

Pengaruh *Hyperparameter Tuning* untuk Efektivitas pada Pendekatan *Hybrid* dalam Mendiagnosis Stres dan Depresi : Tinjauan Studi Literatur

BACHTIAR RAMADHAN¹, SYAFRIAL FACHRI PANE²

^{1,2}Universitas Logistik dan Bisnis Internasional, Indonesia
Email : bachtirramadhan26@gmail.com

Received 30 November 201x | *Revised* 30 Desember 201x | *Accepted* 30 Januari 201x

ABSTRAK

Tinjauan Sistematis Literatur ini mengkaji efektivitas *hyperparameter tuning* dalam pendekatan *hybrid* untuk diagnosis kesehatan mental, menggunakan metode PRISMA untuk evaluasi model prediksi. Penelitian menunjukkan bahwa masih jarang studi yang mengintegrasikan *machine learning* dengan tuning untuk mengidentifikasi variabel krusial dalam diagnosis kesehatan mental. Tujuan utama adalah menganalisis variabel penting, model yang sering digunakan, dan metode tuning terbaik. Hasil menunjukkan bahwa usia dan jenis kelamin adalah variabel kunci, dengan *Random Forest* dan *Tree-Structured Parzen Estimator* dengan *Gradient Boosting* sebagai model *tuning* terbaik, mencapai akurasi 0.986. Penelitian ini menyarankan penggunaan *genetic algorithm* untuk meningkatkan efisiensi dan mengatasi masalah *overfitting* dan *underfitting*, serta mendorong eksplorasi lebih lanjut pada kombinasi model dalam kesehatan mental.

Kata kunci: Diagnosis Kesehatan mental, *Hyperparameter Tuning*, *Hybrid*

ABSTRACT

Systematic Review This literature review examines the effectiveness of hyperparameter tuning in a hybrid approach to mental health diagnosis, using the PRISMA method for prediction model evaluation. The research shows that there are few studies that integrate machine learning with tuning to identify crucial variables in mental health diagnosis. The main objective was to analyze important variables, frequently used models, and the best tuning methods. Results show that age and gender are key variables, with Random Forest and Tree-Structured Parzen Estimator with Gradient Boosting as the best tuning models, achieving an accuracy of 0.986. This study suggests the use of genetic algorithms to improve efficiency and overcome overfitting and underfitting problems, and encourages further exploration of model combinations in mental health.

Keywords: Mental Health Diagnosis, *Hyperparameter Tuning*, *Hybrid*

1. PENDAHULUAN

Gangguan stres dan depresi memiliki dampak serius pada individu dan masyarakat (**Wray & Sullivan, 2021**). Stres adalah respon fisiologis terhadap tekanan, sedangkan depresi ditandai oleh perasaan sedih yang persisten dan kehilangan minat dalam aktivitas sehari-hari (**Chu et al., 2019**). Kedua kondisi ini memiliki gejala tumpang tindih, membuatnya sulit dibedakan (Groen et al., 2020). Menurut WHO, depresi adalah penyebab utama disabilitas global, dan stres semakin umum di masyarakat modern. Di Asia Tenggara, 27% dari 322 juta orang mengalami depresi (**Putra et al., 2022**). Setiap tahun, 800 ribu orang meninggal akibat bunuh diri, sebagian besar dipicu oleh gangguan mental-emosional seperti depresi, dengan Indonesia berada di urutan kelima dengan angka kejadian 3.7%. Korban utama adalah remaja usia 15-29 tahun. Tanpa kesadaran dan intervensi, tingkat depresi akan terus meningkat, berdampak besar pada kesehatan psikis dan kesejahteraan masyarakat (**Levine et al., 2021**). Meskipun diagnosis diperlukan, prosesnya rumit dan panjang, melibatkan penilaian subjektif yang dapat bervariasi dan tidak selalu akurat (**Reed et al., 2022**). Beberapa tantangan dapat membatasi keefektifan proses ini, termasuk variabilitas dalam penilaian subjektif yang dibuat oleh penilai, ketidaktepatan atau ketidaklengkapan penilaian, dan pertimbangan dalam standarisasi tes depresi dan stres. Belum lagi tumpang tindih gejala antara stres dan depresi, serta kompleksitas dari gangguan mental tersebut, menyulitkan proses diagnosis yang akurat. Dengan keterbatasan ini, pada penelitian ini akan mengeksplorasi metode alternatif yang efektif untuk mendiagnosis kesehatan mental.

Dalam beberapa tahun terakhir, *machine learning* telah menjadi alat yang menjanjikan dalam mendiagnosis gangguan kesehatan mental seperti stres dan depresi. Metode ini memungkinkan analisis mendalam terhadap data klinis dan neurobiologis untuk mengidentifikasi pola yang dapat membedakan kondisi tersebut. Teknik seperti klasifikasi, clustering, dan regresi membantu memprediksi dan mendiagnosis stres dan depresi dengan akurasi tinggi (**Shankar K Uthayakumar J, 2022**). Penelitian sebelumnya menunjukkan bahwa model seperti *Support Vector Machine*, *Random Forest*, dan *Deep Learning* efektif dalam diagnosis kesehatan mental (**Abd Rahman et al., 2020; Kundu et al., 2021**). Penelitian (**Choi & Han, 2023**) menunjukkan model *XGBoost* memiliki akurasi 92% dalam mengidentifikasi gangguan kesehatan mental. Penelitian (**Shatte et al., 2019**) juga meninjau deteksi dan diagnosis gangguan mental seperti depresi menggunakan *decision tree*, menunjukkan berbagai manfaat. Namun, efektivitas pendekatan ini dipengaruhi oleh *hyperparameter tuning*, proses menentukan nilai optimal untuk parameter model *machine learning*, seperti *learning rate* dan jumlah layer pada *neural network* (**Iyortsuun et al., 2023**). *Hyperparameter tuning* sangat penting untuk meningkatkan kinerja dan akurasi model dalam mendiagnosis stres dan depresi (**Hossain & Timmer, 2021**).

Penelitian ini menyelidiki bagaimana *hyperparameter tuning* memengaruhi kinerja pendekatan *hybird* dalam diagnosis stres dan depresi. Dengan meninjau pengaruh *hyperparameter tuning* pada model seperti *XGBoost*, *Gradient Boosting*, *Support Vector Machine*, dan *Decision Tree*, penelitian ini bertujuan untuk mengembangkan alat diagnostik yang lebih akurat dan efektif. Pemahaman yang lebih baik tentang peran *hyperparameter tuning* diharapkan meningkatkan pengelolaan dan pengobatan stres dan depresi, serta membuka peluang untuk pengembangan teknologi kesehatan mental yang lebih maju. Penelitian ini terdiri dari beberapa bagian utama: Pendahuluan yang memberikan gambaran umum penelitian dan kriteria pemilihan literatur, termasuk pertanyaan penelitian; Metodologi yang memaparkan metode penelitian yang digunakan; Hasil dan Pembahasan yang mencakup *review* penelitian terkait, seleksi studi, rangkuman, dan jawaban atas pertanyaan penelitian; Kesimpulan yang merangkum temuan

penelitian; Pekerjaan Masa Depan yang membahas potensi penelitian di masa mendatang; dan Ucapan Terima Kasih yang memberikan penghargaan kepada pihak-pihak yang terlibat.

1.1 Kriteria Pemilihan Literatur

Sebelum penyusunan jurnal, pertanyaan penelitian diajukan untuk memungkinkan kesimpulan yang ringkas dan pencarian yang efisien. Pertanyaan-pertanyaan penelitian dirangkum untuk memfokuskan ruang lingkup penelitian sebagai berikut: RQ1: Apa variabel yang paling berpengaruh dalam memprediksi stres dan depresi? RQ2: Model-model apa saja yang menerapkan konsep *hybrid* dan *hyperparameter tuning*? RQ3: Apa metode *hyperparameter tuning* yang paling efektif untuk pendekatan *hybrid*?

1.2 Strategi Pencarian

Penyusunan jurnal dilakukan melalui kata kunci pada beberapa basis data literatur. Kata kunci dalam string tersebut dipilih untuk menghasilkan hasil yang sesuai dengan pertanyaan penelitian. Kueri string kata kunci yang digunakan pada *software watsase* adalah sebagai berikut: ("Stres" ATAU "Depresi") AND ("Diagnosis"). Dalam penelitian ini, publikasi yang diinklusi harus memenuhi beberapa kriteria inklusi tertentu. Berikut adalah kriteria inklusi: (a) publikasi dalam bahasa Inggris, (b) artikel yang melibatkan diagnosis kondisi kesehatan mental dengan menggunakan *machine learning*, dan (c) publikasi yang dipilih dalam kurun waktu 5 tahun terakhir dari tahun 2019 sampai dengan 2020. Berikut adalah kriteria eksklusi: (a) publikasi dalam prosiding konferensi, bab buku, dan disertasi.

2. METODOLOGI PENELITIAN

Metodologi yang digunakan dalam ulasan sistematis ini adalah PRISMA, yang merupakan kerangka kerja yang diterima secara luas untuk melaporkan item dalam ulasan sistematis dan meta-analisis (Page et al., 2020). Proses PRISMA mencakup empat tahap utama: Identifikasi, Penyaringan, Kelayakan, dan Penyertaan. Tahap identifikasi melibatkan pencarian sistematis di berbagai database untuk mengumpulkan sebanyak mungkin studi relevan berdasarkan kata kunci yang ditetapkan. Pada tahap penyaringan, studi yang jelas tidak relevan dieliminasi berdasarkan judul dan abstrak. Studi yang tersisa kemudian diperiksa untuk kelayakan lebih lanjut, di mana teks lengkap dianalisis untuk memastikan mereka memenuhi kriteria inklusi. Studi yang memenuhi semua kriteria inklusi diikutsertakan dalam analisis dan sintesis lebih lanjut.

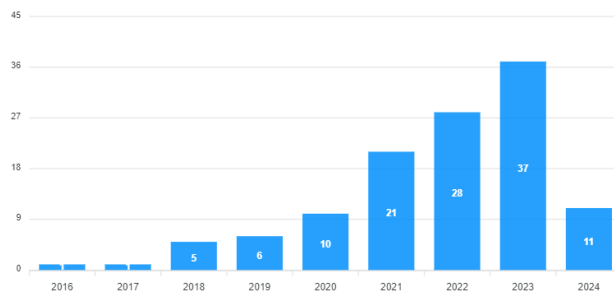
3. HASIL TINJAUAN

Pada bagian ini, disajikan hasil seleksi studi, ringkasan seluruh makalah yang terpilih, dan hasil penilaian kualitas. Temuan dan jawaban atas pertanyaan penelitian (RQ) yang telah ditetapkan dibahas di bagian Diskusi Pertanyaan Penelitian.

3.1 Pencarian dan Seleksi Studi

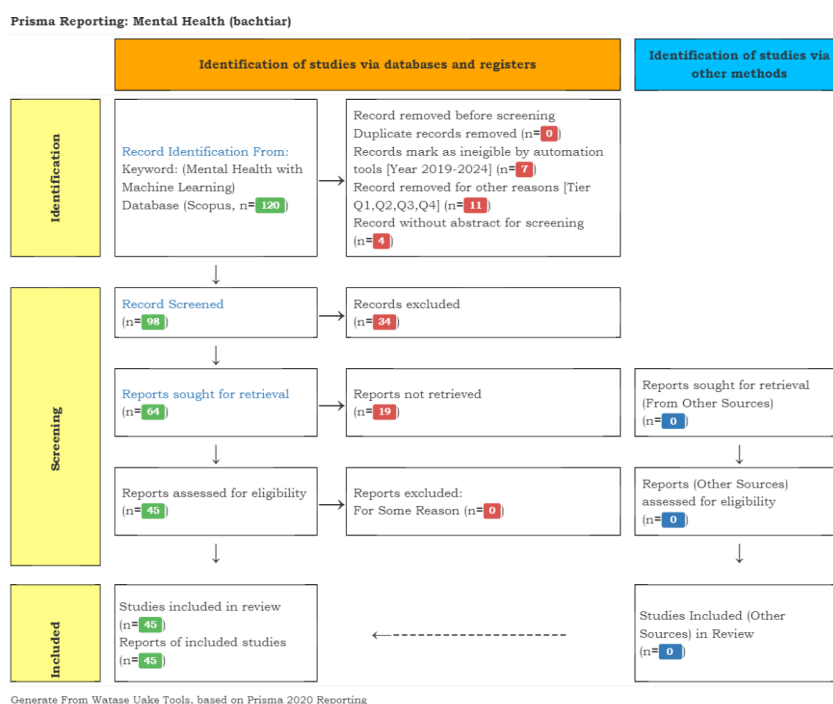
Dalam proses pencarian awal menggunakan basis data Scopus dengan kata kunci "*Mental Health with Machine Learning*", kami mengumpulkan 120 artikel. Dapat dilihat dari gambar 1, penelitian terkait topik ini terus mengalami kenaikan setiap tahunnya. Hal ini menunjukkan bahwa meningkatnya minat terhadap *machine learning* dalam isu kesehatan mental.

Pengaruh *Hyperparameter Tuning* untuk Efektivitas pada Pendekatan *Hybird* dalam Mendiagnosis Stres dan Depresi : Tinjauan Studi Literatur



Gambar 1 Sebaran Tahun Jurnal Penelitian

Setelah proses penilaian kelayakan, 45 laporan dianggap memenuhi kriteria dan diikutsertakan dalam ulasan sistematis ini. Seluruh 45 laporan tersebut dianalisis lebih dalam sesuai dengan kriteria yang telah kami tentukan.

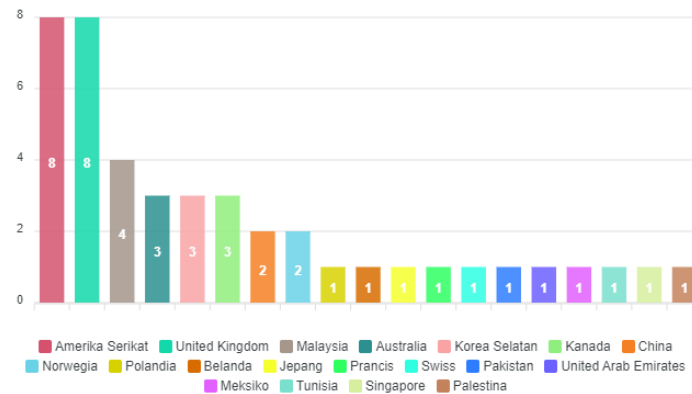


Gambar 2 Diagram PRISMA dalam Memilih Studi Literatur

Diagram alir PRISMA dari tahap awal pencarian hingga jumlah akhir artikel yang dipilih ditampilkan pada Gambar 2, mengikuti standar PRISMA. Analisis 45 artikel yang dipilih didasarkan pada kriteria inklusi dan eksklusi. Ringkasan artikel mencakup sumber data, kata kunci, durasi, lokasi geografis, metode analisis data, tujuan studi, metode ekstraksi fitur, teknik machine learning, dan kinerja pengklasifikasi.

3.2 Deskripsi Studi yang Dipilih

Deskripsi studi yang dipilih pada penelitian ini berfokus pada depresi dan stres. Depresi adalah gangguan kesehatan mental yang menyebabkan perasaan sedih mendalam, kehilangan minat, gangguan tidur, perubahan nafsu makan, kelelahan, perasaan bersalah, dan pikiran tentang kematian atau bunuh diri (**Paiva et al., 2023**). Stres adalah respons tubuh terhadap situasi menantang atau mengancam, melibatkan reaksi fisiologis dan psikologis kompleks. Stres dapat menyebabkan gejala fisik seperti sakit kepala dan ketegangan otot, serta gejala psikologis seperti kecemasan, iritabilitas, dan kesulitan berkonsentrasi (**McEwen & Karatsoreos, 2020**).



Gambar 3 Sebaran Geografis Penelitian di Antar Negara

Studi terkait kesehatan mental dan machine learning tersebar di 20 negara. United Kingdom dan Amerika Serikat memimpin dengan 8 dan 7 artikel masing-masing, diikuti oleh Malaysia dengan 4 artikel, dan beberapa negara lain dengan jumlah lebih sedikit. Ini menunjukkan distribusi global penelitian, dengan UK dan AS sebagai negara dengan penelitian terbanyak. Tidak ditemukan kelompok penelitian terkemuka yang melakukan penelitian berkelanjutan di area ini.

Tabel 1 Survey Jurnal Pemodelan Mental Health

Ref	Tahun	Feature Selection	Metode	Model	Hyperparameter Tuning	F1-S	Rec	Prec	Acc	AUC
(Shatte et al., 2019)	2019	RFE	ML	SVM	Grid Search	-	-	-	-	-
(Dadi et al., 2021)	2021	-	ML	RF	Grid Search	-	-	-	-	-
(Nash et al., 2023)	2021	PCA	ML	SVM	-	-	0.97	0.93	0.95	-
(Kundu et al., 2021)	2021	Lasso Ridge	ML	RF	-	-	-	-	-	-
(Hornstein et al., 2021)	2021	-	ML	RF	Grid Search	-	-	-	0.71	0.60
(Lakshmi et al., 2019)	2021	-	ML	DT	-	-	-	-	-	-
(Zhang et al., 2022)	2022	Genne Fuzzy	ML	RF	Genne Fuzzy Scoring	-	-	-	-	0.93
(Garriga et al., 2022)	2022	SHAP	ML	XGBoost	Bayesian Optimizatin	-	-	-	-	0.79
(Rezapour & Hansen, 2022)	2022	RFE	ML	SMOTE-RF	-	-	-	-	0.96	-
(Chung & Teo, 2022)	2022	-	ML	RF	-	-	-	-	0.88	0.94
(Internation al., 2024)	2022	-	DL	ANN	-	0.91	0.95	0.89	0.96	-
(Yeung et al., 2023)	2022	Correlation-Base	DL	CNN	-	-	-	-	0.86	-
(Haghighi et al., 2023)	2023	SHAP	ML	Gradient Boosting	-	-	-	-	-	0.76
(Rothenberg et al., 2023)	2023	-	ML	RF	-	-	-	-	-	-
(Du et al., 2023)	2023	-	ML	XGBoost	-	-	-	-	0.91	-
(Janczewski & Nitkowski, 2023)	2023	-	ML	RF	-	0.96	0.85	0.94	0.95	-
(Kuhathasan et al., 2023)	2023	SHAP	ML	XGBoost	Grid Search	-	-	-	-	-
(Villa-Pérez et al., 2023)	2023	-	DL	CNN	-	0.87	0.87	0.87	0.88	0.95
(Ismail et al., 2023)	2023	TF-IDF	ML	RF	-	0.96	0.96	0.96	0.96	0.97
(Liu et al., 2023)	2023	-	DL	NN	-	-	0.75	0.82	0.78	-

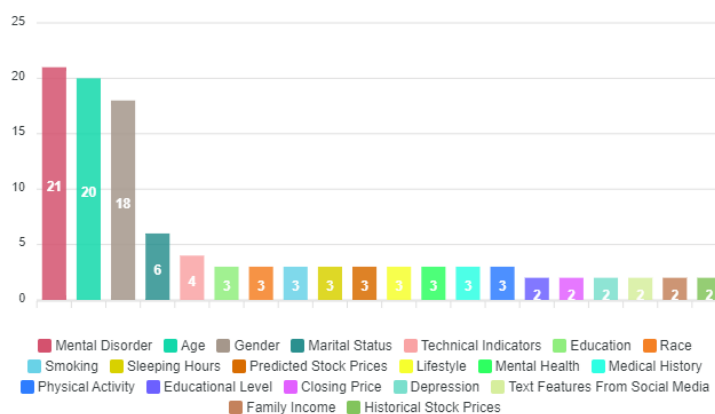
Pengaruh *Hyperparameter Tuning* untuk Efektivitas pada Pendekatan *Hybird* dalam Mendiagnosis Stres dan Depresi : Tinjauan Studi Literatur

(Choi & Han, 2023)	2023	Lasso Ridge	ML	XGBoost	Random Search	-	-	-	0.92	-
(Bjerregaard, 2023)	2023	SHAP	ML	XGBoost	-	-	-	-	-	0.70
(Chung & Teo, 2023)	2023	ETR	ML	Gradient Boosting	-	0.79	0.87	0.82	0.84	-
(Chen, 2023)	2023	PI	ML	RF	-	-	-	-	-	-
(Yu et al., 2021)	2023	SHAP	ML	Gradient Boosting	-	-	-	-	0.60	0.70
(Baba & Bunji, 2023)	2023	SHAP	ML	Gradient Boosting	TPE	0.99	0.98	0.99	0.98	0.85
(Van Mens et al., 2023)	2023	Lasso Ridge	ML	Lasso Regression	-	-	-	-	-	0.80
(Ueda et al., 2023)	2023	-	ML	LSS	-	-	-	-	0.71	-
(Abdul Rahman et al., 2023)	2023	-	ML	SDCA	-	-	-	-	-	-
(Darko et al., 2024)	2024	FI	ML	RF	-	0.84	0.78	0.91	0.68	-
(Khoo et al., 2024)	2024	-	DL	NN	-	-	-	-	-	-
(Radwan et al., 2024)	2024	-	ML	SVM	-	0.84	0.83	0.84	0.86	-

Tabel 1 merangkum penelitian terkait kesehatan mental menggunakan *machine learning* dan *deep learning*. Tabel ini mengkategorikan studi berdasarkan referensi, tahun publikasi, metode seleksi fitur, metode analisis, model *machine learning*, teknik *hyperparameter tuning*, dan metrik kinerja seperti *F1-Score*, *Recall*, *Precision*, Akurasi, dan AUC. Secara keseluruhan, Tabel 1 memberikan gambaran tentang kemajuan terkini dalam pemodelan kesehatan mental, menunjukkan berbagai pendekatan dan hasil dalam diagnosa dan prediksi kondisi kesehatan mental.

3.3 Data Set

Untuk memahami faktor-faktor yang berkontribusi pada stres dan depresi, penting mengidentifikasi variabel-variabel paling berpengaruh (Guan et al., 2019; Shukri et al., 2023; Viertiö et al., 2020). Tantangan utama adalah keterbatasan akses ke dataset berkualitas tinggi dan representatif, karena data kesehatan mental sangat sensitif dan sering diproses untuk anonimitas, yang dapat mengurangi integritas data (Keerie et al., 2023; Nash et al., 2023). Teknik pra-pemrosesan diperlukan untuk meminimalkan distorsi data. Identifikasi variabel kunci memperkuat dasar ilmiah untuk intervensi kesehatan mental yang efektif (García et al., 2021; Rickwood et al., 2019). Peneliti menemukan dataset dengan variabel unik yang memungkinkan analisis komprehensif, membantu merumuskan pola prediktif signifikan untuk stres dan depresi, yang ditampilkan dalam gambar 2 (Abouelmehdi et al., 2021).



Gambar 4 Variabel yang Sering Muncul

Gambar 4 menunjukkan distribusi variabel dalam studi kesehatan mental. "Mental Disorder" dominan dalam 21 studi, menunjukkan pengaruh gangguan mental terhadap berbagai aspek kehidupan. Variabel seperti "Age" dan "Gender" (20 dan 18 studi) menekankan pentingnya faktor demografis, sementara "Marital Status" dan "Smoking" (6 dan 4 studi) menunjukkan dampak perilaku dan sosial. Gambar ini menunjukkan pentingnya variabel tersebut dalam penelitian kesehatan mental. Identifikasi variabel signifikan membantu penyaringan data dan memungkinkan perancangan model prediktif yang akurat (**Nejad & Varathan, 2020**). Variabel *age*, *gender*, *education*, dan *marital status* terbukti berkorelasi erat dengan stres dan depresi. Menggabungkan variabel ini dalam model analisis memperkuat pemahaman risiko dan mengembangkan strategi intervensi yang lebih efektif.

3.4 Feature Selection

Proses *feature selection* merupakan langkah penting dalam pengembangan model prediktif yang akurat untuk memahami faktor-faktor yang mempengaruhi stres dan depresi (**Haghish et al., 2023; Lee & Kim, 2021; Takahashi et al., 2020**). Dalam penelitian ini, *feature selection* mengidentifikasi variabel signifikan. Untuk menjawab pertanyaan penelitian mengenai variabel yang berpengaruh dalam diagnosis stres dan depresi, analisis jurnal ini menggunakan *feature selection* yang ditampilkan pada tabel 1. Metode efektif yang sering digunakan adalah *Shapley Additive Explanations* (SHAP), yang menilai pentingnya setiap fitur dalam prediksi model menggunakan pendekatan berbasis teori permainan.

a. SHapley Additive exPlanations (SHAP)

SHAP (*SHapley Additive exPlanations*) adalah metode interpretasi model prediktif yang berasal dari teori permainan, menggunakan nilai Shapley untuk mengukur kontribusi marginal fitur terhadap prediksi model (**Li, 2022; Nordin et al., 2023; Parsa et al., 2020**). Berikut adalah rumus yang digunakan untuk menghitung nilai SHAP, adalah sebagai berikut.

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} [v(S \cup \{i\}) - v(S)] \quad (1)$$

Rumus (1) menghitung sumbangan rata-rata fitur *i* terhadap output model, mempertimbangkan semua kombinasi fitur lain, memberikan pandangan komprehensif. Dalam tabel 1, SHAP menunjukkan fitur yang paling berpengaruh dan cara pengaruhnya dalam konteks model yang lebih besar (**Haghish et al., 2023**). SHAP telah diterapkan dalam monitoring kesehatan mental jarak jauh (**Garriga et al., 2022; Kuhathasan et al., 2023**) dan prediksi stres (**Baba & Bunji, 2023; Garriga et al., 2022; Yu et al., 2021**).

3.5 Machine Learning dan Hyperparameter Tuning

Peningkatan akurasi dalam mendiagnosis stres dan depresi membutuhkan pemahaman terhadap pola dan fitur klinis (**Lawrance et al., 2022; Tao et al., 2021**). Metode *machine learning*, menggunakan teknik klasifikasi, membantu mendeteksi pola yang sulit terlihat secara manual (**Belal A. Hamed Osman Ali Sadek Ibrahim, 2023; Wani et al., 2021**). Berbagai model dicatat dalam tabel 1. *Random forest* adalah yang paling banyak digunakan (12 kali), sementara beberapa model hanya digunakan sekali. Namun, efektivitas model-model ini bergantung pada *hyperparameter tuning* (**Wu et al., 2020**). Dari tabel 1, model terbaik yang menerapkan konsep *hybrid* adalah Gradient Boosting, Random Forest, dan XGBoost, dengan Random Forest paling sering digunakan (4 penelitian).

a. Gradient Boosting

Gradient Boosting adalah teknik *ensemble* yang menggabungkan beberapa model prediktif sederhana, seperti pohon keputusan, untuk membentuk model prediktif yang kuat (**Chung & Teo, 2023; Sahin, 2020**). Rumus utama dari *Gradient Boosting* adalah.

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad (2)$$

Dimana pada rumus (2), $F_m(x)$ adalah model pada iterasi m , γ_m adalah parameter pembelajaran, dan $h_m(x)$ adalah model dasar (biasanya pohon keputusan).

b. *XGBoost*

XGBoost (Extreme Gradient Boosting) adalah versi yang lebih efisien dan dioptimalkan dari algoritma *Gradient Boosting* (**Kavzoglu & Teke, 2022; Kuhathasan et al., 2023**). Rumus utama *XGBoost* adalah.

$$L(\theta) = \sum_{i=1}^n l(y_i y'_i^{(t)}) + \sum_{k=1}^K \Omega(f_k) \quad (3)$$

Dimana pada rumus (3) l adalah fungsi *loss*, $y'_i^{(t)}$ adalah prediksi pada iterasi t , dan Ω adalah penalti kompleksitas model.

c. *Random Forest*

Random Forest adalah teknik ensemble yang membangun beberapa pohon keputusan dan menggabungkan prediksi mereka untuk meningkatkan akurasi dan stabilitas model (**Chen, 2023; Linhui et al., 2020**). Rumus utama untuk prediksi dengan *Random Forest* adalah.

$$y' = \frac{1}{M} \sum_{m=1}^M h_m(x) \quad (4)$$

Dimana pada rumus (4), y' adalah prediksi akhir, M adalah jumlah pohon dalam hutan, dan $h_m(x)$ adalah prediksi dari pohon m . *Hyperparameter tuning* mengatur nilai-nilai parameter pada model yang tidak dipelajari selama pelatihan tetapi menentukan cara model belajar dan melakukan prediksi (**Dias et al., 2020; Mikail et al., 2021**). Dari hasil kajian literatur, berbagai metode tuning telah diterapkan, seperti yang terlihat pada tabel 1, beberapa model yang sering digunakan dalam diagnosis kesehatan mental, yaitu.

a. *Tree-Structured Parzen Estimator*

TPE adalah sebuah metode untuk optimisasi *hyperparameter*, khususnya sebagai bagian dari optimisasi berbasis *Bayesian* (**Baba & Bunji, 2023**). TPE bertujuan untuk menemukan set *hyperparameter* yang menghasilkan performa model terbaik dengan cara yang lebih efisien dibandingkan dengan metode *grid search* atau *random search* (**Alibrahim & Ludwig, 2021**). TPE mendefinisikan dua model distribusi.

$$l(x) = P(x|y < y^*) : \text{distribusi parameter untuk } loss \text{ rendah} \quad (5)$$

$$l(x) = P(x|y \geq y^*) : \text{distribusi parameter untuk } loss \text{ tinggi} \quad (6)$$

Di sini, y^* adalah nilai ambang batas untuk *loss* yang memisahkan parameter dengan performa baik dan buruk. Pemilihan parameter baru dilakukan dengan memaksimalkan rasio *likelihood* antara $l(x)$ dan $g(x)$, yang dirumuskan sebagai.

$$\arg \max_x \frac{l(x)}{g(x)} \quad (7)$$

b. *Grid Search*

Metode ini melibatkan pemilihan nilai *hyperparameter* dari suatu kumpulan nilai yang telah ditentukan sebelumnya (**Belete & Huchaiah, 2022**). *Grid Search* mencoba semua kombinasi yang mungkin dari nilai-nilai *hyperparameter* yang telah

didefinisikan, dan kemudian memilih kombinasi yang memberikan performa terbaik **(Dadi et al., 2021; Hornstein et al., 2021)**.

c. *Random Search*

Metode ini memilih nilai *hyperparameter* secara acak dari kumpulan nilai yang telah ditentukan sebelumnya **(Andonie & Florea, 2020)**. *Random Search* lebih efisien dalam beberapa kasus karena tidak perlu mencoba semua kombinasi nilai *hyperparameter* **(Shams et al., 2024)**.

d. *Bayesian Optimization*

Bayesian Optimization adalah metode untuk mencari nilai optimal dari fungsi objektif yang mahal, menggunakan model probabilistik seperti *Gaussian Process* untuk memperkirakan fungsi objektif dan menentukan nilai-nilai menjanjikan **(Dadi et al., 2021)**. Metode ini diterapkan luas dalam tuning parameter algoritma machine learning, pengoptimalan *hyperparameter*, dan desain eksperimen **(Lei et al., 2021)**.

$$x^* = \arg \min_{x \in X} f(x) \tag{8}$$

Dimana x^* adalah nilai optimal yang dicari, $f(x)$ adalah fungsi objektif dan X adalah ruang pencarian *hyperparameter*.

Dari beberapa model *tuning* di atas, *Grid search* adalah metode *hyperparameter tuning* favorit dalam diagnosis kesehatan mental, digunakan dalam 4 artikel. Penggabungan model *machine learning* dengan *hyperparameter tuning* menghasilkan model yang lebih kuat untuk mendeteksi pola klinis dalam diagnosis stres dan depresi. Model *hybrid* dengan *hyperparameter tuning* secara signifikan meningkatkan kinerja diagnostik dibandingkan model tunggal, menghasilkan sensitivitas dan spesifisitas lebih tinggi.

3.5 Evaluasi Metrik

Efektivitas diagnosis kesehatan mental dinilai dengan memeriksa nilai keakuratan dalam model prediksi. Lima faktor perhitungan yang digunakan dalam menghitung nilai keakuratan yaitu, *precision*, *recall*, *f1-score*, akurasi, dan nilai AUC. Rentang nilai evaluasi metrik dari 0 sampai dengan 1, semakin mendekati 1 maka nilai evaluasi metrik semakin baik **(Ismail et al., 2023; Powers, 2020; Villa-Pérez et al., 2023)**. Spesifiknya tercantum di bawah ini.

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)} \tag{9} \quad Recall = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FN)} \tag{10}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{11} \quad Accuracy = \frac{True\ Positive\ (TP) + True\ Negative\ (TN)}{Total\ Instances} \tag{12}$$

$$AUC = \sum_{i=1}^{n-1} ((x_{i+1} - x_i) \times \frac{y_i + y_{i+1}}{2}) \tag{13}$$

4. DISKUSI PERTANYAAN PENELITIAN

4.1 RQ1 : Apa Variabel yang Paling Berpengaruh dalam Memprediksi Stres dan Depresi ?

Tinjauan literatur kami menunjukkan bahwa variabel seperti usia, jenis kelamin, dan status sosial-ekonomi memiliki korelasi signifikan dengan prevalensi dan keparahan stres serta depresi. Temuan ini konsisten dengan penelitian yang menggarisbawahi pengaruh faktor demografis ini terhadap kesehatan mental **(Garriga et al., 2022)** Tinjauan literatur kami menunjukkan bahwa variabel seperti usia, jenis kelamin, dan status sosial-ekonomi memiliki

korelasi signifikan dengan prevalensi dan keparahan stres serta depresi. Temuan ini konsisten dengan penelitian yang menggarisbawahi pengaruh faktor demografis ini terhadap kesehatan mental.

4.2 RQ2 : Model-Model Apa Saja yang Menerapkan Konsep *Hybrid* dan *Hyperparameter Tuning* dalam Diagnosis Stres dan Depresi ?

Penelitian ini mengidentifikasi model machine learning yang menggunakan hybrid dan hyperparameter tuning untuk meningkatkan diagnosis stres dan depresi. Model yang sering digunakan termasuk Support Vector Machine, Gradient Boosting, XGBoost, Random Forest, dan Neural Network, dengan Random Forest paling sering dikombinasikan dengan hyperparameter tuning. Teknik ini meningkatkan akurasi dan kemampuan model dalam generalisasi data. Penggunaan Grid Search dalam Random Forest meningkatkan akurasi diagnosis, dengan kombinasi menghasilkan AUC 0.93 (**Zhang et al., 2022**), SVM dan XGBoost juga menunjukkan hasil baik saat di-tuning, dengan XGBoost mencapai akurasi 0.92 (**Choi & Han, 2023**), Hyperparameter tuning membuat model lebih *robust* dan adaptif terhadap data baru, penting untuk aplikasi klinis. Kombinasi model hybrid dan hyperparameter tuning efektif dalam meningkatkan diagnostik kesehatan mental.

4.3 RQ3 : Apa Metode *Hyperparameter Tuning* yang Paling Efektif untuk Pendekatan *Hybrid* dalam Diagnosis Stres dan Depresi ?

Penelitian terkini menunjukkan bahwa Random Search dan Tree-structured Parzen Estimator (TPE) adalah metode hyperparameter tuning yang unggul. Random Search, digunakan dalam model XGBoost, menemukan kombinasi optimal parameter dengan efisien, mencapai akurasi 0.92 (**Choi & Han, 2023**) Ini membuatnya ideal untuk dataset besar dan kompleks. TPE, bagian dari Bayesian optimization, efektif dalam memodelkan dan meminimalkan fungsi kerugian. Dalam penelitian kesehatan mahasiswa (**Baba & Bunji, 2023**), TPE dengan model Gradient Boosting menghasilkan F1-Score 0.999, recall 0.984, precision 0.99, akurasi 0.986, dan AUC 0.857, menunjukkan keakuratan diagnostik yang tinggi. Kedua metode ini meningkatkan akurasi, keandalan, dan adaptabilitas model diagnostik untuk gangguan seperti stres dan depresi. Penelitian lebih lanjut diharapkan membuka peluang baru untuk perbaikan dan personalisasi diagnosis kesehatan mental.

5. KESIMPULAN DAN SARAN

Penelitian ini meninjau aplikasi machine learning dalam diagnosis stres dan depresi, fokus pada pendekatan *hybrid* dan *hyperparameter tuning*. Variabel paling berpengaruh adalah kombinasi faktor psikologis dan biologis seperti umur, jenis kelamin, dan sosial-ekonomi. Model hybrid, seperti *Random Search* dengan *XGBoost* dan *Tree-structured Parzen Estimator* dengan *Gradient Boosting*, meningkatkan akurasi diagnosis dengan metrik evaluasi yang baik. Hyperparameter tuning memungkinkan penyesuaian optimal model, meningkatkan performa dan prediksi. Metode seperti *Bayesian Optimization*, *Grid Search*, *Random Search*, dan *Tree-structured Parzen Estimator* dibandingkan, dengan *Random Search* dan *Tree-structured Parzen Estimator* menunjukkan hasil tuning yang efisien dan efektif. Optimalisasi model *hybrid* dengan *hyperparameter tuning* terbaru harus terus dikembangkan untuk meningkatkan kinerja prediksi. Implementasi model ini dalam lingkungan klinis penting untuk mendukung diagnosis dan intervensi dini dalam sistem kesehatan mental.

Penelitian ini menunjukkan potensi model *hybrid* dalam diagnosis kesehatan mental. Untuk penelitian selanjutnya, disarankan mengeksplorasi kombinasi baru model *machine learning* dengan metode *tuning* yang lebih maju seperti *genetic algorithm* dapat meningkatkan efisiensi, efektivitas, dan membantu mengatasi masalah *overfitting* dan *underfitting*. Kemudian,

penelitian juga dapat lebih fokus pada analisis interaksi fitur, menggunakan teknik seperti *autoencoders* untuk mengidentifikasi fitur-fitur kritikal yang mempengaruhi performa model. Pengembangan dan validasi dataset yang lebih besar dan lebih representatif juga penting, untuk memastikan bahwa model yang dikembangkan memiliki adaptasi ke berbagai setting klinis.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih kepada Universitas Logistik dan Bisnis Internasional atas fasilitas dan dukungannya. Penghargaan juga diberikan kepada *software* Watase yang sangat membantu dalam pengelolaan data dan analisis.

DAFTAR RUJUKAN

- Abd Rahman, R., Omar, K., Noah, S. A. M., Danuri, M. S. N. M., & Al-Garadi, M. A. (2020). Application of machine learning methods in mental health detection: a systematic review. *Ieee Access*, *8*, 183952–183964.
- Abdul Rahman, H., Kwicklis, M., Ottom, M., Amornsriwatanakul, A., H. Abdul-Mumin, K., Rosenberg, M., & Dinov, I. D. (2023). Machine learning-based prediction of mental well-being using health behavior data from university students. *Bioengineering*, *10*(5), 575.
- Abouelmehdi, K., Beni-Hessane, A., & Khaloufi, H. (2021). Big healthcare data: preserving security and privacy. *Journal of Big Data*, *8*(1), Article 73. <https://doi.org/10.1186/s40537-021-00444-6>
- Alibrahim, H., & Ludwig, S. A. (2021). Hyperparameter optimization: Comparing genetic algorithm against grid search and bayesian optimization. *2021 IEEE Congress on Evolutionary Computation (CEC)*, 1551–1559.
- Andonie, R., & Florea, A.-C. (2020). Weighted random search for CNN hyperparameter optimization. *ArXiv Preprint ArXiv:2003.13300*.
- Baba, A., & Bunji, K. (2023). Prediction of mental health problem using annual student health survey: machine learning approach. *JMIR Mental Health*, *10*, e42420.
- Belal A. Hamed Osman Ali Sadek Ibrahim, T. A. E.-H. (2023). Optimizing classification efficiency with machine learning techniques for pattern matching. *Journal of Big Data*, *10*(124). <https://doi.org/10.1186/s40537-021-00456-1>
- Belete, D. M., & Huchaiah, M. D. (2022). Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. *International Journal of Computers and Applications*, *44*(9), 875–886.
- Bjerregaard, S. S. (2023). Exploring predictors of welfare dependency 1, 3, and 5 years after mental health-related absence in danish municipalities between 2010 and 2012 using flexible machine learning modelling. *BMC Public Health*, *23*(1), 224.
- Chen, T. (2023). Investigating the mental health of university students during the COVID-19 pandemic in a UK university: a machine learning approach using feature permutation importance. *Brain Informatics*, *10*(1), 27.
- Choi, J., & Han, S. (2023). Investigation of factors associated with mental health during the early part of the COVID-19 pandemic in South Korea based on machine learning algorithms: A cohort study. *Digital Health*, *9*, 20552076231207572.
- Chu, B., Marwaha, K., Sanvictores, T., & Ayers, D. (2019). *Physiology, Stress Reaction*. StatPearls Publishing.
- Chung, J., & Teo, J. (2022). Mental health prediction using machine learning: taxonomy, applications, and challenges. *Applied Computational Intelligence and Soft Computing*, *2022*, 1–19.

- Chung, J., & Teo, J. (2023). Single classifier vs. ensemble machine learning approaches for mental health prediction. *Brain Informatics*, *10*(1), 1.
- Dadi, K., Varoquaux, G., Houenou, J., Bzdok, D., Thirion, B., & Engemann, D. (2021). Population modeling with machine learning can enhance measures of mental health. *GigaScience*, *10*(10), giab071.
- Darko, A. P., Antwi, C. O., Adjei, K., Zhang, B., & Ren, J. (2024). Predicting determinants influencing user satisfaction with mental health app: An explainable machine learning approach based on unstructured data. *Expert Systems with Applications*, *249*, 123647.
- Dias, L. C., Lemos, C. C., & Inácio, C. M. (2020). Analyzing the Risk Factors for the Spread of Pandemics with Metamodels. *ArXiv Preprint ArXiv:2007.07588*. <https://arxiv.org/abs/2007.07588>
- Du, S., Yao, J., Shen, G. C., Lin, B., Udo, T., Hastings, J., Wang, F., Wang, F., Zhang, Z., Ye, X., & others. (2023). Social Drivers of Mental Health: A US Study Using Machine Learning. *American Journal of Preventive Medicine*, *65*(5), 827–834.
- García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., & Herrera, F. (2021). Big data preprocessing: methods and prospects. *Big Data Analytics*, *6*(1), Article 14. <https://doi.org/10.1186/s41044-021-00054-4>
- Garriga, R., Mas, J., Abraha, S., Nolan, J., Harrison, O., Tadros, G., & Matic, A. (2022). Machine learning model to predict mental health crises from electronic health records. *Nature Medicine*, *28*(6), 1240–1248.
- Groen, R. N., Ryan, O., Wigman, J. T. W., Riese, H., Penninx, B. W. J. H., Giltay, E. J., Wichers, M., & Hartman, C. A. (2020). Comorbidity between depression and anxiety: assessing the role of bridge mental states in dynamic psychological networks. *BMC Medicine*, *18*(308), 1–17.
- Guan, N., Guariglia, A., Moore, P., Xu, F., & Al-Janabi, H. (2019). Financial stress and depression in adults: A systematic review. *PLOS ONE*, *14*(2), e0214031. <https://doi.org/10.1371/journal.pone.0264041>
- Haghighi, E. F., Obaidi, M., Strømme, T., Bjørge, T., & Grønnerød, C. (2023). Mental health, well-being, and adolescent extremism: a machine learning study on risk and protective factors. *Research on Child and Adolescent Psychopathology*, *51*(11), 1699–1714.
- Hornstein, S., Forman-Hoffman, V., Nazander, A., Ranta, K., & Hilbert, K. (2021). Predicting therapy outcome in a digital mental health intervention for depression and anxiety: A machine learning approach. *Digital Health*, *7*, 20552076211060660.
- Hossain, M. R., & Timmer, D. (2021). Machine Learning Model Optimization with Hyper Parameter Tuning Approach. *Global Journal of Computer Science and Technology*, *21*(2).
- International, B. R. (2024). Retracted: Predicting Mental Health of Best Human Capital for Sustainable Organization through Psychological and Personality Health Issues: Shift from Traditional to Novel Machine Learning-Supervised Technique Approach. *BioMed Research International*, *2024*.
- Ismail, L., Shahin, N., Materwala, H., Hennebelle, A., & Frermann, L. (2023). ML-NLPEmot: Machine Learning-Natural Language Processing (NLP) Event-based Emotion Detection Proactive Framework Addressing Mental Health. *IEEE Access*.
- Iyortsuun, N. K., Kim, S.-H., Jhon, M., Yang, H.-J., & Pant, S. (2023). A review of machine learning and deep learning approaches on mental health diagnosis. *Healthcare*, *11*(3), 285.
- Janczewski, C. E., & Nitkowski, J. (2023). Predicting mental and behavioral health service utilization among child welfare-involved caregivers: A machine learning approach. *Children and Youth Services Review*, *155*, 107150.
- Kavzoglu, T., & Teke, A. (2022). Advanced hyperparameter optimization for improved spatial prediction of shallow landslides using extreme gradient boosting (XGBoost). *Bulletin of Engineering Geology and the Environment*, *81*(5), 201.

- Keerie, C., Tuck, C., Milne, G., Eldridge, S., Wright, N., & Lewis, S. C. (2023). Data sharing in clinical trials – practical guidance on anonymising trial datasets. *Trials*, 24(1), Article 56. <https://doi.org/10.1186/s13063-023-07056-6>
- Khoo, L. S., Lim, M. K., Chong, C. Y., & McNaney, R. (2024). Machine Learning for Multimodal Mental Health Detection: A Systematic Review of Passive Sensing Approaches. *Sensors*, 24(2), 348.
- Kuhathasan, N., Ballester, P. L., Minuzzi, L., MacKillop, J., & Frey, B. N. (2023). Predictors of perceived symptom change with acute cannabis use for mental health conditions in a naturalistic sample: A machine learning approach. *Comprehensive Psychiatry*, 122, 152377.
- Kundu, A., Chaiton, M., Billington, R., Grace, D., Fu, R., Logie, C., Baskerville, B., Yager, C., Mitsakakis, N., Schwartz, R., & others. (2021). Machine learning applications in mental health and substance use research among the LGBTQ2S+ population: scoping review. *JMIR Medical Informatics*, 9(11), e28962.
- Lakshmi, S., Kumar, S., Prabhu, B., & Vasudevan, V. (2019). Remote Patient Monitoring Using Radio Frequency Identification (RFID) Technology and Machine Learning for Early Detection of Suicidal Behaviour in Mental Health Facilities. *International Journal of Scientific & Technology Research*, 8(9), 3221–3225.
- Lawrance, E. L., Thompson, R., Vay, J. N. Le, Page, L., & Jennings, N. (2022). The Impact of Climate Change on Mental Health and Emotional Wellbeing: A Narrative Review of Current Evidence, and its Implications. *International Review of Psychiatry*, 34(5), 443–498. <https://doi.org/10.1080/09540261.2022.2128725>
- Lee, C., & Kim, H. (2021). Machine learning-based predictive modeling of depression in hypertensive populations. *PLOS ONE*, 16(5), e0251910. <https://doi.org/10.1371/journal.pone.0251910>
- Lei, B., Kirk, T. Q., Bhattacharya, A., Pati, D., Qian, X., Arroyave, R., & Mallick, B. K. (2021). Bayesian optimization with adaptive surrogate models for automated experimental design. *Npj Computational Materials*, 7(1), 194.
- Levine, G. N., Cohen, B. E., Commodore-Mensah, Y., Fleury, J., Huffman, J. C., Khalid, U., Labarthe, D. R., Lavretsky, H., Michos, E. D., Spatz, E. S., & Kubzansky, L. D. (2021). Psychological Health, Well-Being, and the Mind-Heart-Body Connection: A Scientific Statement From the American Heart Association. *Circulation*, 143(10), e763–e783.
- Li, Z. (2022). Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost. *Computers, Environment and Urban Systems*, 96, 101845.
- Linhui, L., Weipeng, J., & Huihui, W. (2020). Extracting the forest type from remote sensing images by random forest. *IEEE Sensors Journal*, 21(16), 17447–17454.
- Liu, Y. S., Cao, B., & Chokka, P. R. (2023). Screening for adulthood ADHD and comorbidities in a tertiary mental health center using EarlyDetect: a machine learning-based pilot study. *Journal of Attention Disorders*, 27(3), 324–331.
- McEwen, B. S., & Karatsoreos, I. N. (2020). What is stress? *Stress Challenges and Immunity in Space: From Mechanisms to Monitoring and Preventive Strategies*, 19–42.
- Mikail, A., Ozalp, Y., & Emekli, D. (2021). A Hybrid Metaheuristic Approach to Pandemic Modeling. *Informatics*, 8(2), 79. <https://doi.org/10.3390/informatics8020079>
- Nash, C., Nair, R., & Naqvi, S. M. (2023). Machine Learning in ADHD and Depression Mental Health Diagnosis: A Survey. *IEEE Access*.
- Nejad, F. Y., & Varathan, K. D. (2020). Identification of significant climatic risk factors and machine learning models in dengue outbreak prediction. *BMC Medical Informatics and Decision Making*, 20(1), Article number: 176. <https://doi.org/10.1186/s12911-020-01142-2>

- Nordin, N., Zainol, Z., Noor, M. H. M., & Chan, L. F. (2023). An explainable predictive model for suicide attempt risk using an ensemble learning and Shapley Additive Explanations (SHAP) approach. *Asian Journal of Psychiatry*, *79*, 103316.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., & others. (2020). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. <https://doi.org/10.1136/bmj.n71>
- Paiva, T. C., Soares, L., & Faria, A. L. (2023). Depression in Elderly People. *Encyclopedia*, *3*(2), 677–686.
- Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S., & Mohammadian, A. K. (2020). Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. *Accident Analysis & Prevention*, *136*, 105405.
- Powers, D. M. W. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *ArXiv Preprint ArXiv:2010.16061*.
- Putra, I. G. N. S., Sugiharto, J., Windiani, I. G. A. T., Adnyana, I. G. A. N. S., Soetjningsih, & Pratiwi, N. L. S. (2022). Risk factor associated with depressive symptoms among adolescence during COVID-19 pandemic in Denpasar Bali, Indonesia. *Bali Medical Journal*, *11*(2), 1071–1076.
- Radwan, A., Amarneh, M., Alawneh, H., Ashqar, H. I., AlSobeh, A., & Magableh, A. A. A. R. (2024). Predictive Analytics in Mental Health Leveraging LLM Embeddings and Machine Learning Models for Social Media Analysis. *International Journal of Web Services Research (IJWSR)*, *21*(1), 1–22.
- Reed, G. M., First, M. B., Billieux, J., Clotire, M., Briken, P., Achab, S., Brewin, C. R., King, D. L., Kraus, S. W., & Bryant, R. A. (2022). Emerging experience with selected new categories in the ICD-11: complex PTSD, prolonged grief disorder, gaming disorder, and compulsive sexual behaviour disorder. *World Psychiatry*, *21*(2), 165–332.
- Rezapour, M., & Hansen, L. (2022). A machine learning analysis of COVID-19 mental health data. *Scientific Reports*, *12*(1), 14965.
- Rickwood, D., Telford, N., Parker, A., Tanti, C., & McGorry, P. (2019). Prevention and early intervention in youth mental health: is it time for a multidisciplinary and trans-diagnostic model for care? *International Journal of Mental Health Systems*, *13*(1), 23.
- Rothenberg, W. A., Bizzego, A., Esposito, G., Lansford, J. E., Al-Hassan, S. M., Bacchini, D., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., & others. (2023). Predicting adolescent mental health outcomes across cultures: a machine learning approach. *Journal of Youth and Adolescence*, *52*(8), 1595–1619.
- Sahin, E. K. (2020). Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. *SN Applied Sciences*, *2*(7), 1308.
- Shams, M. Y., Elshewey, A. M., El-kenawy, E.-S. M., Ibrahim, A., Talaat, F. M., & Tarek, Z. (2024). Water quality prediction using machine learning models based on grid search method. *Multimedia Tools and Applications*, *83*(12), 35307–35334.
- Shankar K Uthayakumar J, E. M. (2022). An Augmented Artificial Intelligence Approach for Chronic Diseases Prediction. *Frontiers in Artificial Intelligence*, *3*(1). <https://doi.org/10.3389/frai.2022.1076356>
- Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: a scoping review of methods and applications. *Psychological Medicine*, *49*(9), 1426–1448.
- Shukri, M. I. M., Minhat, H. S., Ahmad, N., Ismail, F., Kanthavelu, C., Nurfarahin, D., Ghazali, W. S. W., & Zulkefli, N. A. M. (2023). Prevalence and factors associated with depression, anxiety and stress in Malaysia during COVID-19 pandemic: A systematic review. *PLOS ONE*, *18*(8), e0288618. <https://doi.org/10.1371/journal.pone.0288618>

- Takahashi, Y., Ueki, M., Yamada, M., Tamiya, G., Motoike, I. N., Saigusa, D., Sakurai, M., Nagami, F., Ogishima, S., Koshihara, S., & others. (2020). Variable selection strategies and its importance in clinical prediction modelling. *BMC Public Health*, *20*(1), 1–11.
- Tao, X., Shaik, T. B., Higgins, N., Gururajan, R., & Zhou, X. (2021). Remote patient monitoring using radio frequency identification (RFID) technology and machine learning for early detection of suicidal behaviour in mental health facilities. *Sensors*, *21*(3), 776.
- Ueda, M., Watanabe, K., & Sueki, H. (2023). Emotional distress during COVID-19 by mental health conditions and economic vulnerability: retrospective analysis of survey-linked Twitter data with a semisupervised machine learning algorithm. *Journal of Medical Internet Research*, *25*, e44965.
- Van Mens, K., Lokkerbol, J., Wijnen, B., Janssen, R., de Lange, R., & Tiemens, B. (2023). Predicting Undesired Treatment Outcomes With Machine Learning in Mental Health Care: Multisite Study. *JMIR Medical Informatics*, *11*, e44322.
- Viertiö, S., Kiviruusu, O., Piirtola, M., Kaprio, J., Korhonen, T., Marttunen, M., & Suvisaari, J. (2020). The Prevalence of Depression, Anxiety and Stress and Their Associated Factors in College Students. *International Journal of Environmental Research and Public Health*, *17*(19), 7001. <https://doi.org/10.3390/ijerph17197001>
- Villa-Pérez, M. E., Trejo, L. A., Moin, M. B., & Stroulia, E. (2023). Extracting Mental Health Indicators From English and Spanish Social Media: A Machine Learning Approach. *IEEE Access*, *11*, 128135–128152.
- Wani, J. A., Sharma, S., Muzamil, M., Ahmed, S., Sharma, S., & Singh, S. (2021). Machine Learning and Deep Learning Based Computational Techniques in Automatic Agricultural Diseases Detection: Methodologies, Applications, and Challenges. *Archives of Computational Methods in Engineering*, *29*, 641–677. <https://doi.org/10.1007/s11831-021-09588-55>
- Wray, N. R., & Sullivan, P. F. (2021). Genetic Contributions to Depression: A Review. *Nature Reviews Genetics*, *22*, 539–554. <https://doi.org/10.1038/s41576-021-00318-7>
- Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2020). Graph Neural Networks: A Review of Methods and Applications. *ArXiv Preprint ArXiv:2003.05689*.
- Yeung, H. W., Stolicyn, A., Buchanan, C. R., Tucker-Drob, E. M., Bastin, M. E., Luz, S., McIntosh, A. M., Whalley, H. C., Cox, S. R., & Smith, K. (2023). Predicting sex, age, general cognition and mental health with machine learning on brain structural connectomes. *Human Brain Mapping*, *44*(5), 1913–1933.
- Yu, J., Chiu, C., Wang, Y., Dzubur, E., Lu, W., & Hoffman, J. (2021). A machine learning approach to passively informed prediction of mental health risk in people with diabetes: retrospective case-control analysis. *Journal of Medical Internet Research*, *23*(8), e27709.
- Zhang, X., Lee, J., & Goh, W. W. Bin. (2022). An investigation of how normalisation and local modelling techniques confound machine learning performance in a mental health study. *Heliyon*, *8*(5).