

BAB II

TINJAUAN PUSTAKA

2.1. Penelitian Terkait

Dalam penyusunan penelitian yang dilakukan, penulis mengumpulkan penelitian yang berkaitan dengan latar belakang penelitian dari penelitian-penelitian terdahulu sebagai referensi dan acuan dalam membuat penelitian dengan menggunakan *Support Vector Machine* dan *Adaboost*. Beberapa penelitian terkait yang menjadi *key paper* pada penelitian ini. Berikut adalah penelitian terkait yang menjadi dasar dalam penelitian ini.

Tabel 2. 1. Penelitian Terkait

No	Peneliti	Judul Penelitian	Tahun	Metode	Hasil
1	M. Octa Prasetya, Usman Ependi [3]	Sentimen Analisis Hero Mobile Legends Dengan Algoritma Naive Bayes	2022	Algoritma Naive Bayes	scrapped data berhasil mengumpulkan data yang berasal dari para pengguna game mobile legends di playstore yang bermula sebanyak 56911 data dan di ambil sebanyak 100 data yang di jadikan data sampel pengguna
2	Daurat Sinaga, Cahaya Jatmoko [1]	Analisis Sentimen Untuk Mengetahui	2020	Naïve Bayes Classifier	diperoleh 2500 sentimen, dengan 400 data diantara digunakan sebagai

		Kesan Player Game Mobile Legends Menggunakan Naïve Bayes Classifier			data uji. Setelah proses normalisasi, dilakukan proses case folding, tokenizing, stemming dan filtering. Hasil akurasi pengukuran probabilitas komentar user, diperoleh akurasi sebesar 80%, presisi sebesar 76% dan recall sebesar 90,4 %
3	Selva Indah Nurhafida, Falentino Sembiring [5]	Analisis Sentimen Aplikasi Novel Online Di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)	2022	Support Vector Machine (SVM)	Analisis deskriptif yang diperoleh dengan rating total 4.137 ulasan pengguna Wattpad mengungkapkan bahwa 33,12 persen pengguna sangat tidak suka, 17,04 persen tidak suka, 15,71 persen netral, 11,31 persen suka, dan 22,82 persen suka, sedangkan aplikasi Dreame menerima 10,94 persen pengguna

					sangat tidak suka, 13,24 persen tidak suka, 27,86 persen netral, 21,85 persen suka, dan 26,12 persen pengguna sangat menyukai dari 3.090 ulasan.
4	Herodion Simorangkir, Kemas Muslim Lhaksmana	Analisis Sentimen pada Twitter untuk Games Online Mobile Legends dan Arena of Valor dengan Metode Naïve Bayes Classifier	2018	Naïve Bayes Classifier	Hasil prediksi tweet data Mobile Legends yaitu tweet positif ada sebanyak 33 tweet, dan tweet negatif ada sebanyak 44 tweet. Hasil prediksi Arena of Valor yaitu tweet positif ada sebanyak 54 tweet, dan tweet negatif ada sebanyak 151 tweet. Nilai akurasi, eror, recall, dan precision yang didapat masing-masing sebesar 88,89%, 19,18%, 96,97%, dan 69,57% untuk Mobile Legends sedangkan Arena of

					Valor memiliki nilai akurasi, eror, recall, dan precision masing-masing sebesar 39,02%, 60,98%, 88,89% dan 28,74%
5	Hartati, Deni Hermawan, M. Akhsanal, Zailani Wahyudi, Angga Ariyanto, Dedi Dwi Saputra [7]	Optimasi Analisis Sentimen Pada Twitter Olshop Tokopedia Menggunakan Textmining Dengan Algoritma Naïve Bayes & Adaboost	2022	Algoritma Naïve Bayes & Adaboost	Algoritma Naïve Bayes yang ditambahkan feature Synthetic Minority Over-sampling Technique (SMOTE), yang dioptimasi dengan Adaboost mendapatkan nilai terbaik. Dengan nilai akurasi 94.95%, Presisi 90.86%, Recall 100.00% dan AUC 0,950

Dalam analisis hasil pada penelitian terdahulu dapat terlihat bahwa penggunaan dua algoritma Algoritma Support Vector Machine & Adaboost memiliki hasil akurasi yang relatif sangat baik dari pada penggunaan satu metode dalam sebuah penelitian. Kemudian setelah mengkomparasi jurnal pada penelitian sebelumnya peneliti juga mengkomparasi tiap – tiap algoritma yang menjadi refrensi dalam melakukan penelitian ini.

2.2. Mobile Legend

Mobile Legend merupakan *game* MOBA yang dirilis oleh Moonton. *Mobile Legends* adalah permainan MOBA yang dirancang untuk ponsel. *Game* ini bisa dimainkan di ponsel Android maupun IOS. Pada alur permainan mobile legend terbagi menjadi dua tim dimana Kedua tim masing-masing berisi lima orang berjuang untuk mencapai dan menghancurkan markas musuh sambil mempertahankan markas mereka sendiri untuk mengendalikan tiga jalur, yang dikenal sebagai jalur "atas", "tengah" dan "bawah", yang menghubungkan ke setiap markas. Rata-rata permainan dalam 1 *game* yakni sekitar 15-20 menit. Setiap pemain bisa memilih 1 hero dari puluhan daftar hero yang tersedia. Pada dasarnya, karakter dalam *Mobile Legends* dibedakan menjadi 6 peran (*role*) di antaranya: marksman, fighter, tank, mage, assassin dan support. Setiap peran mempunyai spesialisasi dan kemampuan yang unik. Total jumlah hero per Desember 2022, baik pada server orisinal (*original server*) maupun lanjutan (*advance server*) adalah 118 hero. Pertumbuhan pengguna Mobile Legends pada tahun 2016 terbilang sangat pesat sekali. Mobile Legends dimainkan sebanyak 30 juta orang, sedangkan pada tahun 2018, mencapai 200 juta orang. Pada App Store jumlah pengguna yang mendownload aplikasi pada saat ini mencapai lebih dari 500 juta download sehingga Jumlah tersebut membuat Mobile Legends selalu menduduki peringkat pertama di playstore. Pada ulasan play store sendiri terdapat 31 juta ulasan masyarakat mengenai game mobile legend tersebut. Pada ulasan masyarakat mengenai game mobile legend ini sendiri pasti memiliki komentar dan pendapat masyarakat yang berbeda-beda baik itu saran, support ataupun dukungan dan tak jarang pada ulasan tersebut terdapat komentar-komentar negatif mengenai game ini [8].

2.3. Text Mining

Text mining adalah salah satu teknik yang dapat digunakan untuk melakukan klasifikasi dokumen, clustering, information extraction, analisis sentimen dan information retrieval dimana text mining merupakan variasi dari data mining yang berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar. *Text mining* dapat memberikan solusi dari permasalahan seperti pemrosesan, pengorganisasian atau pengelompokkan dan menganalisa *unstructured*

text dalam jumlah besar. Dalam memberikan solusi, *text mining* mengadopsi dan mengembangkan banyak teknik dari bidang lain, seperti *Data mining*, *Information Retrieval*, Statistik dan Matematik, *Machine Learning*, *Linguistic*, *Natural Language Processing* (NLP), dan *Visualization*. Tujuan dari *Text Mining* adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen, tetapi tujuan utama *text mining* adalah mendukung proses *knowledge discovery* pada koleksi dokumen yang besar. Adapun tugas khusus dari *text mining* antara lain yaitu pengkategorisasian teks (*text categorization*) dan pengelompokkan teks (*text clustering*). *text mining* memiliki peran penting dalam bidang data mining. Dengan mengaplikasikan proses-proses dalam *text mining*, maka akan diperoleh pola-pola data, tren, dan ekstraksi dari pengetahuan-pengetahuan yang potensial dari data teks.[9] Diantara proses yang dapat dilakukan dalam *text mining* adalah klasifikasi teks. Klasifikasi teks dapat didefinisikan sebagai proses untuk menentukan suatu dokumen teks ke dalam suatu kelas tertentu. Untuk melakukan proses klasifikasi teks, ada beberapa algoritma yang dapat digunakan diantaranya *Support Vector Machine* (SVM), *Naive Bayes*, *k-Nearest Neighbor* (KNN), *Decision Tree*, dan *Artificial Neural Networks* (ANN) [9].

2.4. Sentiment Analysis

Sentiment analysis (*opinion mining*) merupakan suatu bidang yang luas dari pengolahan bahasa alami, komputasi linguistik dan *text mining* dimana memiliki tujuan dalam menganalisa pendapat, *sentiment*, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenaan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu [10]. Analisis Sentimen atau bisa disebut dengan *opinion mining* juga dapat memahami proses, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung berpandangan atau beropini negatif atau positif. *Sentiment analysis* memiliki tugas dasar dalam mengelompokkan teks yang ada pada sebuah kalimat atau dokumen kemudian menentukan pendapat yang dikemukakan dalam kalimat atau dokumen tersebut apakah bersifat positif, negatif atau netral [10]

2.5. Support Vector Machine (SVM)

Support Vector Machine (SVM) diperkenalkan oleh Vapnik pada tahun 1992 sebagai suatu teknik klasifikasi yang efisien untuk masalah nonlinear. Support Vector Machine (SVM) juga dikenal sebagai teknik pembelajaran mesin (machine learning) paling mutakhir setelah pembelajaran mesin sebelumnya yang dikenal sebagai Neural Network (NN). Baik SVM maupun NN tersebut telah berhasil digunakan dalam pengenalan pola. Pembelajaran dilakukan dengan menggunakan pasangan data input dan data output berupa sasaran yang diinginkan [11]. Konsep SVM dapat dijelaskan secara sederhana sebagai usaha mencari hyperplane terbaik yang berfungsi sebagai pemisah dua buah kelas pada input space. SVM berusaha menemukan fungsi pemisah (hyperplane) dengan memaksimalkan jarak antar kelas. Dengan cara ini, SVM dapat menjamin kemampuan generalisasi yang tinggi untuk data-data yang akan olah [12].

Dalam referensi [13] Berikut adalah penjelasan lebih rinci tentang algoritma SVM:

1. Konsep Dasar SVM

SVM bekerja dengan mencari hyperplane terbaik yang memisahkan dua kelas data dengan margin maksimal. Hyperplane ini ditemukan dengan memaksimalkan jarak antara hyperplane dan titik-titik terdekat dari kedua kelas, yang disebut vektor pendukung (support vectors). SVM dapat menangani baik masalah klasifikasi biner maupun multikelas, dengan memanfaatkan teknik seperti One-vs-All (OvA) atau One-vs-One (OvO) untuk mengatasi masalah multikelas.

2. Fungsi Kernel

Salah satu keunggulan utama dari SVM adalah kemampuannya untuk menangani data yang tidak linear dengan menggunakan fungsi kernel. Fungsi kernel mengubah data ke dalam ruang fitur yang lebih tinggi sehingga memungkinkan untuk menemukan hyperplane yang dapat memisahkan data dengan baik, meskipun dalam ruang dimensi yang lebih tinggi.

3. Margin dan Dukungan

Margin adalah jarak antara hyperplane dan titik-titik terdekat dari kedua

kelas. SVM berusaha untuk memaksimalkan margin ini untuk meningkatkan generalisasi model. Dukungan (support) vectors adalah titik-titik yang terletak paling dekat dengan hyperplane pembagi dan berperan penting dalam menentukan hyperplane tersebut.

4. Penalti dan Regularisasi

SVM menggunakan parameter C sebagai penalti terhadap kesalahan klasifikasi. Nilai C yang lebih tinggi menyebabkan model lebih memperhatikan kesalahan klasifikasi pada data latih, sementara nilai C yang lebih rendah memprioritaskan margin yang lebih besar. Regularisasi dilakukan untuk mencegah overfitting dengan mengontrol kompleksitas model. Hal ini dicapai dengan mengatur nilai C dan memilih fungsi kernel yang sesuai.

5. Keuntungan dan Keterbatasan

Keuntungan utama dari SVM adalah kemampuannya untuk bekerja dengan baik dalam kasus data yang tidak linear dan memiliki banyak fitur. Namun, SVM memiliki kelemahan dalam kinerja komputasionalnya, terutama pada dataset yang sangat besar. Selain itu, pemilihan parameter seperti C dan pilihan kernel dapat mempengaruhi hasil dari model SVM.

Dengan memahami konsep dasar dan prinsip kerja SVM, peneliti dapat menerapkan algoritma ini secara efektif dalam analisis sentimen untuk mengklasifikasikan teks ke dalam kategori sentimen yang sesuai. Algoritma SVM menjadi salah satu pilihan yang populer dalam analisis sentimen karena kemampuannya untuk menangani masalah klasifikasi dalam ruang fitur yang kompleks [14].

2.6. Adaboost

Boosting merujuk pada *ensemble learning* yang mengkombinasikan beberapa *weak learner* menjadi *strong learner*. Konsep dari metode *boosting* adalah melatih prediktor secara berurutan, yang setiap prediktor mengoreksi prediktor sebelumnya. *Boosting* melakukan pengulangan secara serial, yang mana kesalahan klasifikasi dari iterasi sebelumnya digunakan untuk memengaruhi pelatihan pengklasifikasi pada iterasi berikutnya dengan memberikan bobot lebih berat pada kesalahan klasifikasi tersebut pada langkah selanjutnya. *Adaboost* menjadi metode *boosting*

yang banyak digunakan, karena sederhana dan adaptif. *Adaboost* dikatakan adaptif karena jumlah *update* ditentukan sebagai fungsi dari kesalahan bobot dari *hipoTESIS* yang dihasilkan oleh *weak learner*. *Adaboost* didefinisikan sebagai algoritma pengulangan yang menggunakan pengklasifikasi berbeda untuk set pelatihan yang sama, lalu menyatukannya untuk membuat pengklasifikasi terkuat diakhir. Algoritma *Adaboost* dibuat khusus untuk menangani masalah klasifikasi yang berfungsi untuk meningkatkan akurasi *weak learner*. *Adaboost* memiliki beberapa kelebihan, yaitu sederhana serta mudah diimplementasikan [12].

2.7. Confusion Matrix

Confusion Matrix adalah sebuah metode yang biasa digunakan untuk perhitungan akurasi, *recall*, *precision*, dan *error rate*. Dimana, *precision* mengevaluasi kemampuan sistem untuk menemukan peringkat yang paling relevan, dan didefinisikan sebagai persentase dokumen yang di *retrieve* dan benar-benar relevan terhadap *query*. *Recall* mengevaluasi kemampuan sistem untuk menemukan semua item yang relevan dari koleksi dokumen dan didefinisikan sebagai persentase dokumen yang relevan terhadap *query*. *Accuracy* merupakan perbandingan kasus yang diidentifikasi benar dengan jumlah seluruh kasus dan *error rate* merupakan kasus yang diidentifikasi salah dengan jumlah seluruh kasus. Dalam referensi [14] Rumus untuk *confusion matrix* dapat dilihat pada tabel dibawah ini:

Tabel 2. 2. Rumus *Confusion Matrix*

Document	Nilai Sebenarnya	
	Revant	Non Revant
Retrieved	<i>True Positive (tp)</i> <i>Correct result</i>	<i>False Positive (fp)</i> <i>Unexpected result</i>
Not Retrieved	<i>False Negative (fn)</i> <i>Missing result</i>	<i>True Negative (tn)</i> <i>Corect absence of resul</i>

Keterangan :

1. **TP (True Positive)** : Jumlah prediksi yang benar dari data yang relevant

2. **FP (False Positive)** : Jumlah prediksi yang salah dari data yang tidak relevant
3. **FN (False Negative)** : Jumlah prediksi yang salah dari data yang tidak relevant
4. **TN (True Negative)** : Jumlah prediksi yang benar dari data yang relevant.

Dalam penelitian [15] dijelaskan bahwa penggunaan *multiclass confusion matrix* 3x3 dikarenakan *output sentiment* dari penelitian ini ada 3 *sentiment*, yaitu positif, negatif, netral, sehingga tabel *confusion matrix*-nya adalah sebagai berikut:

Tabel 2. 3. Rumus Multiclass *Confusion Matrix*

		PREDIKSI		
		POSITIF	NEGATIF	NETRAL
AKTUAL	POSITIF	TPos	FPosNeg	FPosNet
	NEGATIF	FNegPos	TNeg	FNegNet
	NETRAL	FNetPos	FNetNeg	TNet

Rumus dalam perhitungan Confusion Matrix sendiri terdiri dari dasar tabel terdiri dari akurasi, precision dan recall. Akurasi merupakan metode pengujian berdasarkan tingkat kedekatan antara nilai prediksi dengan nilai aktual. Dengan mengetahui jumlah data yang diklasifikasikan secara benar maka dapat diketahui akurasi hasil prediksi. Persamaan akurasi ditunjukkan pada persamaan berikut.

$$\text{Akurasi} = \frac{TP + TN}{TP + FN + FP + TN} \times 100\%$$

Kemudian presisi merupakan metode pengujian dengan melakukan perbandingan jumlah informasi relevan yang didapatkan sistem dengan jumlah seluruh informasi yang terambil oleh sistem baik yang relevan maupun tidak. Persamaan precision ditunjukkan pada persamaan berikut

$$\text{Precision} = \frac{TP}{TP + FP}$$

Dan yang terakhir adalah recall merupakan metode pengujian yang membandingkan jumlah informasi relevan yang didapatkan sistem dengan jumlah seluruh informasi relevan yang ada dalam koleksi informasi (baik yang terambil atau tidak terambil oleh sistem). Persamaan recall ditunjukkan pada persamaan berikut

$$\text{Recall} = \frac{TP}{TP + FN}$$

2.8. Google Play Store

Pada awalnya google play adalah Android Market yaitu layanan distribusi digital yang dioperasikan dan dikembangkan oleh Google. Layanan ini berfungsi sebagai toko aplikasi resmi untuk sistem operasi Android, yang memungkinkan pengguna untuk menelusuri dan mengunduh aplikasi yang dikembangkan dengan Android software development kit (SDK) dan diterbitkan melalui Google. Google Play juga berfungsi sebagai toko media digital, yang menawarkan program musik, buku, film, dan televisi. [16] Ini sebelumnya menawarkan perangkat keras Google untuk pembelian sampai diperkenalkannya pengecer perangkat keras online yang terpisah, Google Store, pada 11 Maret 2015, dan juga menawarkan publikasi berita dan majalah sebelum perbaikan Google News pada 15 Mei 2018. Aplikasi tersedia melalui Google Play baik gratis atau dengan biaya. Mereka dapat diunduh langsung pada perangkat Android melalui aplikasi seluler Play Store atau dengan menyebarkan aplikasi ke perangkat dari situs web Google Play. Aplikasi yang mengeksploitasi kemampuan perangkat keras dari suatu perangkat dapat ditargetkan untuk pengguna perangkat dengan komponen perangkat keras tertentu, seperti sensor gerak (untuk game yang bergantung pada gerakan) atau kamera yang menghadap ke depan. Dalam unggahan aplikasi yang disediakan google play terdapat ulasan masyarakat mengenai komentar kepuasan dari penggunaan aplikasi yang disediakan oleh google play store. Data-data ulasan pada google play store tersebut dapat dimanfaatkan untuk membuat analisis review pengguna yang sangat menarik bagi pemilik aplikasi untuk mengambil keputusan di masa depan[17].

