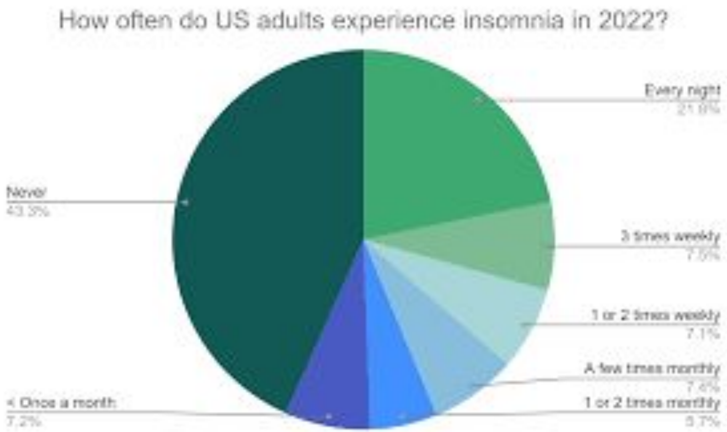
BME 3053C – Computer Applications for BME
Final Project

Sleep Stage Classification for Insomnia Diagnosis Using EEG Data

Group 12 - Sierra, Brianna, Dana, Jennifer, Amber, Ariel

PROBLEM STATEMENT

PROBLEM	Insomnia affects 10-15% of adults, causing health issues (depression, reduced productivity).
CHALLENGE	Polysomnography is invasive, costly, and not scalable for at-home use.
GOAL	Develop a non-invasive EEG-based system to classify sleep stages (Wake, N1, N2, N3, REM) for accurate insomnia diagnosis.



DATASET

- **Source:** Sleepy Driver Data Set (kaggle).
- **Description:** EEG recordings (Fpz-Cz, Pz-Oz channels, 100 Hz) from 4 subjects .
- **Annotations:** 8 minute readings watching the switch between NREM and REM sleep
- **File:** acquiredDataset.csv

Motivation

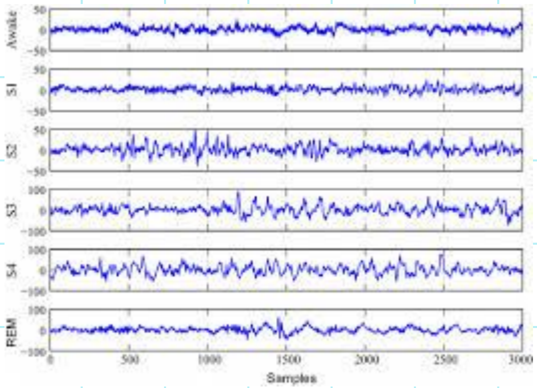
College students commonly experience sleep deprivation.

Insomnia affects:

- Productivity
- Mental health (depression, anxiety)
- Long-term health risks

Better sleep stage monitoring → tailored treatment plans.

Research



Current practices:

- Polysomnography (EEG, EOG, EMG)
- Self-report questionnaires (Insomnia Severity Index)
- Manual sleep stage scoring (time-consuming, subjective)

Machine learning advancements:

- CNNs & RNNs → 80–90% accuracy
- Public datasets (Sleep-EDF, Sleepy Driver)

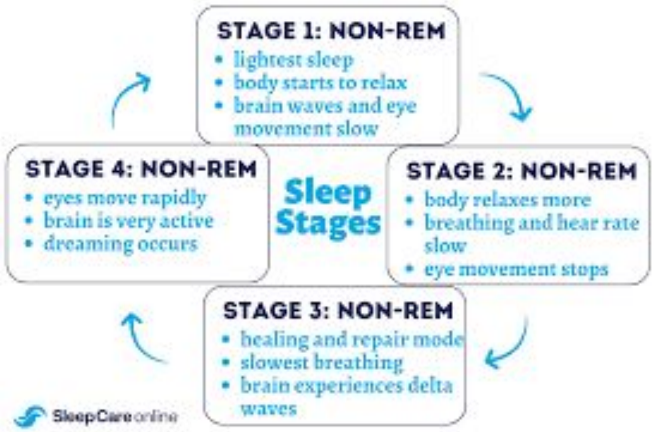
Looking at the Data

Dataset: Sleepy Driver EEG Dataset
Features:

- EEG frequency bands
  - Delta, Theta, Alpha, Beta, Gamma
- Attention & meditation scores

Labels:

- Sleep stages:
  - 0 = Non-REM
  - 1 = REM



Proposed Solution

- Develop a **Random Forest Classifier** for sleep stage classification.
  - a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting

Potential improvements:

- Future use of **Convolutional Neural Networks (CNNs)** for deeper feature extraction.

- Focus on early sleep process analysis to assess insomnia-related issues.

What our Model Does

Data Preprocessing

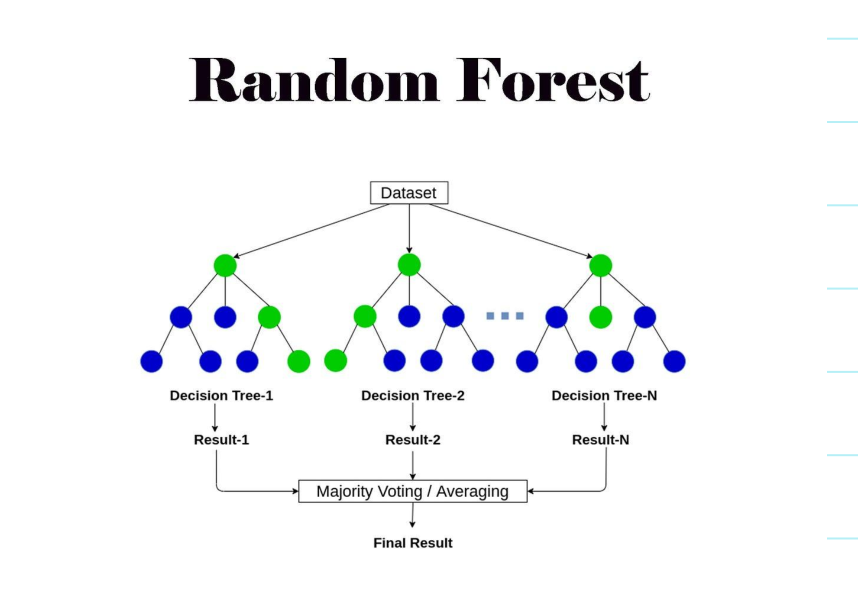
- Feature extraction, Label encoding, and Train-test split (80/20)

Model Training

- Random Forest Classifier (100 trees

Evaluation

- Accuracy
- Precision, Recall, F1-Score



Results

Class	Precision	Recall	F1 Score	Support
0	0.80	0.84	0.82	438
1	0.76	0.70	0.73	309

Discussion

- **Strengths:**
  - Solid baseline accuracy
  - Meaningful features (EEG bands, attention, meditation)
  - Less resource-intensive compared to polysomnography
- **Limitations:**
  - Slight class imbalance (Class 1 underrepresented)
  - Room for improvement with deeper models (CNNs, RNNs)
  - Only has the difference between NREM and REM cycles- future work could decipher between all EEG data

Real World

Improved insomnia diagnosis using EEG-based sleep stage classification.

Personalized treatment plans:

- Sleep aids
- Behavioral therapies
- Lifestyle modifications

Potential for integration with wearable devices for continuous sleep monitoring.

Conclusion

- Developed a machine learning model to classify sleep stages for insomnia diagnosis.
- Achieved **78% accuracy** using EEG frequency bands and attention/meditation features.
- Model lays the foundation for accessible sleep monitoring and improved insomnia treatment.