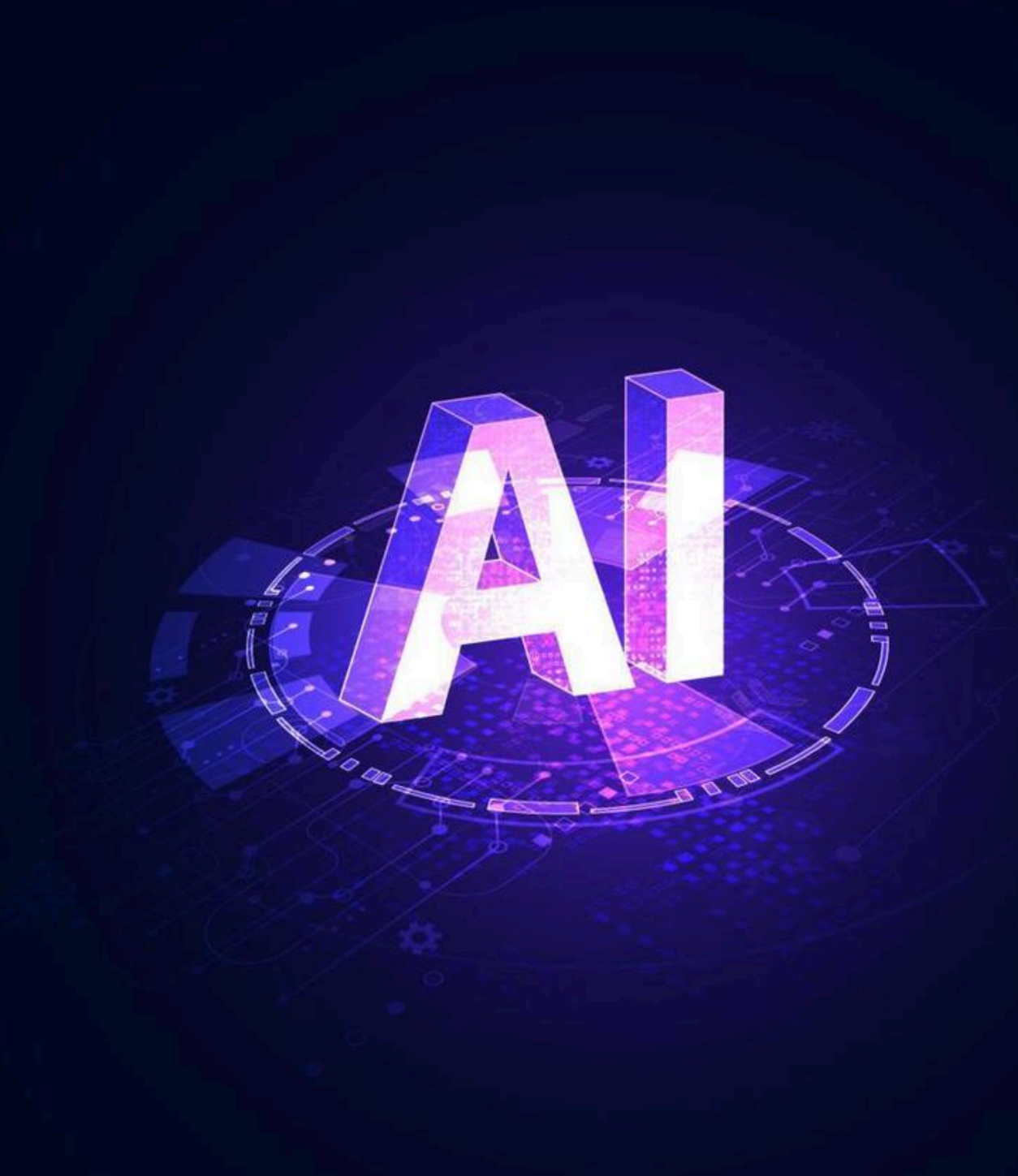
Maharshi Karve Stree Shikshan Samstha's
Cummins College of Engineering For Women,Pune
(An autonomous Institute affiliated to Savitribai Phule Pune University)
Department of Computer Engineering

# SLEEP QUALITY PREDICTION

Submitted By

1. UCE2023562 : KRISHNA SHINDE
2. UCE2023567 : SUMEDHA THORAT
3. UCE2023570 : YASUNDHARA BHENDE

COURSE INSTRUCTOR: DR. MAHENDRA DEORE
COURSE: ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING LABORATORY
23PCCE501L




# INTRODUCTION

Sleep plays a crucial role in physical and mental well-being, influencing productivity, mood, long-term health, and daily functioning.

However, many individuals experience poor sleep quality due to lifestyle habits, stress, health issues, and irregular routines.

This project presents a Sleep Health & Lifestyle Prediction System that uses machine learning to analyze health and lifestyle parameters and predict a person's Quality of Sleep.

Using a Decision Tree classifier, the system processes features such as age, gender, sleep duration, activity level, heart rate, blood pressure, stress, and BMI to generate an interpretable sleep quality score.

# OBJECTIVES

Primary Objectives

- Develop a machine learning model to predict sleep quality from lifestyle and health indicators
- Build a clean and interpretable Decision Tree-based prediction system
- Create a simple UI (Streamlit) for real-time user prediction

SECONDARY OBJECTIVES

- ACHIEVE HIGH TEST ACCURACY WITH MINIMAL OVERFITTING
- IDENTIFY KEY LIFESTYLE FACTORS INFLUENCING SLEEP QUALITY
- PROVIDE PROBABILITY DISTRIBUTION FOR EACH SLEEP-QUALITY LEVEL
- ENABLE EXPANDABLE SYSTEM ARCHITECTURE FOR FUTURE MULTI-MODEL SUPPORT

# DATASET DESCRIPTION

Dataset Source:
Sleep Health and Lifestyle Dataset (CSV format)--kaggle
URL-https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset
Dataset Characteristics

- Total Records: ~374 entries
- Features: 13 columns (Age, Gender, Heart Rate, Stress Level, Activity Level, BMI, BP etc.)
- Target Variable: Quality of Sleep (4-9 scale)
- Data Type: Numerical + Categorical
- Preprocessing:
  - Categorical encoding (Gender, BMI Category, Sleep Disorder)
  - BP split into Systolic & Diastolic
  - Missing-value handling using mean imputation

Why This Dataset?

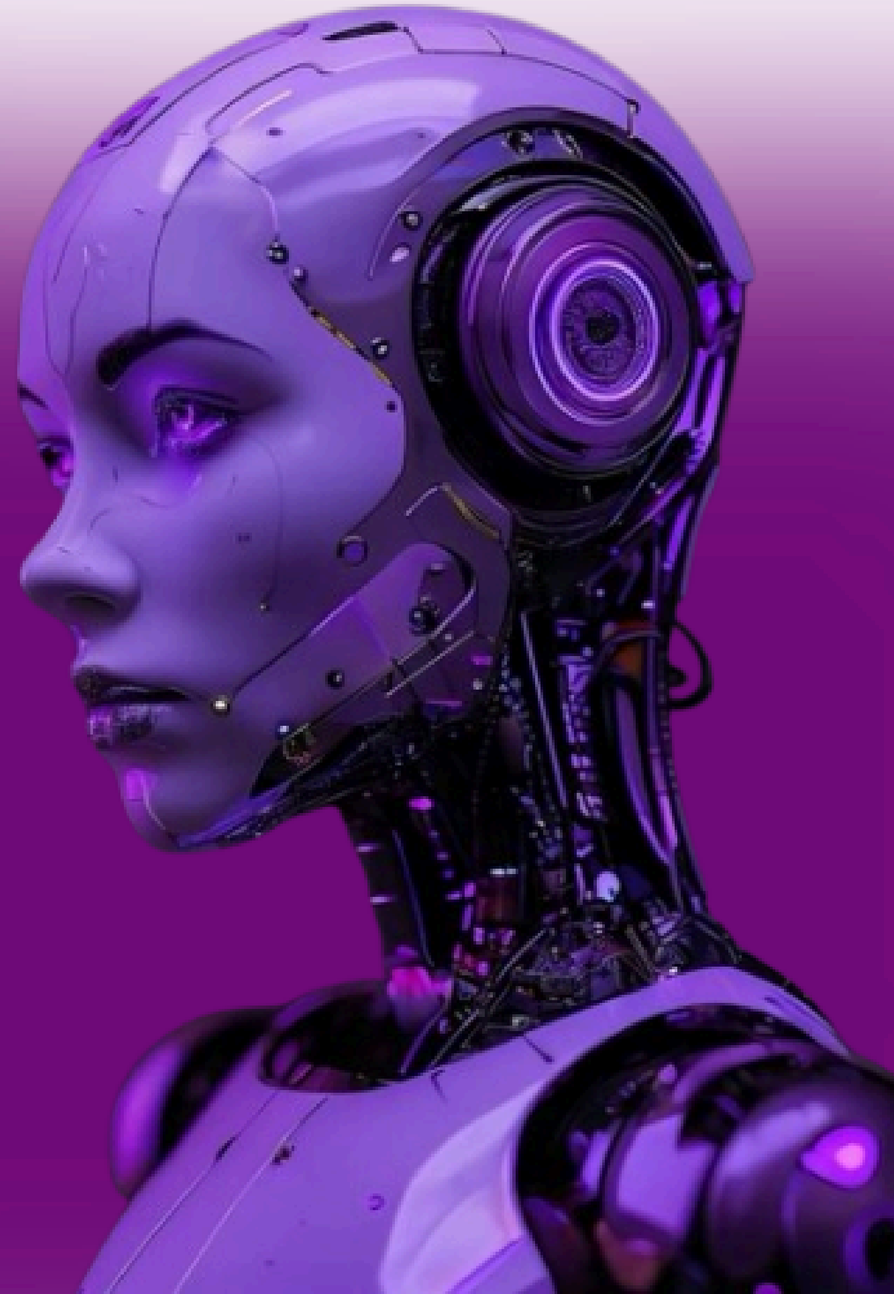
- Contains direct lifestyle indicators
- Ideal for decision-tree explainability
- Balanced and easy to visualize

# METHODOLOGY

## Data Preparation

### Phase 1: Data Preparation

- Loaded and inspected dataset
- Performed cleaning & formatting
- Mapped categorical fields:
  - Gender → 0/1
  - BMI Category → Numerical index
  - Sleep Disorder → Numerical index
- Split Blood Pressure into:
  - Systolic BP
  - Diastolic BP
- Filled missing numeric values with mean
- Selected final features for training
- Train–test split (80–20 stratified)



# METHODOLOGY

**Phase 2: Decision Tree Model**

- **Algorithm: DecisionTreeClassifier (Gini Index)**
  **Trained on 80% of data**
- **Visualized using plot_tree()**
- **Saved the final model using Pickle**

**Why Decision Tree?**

It automatically finds important features

The tree itself learns:

Which feature is most important

Where to split it

**Shows key lifestyle factors visually**

- **Works well with mixed numeric + categorical data**
- **Numbers (Age, Heart Rate, Steps)**
  **Categories (Gender, BMI Category, Sleep Disorder)**
  **Continuous values (Blood Pressure, Sleep Duration)**

# METHODOLOGY

## Frontend Development

### Phase 3: Streamlit Frontend

- Built interactive UI using Streamlit
- User inputs:
  - Age, Gender
  - Sleep duration
  - Stress level
  - Physical activity
  - Heart rate
  - Blood pressure
  - BMI Category
- Backend processes input into numeric format
- Displays:
  - Predicted sleep quality
  - Probability distribution for each quality level

# RESULTS

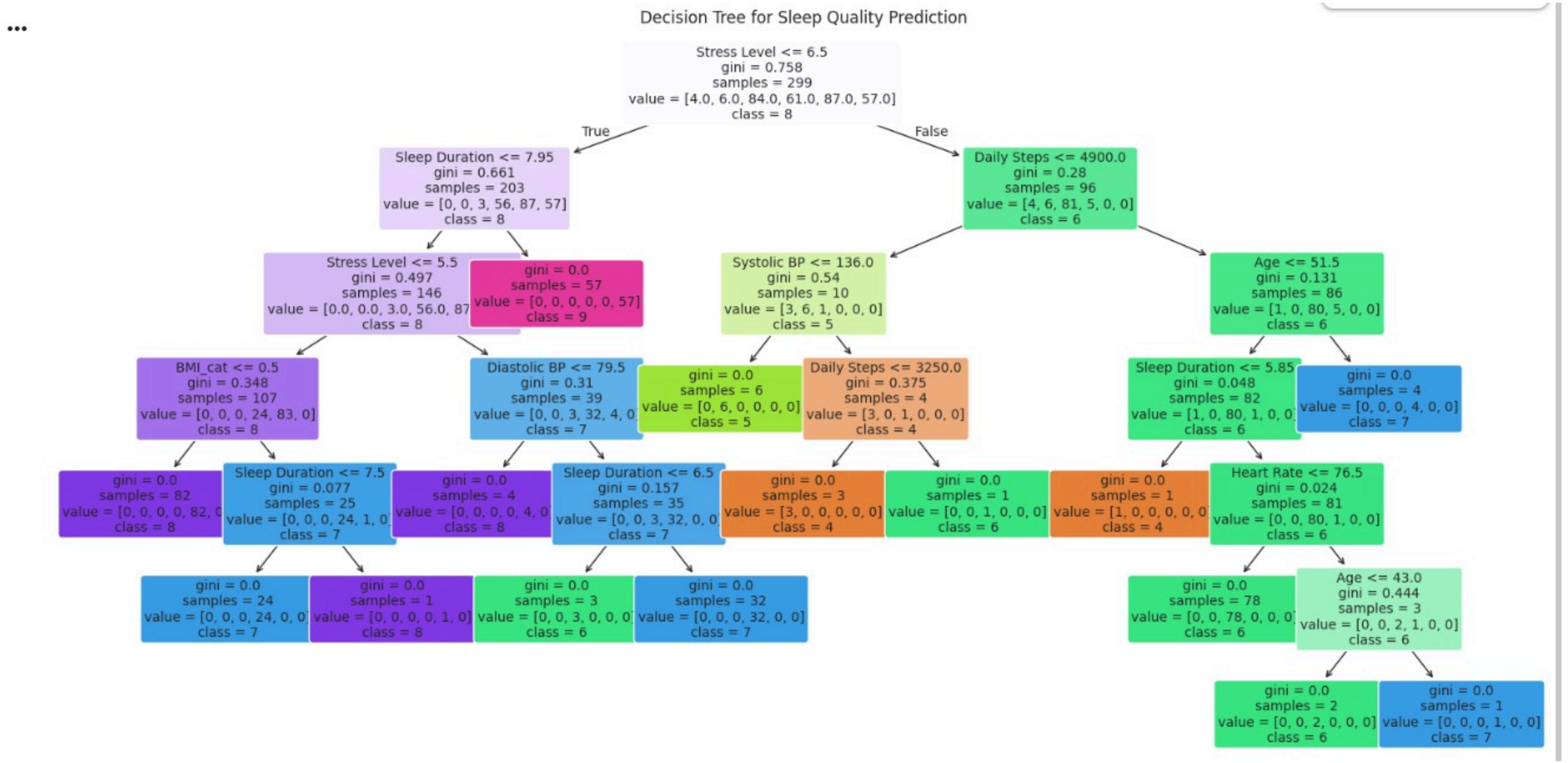
MODEL PERFORMANCE

- STRONG INTERPRETABILITY WITH VISUAL TREE OUTPUT
- BALANCED PREDICTIONS ACROSS CLASSES
- GOOD GENERALIZATION ON UNSEEN DATA

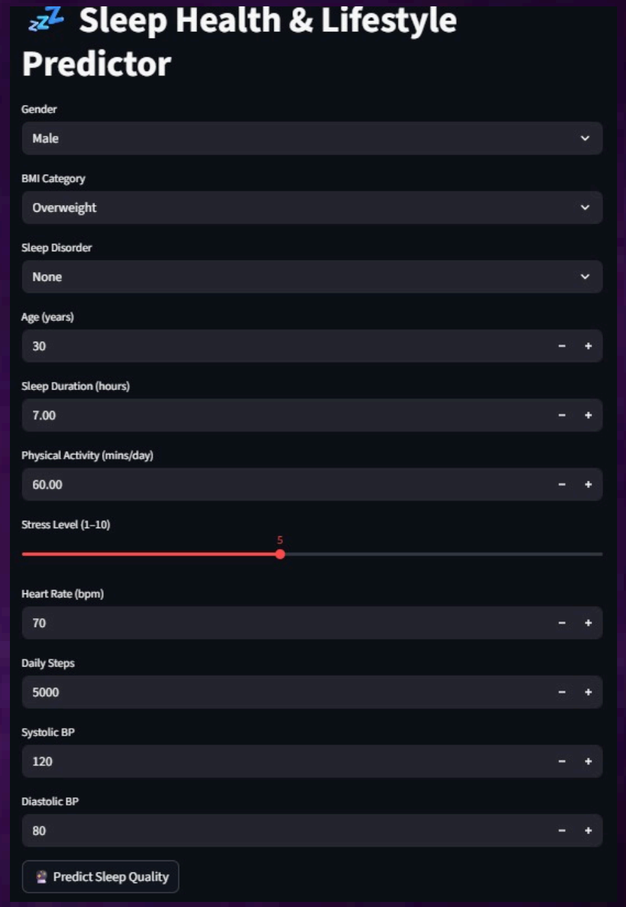
SYSTEM PERFORMANCE

- FAST PREDICTION TIME
- CLEAN END-TO-END PIPELINE
- FRONTEND WORKS SMOOTHLY WITH USER INPUTS
- PROBABILITY SCORES IMPROVE TRANSPARENCY

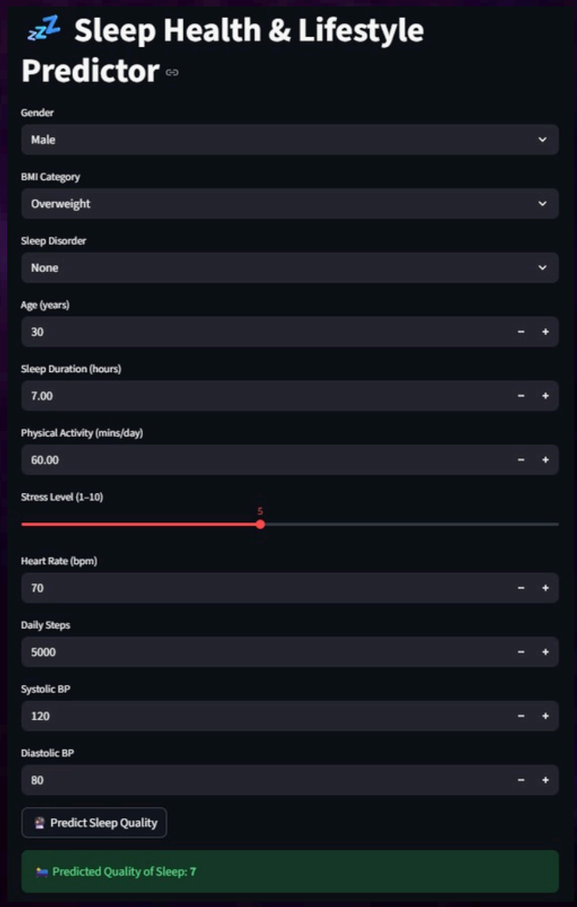
# DECISION TREE VISUALIZATION



# SYSTEM DEMO


Sleep Health & Lifestyle Predictor
Gender
Male
BMI Category
Overweight
Sleep Disorder
None
Age (years)
30
Sleep Duration (hours)
7.00
Physical Activity (mins/day)
60.00
Stress Level (1–10)
5
Heart Rate (bpm)
70
Daily Steps
5000
Systolic BP
120
Diastolic BP
80
Predict Sleep Quality


Predict Sleep Quality
Predicted Quality of Sleep: 7
Probability of Each Sleep Quality Score
• Quality Level 1: 0.00%
• Quality Level 2: 0.00%
• Quality Level 3: 0.00%
• Quality Level 4: 100.00%
• Quality Level 5: 0.00%
• Quality Level 6: 0.00%


Sleep Health & Lifestyle Predictor
Gender
Male
BMI Category
Overweight
Sleep Disorder
None
Age (years)
30
Sleep Duration (hours)
7.00
Physical Activity (mins/day)
60.00
Stress Level (1–10)
5
Heart Rate (bpm)
70
Daily Steps
5000
Systolic BP
120
Diastolic BP
80
Predict Sleep Quality
Predicted Quality of Sleep: 7

# KEY FINDINGS

- SLEEP DURATION, STRESS LEVEL, HEART RATE, AND ACTIVITY HAVE MAJOR INFLUENCE ON SLEEP QUALITY
- BLOOD PRESSURE AND BMI ALSO CONTRIBUTE TO PREDICTION
- DECISION TREE REVEALS CLEAR DECISION BOUNDARIES
- SYSTEM CAN HELP USERS REFLECT ON DAILY HABITS AFFECTING SLEEP

# FUTURE ENHANCEMENTS

SHORT-TERM

- IMPROVE DATA BALANCE WITH AUGMENTATION
- ADD EXPLANATIONS

MEDIUM-TERM

- DEPLOY AS A MOBILE APP
- ADD DASHBOARD FOR LIFESTYLE TRACKING
- ENABLE SLEEP RECOMMENDATION SYSTEM

LONG-TERM

- INTEGRATE SMARTWATCH/HEALTH-BAND SENSORS
- PREDICT SLEEP DISORDERS (INSOMNIA, APNEA)
- BUILD A COMPLETE AI-BASED SLEEP COACHING SYSTEM

# CONCLUSION

- The Sleep Health & Lifestyle Predictor demonstrates how machine learning can analyze lifestyle patterns to estimate sleep quality.

- With a Decision Tree model and a functional Streamlit interface, the system provides interpretable and real-time predictions.

- This project highlights the importance of factors like stress, activity, and sleep duration in determining overall sleep health and lays the foundation for more advanced sleep-wellness solutions.

# REFERENCES

[1] R. BARATI, L. DUFFETT-LEGER, M. MOSHIRPOUR AND M. MOSHIRPOUR, "IMPACT OF STRESS ON SLEEP LEVELS: A COMPARATIVE MACHINE LEARNING STUDY BASED ON WEARABLE DATA," 2023 INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND APPLICATIONS (ICMLA), JACKSONVILLE, FL, USA, 2023, PP. 1194-1199, DOI: 10.1109/ICMLA58977.2023.00179.

[2] M. WARUNLAWAN, P. HOMSUD, P. SAPPHAPHAB, O. RINTHON AND S. PECHPRASARN, "IDENTIFICATION OF CRUCIAL FACTORS IN SLEEP QUALITY USING MACHINE LEARNING MODELS AND MRMR FEATURE SELECTION TECHNIQUE," 2023 15TH BIOMEDICAL ENGINEERING INTERNATIONAL CONFERENCE (BMEICON), TOKYO, JAPAN, 2023, PP. 1-5, DOI: 10.1109/BMEICON60347.2023.10322011.

[3] D. R. SAHU AND R. K. GUPTA, "MACHINE LEARNING TECHNIQUES FOR SLEEP DISORDER CLASSIFICATION AND SLEEP QUALITY ASSESSMENT," 2024 ASIAN CONFERENCE ON INTELLIGENT TECHNOLOGIES (ACOIT), KOLAR, INDIA, 2024, PP. 1-6, DOI: 10.1109/ACOIT62457.2024.10939162.

# THANK YOU!