‘Predicting Stress and Sleep Disorders Using Health and Lifestyle Data’

Objective:

To develop machine learning models that predict stress levels and sleep disorders based on health, lifestyle, and personal attributes, providing early insights.

Team Members:-

1. BHOGARAJU SHANMUKHA SRI KRISHNA
   112201013
2. Bhupathi Varun
   142201034
3. VISHNU SHREERAM M P
   142201014

- **1. Tasks:-**
  - **Data Preprocessing:**
    - Perform feature selection and handle missing values.
    - Normalise numerical data and encode categorical variables.
    - Importance: Ensures data quality and selects the most relevant features for better model performance.
    - Perform EDA on the Dataset to gain insights into the distribution of data.
  - **Classification and Regression:**
    - Build classification models for predicting sleep disorders.
    - Use regression models for predicting stress levels.
      - Importance: Provides insights into health issues, using appropriate models for continuous (stress) and categorical (sleep disorder) predictions.
    - Hyperparameter Tuning:
      - Apply GridSearchCV / RandomizedSearchCV to fine-tune models for optimal performance.
      - Importance: Helps prevent overfitting and enhances accuracy by optimizing model parameters.
    - Ensemble Methods:
      - Test ensemble models (e.g., Random Forest, XGBoost) to improve predictive accuracy.
      - Importance: Combines multiple models to achieve better accuracy and stability.
    - Model Evaluation:
      - Evaluate using metrics such as accuracy for classification and RMSE for regression, with cross-validation.

2. Challenges:

- Trusting the data: Poor data quality can lead to flawed insights and hinder the performance of models.
- Feature Selection vs. Extraction: Balancing complexity and interpretability while reducing overfitting.
- Data Imbalance: Managing imbalanced data in sleep disorder prediction.

3. Tasks checklist:

- ☑ *Data Preprocessing*
- ☐ Model Selection and Evaluation
- ☐ Hyper-Parameter Tuning
- ☐ Ensemble Methods