

2. LANDASAN TEORI

2.1. Tinjauan Pustaka

Berikut berisikan tinjauan pustaka yang digunakan dalam pelaksanaan penelitian ini

2.1.1. Gangguan Mental dalam Media Sosial

Media sosial dapat berperan sebagai sumber informasi dan dukungan yang penting untuk menyelidiki kesehatan mental pengguna berdasarkan tulisan mereka (Freeman, 2022). Namun, media sosial juga dapat memperburuk kondisi kesehatan mental pengguna, dan seringkali pengguna media sosial yang menderita gangguan mental tidak dapat mengambil langkah-langkah yang diperlukan untuk pengobatan mereka. Gangguan-gangguan ini mencakup depresi klinis, kecemasan, gangguan bipolar, gangguan makan, gangguan stres pasca-trauma, dan gangguan obsesif-kompulsif (Haque et al., 2020; Garg et al., 2022; Zhang et al., 2020; Chancellor & Choudhury, 2020). Dengan beragam kondisi yang termasuk di dalamnya seperti depresi klinis, kecemasan, gangguan bipolar, hingga gangguan stres pasca-trauma dan obsesif-kompulsif, gangguan mental memiliki dampak yang signifikan terhadap individu dan masyarakat secara keseluruhan (Garg et al., 2022; Zhang et al., 2020).

Proses klasifikasi gangguan mental merupakan langkah penting dalam pengelolaan dan penanganan kondisi tersebut. Klasifikasi ini bertujuan untuk mengkategorikan gangguan mental ke dalam kelompok atau kategori tertentu berdasarkan gejala, penyebab, atau karakteristik lainnya (Freeman, 2022). Studi terbaru telah mengidentifikasi berbagai metode klasifikasi yang digunakan dalam literatur, mulai dari pendekatan konvensional hingga pendekatan berbasis teknologi seperti machine learning dan deep learning (Chancellor & Choudhury, 2020; Zhang et al., 2020).

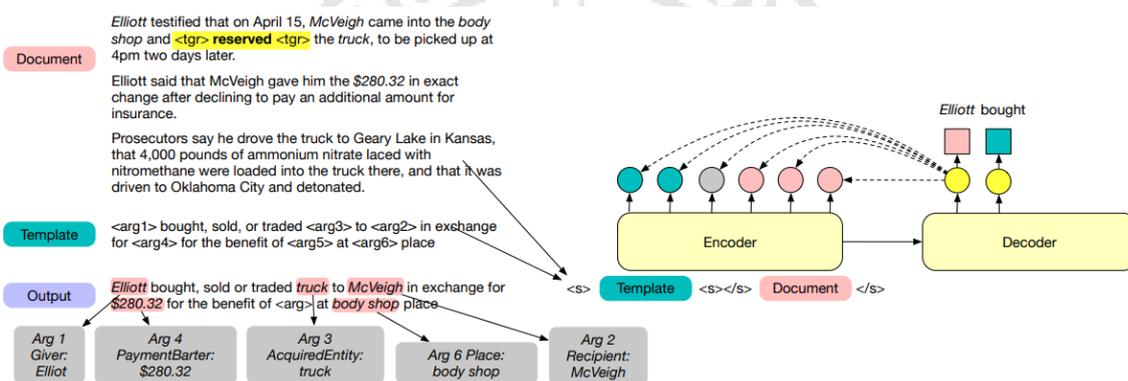
Meskipun pentingnya klasifikasi gangguan mental diakui, tantangan besar muncul ketika melakukan klasifikasi di media sosial. Data teks dari media sosial sering kali mencakup ekspresi yang beragam dari individu, seperti bahasa informal, slang, atau emoji, yang dapat memperumit proses klasifikasi (Zhang et al., 2020). Tantangan ini menyoroti perlunya pendekatan baru dan lebih canggih dalam mengatasi kompleksitas data teks dari media sosial dalam konteks klasifikasi gangguan mental (Chancellor & Choudhury, 2020).

Pendekatan-pendekatan terkini dalam klasifikasi gangguan mental mencakup penggunaan teknologi-teknologi baru seperti machine learning dan deep learning.

Penelitian-penelitian terbaru menunjukkan bahwa teknologi ini mampu meningkatkan akurasi dan efektivitas dalam mendeteksi gangguan mental dari teks media sosial dengan lebih baik daripada metode-metode konvensional (Zhang et al., 2020). Namun demikian, penelitian lebih lanjut diperlukan untuk mengatasi tantangan kompleksitas data teks media sosial dan meningkatkan validitas prediksi secara keseluruhan (Chancellor & Choudhury, 2020).

2.1.2. Conditional Generation Model

Conditional Generation Model (CGM) adalah model yang dirancang untuk melakukan ekstraksi argumen tingkat teks (Li et al., 2021). Model ini merupakan model jaringan saraf yang menghasilkan templat yang akan terisi dengan argumen berdasarkan konteks tertentu dan templat yang tidak terisi. Dikondisikan pada templat yang tidak terisi dan konteksnya, model menghasilkan templat yang terisi dengan argumen.



Gambar 2.1 Cara kerja model *conditional generation*

Sumber: Li et al. (2021). *Document-Level Event Argument Extraction by Conditional Generation*

Proses ini melibatkan prediksi argumen untuk setiap tempat di template berdasarkan informasi konteks (Li et al., 2021). Tidak seperti beberapa pendekatan tradisional yang mengandalkan pengenalan entitas sebagai langkah *preprocess*, CGM tidak memerlukan langkah-langkah ini. Seperti yang ditunjukkan pada gambar 2.1, model ini secara langsung menghasilkan templat yang diisi dengan argumen. CGM menunjukkan fleksibilitas dalam menangani argumen yang hilang dan banyak argumen serta menunjukkan kinerja yang kuat dalam tugas ekstraksi argumen tingkat teks.

CGM memanfaatkan penggabungan kondisi eksternal ke dalam proses pembuatan teks untuk mempengaruhi hasil yang dihasilkan (Guo et al., 2020). Kondisi ini dapat mencakup konteks, topik, emosi, pengetahuan eksternal, dan lainnya. Tidak seperti metode pembuatan

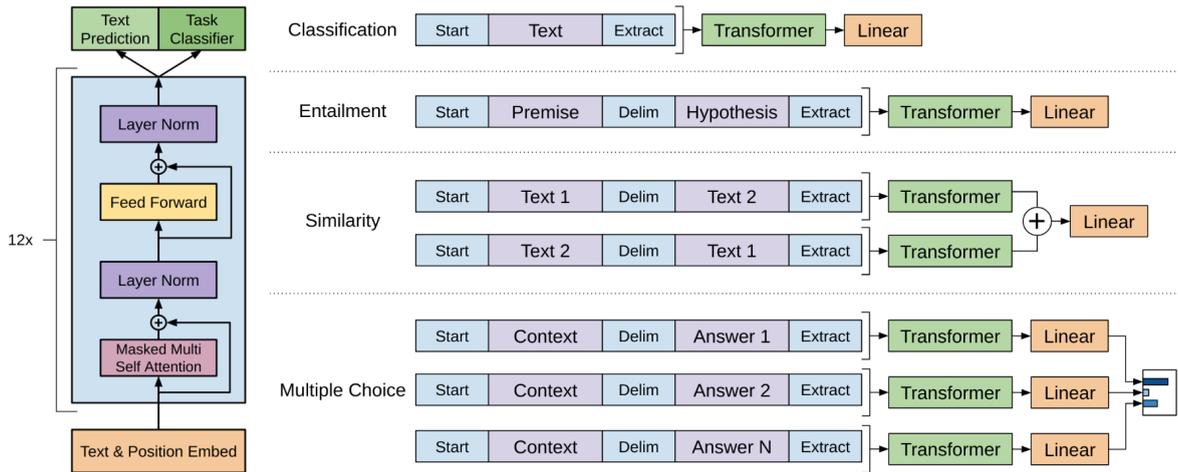
teks tradisional yang hanya berfokus pada konten teks, CGM mempertimbangkan faktor tambahan untuk membuat teks yang dihasilkan lebih beragam dan mirip manusia. Dalam pembuatannya, CGM dapat menggunakan berbagai tipe arsitektur sebagai dasar pembelajarannya, diantaranya model *Seq2Seq*, *Transformer*, RNN, GANs, dan *Variational Autoencoders* (Guo et al., 2020). Salah satu model CGM yang menggunakan pondasi *transformers* adalah *Text-to-Text Transfer Transformer (T5)*.

Model T5 adalah *framework text-to-text* yang menangani berbagai tugas seperti penerjemahan, menjawab pertanyaan, dan klasifikasi dengan memasukkan teks dan melatih model untuk menghasilkan teks target (Raffel et al., 2020). Hal ini memungkinkan penggunaan model, *loss function*, dan *hyperparameter* yang sama di berbagai tugas, memungkinkan perbandingan yang efektif antara tujuan pembelajaran transfer, kumpulan data tidak berlabel, dan faktor lainnya. Dengan memanfaatkan kemampuan tersebut, CGM dapat menghasilkan teks yang mirip dengan teks yang ditulis oleh orang yang mengalami gangguan mental, sehingga dapat digunakan untuk melatih model klasifikasi yang lebih akurat. Selain itu, CGM dapat digunakan untuk memodifikasi data teks media sosial yang tersedia, sehingga dapat mengatasi masalah ketidakseimbangan data yang sering terjadi dalam klasifikasi gangguan mental.

2.1.3. Causal Language Model (CLM)

Causal Language Model (CLM) adalah salah satu jenis teknik model bahasa yang memiliki tujuan untuk memprediksi urutan token selanjutnya dari token sebelumnya (Wolf et al., 2020). Pendekatan ini disebut *autoregressive* karena memungkinkan model menghasilkan teks secara berurutan, satu token pada satu waktu, berdasarkan konteks yang diberikan oleh semua token sebelumnya. Hal ini membuat model CLM seperti GPT sesuai untuk aplikasi seperti generasi teks, peringkasan, dan pembuatan *chatbot* (Naveed et al. 2024).

Perbedaan penting antara model CLM dan model LLM lainnya terletak pada bagian *decoder*. Di CLM, *dekoder* hanya melihat sebagian dari masukan yang di-*encode* pada satu waktu, umumnya token sebelumnya secara berurutan seperti yang ditunjukkan Gambar 2.2 (Wolf et al., 2020; Radford et al. 2021). Hal ini membuat model untuk fokus pada hubungan sebab akibat antar kata dan memprediksi kata berikutnya berdasarkan konteks yang telah dilihat sejauh itu. Mekanisme penutupan/*masking* ini mencegah model menyalin masukan atau menghasilkan keluaran berulang.

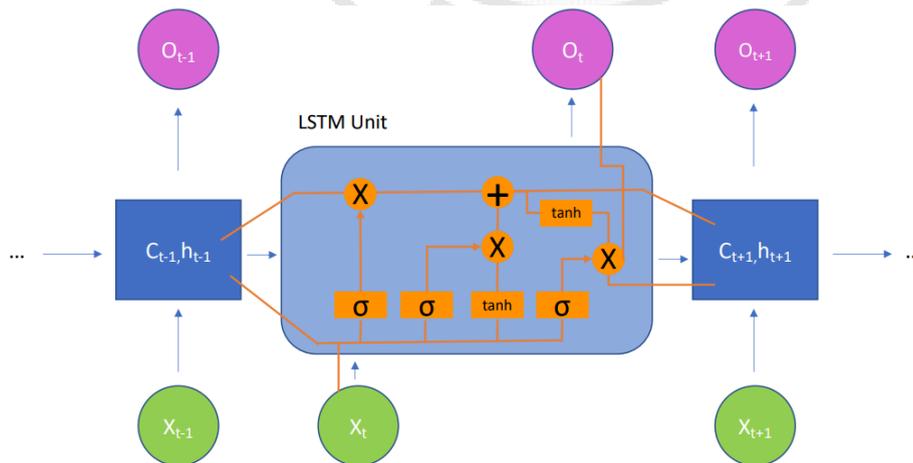


Gambar 2.2 Arsitektur model GPT

Sumber: Radford et al. (2021). *Improving Language Understanding by Generative Pre-Training*

2.1.4. Long Short-Term Memory

Long Short-Term Memory (LSTM) adalah jenis arsitektur *recurrent neural network* (RNN) yang diperkenalkan pada tahun 1997 oleh Hochreiter dan Schmidhuber untuk mengatasi masalah hilangnya gradien yang dihadapi oleh RNN tradisional (Nosouhian et al., 2021). LSTM dirancang untuk menangkap ketergantungan jangka panjang dalam data berurutan dengan menggabungkan sel memori yang dapat menyimpan informasi dalam jangka waktu yang lama (Wu et al., 2020).



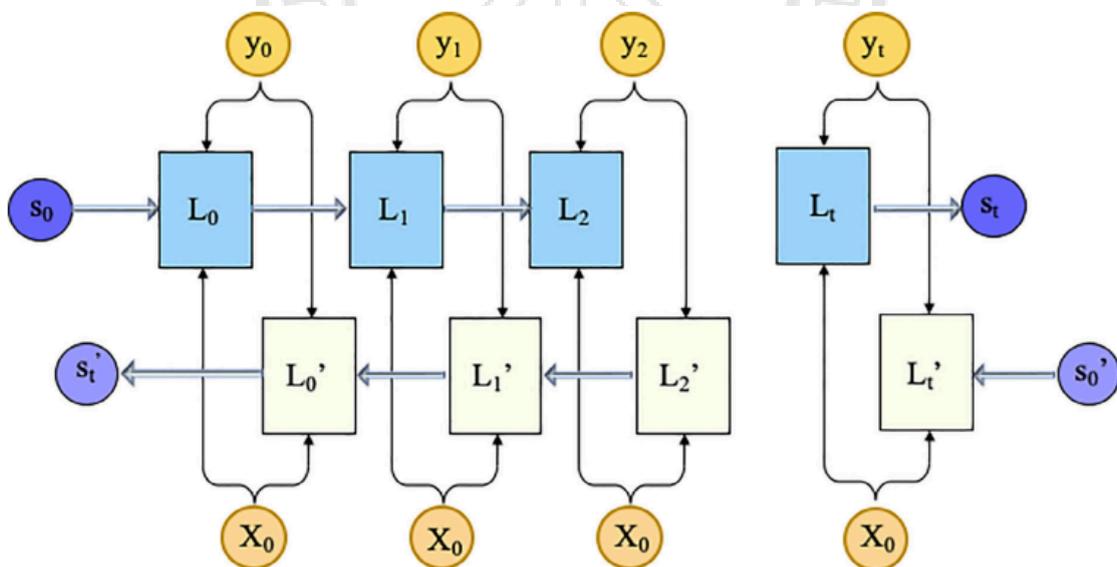
Gambar 2.3 Diagram satu unit LSTM

Sumber: Nosouhian et al., (2021). *A review of recurrent neural network architecture for sequence learning: Comparison between LSTM and GRU*

Seperti yang ditunjukkan pada gambar 2.3, satu unit LSTM menggunakan gerbang-gerbang yang ada dan status sel untuk mengatur aliran informasi melalui jaringan, memungkinkannya mempelajari dan mengingat pola dalam data berurutan selama periode waktu yang lama (Nosouhian et al., 2021). Kemampuan ini menjadikan LSTM sangat efektif dalam tugas-tugas yang memerlukan penangkapan ketergantungan jangka panjang, seperti pemodelan bahasa, pengenalan ucapan, dan prediksi deret waktu.

2.1.5. Bidirectional LSTM

Bidirectional LSTM (BiLSTM) adalah jenis *recurrent neural network* (RNN) yang terdiri dari dua lapisan LSTM: satu memproses urutan masukan dalam arah maju dan yang lainnya memrosesnya dalam arah mundur (Wu et al., 2020). Pemrosesan dua arah ini memungkinkan model untuk menangkap ketergantungan di kedua arah urutan masukan, memungkinkan pemahaman dan prediksi data deret waktu yang lebih baik. Penggunaan BiLSTM dalam model meningkatkan kemampuan untuk menangkap ketergantungan jangka panjang dalam data masukan dan meningkatkan akurasi perkiraan beban jangka pendek dalam tugas klasifikasi.

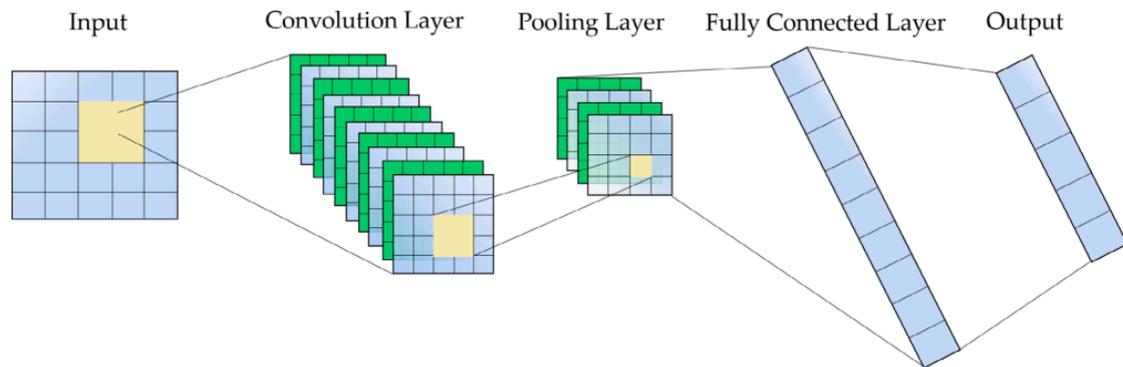


Gambar 2.4 Struktur dari BiLSTM

Sumber: Wu et al., (2020). *An attention-based CNN-LSTM-BiLSTM model for short-term electric load forecasting in integrated energy system*

2.1.6. Convolutional Neural Network

Convolutional Neural Network (CNN) adalah jenis jaringan saraf yang dirancang untuk secara otomatis dan adaptif mempelajari hirarki spasial fitur dari data masukan. Struktur CNN biasanya terdiri dari lapisan konvolusional diikuti oleh lapisan *pooling* dan lapisan yang terhubung sepenuhnya seperti yang ada pada gambar (Wu et al., 2020; Zhang et al., 2022).



Gambar 2.5 Struktur dari CNN

Sumber: Wu et al., (2020). *An attention-based CNN-LSTM-BiLSTM model for short-term electric load forecasting in integrated energy system*

CNN efektif dalam menangkap fitur lokal dalam data teks, sehingga cocok untuk tugas-tugas seperti analisis sentimen dan deteksi emosi (Zhang et al., 2022). Para peneliti telah menggabungkan CNN dengan fitur lain seperti LIWC (*Linguistic Inquiry and Word Count*), TF-IDF (*Term Frekuensi-Invers Dokumen Frekuensi*), BOW (*Bag of Words*), dan POS (*Part-of-Speech*) untuk meningkatkan kinerja model deteksi penyakit mental .

Secara keseluruhan, metode berbasis CNN memainkan peran penting dalam mengekstraksi fitur-fitur bermakna dari data tekstual, memungkinkan identifikasi indikator penyakit mental secara akurat di forum online dan platform media sosial.

2.1.7. Metropolis-Hastings

Algoritma *Metropolis-Hastings* (MH) adalah metode Markov chain Monte Carlo (MCMC) yang biasa digunakan untuk pengambilan sampel dari distribusi probabilitas kompleks (Cho et al., 2022). Dalam konteks augmentasi data, algoritma Metropolis-Hastings digunakan untuk mengambil sampel yang diperbesar dari distribusi target eksplisit (Park et al. 2021). Dalam penelitian yang dilakukan Cho et al. (2022), MH dapat memfasilitasi pembuatan data *multi-label* dari satu label positif melalui pendekatan metode *sampling S2M*. Dengan

memperbarui representasi ruang laten secara berulang menggunakan teknik MCMC, algoritma dapat mengeksplorasi ruang sampel multi-label yang mungkin dan menghasilkan keluaran berkualitas tinggi.

$$\alpha(x', x) = \min \left(1, \frac{p_{(I,J)}(x')/p_G(x')}{p_{(I,J)}(x)/p_G(x)} \right)$$

Rumus 2.1 Probabilitas sampel diterima ke target distribusi

Sumber: Cho et al. (2022). *Mining Multi-Label Samples from Single Positive Labels*

Algoritma ini bekerja dengan mengusulkan sampel baru berdasarkan keadaan *Markov chain* saat ini, menghitung probabilitas penerimaan sampel yang diusulkan, dan kemudian menerima atau menolak proposal berdasarkan rumus probabilitas yang ada pada Gambar 2.5 (Cho et al., 2022; Park et al. 2021). Proses berulang ini memungkinkan rantai untuk mengeksplorasi distribusi target dan memastikan bahwa sampel yang dihasilkan menyatu dengan distribusi target yang diinginkan, memungkinkan pembuatan sampel *multi-label* dari satu label positif sesuai konteks.

2.1.8. BERT

Representasi Encoder Dua Arah dari *transformers* (BERT) (Devlin et al., 2018) adalah model *transformers* yang dimaksudkan untuk mempelajari representasi bahasa alami dua arah yang mendalam dari korpus teks besar tanpa pengawasan. Untuk melatih modelnya, BERT menggunakan dua tugas: *Masked Language Model* (MLM) dan *Next Sentence Prediction* (NSP). Sebelum memasukkan urutan kata ke dalam BERT dalam tugas MLM, umumnya 15% token dijadikan token [MASK], dan model mencoba memprediksi nilai asli dari token yang tercakup berdasarkan istilah yang tidak disamarkan dalam urutan masukan.

Dalam tugas NSP, model BERT menerima sepasang kalimat sebagai masukan dan memprediksi apakah kalimat kedua pada pasangan tersebut merupakan kalimat berikutnya dalam teks asli dengan memahami hubungan antara dua kalimat. Tidak seperti model tradisional, yang hanya melihat urutan teks dari satu arah, mekanisme *encoder attention* BERT menggunakan pelatihan Transformer dua arah, yang mempelajari informasi dari sisi kiri dan kanan sebuah kata, sehingga memungkinkan model memperoleh pemahaman konteks bahasa yang lebih baik.

2.1.8.1. ALBERT

Saat melakukan beberapa tugas NLP, ditemukan bahwa meningkatkan ukuran model pra-pelatihan dapat memberikan hasil yang lebih tinggi (Haque et al., 2020). Namun, karena keterbatasan teknologi, seringkali ukuran model tidak dapat ditingkatkan. Untuk mengatasi masalah ini, Architecture Lite BERT (ALBERT) diperkenalkan. ALBERT memulai dengan memperkenalkan dua matriks alternatif untuk matriks penyematan kata. Dengan memisahkan ini, model sekarang dapat menyertakan berbagai ukuran lapisan tersembunyi, yang tidak berpengaruh pada kata lapisan penyematan. Desain ALBERT memungkinkan pertukaran parameter lintas lapisan. Ini menghentikan peningkatan parameter seiring dengan kedalaman jaringan. Ini juga mengatasi hilangnya prediksi kalimat berikutnya pada arsitektur model BERT. ALBERT telah menunjukkan kehilangan yang diawasi sendiri berdasarkan konsistensi kalimat. Model yang diusulkan mungkin belajar lebih cepat daripada BERT dengan memanfaatkan ALBERT yang telah dilatih sebelumnya.

2.1.8.2. RoBERTa

BERT juga dapat dilatih sebelumnya menggunakan Robust Optimized BERT Pre-Training Approach (RoBERTa). Tujuan dari model ini adalah untuk mengoptimalkan pelatihan arsitektur BERT untuk mengurangi waktu pra-pelatihan. Menurut Murarka et al. model ini dilatih dalam jangka waktu yang lebih lama dan dengan data serta kekuatan pemrosesan 10 kali lebih banyak daripada BERT.

RoBERTa menggabungkan peningkatan *pre-train* lebih lanjut dalam sistem pengawasan mandiri, memungkinkan peningkatan hasil dengan lebih sedikit ketergantungan pada pelabelan data. RoBERTa meningkatkan proses pelatihan dengan menghapus tugas Next Sentence Prediction (NSP) yang digunakan dalam pra-pelatihan BERT dan memperkenalkan penyembunyian dinamis selama pelatihan, sehingga token yang disamarkan bervariasi selama periode pelatihan (Murarka et al., 2021). Ukuran mini-batch yang lebih besar dan kecepatan pembelajaran yang lebih cepat juga terbukti lebih efektif dalam operasi pelatihan.

2.2. Tinjauan Studi

Berikut merupakan tinjauan studi yang berkaitan dengan generasi *data multi-label* untuk pembuatan model klasifikasi *multi-label*.

2.2.1. Classification of mental illnesses on social media using RoBERTa

Penelitian ini membahas tentang penggunaan teknik *Deep Learning*, *Transfer Learning*, dan *Machine Learning* untuk mendeteksi gangguan mental pada individu melalui teks media sosial (Murarka et al., 2021). Penelitian ini juga membahas tentang potensi teknik-teknik ini untuk melengkapi prosedur klinis dalam memprediksi kesehatan mental antara dua kelas individu: mereka yang mencari bantuan *online* dan mereka yang tidak menyadari kondisi mereka. Selain itu, penelitian ini juga membahas tentang *dataset* yang digunakan dan metodologi yang diterapkan untuk mendeteksi gangguan mental pada teks media sosial.

Hasil penelitian menunjukkan bahwa teknik-teknik ini memiliki potensi untuk melengkapi prosedur klinis dalam memprediksi kesehatan mental antara dua kelas individu: mereka yang mencari bantuan online dan mereka yang tidak menyadari kondisi mereka.

Selain itu, penelitian ini juga menunjukkan bahwa model RoBERTa yang telah dilatih sebelumnya dalam transfer learning menghasilkan hasil terbaik dengan akurasi sebesar 0,80 menjadikannya model yang sangat akurat dan berpotensi untuk mendukung prosedur klinis dalam memprediksi kesehatan mental.

Perbedaan penelitian ini dengan skripsi ini adalah hasil klasifikasi yang diprediksi. Pada penelitian ini, teks media sosial diklasifikasi dengan satu hasil label sedangkan skripsi ini akan mengklasifikasi teks media sosial dengan hasil berupa multi-label.

2.2.2. Mining Multi-Label Samples from Single Positive Labels

Penelitian ini membahas tentang pendekatan pengambilan sampel baru yang disebut *single-to-multi-label (S2M)* sampling untuk menghasilkan data multi-label berkualitas tinggi dengan biaya anotasi minimal menggunakan *Generative Adversarial Networks* (Cho et al., 2022).

Kelebihan dari model yang diusulkan adalah kemampuannya untuk menghasilkan data multi-label berkualitas tinggi dengan biaya anotasi minimal. Model ini juga terbukti efektif dan benar dalam eksperimen yang dilakukan pada dataset gambar nyata, bahkan dibandingkan dengan model yang dilatih dengan dataset yang sepenuhnya di-anotasi. Selain itu, model ini dapat diterapkan pada berbagai jenis model generatif, termasuk GAN berbasis pengelompokan dan model difusi.

Perbedaan antara penelitian ini dan skripsi ini terletak pada objek data dan arsitektur model generasi yang digunakan. Dalam penelitian ini, hasilnya berupa *dataset* gambar yang memakai GAN sebagai model yang menghasilkan data, sementara dalam skripsi ini,

pendekatan yang digunakan adalah membuat dataset berjenis teks dengan multi-label dan menggunakan model bahasa *Causal Language Model* sebagai model generasi data. Kemudian, dataset tersebut akan dilatih dengan model RoBERTa agar mampu mendeteksi teks di media sosial lainnya.

2.2.3. BERT-based Ensembles for Detecting Signs of Depression from Social Media Text

Penelitian ini membahas tentang pengembangan model prediktif untuk mendeteksi tanda-tanda depresi dari teks media sosial menggunakan teknik ensemble berbasis BERT (Janatdoust et al., 2022). Model ini kemudian diuji pada tugas *DepSign*, yang merupakan tugas klasifikasi sentimen dalam tiga kategori: "tidak depresi", "sedang depresi", dan "sangat depresi".

Penelitian ini juga membahas tentang klasifikasi sentimen dalam dua kategori dan tiga kategori serta penelitian terkait lainnya dalam bidang prediksi gangguan kesehatan mental dari media sosial. Hasil penelitian menunjukkan bahwa model yang diusulkan mengungguli model dasar dalam semua metrik yang dipertimbangkan dan mencapai *F1-score* sebesar 54% dan akurasi sebesar 61%, yang menempatkannya di peringkat ke-5 pada peringkat untuk tugas *DepSign*.

Perbedaan penelitian ini dengan skripsi ini adalah metode yang dipakai serta hasil klasifikasi yang diprediksi. Pada penelitian ini hanya mendeteksi tingkat depresi seseorang. Sedangkan skripsi ini akan mendeteksi penyakit mental dan dapat mendeteksi multi-label.

2.2.4. An attention-based CNN-LSTM-BiLSTM model for short-term electric load forecasting in integrated energy system

Tujuan dari penelitian ini adalah untuk mengatasi kompleksitas yang terkait dengan peramalan beban listrik jangka pendek dalam sistem energi terintegrasi (Wu et al., 2020). Penelitian ini bertujuan untuk mengusulkan model peramalan baru yang mengintegrasikan arsitektur *Convolutional Neural Network* (CNN), *Long Short-Term Memory* (LSTM), dan *BiDirectional Long Short-Term Memory* (BiLSTM) untuk meningkatkan akurasi peramalan beban.

Kelebihan dari model ini adalah dengan menggabungkan ketiga komponen tersebut, hasil akurasi dapat meningkat. Selain itu, model ini menunjukkan akurasi yang unggul dibandingkan metode tradisional dan teknik regresi lainnya, sehingga menunjukkan efektivitasnya dalam perkiraan beban jangka pendek dalam sistem energi terintegrasi.

Kekurangan dari model ini adalah kompleksitas yang mungkin timbul dari pengintegrasian beberapa tipe arsitektur, yang memerlukan penyesuaian parameter ekstensif untuk kinerja optimal. Ketergantungan pada data historis sebagai fitur masukan dapat membatasi kemampuan adaptasi model terhadap perubahan atau anomali sistem yang tiba-tiba, sehingga berpotensi berdampak pada keakuratan perkiraan beban jangka pendek. Selain itu, sumber daya komputasi yang diperlukan untuk melatih dan menerapkan model dengan beberapa lapisan jaringan saraf, khususnya dalam skenario perkiraan waktu nyata, dapat menimbulkan tantangan praktis.

Perbedaan penelitian ini dengan skripsi ini adalah kombinasi dari arsitektur yang akan dicoba. Skripsi ini akan mencoba kombinasi CNN dengan LSTM, CNN dengan BiLSTM, serta kombinasi dari ketiganya. Selain itu, juga terdapat perbedaan pada konteks *dataset*, penelitian ini menggunakan *dataset* histori data *integrated energy system* (IES) untuk prediksi beban jangka pendek IES, sedangkan penelitian ini menggunakan data gangguan mental bentuk teks

