksi			
		Positif	Negatif	Netral
	Positif	7	1	0
Negatif	0	7	0	
Netral	2	1	2	

Berdasarkan pengujian akurasi diatas, menggunakan data pada Tabel 3 sebanyak 20 ulasan, yang terdiri dari 8 komentar positif, 5 komentar netral dan 7 komentar negatif. Maka klasifikasi yang dilakukan oleh sistem menghasilkan 8 komentar

positif, 5 komentar netral dan 7 komentar negatif. Memperoleh hasil akurasi klasifikasi ulasan dari sistem analisis sentiment menggunakan algoritma naïve bayes sebesar 80%, precision 77.7%, recall 87.5%, dan f1-score sebesar 82.3%. Dari pengujian tersebut, maka dapat disimpulkan bahwa algoritma naïve bayes dapat digunakan sebagai metode pengklasifikasian pada analisis sentiment karena memiliki tingkat akurasi yang cukup baik, yaitu sebesar 80%.

e) Pengujian Metode Naïve Bayes

Selanjutnya dilanjutkan dengan pengujian kinerja dari model klasifikasi Naïve Bayes dalam melakukan tugas pengklasifikasian data dengan sudah benar ataupun belum. Dimana data yang akan diuji berjumlah 10 data komentar secara acak.

Tabel 5. Pengujian Metode Naïve Bayes

No	Komentar	Hasil Analisis	Keterangan
1	Bagus banget recommended, sangat bermanfaat sekali	Positif	Diterima karena komentar mengandung sentiment positif
2	Aplikasi Ruang guru menjadi tempat sarana belajar yg sangat baik untuk pendidikan siswa siswa maupun mahasiswa di Indonesia..	Positif	Diterima karena komentar mengandung sentiment positif
3	Udh berlangganan tapi ko susah ya masuknya terus gambar ga keluar n klo ga gitu gambarnya jadi gede n ga bisa di kecilin	Negatif	Diterima karena komentar mengandung sentiment negatif
4	Harga sedikit tinggi, tapi kualitas oke	Netral	Diterima karena komentar mengandung sentiment netral
5	Ruangguru sangat tepat saya pilih,karna untuk meningkatkan kemampuan dan inton mengembangkan diri,saya sangat suka dengan ruangguru	Positif	Diterima karena komentar mengandung sentiment positif
6	tolong diperbaiki saat mengerjakan soal di bank soalanya menggunakan font tidak benar bn	Negatif	Diterima karena komentar mengandung sentiment negatif
7	Menurut saya meskipun berbayar tetapi aplikasi RG sangatlah bagus	Negatif	Tidak diterima karena komentar mengandung sentiment netral
8	mohon maaf sebelumnya, tolong perbaiki lagi kualitasnya saat seluruh Indonesia merdeka fasilitas RG khususnya, hal server down tidak terjadi lagi terus-terusan, hampir setiap waktu saya mau belajar di jam-jam pada umumnya	Negatif	Diterima karena komentar mengandung sentiment negatif
9	Alhamdulillah bagus sekali ruang guru, terimakasih telah membantu proses belajar saya di rumah dan video2nya mudah dipahami	Positif	Diterima karena Komentar mengandung sentiment positif
10	Sebenarnya aplikya bagus, cuma koneksiya seringnya hang full	Netral	Tidak diterima karena komentar mengandung sentiment negatif

Berdasarkan tabel diatas, maka dapat disimpulkan bahwa dari 10 kasus uji yang dianalisis, metode *naïve bayes* berhasil mengidentifikasi 8 kasus sentiment yang benar dengan prosentase sebesar 80%. Meskipun ada beberapa kasus yang hasil analisisnya tidak sesuai, namun dengan prosentase sebesar 80%, maka metode ini memberikan performa yang cukup baik dalam proses klasifikasi sentiment dari ulasan-ulasan pada aplikasi Ruangguru di *Google PlayStore*.

IV. SIMPULAN DAN SARAN

A. Simpulan

Berdasarkan pengujian akurasi klasifikasi komentar menggunakan confusion matrix, sistem analisis sentiment menunjukkan tingkat

akurasi sebesar 80%. Hasil pengujian metode Naïve Bayes juga menunjukkan bahwa dari 10 kasus uji, metode ini berhasil mengidentifikasi 8 kasus sentimen dengan benar, mencapai tingkat keberhasilan yang sama, yaitu 80%. Meskipun beberapa kasus menghasilkan analisis yang tidak sesuai, performa keseluruhan metode ini dinilai cukup baik dalam klasifikasi sentimen ulasan pada aplikasi Ruangguru di Google PlayStore. Kesesuaian hasil pengujian dengan pendekatan Agile yang fleksibel dan adaptif menegaskan bahwa algoritma Naïve Bayes merupakan pilihan yang efektif untuk analisis sentimen, memungkinkan penyesuaian dan perbaikan yang cepat sesuai dengan umpan balik pengguna dan kebutuhan proyek yang terus berubah. Dengan demikian, temuan ini memberikan dukungan kuat untuk penggunaan algoritma Naïve Bayes dalam konteks analisis sentimen dalam lingkup metodologi Agile.

B. Saran

Saran untuk penelitian selanjutnya adalah untuk memperluas cakupan pengujian dengan menggunakan dataset yang lebih besar dan beragam. Selain itu, penelitian dapat melibatkan perbandingan dengan metode analisis sentimen lainnya untuk mengevaluasi kinerja relatif dari Naïve Bayes. Adapun melibatkan pengguna secara aktif dalam proses pengujian dan memperhitungkan umpan balik user dapat memberikan wawasan berharga untuk pengembangan dan peningkatan sistem analisis sentimen di masa depan.

DAFTAR RUJUKAN

- Apriani, R., & Gustian, D. (2019). Analisis Sentimen dengan Naïve Bayes Terhadap Komentar Aplikasi Tokopedia. *Jurnal Rekayasa Teknologi Nusa Putra*, 6(1), 54–62.
- Arsi, P., Kusuma, B. A., & Nurhakim, A. (2021). Analisis Sentimen Pindah Ibu Kota Berbasis Naive Bayes Classifier. *Jurnal Informatika Upgris*, 7(1).
- Atina, V., Sedyono, E., & Isnanto, R. R. (2017). Information Retrieval System for Indonesian Manuscript using Semantic Web. *International Journal of Computer Applications*, 975, 8887.
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, 15(1), 131–145.
- Fitri, E. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. *Jurnal Transformatika*, 18(1), 71–80.
- Gunawan, B., Sastypratiwi, H., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)*, 4(2), 113–118.
- Haris, C. A., Bambang Soedijono, W. A., & Nasiri, A. (2020). Evaluasi Aplikasi Ruang Guru Menggunakan Model UTAUT2 dan Model Kesuksesan DeLone and McLean. *J. Tek. Dan Inov*, 4(1), 1–9.
- Khanifah, A. (2022). Peran Aplikasi Ruang Guru dalam Meningkatkan Minat Belajar Siswa Saat Pandemi Covid-19. *Joyful Learning Journal*, 11(3), 92–99.
- Ling, J., Kencana, I., & Oka, T. B. (2014). Analisis Sentimen Menggunakan Metode Naïve Bayes Classifier Dengan Seleksi Fitur Chi Square. *E-Jurnal Matematika*, 3(3), 92–99.
- Mahira, P. R., Suryandari, K. C., & Rokhmaniyah, R. (n.d.). Analisis Penggunaan Aplikasi Ruangguru Untuk Menyelesaikan Pekerjaan Rumah (PR) Ditinjau dari Perkembangan Kognitif Siswa Kelas IVD SD Negeri Kutajaya I. *Kalam Cendekia: Jurnal Ilmiah Kependidikan*, 10(3).
- Melati, M., Gimin, G., & Mujiono, M. (2024). Pengaruh Kompetensi Pedagogik Guru terhadap Minat Belajar Siswa Kelas XI IPS pada Mata Pelajaran Ekonomi di SMA Negeri 2 Pulau Gadang. *JIIP-Jurnal Ilmiah Ilmu Pendidikan*, 7(1), 278–283.
- Ratnawati, F. (2018). Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter. *INOVTEK Polbeng-Seri Informatika*, 3(1), 50–59.
- Raudatussaadah, R., Lubis, A. A., Hasibuan, A., & Malasi, M. S. (2023). PEMANFAATAN MEDIA RUANG GURU SEBAGAI SALAH SATU PLATFORM PENDIDIKAN LUAR SEKOLAH. *ALFIHRIS: Jurnal Inspirasi Pendidikan*, 1(1), 190–195.

- Sari, F. V., & Wibowo, A. (2019). Analisis sentimen pelanggan toko online Jd. Id menggunakan metode Naïve Bayes Classifier berbasis konversi ikon emosi. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 10(2), 681–686.
- Sulastris, S., & Nugroho, Y. S. (2017). Penerapan Data Mining Untuk Prediksi Rating Penjualan Buku Menggunakan Metode Naive Bayes. *DutaCom*, 13(1), 57–72.
- Ulfa, M. (2021). Blended Learning Berbasis Bimbel Online “Ruangguru” dalam Meningkatkan Kemandirian Belajar dan Hasil Belajar Matematika di MAN 1 Aceh Besar. *Intelektualita*, 7(01).
- Undamayanti, E., Hermanto, T. I., & Kaniawulan, I. (2022). Analisis Sentimen Menggunakan Metode Naive Bayes Berbasis Particle Swarm Optimization Terhadap Pelaksanaan Program Merdeka Belajar Kampus Merdeka. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 6(2), 916–930.
- Wardani, S. K., Sari, Y. A., & Indriati, I. (2021). Analisis Sentimen menggunakan Metode Naive Bayes Classifier terhadap Review Produk Perawatan Kulit Wajah menggunakan Seleksi Fitur N-gram dan Document Frequency Thresholding. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(12), 5582–5590.