

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kesehatan mental merupakan isu yang semakin mendapat perhatian di Indonesia, mengingat dampaknya yang luas terhadap kualitas hidup, produktivitas, serta hubungan sosial individu di tengah tantangan kehidupan modern dan keberagaman budaya [1], [2]. Salah satu isu kesehatan mental yang kerap dialami oleh generasi muda di Indonesia adalah *quarter-life crises* (QLC) [3]. Masa QLC umumnya terjadi pada rentang usia 20 hingga 30 tahun, seringkali ditandai dengan ketidakpastian dan keraguan terkait arah hidup, karier, dan hubungan personal. Periode ini dapat memicu peningkatan tingkat stres, kecemasan, dan depresi di kalangan individu muda [3], [4]. Berdasarkan laporan dari *World Health Organization* (WHO, 2024), stres kronis pada usia muda dapat memicu gangguan mental yang lebih serius [5]. Meskipun dukungan sosial berperan dalam mengurangi tingkat stres, metode tradisional dalam mengukur stres masih memiliki keterbatasan karena sifatnya yang subjektif. Oleh karena itu, metode berbasis *machine learning* dapat menjadi pendekatan kuantitatif yang lebih objektif dan akurat dalam memprediksi tingkat stres.

Berbagai penelitian telah dilakukan untuk memprediksi tingkat stres menggunakan algoritma *machine learning*. Misalnya, sebuah studi menggunakan teknik data mining lanjutan untuk menganalisis DASS-42 dalam mengidentifikasi tingkat depresi, kecemasan, dan stres pada mahasiswa [6]. Namun, penelitian ini hanya berfokus pada penelitian internal tanpa melibatkan faktor eksternal dari individu yang diteliti. Selain itu, penelitian lain membandingkan berbagai model *machine learning* dalam memprediksi tingkat depresi, kecemasan, dan stres berdasarkan data DASS-42, menunjukkan potensi algoritma seperti *Random Forest* dan *XGBoost* dalam menghasilkan prediksi yang akurat [7]. Meskipun penelitian terkait prediksi gangguan mental dengan *machine learning* telah berkembang, sebagian besar studi lebih berfokus pada depresi dan kecemasan secara umum. Belum banyak penelitian yang secara spesifik mengkaji prediksi tingkat stres pada kelompok usia *quarter-life* dengan mempertimbangkan faktor-faktor psikososial yang relevan. Oleh karena itu, diperlukan pendekatan yang lebih efisien dan akurat dalam memprediksi tingkat stres pada kelompok usia ini.

Penelitian sebelumnya menunjukkan bahwa algoritma *machine learning* dalam prediksi tingkat stres sering menghadapi tantangan dalam ketidakseimbangan kelas data, yang dapat memengaruhi akurasi model. Oleh karena itu, pemilihan algoritma yang tepat,

seperti XGBoost, menjadi faktor kunci dalam meningkatkan performa model prediksi stres[8]. Selain itu, beberapa studi menunjukkan bahwa algoritma seperti *Random Forest* dan *Support Vector Machine* memiliki keterbatasan dalam menangani data yang tidak seimbang dan kompleksitas komputasi yang tinggi [9].

Penelitian ini bertujuan untuk mengatasi kekurangan tersebut dengan mengimplementasikan algoritma XGBoost, sebagai algoritma yang efisien dan skalabel[9], menunjukkan performa stabil dalam prediksi gangguan mental meskipun menghadapi ketidakseimbangan data. Data menggunakan kuesioner DASS-42, yang menyediakan informasi lebih mendetail mengenai kondisi stres responden. Pendekatan ini berbasis pada skor DASS-42, yang secara umum mengklasifikasikan stres ke dalam beberapa kategori. Meskipun algoritma machine learning seperti *Random Forest* dan *Support Vector Machine* telah banyak digunakan dalam prediksi stres, penelitian sebelumnya umumnya hanya mengkaji populasi umum atau mahasiswa tanpa menekankan pada fase *quarter-life crisis* secara spesifik. Selain itu, mayoritas studi mengelompokkan level stres dalam 3–5 kategori saja tanpa mempertimbangkan kebutuhan klasifikasi granular untuk intervensi yang lebih tepat sasaran. Dalam penelitian ini, skala prediksi dikembangkan lebih lanjut menjadi skala 1 hingga 10 untuk memberikan hasil yang lebih detail dan memungkinkan analisis yang lebih mendalam terhadap tingkat stres individu. Studi ini juga mengintegrasikan atribut kepribadian (TIPI), kognitif (VCL), dan demografi secara komprehensif, menjadikannya salah satu studi pertama yang menggabungkan pendekatan ensemble learning pada domain psikologis di Indonesia dengan fokus pada QLC.

Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam bidang kesehatan mental, khususnya dalam upaya deteksi dini dan manajemen stres pada individu di usia *quarter-life*. Penggunaan algoritma XGBoost dan data dari DASS-42 diharapkan dapat menghasilkan model prediksi yang lebih akurat dan efisien, sehingga mendukung pengambilan keputusan dalam intervensi dan penanganan stres secara lebih efektif.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, permasalahan utama dalam penelitian ini berkaitan dengan bagaimana memanfaatkan data kuesioner DASS-42 yang bersifat subjektif dan terbatas dalam interpretasi konvensional untuk memprediksi tingkat stres pada individu usia seperempat abad (20–30 tahun) secara lebih objektif dan akurat. Selain itu, masih terdapat keterbatasan penelitian terdahulu yang lebih banyak berfokus pada

populasi umum atau mahasiswa tanpa menekankan fase *quarter-life crisis* secara spesifik, serta penggunaan algoritma machine learning yang belum optimal dalam menangani ketidakseimbangan data dan kompleksitas atribut psikososial. Oleh karena itu, diperlukan pendekatan berbasis algoritma XGBoost yang diharapkan mampu meningkatkan akurasi prediksi, mengintegrasikan atribut tambahan seperti kepribadian (TIPI), kemampuan kognitif (VCL), dan data demografi, serta menghasilkan klasifikasi granular tingkat stres dalam skala 1–10 untuk mendukung intervensi yang lebih tepat sasaran. Rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana cara memanfaatkan data dari kuesioner DASS-42 secara optimal untuk menghasilkan prediksi tingkat stres yang lebih objektif dan akurat pada individu usia *quarter-life*, mengingat keterbatasan subjektivitas dan interpretasi konvensional dalam skala stres?
2. Apa saja faktor-faktor yang paling berpengaruh dalam menentukan tingkat stres berdasarkan data DASS-42 dan atribut tambahan seperti kepribadian, nilai hidup, dan demografi?
3. Bagaimana pemanfaatan model prediksi stres dapat mendukung intervensi dan manajemen stres yang lebih efektif pada kelompok usia *quarter-life*?

1.3. Tujuan

Penelitian ini bertujuan untuk mengembangkan model prediksi tingkat stres yang objektif dan akurat bagi individu usia seperempat abad menggunakan algoritma *eXtreme Gradient Boosting* (XGBoost) pada data kuesioner DASS-42 yang diperkaya atribut kepribadian (TIPI), kognitif (VCL), dan demografis. Sekaligus mengidentifikasi faktor-faktor psikososial yang paling dominan dalam menentukan tingkat stres melalui analisis *feature importance*, hingga pada akhirnya menghasilkan klasifikasi granular skala 1-10 yang dapat dimanfaatkan sebagai dasar untuk deteksi dini dan strategi intervensi manajemen stres yang lebih efektif pada kelompok usia tersebut.

1.4. Manfaat

Penelitian ini diharapkan dapat memberikan manfaat:

1. Memberikan model prediksi yang dapat digunakan praktisi kesehatan mental dalam mendeteksi dini tingkat stres pada individu muda.
2. Mendukung pengembangan program intervensi yang lebih tepat sasaran dalam manajemen stres, khususnya usia seperempat abad.

3. Membantu individu dalam memahami tingkat stres dan memberikan wawasan dalam pengambilan keputusan untuk menjaga kesehatan mental.

1.5. Ruang Lingkup

Batasan dan ruang lingkup yang diterapkan dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah data hasil kuesioner DASS-42 yang diperoleh dari repositori GitHub (akses melalui <https://github.com/coeusanalytics/dass-42-analysis/tree/master/data>). Data tersebut mencakup respons kuesioner dari individu yang berada pada rentang usia seperempat abad (20-30 tahun).
2. Data fitur responden diperoleh dari hasil pengisian kuesioner DASS-42 yang mencakup 42 item pertanyaan yang mengukur tiga dimensi utama: depresi, kecemasan, dan stres. Fokus penelitian ini adalah pada dimensi stres untuk memprediksi level stres responden.
3. Hasil prediksi adalah level stres pada skala 1 hingga 10, yang memberikan gambaran lebih granular mengenai tingkat stres individu di usia seperempat abad. Hasil prediksi ini diharapkan dapat membantu dalam mengidentifikasi kelompok berisiko tinggi dan mendukung strategi intervensi yang lebih tepat sasaran.