

Analisis Sentimen Ulasan Aplikasi Spotify Untuk Peningkatan Layanan Menggunakan Algoritma Naive Bayes

Sentiment Analysis of Spotify Application Reviews for Service Improvement Using Naive Bayes Algorithm

Mochamad Daffa Rhajendra¹, Nurvita Trianasari²

^{1,2} Universitas Telkom, Bandung

drhajendra@student.telkomuniversity.ac.id¹, nurvitatrianasari@telkomuniversity.ac.id²

Abstrak

Perkembangan medium untuk mendengarkan musik yang di iringi dengan perkembangan teknologi, tidak dapat dipungkiri semakin mengalihkan orang-orang untuk mendengarkan musik melalui perangkat digital. Dengan terjadinya peralihan mendengarkan musik kepada medium digital, hal ini memunculkan peluang bagi para pengembang aplikasi untuk menciptakan produk penyedia layanan musik.

Salah satu penyedia jasa streaming musik, podcast dan video adalah Spotify yang dapat di akses secara digital, melalui platform iOS, Android dan PC. Pada saat ini Spotify menjadi streaming platform yang paling populer di dunia dengan jumlah pengguna spotify sebanyak 345 juta. Kualitas layanan dari sebuah produk yang dimana pada penelitian ini adalah platform musik digital dapat menjelaskan apa-apa saja kebutuhan bagi penggunaannya, sehingga kualitas layanan yang diberikan dapat menjadi parameter yang dapat di kembangkan oleh pihak Spotify.

Melalui penelitian ini akan dilakukan analisis sentimen ulasan pada aplikasi spotify untuk peningkatan layanan dengan menggunakan algoritma Naive Bayes. Metode pengumpulan data dilakukan melalui metode scrapping ulasan yang tersedia pada google play dengan 3600 data. Teknik analisis data yang digunakan dalam penelitian ini adalah melalui algoritma naive bayes dan 5 dimensi e-Servqual. Dengan hasil yang di dominasi sentiment negatif, dan kualitas layanan pada dimensi *reliability*, *responsiveness*, dan *app design* yang perlu di perbaiki.

Kata Kunci : Spotify, Analisis Sentimen, Naive bayes, e-Servqual

Abstract

The development media for listening to music in accordance with technological developments, can not be ignore as people now listen to music through digital devices. By listening to music to this digital medium, it creates opportunities for applications to create music service provider products.

One of the providers of music, podcast and video streaming services is Spotify, which can be accessed digitally, via the iOS, Android and PC platforms. Currently, Spotify is the most popular streaming platform in the world with 345 million Spotify users. The service quality of a product which in this research is a digital music platform can explain what the needs are for its users, so that the quality of the services provided can be a parameter that can be developed by Spotify.

Through this research, sentiment analysis will be carried out on the Spotify application for service improvement using the Naive Bayes algorithm. The data collection method is done through the scrapping review method that available on Google Play with 3600 data. The data analysis technique used in this research is the nave Bayes algorithm and the 5-dimensional e-Servqual. With results that are dominated by negative sentiments, and service quality on the dimensions of reliability, responsiveness, and app design that needs to be improved.

Key Words : Spotify, Analisis Sentimen, Naive bayes, e-Servqual

1. Pendahuluan

Musik yang merupakan kegemaran bagi seluruh kalangan manusia, sering kali di gunakan sebagai suatu alat untuk mewakili emosi atau perasaan pendengarnya. Kegiatan mendengarkan musik telah menjadi kegiatan sehari – hari yang dapat di lakukan setiap saat [5]. Berawal dari cara

konvensional seperti mendengarkan konser secara langsung, berubah menjadi piringan hitam, berubah menjadi kaset, kemudian berbentuk cd, dan berkembang kedalam platform *streaming* digital seperti *spotify*, *apple music*, dan *deezer* yang dapat di putar melalui gawai-gawai yang terkoneksi dengan internet.

Perkembangan medium untuk mendengarkan musik yang di iringi dengan perkembangan teknologi, tidak dapat dipungkiri semakin mengalihkan orang-orang untuk mendengarkan musik melalui perangkat digital. Dengan terjadinya peralihan mendengarkan musik kepada medium digital, hal ini memunculkan peluang bagi para pengembang aplikasi untuk menciptakan produk penyedia layanan musik. Perkembangan pasar dari penyedia layanan *streaming* musik di prediksi dari tingkat pertumbuhan tahunan majemuk akan bertumbuh sebesar 7,47 dollar sepanjang tahun 2021-2025 atau sebesar 19%. Pertumbuhan ini di dorong dengan berkembangnya pengeluaran iklan yang di fokuskan pada gawai yang akan mendorong pertumbuhan pasar[3]. Selain itu pada tahun 2020 pendapatan *streaming* musik menyumbang 62% dari penjualan rekaman musik atau sebesar 13,4 miliar dollar yang dimana adalah pendapatan tertinggi sepanjang masa[1].

Perkembangan dunia musik ke platform *streaming* digital meningkat tingkat konsumsi dan penggunaannya selama pandemi ini berlangsung. Meningkatkan kualitas layanan dapat dengan memanfaatkan *review* atau ulasan dari pengguna aplikasi Spotify yang terdapat pada Playstore, di dalam ulasan yang di berikan pengguna terdapat banyak pengalaman pengguna aplikasi yang baik maupun yang buruk, positif maupun negatif, atau sentimen penggunaannya, selain itu ulasan dari pengguna juga dapat memberikan pengaruh bagi calon pengguna[4]. *Google playstore* merupakan platform yang memberikan layanan konten digital toko aplikasi online milik Google, terdapat produk-produk seperti musik, buku, dan aplikasi. Layanan Google Play dapat diakses melalui tiga cara yaitu melalui web, aplikasi android (PlayStore), dan Google TV. Dengan jumlah download dan ulasan dari 22 juta orang pengguna, Spotify mendapatkan *rating* sebesar 4,4/5. Oleh karena itu menggunakan ulasan yang terdapat pada Googleplay dapat memberikan *insight* mengenai kebutuhan dan keinginan konsumen untuk mengembangkan layanan dan kepuasan konsumen pengguna *spotify*.

2. Dasar Teori dan Kerangka Pemikiran

2.1 Pengertian Manajemen Pemasaran

Manajemen pemasaran adalah gabungan dari seni dan ilmu pengetahuan untuk menentukan target market dan mendapatkannya, mempertahankan, serta mengembangkan hubungan dengan kostumer melalui menciptakan, menyampaikan, dan mengkomunikasikan nilai pelanggan yang unggul[8]. Para pemasar dapat memasarkan barang, jasa, peristiwa, pengalaman, tempat, properti, organisasi, informasi, dan ide di empat pasar yang berbeda yaitu, konsumen, bisnis, global dan nirlaba. Selain gabungan antara seni dan ilmu pengetahuan, manajemen pemasaran juga membangun hubungan yang menguntungkan dengan kostumernya[7].

Manajemen pemasaran adalah seni untuk menumkan apa yang kostumer inginkan dan mampu memberikannya kepada kostumer dengan lebih baik di bandingkan dengan kompetitornya[9].

2.2 Pengertian Kalitas layanan

Kualitas layanan adalah seluruh fitur dan karakterstik sebuah produk atau layanan yang mampu memenuhi atau melebihi standar ekspektasi dari kepuasan dan kebutuhan konsumen[8], yang di berikan penjual atau pemberi layanan melalui produknya. Kualitas layanan adalah suatu kondisi dinamis yang terjadi dalam hubungan antara produk jasa, manusia, dan lingkungan terpenuhi atau melebihi harapan konsumennya[14]. Kualitas layanan sangatlah bervariasi, tergantung dari siapa yang menyediakan, dimana, kapan, dan bagaimana layanan tersebut disediakan[7].

2.3 e-Servqual

Kualitas layanan (*servqual*) adalah model untuk menilai layanan berdasarkan persepsi kinerjanya. Terdapat lima dimensi kualitas layanan umum yaitu *tangibles*, *reliability*, *responsiveness*, *assurance*, dan *empathy*. Model konseptual untuk memahami dan

meningkatkan kualitas layanan dalam bentuk elektronik disebut *e-Servqual*. *e-Servqual* menggabungkan dimensi baru sesuai dengan skala yang diperlukan untuk mengukur kualitas layanan bisnis digital[10].

Tabel 2.1 Dimensi e-Servqual

<i>app design</i>	Daya tarik yang dihadirkan oleh desain antarmuka pengguna untuk pelanggan.
<i>Reliability</i>	kemampuan aplikasi untuk memenuhi pesanan pengiriman dengan benar dan menjaga keamanan informasi pribadi.
<i>Responsiveness</i>	Pelanggan mengharapkan <i>developer</i> untuk segera menanggapi pertanyaan mereka.
<i>Trust</i>	kesediaan pelanggan untuk menerima kerentanan dalam aplikasi dan berdasarkan harapan positif mereka.
<i>Personalization</i>	perhatian individual, dan ketersediaan pesan untuk pertanyaan pelanggan atau komentar.

Sumber: Data diolah penulis (2021)

2.4 Big Data

Big data adalah istilah yang mendeskripsikan volume data yang besar, terstruktur maupun tidak terstruktur. Volume data yang besar itu dapat di ubah menjadi suatu kreasi teknologi yang memfasilitasi perkembangan dan pertumbuhan dari big data itu sendiri[12].

Big data memiliki 3 karakteristik yang di sebut dengan 3 *V's of big data*[2], yaitu :

- Volume (ukuran) : Volume data terus menerus berkembang dari ukuran *terabyte*, ke *petabyte*, kemudian menjadi *zetabyte*, yang menyebabkan di butuhnya penyimpanan yang sangat besar dan dapat di andalkan. Dari volume data yang terus berkembang di gunakan oleh bermacam-macam bisnis untuk, menemukan solusi dari permasalahan yang di alami oleh bisnis tersebut dengan lebih baik.
- Velocity (kecepatan) : Pergerakan dari data sangatlah tinggi dimana proses dan analisa terbentuknya sebuah data terjadi dalam kecepatan yang sangat cepat, dalam 1 menit terdapat 3.3 juta post di facebook, 450 ribu tweet, 400 jam video yang di upload di youtube dan 3.1 juta pencarian dari google.
- Variety (variasi) : variasi dari data di bagi 3 yaitu data yang terstruktur, data yang semi-terstruktur dan, data yang tidak terstruktur. Data yang terstruktur adalah data yang di proses melalui sistem database tradisional seperti, data pegawai, data nasabah bank, dll. Kemudian data yang semi-terstruktur adalah kombinasi dari data yang terstruktur dan tidak terstruktur contohnya adalah XML dan Json, yang di gunakan pada facebook, twitter, dll. Terakhir data yang tidak terstruktur adalah data yang tidak dapat di proses dalam database tradisional seperti excel, word, pdf, jpg dan AVI.

2.5 Text Analytics

Text Analytics adalah proses pengubahan data yang tidak terstruktur menjadi data yang memiliki nilai atau arti, dengan mengolahnya menggunakan *Machine Learning*. *Text Mining* adalah proses pencarian pola dari koleksi teks atau kata – kata yang banyak, sehingga mendapatkan *insight* penting atau informasi yang berkaitan dengan permasalahan yang di hadapi[2].

2.6 Sentiment Analysis

Analisis sentiment atau yang bisa disebut dengan opinion mining adalah proses menentukan pendapat, sikap atau sifat, perasaan dan emosi dari penulis kumpulan kata itu. Analisis sentiment juga dapat menjelaskan dan menentukan perasaan seseorang mengenai suatu

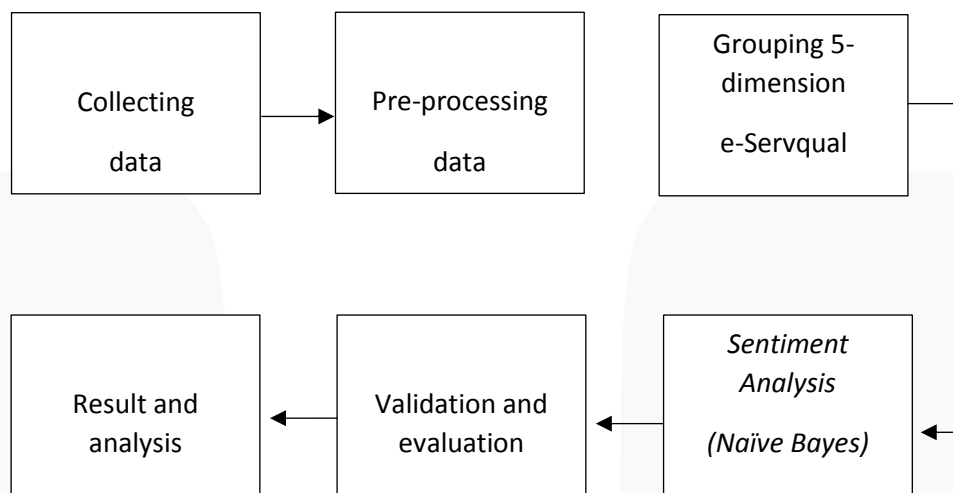
barang dan situasi yang di alaminya. hasil Analisa dari pengolahan tersebut dapat menentukan arti kata yang di tulis oleh seseorang dan memberikan informasi mengenai baik atau buruknya barang atau produk dan kejadian atau situasi tersebut [2].

2.7 Naïve bayes

Naïve Bayes di kemukakan oleh Thomas Bayes yang menciptakan teori Bayes yaitu metode pembelajaran yang diawasi dan ditujukan untuk klasifikasi. Teori tersebut dapat di gunakan untuk memecahkan masalah diagnostik serta memprediksi hasil. Salah satu teknik pengaplikasiannya adalah *Naïve Bayes text classification* dan *spam filtering* pada email [2].

Formula *Naïve Bayes* di gunakan untuk mengkasifikasikan probabilitas yang menghitung bermacam probabilitas dengan memperhitungkan frekuensi dan kombinasi dari nilai pada data set. Teori tersebut menganggap bahwa seluruh variable adalah independent atau berdiri sendiri dan bergantung dari suatu nilai variable kelasnya sendiri[11].

2.8 Kerangka Pemikiran



Gambar 2.1 Kerangka pemikiran

3. Metode Penelitian

Data yang digunakan dalam penelitian ini adalah data hasil *scraping* dari situs Googleplay selama periode 14 februari 2021 – 1 juni 2021. Data di ambil menggunakan google collabs yang kemudian membagi hasil data berdasarkan *rating* yaitu puas dan tidak puas. Data yang terkumpul adalah 3600 dan kemudian di lakukan *cleaning* sebanyak 2.745 ulasan.

3.1 Teknik Analisis Data

1. Pengumpulan data : Pengumpulan data dilakukan dengan melakukan *scraping* menggunakan google collabs Pengambilan data dilakukan secara berulang hingga data yang dibutuhkan terpenuhi. Terdapat empat informasi yang didapatkan dari hasil *scraping* pada google play yaitu review id, usernam, content atau ulasan, dan score yang di berikan dari penulis ulasan.
2. *Preprocessing*: attribute data yang di ambil adalah *review* atau ulasan kemudian di lakukan *cleaning* dan selanjutnya melewati tahapan proses *Case folding*, *punctual removal*, *Tokization*, *Stopword*, dan *stemming*.
3. Klasifikasi dimensi *e-Servqual* : Data yang telah melalui tahapan pre-processing kemudian di klasifikasikan secara manual dengan mengklasifikasikannya terhadap dimensi e-

Servqual. Hal ini dilakukan agar dapat mengukur persepsi kualitas layanan pada ulasan tersebut.

4. Analisis Sentimen : Pengolahan analisis sentimen memerlukan data training dan data testing dari corpus atau kumpulan teks yang digunakan untuk analisis linguistik seperti analisis sentimen. Data training dan data testing pada penelitian ini berasal dari persepsi pengguna spotify pada google play, kemudian dilakukan pemberian label secara manual pada data training yang berisi persepsi positif dan negatif secara manual.
5. Validasi dan Evaluasi: Performa model klasifikasi di ukur dengan menggunakan confusion matrix. confusion matrix yaitu metode yang di gunakan untuk mengukur akurasi pada data. Dimana terdapat 4 nilai yaitu, True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN). Dimana True Positive (TP) diklasifikasikan positif dan menurut machine learning juga positif, True Negative (TN) diklasifikasikan negative, menurut machine learning juga negative, False Positive (FP) diklasifikasikan positif tetapi menurut machine learning negative, dan False Negative (FN) diklasifikasikan negative akan tetapi menurut machine learning di anggap positif.

4. Hasil dan Pembahasan

4.1 Karakteristik data

Data dibagi menjadi *rating* puas dan tidak puas untuk mengetahui alasan pengguna terhadap produk berdasarkan *rating* yang diberikan.

No	Rating	Total Review
1.	Puas (5,4, dan 3)	2400
2.	Tidak Puas (2, dan 1)	1200

4.2 Hasil Pengolahan Data

4.1.1 Hasil performance vector

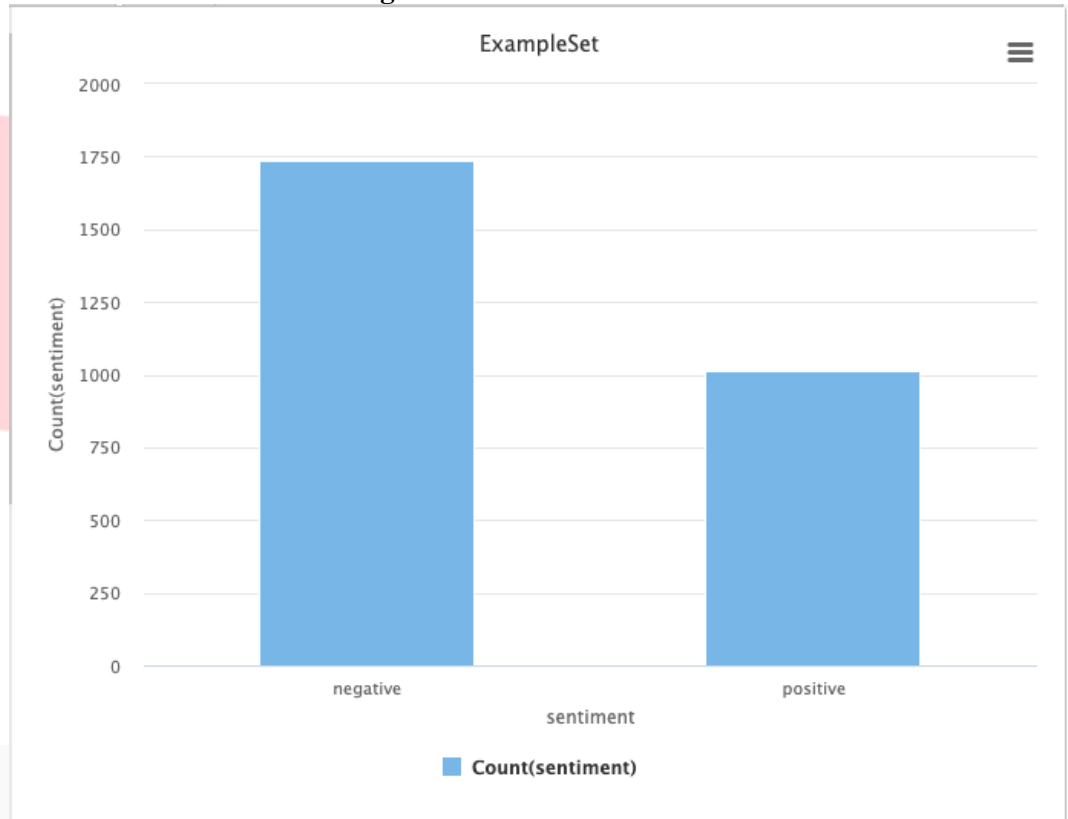
Table View Plot View

accuracy: 74.85% +/- 1.59% (micro average: 74.85%)

	true negative	true positive	class precision
pred. negative	1200	156	88.50%
pred. positive	535	857	61.57%
class recall	69.16%	84.60%	

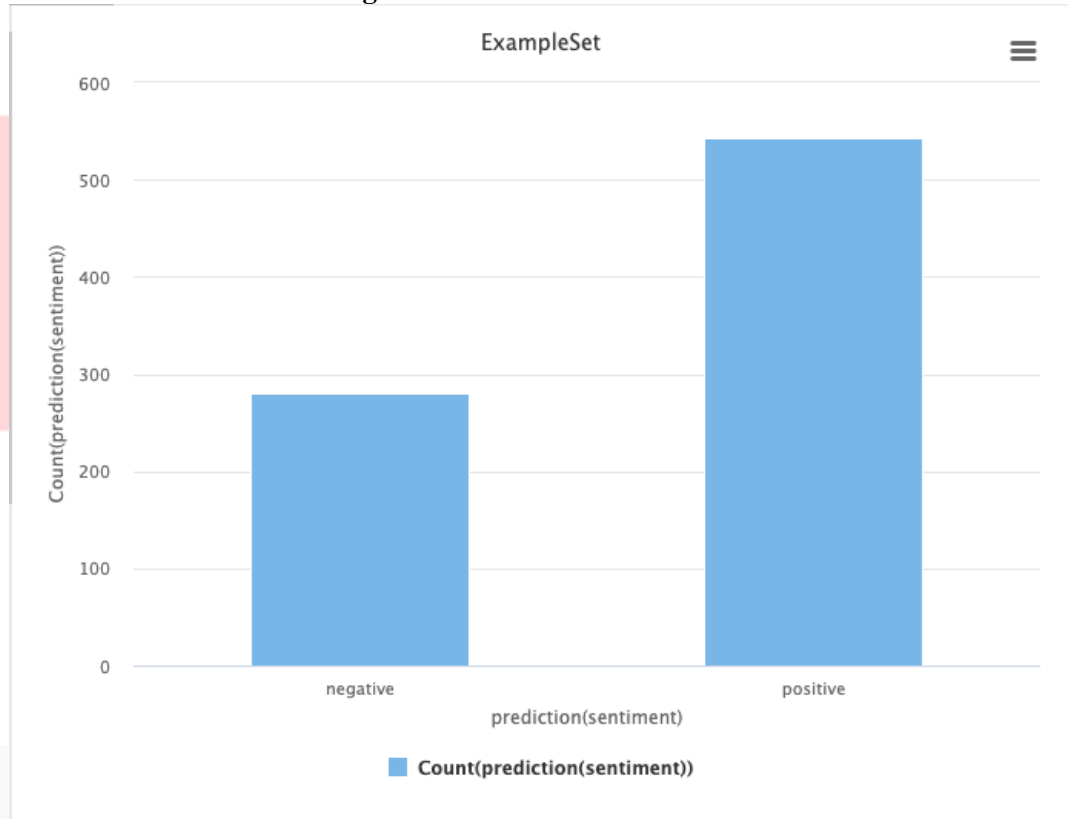
Performance vector di atas menjelaskan bahwa machine learning menyimpulkan bahwa ulasan yang benar - benar negatif sebanyak 1200 dan di anggap positif sebanyak 535 ulasan. Dengan presentasi recall sebesar 69.16% dan presentasi precision sebesar 88.50%. Kemudian untuk ulasan yang benar – benar positif machine learning menyimpulkan bahwa ulasan yang benar – benar positif sebanyak 156 dan yang di anggap negatif sebanyak 857 ulasan. Dengan presentasi recall sebesar 84.60% dan presentasi precision sebesar 61.57%. Untuk accuracy model sentiment analisis spotify dari data ulasan tersebut dapat disimpulkan memiliki nilai sebesar 74.85%.

4.1.2 Grafik sentiment data training



Grafik di atas menunjukkan jumlah sentiment positif dan negatif pada data Training dimana sentiment positifnya lebih tinggi di banding dengan sentiment negatifnya. Untuk sentiment negatif berjumlah 1735 ulasan dan untuk sentiment positifnya berjumlah 1.013 ulasan.

4.1.3 Grafik sentiment data testing



Grafik tersebut menunjukkan jumlah sentiment positif dan negatif pada data testing dimana sentiment positifnya lebih tinggi di banding dengan sentiment negatifnya. Untuk sentiment positifnya berjumlah 543 ulasan dan untuk sentiment negatifnya berjumlah 281 ulasan.

4.1.4 Multiclassification result

Dimensi	Number	Percentage
Reliability	1.201	43.7%
Personatization	851	31%
Responsiveness	389	14.2%
App design	167	6.1%
Trust	138	5%
Total	2.745	100%

Tabel tersebut menunjukkan presentase masing masing dimensi.

4.1.5 Proporsi Sentiment masing – masing Dimensi e-Servqual

Dimensi	Jumlah	Sentiment positif dalam %	Sentiment negatif dalam %
Reliability	1.201	31.8	68.2
Personalization	850	53.6	46.4
Responsiveness	389	5.3	94.6
App Design	167	10.2	89.8
Trust	138	93.4	6.6
Total	2745	194.3	305.6

Tabel diatas menunjukkan hasil multiclass classification dan proporsi sentiment masing – masing dimensi e-Servqual. Dimana pada tabel di atas menjelaskan bahwa terdapat 1.201 dalam dimensi reliability yang menunjukkan perbedaan yang signifikan antara sentiment negatif dan sentiment positif. Dimana sentiment negatif memiliki presentase sebesar 68.2% di bandingkan dengan presentase sentiment positif sebesar 31.8%. Dari kelima dimensi e-Servqual untuk dimensi responsiveness dan app design memiliki nilai sentimen negatif tinggi yang hampir sama, dengan nilai sentiment negatif sebesar 94.6% untuk dimensi responsiveness dan 89.8% untuk dimensi app design. Sedangkan untuk dimensi trust memiliki nilai positif terbesar dengan 93.4%. Kemudian untuk dimensi personalization perbedaan sentimennya tidak terlalu signifikan dengan nilai sentiment positifnya sebesar 53.6% dan sentiment negatifnya sebesar 46.4%.

4.3 Pembahasan hasil penelitian

Hasil tersebut merupakan *costumer insight* yang dapat di gunakan sebagai alat untuk mengidentifikasi kebutuhan secara detail dan dapat di jadikan acuan untuk mendapatkan *competitive advantage* dengan pesaingnya. *consumer insight* adalah meliputi aspek yang luas yang tidak hanya meliputi bagaimana cara konsumen menggunakan sebuah produk tetapi, bagaimana mereka menggunakan media khususnya media digital untuk memberikan *feedback*, kritik dan saran yang dapat menjadi *insight* bagi perusahaan. Konsumen juga tidak hanya sekedar melihat iklan atau *marketing message* sebagai sebuah hal yang di dengar atau di baca saja karena mereka memberikan timbal baliknya. *Insight* yang di dapatkan pada penelitian ini adalah bahwa terdapat beberapa hal yang dapat di lakukan spotify untuk meningkatkan kualitas layanannya seperti memperbaiki aplikasinya dari ke-andalan spotify untuk di gunakan, mengurangi iklan dan memberikan diskon paket berlangganan premium. Kemudian memperbaiki *bug* dan error yang terdapat pada aplikasi, dan terakhir merancang UX dan UI yang lebih sesuai dan dapat di gunakan dengan mudah oleh pelanggannya.

Perusahaan dapat menggunakan database konsumen untuk mengidentifikasi konsumen prospek konsumen dan menentukan penawaran bagi konsumen tertentu. Dimana pada penelitian ini spotify memilih target market penikmat musik yang memiliki gawai android, terkoneksi dengan internet dan membutuhkan atau menginginkan kegiatannya di temani oleh lagu. Kemudian spotify memiliki penawaran tertentu untuk fitur – fitur terbatasnya sesuai dengan segmentasi pasarnya. Dimana spotify memberikan banyak pilihan paket bagi pelanggan yang berminat berlangganan.

5. Kesimpulan

Berdasarkan penelitian di atas, peneliti berhasil menganalisis sentimen dan kualitas layanan dari data tidak terstruktur ulasan mengenai aplikasi spotify menggunakan algoritma naïve bayes dalam situs googleplay dengan kesimpulan sebagai berikut:

1. Analisis sentiment menggunakan algoritma naïve bayes pada penelitian ini berhasil dilakukan dengan hasil di dominasi oleh sentiment negatif, menggunakan aplikasi *machine learning Rapidminer* dengan jumlah data 2.745 data. Dimana terdapat dua kategori data dengan kategori puas dan tidak puas yang di analisis untuk menghasilkan *customer insight* dari aplikasi tersebut untuk meningkatkan layanan.
2. Hasil penelitian ini menunjukkan bahwa *customer insight* pengguna aplikasi spotify memiliki sentiment dan persepsi negatif atau buruk.
3. Spotify perlu memperbaiki serta meningkatkan kinerja aplikasi untuk meningkatkan kualitas layanannya. Seperti memperbaiki aplikasinya dari keandalan spotify untuk di gunakan, mengurangi iklan dan memberikan diskon paket berlangganan premium. Kemudian memperbaiki bug dan error yang terdapat pada aplikasi, dan terakhir merancang UX dan UI yang lebih sesuai dan dapat di gunakan dengan mudah oleh pelanggannya. Selain itu spotify dapat mengidentifikasi konsumen prospek konsumen dan menentukan penawaran bagi konsumen tertentu sesuai dengan segmentasinya.
4. Spotify perlu memperbaiki kualitas layanan personalisasi konsumen, memberikan solusi bagi keluhan konsumen dan merancang aplikasi dengan kesesuaian desain aplikasi yang baik menurut hasil *e-Servqual*.

Referensi :

- [1]Axios. (2021, maret 23). *economy and business*. Retrieved from Axios: <https://www.axios.com/global-streaming-music-revenue-high-bcd41b38-50e3-4605-82d1-0d9952624dcd.html>
- [2]Balusamy, B., Abirami, N., Kadry, S., & Gandomi, A. (2021). *Big Data (Concepts, Technology, and Architecture)*. Hoboken: John Wiley & Sons.
- [3]Business Wire. (20221, March 23). *Research and Market*. Retrieved from Businesswire: <https://www.businesswire.com/news/home/20210323005840/en/Global-Music-Streaming-Market-2021-2025-Market-is-Poised-to-Grow-by-7.47-Billion-Progressing-at-a-CAGR-of-19---ResearchAndMarkets.com>
- [4]Chauhan, C., & Sehgal, S. (2017). SENTIMENT ANALYSIS ON PRODUCT REVIEWS . *International Conference on Computing, Communication and Automation (ICCCA2017)*, 26-30.
- [5]Kompas.com . (2020, 07 21). *Advertorial*. Retrieved from biz.kompas.com: <https://biz.kompas.com/read/2020/07/21/135341828/suka-mendengarkan-musik-berikut-manfaatnya-dalam-kehidupan-sehari-hari>
- [6]Kotler, P. a. (2016). *Prinsip-prinsip Pemasaran* (13nd Edition ed.). Erlangga.
- [7]Kotler, P., & Armstrong, G. (2020). *Principles of Marketing*. Harlow: Pearson Education Limited.
- [8]Kotler, P., & Keller, K. L. (2016). *Marketing Management*. Edinburgh Gate: Pearson Education Limited.
- [9]Ryder, M., & Malik, M. (2020). *Applied Marketing Management*. Mc Caster University.
- [10]Sari. (2018). Measuring e-Commerce Service Quality from Online Customer Review using Sentiment Analysis Case Study: Tokoped. *Journal of Physics: Conference Series*, 1-6.
- [11]Saritas, M. M., & Yasar, A. (2019). Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification. *International Journal of Intelligent Systems and Applications in Engineering*, 88-91.
- [12]Sedkaouri, S., Khelfaoui, M., & Kadi, N. (2021). *BIG DATA ANALYTICS Harnessing Data for New Business Models* . Palm bay: Apple Academic Press.
- [14]Tjiptono, F., & Chandra, G. (2016). *Pemasaran Jasa (Prinsip, Penerapan, dan Penelitian)*. Yogyakarta: Andi.