BAB II

TINJAUAN PUSTAKA

2.1 Studi Sebelumnya

Penelitian terkait mengenai analisis sentimen pada aplikasi X mungkin telah banyak dilakukan dengan berbagai pendekatan dan metode pengembangan. Berikut merupakan beberapa penelitian terdahulu yang relevan dengan penelitian ini:

Fahreza dan rekan-rekannya melakukan analisis mengenai dampak jam kerja terhadap kesehatan mental generasi Z, dengan tujuan untuk meningkatkan kesadaran publik mengenai pentingnya kesehatan mental, khususnya terkait dengan pengaruh jam kerja. Hal ini bertujuan untuk menciptakan suasana kerja yang baik bagi generasi Z dan masyarakat secara keseluruhan, serta memperbaiki kualitas kesehatan mental generasi Z. Penelitian ini menerapkan metode analisis sentimen menggunakan algoritma *Gaussian Naive Bayes* dan *Support Vector Machine*. Temuan dari penelitian menunjukkan bahwa algoritma *Gaussian Naive Bayes* memperoleh presisi 84%, *recall* 84%, *f1-score* 84%, dengan akurasi juga mencapai 84%. Di sisi lain, *Support Vector Machine* berhasil meraih presisi 91%, *recall* 90%, *f1-score* 90%, dan akurasi 91%. Dari hasil tersebut, dapat disimpulkan bahwa model *Gaussian Naïve Bayes* dan *Support Vector Machine* dengan algoritma *Stemming Sastrawi* memberikan hasil yang lebih baik dibandingkan dengan algoritma Stemming Nazief dan Adriani serta Arifin Setiono. Namun, terdapat beberapa perbedaan antara penelitian tersebut dengan penelitian yang dilakukan penulis. Penelitian sebelumnya berfokus pada isu kesehatan mental dan jam kerja Generasi Z, sedangkan penelitian ini berfokus pada opini publik terhadap *game* Roblox pada media sosial X. Selain itu, penelitian ini juga menggunakan karakteristik data yang berbeda karena data komunitas game memiliki banyak istilah slang, singkatan, dan campuran bahasa Indonesia–Inggris yang lebih kompleks dibandingkan data sosial umum. Adapun gap dari penelitian tersebut

adalah jumlah data yang digunakan masih terbatas dan belum melakukan pengujian optimasi parameter model secara mendalam [11].

Cholifah dkk melakukan analisis sentiment terhadap peresmian UU *Omnibus Law* yang akan berdampak terhadap masyarakat luas. Dengan adanya penelitian ini diharapkan dapat membuat sadar para petinggi terhadap kebijakan-kebijakan yang mereka buat, apakah positif atau negatif bagi masyarakat. Dengan hasil yang didapatkan dari penelitian ini adalah hasil klasifikasi *dataset tweets* terhadap UU *Omnibus Law* pada metode *Naïve Bayes* nilai precisionsebesar 77%, *accuracy* sebesar 75%, dan nilai *recall* sebesar 79%. Sedangkan pada metode SVM nilai *accuracy* sebesar 77%, nilai precision sebesar 88% dan nilai *recall* sebesar 67%. maka dapat disimpulkan bahwa metode SVM memiliki hasil *accuracy* yang lebih tinggi dibandingkan metode *Naïve Bayes*. Namun, terdapat gap dengan penelitian yang dilakukan penulis. Penelitian sebelumnya menggunakan jumlah *dataset* yang relatif lebih sedikit serta hanya menggunakan *kernel polynomial* tanpa melakukan perbandingan kernel lainnya. Selain itu, penelitian tersebut belum menerapkan teknik penanganan data *imbalance* seperti SMOTE dan evaluasi model yang lebih mendalam pada setiap kelas sentimen. Sedangkan penelitian ini menggunakan dataset yang lebih besar, menerapkan SMOTE, serta melakukan perbandingan beberapa kernel SVM untuk memperoleh performa klasifikasi yang lebih optimal.[12].

Galih dkk. melakukan penelitian mengenai analisis sentimen ulasan pengguna aplikasi Roblox di Google Play Store menggunakan metode *Latent Dirichlet Allocation* (LDA) dan *Support Vector Machine* (SVM). Penelitian ini menggunakan 40.089 ulasan berbahasa Indonesia dan berhasil mengidentifikasi empat topik utama, yaitu masalah akun dan gangguan teknis, pengalaman bermain dan interaksi sosial, fitur item dan monetisasi, serta performa aplikasi dan koneksi jaringan. Hasil klasifikasi sentimen menggunakan SVM memperoleh akurasi terbaik sebesar 85,22% pada rasio data 70:30. Namun, penelitian tersebut berfokus pada ulasan pengguna Roblox di Google Play Store dan bertujuan

mengidentifikasi topik pembahasan serta sentimen pengguna terhadap aplikasi Roblox secara umum. Sementara itu, penelitian ini berfokus pada perspektif masyarakat Indonesia terhadap Roblox sebagai media permainan anak menggunakan data dari media sosial X. Selain itu, penelitian sebelumnya belum membahas opini publik mengenai manfaat dan risiko Roblox bagi anak-anak, sehingga penelitian ini dilakukan untuk mengisi kesenjangan tersebut[13].

Freddy dkk membuat sebuah riset analisis mengenai sentiment masyarakat di media sosial x terhadap kemenkes dengan metode Naïve Bayes, dan SVM. Dimana pada riset ini ingin mencari tau mengenai opini masyarakat terhadap kinerja Kemenkes atau Kementria Kesehatan selama periode covid 2020. Dan riset ini menghasilkan Hasil menunjukkan SVM mencapai akurasi yang sedikit lebih tinggi (79%) dibandingkan *Naïve Bayes* (77%), menandakan efektivitasnya dalam menangani struktur bahasa yang kompleks, meskipun memerlukan sumber daya komputasi lebih tinggi. Penelitian ini menyoroti potensi SVM dalam analisis sentimen terhadap kebijakan Kesehatan. Adapun *gap* antara penelitian tersebut dengan penelitian yang penulis lakukan terletak pada teknik klasifikasi, jenis pelabelan, serta karakteristik dataset yang digunakan. Penelitian tersebut hanya menggunakan klasifikasi sentimen biner yaitu positif dan negatif dengan pelabelan otomatis menggunakan VADER, sedangkan penelitian ini menggunakan tiga kelas sentimen yaitu positif, netral, dan negatif sehingga analisis sentimen menjadi lebih rinci[14].

Naufal dkk melakukan Perbandingan Performa *Labeling Lexicon InSet* dan VADER pada Analisa Sentimen Rohingya di Aplikasi X dengan SVM**.** Dimana sentiment ini memiliki tujuan untuk melihat opini dan sentiment yang terjadi di antara masyarakat terhadap hadirnya isu penduduk Rohingya yang bermigrasi ke Indonesia. Pada penelitian ini memperoleh sebuah hasil yang didapatkan *lexicon InSet* melabeli 5241 sentimen negatif, 1369 sentimen positif, 521 sentimen netral. Lalu hasil pelabelan dengan VADER menghasilkan 2749 sentimen positif, 2523 sentimen negatif, dan 1881 sentimen netral. Setelah terlabeli dilakukan training

data dan dihasilkan *lexicon InSet* dengan SVM didapatkan rata rata akurasi 87,83% dan VADER dengan akurasi rata rata 87,66%. Terdapat perbedaan antara penelitian ini dengan penelitian yang dilakukan penulis, yaitu terutama pada evaluasi performa metode klasifikasi dan belum mengkaji secara spesifik perspektif masyarakat terhadap Roblox sebagai media permainan anak. Selain itu, belum ditemukan penelitian yang menganalisis opini masyarakat Indonesia mengenai manfaat dan risiko Roblox bagi anak-anak menggunakan data dari media sosial X. Oleh karena itu, penelitian ini dilakukan untuk mengisi research gap tersebut melalui analisis sentimen terhadap diskusi publik mengenai Roblox di media sosial X[15].

Tabel 2. 1 Penelitian Rujukan

No	Peneliti	Tahun	Tujuan	Metode	Hasil
1	Fahreza dkk	2024	meningkatkan kesadaran masyarakat akan pentingnya. kesehatan mental terutama terkait pengaruh jam kerja, guna menciptakan lingkungan kerja yang sehat bagi generasi Z dan masyarakat umumnya, serta meningkatkan kualitas kesehatan mental.	*Naïve Bayes* dan *Support Vector Machine*	Hasil penelitian menunjukkan bahwa algoritma *Gaussian Naive Bayes* mencapai presisi *84%,* recall *84%, dan* f1-score *84%, dengan* akurasi *84%.* Sementara *Support Vector Machine* mencapai presisi 91%, *recall 90%, f1-score 90%,* dan akurasi 91%
2	Cholifah dkk	2023	Menganalisis opini masyarakat atau public terhadap UU baru, yaitu UU Omnibus Law.	*Support Vector Machine* dan *Naïve Bayes*	Hasil menunjukan pada metode *Naïve Bayes* nilai *precision* sebesar 77%, accuracysebesar 75%, dan nilai *recall* sebesar 79%. Sedangkan pada metode SVM nilai *accuracy* sebesar 77%,

					nilai *precision* sebesar 88% dan nilai *recall* sebesar 67%
3	Galih dkk	2026	Melakukan riset terhadap topik ulasan pengguna aplikasi Roblox menggunakan integrasi LDA dan SVM	*Litech Direct Allocation dan Support Vector Machine*	Hasil menunjukan pada emodelan topik menggunakan LDA berhasil mengidentifikasi empat topik utama dalam ulasan pengguna dengan konfigurasi optimal K=4, yang menghasilkan nilai *coherence score* sebesar 0,5547 dan perplexity sebesar 191,57. Model klasifikasi sentimen menggunakan SVM dengan kernel linearmenunjukkan performa terbaik pada skenario pembagian data 70:30, dengan akurasi sebesar 85,22%, presisi 89%, recall 85%, dan *F1-score* sebesar 86%.
4	Freddy dkk	2025	Melakukan riset terhadap kinerja Kemenkes selama periode covid 2020 melalui opini masyarakat yang diambil dari data aplikasi X.	*Support Vector Machine* dan *Naïve Bayes*	Hasil menunjukkan SVM mencapai akurasi yang sedikit lebih tinggi (79%) dibandingkan *Naïve Bayes* (77%), menandakan efektivitasnya dalam menangani struktur bahasa yang kompleks, meskipun memerlukan sumber daya komputasi lebih tinggi. Penelitian ini menyoroti potensi

					SVM dalam analisis sentimen terhadap kebijakan Kesehatan.
5	Naufal dkk	2024	tujuan untuk melihat opini dan sentiment yang terjadi di antara masyarakat terhadap hadirnya isu penduduk Rohingya yang bermigrasi ke Indonesia.	*Support Vector Machine dan VADER*	hasil yang didapatkan lexicon InSet melabeli 5241 sentimen negatif, 1369 sentimen positif, 521 sentimen netral. Lalu hasil pelabelan dengan VADER menghasilkan 2749 sentimen positif, 2523 sentimen negatif, dan 1881 sentimen netral. Setelah terlabeli dilakukan training data dan dihasilkan lexicon InSet dengan SVM didapatkan rata rata akurasi 87,83% dan VADER dengan akurasi rata rata 87,66%

2.2 Dasar Teori

2.2.1 Game Online

Game online merupakan permainan digital yang dimainkan melalui jaringan internet dan memungkinkan interaksi antar pemain secara daring. *Game online* tidak hanya berfungsi sebagai sarana hiburan, tetapi juga sebagai media interaksi sosial, pembentukan komunitas, serta pengembangan kreativitas pengguna. Dalam konteks anak-anak dan remaja, *game online* dapat melatih kemampuan kognitif seperti *problem solving*, pengambilan keputusan, serta berpikir kritis. Namun, sifat *game online* yang interaktif dan terbuka juga berpotensi menimbulkan dampak negatif apabila tidak disertai dengan pengawasan yang memadai, terutama bagi anak-anak. Karena anak-anak jaman sekarang sudah banyak yang mengenal namanya *game online[16].*

2.2.2 Platform Game Roblox

Roblox merupakan *game* daring berbasis platform yang memungkinkan penggunanya, khususnya anak-anak dan remaja, untuk tidak hanya bermain tetapi juga berkreasi dengan membuat, memodifikasi, serta mengembangkan permainan mereka sendiri melalui fitur pengembangan yang disediakan. Selain itu, Roblox juga berfungsi sebagai ruang sosial *virtual*, di mana para pengguna dapat saling berinteraksi, berkomunikasi, dan bekerja sama dengan pengguna lain dari berbagai latar belakang dan wilayah. Keberadaan berbagai dunia virtual dan konten buatan pengguna menjadikan Roblox memiliki variasi permainan yang sangat luas, mulai dari permainan edukatif hingga permainan kompetitif, sehingga mampu menarik minat pengguna dalam jangka waktu yang panjang. Dengan karakteristik tersebut, Roblox tidak hanya dipandang sebagai sarana hiburan semata, tetapi juga sebagai media pembelajaran digital yang dapat mendorong kreativitas, kemampuan berpikir logis, pemecahan masalah, serta keterampilan sosial anak-anak dan remaja[17]. Namun, di sisi lain, kompleksitas ekosistem dan keterbukaan interaksi di dalam platform ini juga menimbulkan tantangan tersendiri, terutama terkait pengawasan, keamanan, dan

perlindungan anak di ruang digital. Oleh karena itu pemerintah juga mempertegas pentingnya peran orang tua dalam melakukan pengawasan kepada anak dalam interaksinya dengan alat elektronik, melalui Peraturan Pemerintah Republik Indonesia sebagai dasar dalam melakukan pengawasan serta melakukan klasifikasi konten digital yang sesuai dengan umur dari anak, memberikan pemahaman akan pentingnya orang tua sebagai garda terdepan dalam melakukan perlindungan terhadap anak yang dimulaidari dalam lingkungan keluarga perlu diperhatikan efektivitas dari regulasi ini sangat bergantung atas kesadaran dan tanggung jawab orang tua[18].

2.2.3 Media Sosial X

Media sosial kini menjadi sarana utama untuk mengungkapkan pendapat secara terbuka dan cepat. Salah satu platform yang banyak dipakai dalam penelitian adalah X (dulu dikenal sebagai Twitter), yang menyediakan format komunikasi singkat dan langsung. Pengguna media sosial sering kali membagikan pengalaman, mengungkapkan kritik, ataupun memberikan pujian terkait berbagai hal, termasuk komentar, kritik, maupun masukan terhadap platform *game* roblox yang dinilai memiliki nilai yang positif dan negatif[19].

2.2.4 Analisis Sentimen

Analisis sentimen merupakan kegiatan untuk mengevaluasi teks demi mengetahui apakah sebuah pendapat bernada positif, negatif, atau netral. Akan tetapi, melakukan analisis sentimen secara manual menghadapi sejumlah kendala, seperti subjektivitas serta ketidakseragaman yang disebabkan oleh bias kognitif dan emosional. Untuk mengatasi masalah ini, pendekatan yang berlandaskan pembelajaran mesin atau kecerdasan buatan (AI) semakin banyak digunakan untuk meningkatkan ketepatan dan efektivitas analisis sentimen.

2.2.5 Deep Translator

Deep Translator merupakan library Python yang digunakan untuk menerjemahkan teks secara otomatis dari satu bahasa ke bahasa lain melalui layanan penerjemahan seperti Google Translate. Dalam penelitian analisis sentimen, Deep Translator sering digunakan untuk mengubah data berbahasa non-Inggris menjadi bahasa Inggris agar dapat diproses oleh metode analisis sentimen yang dikembangkan khusus untuk bahasa Inggris, seperti VADER. Pada penelitian ini, Deep Translator digunakan untuk menerjemahkan tweet berbahasa Indonesia ke bahasa Inggris sebelum proses pelabelan sentimen menggunakan VADER. Hal ini dilakukan agar hasil analisis sentimen lebih sesuai dengan kamus leksikon yang digunakan oleh VADER.

2.2.6 Pre-Processing

Tahap *pre-processing* merupakan tahap awal dalam proses *machine learning* yang bertujuan untuk menyiapkan data sebelum dilakukan klasifikasi. Pada tahap ini, data mentah yang masih belum terstruktur diolah agar menjadi data yang lebih bersih, konsisten, dan siap digunakan dalam proses analisis. Proses *pre-processing* meliputi pembersihan data, penggabungan data, perubahan format data, serta pengurangan data yang tidak diperlukan. Tahapan ini dilakukan untuk mengurangi gangguan (*noise*) dan meningkatkan kualitas data sehingga hasil analisis menjadi lebih akurat. Pada penelitian ini, data ulasan dikumpulkan secara otomatis melalui proses *scraping* dan disimpan dalam format CSV. Selanjutnya, data tersebut diproses melalui beberapa tahapan *pre-processing*, yaitu:

1. **Cleaning**

   Proses pembersihan data dilakukan untuk menghapus data yang tidak relevan, tidak lengkap, atau duplikat agar data lebih siap untuk dianalisis.

2. **Case Folding**

   Tahap ini mengubah seluruh huruf dalam teks menjadi huruf kecil agar penulisan kata menjadi seragam.

3. **Normalization**

   Normalisasi dilakukan untuk mengubah kata tidak baku atau singkatan menjadi

4. **Labeling**

Tahap pelabelan dilakukan dengan memberikan label sentimen pada setiap data ulasan, yaitu positif, negatif, atau netral, untuk mendukung proses klasifikasi. Merupakan serangkaian tahapan *pre-processing*[20].

2.2.7 Metode Support *Vector* Machine

Metode yang menggunakan *machine learning* ini adalah algoritma yang efektif untuk melakukan analisis sentimen. Metode ini biasanya digunakan dalam hal klasifikasi teks dengan mencari *hyperlane* atau garis batas pemisah yang dapat memisahkan dua kelas data misalnya sentimen positif dan negatif. SVM adalah metode learning machine yang bekerja atas prinsip Structural Risk Minimization (SRM) dengan tujuan menemukan *hyperplane* terbaik yang memisahkan dua buah class pada input space. Metode *Support Vector Machine* dikenal sebagai teknik pembelajaran mesin paling mutakhir dan berhasil setelah pembelajaran mesin sebelumnya yakni *Neural Network* dalam pengenalan pola sebagai usaha mencari fungsi pemisah (*hyperlane*) dengan memaksimalkan *margin* atau jarak antar kelas teks. Pada awalnya, metode SVM dikembangkan untuk hal mengenai klasifikasi dua kelas, kemudian dikembangkan kembali untuk klasifikasi bermacam-macam kelas.

2.2.8 Bahasa Pemrograman Python

Python adalah bahasa pemrograman interpretatif multiguna dengan filosofi perancangan yang berfokus pada tingkat keterbacaan kode. Python diklaim sebagai bahasa yang menggabungkan kapabilitas, kemampuan, dengan sintaksis kode yang sangat jelas, dan dilengkapi dengan fungsionalitas pustaka standar yang besar serta komprehensif. Bahasa pemrograman Python adalah bahasa pemrograman tingkat tinggi yang dikembangkan pada tahun 1991 oleh Guido Van Rossum. Bahasa Python memiliki sintaks yang mudah dibaca dan dipahami, biasanya banyak digunakan untuk keperluan pengembangan web hingga analisis data. Bahasa pemrograman Python cocok juga digunakan sebagai *scripting language,* implementasi bahasa pemrograman web, bahkan bisa di

extended ke bahasa pemrograman lain seperti C dan C++. Dalam menerapkan bahasa pemrograman Python, terdapat beberapa aturan yang harus dipenuhi, untuk mengantisipasi adanya *error* pada program yang dibuat. Oleh karena itu, aturan sintaks dari bahasa pemrograman Python yang pertama adalah penulisan suatu perintah atau *statement* yang mendukung kode bahasa pemrograman tersebut.