

Persepsi Publik Tentang Pembelajaran Daring di Indonesia: Studi Menggunakan *ELK Stack* dan Python untuk Analisis Sentimen di Twitter

Andri Oktavianto, dan Satria Fadil Persada

Departemen Manajemen Bisnis, Institut Teknologi Sepuluh Nopember (ITS)

e-mail: satriaafp@gmail.com

Abstrak—Penelitian ini menyelidiki persepsi masyarakat umum tentang aplikasi pembelajaran daring di Indonesia. Banyak studi tentang pembelajaran daring dilakukan di negara maju dan masih sedikit di negara berkembang. Sementara penelitian tradisional menggunakan survei untuk memahami persepsi orang terhadap suatu entitas membutuhkan banyak waktu dan upaya, penelitian ini menggunakan cara yang efisien dan efektif untuk mengumpulkan pendapat di Twitter dan kemudian menganalisis sentimennya menggunakan Elasticsearch, Logstash (ELK stack) dan bahasa pemrograman Python. Algoritma Naïve Bayes digunakan untuk analisis sentimen dan ELK stack untuk mengumpulkan dan menyimpan tweets dari Twitter. Dengan ELK stack, penelitian ini berhasil mengumpulkan 133.477 tweets yang terkait dengan pembelajaran daring. Dari hasil analisis sentimen ditemukan sebanyak 98.3% tweets memiliki sentimen positif terhadap pembelajaran daring, hal ini menjadi bukti bahwa pelajar di Indonesia memiliki persepsi positif terhadap pembelajaran daring. Dari studi ini ditemukan bahwa fitur pembelajaran daring yang paling banyak dibicarakan adalah tampilan, bank soal, dan konten. Hasil evaluasi model pembelajaran mesin menunjukkan bahwa algoritma Naïve Bayes memiliki akurasi sebesar 73% untuk klasifikasi opini dan 75% untuk klasifikasi sentimen. Penelitian ini sebelumnya dipublikasikan di *International Journal of Emerging Technologies in Learning (IJET)* tahun 2020.

Kata Kunci—Analisis Sentimen, ELK stack, Indonesia, Pembelajaran Daring, Twitter

I. PENDAHULUAN

Pendidikan, di masa lampau hingga saat ini, umumnya merupakan proses transfer pengetahuan satu arah: dari guru ke murid. Namun, terdapat kelemahan dari metode belajar tersebut: kurang efektif untuk menekankan kemampuan memecahkan masalah dan berfikir kritis, padahal keduanya dianggap sebagai kemampuan yang penting di abad ke-21 [1]. Transfer pengetahuan satu arah juga kurang sesuai dengan kebutuhan ilmu pengetahuan untuk masyarakat di dunia yang berkembang pesat ini karena kemajuan teknologi [2]. Beruntungnya, siswa-siswi saat ini beruntung hidup di era pesat-pesatnya perkembangan teknologi, dimana internet mudah diakses dan murah. Keberadaan internet membuat siswa tidak terlalu bergantung kepada guru saja sebagai sumber pembelajaran mereka, karena internet memungkinkan siswa untuk belajar subjek atau materi yang mereka ingin pelajari hanya dengan akses internet. Penggunaan teknologi internet untuk pembelajaran disebut pembelajaran daring. Lebih lanjut lagi, pembelajaran daring adalah penggunaan moda daring untuk menyampaikan setidaknya 80% dari isi kursus [3].

Keberadaan pembelajaran daring mengurangi proses pendidikan pada umumnya yang menuntut siswa untuk duduk di kelas dan mendengarkan materi yang mungkin sudah mereka pelajari. Selain itu, keuntungan tertentu dari pembelajaran daring adalah bagaimana guru dan siswa tidak diharuskan untuk online pada waktu bersamaan; ini dapat mengurangi sekitar sepertiga dari total waktu belajar yang dibutuhkan [4]. Namun, keuntungan ini tidak ada artinya ketika siswa tidak memiliki minat atau persepsi yang baik tentang pembelajaran daring. Bahkan hingga persepsi siswa tentang pengalaman belajar daring dapat mempengaruhi keputusan mereka apakah akan melanjutkan belajar daring atau tidak [5]. Karena itu, penting untuk memahami persepsi siswa dengan lebih baik untuk meningkatkan sistem pendidikan kita.

Beberapa studi telah menyelidiki persepsi siswa terhadap pembelajaran daring, dan ada beragam pendapat; beberapa penelitian [6][7][8] menunjukkan persepsi siswa yang positif, sementara yang lain menunjukkan kekhawatiran terhadap pembelajaran daring sebagai metode pembelajaran. Meski demikian, mayoritas siswa dalam studi yang telah disebutkan sebelumnya menunjukkan pendapat positif terhadap pembelajaran daring. Selain itu, studi-studi ini juga hanya dilakukan di negara-negara maju, yang mana membatasi signifikansi praktisnya untuk negara-negara dengan prospek ekonomi yang berbeda. Penelitian ini juga mempertimbangkan aspek sebagaimana disebutkan di atas, menyadari bahwa akses terhadap pembelajaran daring di negara berkembang belum menjadi hal yang umum, sehingga urgensi untuk memahami persepsi siswa tentang pembelajaran daring menjadi penting.

Untuk mengembangkan pembelajaran daring agar diterima lebih baik oleh pelajar di Indonesia, penting untuk memiliki pengetahuan tertentu tentang opini pengguna dan mengerjakan evaluasi dari opini tersebut; analisis sentimen bisa menjadi solusi yang bagus dalam kasus ini. Analisis sentimen dapat membantu penyelenggara pembelajaran daring menganalisis opini pelanggan untuk mengembangkan produk, memberikan layanan yang lebih baik, dan bahkan menemukan peluang bisnis baru. Alih-alih membutuhkan pengaturan khusus, metode ini dapat menangkap pesan dari banyak dan beragam audiens. Hal ini mungkin dilakukan dengan menggunakan analisis sentimen dan Twitter *Streaming API* yang mana telah berhasil dilakukan di banyak penelitian, salah satu penelitian [9] menggunakan algoritma *Naïve Bayes* untuk mengklasifikasikan 45.000 *tweets* dari *dataset* Stanford; sedangkan penelitian lain [10] menggunakan 19.000 *tweets* dari Twitter *Streaming API* dan

beragam algoritma untuk analisis sentimen seperti *Maximum Entropy* dan *Support Vector Machine*. Analisis sentimen dapat menjelaskan sentimen seseorang, sikap, pendapat, dan emosi terhadap produk, individu, topik, organisasi, dan jasa. Dalam banyak penelitian terdahulu, para peneliti tersebut telah menggunakan analisis sentimen untuk mengukur reputasi merek, menemukan wawasan untuk keperluan kampanye politik, dan bahkan memprediksi pergerakan pasar saham menggunakan data Twitter.

Penelitian ini menjadi begitu penting, terutama untuk mengetahui apakah sebenarnya layanan pembelajaran daring telah diterima dengan baik oleh mayoritas pelajar di Indonesia, ataukah baru hanya segelintir saja. Ditambah, hasil dari penelitian ini bisa menjadi salah satu sumber informasi bagi perusahaan layanan pembelajaran daring tentang layanan yang seharusnya diperbaiki dan dipertahankan dengan menggunakan pendapat dari masyarakat umum.

II. TINJAUAN PUSTAKA

A. Pembelajaran Daring

Karena masih belum ada definisi yang jelas dari pembelajaran daring dan rujukan yang samar-samar ke terminologi lain seperti kursus daring, pembelajaran berbasis web, pelatihan berbasis web, atau pembelajaran jarak jauh, maka peneliti menggunakan definisi yang diterima secara umum: pembelajaran daring adalah kursus di mana sebagian besar atau semua konten disampaikan secara daring dan biasanya tidak memiliki pertemuan tatap muka [11].

B. Analisis Sentimen

Analisis sentimen digunakan untuk menentukan apakah sebuah kalimat atau teks mengekspresikan sentimen tertentu: positif atau negatif [12]. Secara umum, ada dua pendekatan untuk melakukan analisis sentimen, pendekatan *lexicon-based* dan pendekatan *machine learning*.

Pendekatan *machine learning* yang digunakan dalam penelitian ini adalah *supervised learning*. *Supervised learning* merupakan pendekatan yang menggunakan data berlabel, yang mana pada penelitian ini data ini digunakan untuk melatih algoritma agar dapat mengklasifikasikan sentimen. Ini berarti bahwa langkah pertama yang harus dilakukan adalah mengumpulkan data teks yang sudah diberi label sentimen positif dan negatif, memecah kalimat utuh menjadi kata-kata (*feature extraction*), dan kemudian melatih algoritma untuk mengklasifikasikan sentimen. Algoritma yang digunakan untuk proses klasifikasi adalah *Naïve Bayes*.

C. Negara Berkembang

Penelitian ini mengadopsi klasifikasi negara berkembang berdasarkan pendapatannya yang diajukan oleh World Bank. Negara-negara diklasifikasikan berdasarkan tingkat perkembangannya yang diukur dengan *gross domestic product (GDP) per capita* yang dinilai setiap tahun dalam dolar Amerika Serikat [13]. Mengacu pada sistem tersebut, negara dikelompokkan sebagai negara berpendapatan tinggi, berpendapatan menengah ke atas, berpendapatan menengah ke bawah, dan berpendapatan rendah. Untuk tahun fiskal 2020, negara-negara dengan *GDP* kurang dari \$1.025 per kapita diklasifikasikan sebagai berpendapatan rendah, negara-negara dengan *GDP* antara \$1.026-3.995 diklasifikasikan sebagai berpendapatan menengah ke bawah,

\$3.996-12.375 diklasifikasikan sebagai berpendapatan menengah ke atas, dan negara yang berpendapatan lebih dari \$12.376 diklasifikasikan sebagai berpenghasilan tinggi [14]. Negara-negara dengan pendapatan menengah ke bawah dianggap sebagai negara berkembang. Indonesia diklasifikasikan sebagai negara berpenghasilan menengah ke bawah oleh World Bank

D. Naïve Bayes

Algoritma *Naïve Bayes* merupakan salah satu algoritma supervised learning yang populer. Algoritma ini berdasar pada Teorema Bayes. Ciri utama *Naïve Bayes* adalah asumsi kuat bahwa fitur dari dataset tidak berhubungan dengan fitur lainnya pada dataset yang sama [15].

$$P(B|A) = P(x_1, x_2, \dots, x_n | A) = \prod_{i=1}^n P(x_i|A)$$

Persamaan diatas merupakan persamaan *Naïve Bayes*. Pada *Naïve Bayes*, fitur pada *dataset* diasumsikan tidak memiliki keterkaitan satu dengan yang lain alias independen. Penelitian ini menggunakan dua jenis algoritma *Naïve Bayes* yang umum digunakan, yaitu: *Multinomial Naïve Bayes* dan *Bernoulli Naïve Bayes*, kedua algoritma tersebut akan dibandingkan performanya. Algoritma dengan akurasi paling tinggi akan digunakan untuk mengklasifikasikan opini dan sentimen *tweets* pada penelitian ini.

E. Confusion Matrix

Confusion Matrix, atau juga dikenal dengan sebutan *contingency matrix*, adalah sebuah metode untuk mengukur performa algoritma *machine learning* menggunakan tabel matriks [16]. Cara kerja *confusion matrix* cukup sederhana, yaitu dengan membagi hasil klasifikasi dari algoritma *machine learning* menjadi empat, yaitu; (1)*True Positive (TP)*: Data positif yang diklasifikasikan positif; (2)*False Negative (FN)*: Data sebenarnya positif, namun diklasifikasikan negative; (3)*True Negative (TN)*: Data negatif yang diklasifikasikan negative; (4)*False Positive (FP)*: Data sebenarnya negatif, namun diklasifikasikan positif.

Penggunaan *confusion matrix* cukup mudah, pengguna hanya perlu menghitung jumlah klasifikasi yang benar dan salah; misalnya dalam penelitian ini terdapat *tweets* yang diklasifikasikan sebagai *tweets* bersentimen positif oleh algoritma *Naïve Bayes*, namun sebenarnya merupakan *tweets* negatif, maka *error* ini dinamakan *False Positive*. Hal ini memberikan wawasan kepada peneliti tidak hanya *error* yang dilakukan oleh algoritma, namun juga jenis *error* yang dibuat.

III. METODE PENELITIAN

A. Data Pipeline

Proses analisis membutuhkan data dari Twitter *Streaming API* yang dikumpulkan menggunakan Logstash, disimpan dalam indeks Elasticsearch, kemudian dibersihkan dan dianalisis menggunakan Python.

B. Pilot Test

Untuk membuat hasil analisis sentimen lebih baik, peneliti memnyaring *tweet* yang bukan merupakan opini terhadap

Tabel 1.
Kata kunci yang digunakan untuk mengambil tweets

Kategori	Kata kunci
Aplikasi Pembelajaran Daring	belajar zenius, pake zenius, belajar ruangguru, belajar RG, pake ruangguru, pake RG, belajar quipper, belajar quiper, pake quipper, pake quiper, belajar youtube, belajar coursera
Pembelajaran Daring Belajar	kelas online, belajar online mau belajar, belajar buat, belajar pake, kamu belajar, belajar apa, belajar dari

Tabel 2.
Jumlah tweets berdasarkan topik

Topik	Jumlah
Tampilan	3.376
Bank Soal	2.260
Konten	1.123
Pengajar	306
Keterjangkauan	150
Kemudahan Belajar	47

Tabel 3.
Jumlah tweets berdasarkan merek

	Zenius	Ruangguru	Quipper
Positif	185	32	19
Negatif	13	5	7
Total	198	37	26

pembelajaran daring dengan melatih algoritma *Naïve Bayes* menggunakan sampel *tweet* agar bisa mengklasifikasikan *tweet* sebagai opini terhadap pembelajaran daring atau bukan. Peneliti memilih 72 *tweet* sampel (masing-masing *tweet* berasal dari pengguna yang berbeda) dan melabel *tweet* tersebut secara manual sebagai opini atau bukan opini. Metode penyaringan ini telah digunakan sebelumnya dalam penelitian serupa [17].

C. Kata Kunci

Berikut ini merupakan kata kunci yang digunakan untuk mendapatkan *tweets* tentang pembelajaran daring: Kata kunci pada Tabel 1 dipilih berdasarkan hasil *pilot test*, dimana sebelumnya data dari *tweet* terkait pembelajaran daring diekstrak dan dihitung frekuensi setiap kata; kata-kata yang berhubungan dengan pembelajaran daring dengan frekuensi tinggi kemudian dipilih sebagai kata kunci. Dari hasil *pilot test* ditemukan bahwa kata kunci satu kata (*unigrams*) dapat menangkap *tweets* lebih banyak dibandingkan dengan dua kata (*bigrams*) atau tiga kata (*trigrams*), tetapi banyak dari *tweets* tersebut tidak terkait langsung dengan pembelajaran daring. Dengan menggunakan *trigrams*, penelitian ini berhasil menangkap *tweets* yang lebih spesifik terkait pembelajaran daring, akan tetapi jumlah *tweets* yang

Tabel 4.
Metrik kunci algoritma dalam mengklasifikasikan opini

	Precision	Recall	F1-Score	Support
Bukan Opini	0.77	0.88	0.82	26
Opini	0.62	0.42	0.50	12
Rata-rata/Total	0.72	0.74	0.72	38

Tabel 5.
Metrik kunci algoritma dalam mengklasifikasikan sentimen

	Precision	Recall	F1-Score	Support
Negatif	0.67	0.40	0.50	10
Positif	0.77	0.91	0.83	22
Rata-rata/Total	0.74	0.75	0.73	32

ditangkap terlalu kecil. Menggunakan *bigrams* juga terbukti menghasilkan hasil yang lebih baik dari algoritma klasifikasi untuk teks subjektif [18].

IV. ANALISIS DAN DISKUSI

A. Pengumpulan Data

Data *tweets* dikumpulkan secara *real-time* dari Twitter yang diakses langsung menggunakan Twitter *Streaming API*. Kata kunci yang digunakan untuk menangkap *tweets* terkait pembelajaran daring telah dijelaskan di bab sebelumnya. Selama periode Maret - Juni 2019 sebanyak 133.477 *tweets* berhasil didapatkan.

B. Preprocessing

Sebelum proses analisis sentimen dapat dilakukan kita perlu membersihkan data yang didapatkan terlebih dahulu agar hasil analisis sentimen semakin akurat. Tahap ini dikenal sebagai *preprocessing*.

1) Normalisasi tweets:

Yang diperlukan untuk melakukan analisis sentimen adalah teks *tweet*, namun didalam isi suatu *tweets* seringkali kita temukan mention kepada pengguna lain, URL, dan hashtag. Keberadaan variabel-variabel tersebut dapat mempengaruhi hasil analisis sentimen, oleh karena itu kita perlu memisahkan variabel tersebut dengan isi *tweets*.

2) Tokenization:

Tokenisasi dilakukan berdasarkan spasi pada penelitian ini. Tokenization adalah proses memecah belah suatu teks menjadi kata per kata, umumnya dipisahkan berdasarkan spasi. Misalnya, kalimat "capek juga belajar seharian" jika dilakukan tokenisasi berdasarkan spasi maka akan menjadi empat kata: "capek", "juga", "belajar", "seharian".

C. Hasil Pilot Test

Perlu kita sadari bahwa dari kumpulan *tweets* yang didapatkan tidak semuanya berkaitan dengan topik pembelajaran daring. Agar hasil analisis sentimen lebih akurat, maka peneliti memisahkan antara *tweets* opini terhadap pembelajaran daring dengan yang bukan opini menggunakan algoritma *Naïve Bayes*. Algoritma *Naïve Bayes* dilatih dengan menggunakan kumpulan *tweets* yang telah dikumpulkan sebelumnya.

Hasilnya, hanya 13.2% dari total *tweets* yang ditandai sebagai opini terhadap pembelajaran daring, hal itu berarti hanya 17.608 *tweets* yang bisa dianalisis sentimennya. Sentimen tiap *tweets* tersebut nantinya akan diklasifikasikan dengan menggunakan algoritma *Naïve Bayes*. Dibawah ini merupakan sampel *tweets* yang diklasifikasikan sebagai opini dan bukan opini terhadap pembelajaran daring.

D. Proses Klasifikasi Sentimen

Agar algoritma *Naïve Bayes* dapat mengklasifikasikan sentimen, maka kita harus melatihnya terlebih dahulu. Penelitian ini menggunakan sampel *tweets* yang telah terkumpul dan melabeli sentimen *tweets* tersebut sebagai positif atau negatif secara manual untuk melatih algoritma *Naïve Bayes*. *Tweets* yang digunakan sebagai sampel adalah *tweets* yang lolos penyaringan pada *pilot test*.

Terdapat total 40 *tweets* yang dilabeli manual dengan penjabaran: 28 *tweets* bersentimen positif dan 12 *tweets* bersentimen negatif yang digunakan untuk melatih algoritma *Naïve Bayes* dalam mengklasifikasikan sentimen suatu *tweets*. *Tweets* yang digunakan dapat dilihat pada bagian Lampiran 2.

Setelah data sampel *tweets* siap digunakan untuk melatih algoritma, data tersebut digunakan untuk melatih dua algoritma *Naïve Bayes* yang berbeda. Hal ini bertujuan untuk membandingkan kinerja dua algoritma *Naïve Bayes* yang umum digunakan, yaitu: *Bernoulli* dan *Multinomial*.

Setelah proses melatih algoritma selesai, didapatkan akurasi masing-masing algoritma dalam mengklasifikasikan *tweets* sebagai opini terhadap pembelajaran daring atau bukan—algoritma ini digunakan pada *pilot test*—masing-masing sebesar: 68% untuk *Bernoulli Naïve Bayes* dan 73% untuk *Multinomial Naïve Bayes*. Cara yang sama juga digunakan dalam proses mengklasifikasikan sentimen, dengan masing-masing akurasi dari algoritma *Bernoulli Naïve Bayes* dan *Multinomial Naïve Bayes* sebesar: 65% dan 75%. Nilai akurasi dari algoritma *Multinomial Naïve Bayes* lebih tinggi daripada algoritma *Bernoulli Naïve Bayes* dalam penelitian ini. Maka dari itu, penelitian ini menggunakan algoritma *Multinomial Naïve Bayes* dalam mengklasifikasikan opini dan sentimen suatu *tweets*.

E. Hasil Analisis Sentimen

Hasil analisis sentimen menunjukkan 98.3% dari *tweets* yang ada memiliki sentimen positif, sedangkan *tweets* lainnya bersentimen negatif. Hasil ini menunjukkan bahwa pelajar di Indonesia—yang menggunakan Twitter—memiliki persepsi dan pengalaman yang bagus terhadap aplikasi pembelajaran daring. Sebagai tambahan, 1.8% dari *tweets* menyatakan bahwa mereka menggunakan aplikasi pembelajaran daring untuk belajar ujian.

Meskipun sebanyak 98.3% dari semua *tweets* pada

penelitian ini menunjukkan sentimen positif terhadap pembelajaran daring, akan tetapi juga terdapat segmen pelajar yang memiliki persepsi negatif terhadap pembelajaran daring. Akan berguna bagi penyelenggara pembelajaran daring untuk mengetahui tidak hanya apa yang pelajar sukai dari pembelajaran daring, namun juga apa fitur yang seharusnya dikembangkan, diperbaiki, atau tidak tersedia dari pembelajaran daring. Salah satu cara yang bisa dicoba yaitu dengan menjelajahi *tweets* bersentimen negatif lebih dalam, dan memverifikasi apakah itu masalah nyata atau bukan.

Untuk membuat analisis yang lebih rinci dilakukan klasifikasi topik *tweets* yang berkaitan dengan pembelajaran daring dan sentimen mereka terhadap topik tersebut dengan enam kategori topik, yaitu: konten, pengajar, bank soal, biaya/keterjangkauan, kemudahan belajar, dan tampilan/*User Interface*.

Tampilan adalah fitur juara satu dari pembelajaran daring dengan sentimen positif terbanyak, diikuti dengan bank soal dan konten. Ketiga fitur utama yang paling banyak dibicarakan mungkin juga yang paling banyak dicari dan diharapkan dari layanan pembelajaran daring.

Saat penelitian ini berlangsung, Zenius adalah merek yang paling banyak disebut pengguna dalam *tweetnya*, diikuti oleh Ruangguru dan Quipper. Zenius menjadi merek pembelajaran daring yang paling banyak diperbincangkan dan paling mengesankan karena jumlah *tweets* bersentimen positif 14x lebih banyak daripada *tweets* bersentimen negatif terhadap merek mereka, namun Zenius juga mendapat sentimen negatif paling banyak, sementara Ruangguru mendapatkan sentimen negatif paling sedikit dibandingkan dengan penyedia pembelajaran daring lainnya. Jumlah *tweets* berdasarkan topik dapat dilihat pada Tabel 2.

F. Evaluasi Algoritma

Berikut ini merupakan nilai metrik kunci hasil performa algoritma *Multinomial Naïve Bayes* dalam mengklasifikasikan *tweets* sebagai opini atau bukan opini terhadap pembelajaran daring. Tingkat akurasi algoritma *Multinomial Naïve Bayes* dalam mengklasifikasikan opini dan sentimen suatu *tweets* dapat dikatakan cukup bagus, dengan masing-masing bernilai: 73% dan 75%.

G. Implikasi Manajerial

Dari hasil analisis kategori topik *tweets* pembelajaran daring, terdapat tiga fitur utama yang paling banyak diperbincangkan, yaitu; (a)Tampilan; (b)Bank Soal, dan; (3)Konten. Tampilan menjadi fitur nomor satu yang paling banyak dibicarakan, hal ini mungkin disebabkan karena rata-rata video pembelajaran daring memiliki tampilan yang menarik. Beberapa sampel *tweets* secara sekilas juga mengatakan bahwa animasi dalam video pembelajaran merupakan hal yang penting untuk mempertahankan motivasi belajar. Penyedia layanan pembelajaran daring harus membuat antar muka aplikasi yang menarik dan memastikan video pembelajaran menggunakan animasi agar disukai oleh pelajar. Jumlah *tweets* berdasarkan merek dapat dilihat pada Tabel 3.

Dengan asumsi rata-rata pengguna pembelajaran daring adalah pelajar SMA dan SMP—mayoritas pelajar SMA, maka tidak heran apabila bank soal merupakan fitur kedua yang paling dicari oleh pengguna, sebab mereka ingin

mempersiapkan diri sebaik mungkin untuk ujian masuk perguruan tinggi agar bisa menjadi mahasiswa di perguruan tinggi favorit. Saran bagi perusahaan penyedia layanan pembelajaran daring yaitu menyediakan bank soal yang bervariasi dan bisa dikerjakan langsung di dalam aplikasi, lebih lanjut lagi perusahaan juga bisa mengadakan *tryout* yang bersifat periodik.

Ketersediaan rangkuman, materi yang mudah dipahami, dan bahan belajar yang mudah diakses juga merupakan poin penting bagi penyedia layanan pembelajaran daring. Penyedia layanan pembelajaran daring harus memastikan bahwa konten belajar mereka mudah dipahami oleh siswa, metrik kunci. Algoritma dalam mengklasifikasikan opini dapat dilihat pada Tabel 4. sebab faktor itu yang membuat salah satu merek aplikasi pembelajaran daring lokal populer.

Dibandingkan dengan merek pembelajaran daring lokal lainnya, Zenius mendapatkan sentimen positif terbanyak. Zenius dikenal dengan cara pengajarannya yang menekankan pada pemahaman fundamental dari topik yang dipelajari, sementara merek lain memiliki daya tariknya sendiri. Siswa memilih merek aplikasi pembelajaran daring berdasarkan preferensi mereka: Ada yang suka mempelajari hal-hal dasar terlebih dahulu dan tidak terlalu mempedulikan aspek visual dari materi pembelajaran, beberapa lebih memilih video belajar yang dilengkapi dengan animasi agar belajar jadi tidak membosankan, sementara yang lain lebih suka materi pembelajaran serta latihan soal yang dapat diakses di ponsel mereka. Secara umum, mayoritas pelajar di Indonesia lebih menyukai Zenius daripada merek lainnya.

Mengikuti langkah Zenius dalam menyampaikan layanan pembelajaran daring mereka mungkin menjadi kunci untuk membuka potensi pertumbuhan pembelajaran daring di negara-negara berkembang seperti Indonesia. Algoritma dalam mengklasifikasikan sentiment dapat dilihat pada Tabel 5.

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Penelitian ini membuktikan bahwa algoritma Naïve Bayes, spesifiknya *Multinomial Naïve Bayes* memiliki akurasi yang cukup baik dalam mengklasifikasikan sentimen tweets tentang pembelajaran daring.

Rekomendasi dari penelitian ini kepada perusahaan pembelajaran daring yaitu: Perhatikan antarmuka aplikasi serta berikan animasi pada video belajar, pastikan terdapat bank soal yang bisa diakses kapanpun, dan tiru langkah Zenius dalam menyampaikan konten belajar. Zenius merupakan merek pembelajaran daring yang paling disukai pelajar di Indonesia berdasarkan jumlah tweets bersentimen positif yang menyebut merek Zenius dalam tweetnya. Padahal, Zenius minim beriklan jika dibandingkan dengan merek kompetitornya—Ruangguru, yang gencar beriklan di Televisi. Setelah mengkurasi tweets yang menyebutkan Zenius, peneliti menemukan karakteristik yang membuat Zenius menonjol, yaitu: Fokus membahas konsep dasar dari materi yang diajarkan. Maka, bagi perusahaan pembelajaran daring lainnya yang ingin sukses Zenius ikutilah mantra berikut ini: Ajarkan pemahaman dasar dari materi yang diajarkan terlebih dahulu supaya pelajar mudah memahami

materi tersebut.

B. Saran

Berikut ini merupakan saran-saran yang dapat bermanfaat untuk penelitian selanjutnya: (1) Menggunakan *server* dalam menarik data dari Twitter *Streaming API* secara otomatis setiap detik. Salah satu kendala yang dihadapi dalam penelitian ini yaitu laptop yang harus menyala setiap saat agar dapat menarik data dari Twitter, hal ini dapat diatasi dengan menyewa *server* yang dapat menyala setiap saat; (2) Meningkatkan jumlah data yang digunakan untuk melatih algoritma klasifikasi karena jumlah data *training* mempengaruhi performa algoritma klasifikasi; (3) Menggunakan algoritma *supervised learning* lainnya atau bahkan *unsupervised learning* dalam mengklasifikasikan sentimen untuk mencari algoritma terbaik dalam mengklasifikasikan sentimen di Twitter; (4) Penelitian selanjutnya dapat dikembangkan dengan cakupan penelitian yang lebih kecil, yaitu menganalisis pendapat pelanggan mengenai layanan atau produk dari tiap-tiap merek pembelajaran daring lokal. Sebagai contoh layanan dari merek Ruangguru antara lain ruangbelajar, ruanguji, dan ruangles; (5) Pembelajaran daring secara *asynchronous* memerlukan komitmen belajar yang tinggi, tidak jarang kita temui pelajar yang pada akhirnya tidak menyelesaikan kursus yang diambil. Masih belum banyak penelitian untuk mengidentifikasi penyebab kejadian tidak terselesaikan atau terputusnya kursus pembelajaran daring.

DAFTAR PUSTAKA

- [1] B. Trilling and C. Fadel, *21st Century Skills: Learning for Life in Our Times*, vol. 47, no. 10. San Francisco: John Wiley & Sons, 2010.
- [2] S. M. Hannafin, M.J., & Land, "The foundations and assumptions of technology-enhanced student-centered learning environments," *Instructional Sci.*, vol. 25, pp. 167–202, 1997.
- [3] I. E. Allen and J. Seaman, *Going the Distance: Online education in the United States*. ERIC, 2011.
- [4] S. Carliner, *An Overview of Online Learning. Human Resource Development*, 2nd ed. USA: HRD Press, INC, 2004.
- [5] S. Carr, "As Distance Education Comes of Age, the Challenge Is Keeping the Students," *Chronicle of Higher Education*, 2000. <https://eric.ed.gov/?id=EJ601725>.
- [6] D. O. Bruff, D. H. Fisher, K. E. McEwen, and B. E. Smith, "Wrapping a mooc: student perceptions of an experiment in blended learning derek," *MERLOT J. Online Learn. Teach.*, vol. 9, no. 2, pp. 187–199, 2013.
- [7] M. Gillingham and C. Molinari, *Online Courses: Student Preferences Survey*, vol. 1, no. 1. Washington, D.C.: New Hampshire Ave., N.W., 2012.
- [8] A. G. Picciano, "Beyond student perceptions: issues of interaction, presence, and performance in an online course," *J. Asynchronous Learn. Networks*, vol. 6, no. 1, pp. 21–40, 2002.
- [9] V. A. and S. S. Sonawane, "Sentiment analysis of twitter data: a survey of techniques," *Int. J. Comput. Appl.*, vol. 139, no. 11, pp. 5–15, 2016, doi: 10.5120/ijca2016908625.
- [10] F. Barbieri, F. Ronzano, and H. Saggion, "Semeval 2015 tasks 10 and 11. sentiment analysis of literal and figurative language in twitter," in *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, 2015, no. SemEval, pp. 704–708, doi: 10.18653/v1/s15-2119.
- [11] I. E. Allen and J. Seaman, *Online ation: Five Years of Growth in Online Learning*. Newburyport: Sloan Consortium, 2007.
- [12] L. Zhang, R. Ghosh, M. Dekhil, M. Hsu, and B. Liu, "Combining lexicon-based and learning-based methods for twitter sentiment analysis," *HP Lab. Tech. Rep.*, no. 89, 2011.
- [13] N. & Fantom and U. Serajuddin, *The World Bank's Classification of Countries by Income.*, no. January. Washington (D.C.): The World Bank, 2016.
- [14] T. W. Bank, "World Bank Country and Lending Groups," 2020. <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519->

- world-bank-country-and-lending-groups.
- [15] H. Zhang, "Exploring conditions for the optimality of naïve bayes," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 19, no. 2, pp. 183–198, 2005, doi: 10.1142/S0218001405003983.
- [16] O. Caelen, "A bayesian interpretation of the confusion matrix," *Ann. Math. Artif. Intell.*, vol. 81, no. 3–4, pp. 429–450, 2017, doi: 10.1007/s10472-017-9564-8.
- [17] W. Chamlerwat, P. Bhattarakosol, T. Rungkasiri, and C. Haruechaiyasak, "Discovering consumer insight from twitter via sentiment analysis," *J. Univers. Comput. Sci.*, vol. 18, no. 8, pp. 973–992, 2012, doi: 10.3217/jucs-018-08-0973.
- [18] S. Wang and C. D. Manning, "Baselines and bigrams: simple, good sentiment and topic classification," *50th Annu. Meet. Assoc. Comput. Linguist. ACL 2012 - Proc. Conf.*, vol. 2, no. July, pp. 90–94, 2012.