BAB II

TINJAUAN PUSTAKA

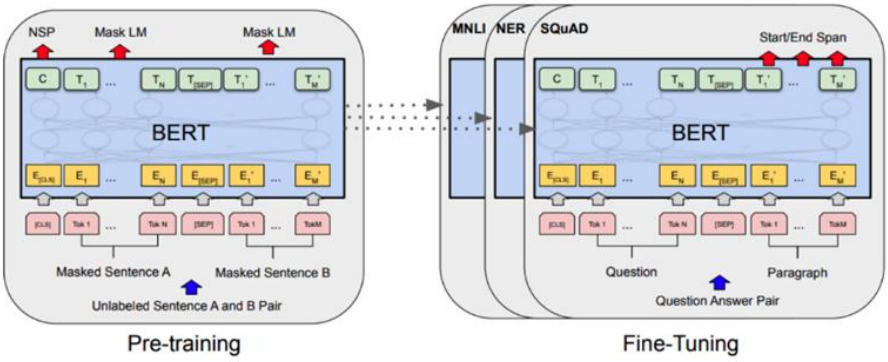
2.1 Landasan Teori

2.1.1 Sarkasme

Sarkasme adalah salah satu jenis majas yang mengandung kata-kata pedas dengan tujuan menyakiti hati orang lain, baik dalam bentuk ejekan maupun cemoohan (Merlina & Dewi, 2022). Kata sarkasme berasal dari bahasa Yunani sarkasmos, yang diturunkan dari kata kerja sarkasein, yang berarti ”merobek-robek daging seperti anjing”, ”menggigit bibir karena marah”, atau ”berbicara dengan kepahitan” (Dinari, 2015). Selain itu, secara etimologis, istilah sarkasme juga berasal dari bahasa Prancis sarcasmus atau sarkazo, yang memiliki makna ”daging atau hati yang tertusuk” (Nasrudin, 2019). Dibandingkan dengan ironi dan sinisme, sarkasme memiliki ungkapan yang lebih tajam, sering disertai nada pahit dan celaan pedas. Gaya bahasa ini bisa bersifat ironis maupun tidak, tetapi selalu memiliki potensi melukai perasaan dan terdengar tidak menyenangkan (Hilmawan, 2022). Sarkasme biasanya digunakan untuk menyindir. Selain itu, sarkasme juga dapat diarahkan pada situasi atau ide. Penggunaan sarkasme sering kali menjadi cara untuk mengungkapkan ekspresi yang tidak dapat disampaikan secara langsung (Dinari, 2015). Sarkasme juga termasuk dalam kategori majas pertentangan, sehingga sering digunakan untuk menyampaikan kalimat yang memiliki makna berlawanan dengan apa yang diucapkan. Pernyataan seperti ini mencerminkan adanya pertentangan antara makna kalimat yang diucapkan dengan maksud yang sebenarnya (Eke dkk., 2021).

2.1.2 IndoBERT

BERT (*Bidirectional Encoder Representations from Transformers*) merupakan model deep learning berbasis arsitektur Transformer (Devlin dkk., 2019). Salah satu pengembangan dari BERT adalah IndoBERT, yang dilatih pada korpus besar berbahasa Indonesia (Koto dkk., 2020). IndoBERT dikembangkan untuk mengatasi keterbatasan model NLP (*Natural Language Processing*) sebelumnya yang kurang optimal dalam memahami struktur dan konteks bahasa Indonesia. Model ini mengintegrasikan lebih dari 220 juta kata yang diambil dari sumber terpercaya berbahasa Indonesia, termasuk surat kabar daring dan Korpus Web Indonesia. IndoBERT telah dikembangkan dengan 2,4 juta langkah atau 180 epoch, sehingga memiliki kinerja yang kuat pada berbagai tugas NLP. Model *pre-train* ini memanfaatkan teknik MLM (*Masked Language Modeling*) dan NSP (*Next Sentence Prediction*), yang memungkinkan model memahami hubungan antar kata dan antar kalimat dalam teks. IndoBERT dilatih menggunakan kosakata WordPiece dalam bahasa Indonesia yang mencakup 31.923 token dan menggunakan arsitektur transformer dengan 12 lapisan dan 768 unit pemrosesan (Asyrofiyyah & Sugiharti, 2024).



Gambar 2.1 *Pre-training* dan *Fine-Tuning* BERT (Devlin dkk., 2019)

Secara garis besar BERT memiliki dua tugas utama, yaitu pre-training dan fine-tuning, seperti ditampilkan pada Gambar 2.1. Selain itu, BERT juga termasuk model yang menggunakan arsitektur Transformer. Arsitektur transformer, menggunakan dua mekanisme terpisah, yaitu encoder dan decoder. Encoder berfungsi untuk membaca input teks, sedangkan decoder digunakan untuk menghasilkan prediksi. Namun, karena tujuan model BERT adalah menghasilkan representasi bahasa, hanya mekanisme encoder yang diperlukan. Selain itu, berbeda dengan model sequential yang membaca input teks secara searah, baik dari kiri ke kanan atau sebaliknya secara berurutan, model transformer membaca input teks secara dua arah. Hal ini memungkinkan model untuk memahami konteks secara lebih menyeluruh (Putra dkk., 2021).

2.1.3 Optimasi *Hyperparameter*

Hyperparameter adalah variabel yang memengaruhi *output* dari sebuah model. Nilai *hyperparameter* harus ditetapkan terlebih dahulu sebelum model menjalani proses *training.* Nilai *hyperparameter* ditentukan dengan eksplisit pada saat pendefinisian model dan tidak diperoleh melalui proses *training* seperti parameter model lainnya (Nugraha & Sasongko, 2022). Optimasi *hyperparameter* atau *hyperparameter tuning* adalah proses pencarian secara sistematis untuk mencari kombinasi *hyperparameter* yang optimal dari model machine learning atau deep learning. Contoh *hyperparameter* yaitu, learning rate, jumlah hidden layers pada neural network, atau jumlah tree pada random forest. Tujuan akhir dari optimasi *hyperparameter* adalah untuk meningkatkan kinerja model pada tugas-tugas seperti klasifikasi atau regresi dengan menemukan konfigurasi yang

meminimalkan *error metrics* (seperti *loss* atau *accuracy*) pada set validasi (Bartz-Beielstein & Zaefferer, 2023).

2.1.4 Bayesian *Optimization*

Bayesian *Optimization* adalah teknik optimasi berbasis model probabilistik yang digunakan untuk menemukan maksimum (atau minimum) dari fungsi *black-box* yang mahal. Fungsi-fungsi ini disebut *"black-box"* karena tidak memiliki bentuk analitis yang diketahui dan evaluasinya dapat memakan biaya tinggi, baik dalam hal waktu maupun sumber daya. Bayesian *Optimization* menggunakan *surrogate* model seperti GP (*Gaussian Process*) (Chuong B. Do, 2007) atau TPE (*Tree Parzen Estimation*) (Bergstra dkk., 2011) untuk memodelkan fungsi objektif berdasarkan data yang telah diamati. Fungsi akuisisi, seperti EI (*Expected Improvement*) (Zhou dkk., 2023) dan UCB (*Upper Confidence bBound*) (Srinivas dkk., 2010), memandu evaluasi berikutnya dengan menyeimbangkan eksplorasi (mencari area baru dengan potensi tinggi) dan eksploitasi (mengoptimalkan area yang sudah diketahui menjanjikan). Bayesian Optimization secara iteratif memperbarui model dan menghitung ulang fungsi akuisisi hingga kriteria penghentian terpenuhi, sehingga mengoptimalkan fungsi *black-box* yang mahal secara efisien (Nguyen, 2019). Algoritma Bayesian *Optimization* dirangkum dalam *Algorithm* 1, menjelaskan langkah–langkah utama penerapannya.

Algoritma 1 Bayesian *Optimization* (Nguyen, 2019).

Input: initial data $\mathcal{D}_0$, #iter T

Output: x_{max}, y_{max}

1 **for** $t = 1$ to T **do**

2 Fit a GP from $\mathcal{D}_t$ and obtain $x_t = \arg\max \alpha(x)$

3 Evaluate $y_t = f(x_t)\ and\ \mathcal{D}_t = \mathcal{D}_{t-1} \cup (x_t, y_t)$

4 **end for**

Data awal $\mathcal{D}_0$ merupakan pasangan nilai *input* dan *output* fungsi objektif (x, y), serta T merupakan banyaknya iterasi. Langkah 1 sampai 4 merupakan proses iterasi sebanyak T kali. Untuk setiap iterasinya, algoritma memodelkan fungsi objektif *f(x)* menggunakan GP (Gaussian *Process*) berdasarkan dataset saat ini $\mathcal{D}_t$. GP menghasilkan prediksi nilai fungsi dan tingkat ketidakpastian di seluruh ruang input. Berdasarkan model ini, algoritma memaksimalkan fungsi akuisisi $\alpha(x)$ untuk memilih nilai x_t berikutnya yang akan dievaluasi. Setelah itu, nilai fungsi $y_t = f(x_t)$ dihitung, dan pasangan data baru (x_t, y_t) ditambahkan ke dataset $\mathcal{D}_t$. Langkah 2 sampai 3 terus diulang sampai jumlah iterasi maksimum. Pada akhir iterasi, algoritma menghasilkan titik terbaik yang ditemukan selama proses optimasi, berupa nilai maksimum fungsi y_{max} dan lokasi x_{max}.

2.1.5 Hyperband

Hyperband adalah algoritma yang digunakan untuk optimasi *hyperparameter*, yang menggabungkan *random search* dengan strategi alokasi sumber daya yang adaptif. Algoritma ini bertujuan mengidentifikasi konfigurasi *hyperparameter* terbaik untuk model *machine learning* secara efisien dengan memanfaatkan teknik *early stopping*. Metode Hyperband dikembangkan untuk mengatasi keterbatasan algoritma SH (*Successive Halving*) (Soper, 2022), khususnya dalam memilih konfigurasi yang relevan secara dinamis (Li dkk., 2018). Tujuan utama Hyperband adalah mencapai keseimbangan optimal antara jumlah

konfigurasi *hyperparameter* (n) dan alokasi *budget* (b) dengan membagi total *budget* (B) menjadi n bagian, di mana setiap bagian dialokasikan ke struktur tertentu (b=B/n). Pendekatan ini SH menjalankan subproses pada setiap struktur yang dihasilkan secara acak, menghapus konfigurasi *hyperparameter* dengan kinerja buruk, dan mengalokasikan lebih banyak sumber daya pada konfigurasi yang menjanjikan, sehingga meningkatkan *accuracy* keseluruhan (Abdellatif dkk., 2022). Algoritma Hyperband dirangkum dalam *Algorithm* 2, menjelaskan langkah-langkah utama penerapannya.

Algorithm 2 Hyperband Algorithm Using SH (Abdellatif dkk., 2022).

Input: budgets b_{min}, b_{max}, and η

Output: best configurations

$s_{max} = \left\lfloor log_{\eta} \frac{b_{max}}{b_{min}} \right\rfloor$

for $s \in (s_{max}, s_{max} - 1, \ldots, 0)$ **do**

    determine budgets $n = \left\lceil \frac{s_{max}+1}{s+1} * \eta^{s} \right\rceil$

    γ = sample configurations(n)

    run SH on γ with $\eta^{s}.b_{max}$ as initial budget

end

return best configuration

Batas *budget* b_{max} dan b_{min} ditentukan oleh keseluruhan data, *budget* yang tersedia, dan jumlah minimum kasus yang diperlukan untuk melatih model yang masuk akal. Pada langkah kedua dan ketiga *Algorithm* 2, ukuran *budget* dan jumlah konfigurasi (n) untuk setiap percobaan dihitung berdasarkan batasan budget

tersebut. Selanjutnya, pada langkah keempat dan kelima, konfigurasi diambil sampelnya berdasarkan alokasi anggaran (b) dan jumlah konfigurasi (n), kemudian dikirim ke model SH. Algoritma SH akan memilih konfigurasi yang menunjukkan kinerja terbaik untuk diteruskan ke iterasi berikutnya, sekaligus mengeliminasi konfigurasi yang berkinerja buruk. Hyperband akan berhenti ketika ditemukan konfigurasi *hyperparameter* terbaik atau ketika jumlah iterasi maksimum tercapai. Kompleksitas waktu Hyperband dihitung menggunakan notasi big-O, dengan nilai kompleksitas sebesar $O(n \log n)$, yang termasuk teknik pencarian dari SH (Li dkk., 2018).

2.1.6 BOHB (Bayesian Optimization Hyperband)

Bayesian *Optimization* Hyperband (BOHB) adalah teknik optimasi yang menggabungkan kelebihan Bayesian *Optimization* (Nguyen, 2019) dan Hyperband (Li dkk., 2018) untuk secara efisien mencari konfigurasi nilai *hyperparameter* yang optimal dalam model *machine learning*. BOHB memanfaatkan *framework* Hyperband, yang mengalokasikan *budget* tetap dan menggunakan *successive halving* untuk secara bertahap mengeliminasi konfigurasi yang kurang menjanjikan. Model probabilistik dari Bayesian Optimization diintegrasikan untuk memandu pengambilan sampel, meningkatkan kemungkinan pemilihan konfigurasi yang optimal. Dengan menyeimbangkan eksplorasi dan eksploitasi, BOHB menunjukkan kinerja yang kuat pada *budget* pendek maupun panjang, menjadikannya alat yang efektif untuk mengoptimalkan *hyperparameter* dalam tugas-tugas *machine learning* yang kompleks (Falkner dkk., 2018). Algoritma

Bayesian *Optimization* Hyperband dirangkum dalam *Algorithm* 3, menjelaskan langkah – langkah utama penerapannya.

Algorithm 3: Pseudocode for Sampling in BOHB (Falkner dkk., 2018).

Input: observations D, fraction of random runs ρ, precentile q, number of samples N_s, minimum number of points N_{min} to build a model, and bandwidth factor b_w

Output: next configuration to evaluate

1 **if** rand() < ρ **then return** random configuration

2 $b = \arg\max\{D_b : |D_b| \geq N_{min} + 2\}$

3 **if** $b = \emptyset$ **then return** random configuration

4 fit KDEs according to Eqs. (2) and (3)

5 draw N_s samples according to *l'(x)* (see next)

6 **return** sample with highest ratio *l(x) / g(x)*

Tahapan pada algoritma ini dimulai dengan memilih konfigurasi secara acak untuk memastikan eksplorasi ruang parameter, dengan probabilitas ρ. Jika konfigurasi tidak dipilih secara acak, proses dilanjutkan ke langkah berikutnya, yaitu memanfaatkan informasi dari observasi sebelumnya. Selanjutnya, menentukan *budget* b yang memiliki jumlah observasi $|D_b|$ terbesar dan memenuhi kriteria $|D_b| \geq N_{min} + 2$. Apabila tidak ada *budget* b yang memenuhi kriteria, algoritma kembali ke langkah awal untuk memilih konfigurasi secara acak. Distribusi probabilitas dari konfigurasi terbaik kemudian diperkirakan menggunakan KDE (*Kernel Density Estimation*). Selanjutnya, diambil N_s sampel dari distribusi berbobot *l'(x)*, yang proporsional terhadap rasio *l(x)/g(x)*. Dari N_s

sampel yang telah diambil, dipilih konfigurasi dengan nilai rasio *l(x)/g(x)* tertinggi. Konfigurasi tersebut kemudian ditetapkan sebagai konfigurasi berikutnya untuk dievaluasi.

2.1.7 Optuna

Optuna adalah *framework* optimasi *hyperparameter* yang dirancang untuk mengotomatiskan proses pencarian *hyperparameter* dalam machine learning atau deep learning. Optuna menggunakan prinsip *define-by-run* yang inovatif, yang memungkinkan pengguna untuk membangun ruang pencarian secara dinamis selama proses pengoptimalan. Fleksibilitas ini memungkinkan alur kerja yang lebih kompleks tanpa perlu mendefinisikan semua parameter sebelumnya. Selain itu, Optuna menyertakan algoritma yang efisien untuk pengambilan sampel *hyperparameter* dan uji coba pemangkasan yang tidak mungkin memberikan hasil yang baik, sehingga membantu mempercepat proses pengoptimalan dengan membuang konfigurasi yang tidak menjanjikan lebih awal (Akiba dkk., 2019).

2.2 State-of-the-Art

State of the art yang disajikan pada Tabel 2.1 berisi tentang hasil penelitian yang sudah dilakukan sebelumnya, berkaitan dengan penerapan judul penelitian yang dilakukan yaitu "Analisis Komparasi Teknik Optimasi Hyperparameter Pada IndoBERT untuk Deteksi Sarkasme Berbahasa Indonesia".

Tabel 2.1 *State-of-the-Art* Penelitian

No	Nama Peneliti/Journal		Hasil Penelitian
1	**Peneliti (tahun)**	Mohan dkk., (2023)	Model BERT-GCN, mengungguli model GCN, Sequential dan LSTM lainnya dan menunjukkan *accuracy* 90,7%, *F1-score* sebesar 89,6% pada dataset Headline dan *accuracy* sebesar 88,3% dengan *F1-score* sebesar 87,7% pada dataset Riloff. Dataset News Headlines berjumlah 3086 sarkastik dan 2914 non-sarkastik. Sedangkan dataset Riloff berjumlah 317 sarkastik dan 1164 non-sarkastik.
	Judul	Sarcasm Detection Using Bidirectional Encoder Representations from Transformers and Graph Convolutional Networks	
	Algoritma/ Metode	GCN, LSTM, Sequential, BERT-GCN	
2	**Peneliti**	Bagate dkk., (2021)	Teknik stacking dengan mengkombinasikan model Regresi Linear dan model LSTM ke XGBoost berhasil mencapai *accuracy* 73%. Metode ini mampu mengenali kalimat sarkasme dan non-sarkasme tanpa adanya hashtag atau konteks lainnya. Pada penelitian deteksi sarkasme ini dataset yang digunakan merupakan dataset dari twitter yang tersedia di kaggle dengan jumlah yang balance untuk kedua kelas sarkasme dan non-sarkasme, yaitu 24452 dan 26736.
	Judul	Sarcasm Detection Of Tweets Without #Sarcasm: Data Science Approach	
	Algoritma/ Metode	LR, LSTM, XGBoost	
3	**Peneliti**	Parkar dkk., (2023)	

No	Nama Peneliti/Journal		Hasil Penelitian
	Judul	Analytical Comparison On Detection Of Sarcasm Using Machine Learning And Deep Learning Techniques	Model BERT memberikan kinerja terbaik di antara model deep learning dan model machine learning dengan nilai *accuracy* 92,73% dan *F1-score* 93% pada dataset News Headlines dan nilai *accuracy* 75% dan *F1-score* 74% pada dataset Reddit. Pada penelitian deteksi sarkasme ini dataset yang digunakan merupakan dataset dari kaggle . Pertama, dataset News Headlines yang terdiri dari 28619 data. Sedangkan dataset kedua yaitu dataset review reddit dengan jumlah data 8000.
	Algoritma/ Metode	NB, SGD, KNN, LR, DT, RF, SVM, GB, EM, LSTM, LSTM + RNN, RNN, CNN, GloVe, Bi-LSTM, BERT	
4	**Peneliti**	Hao dkk., (2023)	Metodologi yang diusulkan mencapai nilai 0,69, 0,70, dan 0,83 dalam hal *accuracy* pada set data Main balanced, Pol balanced dan set data Pol tidak seimbang secara berurutan. Pada penelitian deteksi sarkasme ini dataset yang digunakan merupakan datset SARC dengan 3 variant datasetnya, yaitu: Main balanced, Pol Balanced, Pol imbalanced.
	Judul	Enhanced Semantic Representation Learning for Sarcasm Detection by Integrating Context-Aware Attention and Fusion Network	
	Algoritma/ Metode	Bi-LSTM Encoder + Context-Aware Attention + Fusion Network atau CSDM	
5	**Peneliti**	Hilmawan (2022)	Bidirectional LSTM mendapatkan *accuracy* validasi sebesar 82,55% dan *F1-score* sebesar 80,92% dan algoritma LSTM mendapatkan *accuracy* validasi sebesar 81,90% 1dan *F1-score* sebesar 80,47%. Pada penelitian deteksi sarkasme ini dataset yang digunakan merupakan dataset News Headlines.
	Judul	Deteksi Sarkasme Pada Judul Berita Berbahasa Inggris Menggunakan Algoritma Bidirectional LSTM	
	Algoritma/ Metode	LSTM, BiLSTM	
6	**Peneliti**	Gole dkk., (2023)	

No	Nama Peneliti/Journal		Hasil Penelitian
	Judul	On Sarcasm Detection with OpenAI GPT-based Models	Dalam kasus fine-tuning, model GPT-3 mencapai *accuracy* dan F1-score sebesar 0,81 mengungguli model-model sebelumnya. Dalam kasus zero-shot, salah satu model GPT-4 menghasilkan *accuracy* 0,70 dan *F1-score* 0,75. Model lainnya memiliki skor yang lebih rendah. Penelitian deteksi sarksame ini dilakukan menggunakan dataset SARC 2.0 versi political dan balanced (pol-bal)
	Algoritma/ Metode	GPT-3, InstructGPT, GPT-3.5, GPT-4	
7	**Peneliti**	Misra dkk., (2023)	Arsitektur model Hybrid Neural Network dengan menggunakan attention mechanism yang cukup kompleks pada Dataset News Headlines mencapai kinerja yang sangat baik, mencapai *accuracy* hampir 89%. Penelitian deteksi sarkasme ini dilakukan pada dataset berkualitas tinggi dan berskala besar yang dikumpulkan dari website berita sarkastik dan website berita nyata.
	Judul	Sarcasm Detection Using News Headlines Dataset	
	Algoritma/ Metode	CNN + BiLSTM + Attention Mechanism + MLP	
8	**Peneliti**	Bagate dkk., (2022)	Setelah melatih dan menguji model SVM-RNN untuk beberapa iterasi, model ini memberikan *F1-score* pengujian yang sangat baik sebesar 75.75% dan *accuracy* 80% untuk model yang dijelaskan. untuk deteksi sarkasme dengan weighted average. Penelitian deteksi sarkasme ini dilakukan pada dataset SARC yang terdiri dari 533 juta komentar ditulis dalam bahasa Inggris, dimana 1.3 juta komentar merupakan sarkastik.
	Judul	Sarcasm Detection on Text for Political Domain — An Explainable Approach	
	Algoritma/ Metode	SVM-RNN + weighted average	
9	**Peneliti**	Sharm dkk., (2022)	Model yang diusulkan diverifikasi pada tiga dataset berbeda media sosial dunia nyata , Self-Annotated Reddit Corpus (SARC), dataset News Headlines, dan Dataset Twitter. *Accuracy* yang dicapai adalah 83%, 92%, 90,8% dan 92,80%. Nilai metrik *accuracy* lebih baik daripada state-of-the-art sebelumnya. Penelitian deteksi sarkasme ini dilakukan pada 3 dataset nyata sosial media, yaitu : SARC, headlines dataset, twitter dataset.
	Judul	Sarcasm Detection Over Social Media Platforms Using Hybrid Auto-Encoder-Based Model	
	Algoritma/ Metode	LSTM autoencoder + USE + BERT	
10	**Peneliti**	Jayaraman dkk., (2022)	

No	Nama Peneliti/Journal		Hasil Penelitian
	Judul	Sarcasm Detection in News Headlines using Supervised Learning	Model RoBERTa mencapai hasil yang sebanding dengan context-independent features. Secara khusus, model ini mencapai 93,26% dan 93,11% pada dataset versi 1 untuk validasi dan pengujian. Demikian pula, model ini mencapai 93,36% dan 93,82% dalam dataset versi 2. Penelitian deteksi sarksame ini dilakukan pada 2 versi news headlines dataset. dataset versi 1 terdiri dari 26709 total news headlines yang berisi 11724 sarkastik dan 14985 non-sarkastik. Sementara dataset versi 2 terdiri dari 28619 total news headlines yang berisi 13634 sarkastik dan 14985 non-sarkastik.
	Algoritma/ Metode	NB-SVM, LR, BiGRU, BERT, DistilBERT, RoBERTa, XLNet	
11	**Peneliti**	Fitrianto dkk., (2024)	IndoBERT yang diadaptasi secara khusus mencapai *F1-score* yang tinggi sebesar 95% pada dataset yang kami peroleh dari penelitian sebelumnya yang berisi berbagai macam tweet. Penelitian deteksi sarkasme ini menggunakan dataset twitter yang terdiri dari 8700 tweet dengan jumlah yang seimbang untuk kelas tweet sarkasme dan non-sarkasme.
	Judul	Classification of Indonesian Sarcasm Tweets on X Platform Using Deep Learning	
	Algoritma/ Metode	IndoBERT, RoBERTa, BERT base, BERT Multilingual	
12	**Peneliti**	Sharma dkk., (2023)	Model ensemble dengan mengkombinasikan word embedding dari GloVe, Word2Vec dan BERT kemudian menjadi input untuk Fuzzy Logic ini memilki *accuracy* masing-masing sebesar 90,81%, 85,38%, dan 86,80%. Pada dataset Headlines, "Self-Annotated Reddit Corpus" (SARC) dan dataset aplikasi Twitter secara berurutan.
	Judul	Sarcasm Detection over Social Media Platforms Using Hybrid Ensemble Model with Fuzzy Logic	
	Algoritma/ Metode	(Word2Vec, GloVe, BERT) + Fuzzy Logic	
13	**Peneliti**	Tan dkk., (2023)	Hybrid model Bi-LSTM daan Multi-Task Learning yang digunakan pada penelitian ini mengungguli metode yang sudah ada yang ada dengan selisih 3%, dengan *F1-score* sebesar 94%. Penelitian deteksi sarkasme ini menggunakan dataset dari Reddit, Twitter US airline dan Twitter.
	Judul	Sentiment Analysis and Sarcasm Detection using Deep Multi-Task Learning	

No	Nama Peneliti/Journal		Hasil Penelitian
	Algoritma/ Metode	Bi-LSTM + Multi-Task Learning	
14	**Peneliti**	Farhan dkk., (2023)	Penggunaan model BiLSTM dan word embedding GloVe yang digunakan pada penelitian in telah mencapai *accuracy* 86,35%, dengan recall 88%, Precision, dan F1-score. Penelitian deteksi sarkasme ini menggunakan dataset Twitter (Gosh and Veale), News Headlines, Sarcasm Corpus V2dataset, Combined Dataset (multi-domain).
	Judul	Automatic Sarcasm Detection on Cross-Platform Social Media Datasets: A GLoVe and Bi-LSTM Based Approach	
	Algoritma/ Metode	GloVe + Bi-LSTM	
15	**Peneliti**	Eke dkk., (2021)	Evaluasi teknik ensemble ini pada dua dataset benchmark Twitter mencapai presisi tertinggi 98,5% dan 98,0%, dan dataset IAC-v2 mencapai presisi tertinggi 81,2%. Mengungguli baseline model BiLSTM dan BERT.
	Judul	Context-Based Feature Technique for Sarcasm Identification in Benchmark Datasets Using Deep Learning and BERT Model	
	Algoritma/ Metode	GloVe + BiLSTM, BERT + Feature Fusion	
16	**Peneliti**	Sandor dkk., (2024)	BERT-based model, memiliki kinerja lebih baik dibandingkan dengan model machine learning dan deep learning dengan mencapai *accuracy* 73,1% *F1-score* 72,4% recall 71.3% precision 72,2%. Penelitian deteksi sarkasme ini menggunakan dataset 1,3 juta komentar sosial media, terdiri dari kelas sarkastik dan non-sarkastik
	Judul	Sarcasm Detection In Online Comments Using Machine Learning	
	Algoritma/ Metode	LR, RR, SVM, BiLSTM, BERT	
17	**Peneliti**	Siddiqui dkk., (2021)	SVM memiliki kinerja lebih baik dibandingkan dengan klasifikasi machine

No	Nama Peneliti/Journal		Hasil Penelitian
	Judul	Sarcasm Detection from Social Media Posts using Machine-learning Techniques: A Comparative Analysis	learning konvensional dengan mencapai F1-score sebesar 84%. Penelitian deteksi sarkasme ini menggunakan dataset hasil scrapping dari twitter dengan total 13,882 tweet. Terdiri dari kelas sarkastik 6,382 kelas sementara kelas non-sarkastik 7,500 tweet
	Algoritma/ Metode	SVM, RF, DT, KNN, LR, GB, NB	
18	**Peneliti**	Rosid dkk., (2022)	Kinerja model MHA-CovBi yang diusulkan dievaluasi melalui analisis komparatif terhadap pendekatan yang ada. Model ini berhasil mengungguli state of the art terkini dengan mencapai *accuracy* 94,60% dan *F1-score* 94,38%. Pada penelitian deteksi sarkasme ini dataset yang digunakan merupakan dataset dari scrapping pada platform twitter dengan total jumlah 5000 dataset pada dataset DS1 yang terdiri dari 2450 tweet sarkastik dan 2550 tweet non-sarkastik berbahasa Indonesia-Inggris. Sedangkan dataset DS2 memiliki total 28.619 tweet yang terdiri dari 13.635 tweet sarkastik dan 14984 tweet non-sarkastik berbahasa Inggris.
	Judul	Sarcasm Detection in Indonesian-English CodeMixed Text Using Multihead Attention-Based Convolutional and Bi-Directional GRU	
	Algoritma/ Metode	CNN+BiGRU	
19	**Peneliti**	Suhartono dkk., (2024)	Ditemukan bahwa model bahasa fine-tuning pre-train masih lebih unggul daripada teknik lain, mencapai *F1-score* 62,74% dan 76,92% di Reddit dan Twitter secara berurutan. Ditemukan juga bahwa LLM terbaru gagal melakukan klasifikasi zero-shot untuk deteksi sarkasme dan bahwa mengatasi ketidakseimbangan data membutuhkan pendekatan augmentasi data yang lebih canggih daripada metode dasar kami. Pada penelitian deteksi sarkasme ini menggunakan dataset hasil scrapping dari reddit dan twitter.
	Judul	IdSarcasm: Benchmarking and Evaluating Language Models for Indonesian Sarcasm Detection	
	Algoritma/ Metode	LR, NB, SVM, BERT, XLM, BLOOMZ, mT0	
20	**Peneliti**	Pawestri dkk., (2024)	secara keseluruhan tertinggi dapat dicapai pada fold 6, 0.897 untuk *accuracy* validasi, 0.883 untuk *F1-score* dan presisi, dan 0.882 untuk recall pada model CNN + RoBERTa. Sebaliknya, RNN menunjukkan *accuracy* 0,711, presisi
	Judul	Sarcasm Detection A Comparative Analysis of	

No	Nama Peneliti/Journal		Hasil Penelitian
		RoBERTa-CNN vs RoBERTa-RNN Architectures	0,692, recall sebesar 0.620, *F1-score* 0.654, dan *loss* sebesar 0.564, dengan waktu pemrosesan yang lebih lama waktu 116500 milidetik per epoch. Pada penelitian deteksi sarkasme ini meggunakan dataset News Headlines.
	Algoritma/ Metode	RoBERTa-CNN, RoBERTa-RNN	
21	**Peneliti**	Rahman dkk., (2024)	Model IndoBERTweet yang telah disempurnakan mencapai *accuracy* sebesar 89,00%, dengan demikian mengungguli model-model seperti IndoGPT, IndoBERT, dan LSTM, menegaskan posisinya sebagai yang yang paling canggih. Dengan demikian, dapat disimpulkan bahwa IndoBERTweet menyajikan pendekatan yang menjanjikan untuk klasifikasi sarkasme yang menjanjikan untuk klasifikasi sarkasme dalam korpus data Twitter berbahasa Indonesia. Pada penelitian deteksi sarkasme ini menggunakan dataset Twitter berbahasa Indonesia.
	Judul	IndoBERTweet for Sarcasm Evaluating Domain-Adapted Transformers for Indonesian Twitter Sarcasm Classification	
22	**Peneliti**	Muhaddisi dkk., (2021)	Penelitian ini menghasilkan nilai akurasi pada analisis sentimen tanpa deteksi sarkasme dengan Naive Bayes sebesar 61%, dengan metode Random Forest sebesar 72%. Hasil akurasi pada analisis sentimen dengan deteksi sarkasme menggunakan metode Naïve Bayes – Random Forest sebesar 60% dan metode Random Forest – Random Forest sebesar 71%.
	Judul	Sentiment Analysis With Sarcasm Detection On Politician's Instagram	
23	**Peneliti**	Rahaman dkk., (2021)	Penelitian ini mengungkapkan kumpulan fitur sarkastik dengan model supervised machine learning yang efektif, menunjukan akurasi yang lebih baik. Hasil penelitian menunjukkan bahwa Decision Tree (91,84%) dan Random Forest (91,90%) mengungguli dalam hal akurasi dibandingkan dengan algoritma supervised machine learning lainnya untuk pemilihan fitur yang tepat. Penelitian ini telah menyoroti model supervised machine learning yang sesuai bersama dengan kumpulan fitur yang sesuai untuk mendeteksi
	Judul	Sarcasm Detection in Tweets A Feature-based Approach using Supervised Machine Learning Models	

No	Nama Peneliti/Journal		Hasil Penelitian
			sarkasme dalam tweet.
24	**Peneliti**	Kusumastuti dkk., (2022)	Hasil penelitian pendeteksian kalimat sarkasme menggunakan SentiStrength memperoleh nilai akurasi sebesar 54,52%. Selain itu, penelitian ini berhasil menemukan kelemahan SentiStrength, yaitu kurangnya pembobotan singkatan negatif pada setiap kamus, kurangnya pembobotan kata kata yang kuat, dan bahasa gaul.
	Judul	Detection of Sarcasm Sentences in Indonesian Tweets using SentiStrength	
25	**Peneliti**	Pernanda dkk., (2021)	Hasil eksperimen menunjukkan bahwa metode LSTM dengan Word2Vec mampu mencapai akurasi 82,13% dan f1-score sebesar 61,31% mengungguli TF-IDF konvensional dengan naïve Bayes pendeteksi sarkas.
	Judul	Sarcasm Detection of Tweets in Indonesian Language Using Long Short-Term Memory	
26	**Peneliti**	Arlim dkk., (2022)	Hasil eksperimen menunjukkan bahwa model dengan fitur hiperbola sebagai fitur untuk SVM, RF dan RF+ Bagging mengklasifikasikan lebih banyak tweet dalam data pengujian sebagai sarkasme daripada yang tanpa hiperbola.
	Judul	Sarcasm Detection in Indonesian Tweets Using Hyperbole Features	
27	**Peneliti**	Rosid dkk., (2022)	Hasil eksperimen menunjukkan bahwa kombinasi word embedding fastText dan BiGRU sebagai pengklasifikasi menghasilkan kinerja terbaik, dengan akurasi 93.85%.
	Judul	Pre-Trained Word Embeddings for Sarcasm Detection in Indonesian Tweets A Comparative Study	
28	**Peneliti**	Fu dkk., (2024)	Hasil eksperimen kami menunjukkan peningkatan yang signifikan dalam

No	Nama Peneliti/Journal		Hasil Penelitian
	Judul	Multi-Modal Sarcasm Detection with Sentiment Word Embedding	akurasi deteksi sarkasme dengan menyematkan kata-kata emosional ke dalam vektor fitur. Penilaian dilakukan secara ekstensif pada dataset benchmark publik, yang menunjukkan bahwa metode yang diusulkan, yaitu Desnet + BERT + VIT + BiLSTM melampaui metode dasar yang ada saat ini.
29	**Peneliti**	Rosid dkk., (2023)	Algoritma Support Vector Machine (SVM) telah mencapai akurasi yang baik. Model ini dapat digunakan untuk mendeteksi keberadaan sarkasme dalam teks dengan tingkat keberhasilan 80% pada dimensi 20.
	Judul	Ensemble Machine Learning to Detect Sarcasm in English on Twitter Social Media	
30	**Peneliti**	Azka Fauzi Al-Parisi (2024)	Hasil eksperimen membuktikan performa model IndoSarcasm memiliki akurasi sebesar 84.77% dengan nilai parameter lain yang relatif seimbang yaitu precision sebesar 84.58%, recall sebesar 84.97% dan F1-Score sebesar 84.77%. Hal ini diperkuat dengan hasil eksperimen selama pelatihan, yaitu tidak ditemukannya indikasi overfitting.
	Judul	Deteksi Sarkasme Menggunakan Bidirectional Encoder Representations from Transformers (BERT) Pada Teks Bahasa Indonesia	

Berdasarkan kajian terhadap penelitian sebelumnya diketahui bahwa model *deep learning* memiliki kinerja yang unggul dalam tugas deteksi sarkasme. Selain itu, penelitian deteksi sarkasme pada bahasa yang lebih spesifik juga masih menjadi tantangan. Penelitian deteksi sarkasme berbahasa Indonesia sebelumnya telah dilakukan oleh beberapa peneliti, tetapi penelitiannya masih terbatas sehingga

memberikan ruang untuk pengembangan lebih lanjut . Oleh karena itu, dilakukan penelitian dengan judul “Analisis Komparasi Teknik Optimasi Hyperparameter Pada IndoBERT untuk Deteksi Sarkasme Berbahasa Indonesia”.

2.3 Matriks Penelitian

Tabel 2.2 merupakan matriks penelitian yang berisi perbandingan berbagai penelitian terkait deteksi sarkasme menggunakan teknik *machine learning* dan *deep learning*. Kolom-kolom yang disertakan memberikan gambaran tentang metode yang diterapkan.

Tabel 2. 2 Matriks penelitian

No	Peneliti/ Tahun	Judul	Ruang Lingkup																			
			Model *Machine Learning*													Representasi Teks				Dataset		
			Logistic Regression	***Naive Bayes***	***Support Vector Machine***	***K-Nearest Neighbors***	***Decision Tree***	***Random Forest***	***Passive Aggressive Classifier***	***Expectation Maximization***	***Gradient Boosting***	***CNN***	***RNN***	***LSTM***	***BiGru***	***BERT***	***XLNet***	***FaxtText***	***Word2Vec***	***GloVe***	***TF-IDF***	***Twitter***	***Reddit***	***News Headlines***
1	(Mohan dkk., 2023)	Sarcasm Detection Using Bidirectional Encoder Representations from												✓		✓							✓	✓

		Transformers and Graph Convolutional Networks																					
2	(Bagate, 2021)	Sarcasm Detection of Tweets Without #Sarcasm: Data Science Approach	✓								✓			✓								✓		
3	(Parkar & Bhalla, 2024)	Analytical Comparison On Detection Of Sarcasm Using Machine Learning And Deep Learning Techniques	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓		✓							✓	✓
4	(Hao dkk., 2023)	Enhanced Semantic Representation Learning for Sarcasm Detection by Integrating Context-Aware Attention and Fusion Network												✓				✓					✓	
5	(Hilmawan, 2022)	Deteksi Sarkasme Pada Judul Berita Berbahasa Inggris Menggunakan Algoritma Bidirectional LSTM												✓						✓				✓
6	(Gole dkk., 2023)	On Sarcasm Detection with OpenAI GPT-based Models																					✓	
7	(Misra, 2023)	Sarcasm Detection Using News Headlines Dataset										✓		✓					✓					✓

8	(Bagate & Suguna, 2022)	Sarcasm Detection on Text for Political Domain — An Explainable Approach			✓								✓									✓	
9	(Sharma, 2022)	Sarcasm Detection Over Social Media Platforms Using Hybrid Auto-Encoder-Based Model												✓		✓						✓	✓	✓
10	(Jayaraman dkk., 2022)	Sarcasm Detection in News Headlines using Supervised Learning	✓	✓	✓										✓	✓			✓	✓				✓
11	(Fitrianto dkk., 2024)	Classification of Indonesian Sarcasm Tweets on X Platform Using Deep Learning														✓						✓		
12	(Sharma dkk., 2023)	Sarcasm Detection over Social Media Platforms Using Hybrid Ensemble Model with Fuzzy Logic														✓			✓	✓	✓	✓	✓	✓
13	(Tan dkk., 2023)	Sentiment Analysis and Sarcasm Detection using Deep Multi-Task Learning										✓		✓					✓			✓	✓	✓
14	(Farhan dkk., 2024)	Automatic Sarcasm Detection on Cross-Platform Social Media Datasets: A GLoVe and												✓						✓		✓		✓

		Bi-LSTM Based Approach																					
15	(Eke dkk., 2021)	Context-Based Feature Technique for Sarcasm Identification in Benchmark Datasets Using Deep Learning and BERT Model												✓		✓				✓		✓		
16	(Šandor & Bagić Babac, 2024)	Sarcasm Detection in Online Comments Using Machine Learning	✓		✓									✓		✓							✓	
17	(Siddiqui dkk., 2021)	Sarcasm Detection from Social Media Posts using Machine-learning Techniques: A Comparative Analysis	✓	✓	✓	✓	✓	✓			✓											✓		
18	(Rosid, 2022)	Sarcasm Detection in Indonesian-English CodeMixed Text Using Multihead Attention-Based Convolutional and Bi-Directional GRU										✓			✓				✓		✓			
19	(Suhartono dkk., 2024)	IdSarcasm: Benchmarking and Evaluating Language	✓	✓	✓											✓				✓		✓	✓	

		Models for Indonesian Sarcasm Detection																					
20	(Pawestri & Auzan, 2024)	Sarcasm Detection a Comparative Analysis of RoBERTa-CNN vs RoBERTa-RNN Architectures									✓		✓		✓								✓
21	(Fitrahtur Rahman & Abba Suganda Girsang, 2024)	IndoBERTweet for Sarcasm Evaluating Domain-Adapted Transformers for Indonesian Twitter Sarcasm Classification													✓						✓		
22	(Muhaddisi dkk., 2021)	Sentiment Analysis With Sarcasm Detection On Politician's Instagram		✓				✓															
23	(Rahaman dkk., 2021)	Sarcasm Detection in Tweets A Feature-based Approach using Supervised Machine Learning Models	✓		✓		✓	✓													✓		
24	(Kusumastuti dkk., 2022)	Detection of Sarcasm Sentences in Indonesian Tweets using SentiStrength																			✓		
25	(Pernanda dkk., 2021)	Sarcasm Detection of Tweets in Indonesian		✓									✓					✓		✓			

		Language Using Long Short-Term Memory																					
26	(Arlim dkk., 2022)	Sarcasm Detection in Indonesian Tweets Using Hyperbole Features						✓													✓		
27	(Rosid dkk., 2022)	Pre-Trained Word Embeddings for Sarcasm Detection in Indonesian Tweets A Comparative Study										✓		✓	✓	✓		✓			✓		
28	(Fu dkk., 2024)	Multi-Modal Sarcasm Detection with Sentiment Word Embedding												✓		✓							
29	(Rosid dkk., 2023)	Ensemble Machine Learning to Detect Sarcasm in English on Twitter Social Media	✓	✓	✓		✓														✓		
30	(Azka Fauzi Al-Parisi, 2024)	Deteksi Sarkasme Menggunakan Bidirectional Encoder Representations from Transformers (BERT) Pada Teks Bahasa Indonesia														✓					✓		
31	**(Ajeng Alya**	**Analisis Komparasi Teknik Optimasi *Hyperparameter* Pada**														✓					✓		

	Kartika Sari, 2025)	**IndoBERT untuk Deteksi Sarkasme Berbahasa Indonesia**																					

Di samping penggunaan *pre-trained* model seperti IndoBERT yang telah terbukti efektif dalam mendeteksi sarkasme berbahasa Indonesia, penerapan teknik optimasi *hyperparameter* secara otommatis dan menganalisis pengaruhnya terhadap peningkatan kinerja model belum banyak diterapkan. Penerapan teknik optimasi tersebut berpotensi meningkatkan kinerja model IndoBERT secara keseluruhan (Anugerah Simanjuntak dkk., 2024), sehingga mendukung pengembangan sistem yang lebih andal dan efektif dalam mengidentifikasi kalimat sarkasme dalam bahasa Indonesia.