Dimension Reduction Methods Applied to Sleep Stage Analysis

Zekuan Liu

liuzekuan@bupt.edu.cn

May 18, 2021

Outline

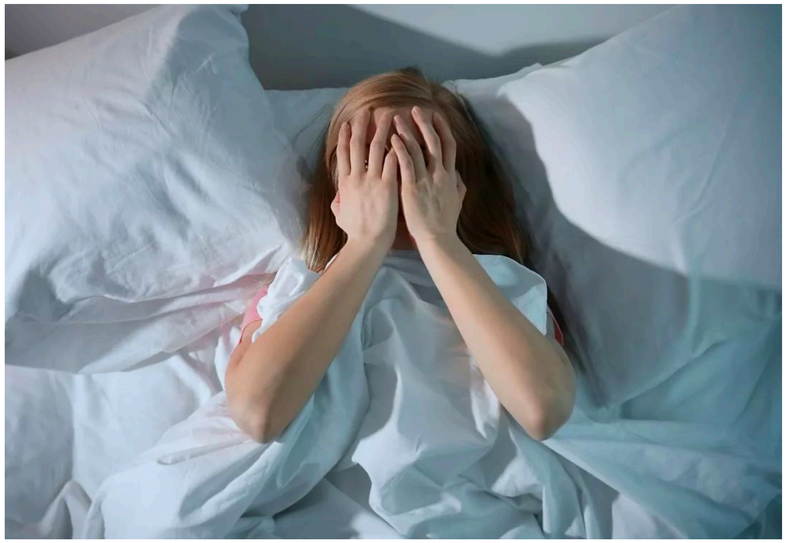
- **Problem Statement**
- Methodology
- Experiments
- Conclusion

Problem Statement

Background

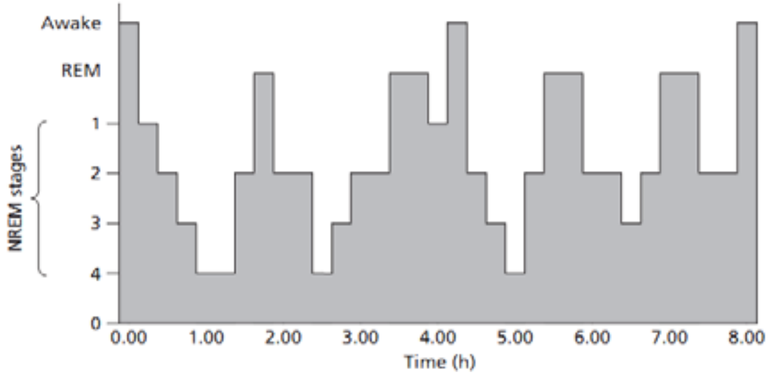
- Sleep is important for the human body
- Sleep disorders become an increasingly frequent problem
- **Sleep stage analysis** helps to address sleep disorders
- **Polysomnography** (PSG) is widely adopted in sleep stage analysis

(a) Sleep disorder



Source: https://stock.tuchong.com/

(b) A Sample of Hypnogram



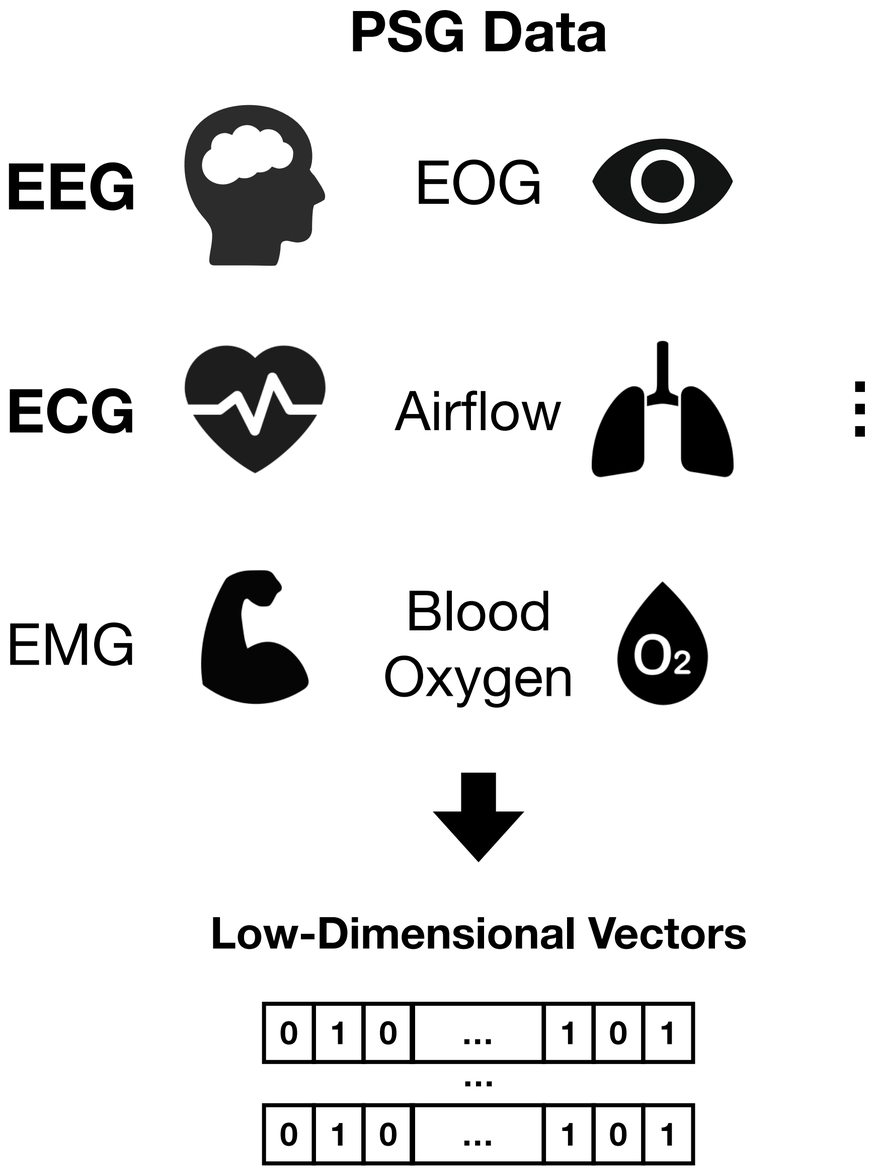
Noviyanto A, Isa SM, Wasito I, et al. (2011) Selecting features of single lead ECG signal for automatic sleep stages classification using correlation-based feature subset selection. *International Journal of Computer Science Issues* 8(5): 139.

Problem Statement

Motivation

- Multiple measurements and features of PSG add data redundancy
- Features extracted from physiological data can have very high dimensions

- Reduce data redundancy
- Save computational resource
- Avoid over-fitting



Problem Statement

Achievements

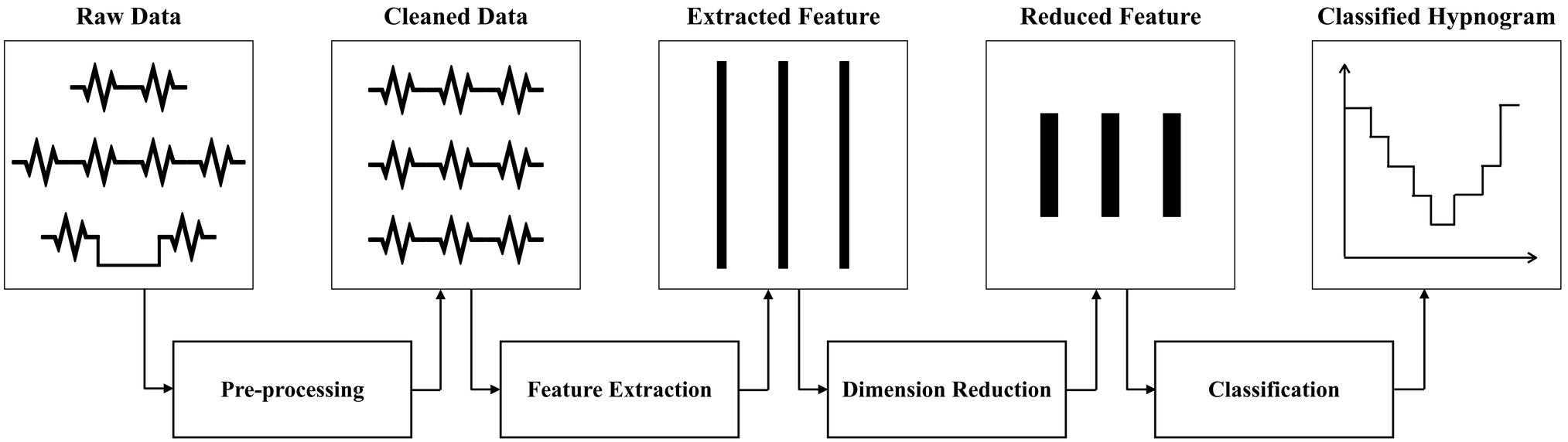
1. Establishing of pre-processed dataset
2. Implementation of sleep stage analysis procedure
3. Verification of the necessity and superiority of dimension reduction
4. Design of a variable weighted PCA based on mutual information
5. Proposal of a novel form of hypnogram

Outline

- Problem Statement
- **Methodology**
- Experiments
- Conclusion

Methodology

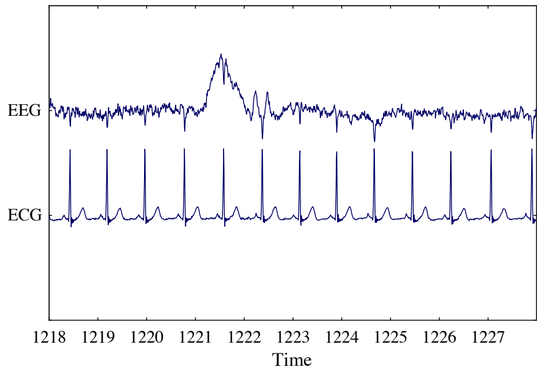
Workflow

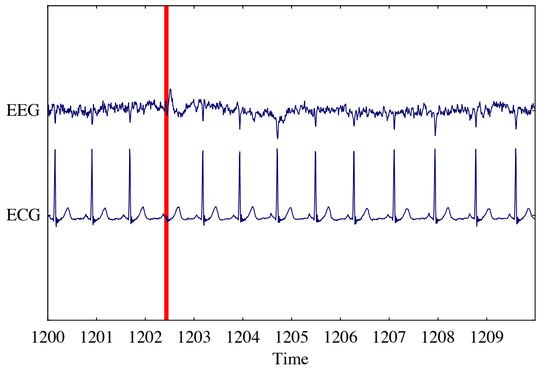


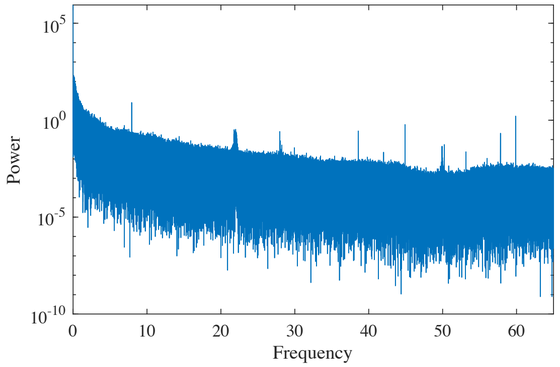
Methodology

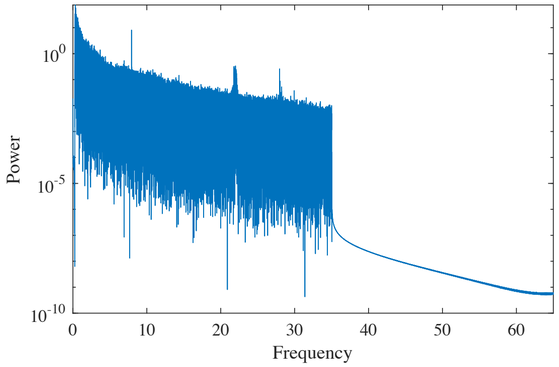
Pre-processing

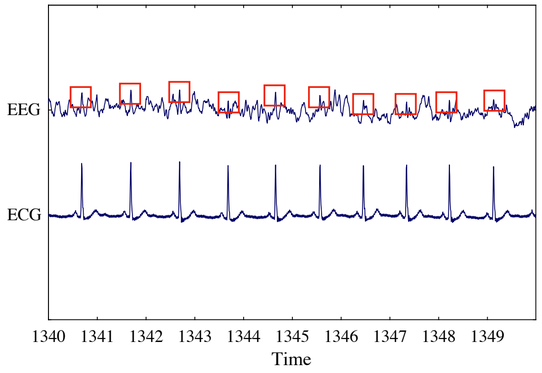
- Removing bad segments (RMD)
- Filtering (FLT)
- Independent Component Analysis
- Normalization
- Segmentation

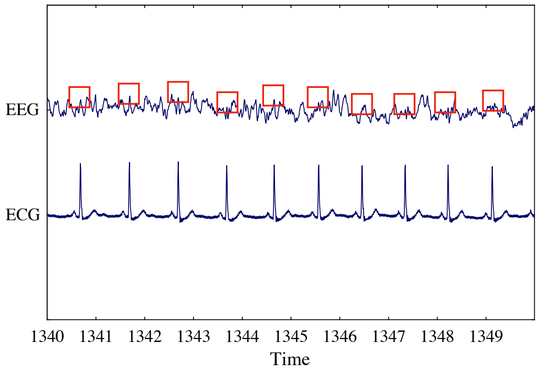












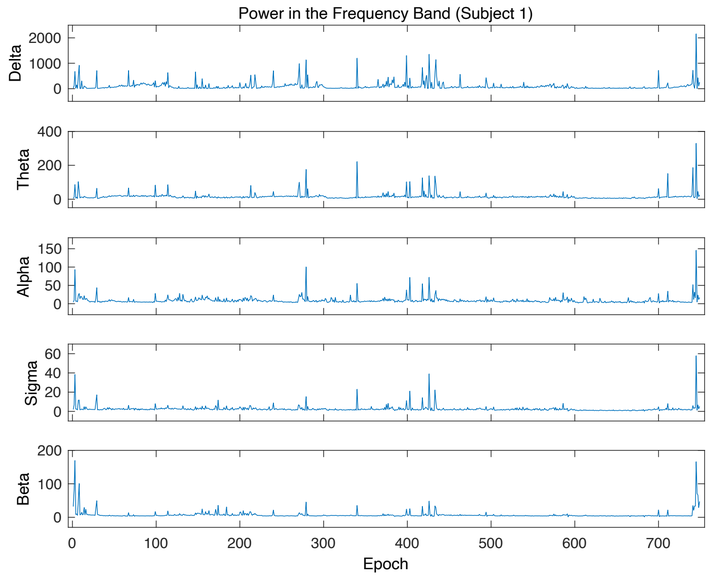
Methodology
Feature Extraction

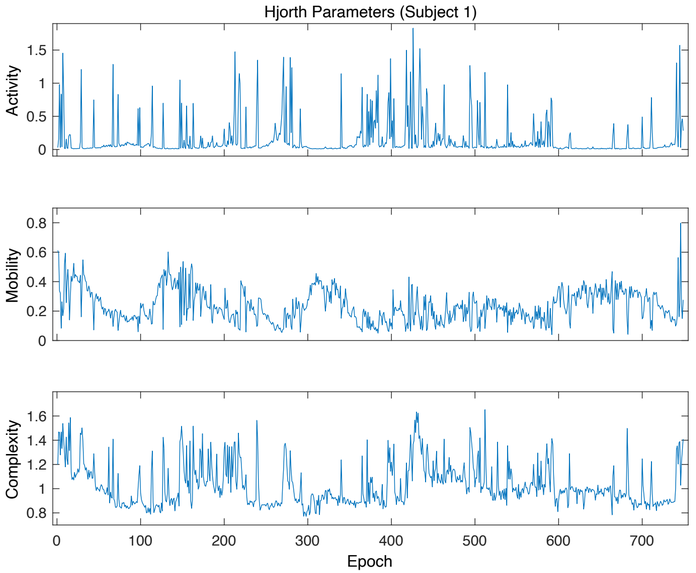
EEG (8 dimension)

- Power in frequency band
- Hjorth Parameters

ECG (30 dimension)

- Autoregressive Coefficients
- Shannon Entropy (WPT)
- Wavelet Leader (DWT)

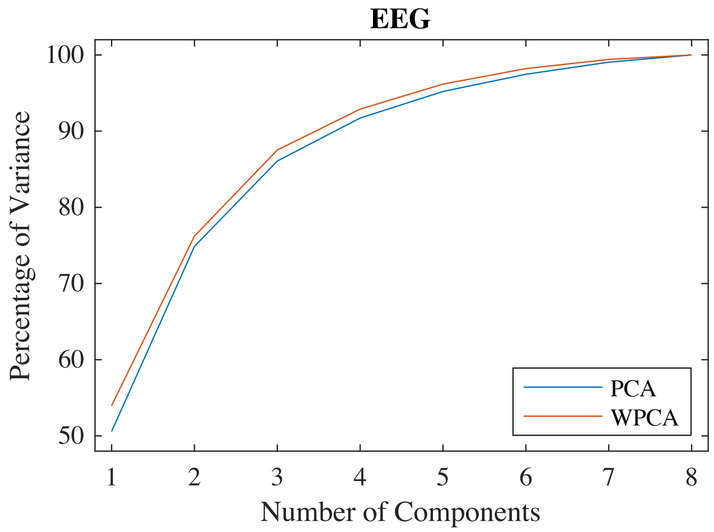


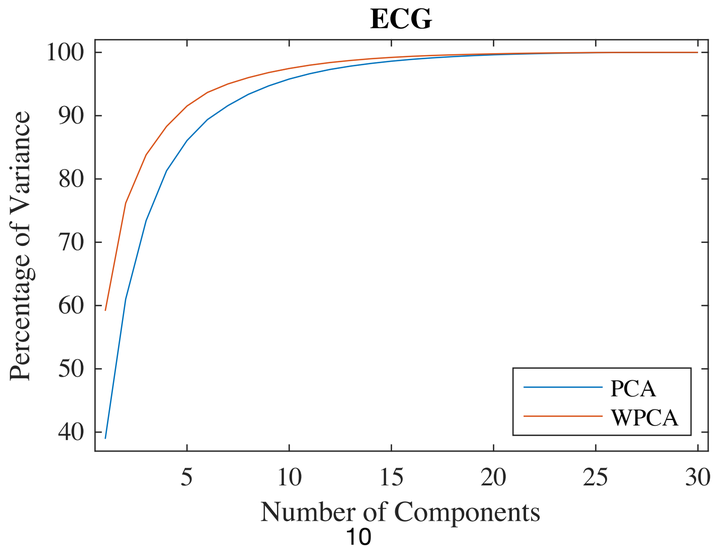


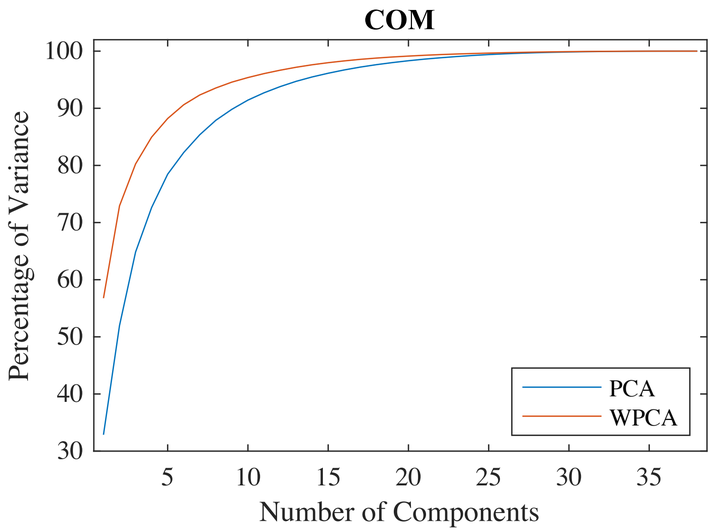
Methodology
Dimension Reduction

- Principal Component Analysis (PCA)
- Weighted PCA (WPCA)

$$w_i = \sum_{j,k} P(X_i = x_j, Y = y_k) \log \frac{P(X_i = x_j, Y = y_k)}{P(X_i = x_j)P(Y = y_k)}$$







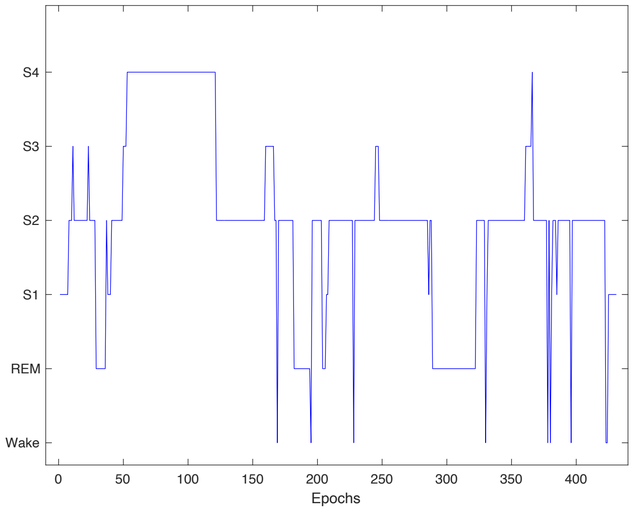
Methodology

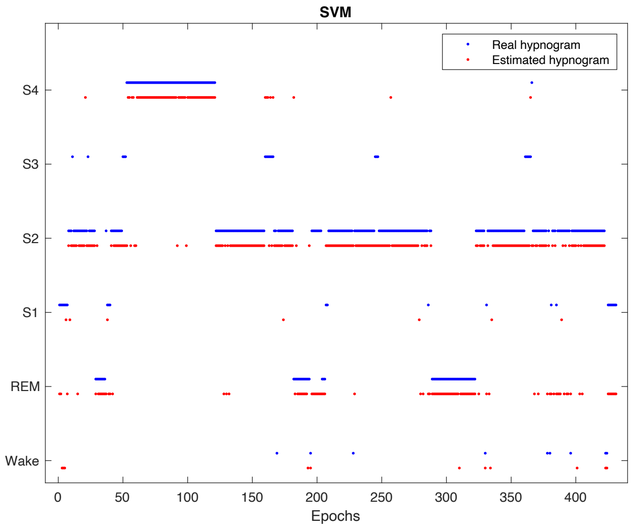
Classification

- Leave-one-out method

Classifiers:

- Linear Discriminant Analysis (LDA)
- Quadratic Discriminant Analysis (QDA)
- Naive Bayes (NB)
- *k* Nearest Neighbor (*k*NN)
- Support Vector Machine (SVM)





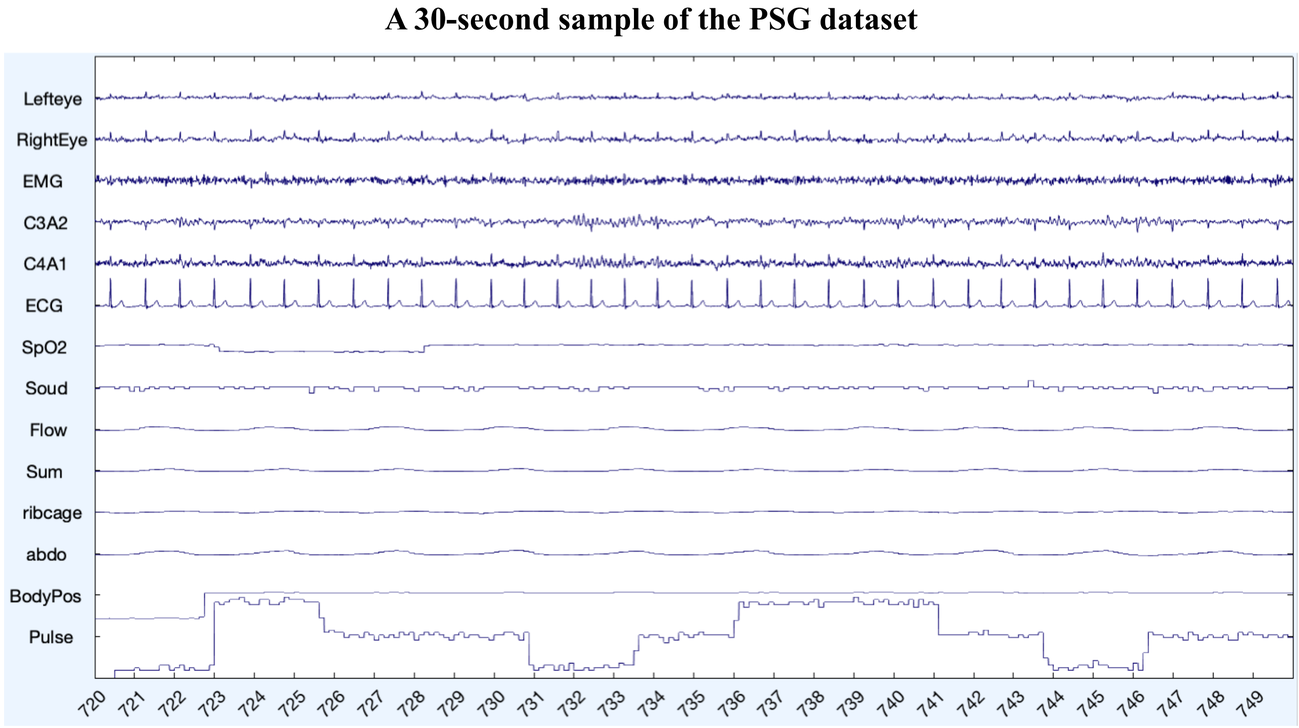
Outline

- Problem Statement
- Methodology
- **Experiments**
- Conclusion

Experiments

Dataset

- St. Vincent's University Hospital / University College Dublin Sleep Apnea Dataset (UCDDB)
- Overnight PSG data
- 25 subjects in total
- Only EEG channel C3A2 and ECG channel is used
- Open access in https://doi.org/10.13026/C26C7D.



Experiments

Result

- SVM outperform other classifiers
- Both pre-processing, filtering and removing bad segments, are effective

ACC	WRN	WRLS	ALL
LDA	74.1	62.5	53.2
QDA	68.3	55.7	48.1
***k*NN**	71.4	60.0	48.3
NB	71.0	61.9	54.6
SVM	**77.2**	**68.4**	**60.3**

ACC	WRN	WRLS	ALL
RAW	72.5	61.8	52.6
FLT	73.0	64.2	54.5
RMD	75.4	65.4	57.6
FULL	**77.2**	**68.4**	**60.3**

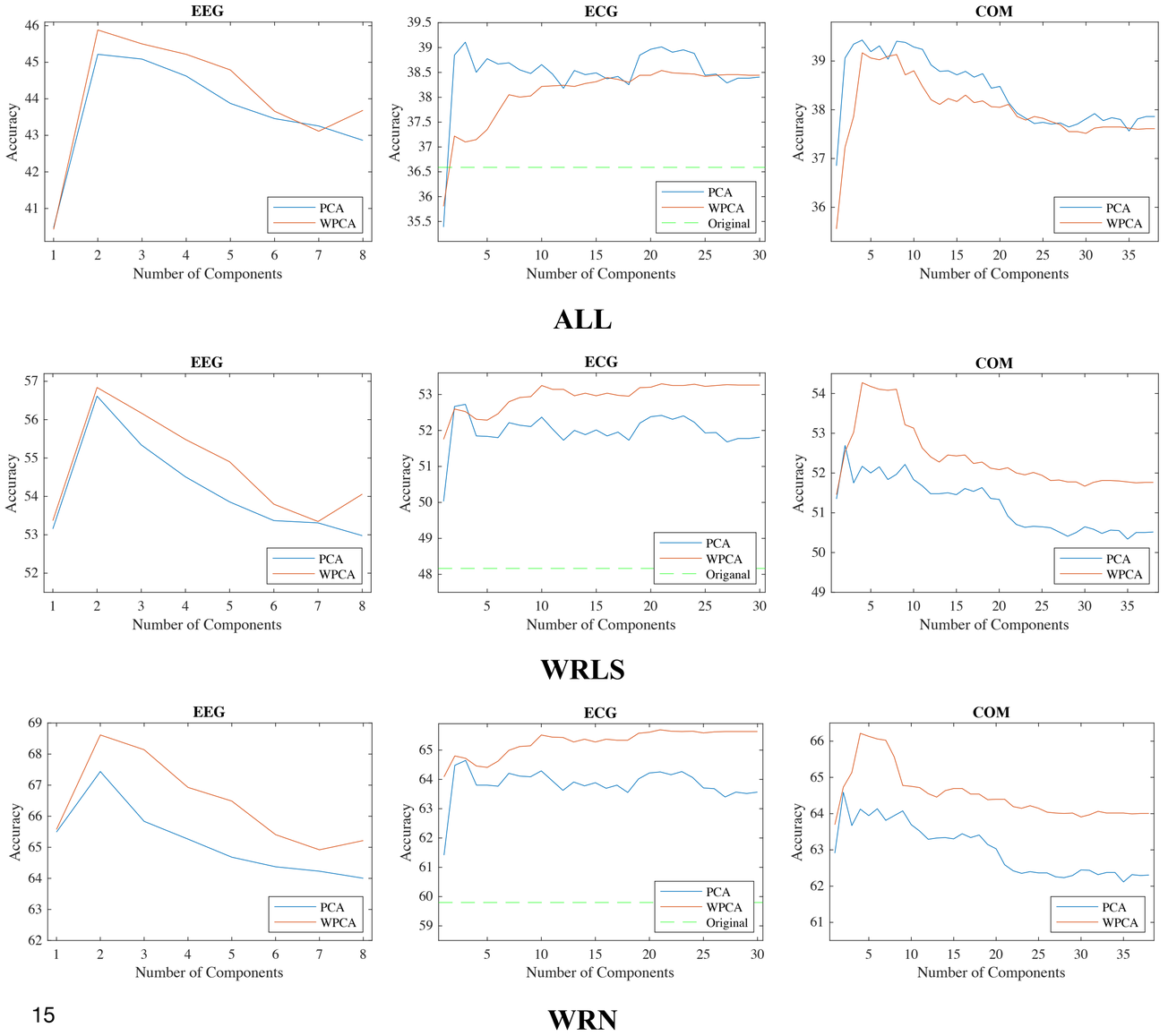
WRN	Wake	REM	NREM			
WRLS	Wake	REM	LS		SWS	
ALL	Wake	REM	S1	S2	S3	S4

Experiments

Result

- Dimension reduction is effective
- Dimension reduction improves the performance in some cases
- Weighted PCA outweighs conventional PCA in most times

	WRN			WRLS			ALL		
	EEG	**ECG**	**COM**	**EEG**	**ECG**	**COM**	**EEG**	**ECG**	**COM**
Original	75.8	59.8	74.1	69.1	48.2	66.0	60.3	36.6	55.2
PCA	67.4	64.6	64.6	56.6	52.7	52.7	45.2	**39.1**	39.4
WPCA	68.6	**65.7**	66.2	56.8	**53.3**	54.3	45.9	38.5	39.2



Outline

- Problem Statement
- Methodology
- Experiments
- **Conclusion**

Conclusion

Summary

1. The complete procedure of sleep stage analysis
    - pre-processing, feature extraction, **dimension reduction**, and classification
2. Experiments demonstrated
    - Effectiveness of data pre-processing
    - Necessity and superiority of dimension reduction in both general and specific cases
    - Improvement of a variable weighted PCA based on mutual information
3. A novel form of hypnogram

Conclusion

Future Work

- Introducing deep learning
- Considering time series
- Including more types of data

Thank you!

Q&A