KLASIFIKASI TERHADAP KONDISI KESEHATAN MENTAL MENGGUNAKAN ALGORITMA *XGBOOST* DENGAN *SEMANTIC*, *SYNTACTIC*, DAN *STRUCTURAL FEATURES*

SKRIPSI

Disusun Oleh:

SYARIFATUL UMMAH

09040621070

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA
2025

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama	: Syarifatul Ummah
NIM	: 09040621070
Program Studi	: Sistem Informasi
Angkatan	: 2021

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul: "KLASIFIKASI TERHADAP KONDISI KESEHATAN MENTAL MENGGUNAKAN ALGORITMA *XGBOOST* DENGAN *SEMANTIC, SYNTACTIC,* DAN *STRUCTURAL FEATURES*". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 06 Juni 2025
Yang menyatakan,

Syarifatul Ummah
NIM. 09040621070

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh,

NAMA	:	SYARIFATUL UMMAH
NIM	:	09040621070
JUDUL	:	KLASIFIKASI TERHADAP KONDISI KESEHATAN MENTAL MENGGUNAKAN ALGORITMA *XGBOOST* DENGAN *SEMANTIC, SYNTACTIC,* DAN *STRUCTURAL FEATURES*

Ini telah diperiksa dan disetujui untuk diujikan.

Surabaya, 02 Juni 2025

Dosen Pembimbing 1	Dosen Pembimbing 2
Khalid, M.Kom	Dr. Eng. Anang Kunaefi, M. Kom
NIP. 197906092014031002	NIP. 197911132014031001

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi Syarifatul Ummah ini telah dipertahankan
di depan tim penguji skripsi
di Surabaya, 13 Juni 2025

Mengesahkan,

Dewan Penguji

Penguji 1

Dwi Rolliawati, M.T

NIP. 197909272014032001

Penguji 2

Ahmad Yusuf, M.Kom

NIP. 199001202014031003

Penguji 3

Khalid, M.Kom

NIP. 197906092014031002

Penguji 4

Dr. Eng. Anang Kunaefi, M.Kom

NIP. 197911132014031001

Mengetahui,

Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya

Dr. A. Saepul Hamdani, M.Pd

NIP. 196507312000031002

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN
Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : SYARIFATUL UMMAH
NIM : 09040621070
Fakultas/Jurusan : SAINS DAN TEKNOLOGI / SISTEM INFORMASI
E-mail address : syarifaum@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :
■ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :

KLASIFIKASI TERHADAP KONDISI KESEHATAN MENTAL MENGGUNAKAN ALGORITMA *XGBOOST* DENGAN *SEMANTIC, SYNTACTIC,* DAN *STRUCTURAL FEATURES*

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Ekslusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara ***fulltext*** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 20 Juni 2025

Penulis

Syarifatul Ummah

ABSTRAK

KLASIFIKASI TERHADAP KONDISI KESEHATAN MENTAL MENGGUNAKAN ALGORITMA *XGBOOST* DENGAN *SEMANTIC, SYNTACTIC*, DAN *STRUCTURAL FEATURES*

Oleh:
Syarifatul Ummah

Kesehatan mental tidak kalah pentingnya dengan kesehatan fisik, namun sering kurang mendapat perhatian yang memadai. Kesehatan mental kini menjadi isu global yang semakin mendesak dengan prevalensi gangguan mental yang terus meningkat di seluruh dunia. Penelitian ini bertujuan untuk mengklasifikasikan kondisi kesehatan mental berdasarkan pernyataan individu menggunakan algoritma XGBoost dengan ekstraksi fitur *semantic, syntactic*, dan *structural.* Model diuji secara tunggal dan kombinasi untuk menentukan representasi fitur terbaik. Hasil evaluasi menunjukkan bahwa model XGBoost pada kombinasi fitur *semantic syntactic,* dan *structural* memiliki nilai akurasi tertingi sebesar 0.74. Fitur *semantic* memberikan hasil terbaik ketika digunakan secara tunggal sebesar 0.69, sementara *syntactic* sebesar 0.44 dan *structural* sebesar 0.43. Kombinasi fitur *semantic* dan *syntactic* sebesar 0.71, kombinasi fitur *semantic* dan *structural* sebesar 0.73, sementara kombinasi *syntactic* dan *structural* lebih rendah dengan hasil akurasi 0.50. Hasil penelitian menunjukkan bahwa efektivitas XGBoost tidak hanya bergantung pada algoritma, tetapi juga dipengaruhi oleh jenis ekstraksi fitur yang digunakan. Kombinasi fitur *semantic, syntactic* dan *structural* terbukti memberikan representasi data yang paling optimal untuk klasifikasi kondisi kesehatan mental.

Kata Kunci: Klasifikasi Teks, Kesehatan Mental, *Machine Learning*, XGBoost, *Feature Extraction*.

ABSTRACT

CLASSIFICATION OF MENTAL HEALTH CONDITIONS USING XGBOOST ALGORITHM WITH SEMANTIC, SYNTACTIC, AND STRUCTURAL FEATURES

***By*:**

Syarifatul Ummah

Mental health is no less important than physical health, but often receives insufficient attention. Mental health has become an increasingly urgent global issue, with the prevalence of mental disorders continuing to rise worldwide. This study aims to classify mental health conditions based on individual statements using the XGBoost algorithm with semantic, syntactic, and structural feature extraction. The model was tested individually and in combination to determine the best feature representation. Evaluation results showed that the XGBoost model with a combination of semantic, syntactic, and structural features achieved the highest accuracy of 0.74. Semantic features give the best results when used individually at 0.69, while syntactic yielded 0.44 and structural 0.43. The combination of semantic and syntactic features yielded 0.71, the combination of semantic and structural features yielded 0.73, while the combination of syntactic and structural yielded a lower accuracy of 0.50. The results of the study show that the effectiveness of XGBoost does not only depend on the algorithm, but is also influenced by the type of feature extraction used. The combination of semantic, syntactic and structural features has proven to provide the most optimal data representation for the classification of mental health conditions.

***Keywords**: Text Classification, Mental Health, Machine Learning, XGBoost, Feature Extraction.*

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR PUSTAKA

Adriani, M., Asian, J., Nazief, B., Tahaghoghi, S. M. M., & Williams, H. E. (2007). Stemming Indonesian: A Confix-Stripping Approach. *Acm Transactions On Asian Language Information Processing (Talip)*, *6*(4), 1–33. Https://Doi.Org/10.1145/1316457.1316459

Ahmad, M., Wahid, N., Hamid, R. A., Sadiq, S., Mehmood, A., & Choi, G. S. (2022). Decision Level Fusion Using Hybrid Classifier For Mental Disease Classification. *Computers, Materials And Continua*, *72*(3), 5041–5058. Https://Doi.Org/10.32604/Cmc.2022.026077

Al-Shatnwai, A. M., & Faris, M. (2020). Predicting Customer Retention Using Xgboost And Balancing Methods. *International Journal Of Advanced Computer Science And Applications (Ijacsa)*, *11*(7), 704–712. Https://Doi.Org/10.14569/Ijacsa.2020.0110785

Ali, Z. A., Abduljabbar, Z. H., Tahir, H. A., Sallow, A. B., & Almufti, S. M. (2023). Exploring The Power Of Extreme Gradient Boosting Algorithm In Machine Learning: A Review. *Academic Journal Of Nawroz University*, *12*(2), 320–334. Https://Doi.Org/10.25007/Ajnu.V12n2a1612

Aliwy, A. H. (2012). Tokenization As Preprocessing For Arabic Tagging System. *International Journal Of Information And Education Technology*, *January 2012*, 348–353. Https://Doi.Org/10.7763/Ijiet.2012.V2.149

Alnaggar, M., Handosa, M., Medhat, T., & Z. Rashad, M. (2023). Thyroid Disease Multi-Class Classification Based On Optimized Gradient Boosting Model. *Egyptian Journal Of Artificial Intelligence*, *2*(1), 1–14. Https://Doi.Org/10.21608/Ejai.2023.205554.1008

Anggraini, N., Putra, S. J., Wardhani, L. K., Arif, F. D. U., Hakiem, N., & Shofi, I. M. (2024). A Comparative Analysis Of Random Forest, Xgboost, And Lightgbm Algorithms For Emotion Classification In Reddit Comments. *Jurnal Teknik Informatika*, *17*(1), 88–97. Https://Doi.Org/10.15408/Jti.V17i1.38651

Ardana, A. (2023). Performance Analysis Of Xgboost Algorithm To Determine The Most Optimal Parameters And Features In Predicting Stock Price Movement. *Telematika: Jurnal Informatika Dan Teknologi Informasi*, *20*(1), 91–102. Https://Doi.Org/10.31315/Telematika.V20i1.9329

Ardeshirifar, R. (2024). Automated Classification Of Dry Bean Varieties Using Xgboost And Svm Models. *Arxiv Preprint Arxiv:2408.01244*, 1–8.

Bell, R. D., Brendel, M., Konnaris, M. A., Xiang, J., Otero, M., Fontana, M. A., Bai, Z., Raychaudhuri, S., Rao, D. A., Pitzalis, C., Wei, K., Watts, G. F. M., Utz, P. J., Turner, J. D., Tabechian, D., Sutherby, D., Slowikowski, K., Seshadri, A., Seifert, J., … Wang, F. (2024). Automated Multi-Scale Computational Pathotyping (Amscp) Of Inflamed Synovial Tissue. *Nature Communications*, *15*(7503). Https://Doi.Org/10.1038/S41467-024-51012-6

Chen, T., & Guestrin, C. (2016). Xgboost: A Scalable Tree Boosting System.

Proceedings Of The Acm Sigkdd International Conference On Knowledge Discovery And Data Mining, *13-17-Augu*, 785–794. Https://Doi.Org/10.1145/2939672.2939785

Cheng, J.-P., & Haw, S.-C. (2023). Mental Health Problems Prediction Using Machine Learning Techniques. *International Journal On Robotics, Automation And Sciences*, *5*(2), 59–72. Https://Doi.Org/10.33093/Ijoras.2023.5.2.7

Chiche, A., & Yitagesu, B. (2022). Part Of Speech Tagging: A Systematic Review Of Deep Learning And Machine Learning Approaches. *Journal Of Big Data*, *9*(1). Https://Doi.Org/10.1186/S40537-022-00561-Y

Dash, A., Zhang, D., & Zhou, L. (2021). Personalized Ranking Of Online Reviews Based On Consumer Preferences In Product Features. *International Journal Of Electronic Commerce*, *25*(1), 29–50. Https://Doi.Org/10.1080/10864415.2021.1846852

Dave, N. R., Mehta, M. A., & Kotecha, K. (2024). Effective Stemmers Using Trie Data Structure For Enhanced Processing Of Gujarati Text. *International Journal Of Computational Intelligence Systems*, *17*(1). Https://Doi.Org/10.1007/S44196-024-00679-2

Donga, B. C., Pitroda, P. D., Domadiya, D. H. B., & Domadiya, D. D. H. (2024). Decoding The Future: A Comprehensive Review Of Machine Learning Innovations And Applications. *International Journal For Research In Applied Science And Engineering Technology*, *12*(7), 868–873. Https://Doi.Org/10.22214/Ijraset.2024.63667

Dwinanda, M. W., Satyahadewi, N., & Andani, W. (2023). Classification Of Student Graduation Status Using Xgboost Algorithm. *Barekeng: Jurnal Ilmu Matematika Dan Terapan*, *17*(3), 1785–1794. Https://Doi.Org/10.30598/Barekengvol17iss3pp1785-1794

English Penn Treebank Pos Tagset | Sketch Engine. (2015). Https://Www.Sketchengine.Eu/Penn-Treebank-Tagset/

Fakhriyani, D. V. (2019). Kesehatan Mental. In *Duta Media Publishing*. Https://Www.Researchgate.Net/Profile/Diana-Fakhriyani/Publication/348819060_Kesehatan_Mental/Links/60591b56458515e834643f66/Kesehatan-Mental.Pdf

Ferrouhi, E. M., & Bouabdallaoui, I. (2024). A Comparative Study Of Ensemble Learning Algorithms For High-Frequency Trading. *Scientific African*, *24*(February), E02161. Https://Doi.Org/10.1016/J.Sciaf.2024.E02161

Firdaus, I. A., Rolliawati, D., Kunaefi, A., & Firdaus. (2023). Classifying Product Review Quality Based On Semantic And Structural Features. *Indonesian Journal Of Electrical Engineering And Computer Science*, *32*(3), 1495–1502. Https://Doi.Org/10.11591/Ijeecs.V32.I3.Pp1495-1502

Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *Annals Of Statistics*, *29*(5), 1189–1232.

Https://Doi.Org/10.1214/Aos/1013203451

Gautam, S., Jain, A., Chaudhary, J., Gautam, M., Gaur, M., & Grover, S. (2024). Concept Of Mental Health And Mental Well-Being, It's Determinants And Coping Strategies. *Indian Journal Of Psychiatry*, *66*(Suppl 2), S231–S244. Https://Doi.Org/10.4103/Indianjpsychiatry.Indianjpsychiatry_707_23

Goyal, P., Rani, R., & Singh, K. (2024). An Efficient Ranking-Based Ensembled Multiclassifier For Neurodegenerative Diseases Classification Using Deep Learning. *Journal Of Neural Transmission*, *132*, 67–93. Https://Doi.Org/Https://Doi.Org/10.1007/S00702-024-02830-X

Gunawan, D. (2016). Evaluasi Performa Pemecahan Database Dengan Metode Klasifikasi Pada Data Preprocessing Data Mining. *Khazanah Informatika : Jurnal Ilmu Komputer Dan Informatika*, *2*(1), 10–13. Https://Doi.Org/10.23917/Khif.V2i1.1749

Harmer, B., Lee, S., Rizvi, A., & Saadabadi, A. (2024). Suicidal Ideation. *Acute Medicine: A Symptom-Based Approach*, 415–420. Https://Doi.Org/10.1007/9781139600354.061

Hidayati, L. N., & Harsono, M. (2021). Tinjauan Literatur Mengenai Stres Dalam Organisasi. *Jurnal Ilmu Manajemen*, *18*(1), 20–30. Https://Journal.Uny.Ac.Id/Index.Php/Jim/Article/View/39339/15281

Hong, S., Park, C., & Cho, S. (2021). A Rail-Temperature-Prediction Model Based On Machine Learning: Warning Of Train-Speed Restrictions Using Weather Forecasting. *Sensors (Basel, Switzerland)*, *21*(4606). Https://Doi.Org/10.3390/S21134606

Ikram, M., Shaikh, N. F., Vishwanatha, J. K., & Sambamoorthi, U. (2023). Leading Predictors Of Covid-19-Related Poor Mental Health In Adult Asian Indians: An Application Of Extreme Gradient Boosting And Shapley Additive Explanations. *International Journal Of Environmental Research And Public Health*, *20*(1). Https://Doi.Org/10.3390/Ijerph20010775

Islam, M. M., Hassan, S., Akter, S., Jibon, F. A., & Sahidullah, M. (2024). A Comprehensive Review Of Predictive Analytics Models For Mental Illness Using Machine Learning Algorithms. *Healthcare Analytics*, *6*(June), 100350. Https://Doi.Org/10.1016/J.Health.2024.100350

Islam, S. F. N., Sholahuddin, A., & Abdullah, A. S. (2021). Extreme Gradient Boosting (Xgboost) Method In Making Forecasting Application And Analysis Of Usd Exchange Rates Against Rupiah. *Journal Of Physics: Conference Series*, *1722*(1). Https://Doi.Org/10.1088/1742-6596/1722/1/012016

Javed, T., Raza, A., Maqbool, H. F., Zafar, S., Taborri, J., & Rossi, S. (2024). Multi-Class Classification Of Human Activity And Gait Events Using Heterogeneous Sensors. *Journal Of Sensor And Actuator Networks*, *13*(6). Https://Doi.Org/10.3390/Jsan13060085

Kapusta, J., Drlik, M., & Munk, M. (2021). Using Of N-Grams From Morphological Tags For Fake News Classification. *Peerj Computer Science*,

7, 1–27. Https://Doi.Org/10.7717/Peerj-Cs.624

Kasome, R. S. (2020). Hubungan Keakraban Orangtua Dan Bullying Dengan Depresi Pada Remaja Kota Bukittinggi. *Human Care Journal*, *5*(3), 798. Https://Doi.Org/10.32883/Hcj.V5i3.840

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). Lightgbm: A Highly Efficient Gradient Boosting Decision Tree. *Advances In Neural Information Processing Systems*, *2017-Decem*(Nips), 3147–3155.

Kementerian Kesehatan Ri. (2024). Sehat Negeriku. Https://Sehatnegeriku.Kemkes.Go.Id/Baca/Umum/20240917/2446492/Cegah-Bunuh-Diri-Kemenkes-Ajak-Remaja-Bicara-Soal-Kesehatan-Mental/

Kesuma, F. I. M., & Kalifia, A. D. (2024). Pengaruh Media Sosial Terhadap Tingkat Anxiety Pada Remaja : Sebuah Analisis Dengan Rapidminer. *Jurnal Ilmiah Multidisiplin*, *1*(3), 177–181. Https://Doi.Org/10.62017/Merdeka

Khadim, A. (2018). An Evaluation Of Preprocessing Techniques For Text Classification. *International Journal Of Computer Science And Information Security*, *16*(6), 22–32. Https://Www.Researchgate.Net/Publication/329339664_An_Evaluation_Of_Preprocessing_Techniques_For_Text_Classification

Khan, A. S., Ahmad, H., Asghar, M. Z., Saddozai, F. K., Arif, A., & Khalid, H. A. (2020). Personality Classification From Online Text Using Machine Learning Approach. *International Journal Of Advanced Computer Science And Applications*, *11*(3), 460–476. Https://Doi.Org/10.14569/Ijacsa.2020.0110358

Maldini, N., Utomo, D. W., & Tresyani, R. P. (2025). Deteksi Dini Gangguan Kesehatan Mental Berdasarkan Sentimen Menggunakan Metode Stacking. *Sistemasi: Jurnal Sistem Informasi*, *14*(1), 271–280.

Markoulidakis, I., Rallis, I., Georgoulas, I., Kopsiaftis, G., Doulamis, A., & Doulamis, N. (2021). Multiclass Confusion Matrix Reduction Method And Its Application On Net Promoter Score Classification Problem. *Technologies*, *9*(81). Https://Doi.Org/10.3390/Technologies9040081

Mchugh, C. M., Corderoy, A., Ryan, C. J., Hickie, I. B., & Large, M. M. (2019). Association Between Suicidal Ideation And Suicide: Meta-Analyses Of Odds Ratios, Sensitivity, Specificity And Positive Predictive Value. *Bjpsych Open*, *5*(2), 1–12. Https://Doi.Org/10.1192/Bjo.2018.88

Mcmillan, S. S., Stewart, V., Wheeler, A. J., Kelly, F., & Stapleton, H. (2020). Medication Management In The Context Of Mental Illness: An Exploratory Study Of Young People Living In Australia. *Bmc Public Health*, *20*(1), 1–13. Https://Doi.Org/10.1186/S12889-020-09237-9

Meilong, S., He, P., Xiao, H., Li, H., & Zeng, C. (2020). An Approach To Semantic And Structural Features Learning For Software Defect Prediction. *Mathematical Problems In Engineering*, *2020*. Https://Doi.Org/10.1155/2020/6038619

Muhammad, Y. I., Salim, N., & Zainal, A. (2024). Joint Intent Detection And Slot Filling With Syntactic And Semantic Features Using Multichannel Cnn-Bilstm. *Peerj Computer Science*, *10*. Https://Doi.Org/10.7717/Peerj-Cs.2346

Naeem, S., Ali, A., Anam, S., & Ahmed, M. M. (2023). An Unsupervised Machine Learning Algorithms: Comprehensive Review. *International Journal Of Computing And Digital Systems*, *13*(1), 911–921. Https://Doi.Org/10.12785/Ijcds/130172

Naidu, P. B., Ruchitha, M., Yaswanth, P., Harika, B., Prabhu, P., & Deepthi Sree, G. V. (2024). Mental Health Detection Using Machine Learning. *International Journal Of Innovative Science And Research Technology (Ijisrt)*, *9*(4), 760–766. Https://Doi.Org/10.38124/Ijisrt/Ijisrt24apr701

Novakovic, J. D., Veljovi, A., Siniša, S. I., Papic, Z., & Tomovic, M. (2017). Evaluation Of Classification Models In Machine Learning. *Theory And Applications Of Mathematics & Computer Science*, *7*(1), 39–46.

Nti, I. K., Nyarko-Boateng, O., Aning, J., Fosu, G. K., Pokuaa, H. A., & Kyeremeh, F. (2022). Early Detection Of Stroke For Ensuring Health And Well-Being Based On Categorical Gradient Boosting Machine. *Journal Of Ict Research And Applications*, *16*(3), 313–332. Https://Doi.Org/10.5614/Itbj.Ict.Res.Appl.2022.16.3.8

Nugraha, W., & Sasongko, A. (2022). Hyperparameter Tuning Pada Algoritma Klasifikasi Dengan Grid Search. In *Sistemasi: Jurnal Sistem Informasi* (Vol. 11, Issue 2, Pp. 391–401). Http://Sistemasi.Ftik.Unisi.Ac.Id

Nugroho, A. B., Al Asri, H. B., & Pramesti, A. A. (2022). Survei Kesadaran Mental Mahasiswa Upn Veteran Yogyakarta Di Era Digital Dan Covid-19. *Jurnal Kesehatan Masyarakat (Undip)*, *10*(1), 38–42. Https://Doi.Org/10.14710/Jkm.V10i1.31611

Pirizadeh, M., Alemohammad, N., Manthouri, M., & Pirizadeh, M. (2021). A New Machine Learning Ensemble Model For Class Imbalance Problem Of Screening Enhanced Oil Recovery Methods. *Journal Of Petroleum Science And Engineering*, *198*(December 2020), 108214. Https://Doi.Org/10.1016/J.Petrol.2020.108214

Porter, M. F. (1980). An Algorithm For Suffix Stripping. *Program*, *14*(3), 130–137. Https://Doi.Org/10.1108/Eb046814

Pratama, C. H., & Findawati, Y. (2024). Klasifikasi Hate Speech Dan Emosi Dalam Teks Berbahasa Indonesia Pada Pengguna Twitter Menggunakan Metode Naïve Bayes Classifier. *Indonesian Journal Of Applied Technology*, *1*(3), 10. Https://Doi.Org/10.47134/Ijat.V1i3.3105

Pratiwi, T. S. (2024). Kesehatan Mental Perempuan Dan Sinergitas Kebijakan Kesehatan Mental Di Indonesia Terhadap Norma Internasional. *Jurnal Ilmiah Hubungan Internasional*, *1*(1), 66–86. Https://Doi.Org/10.26593/Jihi.V1i1.7814.66-86

Prawita, E., & Haerunnisa, A. (2024). Tipe- Tipe Gangguan Kepribadian Pada

Psikologi Abnormal. *Jurnal Ilmiah Penenlitian Multi Disiplin Ilmu*, *2*, 10–15. Https://Journal.Uinmataram.Ac.Id/Index.Php/Altazkiah/Article/View/107

Putri, K., & Raharja, A. M. (2023). Implementasi Algoritma Support Vector Machine Dalam Klasifikasi Deteksi Depresi Dari Postingan Pada Media Sosial. *Jurnal Nasional Teknologi Informasi Dan Aplikasinya*, *2*(1), 193–202.

Rahayu, K., Fitria, V., Septhya, D., Rahmaddeni, & Efrizoni, L. (2023). Klasifikasi Teks Untuk Mendeteksi Depresi Dan Kecemasan Pada Pengguna Twitter Berbasis Machine Learning. *Institut Riset Dan Publikasi Indonesia (Irpi) Malcom: Indonesian Journal Of Machine Learning And Computer Science*, *3*(2), 108–114. Https://Journal.Irpi.Or.Id/Index.Php/Malcom/Article/View/780

Raschka, S. (2018). *Model Evaluation, Model Selection, And Algorithm Selection In Machine Learning*. Http://Arxiv.Org/Abs/1811.12808

Rizki, A. N., Padilah, F. I., Rahma, R., & Nurfalah, S. T. (2023). Pengaruh Kaderisasi Terhadap Kesehatan Mental Mahasiswa Baru Fakultas Pendidikan Olahraga Dan Kesehatan Universitas Pendidikan Indonesia. *Jurnal Kesehatan Masyarakat Indonesia (Jkmi)*, *1*(2), 29–33. Https://Doi.Org/10.62017/Jkmi.V1i2.514

Sarkar, S. (2024). *Sentiment Analysis For Mental Health*. Kaggle. Https://Www.Kaggle.Com/Datasets/Suchintikasarkar/Sentiment-Analysis-For-Mental-Health

Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications And Research Directions. *Sn Computer Science*, *2*(3), 1–21. Https://Doi.Org/10.1007/S42979-021-00592-X

Sheikhi, S., & Kostakos, P. (2024). Safeguarding Cyberspace: Enhancing Malicious Website Detection With Pso[Sbnd]Optimized Xgboost And Firefly-Based Feature Selection. *Computers And Security*, *142*(February), 103885. Https://Doi.Org/10.1016/J.Cose.2024.103885

Shi, R., Xu, X., Li, J., & Li, Y. (2021). Prediction And Analysis Of Train Arrival Delay Based On Xgboost And Bayesian Optimization. *Applied Soft Computing*, *109*, 107538. Https://Doi.Org/10.1016/J.Asoc.2021.107538

Singh, A. (2024). *The Early Detection And Diagnosis Of Mental Health Status Employing Nlp-Based Methods With Ml Classifiers*. *17*(17), 56–66. Https://Doi.Org/10.37648/Ijrmst.V17i01.009

Song, F., & Croft, W. B. (1999). General Language Model For Information Retrieval. *International Conference On Information And Knowledge Management, Proceedings*, 316–321. Https://Doi.Org/10.1145/319950.320022

Suliati, S., Achmadi, S., & Rudhistiar, D. (2022). Penerapan Sistem Pakar Untuk Deteksi Dini Mental Illness Dengan Menggunakan Metode Forwad Chaining Dan Certainty Factor Berbasis Website. *Jati (Jurnal Mahasiswa Teknik Informatika)*, *6*(2), 1087–1095. Https://Doi.Org/10.36040/Jati.V6i2.5464

Suthaharan, S. (2015). Machine Learning Models And Algorithms For Big Data Classification: Thinking With Examples For Effective Learning. In *Springer New York, Ny*.

Syarif, I., Prugel-Bennett, A., & Wills, G. (2016). Svm Parameter Optimization Using Grid Search And Genetic Algorithm To Improve Classification Performance. *Telkomnika (Telecommunication Computing Electronics And Control)*, *14*(4), 1502. Https://Doi.Org/10.12928/Telkomnika.V14i4.3956

Syisnawati, Musdalifah, & Maulinda. (2023). Analisis Faktor Yang Berhubungan Dengan Kekambuhan Pasien Gangguan Jiwa. *Jurnal Keperawatan Jiwa*, *11*(1), 19–26. Https://Jurnal.Unimus.Ac.Id/Index.Php/Jkj/Article/View/10946

Tan, K. L., Lee, C. P., & Lim, K. M. (2023). Roberta-Gru: A Hybrid Deep Learning Model For Enhanced Sentiment Analysis. *Applied Sciences (Switzerland)*, *13*(6). Https://Doi.Org/10.3390/App13063915

Taufik, I. (2017). *Analisis Sentimen Terhadap Tokoh Publik Menggunakan Algoritma Support Vector Machine (Svm)*.

Tran, A. T., Desser, D., Zeevi, T., Abou Karam, G., Zietz, J., Dell'orco, A., Chen, M. C., Malhotra, A., Qureshi, A. I., Murthy, S. B., Majidi, S., Falcone, G. J., Sheth, K. N., Nawabi, J., & Payabvash, S. (2025). Optimizing Automated Hematoma Expansion Classification From Baseline And Follow-Up Head Computed Tomography. *Applied Sciences (Switzerland)*, *15*(1), 111. Https://Doi.Org/10.3390/App15010111

Ünalan, S., Günay, O., Akkurt, I., Gunoglu, K., & Tekin, H. O. (2024). A Comparative Study On Breast Cancer Classification With Stratified Shuffle Split And K-Fold Cross Validation Via Ensembled Machine Learning. *Journal Of Radiation Research And Applied Sciences*, *17*(4), 101080. Https://Doi.Org/10.1016/J.Jrras.2024.101080

Vargas, V. W. De, Aranda, J. A. S., Costa, R. Dos S., Pereira, P. R. Da S., & Barbosa, J. L. V. (2023). Imbalanced Data Preprocessing Techniques For Machine Learning: A Systematic Mapping Study. *Knowledge And Information Systems*, *65*(1), 31–57. Https://Doi.Org/10.1007/S10115-022-01772-8

Vitoasmara, K., Vio Hidayah, F., Yuna Aprillia, R., & Dyah Dewi, L. A. (2024). Gangguan Mental (Mental Disorders). *Student Research Journal*, *2*, 57–68. Https://Doi.Org/10.55606/Srjyappi.V2i3.1219

Wibawa, P. W. A., & Pramartha, C. (2023). Systematic Literature Review: Machine Learning Methods In Emotion Classification In Textual Data. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, *12*(3), 425–433. Https://Doi.Org/10.32736/Sisfokom.V12i3.1787

Wongvorachan, T., He, S., & Bulut, O. (2023). A Comparison Of Undersampling, Oversampling, And Smote Methods For Dealing With Imbalanced Classification In Educational Data Mining. *Information*, *14*(54). Https://Doi.Org/10.3390/Info14010054

World Health Organization. (2013). Mental Health Action Plan 2013-2020. In

Journal Of Personality Assessment. Https://Doi.Org/Https://Www.Who.Int/Publications/I/Item/9789241506021

Wu, Y., Wang, L., Tao, M., Cao, H., Yuan, H., Ye, M., Chen, X., Wang, K., & Zhu, C. (2023). Changing Trends In The Global Burden Of Mental Disorders From 1990 To 2019 And Predicted Levels In 25 Years. *Epidemiology And Psychiatric Sciences*, *32*. Https://Doi.Org/10.1017/S2045796023000756

Xgboost Parameters. (N.D.). Retrieved March 30, 2025, From Https://Xgboost.Readthedocs.Io/En/Stable/Parameter.Html

Yaqin, A., & Ramadhani, G. (2023). Penilaian Kredit Menggunakan Algoritma Xgboost Dan Logistic Regression. *Jurnal Informatika: Jurnal Pengembangan It*, *8*(1), 4–10. Https://Doi.Org/10.30591/Jpit.V8i1.4337

Yulianti, S. E. H., Soesanto, O., & Sukmawaty, Y. (2022). Penerapan Metode Extreme Gradient Boosting (Xgboost) Pada Klasifikasi Nasabah Kartu Kredit. *Journal Of Mathematics: Theory And Applications*, *4*(1), 21–26. Https://Doi.Org/10.31605/Jomta.V4i1.1792

Yusrani, K. G., Aini, N., Maghfiroh, S. A., & Istanti, N. D. (2023). Tinjauan Kebijakan Kesehatan Mental Di Indonesia: Menuju Pencapaian Sustainable Development Goals Dan Universal Health Coverage. *Jurnal Medika Nusantara*, *1*(2), 89–107. Https://Doi.Org/10.59680/Medika.V1i2.281

Yuwono, D. T., Fadlil, A., & Sunardi, S. (2019). Implementasi Metode Dempster Shafer Pada Sistem Pakar Diagnosa Gangguan Kepribadian. *Jurnal Sistem Informasi Bisnis*, *9*(1), 25. Https://Doi.Org/10.21456/Vol9iss1pp25-31

Zhang, Q., Tsang, C. C. E., He, Q., & Guo, Y. (2023). Ensemble Of Kernel Extreme Learning Machine Based Elimination Optimization For Multi-Label Classification. *Knowledge-Based Systems*, *278*, 110817. Https://Doi.Org/Https://Doi.Org/10.1016/J.Knosys.2023.110817