# **2. TINJAUAN PUSTAKA**

## **2.1 Insomnia**

Insomnia adalah gangguan tidur yang ditandai dengan kesulitan untuk memulai tidur, mempertahankan tidur, atau bangun terlalu pagi dan tidak bisa kembali tidur, sehingga menyebabkan kualitas tidur menjadi buruk dan mengganggu aktivitas sehari-hari (American Academy of Sleep Medicine, 2014). Gangguan ini dapat berlangsung sementara atau kronis, dan sering kali disertai rasa lelah, mudah marah, gangguan konsentrasi, serta penurunan produktivitas.

Menurut American Psychiatric Association (2013), insomnia terjadi meskipun terdapat waktu dan kesempatan yang memadai untuk tidur. Hal ini menunjukkan bahwa insomnia bukan hanya tentang kurang tidur secara kuantitas, tetapi juga tentang kualitas tidur yang buruk.

Insomnia menjadi salah satu gangguan tidur yang paling umum di dunia. Data dari World Health Organization (2021) menunjukkan bahwa sekitar 30% populasi dunia mengalami insomnia akut, dan 10% mengalami insomnia kronis. Insomnia juga sering menjadi gejala dari gangguan psikologis lain, seperti gangguan kecemasan, depresi, atau stres berkepanjangan.

### **2.1.1 Jenis-jenis Insomnia**

Insomnia dapat diklasifikasikan menjadi beberapa jenis berdasarkan durasi, pola gangguan tidur, dan penyebabnya.

1. Berdasarkan durasinya

insomnia dibagi menjadi dua, yaitu insomnia akut dan insomnia kronis. Insomnia akut bersifat jangka pendek, biasanya berlangsung selama beberapa hari hingga beberapa minggu, dan sering kali disebabkan oleh stres, perubahan lingkungan, atau gangguan emosional sementara (National Sleep Foundation, 2020). Sementara itu, insomnia kronis terjadi sedikitnya tiga malam dalam seminggu selama minimal tiga bulan, dan sering dikaitkan dengan kondisi psikologis atau medis yang berlangsung lama, seperti depresi atau gangguan kecemasan (American Psychiatric Association, 2013).

1. Berdasarkan pola gangguan tidurnya

insomnia dibagi menjadi tiga jenis utama, yaitu sleep-onset insomnia, sleep-maintenance insomnia, dan early morning awakening. Sleep-onset insomnia adalah kondisi ketika seseorang mengalami kesulitan untuk memulai tidur, biasanya memerlukan waktu lebih dari 30 menit untuk tertidur. Sleep-maintenance insomnia ditandai dengan sering terbangun di malam hari dan sulit kembali tidur. Sedangkan early morning awakening adalah kondisi di mana seseorang bangun terlalu pagi, seperti pukul tiga atau empat dini hari, dan tidak bisa tidur kembali meskipun waktu tidur belum mencukupi.

1. Berdasarkan penyebabnya (etiologi)

Insomnia dibagi menjadi dua kategori, yaitu primary insomnia dan secondary insomnia. Primary insomnia adalah jenis insomnia yang tidak terkait langsung dengan kondisi medis, psikologis, atau lingkungan tertentu, sehingga sering kali sulit untuk ditentukan penyebab pastinya. Sebaliknya, secondary insomnia terjadi akibat adanya kondisi kesehatan atau faktor eksternal lain, seperti gangguan kecemasan, nyeri kronis, efek samping obat-obatan, atau konsumsi zat tertentu seperti alkohol, nikotin, dan kafein dalam jumlah berlebih.

### **2.1.2 Faktor Risiko Insomnia**

Insomnia merupakan gangguan tidur yang dapat dipengaruhi oleh berbagai faktor risiko, baik yang bersifat psikologis, fisiologis, lingkungan, maupun gaya hidup. Pemahaman terhadap faktor-faktor risiko ini penting untuk mendeteksi dan mencegah terjadinya insomnia sejak dini.

Menurut Morin dan Benca (2012), beberapa faktor risiko utama insomnia antara lain:

1. **Stres dan Tekanan Psikologis**

Kondisi stres yang berkepanjangan akibat tekanan pekerjaan, masalah keuangan, atau konflik interpersonal dapat meningkatkan risiko terjadinya insomnia. Individu yang memiliki gangguan kecemasan atau depresi juga cenderung mengalami gangguan tidur (Baglioni et al., 2011).

1. **Pola Tidur yang Tidak Teratur**

Kebiasaan tidur yang tidak konsisten, seperti sering begadang, tidur siang berlebihan, atau perubahan jam tidur karena pekerjaan shift, dapat mengganggu ritme sirkadian tubuh dan menyebabkan kesulitan tidur pada malam hari (American Academy of Sleep Medicine, 2014).

1. **Gaya Hidup Tidak Sehat**

Konsumsi kafein, alkohol, atau nikotin berdekatan dengan waktu tidur terbukti dapat mengganggu proses tidur alami. Kurangnya aktivitas fisik harian juga dapat memengaruhi kualitas dan durasi tidur (National Sleep Foundation, 2020).

1. **Kondisi Medis dan Fisiologis**

Penyakit kronis seperti hipertensi, diabetes, asma, nyeri kronis, dan gangguan hormonal dapat menyebabkan gangguan tidur. Selain itu, gangguan seperti sleep apnea dan restless legs syndrome juga sering dikaitkan dengan insomnia (Roth, 2007).

1. **Faktor Usia dan Jenis Kelamin**

Prevalensi insomnia cenderung meningkat seiring bertambahnya usia, terutama pada lansia. Selain itu, perempuan memiliki risiko lebih tinggi mengalami insomnia, yang dipengaruhi oleh perubahan hormon selama menstruasi, kehamilan, atau menopause (Zhang & Wing, 2006).

1. **Lingkungan Tidur yang Tidak Mendukung**

Suara bising, pencahayaan berlebih, suhu ruangan yang tidak nyaman, atau penggunaan gawai sebelum tidur dapat mengganggu proses tidur alami seseorang (Hirshkowitz et al., 2015).

Dengan mengenali berbagai faktor risiko tersebut, intervensi pencegahan dan edukasi mengenai pentingnya gaya hidup sehat serta manajemen stres dapat menjadi langkah strategis dalam menurunkan prevalensi insomnia, terutama di kelompok usia produktif.

### **2.2 Machine Learning**

Machine Learning (ML) atau pembelajaran mesin adalah cabang dari kecerdasan buatan (Artificial Intelligence/AI) yang memungkinkan sistem komputer untuk belajar secara otomatis dari data dan meningkatkan kinerjanya dalam menyelesaikan tugas tertentu tanpa harus diprogram secara eksplisit (Alzubaidi et al., 2021). ML berfokus pada pengembangan algoritma yang mampu mengidentifikasi pola serta membuat prediksi atau keputusan berdasarkan data historis.

Dalam penelitian ini, pendekatan yang digunakan adalah **supervised learning**, yaitu metode machine learning yang melibatkan pelatihan model menggunakan data yang telah diberi label. Menurut Zhang et al. (2021), supervised learning bekerja dengan membangun model matematis dari data input, kemudian memanfaatkannya untuk memprediksi output yang telah diketahui sebelumnya.

Data yang digunakan dalam sistem deteksi risiko insomnia ini telah memiliki label, seperti kategori “berisiko” dan “tidak berisiko”. Model dilatih untuk mengenali hubungan antara fitur-fitur input, seperti durasi tidur, tingkat stres, detak jantung, dan faktor-faktor lainnya, dengan label output tersebut. Setelah model selesai dilatih, sistem dapat digunakan untuk memprediksi risiko insomnia berdasarkan data baru yang dimasukkan oleh pengguna.

Machine learning, khususnya supervised learning, telah banyak diterapkan dalam bidang kesehatan, seperti prediksi penyakit jantung, deteksi kanker, diagnosis diabetes, dan pemantauan kondisi mental. Penerapan ML memungkinkan peningkatan efisiensi dalam diagnosis, pengambilan keputusan medis berbasis data, dan mendukung personalisasi layanan kesehatan. Hal ini menjadikan machine learning sebagai teknologi pendukung yang penting dalam pengembangan sistem deteksi dini terhadap gangguan kesehatan, termasuk gangguan tidur seperti insomnia.

Dalam penelitian ini digunakan dua algoritma supervised learning, yaitu Random Forest dan Gradient Boosting, yang merupakan bagian dari teknik ensemble. Kedua algoritma ini dikenal mampu menangani data klasifikasi yang kompleks dan sering digunakan dalam prediksi medis karena akurasinya yang tinggi (Wang et al., 2020).

### **2.2.1 Algoritma Random Forest**

Random Forest adalah algoritma ensemble yang menggunakan pendekatan bagging, yaitu membangun sejumlah pohon keputusan (decision tree) dari subset data yang berbeda, kemudian menggabungkan hasil prediksinya melalui voting mayoritas untuk klasifikasi (Breiman, 2001). Algoritma ini memiliki keunggulan dalam mengatasi overfitting dan memberikan hasil yang stabil meskipun data memiliki noise atau fitur yang tidak terlalu relevan.

Dalam konteks deteksi risiko insomnia, Random Forest membantu memetakan hubungan antara berbagai fitur input seperti kebiasaan tidur, tingkat stres, dan kondisi fisik dengan tingkat risiko yang mungkin dimiliki seseorang.

### **2.2.2 Algoritma Gradient Boosting**

Gradient Boosting adalah algoritma ensemble yang menggunakan pendekatan boosting, di mana model dibangun secara berurutan. Setiap model baru dilatih untuk memperbaiki kesalahan prediksi dari model sebelumnya, sehingga model akhir menjadi sangat kuat (Friedman, 2001). Teknik ini sangat efektif untuk meningkatkan akurasi dalam tugas klasifikasi yang kompleks.

Gradient Boosting bekerja dengan cara meminimalkan fungsi loss secara bertahap dan cocok digunakan ketika dibutuhkan prediksi yang presisi tinggi. Dalam penelitian ini, algoritma ini digunakan untuk meningkatkan performa klasifikasi risiko insomnia berdasarkan data yang diberikan oleh pengguna.

### **2.2.3 Metode Ensemble Learning**

Ensemble learning adalah pendekatan dalam machine learning yang menggabungkan beberapa model prediktif untuk meningkatkan akurasi dan kestabilan hasil prediksi. Alih-alih mengandalkan satu model tunggal, metode ensemble bekerja dengan menggabungkan kekuatan banyak model dasar (base learners) untuk menghasilkan keputusan akhir yang lebih baik (Zhou, 2012).

Terdapat dua teknik utama dalam ensemble learning, yaitu bagging dan boosting. Bagging (bootstrap aggregating) bertujuan mengurangi variansi model dengan cara melatih beberapa model pada subset data yang berbeda secara acak. Sementara boosting bekerja secara iteratif, di mana model baru dilatih untuk memperbaiki kesalahan dari model sebelumnya (Dietterich, 2000).

Dalam penelitian ini, digunakan algoritma Random Forest (metode bagging) dan Gradient Boosting (metode boosting), yang keduanya telah terbukti efektif dalam menangani permasalahan klasifikasi dengan data medis yang kompleks (Kourou et al., 2015).

## **2.3 Bahasa Pemrograman Python**



Gambar 2. 1 Gambar Logo Python

Python adalah bahasa pemrograman tingkat tinggi yang banyak digunakan dalam pengembangan sistem berbasis data, termasuk dalam bidang kecerdasan buatan dan machine learning. Bahasa ini memiliki sintaksis yang sederhana dan mudah dipahami, serta didukung oleh berbagai pustaka (library) yang memudahkan dalam proses analisis data, pelatihan model, serta pembuatan antarmuka aplikasi (Lutz, 2013).

Python sangat populer di kalangan peneliti dan praktisi data science karena kemampuannya dalam menangani data dalam jumlah besar, fleksibilitas sintaks, serta komunitas pengembang yang aktif. Pustaka seperti pandas, numpy, dan scikit-learn menyediakan berbagai fungsi penting untuk mengolah data, membuat model prediksi, dan melakukan evaluasi model secara efisien (McKinney, 2017; Pedregosa et al., 2011).

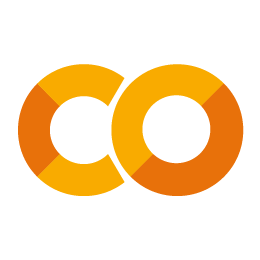
Dalam penelitian ini, Python digunakan sebagai alat utama dalam proses pengolahan data dan pengembangan model deteksi risiko insomnia. Proses pelatihan dan pengujian model dilakukan menggunakan pustaka pandas untuk manipulasi data, numpy untuk perhitungan numerik, dan scikit-learn untuk penerapan algoritma machine learning. Pustaka-pustaka ini memungkinkan proses pelatihan model prediksi menjadi lebih cepat, terstruktur, dan dapat direplikasi.

Selain itu, proses visualisasi data dilakukan menggunakan pustaka seperti matplotlib dan seaborn, yang membantu dalam memahami pola hubungan antar variabel, seperti korelasi antara tingkat stres, durasi tidur, kualitas tidur, dan aktivitas fisik terhadap kejadian insomnia. Visualisasi ini penting untuk mengidentifikasi tren data sebelum memasuki tahap pemodelan.

Pada tahap pemodelan, algoritma klasifikasi seperti Random Forest digunakan untuk memprediksi apakah seseorang berisiko mengalami insomnia berdasarkan sejumlah fitur input, seperti usia, jenis kelamin, tingkat stres, kategori BMI, dan kualitas tidur. Python melalui scikit-learn menyediakan pipeline lengkap untuk membagi data (train-test split), melatih model, melakukan validasi silang (cross-validation), serta mengukur performa model menggunakan metrik seperti akurasi, presisi, recall, dan confusion matrix (Pedregosa et al., 2011).

Secara keseluruhan, penggunaan Python dalam penelitian ini memberikan kemudahan dan efisiensi dalam mengembangkan sistem berbasis machine learning untuk mendeteksi risiko insomnia. Kemampuan Python dalam menangani seluruh siklus proyek data — mulai dari pengolahan hingga evaluasi model — menjadikannya pilihan yang tepat dalam konteks penelitian ini.

## **2.4 Google Colaboratory (Google Collab)**



Gambar 2. 2 Logo Google Collab

Google Colaboratory, atau lebih dikenal sebagai Google Colab, adalah sebuah platform pemrograman berbasis cloud yang dikembangkan oleh Google. Layanan ini memungkinkan pengguna untuk menulis dan mengeksekusi kode Python secara langsung melalui peramban web tanpa perlu menginstal perangkat lunak tambahan atau mengonfigurasi lingkungan pengembangan secara lokal. Google Colab menggunakan antarmuka Jupyter Notebook, yang mendukung kombinasi antara kode program, visualisasi, dan penjelasan dalam format teks secara interaktif dalam satu dokumen yang terintegrasi (Bisong, 2020).

Salah satu keunggulan utama dari Google Colab adalah tersedianya akses gratis terhadap komputasi berbasis GPU (Graphics Processing Unit) dan TPU (Tensor Processing Unit). Fitur ini sangat membantu dalam menjalankan proses komputasi intensif seperti pelatihan model machine learning dan deep learning, terutama ketika pengguna tidak memiliki perangkat keras dengan spesifikasi tinggi. Selain itu, Google Colab memungkinkan instalasi pustaka Python tambahan melalui perintah seperti !pip install, sehingga sangat fleksibel untuk digunakan dalam berbagai proyek penelitian dan pengembangan.

Google Colab juga terintegrasi langsung dengan Google Drive, yang memudahkan pengguna dalam menyimpan, membagikan, dan mengelola file notebook mereka secara daring. Fitur kolaborasi real-time memungkinkan lebih dari satu pengguna untuk mengakses dan mengedit notebook yang sama secara bersamaan, menjadikan platform ini sangat efektif untuk kegiatan pembelajaran dan penelitian kolaboratif.

Dalam konteks penelitian ini, Google Colab digunakan sebagai lingkungan pengembangan untuk membangun, melatih, dan mengevaluasi model machine learning secara efisien. Platform ini dipilih karena kemudahannya dalam visualisasi data, pengolahan dataset, serta dukungannya terhadap berbagai pustaka machine learning seperti TensorFlow, scikit-learn, dan XGBoost. Menurut Prasanna et al. (2021), Google Colab merupakan salah satu platform paling banyak digunakan di dunia akademik karena aksesibilitas, fleksibilitas, dan kemampuan integrasinya yang tinggi terhadap berbagai teknologi komputasi modern.

## **2.5 Streamlit**



Gambar 2. 3 Logo Streamlit

Streamlit adalah sebuah framework open-source berbasis Python yang dirancang khusus untuk membangun antarmuka pengguna (UI) aplikasi data science dan machine learning secara cepat dan interaktif. Dengan sintaks yang sederhana dan deklaratif, Streamlit memungkinkan pengembang untuk mengubah skrip Python menjadi aplikasi web interaktif hanya dengan beberapa baris kode, tanpa memerlukan pengetahuan mendalam tentang front-end development seperti HTML, CSS, atau JavaScript (Carvalho & Welling, 2022).

Salah satu keunggulan utama dari Streamlit adalah kemudahannya dalam menyajikan visualisasi data dan hasil model machine learning dalam bentuk yang mudah dipahami oleh pengguna akhir. Pengembang cukup menulis fungsi-fungsi Python seperti st.line\_chart(), st.write(), atau st.slider() untuk menampilkan grafik, teks, atau kontrol input secara langsung di aplikasi web. Streamlit juga mendukung integrasi dengan pustaka visualisasi populer seperti Matplotlib, Plotly, dan Seaborn, serta dapat dihubungkan dengan framework machine learning seperti Scikit-learn, TensorFlow, dan PyTorch.

Dalam konteks penelitian ini, Streamlit digunakan sebagai *frontend* dari sistem prediksi risiko insomnia, di mana pengguna dapat menginput data seperti durasi tidur, tingkat stres, detak jantung, serta gaya hidup lainnya, dan kemudian memperoleh hasil prediksi secara instan melalui tampilan yang interaktif dan mudah digunakan. Streamlit juga memungkinkan pengembang untuk menambahkan elemen edukatif seperti grafik, deskripsi prediksi, dan saran berdasarkan hasil model, yang menjadikannya tidak hanya sebagai alat teknis, tetapi juga media edukatif.

Menurut Wang et al. (2021), Streamlit telah menjadi salah satu framework yang paling banyak digunakan dalam prototipe sistem berbasis machine learning karena kecepatannya dalam pengembangan serta kemampuannya menghadirkan aplikasi yang responsif dan intuitif. Selain itu, aplikasi Streamlit dapat dijalankan secara lokal maupun di-*deploy* ke cloud platform seperti Heroku atau Streamlit Cloud, menjadikannya fleksibel dan mudah diakses oleh berbagai kalangan pengguna.

## **2.6 Penelitian Terdahulu**

Penelitian tentang deteksi gangguan tidur, khususnya insomnia, dengan pendekatan teknologi terus berkembang seiring meningkatnya kebutuhan akan solusi yang cepat, akurat, dan mudah diakses. Berbagai studi sebelumnya telah menunjukkan potensi penggunaan machine learning dalam mendeteksi risiko insomnia berdasarkan data klinis maupun non-klinis.

Penelitian oleh Yang et al. (2024) menggunakan sinyal EEG (Electroencephalogram) untuk mendeteksi pasien dengan insomnia tanpa memerlukan anotasi tahap tidur. Penelitian tersebut menunjukkan bahwa pendekatan machine learning berbasis EEG memiliki akurasi tinggi dalam membedakan individu dengan gangguan tidur dan yang tidak, namun tetap memiliki keterbatasan dalam hal ketersediaan perangkat EEG yang mahal dan tidak praktis untuk penggunaan umum. Studi ini menekankan pentingnya pengembangan metode non-invasif untuk skrining awal insomnia.

Selanjutnya, Liu et al. (2023) mengembangkan platform berbasis kecerdasan buatan untuk memetakan risiko gangguan tidur pada mahasiswa. Penelitian ini menggunakan kuesioner berbasis persepsi sebagai input, dan hasilnya divisualisasikan secara daring. Meskipun berhasil menunjukkan hasil yang informatif, cakupan input data masih terbatas dan belum mempertimbangkan aspek fisiologis seperti tekanan darah atau detak jantung yang juga berpengaruh terhadap kualitas tidur.

Sementara itu, Airlangga (2024) meneliti efektivitas berbagai algoritma machine learning dalam mendeteksi gangguan tidur berdasarkan data gaya hidup. Hasil penelitian menunjukkan bahwa metode ensemble seperti Random Forest dan Gradient Boosting memiliki performa yang unggul dalam klasifikasi data kompleks. Hal ini memperkuat pilihan algoritma yang digunakan dalam penelitian ini, mengingat kedua metode tersebut mampu menangani fitur yang bervariasi serta memiliki ketahanan terhadap overfitting.

Berdasarkan kajian terhadap penelitian-penelitian sebelumnya, terlihat bahwa meskipun telah banyak pendekatan yang dilakukan, sebagian besar masih terbatas pada data klinis atau kuesioner sederhana. Penelitian ini hadir sebagai alternatif dengan memadukan pendekatan praktis berbasis data non-klinis seperti gaya hidup dan data fisiologis, serta menerapkan metode ensemble yang terbukti akurat dan efisien untuk membangun sistem prediksi risiko insomnia yang mudah digunakan oleh masyarakat umum.