

BAB I

PENDAHULUAN

1.1 Latar Belakang

Tidur merupakan kebutuhan biologis fundamental yang berperan penting dalam menjaga stabilitas fisiologis dan psikologis manusia [1][2][3]. Secara fisiologis, tidur berkontribusi dalam proses pemulihan energi, regulasi hormon, penguatan sistem imun, serta pemeliharaan fungsi metabolik tubuh [1][3]. Selain itu, tidur juga memiliki peran penting dalam fungsi kognitif, seperti konsolidasi memori, pemrosesan informasi, regulasi emosi, stabilitas suasana hati, serta fungsi pemulihan tubuh [2][3][4]. Kualitas tidur yang baik tidak hanya ditentukan oleh durasi tidur, tetapi juga oleh kontinuitas tidur, struktur tidur, latensi tidur, gangguan tidur, fungsi siang hari, serta efisiensi tidur [4][5]. Salah satu indikator yang dapat digunakan untuk menggambarkan kualitas tidur secara kuantitatif adalah *sleep efficiency*, yaitu perbandingan antara waktu tidur efektif dengan total waktu yang dihabiskan di tempat tidur [4][5]. Dalam instrumen *Pittsburgh Sleep Quality Index (PSQI)*, habitual *sleep efficiency* juga termasuk salah satu komponen yang digunakan untuk menilai kualitas tidur [4]. Nilai *sleep efficiency* yang tinggi menunjukkan bahwa waktu istirahat digunakan secara lebih efektif untuk tidur, sedangkan nilai yang rendah dapat menunjukkan adanya gangguan dalam mempertahankan tidur atau kualitas istirahat yang kurang optimal [4][5]. Karena berbentuk nilai numerik kontinu, *sleep efficiency* dapat digunakan sebagai target prediksi untuk memahami kualitas tidur secara lebih terukur melalui pendekatan regresi.

Pemantauan kualitas tidur secara klinis umumnya masih menggunakan metode seperti *polysomnography (PSG)* [1][4][6]. Meskipun PSG dianggap sebagai metode yang akurat dalam mengevaluasi tidur, metode ini membutuhkan peralatan khusus, biaya yang tinggi, kurang portabel, serta umumnya dilakukan dalam lingkungan klinis tertentu [1][4][6]. Selain itu, PSG umumnya hanya dilakukan dalam periode pengamatan yang terbatas, sehingga

kurang praktis untuk pemantauan tidur secara berkelanjutan dalam kehidupan sehari-hari [4][7]. Keterbatasan tersebut mendorong perlunya pendekatan pemantauan tidur yang lebih fleksibel, non-invasif, dan dapat digunakan dalam konteks aktivitas harian pengguna [4][7][8]. Perkembangan *wearable device* seperti *smartwatch* dan *fitness tracker* membuka peluang baru dalam pemantauan kualitas tidur karena perangkat tersebut mampu merekam berbagai data fisiologis dan aktivitas harian secara berkelanjutan [8][9]. Data yang dapat direkam mencakup jumlah langkah, durasi olahraga, detak jantung, variabilitas detak jantung, kadar oksigen dalam darah, tingkat stres, konsumsi kafein, suasana hati, serta indikator tidur [8][9]. Data tersebut bersifat *time series* karena dikumpulkan secara berurutan berdasarkan waktu, sehingga berpotensi digunakan untuk menganalisis hubungan antara pola aktivitas harian, kondisi fisiologis, perilaku pengguna, dan *sleep efficiency* [9][10].

Perkembangan *machine learning* dan analitik prediktif juga membuka peluang untuk meningkatkan pemantauan dan prediksi masalah tidur melalui pemanfaatan data berskala besar [9][10][11]. Data yang digunakan dalam pemodelan tidur dapat mencakup pola tidur, perilaku, data fisiologis, serta faktor lingkungan yang dikumpulkan dari berbagai sumber, termasuk *wearable device* [8][9][11]. *Wearable device* dinilai potensial karena mampu mendukung pemantauan tidur secara berkelanjutan dan *real-time* di luar lingkungan klinis [8][9]. Namun, penerapan model prediksi tidur masih menghadapi beberapa tantangan, seperti reliabilitas data sensor, keterbatasan keberagaman *dataset*, sinkronisasi data, kemampuan generalisasi model pada populasi yang berbeda, serta interpretabilitas model yang masih terbatas [11][12]. Kondisi tersebut menunjukkan bahwa pengembangan model prediksi berbasis *wearable* tidak hanya perlu memperhatikan performa prediksi, tetapi juga perlu mempertimbangkan karakteristik data dan transparansi model [11][13].

Beberapa penelitian sebelumnya telah memanfaatkan data *wearable* untuk menganalisis dan memprediksi kualitas tidur menggunakan pendekatan *machine learning* maupun *Deep Learning* [9][10]. Hasil penelitian tersebut menunjukkan bahwa aktivitas fisik, detak jantung, serta pola fisiologis harian

memiliki keterkaitan dengan kondisi tidur individu [9][10]. Penelitian lain juga menunjukkan bahwa kualitas tidur berkaitan dengan aspek psikologis seperti kecemasan dan depresi, serta dapat dipengaruhi oleh durasi dan intensitas aktivitas fisik [3]. Selain itu, kondisi stres dan tekanan kerja juga dilaporkan memiliki hubungan dengan kualitas tidur, sehingga faktor psikologis dan perilaku harian perlu dipertimbangkan dalam analisis kualitas tidur [12]. Temuan-temuan tersebut menunjukkan bahwa kualitas tidur dipengaruhi oleh berbagai faktor yang saling berkaitan, mulai dari aktivitas harian, kondisi fisiologis, perilaku, hingga faktor psikologis [3][9][10][12].

Meskipun penelitian terkait prediksi tidur telah berkembang, masih terdapat beberapa kekurangan yang perlu diperhatikan. Sebagian penelitian masih berfokus pada klasifikasi kualitas tidur atau gangguan tidur, sedangkan prediksi *sleep efficiency* sebagai nilai numerik kontinu masih belum banyak dibahas secara khusus [9][10][13]. Selain itu, beberapa penelitian masih menggunakan fitur yang terbatas, belum sepenuhnya memanfaatkan pola historis dari data *time series*, serta belum banyak menggabungkan evaluasi performa model dengan interpretasi kontribusi fitur secara jelas [9][10][13]. Padahal, data *wearable* memiliki karakteristik yang kompleks karena terdiri dari berbagai informasi aktivitas harian, kondisi fisiologis, perilaku pengguna, tingkat stres, suasana hati, dan riwayat tidur [8][9][10]. Data tersebut juga dapat bervariasi antar pengguna dan berpotensi mengandung *noise*, sehingga dibutuhkan metode yang dapat menangkap hubungan temporal serta pola data yang tidak bersifat linear [9][10][11]. Keterbatasan tersebut menunjukkan adanya ruang penelitian untuk mengembangkan model prediksi *sleep efficiency* berbasis data *wearable* yang tidak hanya berorientasi pada performa prediksi, tetapi juga memperhatikan pemanfaatan pola historis dan interpretasi kontribusi fitur.

Ruang penelitian tersebut menjadi dasar dilakukannya penelitian ini. Penelitian ini memprediksi nilai *sleep efficiency* menggunakan data *wearable* berbasis *time series* dengan pendekatan regresi. Model yang dibandingkan adalah *LSTM*, *CNN-BiLSTM*, dan *CNN-BiLSTM* dengan *Attention* untuk

mengetahui model yang paling sesuai dalam memprediksi nilai *sleep efficiency* berdasarkan data harian pengguna. Performa model dievaluasi menggunakan metrik regresi, yaitu *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan koefisien determinasi (R^2). Selain itu, SHAP digunakan untuk menginterpretasikan kontribusi fitur terhadap hasil prediksi agar model tidak hanya dinilai berdasarkan nilai kesalahan prediksi, tetapi juga dapat dijelaskan berdasarkan fitur-fitur yang memengaruhi output model [13].

Kontribusi penelitian ini terletak pada penerapan prediksi *sleep efficiency* sebagai target regresi berbasis data *wearable* harian, sehingga hasil prediksi tidak hanya berupa kategori kualitas tidur, tetapi berupa nilai numerik yang lebih terukur. Penelitian ini juga membandingkan beberapa model *Deep Learning* pada data *time series* untuk mengetahui model yang paling sesuai dengan karakteristik data *wearable*. Selain itu, penelitian ini menerapkan SHAP untuk menjelaskan kontribusi fitur terhadap hasil prediksi, sehingga hasil penelitian tidak hanya memberikan informasi mengenai model dengan performa terbaik, tetapi juga memberikan gambaran mengenai faktor-faktor yang berperan dalam prediksi efisiensi tidur. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan model prediksi kualitas tidur berbasis data *wearable* yang lebih terukur, memperhatikan pola historis pengguna, dan lebih mudah diinterpretasikan.

1.2 Rumusan Masalah

Dari permasalahan yang telah dipaparkan sebelumnya, fokus penelitian ini diarahkan pada beberapa pertanyaan berikut:

1. Bagaimana hasil kinerja model *Long Short-Term Memory* (LSTM) dalam memprediksi nilai *sleep efficiency* berdasarkan data *wearable* berbasis *time series*?
2. Bagaimana hasil perbandingan performa model LSTM, CNN-BiLSTM, dan CNN-BiLSTM dengan Attention dalam memprediksi nilai *sleep efficiency*?

3. Fitur-fitur apa saja dari data *wearable* yang paling berkontribusi terhadap hasil prediksi *sleep efficiency*?
4. Bagaimana hasil penggunaan metode *Explainable Artificial Intelligence* berbasis SHAP dapat membantu menginterpretasikan hasil prediksi model *Deep Learning* pada data *wearable*?

1.3 Batasan Masalah

Agar penelitian ini terfokus dan tidak menyimpang dari tujuan yang telah ditetapkan, maka diberikan batasan masalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini merupakan data *time series* harian yang diperoleh dari *wearable device*.
2. *Dataset* yang digunakan berasal dari platform Kaggle, yaitu *dataset Health + Wearables + Stress/Sleep Tracking* yang berisi data pemantauan kesehatan harian pengguna *wearable device* selama kurang lebih enam bulan.
3. Penelitian ini difokuskan pada prediksi nilai *sleep efficiency* sebagai indikator efisiensi tidur, bukan diagnosis gangguan tidur secara klinis.
4. Target prediksi dalam penelitian ini adalah variabel numerik *sleep_efficiency*, sehingga penelitian menggunakan pendekatan regresi.
5. Model yang digunakan dalam penelitian ini dibatasi pada *LSTM*, *CNN-BiLSTM*, dan *CNN-BiLSTM* dengan *Attention*.
6. Metode SHAP digunakan untuk membantu menjelaskan hasil prediksi model dengan melihat kontribusi masing-masing fitur terhadap output prediksi.
7. Evaluasi performa model dilakukan menggunakan metrik regresi, yaitu *Mean Squared Error (MSE)*, *Root Mean Squared Error (RMSE)*, *Mean Absolute Error (MAE)*, dan koefisien determinasi (R^2).
8. Penelitian ini tidak melakukan tahap *deployment* sistem, karena fokus penelitian hanya pada pembangunan model, evaluasi performa, dan interpretasi hasil.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Tujuan utama penelitian ini adalah membangun, mengevaluasi, dan menginterpretasikan model prediksi *sleep efficiency* berbasis data *wearable time series* untuk mengetahui model *Deep Learning* yang paling sesuai dalam memprediksi efisiensi tidur pengguna. Penelitian ini menggunakan pendekatan regresi karena target yang diprediksi berupa nilai numerik kontinu, yaitu *sleep_efficiency*, bukan kategori kualitas tidur.

Secara lebih khusus, tujuan penelitian ini adalah sebagai berikut:

1. Membangun model prediksi nilai *sleep_efficiency* berdasarkan data *wearable* harian berbasis *time series*.
2. Menganalisis kinerja model LSTM dalam memprediksi nilai *sleep_efficiency* berdasarkan data historis pengguna.
3. Membandingkan performa model LSTM, CNN-BiLSTM, dan CNN-BiLSTM dengan *Attention* dalam memprediksi nilai *sleep_efficiency*.
4. Mengidentifikasi fitur-fitur dari data *wearable* yang paling berkontribusi terhadap hasil prediksi *sleep_efficiency*.
5. Menerapkan metode SHAP sebagai pendekatan *Explainable Artificial Intelligence* untuk membantu menginterpretasikan kontribusi fitur terhadap hasil prediksi model.

Dengan demikian, penelitian ini tidak hanya berfokus pada pencarian model dengan nilai kesalahan prediksi paling rendah, tetapi juga berupaya menjelaskan fitur-fitur yang berkontribusi terhadap hasil prediksi model.

1.4.2 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi pada beberapa aspek berikut:

1. Penelitian ini diharapkan dapat memberikan kontribusi dalam penerapan model *Deep Learning* berbasis *time series*, khususnya LSTM, CNN-BiLSTM, dan CNN-BiLSTM dengan *Attention*, untuk prediksi nilai *sleep efficiency* menggunakan data *wearable*.

2. Penelitian ini diharapkan dapat menjadi referensi dalam pengembangan analisis kualitas tidur berbasis data *wearable* yang menggunakan pendekatan regresi, bukan hanya klasifikasi kualitas tidur.
3. Penelitian ini diharapkan dapat memberikan gambaran mengenai fitur-fitur *wearable* yang berkontribusi terhadap prediksi efisiensi tidur, seperti aktivitas harian, kondisi fisiologis, tingkat stres, perilaku pengguna, dan riwayat tidur.
4. Penelitian ini diharapkan dapat mendukung penerapan *Explainable Artificial Intelligence* dalam model prediksi berbasis kesehatan digital, sehingga hasil prediksi tidak hanya dilihat dari performa model, tetapi juga dari interpretasi kontribusi fitur terhadap hasil prediksi.
5. Hasil penelitian ini dapat digunakan sebagai dasar pengembangan sistem pemantauan efisiensi tidur berbasis data *wearable*. Model prediksi yang dibangun berpotensi dikembangkan pada aplikasi atau *dashboard* yang terhubung dengan perangkat pemantau kesehatan untuk memperkirakan nilai *sleep_efficiency* pengguna. Namun, pemanfaatan tersebut masih memerlukan pengujian lanjutan dan validasi menggunakan data nyata dari perangkat pengguna.

1.5 Sistematika Penulisan

Sistematika penulisan skripsi ini disusun sebagai berikut:

1. Bab I Pendahuluan
Bab ini membahas latar belakang penelitian, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, serta sistematika penulisan.
2. Bab II Landasan Teori
Bab ini memuat penelitian terdahulu, teori-teori yang berkaitan dengan kualitas tidur dan *sleep efficiency*, penggunaan *wearable device* untuk pemantauan tidur, serta algoritma yang digunakan dalam penelitian, yaitu LSTM, CNN-BiLSTM, CNN-BiLSTM dengan *Attention*, dan *Explainable Artificial Intelligence* berbasis SHAP.
3. Bab III Metodologi Penelitian

Bab ini menjelaskan tahapan penelitian yang dilakukan menggunakan framework CRISP-DM, mulai dari *business understanding*, *data understanding*, *data preparation*, *modeling*, hingga *evaluation*. Bab ini juga menjelaskan proses pengumpulan data, *preprocessing*, *feature engineering*, pembentukan *sequence time series*, pemodelan menggunakan LSTM, CNN-BiLSTM, dan CNN-BiLSTM dengan *Attention*, evaluasi model menggunakan metrik regresi, serta interpretasi fitur menggunakan SHAP.

4. Bab IV Analisis dan Hasil Penelitian

Bab ini menyajikan hasil pelaksanaan setiap tahapan CRISP-DM, hasil eksplorasi data, proses persiapan data, hasil pemodelan, evaluasi performa model, visualisasi hasil prediksi, interpretasi fitur menggunakan SHAP, serta pembahasan terhadap hasil penelitian.

5. Bab V Simpulan dan Saran

Bab ini berisi simpulan dari hasil penelitian berdasarkan rumusan masalah yang telah ditetapkan, serta saran untuk pengembangan penelitian selanjutnya.

