End Sem Project Seminar

Sem 3 MTech AI

Presented by: Tanmay Rathod

Under the guidance: Dr. Santosh Kumar Satapathy

Assistant Professor, Information and Communication Technology

Pandit Deendayal Energy University (PDEU)

December 2024

TABLE OF CONTENTS

Abstract

- Sleep stage classification is crucial for diagnosing sleep-related disorders.
- Processing raw EEG signals is computationally expensive and resource-intensive, especially with large datasets.

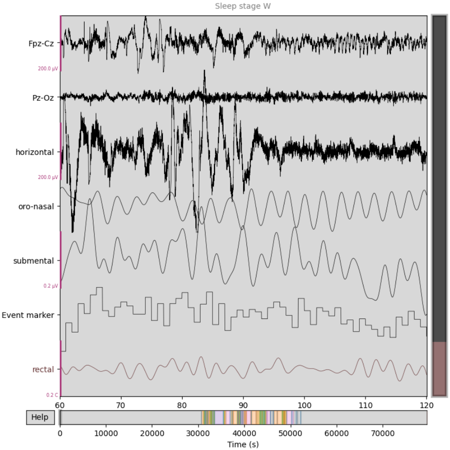

Figure: Visualization of EEG signal.

Abstract

- This article discusses a method for preparing the Sleep-EDF dataset:
  - Extracting, segmenting, and labeling PSG data.
  - Converting data into Python pickle (.pkl) format for easy handling with deep learning frameworks.
- Used Annotations descriptions:

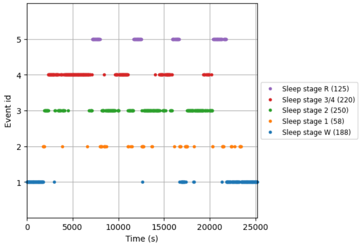

Figure: Sleep stage event plot.

Sleep Stages and Characteristics

Sleep Stage	Frequency Range (Hz)	Description
Wake (Beta)	12-30	Active, alert state; engaged in cognitive activities.
N1 (Light Sleep)	4-8	Transition stage between wakefulness and sleep; easy to wake up.
N2 (Moderate Sleep)	4-6	Sleep spindles and K-complexes
N3 (Deep Sleep)	0.5-4	Slow-wave sleep; restorative processes occur.
REM (Theta)	4-6	Associated with dreaming; brain activity resembles wakefulness.

Problem Statement

Multi-Channel EEG Analysis for Advanced Sleep Disorder Characterization Using Deep Learning

Motivation

- Enhanced efficiency and scalability for sleep disorder diagnosis.
- Integration potential with consumer-grade devices like wearables for sleep monitoring.
- Opportunities to improve sleep quality analysis using advanced deep learning techniques and personalized health recommendations & disease classification.

Literature Survey - Part 1

No.	Author/Title/Journal	Technique(s) Used	Database Used	Advantages/Limitations
1	Yan, R., Zhang, C., Spruyt, K., Wei, L., Wang, Z., Tian, Cong, F. (2019).	Multi-modality PSG signals + Cyclic Alternating Pattern (CAP)	PhysioNet Database	Advantages: Automatic sleep scoring methods by fusing four modalities of PSG signals. Limitations: Stage S1 is often misclassified as wakefulness and REM by automatic sleep scoring.
2	Zhou, J., Tian, Y., Wang, G., Liu, J., Wu, D., Xu, W., Hu, Y. (2020).	Single-channel EEG signal + Stacked Ensemble Model	Sleep-EDF Expanded	Advantages: Effective sleep stage classification. Limitations: Only considered healthy controlled subjects. Not analyzed EEG signals in detail.

Literature Survey - Part 2

No.	Author/Title/Journal	Technique(s) Used	Database Used	Advantages/Limitations
3	Shen, H., Ran, F., Xu, M., Guez, A., Li, A., Guo, A. (2020).	Improved model based on essence features + Grid-search strategy	ISRUC-Sleep dataset	Advantages: Improved model for sleep staging. Limitations: Misclassification ratio was higher for S2 stage.
4	Huang, W., Guo, B., Shen, Y., Tang, X., Zhang, T., Li, D., Jiang, Z. (2019).	Multi-channel data adding + Multi-feature screening	Sleep-EDF dataset	Advantages: Multi-channel signal superposition method reduces noise. Limitations: Ineffective performance with heterogeneous signal combinations.

Methodology

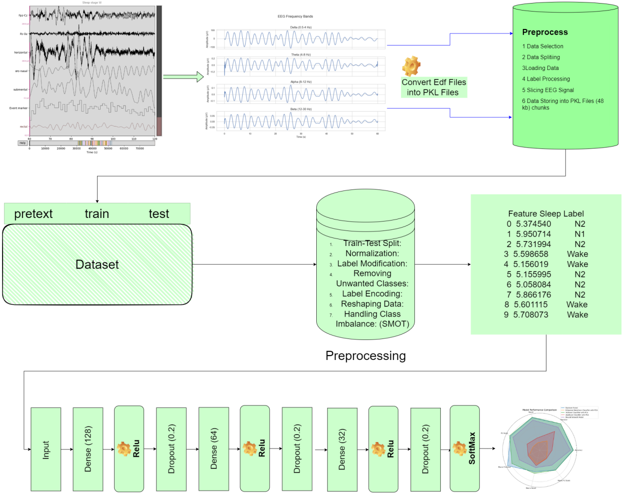

Figure: Proposed Architecture

Methodology

Preprocessing Workflow

- **Loading Data**: Sleep stage data is loaded from '.pkl' files in preprocessed directories.
- **Handling Test Sets**: Ensured test data availability by splitting the training set if necessary.
- **Normalization**: Standardized input features to have zero mean and unit variance.
- **Label Encoding**: Converted sleep stage labels (e.g., 'W', 'N1', 'N2') into numerical form.
- **One-Hot Encoding**: Transformed numerical labels into one-hot vectors for classification tasks.
- **Reshaping**: Reshaped input features for compatibility with machine learning models.

Methodology

Sleep Stages Data Visualization

Table: Summary of Sleep Stage Data

Feature (x)	Label (y)
[0.059, 0.596, -0.193, ..., -0.601, 0.201]	W
[-0.022, -0.107, -0.135, ..., 0.038, 0.103]	W
[... (Additional data)]	N1

Note: Data truncated for brevity, showing representative examples of preprocessed input and labels.

Methodology: Class Mapping

Class Mapping Process

- Original sleep stage labels were transformed using a predefined mapping:
  - `'1'` → `'N1'` (Light Sleep)
  - `'2'` → `'N2'` (Intermediate Sleep)
  - `'3'`, `'4'` → `'N3'` (Deep Sleep)
  - `'W'` → `'Wake'`
  - `'R'` → `'REM'` (Rapid Eye Movement)
  - `'e'` → Removed (unnecessary class)

Mapping ensures consistency in class representation across the dataset.

Methodology: Cleaning and Encoding

Cleaning and Label Encoding

- **Mask Application:**
  - Instances of unnecessary classes (`'e'`) were removed using masks.
- **Label Encoding:**
  - Transformed remaining classes into numerical labels using `LabelEncoder`.

Cleaning and encoding prepared the data for machine learning models.

Methodology : Neural Network Architecture

Table: Simple Neural Network Architecture

Layer Type	Neurons/Units	Activation Function
Input Layer	`X_train`	-
Dense Layer	128	ReLU
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	64	ReLU
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	32	ReLU
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	Number of classes (e.g., 5)	Softmax

Methodology: Simple RNN Neural Network Architecture

Table: Simple RNN Neural Network Architecture for Sleep Disorder

Layer Type	Neurons/Units	Activation Function
Input Layer	`(X_trains)`	-
RNN Layer 1	128	Tanh
Dropout Layer	-	0.2 (Dropout Rate)
RNN Layer 2	64	Tanh
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	64	ReLU
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	32	ReLU
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	Number of classes	Softmax

Methodology: LSTM Neural Network Architecture

Table: LSTM Neural Network Architecture

Layer Type	Neurons/Units	Activation Function
Input Layer	`(X_train)`	-
LSTM Layer 1	128	Tanh
Dropout Layer	-	0.2 (Dropout Rate)
LSTM Layer 2	64	Tanh
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	64	ReLU
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	32	ReLU
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	Number of classes (e.g., 5)	Softmax

Methodology: Bidirectional LSTM Neural Network Architecture

Table: Bidirectional LSTM Neural Network Architecture for Sleep Disorder

Layer Type	Neurons/Units	Activation Function
Input Layer	(X_train)	-
Bi-LSTM Layer 1	128	Tanh
Dropout Layer	-	0.2 (Dropout Rate)
Bi-LSTM Layer 2	64	Tanh
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	64	ReLU
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	32	ReLU
Dropout Layer	-	0.2 (Dropout Rate)
Dense Layer	Number of classes	Softmax

Simple Neural Network : Training Vs Loss Plot

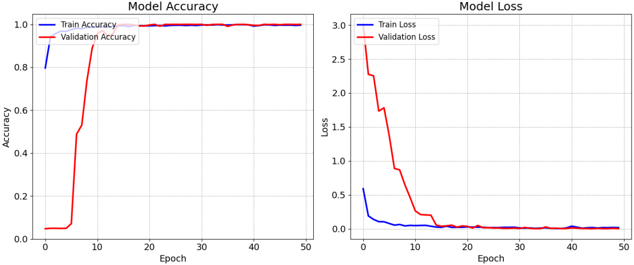

Figure: Training Loss Plot for Deep Learning Model

Model Training Results

Final Epoch: 259/259

Accuracy: 0.9955 **Loss:** 0.0208

Validation Accuracy: 0.9995 **Validation Loss:** 0.0067

RNN Model: Training Vs Loss Plot

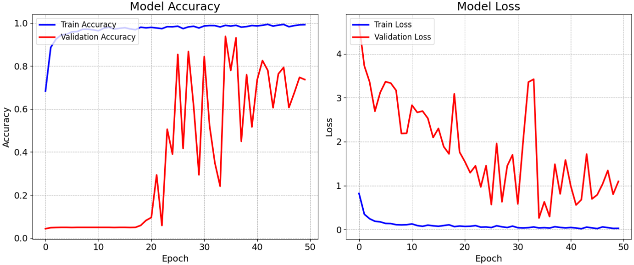

Figure: Training Loss Plot for RNN Model

RNN Model Training Results

Final Epoch: 259/259

Accuracy: 0.9855 **Loss:** 0.0508

Validation Accuracy: 0.9975 **Validation Loss:** 0.0157

LSTM Model: Training Vs Loss Plot

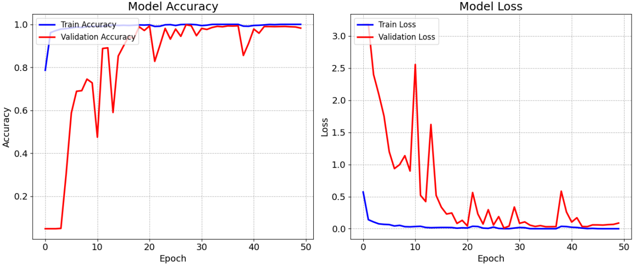

Figure: Training Loss Plot for LSTM Model

LSTM Model Training Results

Final Epoch: 259/259

Accuracy: 0.9900 **Loss:** 0.0182

Validation Accuracy: 0.9990 **Validation Loss:** 0.0080

BiLSTM Model: Training Vs Loss Plot

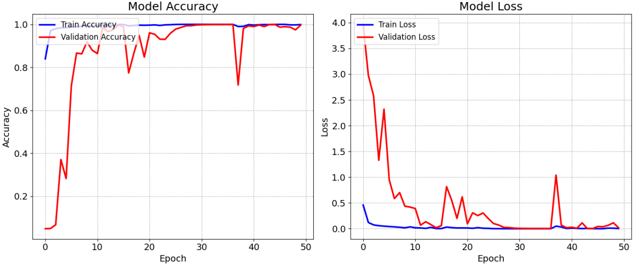

Figure: Training Loss Plot for BiLSTM Model

BiLSTM Model Training Results

Final Epoch: 259/259

Accuracy: 0.9925 **Loss:** 0.0124

Validation Accuracy: 0.9992 **Validation Loss:** 0.0056

Machine Learning Model Performance Comparison

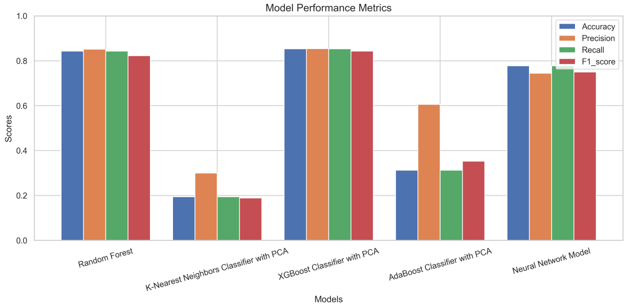

Figure: Comparison of Accuracy, Precision, Recall, and F1 Score for Different Models

Neural Network Model Performance Comparison

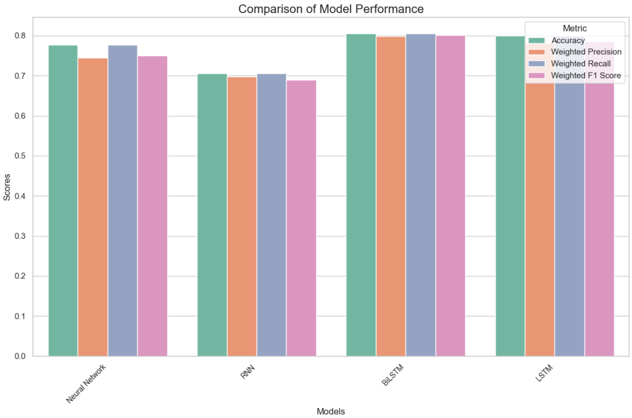

Figure: Comparison of Accuracy, Precision, Recall, and F1 Score for Different Models

Recent Achievement: Paper Accepted at IATMSI-2025

Paper Title

Automated Sleep Staging System using EEG Signal Feature-Based Classification by Machine Learning Techniques

Conference Details

Conference: IEEE IATMSI-2025
Venue: ABV-IIITM Gwalior, India (Hybrid Mode)
Dates: 06th – 08th March 2025

Future Plan: Main Problem Statement

1. Transfer Learning for Cross-Dataset Sleep Stage Classification Using EEG

We aim to explore transfer learning techniques to apply pre-trained models on different datasets to enhance the accuracy of sleep stage classification using EEG signals.

2. Personalized Sleep Quality Assessment Using EEG-based Deep Learning Models

We aim to build personalized sleep quality assessment models by leveraging EEG-based deep learning methods. This will ensure better customization for individuals with different sleep patterns.

Future Plan: Research Goals

1. Transfer Learning with CNN for Sleep EDF Data

We are planning to apply transfer learning techniques using CNN neural networks on the Sleep EDF dataset. The goal is to improve classification accuracy by fine-tuning pre-trained models for EEG signal processing.

2. Exploration of CNN, LSTM, and Hybrid Models

We aim to investigate various deep learning models like CNN, LSTM, and hybrid models to better capture temporal dependencies in the data, improving performance beyond existing benchmarks.

Conclusion: Key Achievements

Application of Machine Learning and Deep Learning

We have successfully applied both machine learning and deep learning techniques on the Sleep EDF EEG signals, achieving reliable results with simple yet effective architectures.

Strength of Simple Architectures

Our analysis demonstrates the strength of using simple architectures in deep learning models, showing that even basic models can achieve significant improvements in performance.

Conclusion: Approaches and Findings

Direct EEG Signals for Machine Learning

We initially applied machine learning models directly on EEG signals, yielding promising results and highlighting the potential of raw EEG data for classification tasks.

Signal-to-PKL Conversion for Model Optimization

Another key finding was the conversion of EEG signals into PKL files, which allowed us to treat the data as simpler inputs. This method enabled more efficient model training and optimization, particularly in devices requiring lower computational resources.

Conclusion: Limitations and Future Directions

Limitations of Direct Signal Approach

Despite the promising results, we faced some limitations with the direct EEG signal approach, particularly in terms of signal quality and noise interference.

Future Plan: Hybrid Models and Transfer Learning

To overcome these limitations, we plan to explore hybrid models and fine-tune transfer learning techniques to further improve accuracy and generalize the model's performance across different datasets.

References

- Yan, R., Zhang, C., Spruyt, K., Wei, L., Wang, Z., Tian, L., … Cong, F. (2019). Multi-modality of polysomnography signals' fusion for automatic sleep scoring. *Biomedical Signal Processing and Control, 49*, 14–23. doi:10.1016/j.bspc.2018.10.001
- Zhou, J., Tian, Y., Wang, G., Liu, J., Wu, D., Xu, W., … Hu, Y. (2020). Automatic Sleep Stage Classification with Single Channel EEG Signal Based on Two-layer Stacked Ensemble Model. *IEEE Access, 1–1*. doi:10.1109/access.2020.2982434
- Shen, H., Ran, F., Xu, M., Guez, A., Li, A., & Guo, A. (2020). An Automatic Sleep Stage Classification Algorithm Using Improved Model Based Essence Features. *Sensors, 20*(17), 4677. doi:10.3390/s20174677
- Huang, W., Guo, B., Shen, Y., Tang, X., Zhang, T., Li, D., & Jiang, Z. (2019). Sleep staging algorithm based on multichannel data adding and multi feature screening. *Computer Methods and Programs in Biomedicine, 105253*. doi:10.1016/j.cmpb.2019.105253
- Ghimatgar, H., Kazemi, K., Helfroush, M. S., & Aarabi, A. (2019). An automatic single-channel EEG-based sleep stage scoring method based on hidden Markov model. *Journal of Neuroscience Methods, 108320*. doi:10.1016/j.jneumeth.2019.108320
- Ghimatgar, H., Kazemi, K., Helfroush, M. S., Pillay, K., Dereymaeker, A., Jansen, K., … Aarabi, A. (2020). Neonatal EEG sleep stage classification based on deep learning and HMM. *Journal of Neural Engineering*. doi:10.1088/1741-2552/ab965a

Questions and Answers

Thank You!