

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kecemasan sosial yakni rasa cemas yang timbul secara berlebihan ketika seseorang berada dalam situasi sosial. Individu yang mengalami kondisi ini yang mengalaminya cenderung terus-menerus khawatir akan dinilai buruk oleh lingkungannya, sehingga muncul kecenderungan untuk menarik diri dari pergaulan Association (2022). Persoalan ini menjadi relevan untuk dibahas pada Generasi Z, yaitu mereka yang lahir dalam periode 1997–2012. Berbeda dari generasi sebelumnya, kelompok ini besar dan berkembang bersamaan dengan pesatnya kemajuan teknologi digital serta keterbukaan akses terhadap media sosial BPS Kabupaten Gorontalo (2025). Mereka sering disebut sebagai *digital native* karena menghabiskan sebagian besar waktu mereka untuk berinteraksi di dunia maya. Jumlah mereka diperkirakan mencapai 74,93 juta orang di Indonesia, atau sekitar 27,94% dari total populasi negara tersebut Wikipedia (2026), yang menjadikan mereka kelompok yang signifikan dalam penelitian tentang kesehatan mental.

Tingginya intensitas mengakses media sosial menjadi salah satu ciri utama gaya hidup digital Generasi Z. We Are Social (2024), mencatat bahwa rata-rata warga Indonesia menghabiskan hampir 7 jam 38 menit setiap harinya untuk bermedia sosial. Meskipun teknologi memberikan berbagai kemudahan, paparan yang berlebihan terhadap konten di media sosial, seperti pencapaian *public figure* dan rekan sebaya yang sering kali memamerkan kesuksesan, yang dapat memicu perbandingan sosial dan Fear of Missing Out (FoMO) Santoso et al. (2025). Hal ini secara keseluruhan dapat mendapatkan resiko kecemasan sosial. Penelitian menunjukkan bahwa pola hidup digital ini memiliki hubungan yang signifikan dengan tingkat kecemasan Gen Z, bahkan bisa berujung pada depresi serta keinginan untuk menyakiti diri jika tidak segera diatasi Santoso et al. (2025).

Deteksi awal terhadap kecemasan sosial sangat penting agar tindakan dapat diambil sebelum situasi menjadi lebih buruk. Namun, metode diagnostik konvensional yang mengandalkan penilaian dari tenaga ahli sering terkendala oleh

faktor biaya, waktu, dan stigma yang ada. Di titik inilah pendekatan *machine learning* menawarkan solusi alternatif, sebab mampu menganalisis serta memprediksi kondisi kesehatan mental hanya dari data perilaku digital seseorang Sahu et al. (2024). Berkat kinerja yang sangat baik dan tingkat interpretabilitas yang tinggi, algoritma seperti Naïve Bayes, KNN, dan Random Forest banyak digunakan dalam penelitian kesehatan mental Dheenathayalan (2025). Dewi et al. (2025) telah menunjukkan efektivitas *Naïve Bayes* dalam memprediksi kecemasan di kalangan mahasiswa, sedangkan Hapriyanti et al. (2021) mencatat akurasi 87,3% dengan KNN. Menurut sebuah penelitian yang dilakukan oleh Setiawan et al. (2025) model Random Forest lebih akurat dalam memprediksi tingkat depresi mahasiswa dibandingkan dengan model *Decision Tree*, Naïve Bayes, dan KNN. Namun, penelitian komparatif antara ketiga algoritma khusus untuk memprediksi kecemasan sosial pada generasi Z, berdasarkan gaya hidup digital di Indonesia masih sangat sedikit.

Berdasarkan penjelasan diatas, penelitian ini disusun untuk membandingkan tingkat performa algoritma Naïve Bayes, K-Nearest Neighbor, dan Random Forest dalam memprediksi tingkat kecemasan sosial Gen Z ditinjau dari faktor gaya hidup digital. Penelitian ini bertujuan untuk meneliti pola interaksi antara berbagai faktor gaya hidup digital dan kecemasan sosial, guna menemukan algoritma yang paling efektif. Menggunakan instrumen *Liebowtiz Social Anxiety Scale-Self Report* (LSAS-SR), diharapkan penelitian ini dapat menghasilkan model skrining dini yang efektif dan juga mengidentifikasi faktor gaya hidup digital yang paling berpengaruh terhadap kecemasan sosial generasi muda di Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana tingkat kecemasan sosial Gen Z dapat diprediksi menggunakan algoritma Naïve Bayes, KNN, dan Random Forest berdasarkan faktor gaya hidup digital?
2. Seberapa akuratkah algoritma *Naive Bayes*, *K-Nearest Neighbor* dan *Random Forest* dalam memprediksi tingkat kecemasan sosial pada Gen Z?

3. Faktor gaya hidup digital apa yang paling berpengaruh terhadap tingkat kecemasan sosial pada Gen Z?

1.3 Batasan Masalah

Agar pembahasan tetap terarah, penelitian ini dibatasi pada hal-hal berikut:

1. Subjek Penelitian: Generasi Z di Indonesia dengan rentan usia antara 13-28 tahun dan aktif dalam penggunaan media sosial.
2. Variabel Independen (Gaya Hidup Digital): Diukur melalui pola respons pada instrument kuesioner adaptasi *Liebowitz Social Anxiety Scale (LSAS)* yang mencerminkan intensitas perilaku digital responden, tanpa memisahkan secara eksplisit ke dalam sub-dimensi durasi perbandingan sosial daring, *Fear of Missing Out (FoMO)*, dan durasi penggunaan media sosial.
3. Variabel Dependen (Kecemasan Sosial): Versi laporan mandiri (*self-report*) *Liebowitz Social Anxiety Scale (LSAS)* digunakan untuk mengukur variabel terkait.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengembangkan sistem untuk memprediksi tingkat kecemasan sosial Generasi Z berdasarkan unsur-unsur gaya hidup digital dengan menggunakan algoritma *Naive Bayes*, *K-Nearest Neighbor*, dan *Random Forest*.
2. Membandingkan dan mengevaluasi tingkat akurasi ketiga algoritma tersebut dalam memprediksi kecemasan sosial pada Gen Z.
3. Mengidentifikasi faktor gaya hidup digital yang paling berpengaruh terhadap tingkat kecemasan sosial Gen Z.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan dalam ilmu komputer dalam machine learning, dengan fokus pada penerapan algoritma *Naïve Bayes*, *K-Nearest Neighbor*, dan *Random Forest* untuk menganalisis kesehatan mental yang berbasis data digital.

2. Manfaat Praktis

Manfaat praktis dari penelitian ini meliputi:

1. Menghasilkan model prediksi yang dapat digunakan sebagai alat skrining awal kecemasan sosial bagi Generasi Z.
2. Memberikan informasi bagi pengembang aplikasi kesehatan mental dalam merancang intervensi berbasis teknologi.
3. Meningkatkan kesadaran masyarakat tentang pentingnya kesehatan mental dan dampak gaya hidup digital.

1.6 Sistematika Penulisan

Sistematika penulisan penelitian ini disusun sebagai berikut:

BAB I PENDAHULUAN Pada bab ini memuat latar belakang, perumusan masalah, batasan masalah, tujuan, manfaat, sistematika penulisan dan penelitian terkait.

BAB II TINJAUAN PUSTAKA Bab ini menjelaskan tentang kecemasan sosial, Generasi Z, gaya hidup digital, *machine learning*, algoritma *naïve bayes*, algoritma *k-nearest neighbor*, dan algoritma *random forest* dan metode evaluasi yang digunakan.

BAB III METODOLOGI PENELITIAN Bab ini memaparkan langkah-langkah penelitian: pengumpulan data, pengelolaan data, pembentukan model, pengujian model, dan evaluasi hasil.

BAB IV HASIL DAN PEMBAHASAN Bab ini membahas tahapan implementasi dan analisis perbandingan.

BAB V KESIMPULAN Bab ini berisi kesimpulan dari pembahasan dan serta saran yang bermanfaat untuk pengembangan penelitian.

1.7 Penelitian Terkait

Tabel 1. 1 Penelitian Terkait

Penulis dan Tahun	Judul	Hasil/Pencapaian
Dewi, Utami, & Hasanah (2025)	Deteksi Resiko Depresi dan Kecemasan Mahasiswa Menggunakan Algoritma <i>Naïve Bayes</i>	Menunjukkan efektivitas algoritma <i>Naïve Bayes</i> dalam mendeteksi resiko depresi dan kecemasan pada mahasiswa.
Faozi, Nisa, & Sari (2023)	Analisis Faktor Penyebab Kecemasan Sosial pada Remaja Pengguna Media Sosial	Mengidentifikasi faktor kunci: perbandingan sosial, FOMO, <i>cyberbullying</i> , dan ketergantungan validasi online sebagai faktor penting dalam perkembangan kecemasan sosial.
Fitri, Septyani, & Ramadhan (2025)	Penggunaan <i>Random Forest</i> Dalam Klasifikasi Kecemasan Pada Generasi Z	Membangun sistem klasifikasi kecemasann pada Generasi Z dengan akurasi 97,71% menggunakan <i>Random Forest</i> .
Haspriyanti & Prasetyaningrum (2021)	Penerapan Algoritma <i>K-Nearest Neighbor</i> untuk Deteksi Dini Gangguan Mental pada Mahasiswa	Mengimplementasikan KNN untuk deteksi gangguan mental dengan akurasi 87.3% menggunakan $k=7$, menunjukkan keunggulan KNN dalam menangkap pola non-linear.

Hidayat & Rizal (2023)	Analisis Performa Algoritma <i>Machine Learning</i> dalam Prediksi Gangguan Kesehatan Mental Berbasis Data Digital	Membandingkan beberapa algoritma (<i>Naïve Bayes</i> , KNN, SVM, <i>Decision Tree</i>) dan menemukan bahwa tidak ada algoritma yang unggul di semua metrik, pemilihan bergantung pada karakteristik dataset.
Kusuma, Dewi, & Sari (2024)	Hubungan Intensitas Penggunaan Media Sosial dengan Kecemasan Sosial pada Generasi Z di Indonesia	Menemukan korelasi positif signifikan antara waktu di media sosial dengan tingkat kecemasan sosial pada Gen Z, terutama pada aspek ketakutan evaluasi negatif dan penghindaran situasi sosial.
Raharja, Sari, & Wijaya (2024)	Teknik <i>Data Splitting</i> dan Evaluasi Model <i>Machine Learning</i> untuk Data Kesehatan Mental	Mengembangkan <i>framework</i> evaluasi komprehensif menggunakan multiple metrics (akurasi, presisi, recall, F1-score) dan cross-validation untuk estimasi performa yang robust.
Sumantri, Haryanto, & Kurniawan (2023)	Penerapan <i>Machine Learning</i> untuk Analisis Data Kesehatan Mental Berbasis Media Sosial	Menunjukkan bahwa algoritma <i>machine learning</i> dapat mengidentifikasi pola kompleks dalam data digital yang berkaitan dengan kondisi kesehatan mental dengan tingkat akurasi yang baik.

Yolanda, Septyani, & Pratama (2021)	Implementasi <i>K-Nearest Neighbor</i> untuk Prediksi Tingkat Stres Mahasiswa	Menemukan bahwa $k=5$ memberikan keseimbangan terbaik antara akurasi dan efisiensi. Normalisasi data dengan Min-Max Scaling meningkatkan performa KNN secara signifikan.
Setiawan, Yasin, & Desianti (2025)	Komparasi Kinerja Algoritma <i>Random Forest</i> , <i>Decision Tree</i> , <i>Naïve Bayes</i> , dan KNN dalam Prediksi Tingkat Depresi Mahasiswa menggunakan <i>Student Depression Dataset</i>	Membandingkan kinerja empat algoritma klasifikasi (<i>Random Forest</i> , <i>Decision Tree</i> , <i>Naïve Bayes</i> , dan KNN) untuk prediksi depresi mahasiswa. Hasil menunjukkan <i>Random Forest</i> memiliki performa terbaik dalam menyeimbangkan <i>Precision</i> dan <i>Recall</i> , sementara <i>Naïve Bayes</i> unggul pada aplikasi yang memprioritaskan <i>Precision</i> tinggi.