



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

FORM PENGAJUAN JUDUL



Nama : Edbert

NIM : 211402004

Judul diajukan oleh\* : ☒ Dosen  
☒ Mahasiswa

Bidang Ilmu (tuliskan dua bidang) : 1. Data Science and Intelligent System  
2. Computer Graphics and Vision

Uji Kelayakan Judul\*\* : ☐ Diterima ☐ Ditolak

Hasil Uji Kelayakan Judul :

Calon Dosen Pembimbing I: Dedy Arisandi, S.T., M.Kom.  
(Jika judul dari dosen maka dosen tersebut berhak menjadi pembimbing I)

Calon Dosen Pembimbing II: Dr. Erna Budhiarti Nababan M.IT

Paraf Calon Dosen Pembimbing I

Medan, 26 Juli 2024

Ka. Laboratorium Penelitian,

\* Centang salah satu atau keduanya

\*\* Pilih salah satu

(Fanindia Purnamasari, S.TI., M.IT)

NIP. 198908172019032023





KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

RINGKASAN JUDUL YANG DIAJUKAN

\*Semua kolom di bawah ini diisi oleh mahasiswa yang sudah mendapat judul

<b>Judul / Topik Skripsi</b>	<b>Implementasi Algoritma LSTM dan IndoBERTweet Embedding untuk Deteksi Depresi Berdasarkan Postingan Media Sosial</b>
<b>Latar Belakang dan Penelitian Terdahulu</b>	<p><b>Latar Belakang</b></p> <p>Depresi merupakan gangguan mental yang ditandai dengan kesedihan yang mendalam dan terus-menerus, perasaan tidak berharga, bersalah, dan kehilangan minat dalam kehidupan sehari-hari, yang dapat mempengaruhi kondisi fisik dan kehidupan sosial seseorang. Gejala-gejala depresi sering kali meliputi perasaan mudah tersinggung, tertekan, cemas, tidak bersemangat, sedih, mood depresif, kehilangan kesenangan, perasaan bersalah, pola tidur dan nafsu makan terganggu, dan energi rendah. Dampak negatifnya bisa berupa kesulitan berkonsentrasi dan munculnya pikiran negatif. (Riziana et al., 2023)</p> <p>Ada beberapa teknik mengdiagnosis depresi, seperti <i>clinical interview method</i> dan <i>self-report questionnaires</i>. Kedua teknik tersebut memiliki kekurangan masing-masing yaitu tingkat akurasi yang tidak pasti dengan kemungkinan kesalahan mengdiagnosa akibat gejala yang mirip dengan penyakit mental lainnya (Smith et al., 2013). Selain itu, subjektivitas dan kemampuan pasien untuk mendeskripsikan gejala sangat berpengaruh terhadap hasil akhir diagnosis yang menjadi salah satu kesulitan dalam melakukan deteksi depresi. (Nordgaard et al., 2013)</p> <p>Menurut laporan Kementerian Kesehatan, angka penduduk yang berumur 15 tahun ke atas yang mengalami depresi pada tahun 2018 adalah 6.2% dan pada tahun 2023, angka ini hanya menurun 0.1%. Tidak hanya itu, berdasarkan laporan tersebut, hanya 10.4% penderita depresi yang melakukan pengobatan. Penderita depresi yang tidak diobati dapat mengakibatkan tindakan drastis seperti penggunaan narkoba ataupun bunuh diri. Oleh karena itu, depresi perlu dideteksi secara dini agar dapat dilakukan pengobatan.</p> <p>Teknologi <i>Machine Learning</i> dan <i>Deep Learning</i> bisa membantu dalam mendeteksi depresi. Penelitian sebelumnya yang dilakukan oleh Fine et al. (2020) yang berjudul “<i>Assessing population-level symptoms of anxiety, depression, and suicide risk in real time using NLP applied to social media data</i>”. Penelitian ini melakukan analisis tingkat kecemasan, depresi, dan bunuh diri dari tanggal 1 Januari 2020 hingga 1 Juni 2020 dengan data yang dikumpulkan dari Twitter (X) di negara Amerika Serikat. Model yang digunakan adalah <i>Bi-directional Long Short-Term Memory (BiLSTM)</i> dengan <i>GloVe</i> sebagai word embedding untuk meminimalisir overfitting. Akurasi yang dihasilkan untuk memprediksi depresi dan bunuh diri adalah 72% dan 73%.</p> <p>Penelitian yang dilakukan oleh Amanat et al. (2022) menggunakan <i>one-hot encoding</i> dan <i>principal component analysis (PCA)</i> sebagai teknik pre-processing yang berjudul “<i>Deep Learning for Depression Detection from Textual Data</i>”. Penelitian ini menggunakan algoritma LSTM-RNN yang menggunakan 60 units dan hasil akurasi yang didapatkan berupa 0.98 precision, 0.99 recall, dan 0.98 f1-measure.</p> <p>Penelitian lain yang menggunakan <i>Long Short-Term Memory (LSTM)</i> untuk mendeteksi depresi dilakukan oleh Singh et al. (2022) yang berjudul “<i>An embedded LSTM based scheme for depression detection and analysis</i>”. Dataset yang digunakan berasal dari hasil scraping menggunakan PushShift API untuk mendapatkan postingan dari Reddit yang berhubungan</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

**PROGRAM STUDI S1 TEKNOLOGI INFORMASI**

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

dengan kesehatan mental dari beberapa *subreddit*. Penelitian ini melakukan *preprocessing* berupa *tokenization*, *data cleaning*, dan *stemming*. Kemudian, data tersebut diubah menjadi vektor menggunakan GloVe *embedding*. Akurasi yang didapatkan menggunakan algoritma LSTM dan GloVe *embedding* adalah 94.88%.

Kemudian, terdapat penelitian yang dilakukan oleh Jain et al. (2022) yang berjudul “*Depression and Suicide Analysis Using Machine Learning and NLP*” menggunakan model *Logistic Regression*, *Naïve Bayes*, *Support Vector Machines*, dan *Random Forest* untuk memprediksi apakah sebuah text berunsur depresi atau bunuh diri. Model yang menghasilkan nilai akurasi tertinggi adalah *Random Forest* dengan nilai 77.298%

Penelitian yang menggunakan *Random Forest Classifier*, *XGBoost*, dan *Support Vector Machine* untuk melakukan diagnosa depresi yang bisa membedakan dengan PTSD (*Post-Traumatic Stress Disorder*) dilakukan oleh Lorenzoni et al. (2024) yang berjudul “*Assessing ML Classification Algorithms and NLP Techniques for Depression Detection: An Experimental Case Study*”. Dataset yang digunakan berasal dari *Distress Analysis Interview Corpus - Wizard-of-Oz* yang berisi data audio, video, dan transkrip dari interview yang dilakukan. Akurasi paling tinggi dihasilkan oleh model *Random Forest Classifier* dan *XGBoost* dengan nilai akurasi 83.8%.

*Word embedding* yang digunakan dalam mendeteksi depresi tidak hanya terbatas pada GloVe tetapi juga bisa menggunakan salah satu *contextualized word embedding* yaitu BERT seperti yang dilakukan oleh Devika et al. (2023) yang berjudul “*BERT-Based Approach for Suicide and Depression Identification*”. Penelitian ini melakukan *preprocessing* berupa *tokenization*, *removing stop words*, dan *part-of-speech tagging*. Lalu, data tersebut diubah menjadi vektor menggunakan BERT *embedding* dan menggunakan algoritma *Bi-LSTM* untuk melakukan prediksi. Dataset yang digunakan berasal dari media sosial Reddit. Akurasi yang didapatkan ketika menggunakan *Bi-LSTM* adalah 81%, sedangkan BERT model menghasilkan akurasi sebesar 72%.

BERT *embedding* yang dikembangkan oleh Google banyak digunakan oleh peneliti-peneliti sehingga ada beberapa model yang dikembangkan dari BERT seperti ARABERT dan MARBERT yang digunakan oleh El-Ramly et al. (2021) pada penelitian yang berjudul “*CairoDep: Detecting Depression in Arabic Posts Using BERT Transformers*”. Dataset yang digunakan berasal dari CairoDep v1.0 yang memuat 7000 data depresi dan non-depresi dalam bahasa arab. Kemudian, dilakukan *data cleaning* dan *data preprocessing* sebelum digunakan untuk training pada model ARABERT dan MARBERT untuk melakukan prediksi. Akurasi yang diperoleh ARABERT adalah 96.93% sedangkan MARBERT menghasilkan akurasi yaitu 96.07%.

Penelitian lain yang menggunakan BERT dalam mendeteksi depresi dilakukan Hidayat & Maharani (2022) yang berjudul “*General Depression Detection Analysis Using IndoBERT Method*”. Penelitian ini menggunakan data yang dikumpulkan dari Twitter (X). Selanjutnya, dilakukan *preprocessing data* agar data dapat digunakan untuk training pada IndoBERT model. Akurasi yang didapatkan masih pada angka 51%. Hal ini disebabkan oleh beberapa kata yang masih belum bisa diproses oleh IndoBERT seperti ‘wkwkwk’ dan kata-kata diluar Bahasa Indonesia.

Berdasarkan penelitian yang telah dilakukan, penulis mengusulkan penerapan IndoBERTweet *Embedding* dan algoritma *Long Short-Term Memory* untuk mendeteksi postingan pada media sosial yang memiliki unsur depresi dan keinginan bunuh diri. Penulis memberikan penelitian ini



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

**PROGRAM STUDI S1 TEKNOLOGI INFORMASI**

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

judul “Implementasi Algoritma LSTM dan IndoBERTweet Embedding untuk Deteksi Depresi Berdasarkan Postingan Media Sosial”.

**Penelitian Terdahulu**

No.	Penulis	Judul	Tahun
1.	Alex B. Fine, Patrick Crutchley, Jenny Blase, Joshua Carroll, dan Glen Coppersmith	Assessing population-level symptoms of anxiety, depression, and suicide risk in real time using NLP applied to social media data	2020
2.	Amna Amanat, Muham mad Rizwan, Abdul Rehman Javed, Maha Abdelhaq, Raed Alsaqour, Sharnil Pandya, dan Mueen Uddin	Deep Learning for Depression Detection from Textual Data	2022
3.	Jaskaran Singh, Mohammad Wzid, D.P. Singh, dan Sumit Pundir	An embedded LSTM based scheme for depression detection and analysis	2022
4.	Pratyaksh Jain , Karthik Ram Srinivas, dan Abhishek Vichare	Depression and Suicide Analysis Using Machine Learning and NLP	2022
5.	Guiliano Lorenzoni, Cristina Tavares, Nathanlia Nascimento, Paulo Alencar, Donald Cowan David R. Cheriton	Assessing ML Classification Algorithms and NLP Techniques for Depression Detection: An Experimental Case Study	2024



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	6.	S.P. Devika, M.R. Pooja, M.S. Arpitha, dan Ravi Vinayakumar	BERT-Based Approach for Suicide and Depression Identification	2023
	7.	Mohammad El- Ramly, hager Abu-Elyazid, Youseef Mo'men, Gameel Alshaer, Nardine Adib, Kareen Alla Eldeen, dan Mariam El- Shazly	CairoDep: Detecting depression in Arabic posts using BERT transformers	2021
	8.	Ilham Rizki Hidayat dan Warih Maharani	General Depression Detection Analysis Using IndoBERT Method	2022
<b>Rumusan Masalah</b>	<p>Depresi merupakan masalah kesehatan mental yang sangat serius dan perlu dideteksi serta dicegah secara dini. Teknik manual seperti interview atau kuesioner yang digunakan pakar masih belum bisa mendiagnosis individu secara akurat. Banyak individu yang mengalami depresi tidak mengungkapkan masalah mereka secara terbuka, melainkan memilih untuk mencurahkan perasaan dan kondisi mental mereka melalui media sosial seperti Twitter (X). Postingan di media sosial yang berhubungan dengan depresi sering kali mengandung ungkapan tentang masalah pribadi seperti <i>bullying</i>, kekerasan, tekanan stres, masalah keluarga, dan lain-lain. Namun, mendeteksi postingan yang mengindikasikan depresi pada media sosial secara manual bersifat subjektif dan memakan waktu lama dalam memahami konteks setiap postingan. Hal ini menjadi pemicu dibutuhkannya suatu pendekatan yang efektif untuk mendeteksi depresi pada postingan di media sosial.</p>			



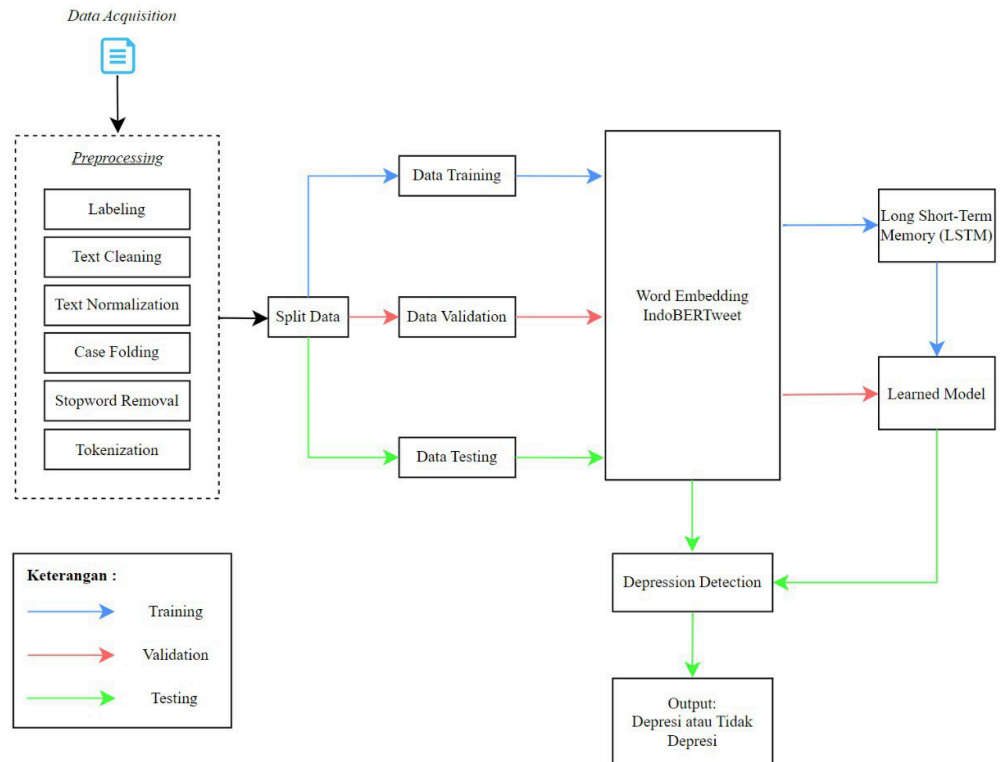
# KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

## PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

### Metodologi



### Tahapan Penelitian:

#### 1. Data Acquisition

Pada tahap ini, dilakukan pengumpulan data postingan dari media sosial pada platform Twitter (X). Data yang dikumpulkan berasal dari Kaggle dan teknik *scraping*.

#### 2. Preprocessing

Setelah melakukan *labeling*, akan dilanjutkan dengan tahapan *preprocessing* lainnya untuk menghasilkan data yang lebih mudah dimengerti oleh model. Tahapan *preprocessing* tersebut terdiri dari *labeling*, *text cleaning*, *text normalization*, *case folding*, *stopword removal*, dan *tokenization*.

#### 3. Labeling

Pada tahap ini, dilakukan proses *labeling* untuk data *training*, *validation*, dan *testing*. Untuk setiap postingan yang mengandung pernyataan depresi akan diberikan kelas 1, sedangkan untuk postingan yang tidak mengandung pernyataan depresi akan diberikan kelas 0.

#### 4. Text Cleaning

Tahap *text cleaning* dilakukan untuk membersihkan data dari karakter-karakter tertentu yang tidak diperlukan dalam proses deteksi seperti emoji, spasi yang berlebihan, link URL, dan lainnya. Tahap ini penting untuk dilakukan untuk mengurangi tingkat kegagalan dalam proses mendeteksi.

#### 5. Text Normalization

Tahap *text normalization* dilakukan untuk mengubah kata-kata *slang*, tidak baku, akronim, dan bahasa kolokial menjadi bahasa yang baku sesuai dengan KBBI. Tahap ini juga memastikan untuk memperbaiki kata-kata yang memiliki kesalahan ejaan agar dapat diubah ke kata yang benar.

#### 6. Case Folding

Pada tahap ini, dilakukan transformasi untuk setiap kata yang mengandung huruf besar menjadi huruf kecil untuk menghindari kesalahan dalam *tokenization* akibat





KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	<p>perbedaan huruf. Contohnya: Sedih dan sedih akan memiliki token yang berbeda jika tidak diubah menjadi huruf kecil.</p> <p>7. <b>Stopword Removal</b></p> <p>Tahap <i>stopword removal</i> dilakukan untuk menghilangkan kata-kata yang tidak berpengaruh secara signifikan pada makna kalimat secara keseluruhan. Contoh dari kata-kata <i>stopword</i> aku, kamu, dia, dan, atau, tetapi, ke-, dari, dan lainnya.</p> <p>8. <b>Tokenization</b></p> <p>Tahap <i>tokenization</i> merupakan tahap terakhir dari <i>preprocessing</i> yang bertujuan untuk mengubah kalimat menjadi unit kecil yang disebut token. Untuk <i>word embedding</i> IndoBERTweet, diperlukan token khusus [CLS] pada awal kalimat dan [SEP] untuk akhir kalimat. Setelah tahap <i>tokenization</i>, akan dilanjutkan dengan tahap <i>word embedding</i> menggunakan IndoBERTweet.</p> <p>9. <b>Split Data</b></p> <p>Setelah melakukan tahap <i>preprocessing</i>, data akan dibagi menjadi 3, yaitu data <i>training</i>, <i>validation</i>, dan <i>testing</i>. Rasio yang digunakan untuk pembagian tersebut adalah 50% data <i>training</i>, 20% data <i>validation</i>, dan 30% data <i>testing</i>.</p> <p>10. <b>Word Embedding</b></p> <p>Setelah dilakukan <i>split data</i>, data akan diubah menjadi vektor dengan <i>word embedding</i> IndoBERTweet. IndoBERTweet embedding adalah model yang dikembangkan dari model BERT dan BERTweet yang menghasilkan vektor berdasarkan makna kata yang terdapat pada keseluruhan kalimat.</p> <p>11. <b>Model</b></p> <p>Pada tahap ini, pelatihan model dilakukan menggunakan data vektor yang dibentuk oleh <i>word embedding</i> IndoBERTweet menggunakan algoritma <i>Long Short-Term Memory</i> (LSTM). Setelah melakukan <i>training</i> dan <i>validation</i>, dihasilkan <i>learned model</i>.</p> <p>12. <b>Output</b></p> <p><i>Learned model</i> yang dihasilkan akan digunakan untuk melakukan deteksi postingan yaitu Depresi atau Tidak Depresi.</p>
Referensi	<p>Amanat, A., Rizwan, M., Javed, A. R., Abdelhaq, M., Alsaqour, R., Pandya, S., &amp; Uddin, M. (2022). Deep Learning for Depression Detection from Textual Data. <i>Electronics (Switzerland)</i>, 11(5). <a href="https://doi.org/10.3390/electronics11050676">https://doi.org/10.3390/electronics11050676</a></p> <p>Devika, S. P., Pooja, M. R., Arpitha, M. S., &amp; Vinayakumar, R. (2023). BERT-Based Approach for Suicide and Depression Identification. <i>Proceedings of Third International Conference on Advances in Computer Engineering and Communication Systems</i>, 435–444.</p> <p>El-Ramly, M., Abu-Elyazid, H., Mo'men, Y., Alshaer, G., Adib, N., Eldeen, K. A., &amp; El-Shazly, M. (2021). CairoDep: Detecting Depression in Arabic Posts Using BERT Transformers. <i>Proceedings - 2021 IEEE 10th International Conference on Intelligent Computing and Information Systems, ICICIS 2021</i>, 207–212. <a href="https://doi.org/10.1109/ICICIS52592.2021.9694178">https://doi.org/10.1109/ICICIS52592.2021.9694178</a></p> <p>Fine, A. B., Crutchley, P., Blase, J., Carroll, J., &amp; Coppersmith, G. (2020). Assessing population-level symptoms of anxiety, depression, and suicide risk in real time using NLP applied to social media data. <i>Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science</i>, 50–54.</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN  
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

**PROGRAM STUDI S1 TEKNOLOGI INFORMASI**

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155  
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

- Hidayat, I. R., & Maharani, W. (2022). General Depression Detection Analysis Using IndoBERT Method. *International Journal on Information and Communication Technology (IJoICT)*, 8(1), 41–51. <https://doi.org/10.21108/ijoiect.v8i1.634>
- Jain, P., Srinivas, K. R., & Vichare, A. (2022). Depression and Suicide Analysis Using Machine Learning and NLP. *Journal of Physics: Conference Series*, 2161(1). <https://doi.org/10.1088/1742-6596/2161/1/012034>
- Lorenzoni, G., Tavares, C., Nascimento, N., Alencar, P., & Cowan, D. (2024). *Assessing ML Classification Algorithms and NLP Techniques for Depression Detection: An Experimental Case Study*. <http://arxiv.org/abs/2404.04284>
- Nordgaard, J., Sass, L. A., & Parnas, J. (2013). The psychiatric interview: Validity, structure, and subjectivity. *European Archives of Psychiatry and Clinical Neuroscience*, 263(4), 353–364. <https://doi.org/10.1007/s00406-012-0366-z>
- Riziana, K. F., Fatmawati, & Darmawan, A. (2023). HUBUNGAN TINGKAT GEJALA DEPRESI DENGAN IDE BUNUH DIRI PADA REMAJA SEKOLAH MENENGAH ATAS. *Journal of Medical Studies*, 3(1), 39–47.
- Singh, J., Wazid, M., Singh, D. P., & Pundir, S. (2022). An embedded LSTM based scheme for depression detection and analysis. *Procedia Computer Science*, 215, 166–175. <https://doi.org/10.1016/j.procs.2022.12.019>
- Smith, K. M., Renshaw, P. F., & Bilello, J. (2013). The diagnosis of depression: Current and emerging methods. *Comprehensive Psychiatry*, 54(1), 1–6. <https://doi.org/10.1016/j.comppsy.2012.06.006>

Medan, 26 Juli 2024  
Mahasiswa yang mengajukan,

(Edbert)

NIM. 211402004