

BAB II

TINJAUAN PUSTAKA

2.1 Kajian Penelitian Terdahulu

Beberapa penelitian berikut digunakan sebagai acuan dalam penyusunan tugas akhir ini.

Tabel 1. Penelitian Terdahulu

No.	Peneliti	Judul	Tahun	Metode
1.	Aryan Rastogi1, Qian Liu2, Erik Cambria2	Stress Detection from Social Media Articles : New Dataset Benchmark and Analytical Study	2022	BERT, RoBERTa, dan DistilBERT
2.	Junanda Patihullah, Edi Winarko	Hate Speech Detection for Indonesia Tweets Using Word Embedding And Gated Recurrent Unit	2019	Word Embedding dan Gated Recurrent Unit
3.	Clinton Imanuel Kawatu	THE APPLICATION OF GATED RECURRENT UNIT IN SOCIAL MEDIA SENTIMENT ANALYSIS	2019	Gated Recurrent Unit

Pada Tabel 1. Menunjukkan penelitian sebelumnya yang memiliki keterkaitan dengan penelitian yang sedang dilakukan oleh peneliti yang berjudul “Deteksi Stres pada Sosial Media Reddit Menggunakan Metode Gated Recurrent Unit (GRU)”. Berikut adalah penjelasan dari tabel tersebut.

1. Penelitian pertama yang dilakukan pada tahun 2022, memiliki kelebihan dengan menyediakan empat dataset berkualitas tinggi dari Twitter dan Reddit, yang dirancang khusus untuk deteksi stres, sekaligus memberikan deskripsi lengkap

tentang proses pengumpulan, praproses, dan anotasi data. Penelitian ini berhasil mengidentifikasi kelemahan pendekatan polaritas sentimen yang tidak cukup efektif untuk deteksi stres, sehingga mendorong pengembangan metode yang lebih spesifik. Namun, penelitian ini memiliki kekurangan, seperti keterbatasan adaptasi metode yang ada untuk tugas ini, kebutuhan perbaikan lebih lanjut dalam kualitas dataset dan model, serta ketergantungan pada dataset tertentu yang mungkin tidak mencakup berbagai situasi dalam deteksi stres [5].

2. Penelitian kedua yang dilakukan pada tahun 2019, bertujuan untuk mendeteksi ujaran kebencian secara otomatis menggunakan metode deep learning, yaitu Gated Recurrent Unit (GRU). Kelebihan penelitian ini adalah kemampuannya menggunakan metode GRU untuk mendeteksi ujaran kebencian di media sosial dengan akurasi mencapai 92,96%. Namun, terdapat kekurangan pada fitur word2vec terhadap jumlah data latih yang besar agar dapat merepresentasikan kata-kata secara akurat. Dengan kata lain, semakin banyak data latih yang tersedia, semakin baik performa word2vec, sehingga pendekatan ini menjadi kurang optimal ketika data latih terbatas [6].
3. Penelitian ketiga yang dilakukan pada tahun 2019, memiliki keunggulan dalam meningkatkan efisiensi pengolahan data dengan mengotomatisasi analisis sentimen menggunakan algoritma Gated Recurrent Unit (GRU), yang efektif menangkap konteks kalimat panjang dan kompleks, sehingga hasilnya lebih akurat. Namun, penelitian ini juga memiliki kekurangan, di mana pengembangan lebih lanjut masih diperlukan, khususnya dalam optimalisasi penggunaan algoritma GRU, seperti penyesuaian jumlah unit yang digunakan, serta proses pembersihan dataset yang lebih menyeluruh. Hal ini penting untuk memastikan hasil akurasi yang lebih tinggi dan mengurangi risiko overfitting [8].

2.2 Deteksi Stres

Deteksi stres adalah bagian khusus dari klasifikasi emosi dimana sistem cerdas dirancang untuk mengidentifikasi stres atau depresi pada seseorang. Stres sendiri didefinisikan sebagai kondisi ketika ada ketidakseimbangan antara kebutuhan internal seseorang dan kemampuannya untuk memenuhinya, dan hal ini dianggap sebagai masalah medis yang serius. Selama pandemi Covid-19, berbagai survei menunjukkan bahwa stres dialami oleh orang-orang dari berbagai latar belakang. Dengan demikian, deteksi stres menjadi penting untuk menilai kesehatan mental suatu populasi dan

memberikan informasi berharga kepada pemerintah serta perusahaan tentang faktor-faktor yang dapat menyebabkan stres, sehingga mereka dapat membuat kebijakan yang lebih baik dengan memperhatikan kesehatan mental masyarakat. Meskipun banyak penelitian telah dilakukan untuk mendeteksi stres berdasarkan faktor-faktor fisiologis dan biologis, penelitian tentang deteksi stres dari teks tertulis kini menjadi semakin populer.

Orang-orang banyak menghabiskan waktu di media sosial, dan sering membagikan pandangan, pengalaman, dan emosi pribadi mereka dengan pengikut dan teman-teman mereka. Karena itu, data dari media sosial bisa digunakan untuk membangun sistem yang bisa mendeteksi stres menggunakan model pembelajaran mesin. Namun, ada beberapa tantangan dalam membuat dataset yang berkualitas untuk mendeteksi stres. Salah satu masalah utamanya adalah data dari media sosial sering kali bercampur dengan informasi yang tidak relevan, yang bisa mempengaruhi hasil prediksi model secara negatif. Untuk mengatasi masalah ini, penelitian ini berfokus pada cara-cara untuk membuat sistem deteksi stres yang lebih cerdas dari artikel di media sosial. Penelitian ini membangun dataset dari platform media sosial populer yaitu Reddit [5].

2.3 Sosial Media

Kaplan et al. (2010) dalam Anwar (2017) [9] menjelaskan bahwa media sosial adalah kumpulan aplikasi di internet yang menggunakan teknologi Web 2.0, memungkinkan pengguna untuk membuat dan berbagi konten satu sama lain. Menurut teori Uses and Gratification oleh McQuail, ada empat tujuan utama dalam menggunakan media, yaitu mencari informasi, membangun identitas pribadi, berinteraksi dan berintegrasi sosial, serta mencari hiburan [10]. Sebagian besar orang menggunakan platform media sosial untuk mencari informasi. Media sosial adalah alat di internet yang memungkinkan pengguna untuk menampilkan diri, berinteraksi, berkolaborasi, berbagi, berkomunikasi dengan pengguna lain, dan membentuk hubungan sosial virtual [11].

Media sosial semakin mudah digunakan oleh siapa saja untuk mencari, mendapatkan, dan memanfaatkan berbagai informasi tanpa hambatan. Sebagai platform yang fleksibel, media sosial digunakan oleh orang-orang dari berbagai latar belakang sosial, ekonomi, dan politik. Fleksibilitas ini memungkinkan media sosial untuk

memperkuat dan membangun hubungan antara individu atau kelompok di dunia maya, tanpa terpengaruh oleh perbedaan status dalam masyarakat [12].

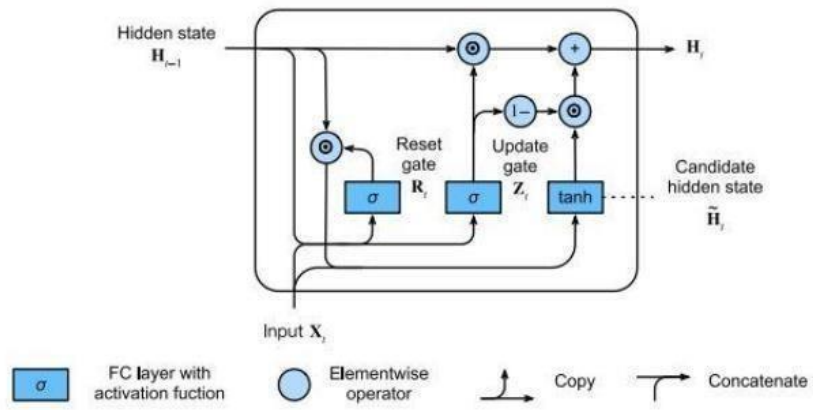
2.4 Reddit

Reddit adalah platform media sosial berbentuk forum yang diluncurkan pada tahun 2005. Pengguna bisa berbagi teks, foto, dan video serta memberikan suara dan komentar pada postingan. Ada komunitas khusus di Reddit yang disebut Subreddit, seperti r/science atau r/memes, yang bisa dibuat oleh siapa saja yang telah mendaftar, asalkan mengikuti aturan Reddit. Pengguna biasanya mengikuti komunitas berdasarkan minat mereka, bukan akun individu. Dengan banyaknya subreddit yang ada, Reddit menjadi tempat yang sangat beragam untuk berbagi dan mendapatkan informasi [9].

Diskusi di Reddit umumnya bisa diakses oleh siapa saja, baik yang punya akun atau tidak, kecuali di subreddit privat. Konten dan komentar yang diposting bisa mendapatkan vote dari pengguna, yang mempengaruhi visibilitasnya. Untuk membuat akun Reddit, pengguna hanya perlu memilih nama pengguna dan kata sandi tanpa verifikasi email, dan harus berusia minimal 13 tahun. Reddit mendorong penggunaan nama samaran untuk melindungi privasi, dan riwayat aktivitas pengguna bisa dilihat oleh publik [13].

2.5 Gated Recurrent Unit (GRU)

GRU adalah sebuah arsitektur yang dikembangkan oleh Kyunghun Cho pada tahun 2014 [7]. Model Gated Recurrent Unit (GRU) merupakan versi yang lebih sederhana dari model Recurrent Neural Network (RNN). GRU memiliki dua gerbang utama, yaitu update gate dan reset gate. Walaupun strukturnya lebih sederhana dibandingkan model lain, tetapi GRU tetap memberikan keakuratan prediksi dengan model deep learning lainnya. Reset gate berfungsi untuk memilih informasi yang perlu dihapus atau diatur ulang, sedangkan update gate berfungsi untuk memilih informasi yang harus dipertahankan. Setelah itu, hidden state dihitung untuk menyimpan informasi penting dari sebelumnya dengan bantuan update gate [14].



Gambar 1. Model GRU

