

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Teori

2.1.1 Kecemasan Sosial (*Social Anxiety Disorder*)

Kecemasan sosial adalah rasa cemas yang timbul secara berlebihan serta bersifat menetap terkait kemungkinan mengalami rasa malu atau menerima penilaian negatif dari orang lain ketika terlibat dalam interaksi sosial. Individu yang mengalami kondisi ini umumnya memiliki kecenderungan menghindari pertemuan sosial dan merasa lebih nyaman ketika sendirian, sehingga kondisi ini dapat mengganggu aktivitas sehari-hari mereka Association (2022).

Menurut Faozi et al. (2023), kecemasan sosial dapat dibedakan menjadi tiga jenis:

1. Perilaku: Merasakan kegelisahan, ketegangan dalam tubuh, respons mengejutkan, bicara lebih cepat dari biasanya, kurangnya koordinasi, mudah mengalami cedera, menarik diri dari interaksi sosial, serta kewaspadaan yang berlebihan.
2. Kognitif: Kesulitan dalam memfokuskan perhatian, penurunan konsentrasi, mudah lupa, pemikiran yang terhambat, penurunan kreativitas dan produktifitas, kebingungan, kewaspadaan yang meningkat, takut kehilangan control, serta muncul mimpi buruk.
3. Emosi/Afektif: Mudah terganggu, tidak sabaran, merasakan kegelisahan, ketegangan, perasaan canggung, rasa takut yang berlebihan, merasa bersalah, serta rasa malu.

Tingkat kecemasan diukur dari jumlah skor jawaban semua pertanyaan. Zahra (2024) menyatakan bahwa tingkat kecemasan dalam skala ini dibagi menjadi empat kategori, yaitu Non-Klinis, Ringan, Sedang, Berat.

2.1.2 Generasi Z

Generasi Z (Gen Z) mencakup individu-individu yang lahir pada rentang tahun 1997 hingga 2012 dan dikenal sebagai generasi pertama yang tumbuh secara penuh dalam lingkungan digital. Tidak seperti generasi sebelumnya, Gen Z telah memiliki akses ke internet, gadget pintar, dan media sosial sejak usia yang relatif muda, sehingga mengembangkan pola perilaku, komunikasi dan interaksi sosial yang sangat dipengaruhi oleh lingkungan digital Chang et al. (2023).

Di Indonesia, Gen Z adalah kelompok pengguna internet yang sangat aktif. Asosiasi Penyelenggara Jasa Internet Indonesia (APJII, 2023) menyatakan bahwa 34,4% dari seluruh pengguna internet di negara ini berusia 13-28 tahun. Ini menunjukkan bahwa Gen Z berkaitan erat dengan meningkatnya resiko kecemasan sosial, yang disebabkan oleh perbandingan sosial daring dan Fear of Missing Out (FoMO) Santoso et al. (2025).

2.1.3 Gaya Hidup Digital

Gaya hidup digital memicu pada perilaku dan aktivitas harian individu yang dipengaruhi oleh pemanfaatan teknologi digital, seperti internet, media sosial, gadget pintar dan platform online. Ini mencakup berbagai aspek kehidupan, mulai dari cara orang berinteraksi, berbelanja, bekerja, hingga mendapatkan hiburan dan informasi. Untuk Generasi Z, gaya hidup digital menjadi ciri yang menonjol karena mereka merupakan generasi pertama yang tumbuh besar sepenuhnya dalam era digital, sehingga mayoritas interaksi sosial, proses pembelajaran dan aktivitas sehari-hari mereka terjadi melalui perangkat dan platform online Wahyu et al. (2020).

Salah satu komponen utama gaya hidup digital yang sering diteliti adalah seberapa sering pengguna media sosial, yang ternyata sangat berhubungan dengan kondisi mental penggunanya. Penelitian Permadi (2022) tingkat kecemasan sosial remaja berkorelasi dengan lamanya waktu yang mereka habiskan di media sosial; semakin banyak waktu yang mereka habiskan di platform-platform tersebut, semakin besar kemungkinan mereka mengalami

gejala kecemasan saat berinteraksi dengan orang lain. Trikindini et al. (2021), melaporkan hasil yang serupa, yaitu adanya korelasi yang kuat antara tingkat kecemasan mahasiswa dan seberapa sering mereka menggunakan media sosial.

Keterikatan antara gaya hidup digital dan kecemasan sosial semakin jelas dengan adanya penelitian yang fokus pada Generasi Z. Nurlina at al. (2022) menganalisis hubungan antara frekuensi pengguna media sosial dan tingkat kecemasan di kalangan mahasiswa Gen Z. Hasil temuannya menunjukkan bahwa pengguna media sosial yang intensif berkontribusi pada peningkatan kecemasan, terutama akibat kecenderungan membandingkan diri dengan orang lain dan rasa takut akan penilaian negative. Temuan ini menekankan bahwa gaya hidup digital, khususnya penggunaan media sosial adalah faktor penting yang perlu diteliti lebih dalam terkait dengan kecemasan sosial pada Gen Z di Indonesia.

2.1.4 Machine Learning

Machine learning adalah teknologi yang menggabungkan komputer dan matematika, memanfaatkan data sebagai sarana pembelajaran bagi sistem komputer untuk menghasilkan prediksi di masa depan. Teknologi ini muncul sebagai solusi untuk melakukan deteksi awal terhadap suatu penyakit. Prediksi diagnostic menggunakan machine learning telah terbukti menjadi salah satu metode paling efektif dalam berbagai tantangan di kehidupan. Metode machine learning dibangun melalui beragam alat dan teknik yang dapat mengeksplorasi kumpulan data besar, sehingga menghasilkan pola-pola signifikan yang bermanfaat bagi perkembangan ilmu pengetahuan Hidayat (2023).

2.1.5 Algoritma Naïve Bayes

Naïve Bayes merupakan salah satu metode dalam machine learning yang mengandalkan perhitungan probabilitas berdasarkan statistik yang pertama kali diperkenalkan ilmuwan asal Inggris, Thomas Bayes. Metode ini bekerja dengan cara memprediksi kemungkinan masa depan berdasarkan kejadian yang terjadi sebelumnya Pebdika (2023). Algoritma ini dikenal luas sebagai salah satu teknik data mining yang praktis digunakan dan memiliki kecepatan pemrosesan yang tinggi, serta mudah diterapkan karena strukturnya yang relatif sederhana namun

tetap menunjukkan tingkat efektivitas yang baik Punkastyo et al. (2024). *Naïve Bayes* termasuk dalam kategori machine learning yang menerapkan perhitungan probabilitas berdasarkan konsep pendekatan *Bayes*. Dalam algoritma *Naïve Bayes*, penerapan teorema Bayes melibatkan penggabungan probability dan conditional probability dalam sebuah rumus guna menghitung tingkat probabilitas dari masing-masing kemungkinan klasifikasi Nasokha (2024).

Rumus dasar *Naïve Bayes*:

$$P(C|X) = \frac{P(X|C).P(C)}{P(x)} \quad (1)$$

Dimana:

1. X = data yang kelasnya (class) belum diketahui
2. C = hipotesis bahwa data tersebut termasuk ke dalam kelas tertentu
3. $P(C|X)$ = nilai probabilitas dari suatu hipotesis dengan mempertimbangkan kondisi tertentu (*posterior probability*)
4. $P(C)$ = nilai probabilitas awal dari suatu hipotesis (*prior probability*)
5. $P(X|C)$ = nilai probabilitas munculnya suatu kondisi dengan mengacu pada hipotesis yang ada
6. $P(X)$ = nilai probabilitas dari kemunculan data X

2.1.6 Algoritma *K-Nearest Neighbor*

Algoritma *K-Nearest Neighbor* (KNN) Algoritma *K-Nearest Neighbor* (KNN) adalah suatu metode yang bekerja dengan mengklasifikasikan objek berdasarkan kedekatannya dengan data latih yang tersedia. Dalam proses ini, kumpulan data pelatihan digambarkan dalam bentuk ruang multidimensi, di mana masing-masing dimensi mencerminkan satu fitur atau karakteristik data tertentu. Ruang tersebut selanjutnya terbagi menjadi beberapa area yang disesuaikan dengan pengelompokan data latih yang sudah ada. Suatu titik akan digolongkan ke dalam kelas c apabila kelas tersebut merupakan kelas yang paling dominan atau paling sering ditemukan di antara k tetangga yang jaraknya paling dekat dengan titik itu. Adapun tingkat kedekatan maupun jauhnya jarak antar titik

umumnya diukur menggunakan perhitungan jarak Euclidean, sesuai dengan formula yang tercantum pada persamaan terkait Hilmi et al. (2024) .

Rumus *K-Nearest Neighbor*:

$$d(x,y) = \sqrt{\sum_{i=1}^n (x_{i \text{ training}} - y_{i \text{ testing}})^2} \quad (2)$$

Dimana:

1. $d(x,y)$ = jarak
2. $x_{i \text{ training}}$ = Data training
3. $y_{i \text{ testing}}$ = Data testing
4. i = variable data
5. n = dimensi data

2.1.7 Algoritma *Random Forest*

Random Forest adalah salah satu algoritma machine learning berlandaskan ensemble learning, di mana sejumlah besar pohon keputusan (decision tree) secara bersamaan untuk menghasilkan satu model prediksi yang lebih akurat dan stabil. Masing-masing pohon dalam Random Forest dilatih menggunakan sampel data yang diambil secara acak dengan pengembalian dari data latih (*bootstrap aggregating* atau *bagging*), sementara pemilihan fitur pada setiap percabangan pohon juga dilakukan secara acak dari sebagian subset fitur yang tersedia. Kombinasi kedua proses acak ini membuat setiap pohon acak ini membuat setiap pohon memiliki karakteristik yang berbeda-beda, sehingga mengurangi kolerasi antar pohon dan menurunkan resiko overfitting yang kerap muncul pada model pohon keputusan tunggal Arumnisaa et al. (2023).

Pada tahap klasifikasi, setiap pohon dalam hutan (*forest*) akan memberikan satu suara (vote) terhadap kelas prediksinya, dan penentuan hasil klasifikasi akhir dilakukan melalui majority voting, yaitu kelas yang paling banyak dipilih oleh keseluruhan pohon yang ada. Keunggulan utama Random Forest terletak pada kapasitasnya dalam mengolah data berdimensi tinggi, tetap stabil terhadap data yang mengandung derau (noise) maupun data yang tidak seimbang, serta mampu menghasilkan tingkat akurasi yang tinggi dibandingkan algoritma perbandingan lainnya Juliandewi et al. (2025).

2.1.8 Evaluasi Model

Proses evaluasi memegang peranan yang sangat penting dalam machine learning karena berfungsi untuk mengukur tingkat kemampuan suatu model dalam menghasilkan prediksi yang tepat terhadap data baru yang belum pernah diproses sebelumnya. Evaluasi ini dilakukan menggunakan metrik-metrik kuantitatif yang dapat menggambarkan performa model secara komprehensif. Dalam konteks klasifikasi seperti penelitian ini terdapat beberapa metrik evaluasi yang umum digunakan:

1. Akurasi (*Accuracy*): Metrik ini digunakan untuk mengukur besarnya proporsi prediksi yang tepat dibandingkan dengan keseluruhan prediksi yang dilakukan oleh model. Meskipun mudah diinterpretasi, akurasi dapat menyesatkan jika dataset tidak seimbang (*imbalanced class*).

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Prediksi}} \times 100\% \quad (3)$$

2. Presisi (*Precision*): Metrik ini digunakan untuk mengukur tingkat ketepatan model dalam menghasilkan prediksi positif. Presisi tinggi berarti model jarang memprediksi positif palsu (*false positive*).

$$\text{Presisi} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (4)$$

3. Recall (*Sensitivity/True Positive Rate*): Metrik ini digunakan untuk menilai seberapa baik kemampuan model dalam mendeteksi seluruh kasus positif yang sesungguhnya ada. Recall tinggi berarti model jarang melewatkan kasus positif (*false negative*).

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (5)$$

4. F1-Score: Merupakan nilai rata-rata harmonik yang diperoleh dari kombinasi antara presisi dan recall. Metrik ini sangat berguna ketika perlu

dicari keseimbangan antara kedua metrik tersebut, terutama ketika distribusi kelas dalam data tidak seimbang.

$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (6)$$

5. *Confusion Matrix*: Merupakan suatu tabel yang berfungsi untuk menjelaskan kinerja model klasifikasi melalui perbandingan hasil prediksi terhadap nilai sebenarnya. Tabel ini terdiri dari empat elemen:
 1. True Positive (TP): Kondisi ketika model memberikan hasil prediksi positif yang sesuai dengan kenyataan
 2. True Negative (TN): Kondisi ketika model memberikan hasil prediksi negatif yang sesuai dengan kenyataan
 3. False Positive (FP): Kondisi ketika model memberikan hasil prediksi positif yang ternyata keliru
 4. False Negative (FN): Kondisi ketika model memberikan hasil prediksi negatif yang ternyata keliru

2.2 Kerangka Teori

Kerangka teori dalam penelitian ini menjelaskan hubungan sistematis antara variabel-variabel yang diteliti, yaitu faktor gaya hidup digital sebagai variabel dependen, dengan algoritma *machine learning Naïve Bayes*, *K-Nearest Neighbor*, dan *Random Forest* yang masing-masing berfungsi mengidentifikasi keterhubungan tersebut secara akurat.

2.2.1 Model Hubungan Variabel

Model ini menjelaskan perihal bagaimana Gaya Hidup Digital (Variabel Independen) berpotensi memengaruhi Tingkat Kecemasan Sosial (Variabel Dependen) melalui serangkaian Mekanisme Psikologis Perantara.

- a. Gaya Hidup Digital (Variabel Independen)

Pada konteks ini, terdapat tiga komponen pokok yang membentuk gaya hidup digital, di mana ketiganya memiliki keterkaitan satu sama lain:

1. **Durasi Penggunaan Media Sosial:** Total durasi seseorang mengakses platform media sosial. Aspek ini begitu erat dengan interaksi konten sosial yang berkelanjutan dan intensif, membentuk landasan dua bagian selanjutnya.
2. **Fear of Missing Out (FoMO):** Kondisi psikologis berupa kecemasan, kekosongan, sekaligus kekhawatiran karena merasa tertinggal dari momen berharga tanpa kehadiran dirinya. Ini merupakan respons emosional terhadap paparan media sosial yang terus-menerus.
3. **Perbandingan Sosial Daring:** Keinginan seseorang dalam menakar pribadinya, capaian, serta keseharian lewat cara menandingkannya dengan potret digital orang lain yang kerap ditata dan disunting sedemikian rupa.

b. Mekanisme Psikologis Perantara

Kombinasi dari paparan berlebihan, FoMO, dan perbandingan sosial daring memunculkan proses psikologis internal yang merusak, yaitu:

1. Penurunan Harga Diri (*Self-Esteem*): Membandingkan hidup pribadi dengan "highlight reel" orang lain bisa memicu perasaan tidak mampu, tidak menarik, dan tidak cukup, sehingga mengikis rasa percaya diri.
2. Distorsi Kognitif: Pola pikir tidak rasional yang berkembang, seperti berpikir hitam-putih ("Jika aku tidak populer di media sosial, berarti aku gagal"), menyalahkan diri sendiri, atau melebih-lebihkan penilaian orang lain.
3. Tekanan untuk Konformitas Digital: Hasrat mendesak untuk mengikuti trend, standar kecantikan, atau opini populer daring agar diterima dan tidak dikucilkan, menciptakan beban psikologis tambahan.
4. Kelelahan Kognitif Digital: Lelah mental akibat stimulasi informasi yang berlebihan, multitasking digital, dan tuntutan untuk selalu "terhubung" dan responsif, sehingga menghambat kapasitas dalam meredam tekanan emosi.

Mekanisme perantara ini secara kolektif membebani kesehatan mental individu, yang pada akhirnya mengakibatkan:

c. **Tingkat Kecemasan Sosial (Variabel Dependen)**

Kecemasan sosial yang dihasilkan dapat diklasifikasikan dalam spektrum keparahan, yang sering diukur dengan skala seperti *Liebowitz Social Anxiety Scale-Self Report (LSAS-SR)*:

1. **Tingkat Non-Klinis (Skor 0-20):** Indikasi kecemasan sosial sangat tipis bahkan nihil, tidak terlalu mengganggu aktivitas kehidupan sehari-hari.
2. **Tingkat Ringan (Skor 21-38):** Rasa cemas mulai muncul serta sedikit mengganggu beberapa aktivitas sosial, namun tetap bisa dikendalikan.
3. **Tingkat Sedang (Skor 39-50):** Gejala kecemasan sosial yang jelas dan cukup mengganggu, sering memicu sikap menarik diri dari lingkungan serta menciptakan tekanan batin.
4. **Tingkat Berat (Skor 51-60):** Kecemasan sosial yang ekstrem, melumpuhkan, dan menyeluruh. Menghindari interaksi begitu kuat, merusak stabilitas hidup (akademik, karier, dan relasi), dan seringkali membutuhkan penanganan profesional.

2.2.2 Landasan Teori Pendukung

a. **Teori Uses and Gratification**

Menurut Katz, Blumler, dan Gurevitch (1973), mengemukakan bahwa teori ini menerangkan bagaimana setiap individu secara sadar dan aktif menentukan pilihan media guna memuaskan kebutuhan psikologis maupun sosial yang dimilikinya. Dalam konteks media sosial, Generasi Z menggunakan platform digital untuk memenuhi kebutuhan:

1. Kebutuhan kognitif (mencari informasi)
2. Kebutuhan afektif (pencarian dukungan emosional)
3. Kebutuhan integrasi personal (penguatan identitas)
4. Kebutuhan integrasi sosial (interaksi dengan kelompok)

Namun, ketika pengguna media sosial berlebihan, dapat terjadi *gratification displacement* di mana kebutuhan yang seharusnya terpenuhi justru menimbulkan efek negatif seperti kecemasan sosial.

b. Social Comparison Theory

Menurut Festinger (1954), melalui teori perbandingan sosial mengungkapkan bahwa setiap individu pada dasarnya memiliki kecenderungan untuk menilai dirinya sendiri dengan cara membandingkannya terhadap pihak lain di sekitarnya. Di era digital, perbandingan ini terjadi secara konstan melalui:

1. Upward Comparison: Menempatkan diri secara sejajar dengan pihak yang dinilai memiliki kondisi lebih baik.
2. Downward Comparison: Menempatkan diri secara sejajar dengan pihak yang dianggap kurang baik.

Media sosial memperkuat upward comparison karena menampilkan *highlight reels* kehidupan orang lain, yang dapat menurunkan self-esteem dan meningkatkan kecemasan sosial.

c. Theory of Planned Behavior

Menurut Ajzen (1991), menerangkan melalui teorinya bahwa kecenderungan seseorang untuk berperilaku di media sosial dipengaruhi oleh:

1. Sikap terhadap perilaku online
2. Norma subjektif (pressure dari lingkungan sosial)
3. Persepsi kontrol perilaku

Ketika individu merasa kurang kontrol atas pengalaman digital mereka (misalnya, ketidakmampuan menghentikan scrolling), dapat timbul kecemasan yang berujung pada gangguan kecemasan sosial.

d. Digital Wellbeing Framework

Menurut Przybylski & Weinstein (2017), Kerangka kesejahteraan digital menekankan pentingnya keseimbangan antara manfaat teknologi

dan dampak psikologisnya. Faktor-faktor yang mempengaruhi wellbeing digital termasuk:

1. Kualitas penggunaan (bukan sekadar kuantitas)
2. Motivasi penggunaan (intrinsik vs ekstrinsik)
3. Kemampuan regulasi diri digital

2.2.3 Hipotesis Penelitian

Mengacu pada kerangka teori yang dikembangkan, dirumuskan hipotesis penelitian:

a. Hipotesis 1:

H₁: erdapat perbedaan signifikan dalam performa ketiga model klasifikasi yang digunakan dalam memprediksi tingkat kecemasan sosial pada Gen Z berdasarkan gaya hidup digital.

b. Hipotesis 2:

H₂: Faktor *Fear of Missing Out* (FoMO) dan intensitas perbandingan sosial daring merupakan prediktor paling dominan terhadap tingkat kecemasan sosial pada Gen Z dibandingkan faktor durasi pengguna media sosial dan jumlah platform yang digunakan.

c. Hipotesis 3:

H₃: Model terbaik yang dihasilkan dari ketiga model klasifikasi tersebut akan memiliki nilai presisi dan recall minimal 75% untuk setiap kategori kecemasan sosial.

2.2.4 Definisi Operasional Variabel

2.2.4.1 Variabel Independen

Tabel 2. 1 Variabel Independen

Variabel	Definisi Operasional	Skala	Instrumen Pengukuran
Durasi Penggunaan Media Sosial	Jumlah waktu (dalam jam) yang dihabiskan menggunakan platform media sosial per hari	Rasio	Self-report melalui kuesioner
Fear of Missing Out (FoMO)	Tingkat kecemasan individu tentang kemungkinan kehilangan pengalaman positif yang dialami orang lain	Interval (FoMO Scale, 10 item, skala 1-5)	Total skor FoMO Scale (10-50)
Perbandingan sosial Daring	Frekuensi individu membandingkan diri dengan orang lain di media sosial.	Interval (5 item, skala 1-5)	Rata-rata skor perbandingan sosial

2.2.4.2 Variabel Dependen

Tabel 2. 2 Variabel Dependen

Variabel	Definisi Operasional	Skala	Instrumen Pengukuran
Tingkat Kecemasan	Tingkat keparahan gejala kecemasan sosial berdasarkan kriteria diagnostik	Ordinal (kategorikal)	Skor total LSAS-SR (0-60) dikategorikan: Non-Klinis (0-20), Ringan (21-38), Sedang (39-50), Berat (51-60)