**IV HASIL DAN PEMBAHASAN**

1. **Data selection**

Seleksi data atau fitur yang dilakukan dalam penelitian ini digunakan untuk melihat gambaran umum mengenai informasi dalam bentuk *tweet* yang berkaitan dengan *brand* *Ruangguru* yang didapat dari *twitter* menggunakan API milik *twitter* dan *software* (perangkat lunak) *Orange*. Data yang didapat dari hasil proses scraping yang dilakukan berbentuk atau berformat *excel*(.xlxs). **Tabel 2.** Merupakan contoh data *tweet* yang didapat dan menggambarkan bentuk data yang terdapat pada data dimana fitur atau kolom yang terdapat dalam data terdiri dari *author, content, date, languange, location, number of likes, number of retweet, in reply to, author name, author description, statuses count, favourites count, friends count, followers count, listed count,* dan *author verified*.

**Tabel 2.** Fitur data *tweets* hasil proses *scraping*

|  |  |
| --- | --- |
| **Label** | **Penjelasan** |
| Author | *Username* penulis |
| Content | Isi *tweet* |
| Date | Tanggal *tweet* dibuat |
| Languange | Bahasa |
| Location | Lokasi *tweet* dibuat |
| Number of likes | Jumlah *like* terhadap *tweet* |
| Number of retweet | Jumlah *retweet* terhadap *tweet* |
| In reply to | *Username* tujuan balasan *tweet* |
| Author name | Nama penulis |
| Author description | Deskripsi penulis |
| Statuses count | Jumlah *tweet* milik penulis |
| Favourites count | Jumlah favorit *tweet* milik penulis |
| Friends count | Jumlah teman penulis |
| Followers count | Jumlah pengikut penulis |
| Listed count | Jumlah akun yang diikuti |
| Author verified | Status verifikasi akun penulis |

Dari beberapa fitur tersebut, dilakukan proses pembuangan fitur-fitur atau kolom yang tidak diperlukan dimana dalam proses analisis sentimen data yang diperlukan hanya berupa konten atau tweets yang terdapat dalam data tersebut, jumlah pengguna yang menyukai tweets dan jumlah retweet yang dilakukan terhadap tweet tersebut.

1. **Preprocessing Data**

Sebelum dilakukannya proses klasifikasi data teks, perlu dilakukan proses *preprocessing data* atau pra-pemrosesan data karena data yang sebelumnya telah diperoleh dengan proses *crawling* tersebut belum siap untuk digunakan dalam proses klasifikasi yang dikarenakan oleh datamasih tidak terstruktur dan banyak terdapat *noise.* Data tweets yang diperoleh terdapat tanda-tanda baca, angka, tautan (link) serta kata-kata yang tidak penting sehingga perlu dihilangkan untuk mengurangi volume kosakata sehingga data menjadi lebih terstruktur. Tahap preprocessing data diawali dengan proses *spelling* *normalization*, yang kemudian dilanjutkan ke proses *case folding, tokenizing* hingga tahapan *filtering.* Tahapan preprocessing data pada penelitian ini dilakukan dengan menggunakan aplikasi orange dan *script* yang dibuat dengan menggunakan bahasa pemrograman *python.*

1. ***Spelling normalization* (Normalisasi pengejaan)**

Proses spelling normalization adalah tahapan awal yang perlu dilakukan dalam proses preprocessing data dimana kata-kata yang terdapat dalam data teks dilakukan perbaikan pengeejaan terhadap kata-kata yang salah eja atau kata yang disingkat dalam bentuk tertentu. Data tweets yang didapat juga sering terdapat campuran bahasa inggris sehingga perlu dilakukan proses translasi ke bahasa indonesia. Contoh hasil proses normalisasi pengejaan dapat dilihat pada **Tabel 3**.

**Tabel 3.** Proses *spelling normalization*

|  |  |
| --- | --- |
| **Data input** | **Data hasil normalisasi** |
| **Thank you** @ruangguru @adamasbelva semoga sukses terus **dlm** memajukan dunia pendidikan di indonesia  | **Terima kasih** @ruangguru @adamasbelva semoga sukses terus **dalam** memajukan dunia pendidikan di indonesia  |

1. ***Case folding***

*Case folding* adalah proses penyeragaman bentuk huruf dimana dalam proses ini hanya menerima huruf latin dari ‘a’ hingga ‘z’ dalam bentuk huruf kecil secara keseluruhan. Hal ini dilakukan dengan tujuan agar kata yang memiliki huruf kapital tidak terdeteksi memiliki arti yang berbeda dengan kata yang sama namun tidak memiliki huruf kapital. Contoh hasil proses case folding dapat dilihat pada **Tabel 4.**

**Tabel 4.** Proses *case folding*

|  |  |
| --- | --- |
| **Data input** | **Data hasil normalisasi** |
| **Terima** kasih @ruangguru @adamasbelva semoga sukses terus dalam memajukan dunia pendidikan di indonesia | **terima** kasih @ruangguru @adamasbelva semoga sukses terus dalam memajukan dunia pendidikan di indonesia |

1. ***Tokenizing***

*Tokenizing* atau tokenisasi adalah proses pemisahan antar kata pada suatu dokumen teks menjadi kata yang tidak bergantung dengan kata lain (independen) yang akan dijadikan sebagai entitas yang memiliki nilai dalam penyusunan matriks dokumen teks pada proses selanjutnya serta memiliki frekuensi kemunculan kata tersebut dalam sebuah corpus atau data teks.

**Tabel 5.** Proses *tokenizing*

|  |  |
| --- | --- |
| **Data input** | **Data hasil normalisasi** |
| terima kasih @ruangguru @adamasbelva semoga sukses terus dalam memajukan dunia pendidikan di indonesia | terima, kasih, @ruangguru, @adamasbelva, semoga, sukses, terus, dalam, memajukan, dunia, pendidikan, di, indonesia |

1. ***Filtering***

*Filtering* atau tahap penyaringan merupakan tahap yang dilakukanya proses pemilihan kata pada suatu dokumen yang bertujuan untuk mengurangi dimensi atau jumlah kata yang terdapat dalam corpus atau data teks. Proses ini dilakukan dengan menggunakan stopwords yang merupakan kata-kata yang tidak penting atau tidak terlalu berpengaruh dalam yang sering muncul dalam suatu dokumen. Kata-kata yang dimaksud berupa kata penghubung, kata ganti orang, kata ajakan, dan kata lainnya yang tidak memiliki arti dalam penentuan kelas suatu data teks.

**Tabel 6.** Proses Filtering

|  |  |
| --- | --- |
| **Data input** | **Data hasil normalisasi** |
| terima, kasih, @ruangguru, @adamasbelva, semoga, sukses, terus, dalam, memajukan, dunia, pendidikan, di, indonesia | terima, kasih, sukses, maju, dunia, didik |

Dari hasil proses preprocessing data yang telah dilakukan, dapat digunakan fitur *word cloud* yang terdapat dalam aplikasi *orange* yang dapat menggambarkan kata apa yang paling sering muncul dalam data yang ada yang mana data tersebut akan digunakan untuk proses klasifikasi. Frekuensi kata yang paling sering muncul dapat dilihat pada G**ambar 8**.



**Gambar 8.** Kata yang paling sering muncul

1. **Pelabelan Data**

Setelah data dilakukan proses preprocessing, dilanjutkan dengan melakukan analisis sentimen yang bertujuan untuk memberikan label terhadap setiap data. Proses pelabelan data dilakukan secara otomatis menggunakan kamus lexicon yang terdiri dari kamus kata positif dan kamus kata negatif yang kemudian dihitung nilai sentimen dari data dengan menghitung nilai rata-rata dari kata positif dengan kata negatif yang terdapat dalam data. Semakin banyak jumlah kata positif yang terdapat dalam dokumen, maka semakin tinggi nilai sentimen dari data tersebut. Dan sebaliknya semakin banyak kata negatif yang muncul dibandingkan dengan kata positif, maka semakin rendah nilai sentimen data tersebut.

$$Nilai sentimen=\frac{jumlah kata positif+jumlah kata negatif}{total kata}\*100\%$$

 Setelah mendapatkan nilai sentimen dari masing-masing data, selanjutnya dilakukan proses pengelompokkan data kedalam kelas positif, negatif, dan netral. Jika nilai sentimen dari suatu teks lebih kecil dari 0, maka data tersebut akan dikelompokkan kedalam kelas negatif. Jika nilai sentimen dari suatu teks sama dengan 0, maka data tersebut akan dikelompokkan kedalam kelas netral. Dan jika nilai sentimen dari suatu teks lebih besar dari 0, maka data tersebut akan dikelompokkan kedalam kelas negatif. Adapun hasil pelabelan data yang dilakukan diperoleh jumlah data seperti pada **Tabel 7.**

**Tabel 7.** Perbandingan jumlah data tweet pada kelas sentimen

|  |  |
| --- | --- |
| **Sentimen** | **Jumlah Tweet** |
| Positif | 1680 |
| Netral | 341 |
| Negatif | 238 |

Berdasarkan dari **Tabel 7,** hasil pelabelan kelas sentimen memperlihatkan bahwa jumlah data tweet yang tergabung kedalam kelas sentimen positif memiliki frekuensi yang lebih tinggi dibandingkan dengan jumlah data tweet pada kelas netral maupun negatif dimana jumlah kelas positif sebanyak 1680 tweet (75%), kelas netral sebanyak 341 tweet (15%) dan kelas negatif hanya sebanyak 238 tweet (10%).

* + 1. **Simulasi Perhitungan Nilai Sentimen**

Berdasarkan data tweet “enak nyaman semangat belajar hebat guru asyik”, terdapat 5 kata yang terdeteksi sebagai kata yang bersifat positif dan tidak terdapat kata yang bersifat negatif seperti yang digambarkan pada **Tabel 8.**

**Tabel 8.** Simulasi perhitungan nilai sentimen

|  |  |  |
| --- | --- | --- |
| **Tweet** | **Kata Positif** | **Kata Negatif** |
| enak nyaman semangat belajar hebat guru asyik | Enak, nyaman, semangat, hebat, asyik | **-** |

Dengan menggunakan rumus menghitung nilai sentimen, didapatlah perhitungan sebagai berikut:

* Kata positif = 5
* Kata negatif = 0
* Jumlah kata = 7

$$Nilai sentimen=\frac{Positif+Negatif}{Jumlah Kata}\*100\%$$

$$Nilai sentimen=\frac{5+0}{7}\*100\%$$

$$Nilai sentimen=71,4 \%$$

Skor akhir yang didapatkan dari proses simulasi perhitungan nilai sentimen yang dilakukan diatas memiliki nilai positif sebesar 71,4%, sehingga dapat dikelompokkan kedalam kelas sentimen positif.

1. **Pembagian Data Latih dan Data Uji**

Data latih digunakan pada algoritma klasifikasi dalam membentuk sebuah model yang mana model tersebut merupakan sebuah penggambaran atau representasi suatu pengetahuan yang akan digunakan untuk memprediksi kelas data baru yang belum memiliki kelas. Semakin banyak jumlah data latih yang digunakan, maka mesin klasifikasi akan semakin baik dalam memahami pola data. Data uji digunakan untuk mengukur sejauh mana mesin klasifikasi berhasil melakukan klasifikasi dengan benar dan akurat. Data yang digunakan sebagai data latih dan data uji merupakan data yang telah memiliki label kelas, dimana berdasarkan *Paretto Principle* perbandingan jumlah data latih dan data uji yang paling umum digunakan adalah sebesar 80% : 20%. Proses pembagian data latih dan data uji yang dilakukan pada penelitian ini dilakukan langsung oleh mesin pada aplikasi Orange. Perbandingan data latih dan data uji masing-masing kelas dapat dilihat pada **Tabel 9.**

**Tabel 9.** Perbandingan Data latih dan Data uji

|  |  |  |  |
| --- | --- | --- | --- |
| **Kelas** | **Jumlah** | **Data Latih** | **Data Uji** |
| Positif | 1680 | 1344 | 335 |
| Netral | 341 | 272 | 68 |
| Negatif | 238 | 190 | 48 |
| **Total** | **2259** | **1808** | **451** |

1. **Proses Klasifikasi Menggunakan SVM**

Model klasifikasi masing-masing algoritma *machine learning* didapat berdasarkan hasil dari proses pelatihan data yang telah dilakukan. Dimana model-model tersebut akan dilakukan proses pengujian untuk mendapatkan tingkat akurasi mengenai sejauh mana model tersebut dapat melakukan proses klasifikasi data yang diujikan. Dalam penelitian ini dilakukan percobaan menggunakan empat kernel SVM yang terdapat pada aplikasi *Orange* yaitu adalah kernel *RBF, Linear, Polynomial, dan Sigmoid* untuk melihat kernel mana yang memiliki tingkat akurasi yang paling tinggi dari proses klasifikasi yang dilakukan.



**Gambar 9.** Alur proses klasifikasi pada aplikasi Orange

Proses klasifikasi yang digambarkan pada **Gambar 9.** dilakukan menggunakan *software orange* diawali dengan menginputkan corpus atau file data. Kemudian menentukan fitur yang akan digunakan sebagai target kelas klasifikasi menggunakan *widget select column* yang nantinya data tersebut dihubungkan dengan *Bag-of-Word* yang selanjutnya dihubungkan dengan widget data sampler dimana data dibagi menjadi data latih dan data uji sebelum dilakukannya proses klasifikasi. Teknik proses klasifikasi dilakukan dengan teknik membuat *machine learning* menggunakan data latih dan data uji secara acak dengan melakukan lima kali perulangan. Adapun hasil dari proses klasifikasi yang dilakukan ditunjukkan pada **Tabel 9.**

**Tabel 9.** Perbandingan akurasi kernel pada klasifikasi SVM

|  |  |
| --- | --- |
| **Kernel** | **Akurasi** |
| RBF | 72.1% |
| Linear | 76.7% |
| Polynomial | 74.1% |
| Sigmoid | 74.5% |

 Dari **Tabel 9.** menggambarkan bahwa dari keempat kernel klasifikasi SVM yang digunakan, kernel *linear* memiliki tingkat akurasi yang relatif lebih besar dibandingkan dengan metode kernel lainnya dengan tingkat akurasi sebesar 76.7%. Maka dari itu pada penelitian ini digunakan kernel linear dalam melakukan proses klasifikasi.

*Confusion matrix* merupakan salah satu *tools* yang penting dalam metode evaluasi yang pada *machine learning* yang biasanya memuat dua atau lebih kategori. Setiap unsur matriks menggambarkan jumlah data uji terhadap kelas sebenarnya dalam bentuk baris, sedangkan kelas yang diprediksi digambarkan dalam bentuk kolom (Praptiwi, 2018).

**Tabel 10.** *Confusion matrix* hasil klasifikasi

|  |
| --- |
| **Prediksi** |
| **Data Asli** |  | **Negatif** | **Netral** | **Positif** | **Jumlah** |
| **Negatif** | 23 | 11 | 14 | **48** |
| **Netral** | 10 | 10 | 48 | **68** |
| **Positif** | 6 | 16 | 313 | **335** |
| **Jumlah** | **39** | **37** | **375** | **451** |

Berdasarkan dari **Tabel 10.**, dapat digambarkan seberapa banyak jumlah data yang diprediksi kedalam kelas yang benar. Besarnya data klasifikasi berkategori negatif yang mampu diprediksi dengan tepat kedalam kelas negatif disebut *true negatif.* Besarnya data klasifikasi berkategori netral yang mampu diprediksi dengan tepat kedalam kelas netral disebut *true netral.* Besarnya data klasifikasi berkategori positif yang mampu diprediksi dengan tepat kedalam kelas positif disebut *true positif.* Banyaknya data observasi yang berkategori positif namun terdapat kesalahan prediksi disebut *false positif*. Banyaknya data observasi yang berkategori netral namun terdapat kesalahan prediksi disebut *false netral*. Banyaknya data observasi yang berkategori negatif namun terdapat kesalahan prediksi disebut *false negatif.*

Dari hasil klasifikasi yang dilakukan pada kelas negatif, dari total 48 *tweet* negatif yang diuji, terdapat 23 *tweet* yang telah diklasifikasikan dengan benar dan terdapat kesalahan prediksi sebesar 11 *tweet* yang masuk kedalam kelas netral dan 14 *tweet* masuk kedalam kelas positif. Pada kelas netral, dari total 68 *tweet* netral yang diuji, terdapat 10 *tweet* yang telah diklasifikasikan dengan benar dan terdapat kesalahan prediksi sebesar 10 *tweet* yang masuk kedalam kelas negatif dan 48 *tweet* masuk kedalam kelas positif. Sedangkan pada kelas positif, dari total 335 *tweet* positif yang diuji, terdapat 313 *tweet* yang telah diklasifikasikan dengan benar dan terdapat kesalahan prediksi sebesar 6 *tweet* yang masuk kedalam kelas negatif dan 16 *tweet* masuk kedalam kelas netral. Dari nilai confusion matrix yang didapat tersebut dapat diketahui bahwa dari 451 data *tweet* yang diuji, terdapat 346 *tweet* yang telah diklasifikasikan kedalam kelas yang benar dimana tingkat akurasi dari proses klasifikasi yang telah dilakukan sebesar 76,7%. Dimana didapat dihitung tingkat akurasinya dengan menggunakan rumus berikut :

$$Tingkat Akurasi=\frac{jumlah data yang diprediksi benar}{jumlah data yang diuji}\*100\%$$

Sehingga diperoleh perhitungan sebagai berikut :

$$Tingkat Akurasi=\frac{23+10+313}{48+68+335}\*100\%$$

$$Tingkat Akurasi=\frac{346}{451}\*100\%$$

$$Tingkat Akurasi=76,7\%$$

1. **Visualisasi hasil klasifikasi**

Visualisasi perlu dilakukan terhadap masing-masing kelas klasifikasi yang bertujuan untuk mengekstraksi informasi seperti topik yang paling sering diperbincangkan serta melihat kata apa yang paling sering muncul dalam perbincangan tersebut. Tahapan visualisasi hanya dilakukan untuk kelas sentimen positif dan sentimen negatif, dikarenakan data yang terdapat dalam kelas netral cenderung tidak memiliki arti. Berikut penjelasan hasil proses visualisasi dari kelas positif dan kelas negatif.

1. **Visualisasi sentimen positif**

Data sentimen positif yang digunakan adalah data hasil proses pelabelan yang dilakukan menggunakan kamus *lexicon.* Proses ekstraksi informasi yang dilakukan terhadap data sentimen positif dilakukan secara berulang untuk mendapatkan informasi mengenai *tweet* yang memiliki sentimen positif yang diidentifikasi dengan berdasarkan kemunculan atau frekuensi kata yang memiliki sentimen positif yang paling sering muncul. Visualisasi hasil proses ekstraksi informasi yang didapat tersebut dapat dilihat pada **Gambar 10.**

****

**Gambar 10.** Kata positif yang paling sering muncul

Dari hasil klasifikasi sentimen positif dari *tweets* mengenai Ruangguru sebanyak 1680 *tweets,* didapatkan kata yang paling sering muncul diantaranya yaitu kata “bisa” dengan jumlah frekuensi kemunculan sebanyak 607 kali, kata “gratis” sebanyak 544 kali, kata “kasih” sebanyak 399 kali, dan seterusnya. Kata-kata yang tampil pada **Gambar 10.** Merupakan kata-kata yang memiliki sentimen positif yang menjadi topik pembicaraan yang paling banyak digunakan oleh pengguna twitter terhadap Ruangguru. Kumpulan kata-kata yang sering muncul tersebut dapat divisualisasikan dengan berbentuk *word cloud* dimana visualisasi *word cloud* dapat memberikan gambaran yang lebih jelas mengenai topik atau kata positif yang sering muncul. Semakin sering sebuah kata digunakan, maka kata tersebut semakin besar ditampilkan pada *word cloud, s*eperti yang terlihat pada **Gambar 11.**



**Gambar 11.** Visualisasi kata positif pada *word cloud.*

1. **Visualisasi sentimen negatif**

Proses ekstraksi pada *tweet* yang memiliki sentimen negatif dilakukan secara berulang dalam mendapatkan informasi terhadap kata negatif pada *tweet* mengenai ruangguru yang paling sering muncul. Dari total 2259 *tweet* yang digunakan pada penelitian, sebanyak 238 *tweet* tergolong kedalam kelas sentimen negatif. Pada hasil klasifikasi terhadap *tweet* yang tergabung dalam kelas sentimen negatif terdapat kata yang sering muncul dimana kata “salah” digunakan sebanyak 42 kali, kata “belajar” muncul sebanyak 41 kali, kata “lupa” digunakan sebanyak 24 kali dan seterusnya yang dapat dilihat pada **Gambar 12.**

****

**Gambar 12.** Kata negatif yang paling sering muncul

 Kata-kata yang muncul merupakan kata yang terdapat dalam *tweet* yang memiliki sentimen negatif yang merupakan topik pembicaraan yang paling banyak diperbincangkan terhadap aplikasi atau layanan ruangguru. Kumpulan kata-kata yang muncul tersebut dapat digambarkan menggunakan *word cloud* seperti pada **Gambar 13.** Visualisasi menggunakan *word cloud* memberikan gambaran yang lebih jelas tentang topik atau kata-kata yang paling sering digunakan pengguna twitter terhadap aplikasi atau layanan ruangguru.



**Gambar 13.** Visualisasi kata negatif pada *word cloud.*

**Tabel 11.** Kata yang sering muncul berdampingan dengan kata kunci

|  |  |  |
| --- | --- | --- |
| No. | Kata kunci | Kata yang sering muncul |
| 1 | Belajar | Aplikasi |
|  |  | Salah |
|  |  | Pakai |
| 2 | Layanan | Salah |
|  |  | Kesal |
|  |  | Masalah |
| 3 | Materi | Pakai |
|  |  | Bingung |
|  |  | Buku |
| 4 | Aplikasi | Belajar |
|  |  | Buka |
|  |  | Bingung |
| 5 | Handphone | Unggah |
|  |  | Lambat |
|  |  | Kesal |
| 6 | Langganan | Aplikasi |
|  |  | Takut |
|  |  | Sekolah |
| 7 | Beli | Tipu |
|  |  | Kuota |
|  |  | Krisis |

 **Tabel 11.** Menampilkan kata-kata yang sering muncul yang berdampingan dengan kata kunci seperti belajar, layanan, materi, aplikasi, handphone dan langganan. Berdasarkan dari tabel tersebut dapat diperoleh beberapa informasi sebagai berikut.

 Kata-kata yang sering muncul yang berdampingan dengan kata “belajar” pada tweet bersentimen negatif memberikan informasi mengenai pengguna yang salah atau bingung dalam memakai aplikasi. Dan kata yang sering muncul berdampingan dengan kata kunci “layanan” memberikan informasi mengenai pengguna yang merasa kesal terhadap layanan yang diberikan atau tidak sesuai dengan apa yang diharapkan oleh pengguna.

 Kata-kata yang sering muncul yang berdampingan dengan kata “materi” pada sentimen negatif memberikan informasi mengenai pengguna yang bingung terhadap pelajaran ataupun sumber materi yang digunakan pada proses pembelajaran yang dilakukan menggunakan aplikasi ruangguru. Dan kata yang sering muncul berdampingan dengan kata kunci “langganan” memberikan gambaran informasi mengenai pengguna yang takut jika materi yang dipelajari tidak sesuai dengan pelajaran disekolah yang menjadikan pengguna takut atau ragu untuk berlangganan.

 Kata yang sering muncul berdampingan dengan kata “aplikasi” memberikan informasi tentang pengguna yang kebingunan dalam membuka dan menggunakan aplikasi ruangguru. Kata yang sering muncul berdampingan dengan kata “*handphone*” memberikan informasi mengenai pengguna yang kesal akibat *handphone* atau perangkat yang digunakan menjadi lambat setelah mengunggah aplikasi atau saat menggunakan aplikasi. Dan kata yang paling sering muncul berdampingan dengan kata “beli” memberikan informasi mengenai beberapa pengguna yang tertipu saat melakukan pembelian dari pihak ketiga serta pengguna yang memiliki kuota yang sedikit untuk menggunakan aplikasi.

**4.7** **Diagram *Fishbone***

Berdasarkan dari hasil klasifikasi terhadap *tweet* yang memiliki nilai sentimen negatif, maka dapat diperoleh informasi mengenai masalah yang dibahas atau diperbincangkan terhadap aplikasi atau layanan Ruangguru.



**Gambar 14.** Diagram fishbone sentimen negatif terhadap ruangguru

 Dari **Gambar 14.** dapat ditarik informasi mengenai faktor-faktor yang menyebabkan layanan pembelajaran ruangguru mempunyai sentimen negatif dari segi *price, people, process, promotion, place,* dan *product.* Langkah selanjutnya yaitu menentukan bagaimana memecahkan permasalahan tersebut. Adapun rencana pemecahan masalah terhadap layanan ruangguru dapat digambarkan sebagai berikut.

**Tabel 12.** Solusi pemecahan masalah layanan ruangguru.

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Faktor | Permasalahan | Solusi |
| 1 | *Price* | Takut untuk berlangganan | Memperjelas proses berlangganan serta harga yang ditawarkan |
| Kuota kritis | Bekerja sama dengan penyedia layanan internet untuk memberikan bantuan kuota |
| 2 | *People* | Kesal terhadap layanan yang diberikan | Meningkatkan kualitas layanan yang diberikan |
| Pengajar kurang kompeten | Memberikan pelatihan terhadap tenaga pengajar |
| 3 | *Process* | Perangkat lambat saat menggunakan aplikasi | Membuat versi *lite* (ringan) dari aplikasi |
| Bingung menggunakan aplikasi | Memperbaiki *UI/UX* dari aplikasi |
| 4 | *Promotion* | Penipuan pembelian pihak ketiga | Memantau lebih ketat penjual/penyedia layanan ilegal |
| 5 | *Place* | Akses server lambat | Melakukan *maintenance* secara rutin atau meng-*upgrade* *server* |
| 6 | *Product* | Materi belajar tidak sesuai sekolah | Menyesuaikan materi pelajaran dengan kurikulum sekolah yang berlaku |
| Bingung terhadap materi pelajaran | Menyediakan *soft copy* atau sumber yang digunakan dalam pelajaran |
| Materi pelajaran sulit dipahami | Memperbaiki teknik mengajar dari tenaga pengajar |

**4.8** **Pembahasan Hasil Penelitian**

 Dari hasil dari penelitian yang dilakukan dapat diketahui bahwa arah sentimen dari pengguna twitter terhadap *brand* atau aplikasi ruangguru cenderung mengarah kepada sentimen positif yang dapat dilihat dari perbandingan sentimen positif terhadap sentimen negatif dan netral pada dataset sebanyak 2259 data *tweet*.Akan tetapi proses pengelompokkan data kedalam kelas yang menggunakan kamus lexicon terdapat beberapa kekurangan dimana jika *tweet* terdapat ekspresi positif akan tetapi menggunakan kata negatif seperti kata ‘gila’, metode pengelompokkan data tidak dapat menentukan sentimen yang akurat dari tweet tersebut.

Proses pengumpulan data yang dilakukan menggunakan API twitter dengan akun yang bersifat tidak berbayar atau gratis dimana pihak twitter membatasi akses terhadap data tweet sehingga menyebabkan tidak memungkinkan untuk mengakses data tweet dalam jumlah yang lebih besar sehingga menyebabkan jumlah data yang didapat relatif kecil sehingga penelitian yang dilakukan tidak dapat memberikan hasil yang lebih akurat.

Metode SVM dalam mengklasifikasikan data yang digunakan telah dapat mengklasifikasikan data dengan cukup baik dimana tingkat akurasi yang didapatkan relatif sama terhadap penelitian sebelumnya dilakukan oleh Novintirani et al (2020) yang menganalisis sentimen mengenai transportasi umum pada media sosial twitter.

Dari data tweet yang telah dianalisis dapat dilihat bahwa informasi yang terkandung dalam masing-masing tweet tersebut cenderung hanya memuat ekspresi yang tidak konstruktif dan tanpa alasan mengapa ekspresi tersebut diberikan kepada pihak ruangguru. Sehingga proses pengumpulan informasi khususnya ekspresi yang bersentimen negatif kurang efektif.