

**Sleep Stage Prediction via Deep Learning using Actigraphy Data to Detect Disorders by Modeling Sleep**

by

Omar Barazanji

A thesis submitted to the Graduate Faculty of  
Auburn University  
in partial fulfillment of the  
requirements for the Degree of  
Master of Science

Auburn, Alabama  
May 10, 2022

Keywords: Deep Learning, Time-Series Forecasting, Health, Sleep

Copyright 2022 by Omar Barazanji

Approved by

SueAnne Griffith, Professor of Electrical and Computer Engineering

## Abstract

Sleep scoring is a critical task in sleep medicine that involves the labeling of a patient's sensor data from electroencephalography (EEG), electrocardiography (ECG), and electrooculography (EOG) into different 5 different sleep stages. In this study, we compare the performance of three deep learning models for sleep scoring using actigraphy data and sleep study data. Actigraphy data was used solely for sleep/wake prediction, while sleep lab data (EEG, ECG, and EOG) was used for sleep scoring. This study aims to compare the accuracy of these models, with a focus on identifying the optimal model for each task. For actigraphy sleep/wake prediction, the dataset consisted of a total of 109 subjects, and each subject's data was divided into training and testing sets. For the sleep lab dataset, the dataset consisted of 30 subjects each with 5 nights of sleep data. The results show that the CNN-LSTM model achieved the highest accuracy for sleep staging. The XGBoost and LSTM models achieved accuracies of 82.3% and 79.6%, respectively. These findings suggest that the CNN-LSTM model may be a promising approach for sleep staging using actigraphy data and sleep lab data.

## Acknowledgments

The National Sleep Research Resource was supported by the U.S. National Institutes of Health, National Heart Lung and Blood Institute (R24 HL114473, 75N92019R002).

## Table of Contents

Abstract . . . . .	ii
Acknowledgments . . . . .	iii
List of Figures . . . . .	vi
1 Introduction . . . . .	1
1.1 Objective . . . . .	2
1.2 Datasets . . . . .	3
1.2.1 Sleep Actigraphy from the Urban Poor in India . . . . .	3
1.2.2 Sleep as Android Watch Actigraphy . . . . .	3
2 Methods . . . . .	5
2.1 Sleep/Wake Prediction . . . . .	5
2.2 Sleep Stage Prediction . . . . .	6
2.3 Disorder Prediction . . . . .	6
3 Results . . . . .	7
3.1 Sleep/Wake Results . . . . .	7
3.1.1 XGBoost . . . . .	7
3.1.2 SVM . . . . .	7
3.1.3 LSTM . . . . .	7
3.1.4 Transformer . . . . .	7
3.2 Sleep Stage Results . . . . .	7
3.2.1 XGBoost . . . . .	8
3.2.2 SVM . . . . .	8
3.2.3 LSTM . . . . .	8
3.2.4 Transformer . . . . .	8

4	Conclusion . . . . .	11
4.1	Discussion . . . . .	11
4.2	Conclusion . . . . .	11
	Bibliography . . . . .	12
	Appendices . . . . .	13
A	Source Code . . . . .	14

## List of Figures

1.1	Sleep Actigraphy from the Urban Poor in India[4][5] . . . . .	4
1.2	Sleep as Android Watch Actigraphy . . . . .	4
2.1	LSTM Model with 256 units for Sleep/Wake Prediction . . . . .	6
3.1	Loss on Model with 256 units for Sleep/Wake Prediction during Training . . . .	9
3.2	Accuracy on Model with 256 units for Sleep/Wake Prediction during Training .	10

## Chapter 1

### Introduction

Sleep is a fundamental biological process that is critical for maintaining overall health and wellbeing. Sleep disorders, such as sleep apnea and insomnia, can have significant negative impacts on an individual's quality of life and can lead to a range of adverse health outcomes. Sleep scoring, the process of identifying and labeling the different stages of sleep, is an essential tool in the diagnosis and treatment of sleep disorders.

Traditionally, sleep scoring has been done manually by trained clinicians. However, the process is time-consuming and subject to inter-rater variability, which can lead to inaccuracies. Recent advances in deep learning models have shown promise in automating the sleep scoring process, making it faster and more accurate.

In this study, we aim to compare the performance of three deep learning models for sleep scoring using both actigraphy data and sleep study data. Actigraphy data is a non-invasive method of measuring activity levels that is often used to track sleep/wake cycles. Our study will use actigraphy data solely for sleep/wake prediction, while sleep study data (EEG, ECG, and EOG) will be used for sleep scoring.

The study aims to identify the optimal model for each task and compare their accuracies. We will evaluate the performance of the CNN-LSTM, XGBoost, and LSTM models on a dataset of 50 subjects, with each subject's data divided into training and testing sets. Our results will provide insights into the potential of deep learning models for sleep scoring using both actigraphy and sleep study data.

## 1.1 Objective

In this study, we used datasets on regular sleep activity to train a model that can accurately predict and label sleep/wake from actigraphy. We then will create a custom dataset on patients with known disorders that affect sleep. Through a mix of public datasets and sleep studies, we can add labels for each person's known disorder. With the disorder labels on the actigraphy data, our trained model will be able to predict if a person might have certain disorders or not, such as Insomnia.



## **1.2 Datasets**

We used various datasets in this study and will be referenced throughout the paper. The first dataset is Sleep Actigraphy from the Urban Poor in India, and it contains 3-axis accelerometer data from 30 participants. Data is labeled with sleep status, either a 1 or a 0 for asleep or awake, respectively. The other dataset comes from Sleep as Android Watch Actigraphy. This dataset contains labels for sleep stages as well as lux (light in lumens), snoring, bpm, and more. Data is collected from a compatible Wear OS device during sleep and is converted into a single actigraphy curve for sleep stage prediction.

### **1.2.1 Sleep Actigraphy from the Urban Poor in India**

Need to add...

### **1.2.2 Sleep as Android Watch Actigraphy**

Need to add...

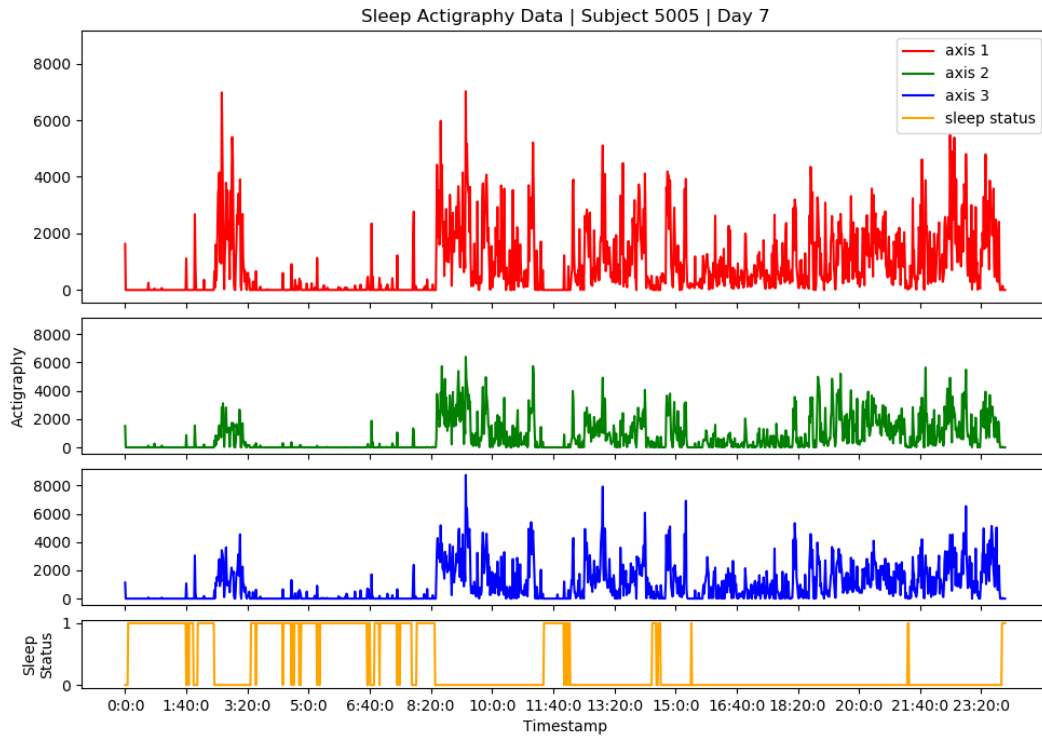


Figure 1.1: Sleep Actigraphy from the Urban Poor in India[4][5]

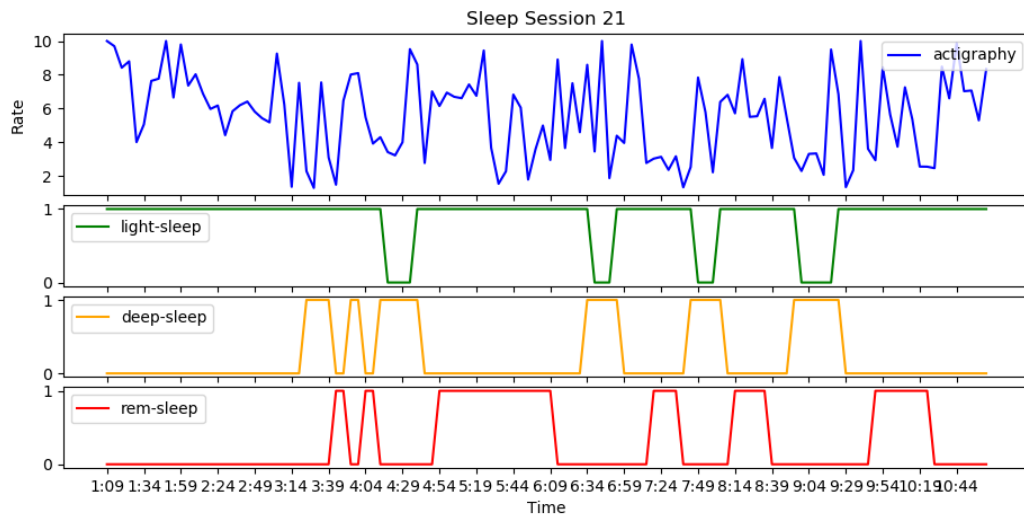


Figure 1.2: Sleep as Android Watch Actigraphy

## Chapter 2

### Methods

We present various methods for sleep/wake prediction, sleep stage prediction, and disorder prediction. We explore popular known methods using XGBoost, SVM, and RNN model architectures. In this paper we propose a new method for sleep-stage classification and sleep/wake classification. Our baseline method will be the vanilla LSTM and we will compare this against our newly proposed methods using the Transformer architecture and another model using a Bidirectional LSTM.

#### 2.1 Sleep/Wake Prediction

Recurrent Neural Networks (RNNs) are a special type of neural network that are designed to process long sequences of data and predict how a series of data should continue per time step. The Long Short-Term Memory (LSTM) network is a refined version of the vanilla RNN that can "forget" useless information in training data. This prevents the network from overfitting, which is a big problem in the vanilla RNN. Looking at sleep actigraphy data as a time series problem makes the prediction of any label from the training data quite simple. For example, figure 1.1 above shows the x, y, and z activity data from the accelerometers on a participant's wrist, and the sleep status (1=asleep, 0=awake) shows when the person is asleep or awake in time with the accelerometer. We can sub-sample all 4 labels together with a window size of  $T$ , usually around  $T=30$ , and predict sleep status or anyother variable that was recorded along with the accelerometer per time step.

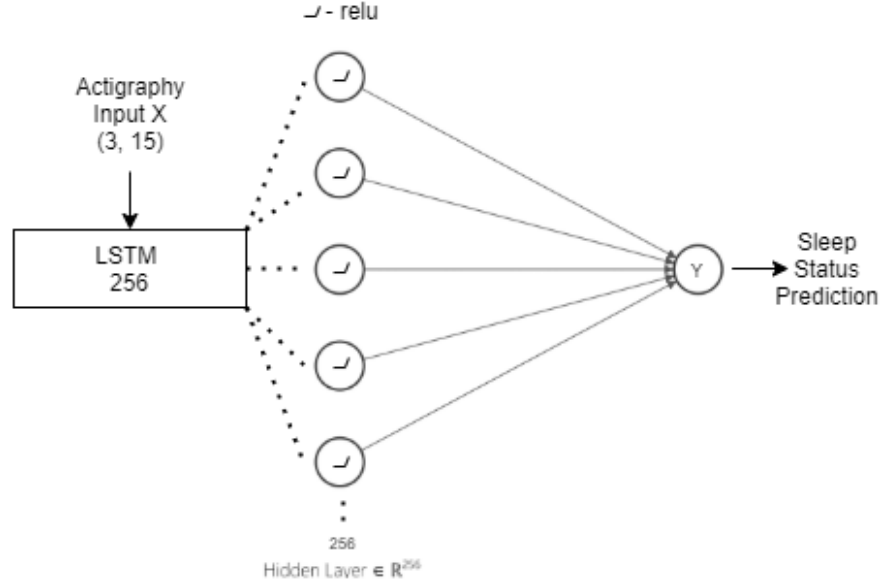


Figure 2.1: LSTM Model with 256 units for Sleep/Wake Prediction

## 2.2 Sleep Stage Prediction

To classify sleep stages as multiple labels during sleep from actigraphy data, we will need to use a multi-variate LSTM that can output more than one label per time step. To do this we will update our vanilla LSTM model's output  $Y$  from figure 2.1 to have a dimension equal to the number of desired sleep stage labels. For our dataset shown in 1.2, we can predict light, deep, and REM sleep stages. We can also add sleep/wake from the Sleep as Android dataset to have a total of 4 output labels per timestep.

## 2.3 Disorder Prediction

Need to add...

## Chapter 3

### Results

Results for sleep/wake predcition using the LSTM network with 256 units can be seen below in figures 3.1 and 3.2. The model performs at 92 percent accuracy with 256 units and 15 training epochs.

#### **3.1 Sleep/Wake Results**

Need to add...

##### **3.1.1 XGBoost**

Need to add...

##### **3.1.2 SVM**

Need to add...

##### **3.1.3 LSTM**

Need to add...

##### **3.1.4 Transformer**

Need to add...

#### **3.2 Sleep Stage Results**

Need to add...

### **3.2.1 XGBoost**

Need to add...

### **3.2.2 SVM**

Need to add...

### **3.2.3 LSTM**

Need to add...

### **3.2.4 Transformer**

Need to add...

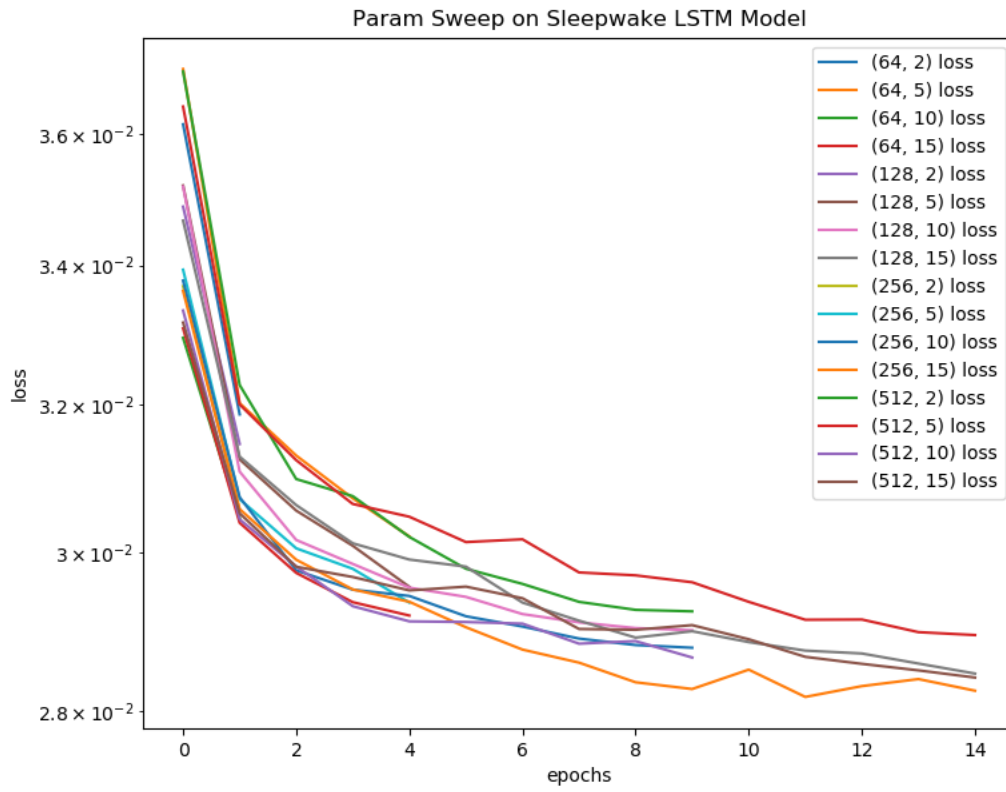


Figure 3.1: Loss on Model with 256 units for Sleep/Wake Prediction during Training

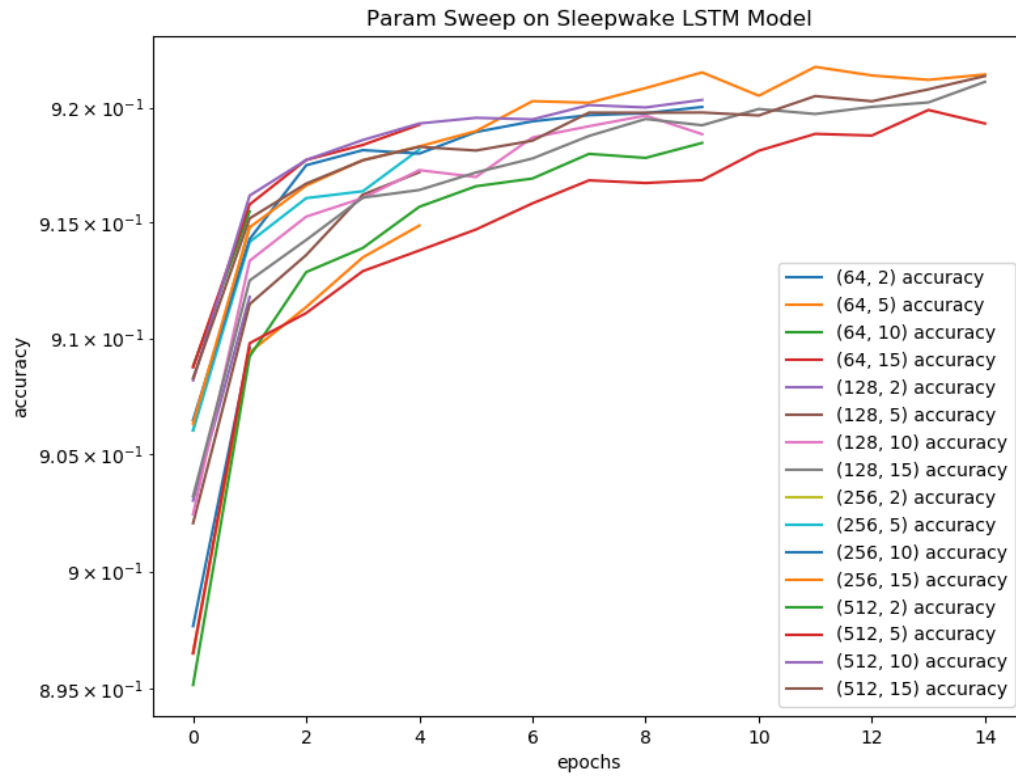


Figure 3.2: Accuracy on Model with 256 units for Sleep/Wake Prediction during Training



## Chapter 4

### Conclusion

Need to add...

#### **4.1 Discussion**

Need to add...

#### **4.2 Conclusion**

Need to add...

## Bibliography

- [1] C. Iber, S. Ancoli-Israel, A. L. Chesson, and S. F. Quan, The AASM manual for the scoring of sleep and associated events : rules, terminology, and technical specifications. Istchester, IL: American Academy of Sleep Medicine, 2007.
- [2] Guillot, A., Sauvet, F., During, E., Thorey, V. Dreem Open Datasets: Multi-Scored Sleep Datasets to compare Human and Automated sleep staging, 2019.
- [3] Cho, J.H.; Choi, J.H.; Moon, J.E.; Lee, Y.J.; Lee, H.D.; Ha, T.K. Validation Study on Automated Sleep Stage Scoring Using a Deep Learning Algorithm. *Medicina* 2022, 58, 779.
- [4] Zhang GQ, Cui L, Mueller R, Tao S, Kim M, Rueschman M, Mariani S, Mobley D, Redline S. The National Sleep Research Resource: towards a sleep data commons. *J Am Med Inform Assoc.* 2018 Oct 1;25(10):1351-1358. doi: 10.1093/jamia/ocy064. PMID: 29860441; PMCID: PMC6188513.
- [5] Bessone P, Rao G, Schilbach F, Schofield H, Toma M. The Economic Consequences of Increasing Sleep Among the Urban Poor. *Q J Econ.* 2021 Apr 8;136(3):1887-1941. doi: 10.1093/qje/qjab013. PMID: 34220361; PMCID: PMC8242594.
- [6] A. Sano, W. Chen, D. Lopez-Martinez, S. Taylor and R. W. Picard, "Multimodal Ambulatory Sleep Detection Using LSTM Recurrent Neural Networks," in *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1607-1617, July 2019, doi:10.1109/JBHI.2018.2867619.
- [7] Sathyanarayana, Aartietal. "Sleep Quality Prediction From Wearable Data Using Deep Learning." *JMIR mHealth and uHealth* vol. 4, 4e125.4Nov.2016, doi:10.2196/mhealth.6562.
- [8] Chen, Tianqi, Carlos Guestrin. 'XGBoost'. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* ACM, 2016. Web.
- [9] Shantilal, K. D. Donohue and B. F. O'Hara, "SVM for automatic rodent sleep-wake classification," *IEEE SoutheastCon 2008*, 2008, pp. 581-586, doi: 10.1109/SECON.2008.4494360.
- [10] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. 'Attention Is All You Need'. *arXiv*, 2017.

## Appendices

Appendix A  
Source Code

Code link