

Analisis Sentimen Kepuasan Pemain PUBG Berdasarkan Komentar di Media Sosial

Wahit Desta P¹, Fandli Supandi², Yuni Ambar S³

¹Universitas Alma Ata, ²Universitas Muhammadiyah Gorontalo

e-mail: ¹destawahit@gmail.com, ²fandli.Supandi10@gmail.com, ³ambarsetianto@gmail.com

Abstract

Abstrak- Internet telah menjadi bagian dari kehidupan manusia. Internet sering digunakan untuk komunikasi, interaksi sosial, transaksi bisnis, bahkan hiburan. Sejalan dengan perkembangan internet, game di komputer juga berkembang menjadi game online. Salah satu game online yang sedang populer saat ini adalah game battle royale yaitu Player Unknown's Battle Ground (PUBG). Hanya dalam kurun waktu sekitar satu tahun game tersebut terjual dengan jumlah yang fantastis. Dengan adanya game baru ini, perlu dilakukan estimasi kepuasan pengguna game PUBG. Penelitian yang dilakukan dan dibahas adalah tentang analisis sentimen dari penggemar serta pemain game PUBG Mobile melalui komentar di fanpage Facebook. Metode yang digunakan adalah menggunakan klasifikasi dengan algoritma Naive Bayes Classifier menggunakan statistik yang mendasar dalam pengenalan pola (pattern recognition). Pertimbangan antara menghitung trade-off antara berbagai keputusan yang menggunakan, dan probabilitas yang muncul dalam keputusan tersebut. Hasil pengujian dengan menggunakan metode Naive Bayes pada percobaan pertama diperoleh sebesar 41,7% dan pada percobaan kedua dengan jumlah data yang lebih banyak meningkat menjadi 43,429%.

Kata Kunci : Sentimen, Klasifikasi, Naive Bayes.

I. PENDAHULUAN

Semakin maraknya penggunaan jejaring sosial membuat jejaring sosial menjadi data yang sangat besar. Salah satu kegunaan data ini adalah untuk mengetahui pendapat atau sentimen pengguna jejaring sosial di halaman penggemar Facebook. Pemilik game online juga memanfaatkan fanpage ini sebagai sarana komunikasi dengan para pengguna game online serta sebagai sarana untuk menyampaikan pendapat, saran atau keluhan dari pengguna game tersebut.

Keluhan merupakan bentuk ekspresi formal tentang ketidakpuasan atau ketidakpuasan terhadap beberapa aspek yang diterima seseorang. Keluhan sangat dapat diterima perusahaan karena aspek pelayanan yang diberikan belum dapat memuaskan keinginan pelanggan. Oleh karena itu, perusahaan harus siap dengan segala bentuk pengaduan yang dapat diterima sewaktu-waktu. Di era media sosial saat ini, penyampaian pengaduan sangat mudah, pelanggan hanya perlu mengakses media sosial dan berharap pengaduan dapat ditindaklanjuti dengan cepat.. [1]

Analisis sentimen disebut juga opinion mining, merupakan bidang ilmu yang menganalisis pendapat,

sentimen, evaluasi, penilaian, sikap dan emosi masyarakat terhadap entitas seperti produk, jasa, organisasi, individu, masalah, peristiwa, topik, dan atributnya Opini yang dibutuhkan untuk melakukan analisis berasal dari komentar dari halaman fanpage Pubg mobile di Facebook.

Pada kajian analisis sentimen ini dibahas bagaimana penerapan analisis sentimen untuk mengetahui kecenderungan opini publik terhadap game pubg mobile dengan menggunakan metode klasifikasi naive bayes. Naive Bayes merupakan metode machine learning yang memiliki model dalam membentuk probabilitas dan peluang. Oleh karena itu, naive bayes akan menghitung probabilitas kemunculan kata yang menghadirkan komentar berdasarkan kelas positif dan negatif.

II. KAJIAN TERDAHULU

Beberapa penelitian yang telah dilakukan sebelumnya dan dibandingkan dengan penelitian yang akan dilakukan. Penelitian pertama adalah penelitian yang dilakukan oleh Vishal et al. Penelitian ini berfokus pada analisis sentimen terhadap data twitter yang tidak terstruktur, heterogen, positif, negatif, atau netral. Penelitian ini menggunakan metode survey dan analisis komparatif serta menggunakan teknik data mining dengan pendekatan berbasis leksikon. Algoritma yang digunakan adalah Naive Bayes, Max Entropy, dan Support Vector Machines. Hasil yang diperoleh dari penelitian ini adalah metode berbasis leksikon sangat efektif dalam beberapa kasus, terutama data yang telah diawasi. [4]

Penelitian selanjutnya dilakukan oleh Rizal, dkk. Penelitian ini menggunakan kombinasi fitur tekstual dan nontekstual untuk meningkatkan performa prediksi sentimen. Metode yang digunakan adalah Naive Bayes untuk klasifikasi tekstual dan Fisher Score untuk menentukan fitur nontekstual (likes dan retweets). Hasil dari penelitian ini adalah nilai optimal dari α dan β , kinerja evaluasi menggunakan ukuran F1 memberikan akurasi sebesar 0,838 dengan α dan β masing-masing 0,6 dan 0,4. [5]

Penelitian selanjutnya dilakukan oleh David, et al. Pada penelitian ini data yang digunakan adalah data twitter dengan menganalisis sentimen dari sebuah tweet. Pendekatan yang digunakan adalah menggunakan TSA. Penelitian ini menggunakan lima set data yang berbeda

untuk mengantisipasi kesalahan. Hasil penelitian ini ditemukan 70% akurat. [6]

Penelitian selanjutnya dilakukan oleh Hutto, et al. Penelitian ini menggunakan VADER, model berbasis aturan sederhana untuk umum analisis sentimen, dan membandingkan keefektifannya dengan sebelas tolok ukur termasuk LIWC, ANEW, Senti Word Net, dan teknik penambangan data yang mengandalkan Naive Bayes, Algoritma Entropi Maksimum, dan Mesin Vektor Dukungan (SVM). Hasilnya menggabungkan fitur leksikal ini dengan pertimbangan lima aturan umum yang mewujudkan konvensi gramatikal dan sintaksis untuk mengekspresikan dan menekankan intensitas sentimen. dengan VADER mengungguli penilai manusia individual (Keakuratan Klasifikasi F1 = 0,96 dan 0,84, masing-masing). [7]

Penelitian selanjutnya dilakukan oleh Wang, et al. Studi ini menjelaskan tentang sistem analisis sentimen publik secara real-time terhadap calon presiden pada Pemilu AS 2012 seperti diungkapkan di Twitter, layanan microblogging. Hasil yang diperoleh dari penelitian ini adalah pengolahan data secara realtime pada infrastruktur dan model sentimen statistik dengan mengevaluasi perubahan sentimen masyarakat dalam menanggapi peristiwa dan berita politik yang muncul saat pemerintahan berlangsung. Arsitektur dan metode bersifat umum, dan dapat dengan mudah diadopsi dan diperluas ke domain lain (misalnya, penggunaan sistem untuk mengukur sentimen tentang film dan aktor atau nominasi dan pemilihan Oscar). [8]

Penelitian selanjutnya dilakukan oleh Songbo et al. Penelitian ini melakukan klasifikasi sentimen untuk dokumen berbahasa Mandarin. Untuk memperbaiki kekurangan ini, makalah ini menyajikan studi empiris kategorisasi sentimen dalam dokumen Cina. Empat metode seleksi fitur (MI, IG, CHI dan DF) dan lima metode pembelajaran (klasifikasi centroid, K-NN, Naive Bayes dan SVM) menggunakan sentimen bahasa China dengan ukuran 1.021 dokumen. Hasil percobaan menunjukkan bahwa IG melakukan yang terbaik untuk pemilihan istilah sentimental dan SVM menunjukkan kinerja terbaik untuk mengklasifikasikan sentimen dan mengklasifikasikan sentimen sangat bergantung pada domain atau topik. [9]

Penelitian selanjutnya dilakukan oleh Alexander, et al. Studi ini menggunakan Twitter, platform microblogging paling populer, untuk tugas analisis sentimen. Cara mengumpulkan korpus secara otomatis untuk analisis sentimen dan tujuan penambangan data. Analisis yang dilakukan adalah analisis linguistik terhadap koleksi yang dikumpulkan dan. Menggunakan korpus, mengklasifikasikan sentimen, kemudian menentukan sentimen positif, negatif, dan netral untuk dokumen. Hasilnya menunjukkan bahwa teknik yang diusulkan efisien dan berkinerja lebih baik daripada metode yang diusulkan sebelumnya. [10]

III. KAJIAN TEORI

a) Sentiment analysis

Tugas dasar dalam analisis sentimen adalah mengklasifikasikan polaritas teks dalam dokumen, kalimat, atau tingkat fitur/aspek dan menentukan apakah opini dinyatakan dalam dokumen, kalimat atau fitur entitas/aspek positif, negatif atau netral. [2]

b) Metode Naive Bayes

Naive Bayes adalah teknik prediksi berbasis probabilistik sederhana yang didasarkan pada penerapan teorema Bayes (aturan Bayes) dengan asumsi kemandirian (bukan ketergantungan) yang kuat (naif). Dengan kata lain, dalam Naive Bayes model yang digunakan adalah model fitur independen. Naive Bayes Classifier adalah salah satu algoritma yang digunakan untuk klasifikasi teks dan merupakan metode Machine Learning yang menggunakan perhitungan probabilitas dan statistik yang dikemukakan oleh Thomas Bayes. Algoritma digunakan untuk memprediksi probabilitas di masa depan berdasarkan pengalaman masa lalu. [3] Dasar dari naive bayes yang digunakan adalah rumusnya:

$$P(C|X) = \frac{P(x|c)P(c)}{P(x)} \quad (1)$$

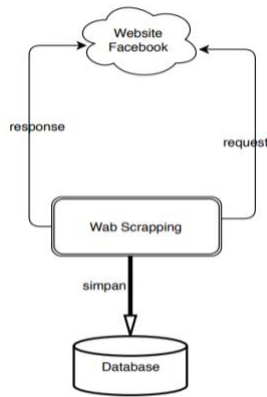
- x : Data dengan kelas yang tidak diketahui
- c : Hipotesis data adalah kelas tertentu
- $P(c | x)$: Probabilitas suatu hipotesis berdasarkan kondisi (posteriori)
- $P(c)$: Probabilitas hipotesis (probabilitas sebelumnya)
- $P(x | c)$: Probabilitas berdasarkan kondisi pada hipotesis
- $P(x)$: Probabilitas c

IV. METODE PENELITIAN

Pada metode penelitian ini akan dijelaskan langkah-langkah yang dilakukan untuk merancang aplikasi analisis sentimen pada fanpage Facebook sebagai berikut:

a) Metode Pengumpulan Data

Pada penelitian ini pengumpulan data diawali dengan penarikan data fanpage facebook dari website facebook yang kemudian disimpan ke dalam database. Penarikan data tweet dilakukan dengan menggunakan fasilitas Web Scrapping. Web Scrapping adalah teknik untuk mengutip dan mengekstraksi data atau informasi dari situs web menggunakan HTML tingkat rendah. Web Scrapping ini mengambil data kotor secara realtime dari website Facebook, yang kemudian akan dipilih sebagai data komentar bersih. Data komentar ini akan disimpan dalam database. Proses pengumpulan data seperti pada Gambar 1.



Gambar 1. Proses pengumpulan data

Data diperoleh dari komentar fanpage Facebook yang dilakukan sejak Januari 2019. Setelah proses pengambilan data selesai. Data komentar akan digunakan dengan membagi semua pelatihan dengan 70% dan data uji dengan 30%. Untuk proses pembuatan data latih, klasifikasi data dilakukan secara manual menjadi tiga kategori yaitu positif, negatif, dan netral.

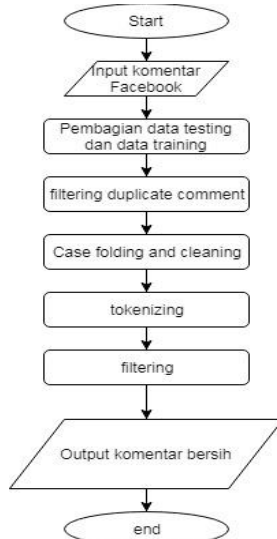
3.2. Metode Pengolahan Data

a. Preprocessing

Preprocessing dilakukan untuk menghindari data yang tidak sempurna, interferensi data, dan data yang tidak konsisten. Dataset yang digunakan dalam penelitian adalah crawling data pada fanpage Facebook PUBG periode Januari 2019 dengan total raw data 2000. Tahapan text preprocessing yang dilakukan adalah:

1. Filtering duplicate
2. Case folding
3. Cleaning
4. The tokenizing / parsing stage
5. Filtering

Preprocessing dapat digambarkan melalui flowchart seperti pada Gambar 2.

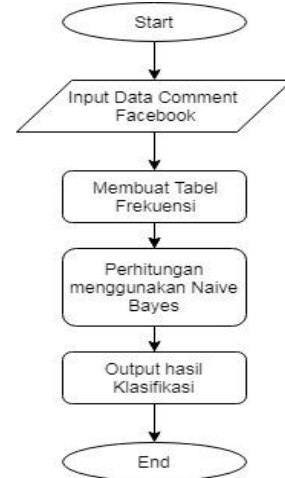


Gambar 2. Preprocessing

Setelah dibersihkan diperoleh hasil data sebanyak 1500 yang selanjutnya dilakukan pemilihan data yang memiliki sentimen terhadap komentar di fanpage PUBG facebook.

b. Classification Model

Proses menggunakan Metode Naive Bayes seperti Flowchart pada Gambar 3.



Gambar 3. Model Klasifikasi

Perancangan flowchart ini bertujuan untuk menggambarkan bagaimana proses klasifikasi komentar dengan perhitungan algoritma Naive Bayes Classifier mulai dari memasukkan data komentar yang akan diuji hingga aplikasi dapat menampilkan hasil akhir klasifikasi komentar.

IV. PEMBAHASAN

5.1. Dataset

Dataset yang digunakan dalam penelitian adalah data crawling data pada fanpage PUBG mobile di Facebook periode Januari 2019 dengan jumlah data sebanyak 1000 data pada percobaan pertama dan kemudian 1500 data pada percobaan berikutnya. Proses crawling data dilakukan menggunakan software facepager dengan mengambil data komentar untuk setiap komentar pada fanpage PUBG Mobile di Facebook sebanyak 10 halaman..

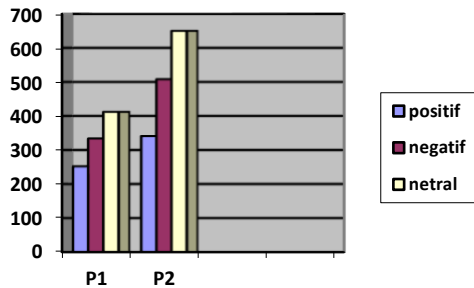
5.2. Data Processing

Data komentar yang diperoleh dari proses crawling data kemudian dilakukan pembersihan data, kemudian diberi label secara manual menjadi 3 bagian yaitu sentimen positif, negatif dan netral. Dataset yang telah diberi label dapat dilihat pada Gambar 4.

KOMENTAR	SENTIMEN
Makin semangat memulai pertandingan dengan skin parasut yang keren	positif
Min tolong kalau bisa adain fitur replay min jadi kita bisa tau min yang pake cheat bisa ke deteksi bisa di liat ulang. Biar ketauar	negatif
Tolong kasih death replay Min saya tadi mati gak jelas udah lari di balik tembok tetep masih kena tembak dan musuh kayaknya n	negatif
Hallo min tanggap dongg Coba minin bhat di youtoube. Banyak yang ngasih tutorial cheat cheater Gamers lain juga kesel min	negatif
Min cheater nya parah ya Gua lupa Screenshot. Dia bangga. Pake cheater min ngaku ngaku anak buah digira. Sakit gua min la	negatif
You know lah...klo wibu bakal pilih yg mana	netral
Apalah daya free player,parasutnya yang ada cuma gambar listrik	netral
min mau nanya di akun gw kok gada mistery air dropnya ?	netral
Banned min sering LS di FB nih orang	negatif
Yang kanan gak dapet2 njing	netral
Masuk gw masuk ke castle tiba gw ada di dalam tanah	negatif
Makin mantap abis dikil citer gak bisa di report	negatif
Baikin voice impresif plississ	negatif
Bau bawangv	negatif
Cara dptin yg kanan bagaimana min?	netral
Ini gimana ya saya gak bisa login cuman gini doang trs bisa itu kembali keluar lagi	negatif
Banned min bikin game gak seru aja Winrate jadi rusak pula	negatif
paling ga lama lagi player pubg pada minggat krn pihak pubg terkesan acuh dlm menangan ceater cuma bann akun ya kalo perm	negatif
Min ini knapa hadiah dance nya g bisa di claim padahal udah dapet	negatif
Next sessions system matching dirubah dong. Crown 1 bisa mabar dgn Ace.. Diamond 1 bisa Mabar dgn crown2...sya imbang n	netral
Min tolong dong di rank Crown 2 - Crown 1 cheater pada berkellaran masa iya nembak gk kelihatan posisinya	negatif
Kalau beli RP / topup sekarang nyesal gak? Bisa 16 hari lagi mohon info gank	netral
Kak mau tanya beat bisa streaming di grub ini caranya gmn.?	netral
Mau sebagai apa parasut atau skin pesawat... Permainan tetap di mulai saat sudah landing. Spanjang perjalanan jg trs mem	negatif
Min kenapa pubg sekarang suka framedrop klo ketemu musuh jarak dekat?	negatif
Ini barang dari tiket kok langsung kadaluarsa. Udah banyak yg kadaluarsa min. Mayan kan naik royal pass	netral

Gambar 4. Dataset Komentar Facebook

Data pengolahan data pada percobaan pertama menggunakan 1000 data dan pada percobaan kedua menggunakan 1500 data dapat dilihat pada Gambar 1 dimana P1 = Percobaan Pertama dan P2 = Percobaan Kedua. Grafik sentimen untuk percobaan pertama dan kedua dapat dilihat pada Gambar 5.



Gambar 5. Grafik Pengalaman Sentimen 1 dan 2

Eksperimen pertama menggunakan 1000 data komentar diperoleh sentimen positif = 252, negatif = 334 dan netral = 412 sedangkan pada eksperimen kedua menggunakan 1500 dataset secara acak diperoleh positif = 341, negatif = 508 dan netral = 650

5.1. Metode Naive Bayes

Penerapan metode Naive Bayes digunakan untuk menguji akurasi setiap data berlabel. Pada percobaan pertama menggunakan 1000 dataset, percobaan dilakukan dengan cross-validation dengan jumlah folds = 10 diperoleh akurasi sebesar 41,7% dengan akurasi detail seperti pada Tabel 1.

Tabel 1. Ketepatan detail untuk Eksperimen Pertama

Precision	Recall	F-Measure	Class
0.600	0.12	0.023	Positif
0.800	0.12	0.024	Negatif
0.414	0.993	0.584	Netral
0.590	0.417	0.255	

Karena akurasi yang rendah pada percobaan pertama, percobaan kedua dilakukan dengan asumsi persentase akurasi akan meningkat. Oleh karena itu, pada percobaan kedua digunakan dataset sebanyak 1500 data dengan cara yang sama seperti percobaan pertama. Dari hasil percobaan kedua didapatkan nilai akurasi sebesar 43,429% meningkat sebesar 2% dari percobaan pertama. Rincian akurasi pada percobaan kedua dapat dilihat pada Tabel 2.

Tabel 2. Akurasi detail untuk Eksperimen Kedua

Precision	Recall	F-Measure	Class
0.500	0.018	0.034	Positif
0.333	0.004	0.008	Negatif
0.434	0.989	0.603	Netral
0.415	0.434	0.272	

V. KESIMPULAN

Berdasarkan penelitian yang dilakukan dapat disimpulkan bahwa pada percobaan pertama dengan menggunakan 1000 dataset komentar diperoleh sentimen positif = 252, negatif = 334 dan netral = 412 sedangkan pada percobaan kedua dengan menggunakan 1500 dataset secara acak diperoleh positif = 341, negatif = 508 dan netral = 650. Dan untuk pengujian akurasi menggunakan metode Naive Bayes pada percobaan pertama akurasi sebesar 41,7% dan pada percobaan kedua dengan akurasi data yang lebih meningkat menjadi 43,429%.

VI. SARAN

Untuk mendapatkan akurasi yang lebih besar disarankan untuk menggunakan kumpulan data yang lebih banyak karena komentar di media sosial Facebook sangat bervariasi.

REFERENSI

- [1] Lovelock, Christopher. & Wright, Lauren. (2002). *Principle of Service Marketing and Management* (2nded). New Jersey: Pearson Education International.
- [2] Liu, B, (2012), *Opinion Mining*. Chicago, United States of America.
- [3] Daniela, Christopher J, and Roger G, 2009, Naive Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages, *IJCSI International Journal of Computer Science Issues*, Vol. 4, No. 1, 2009
- [4] Vishal A. Kharde, S.S. Sonawane, 2016, Sentiment Analysis of Twitter Data: A Survey of Techniques, *International Journal of Computer Applications* (0975 – 8887) Volume 139 – No.11, April 2016
- [5] Rizal Setya Perdana and Aryo Pinandito, 2018, Combining Likes-Retweet Analysis and Naive Bayes Classifier within Twitter for Sentiment Analysis, e-ISSN: 2289-8131 Vol. 10 No. 1-8
- [6] David Zimbra, Abbasi, Daniel Zenk and Hsinchun Chen, 2018, The State-of-the-Art in Twitter Sentiment Analysis: A Review and Benchmark Evaluation, *ACM Transactions on Management Information Systems*, Vol. 9, No. 2, Article 5. Publication date: August 2018.
- [7] C.J. Hutto, Eric Gilbert, 2014, VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text, *Proceedings of the Eighth International AAI Conference on Weblogs and Social Media*
- [8] Hao Wang, Dogan Can, Abe Kazemzadeh, François Bar and Shrikanth Narayanan, 2012, A System for Real-time Twitter Sentiment Analysis of 2012 U.S. Presidential Election Cycle, *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, pages 115–120, Jeju, Republic of Korea, 8-14 July 2012. c 2012 Association for Computational Linguistics.

- [9] Songbo Tan JinZhang 2008, An empirical study of sentiment analysis for chinese documents Expert Systems with Applications Volume 34, Issue 4, May 2008, Pages 2622-2629
- [10] Alexander Pak, Patrick Paroubek, 2016, Twitter as a Corpus for Sentiment Analysis and Opinion Mining, Universit e de Paris-Sud, Laboratoire LIMSIS-CNRS, B atiment 508, F-91405 Orsay Cedex, France