Functional EEG Network Analysis for Cognitive Diagnosis of Alzheimer's Disease

EEG DATA ANALYSIS WITH MACHINE LEARNING PROJECT

Project Mentor

EEG DATA ANALYSIS PROJECT

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Xiaopeng ZHAO
Associate Professor
University of Tennessee, Knoxville
Mechanical, Aerospace and Biomedical Engineering

Program Supervisor

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Kwai L. WONG
Research Assistant Professor
University of Tennessee, Knoxville
Mechanical, Aerospace and Biomedical Engineering

Group Members

EEG DATA ANALYSIS PROJECT

Jeremy LIU
City University of Hong Kong
College of Engineering

Department of Electronic and Communication Engineering

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Joshua Dunkley
University of Tennessee, Knoxville
College of Engineering

Department of Electrical Engineering and Computer Science

Arden GUO
City University of Hong Kong
College of Science

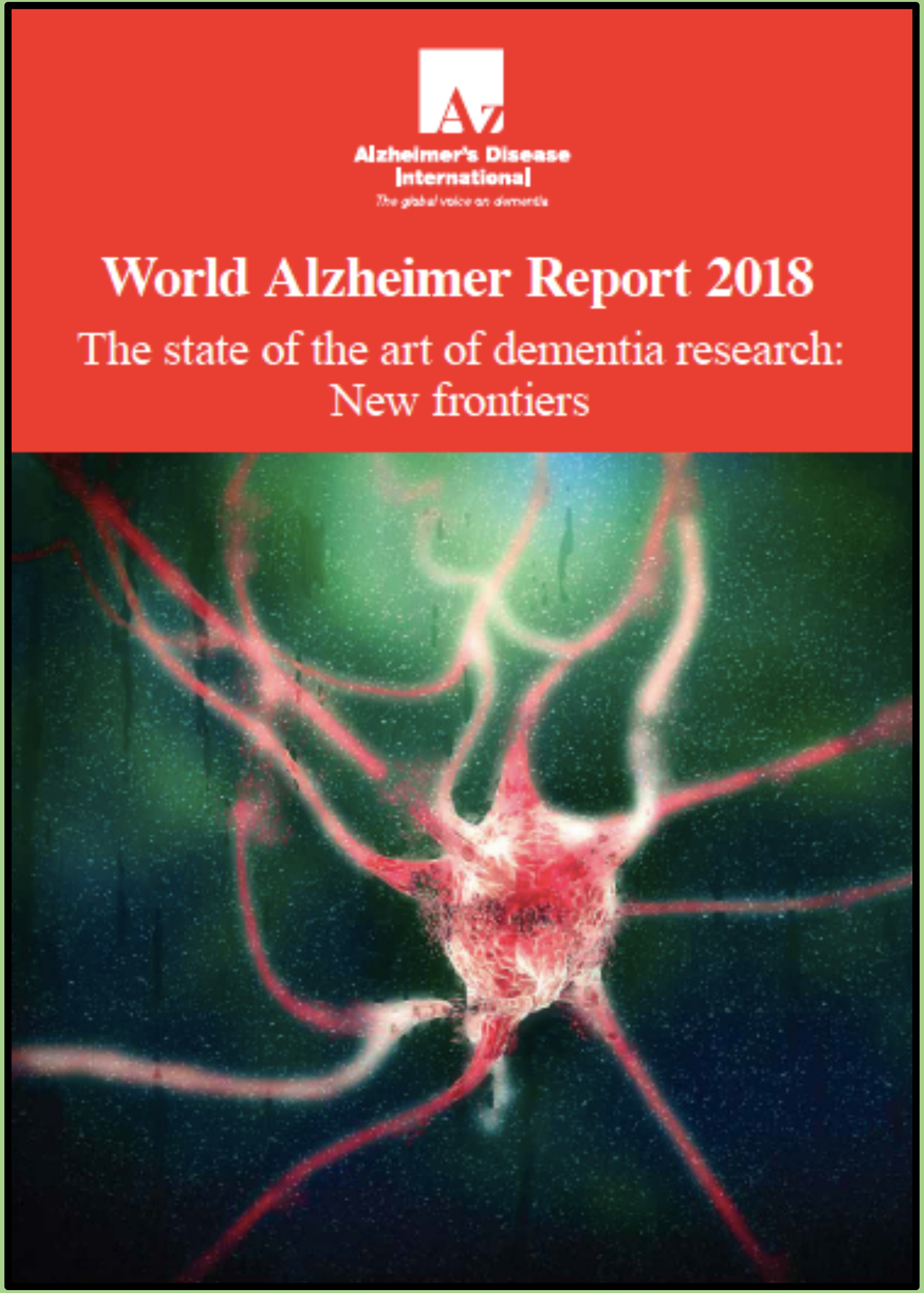
Department of Mathematics

Background:

Harms:

- Irreversible
- Mentation
- Memory
- The ability to conduct the simplest tasks

The World Alzheimer Report of 2018 shows that the number of people living with dementia is estimated to be 50 million today.

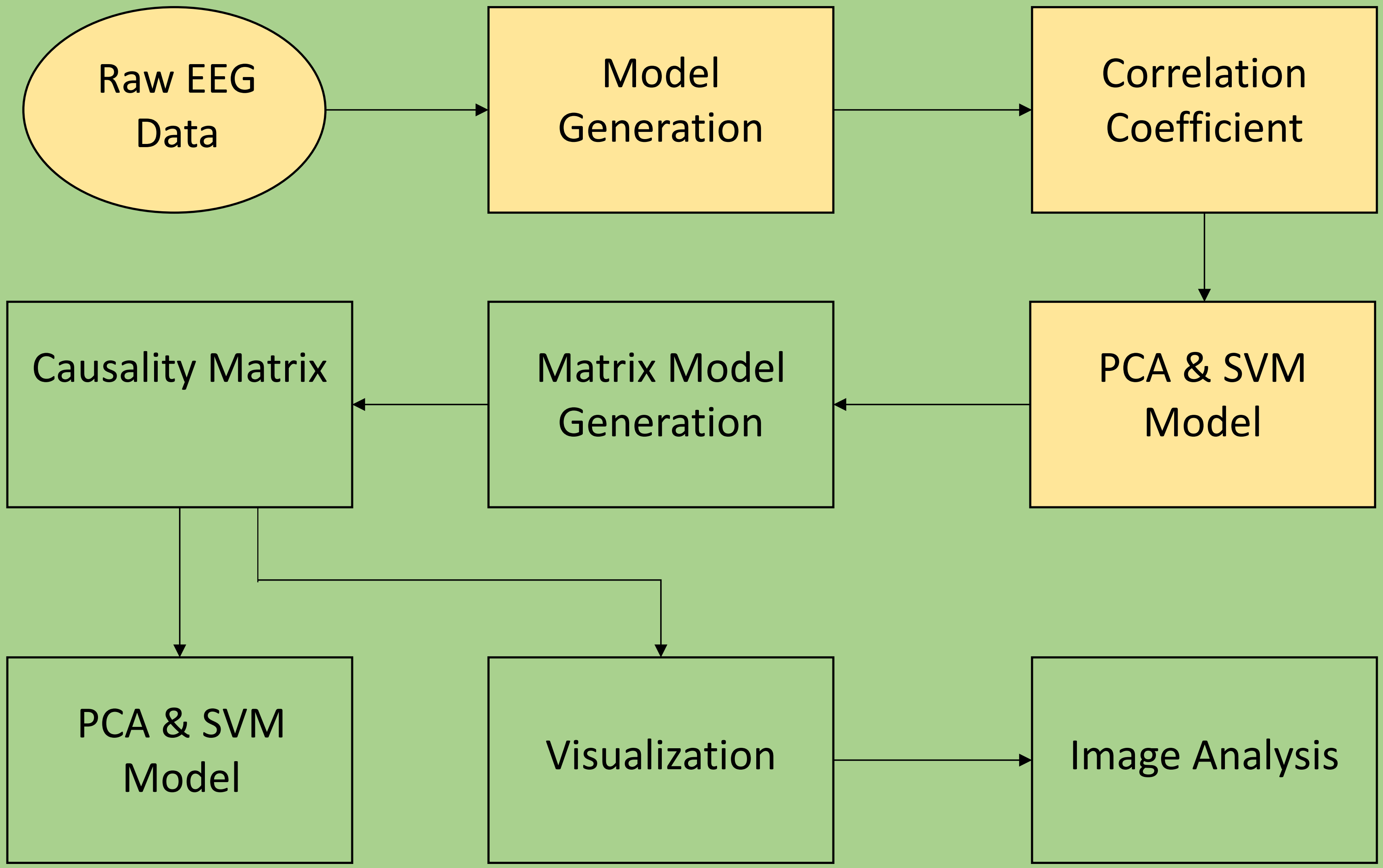


Background:

Electroencephalography(EEG)

- EEG has been introduced as a tool of documenting human brain activity
- Efficient and considerably low cost

- A common problem with this is that numerous factors lead to the loss of data, which yields inaccurate detection for AD.

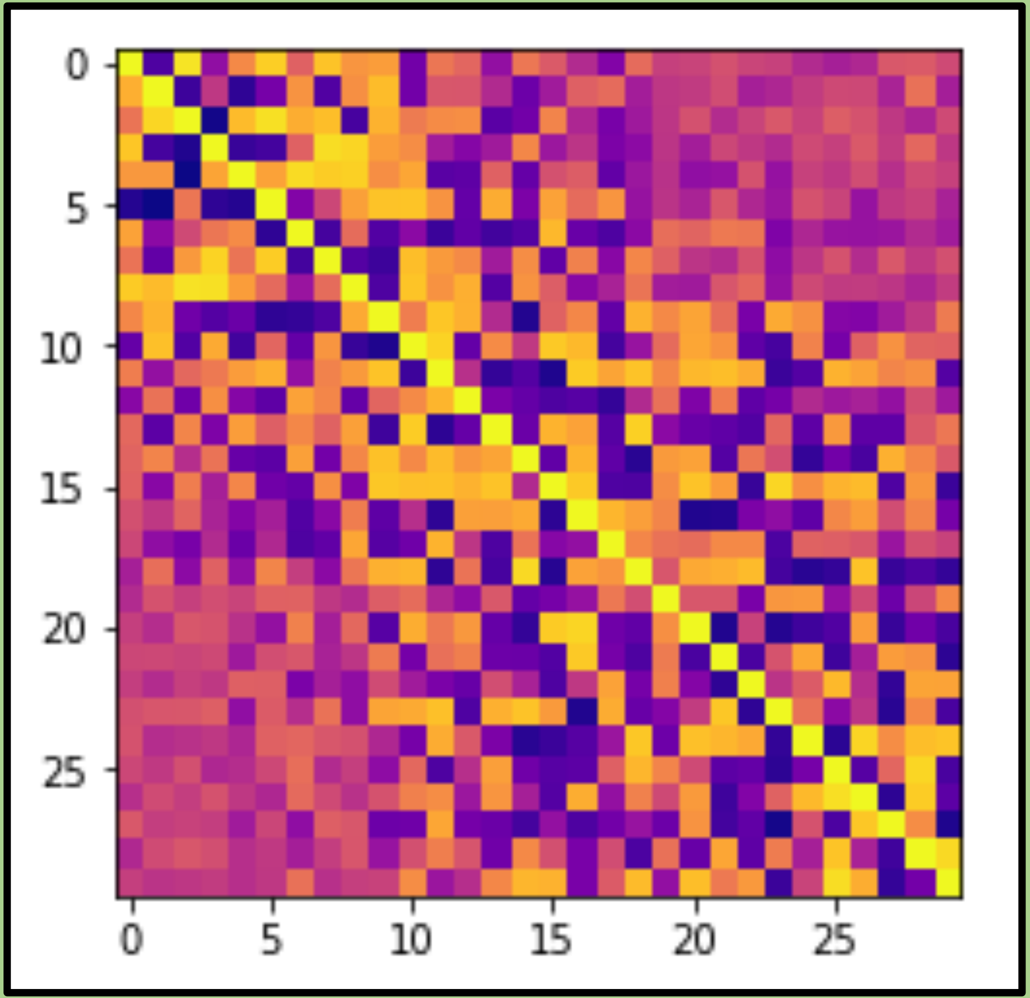

Raw EEG Data
Model Generation
Correlation Coefficient
PCA & SVM Model
Matrix Model Generation
Causality Matrix
PCA & SVM Model
Visualization
Image Analysis

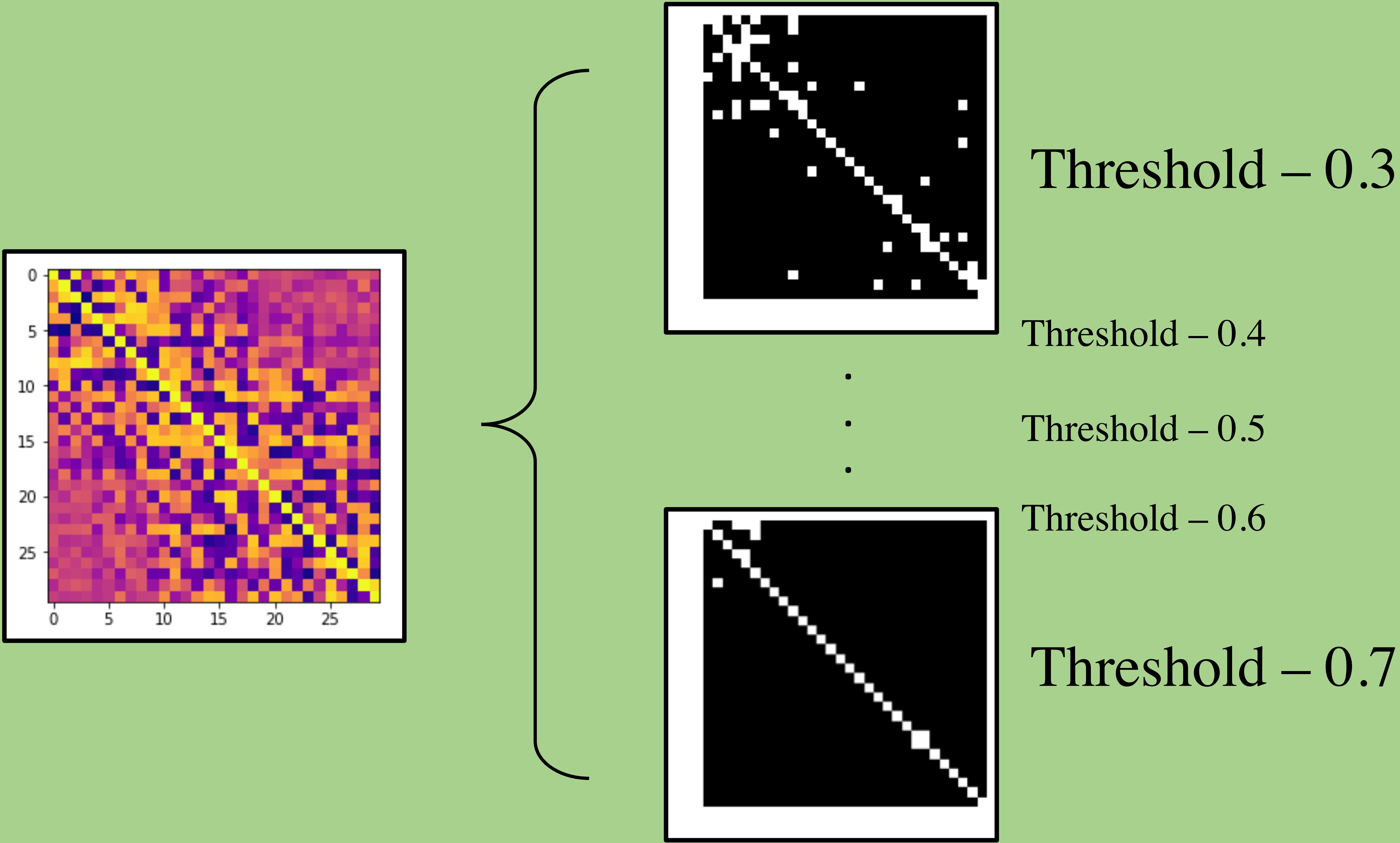
Causality MATRIX:

$$
\begin{array}{cc}
 & \begin{array}{cccccccc} 1 & 2 & 3 & \dots & j & \dots & 29 & 30 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ . \\ . \\ i \\ . \\ 29 \\ 30 \end{array} &
\left[\begin{array}{cccccccc}
* & * & * & . & . & . & * & * \\
* & * & & & & & & \\
* & & * & & & & & \\
. & & & . & & & & \\
. & & & & & & & \\
 & & & & & & & \\
. & & & & & . & & \\
* & & & & & & * & \\
* & & & & & & & *
\end{array}\right]
\end{array}
$$

Image:

$$
\begin{array}{c|cccccccc}
 & 1 & 2 & 3 & \dots & i & \dots & 29 & 30 \\
\hline
1 & * & * & * & . & . & . & * & * \\
2 & * & * & & & & & & \\
3 & * & & * & & & & & \\
\vdots & . & & & . & & & & \\
j & . & & & & . & & & \\
\vdots & . & & & & & . & & \\
29 & * & & & & & & * & \\
30 & * & & & & & & & *
\end{array}
$$




0
5
10
15
20
25
0
5
10
15
20
25
Threshold – 0.3
Threshold – 0.4
Threshold – 0.5
Threshold – 0.6
Threshold – 0.7

Causality MATRIX:

$$
\begin{array}{cc}
 & \begin{array}{cccccccc} 1 & 2 & 3 & \dots & i & \dots & 29 & 30 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ . \\ . \\ j \\ . \\ 29 \\ 30 \end{array} &
\left[\begin{array}{cccccccc}
* & * & * & . & . & . & * & * \\
* & * & & & & & & \\
* & & * & & & & & \\
. & & & . & & & & \\
 & & & & & & & \\
. & & & & . & & & \\
. & & & & & . & & \\
* & & & & & & * & \\
* & & & & & & & *
\end{array} \right]
\end{array}
$$

Feature Reduction:

Column Means $\qquad\qquad$ Eigen Values

$$\underbrace{\begin{array}{cccccccc} 1 & 2 & 3 & \ldots & i & \ldots & 29 & 30 \\ \end{array} \quad \begin{array}{cccccccc} 1 & 2 & 3 & \ldots & i & \ldots & 29 & 30 \\ \end{array}}_{}$$

$$\underbrace{\begin{bmatrix} * & * & * & . & . & . & * & * \end{bmatrix} \quad \begin{bmatrix} * & * & * & . & . & . & * & * \end{bmatrix}}_{\text{PCA \& SVM}}$$

Data:

- Data was obtained from the University of Kentucky’s College of Medicine
- Data was preprocessed before the initial start of the research for noise and artifacts
- Patients were already diagnosed prior to the research

Research Plans:

- Steps

1. Separate the subjects into training sets based on the cognitive state
   - Keep a record of the subject ID, subject data, and subject cognitive state in a class object
2. Build Reconstruction models for all 3 training sets to save time with building models during actual training
3. Loop through the NC-training set and build reconstruction models for Nc_i using the data from the other n subjects in the NC-training set
4. Make predictions for each subjects along the way and calculate the correlation coefficient between the original data and the predicted data.
5. Repeat Steps 3-4 for the MCI- and AD-training sets.
6. Reorganize the correlation coefficients for all the subjects into a matrix, and perform feature reduction and PCA on the newly restructured data.
7. Create an SVM Model for principle components

Research Plans:

- Steps

1) Build a 30x30 matrix where each element in the causality matrix contains a reconstruction model for reconstruction of channel j using channel i.

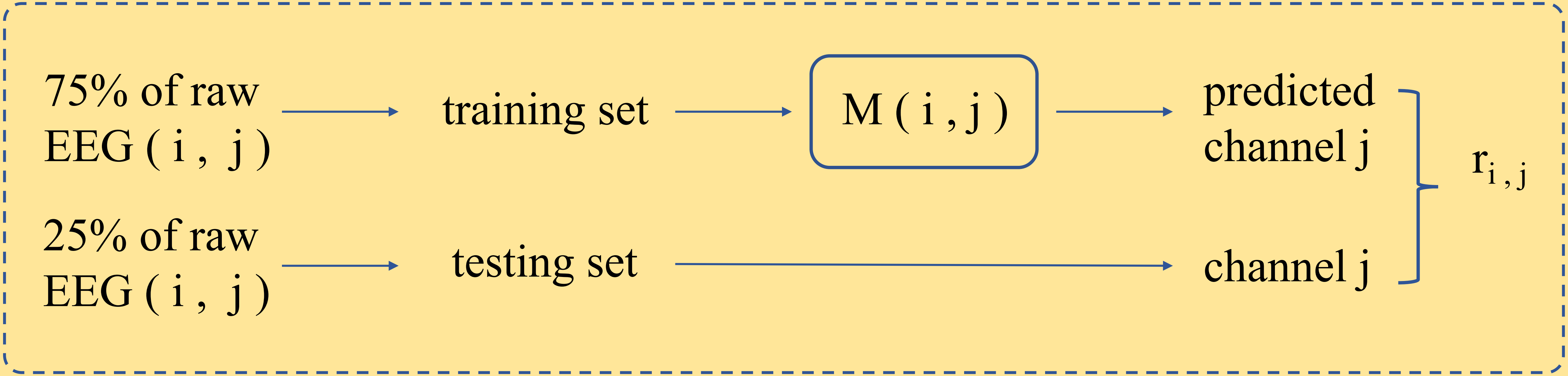
2) Use reconstructed EEG data to make a 30x30 correlation matrix of the causal relationships between EEG channels i and j.

3) Create a color map using the values of the correlation matrix and perform image classification

4) Reduce the number of features and classify using an SVM model

$$
\begin{array}{c|cccccccc|}
 & 1 & 2 & 3 & \dots & i & \dots & 29 & 30 \\
\hline
1 & * & * & * & . & . & . & * & * \\
2 & * & * & & & & & & \\
3 & * & & * & & & & & \\
\vdots & . & & & . & & & & \\
\vdots & . & & & & . & & & \\
j & . & & & & & . & & \\
29 & * & & & & & & * & \\
30 & * & & & & & & & * \\
\hline
\end{array}
$$

STEPS:


75% of raw
EEG (i , j)
training set
M (i , j)
predicted
channel j
25% of raw
EEG (i , j)
testing set
channel j
$r_{i,j}$

Research Plans:

- Steps

1) Build a 30x30 matrix where each element in the causality matrix contains a reconstruction
model for reconstruction of channel j using channel i.

2) Use reconstructed EEG data to make a 30x30 correlation matrix of the causal relationships between EEG channels i and j.

3) Create a color map using the values of the correlation matrix and perform image classification

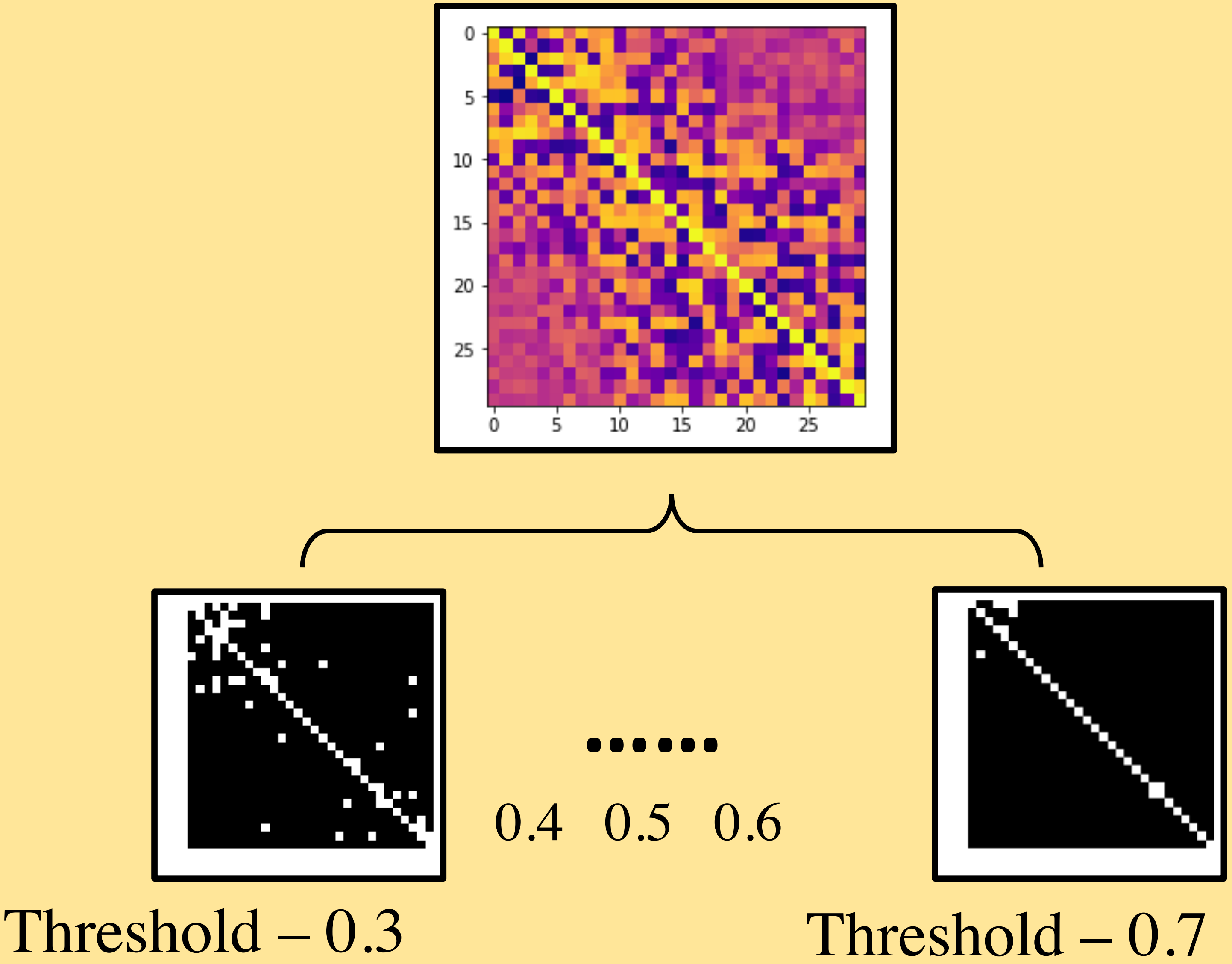
4) Reduce the number of features and classify using an SVM model

$$
\begin{array}{c|cccccccc}
 & 1 & 2 & 3 & \ldots & i & \ldots & 29 & 30 \\
\hline
1 & * & * & * & . & . & . & * & * \\
2 & * & * & & & & & & \\
3 & * & & * & & & & & \\
. & . & & & . & & & & \\
. & . & & & & . & & & \\
j & . & & & & & . & & \\
. & & & & & & & & \\
29 & * & & & & & & * & \\
30 & * & & & & & & & *
\end{array}
$$

Research Plans:

- Steps

1) Build a 30x30 matrix where each element in the causality matrix contains a reconstruction
model for reconstruction of channel j using channel i.
2) Use reconstructed EEG data to make a 30x30 correlation matrix of the causal relationships between EEG channels i and j.
3) Create a color map using the values of the correlation matrix and perform image classification
4) Reduce the number of features and classify using an SVM model



Research Plans:

- Steps

1) Build a 30x30 matrix where each element in the causality matrix contains a reconstruction
model for reconstruction of channel j using channel i.
2) Use reconstructed EEG data to make a 30x30 correlation matrix of the causal relationships between EEG channels i and j.
3) Create a color map using the values of the correlation matrix and perform image classification
4) Reduce the number of features and classify using an SVM model

$$
\begin{array}{c c}
 & \begin{array}{cccccccc} 1 & 2 & 3 & \ldots & i & \ldots & 29 & 30 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ . \\ . \\ j \\ . \\ 29 \\ 30 \end{array} &
\left[\begin{array}{cccccccc}
* & * & * & . & . & . & * & * \\
* & * & & & & & & \\
* & & * & & & & & \\
. & & & . & & & & \\
. & & & & . & & & \\
. & & & & & . & & \\
 & & & & & & & \\
* & & & & & & * & \\
* & & & & & & & *
\end{array}\right]
\end{array}
$$

$$
\begin{array}{cccccccc} 1 & 2 & 3 & \ldots & i & \ldots & 29 & 30 \end{array}
$$

$$
\left[\begin{array}{cccccccc} * & * & * & . & . & . & * & * \end{array}\right]
$$

Research Plans:

- Algorithms/Math
  - **Leave-one-out principle**
  - Correlation Coefficient
  - Eigen Values
  - Principle Component Analysis (PCA)
  - SVM Model

Research Plans:

- Algorithms/Math
  - Leave-one-out principle
  - **Correlation Coefficient**
  - Eigen Values
  - Principle Component Analysis (PCA)
  - SVM Model

$$r = \frac{Cov\ (X,Y)}{\sigma_X \sigma_Y}$$

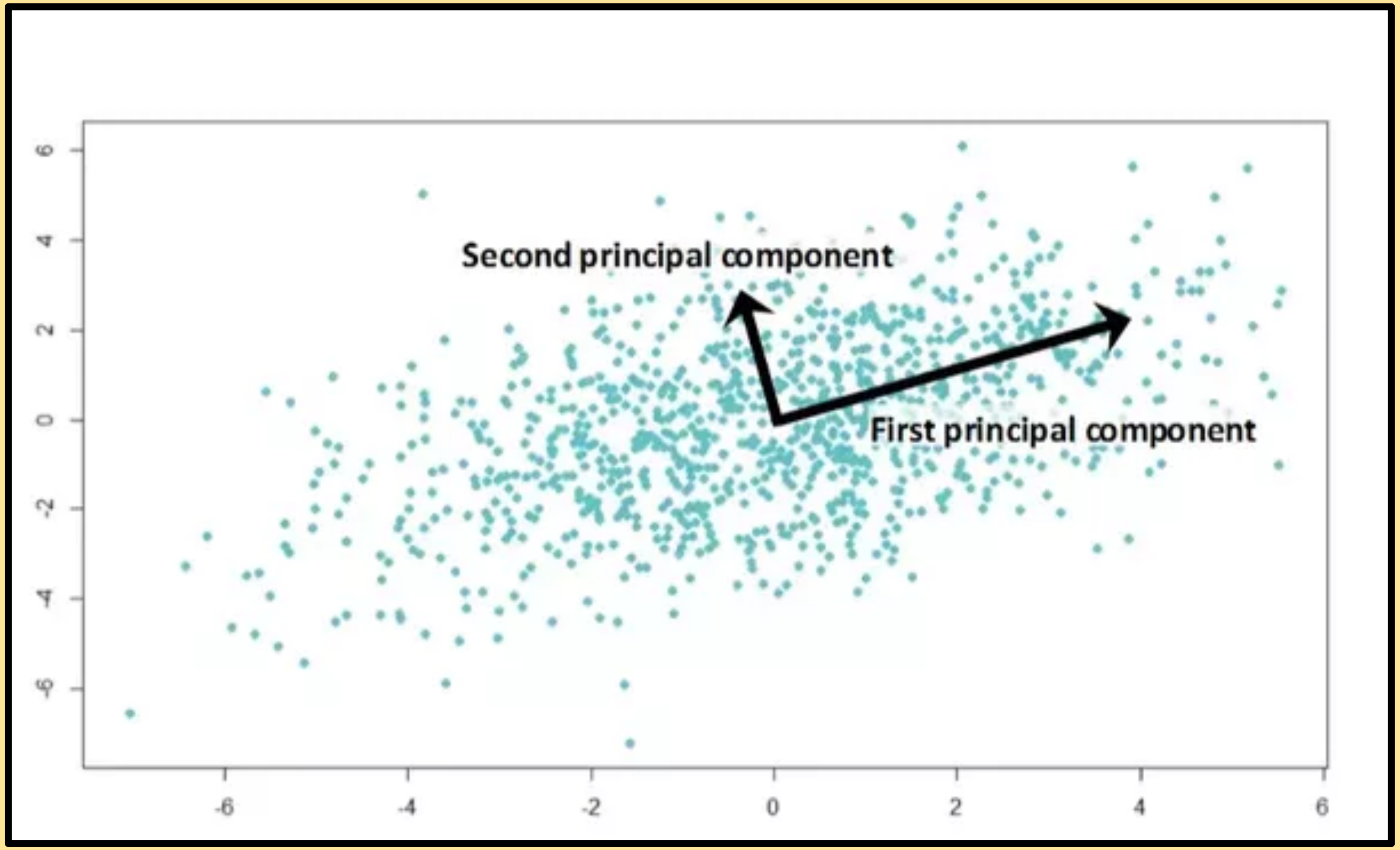
Research Plans:

- Algorithms/Math
  - Leave-one-out principle
  - Correlation Coefficient
  - **Eigen Values**
  - Principle Component Analysis (PCA)
  - SVM Model

$$AX = \lambda X$$

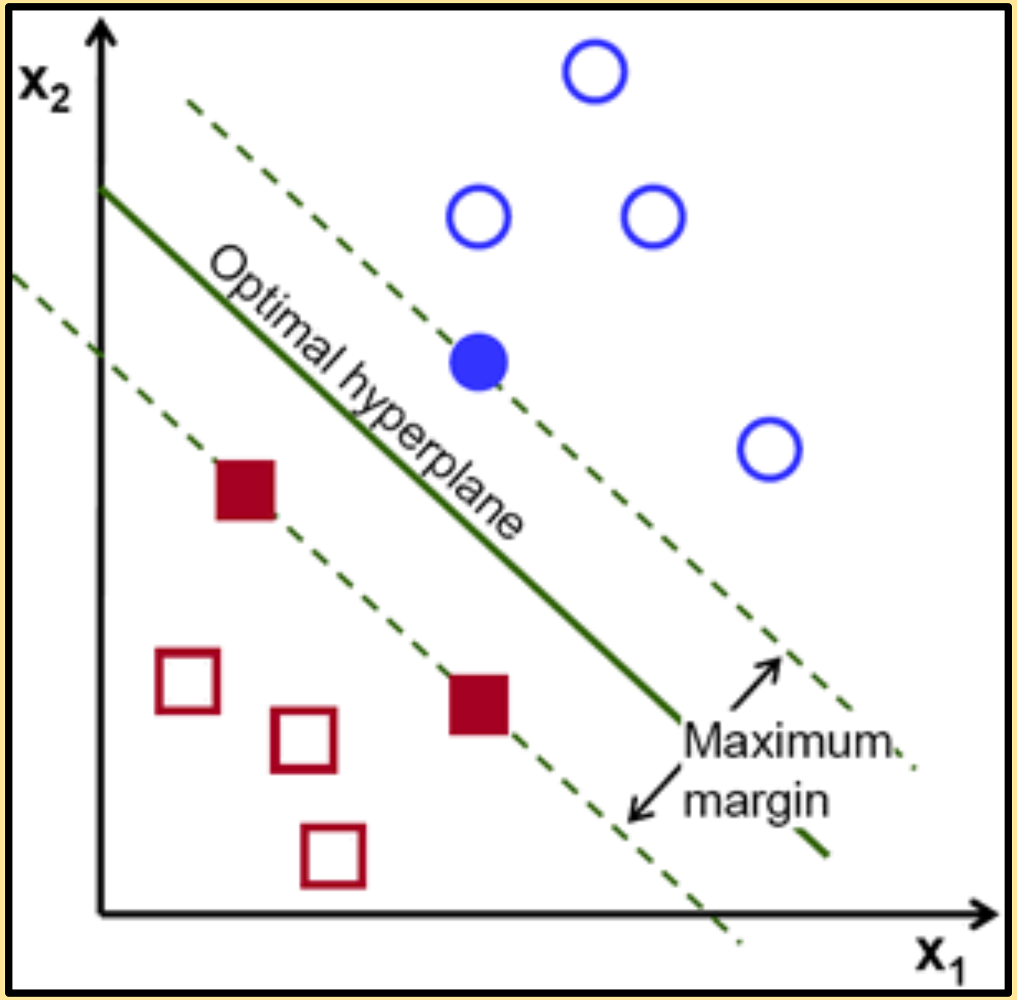
Research Plans:

- Algorithms/Math
  - Leave-one-out principle
  - Correlation Coefficient
  - Eigen Values
  - **Principle Component Analysis (PCA)**
  - SVM Model



Research Plans:

- Algorithms/Math
  - Leave-one-out principle
  - Correlation Coefficient
  - Eigen Values
  - Principle Component Analysis (PCA)
  - **SVM Model**

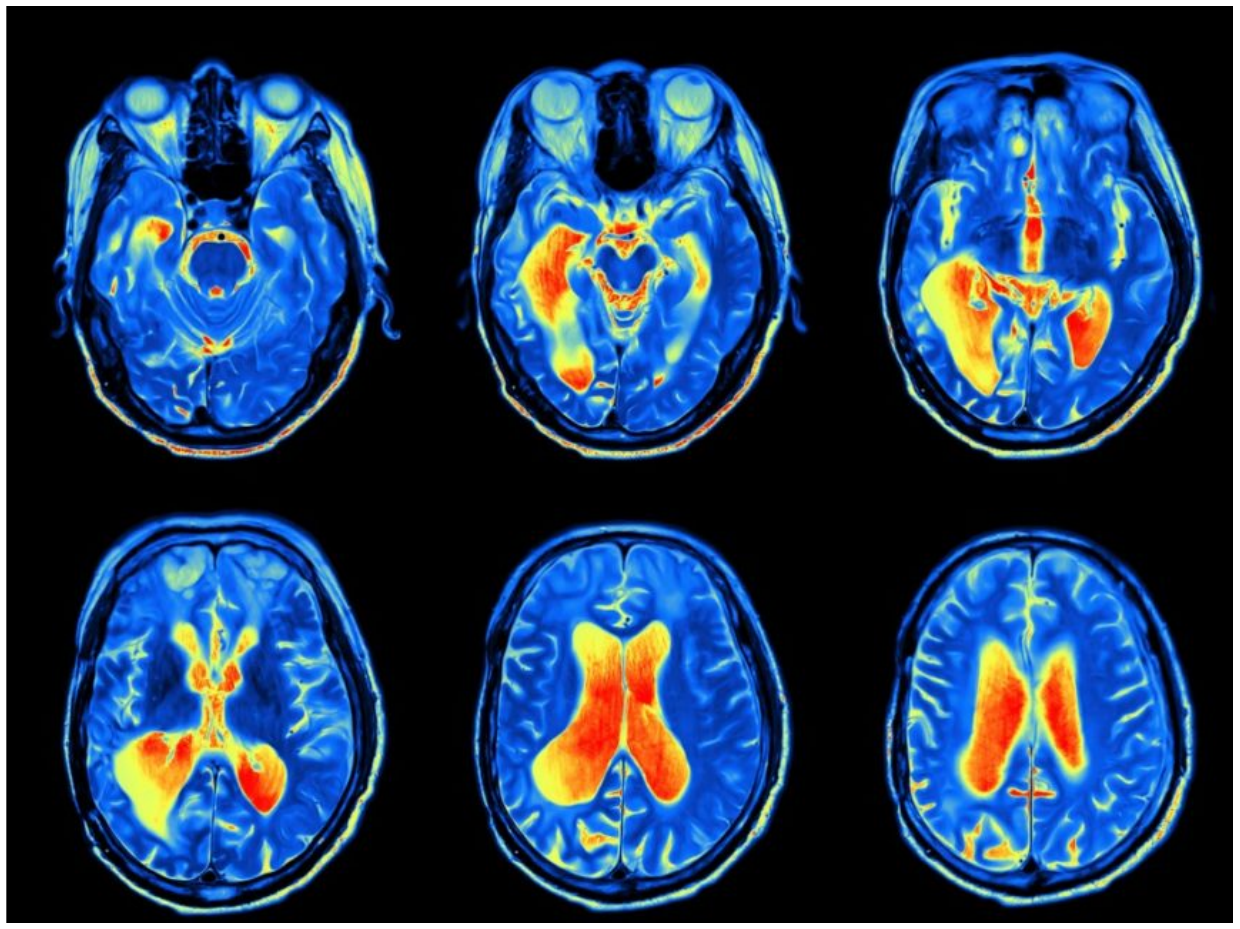


Research Plans:

- Implementation
  - Currently working with Python script and IPython Notebooks in Google Collaboratory
  - Built Reconstruction Models using Keras with a Tensorflow backend and using Tensorflow itself
  - Made graphs and figures using Matplotlib
  - Used sklearn and numpy to normalize the data, and perform feature reduction and Principle Component Analysis on reconstructed data
  - Used the Deep Neural Network Toolbox in MATLAB to perform image classification

Research Plans:

- Future works
  - Correlate EEG Causality Network to fMRI
  - Collect EEG data from more subjects
  - Real-time diagnosis of cognitive deficit

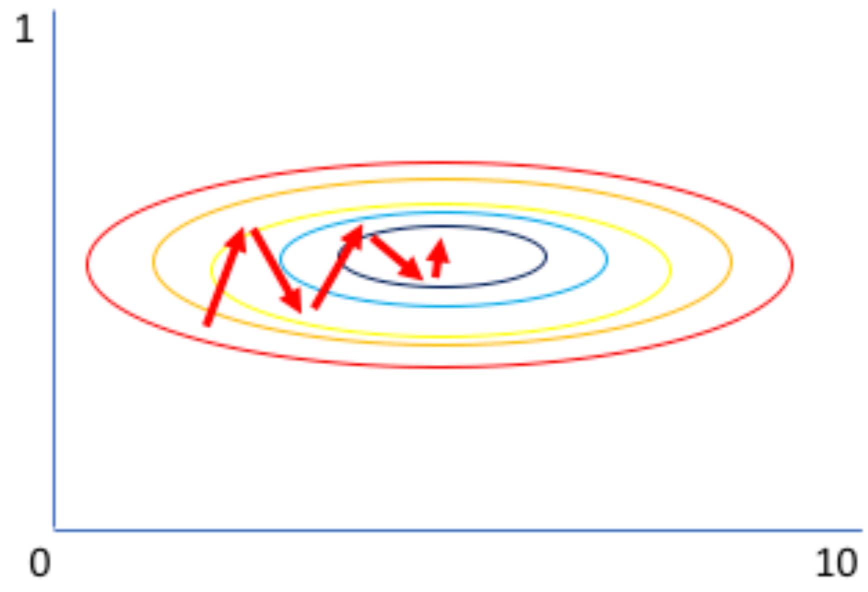


Data Preprocessing:

Standardization

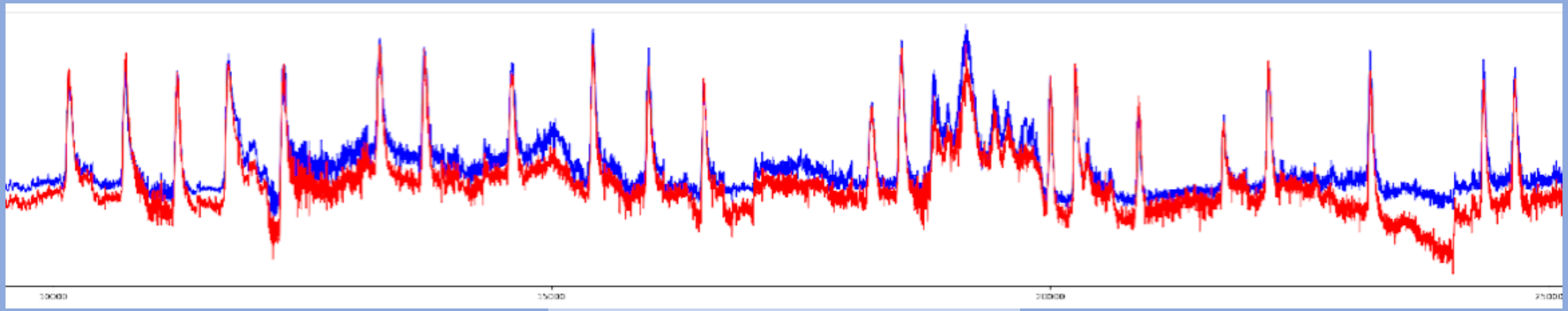
Standardize features by removing the mean and scaling to unit variance

- Faster learning speed
- a feature has a variance that is orders of magnitude larger that others, it might dominate the objective function, causing the estimator unable to learn from other features correctly as expected.

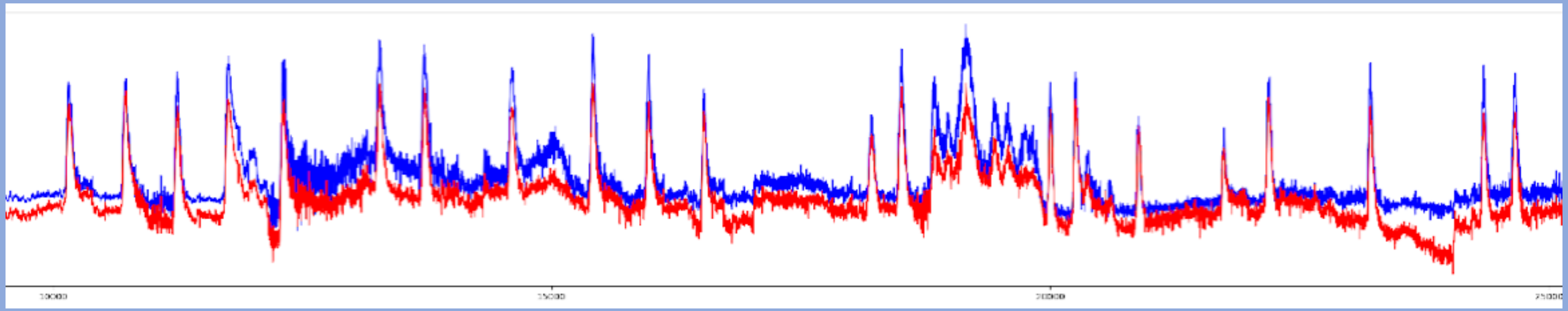


Gradient of larger parameter dominates the update

Reconstruction Results



Normalized data



Without preprocessing

Original Data

Predicted data

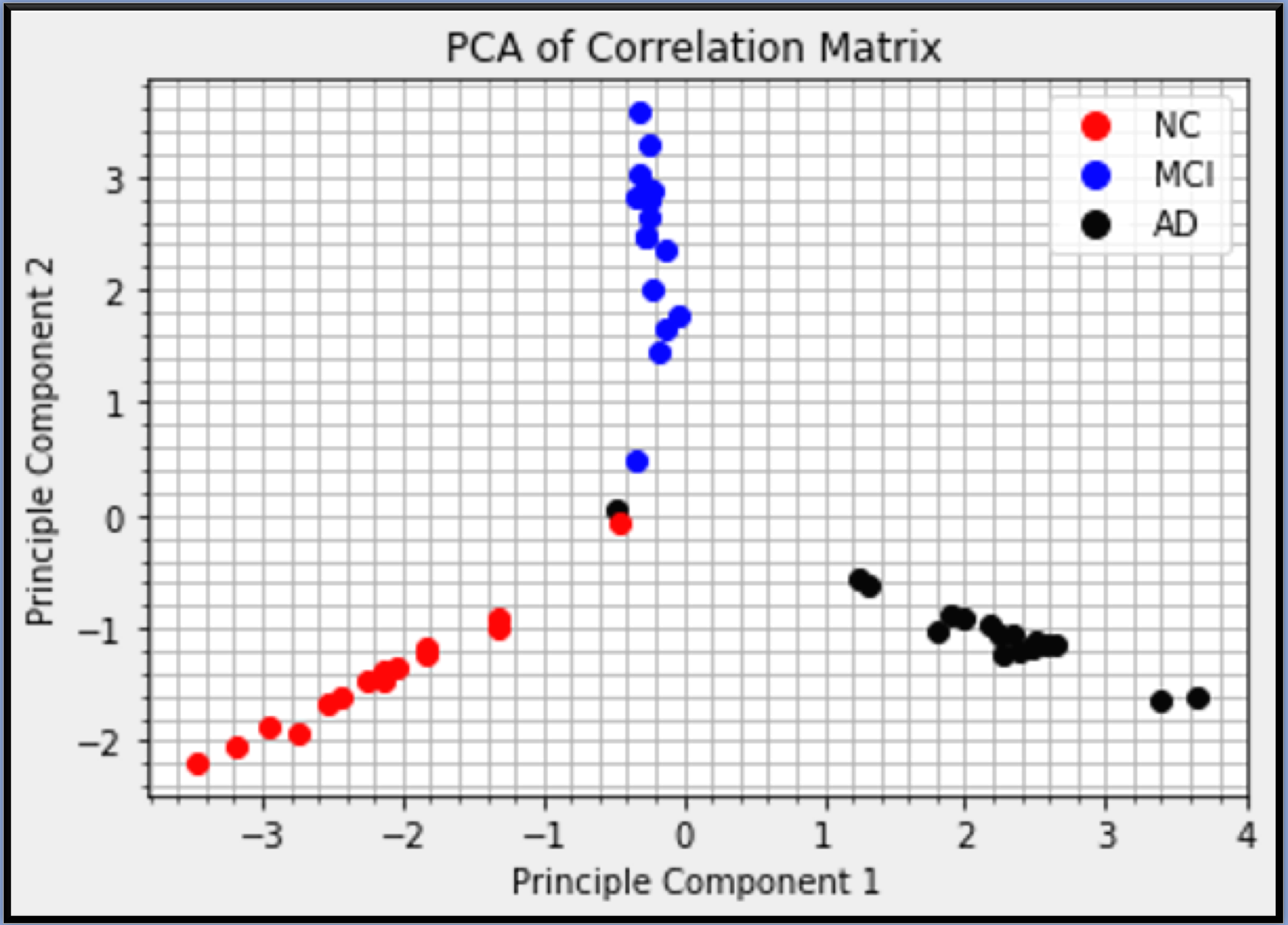
PCA Results:

Correlation Matrix

- Row = 48 subjects
- Column = 3 states
  - 30 values for AD
  - 30 values for MCI
  - 30 values for NC

PCA result

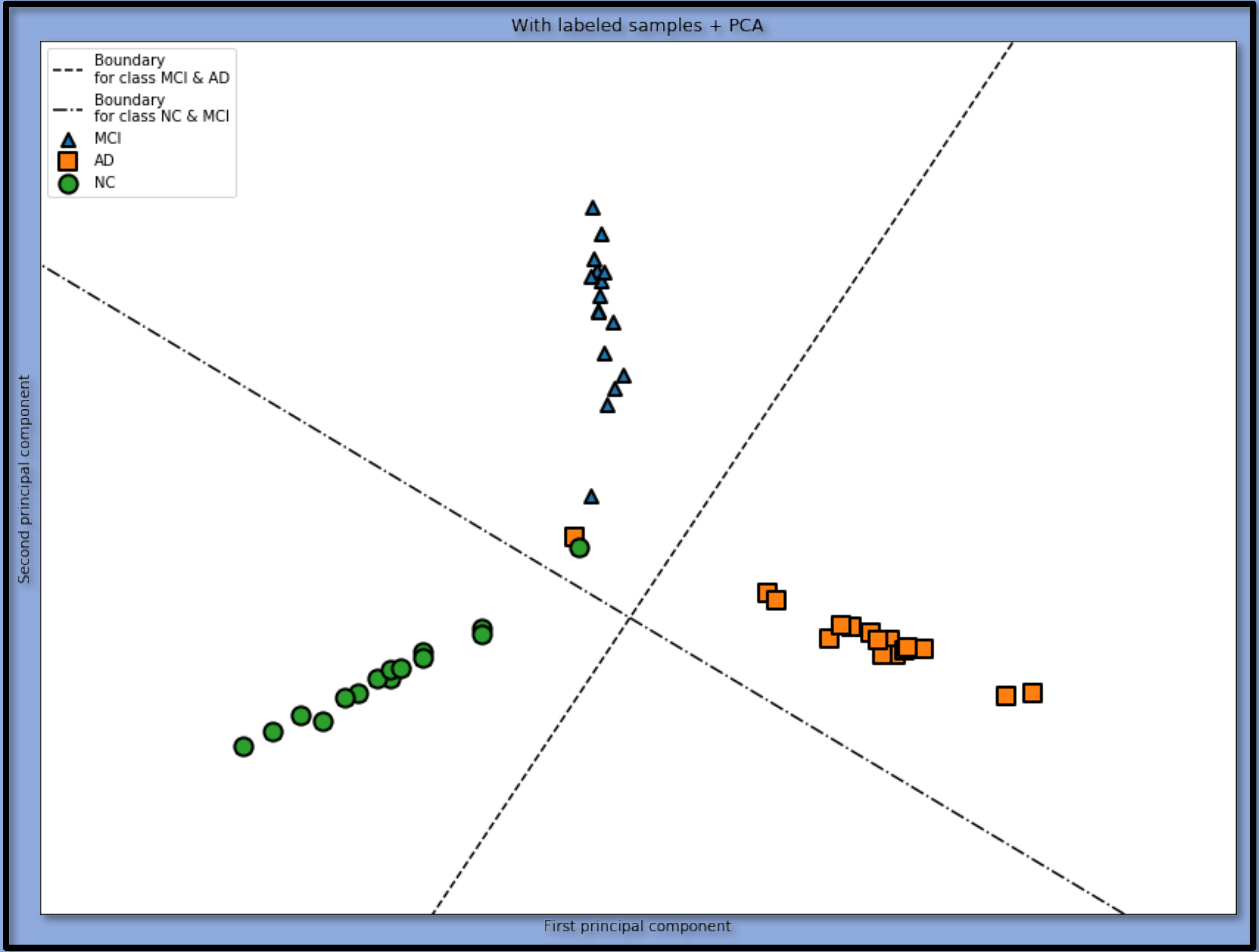
- 2 principal components
- Plot 2D diagram
- Most of subjects clustered together



SVM Model:

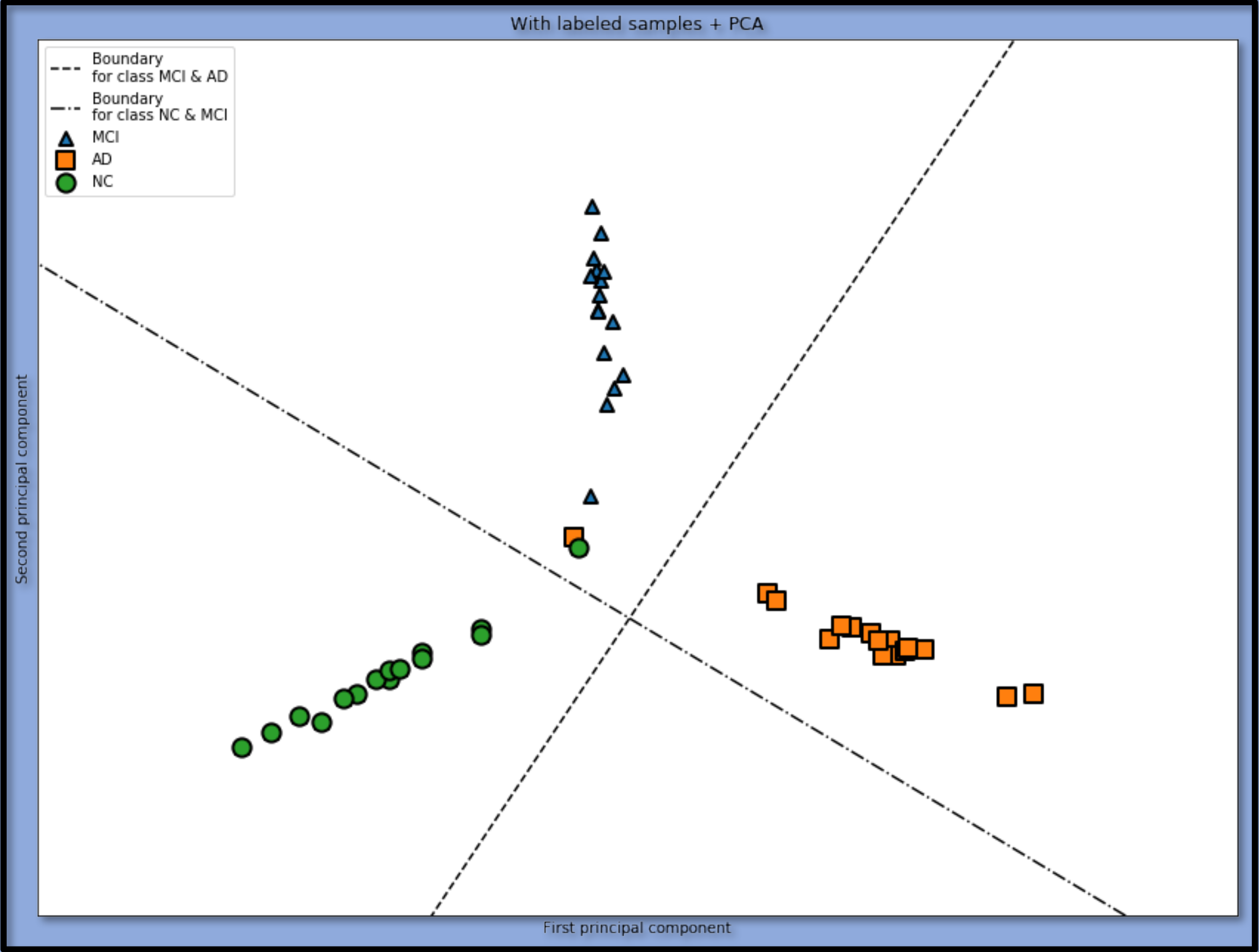
Support Vector Machine:

- Based on PCA results
- Linear SVM



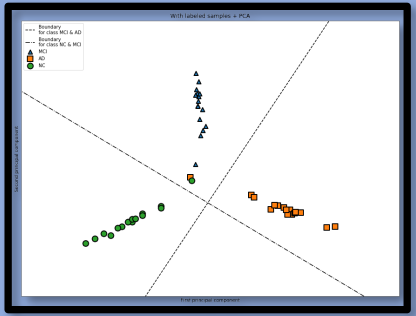
SVM Model:

		Predicted classes			
		NC	MCI	AD	
True classes	NC	14	1	0	93.3%
	MCI	0	15	0	100%
	AD	0	1	15	93.8%
		100%	88.2%	100%	Overall Acc: 95.8%



Accuracy assessment:

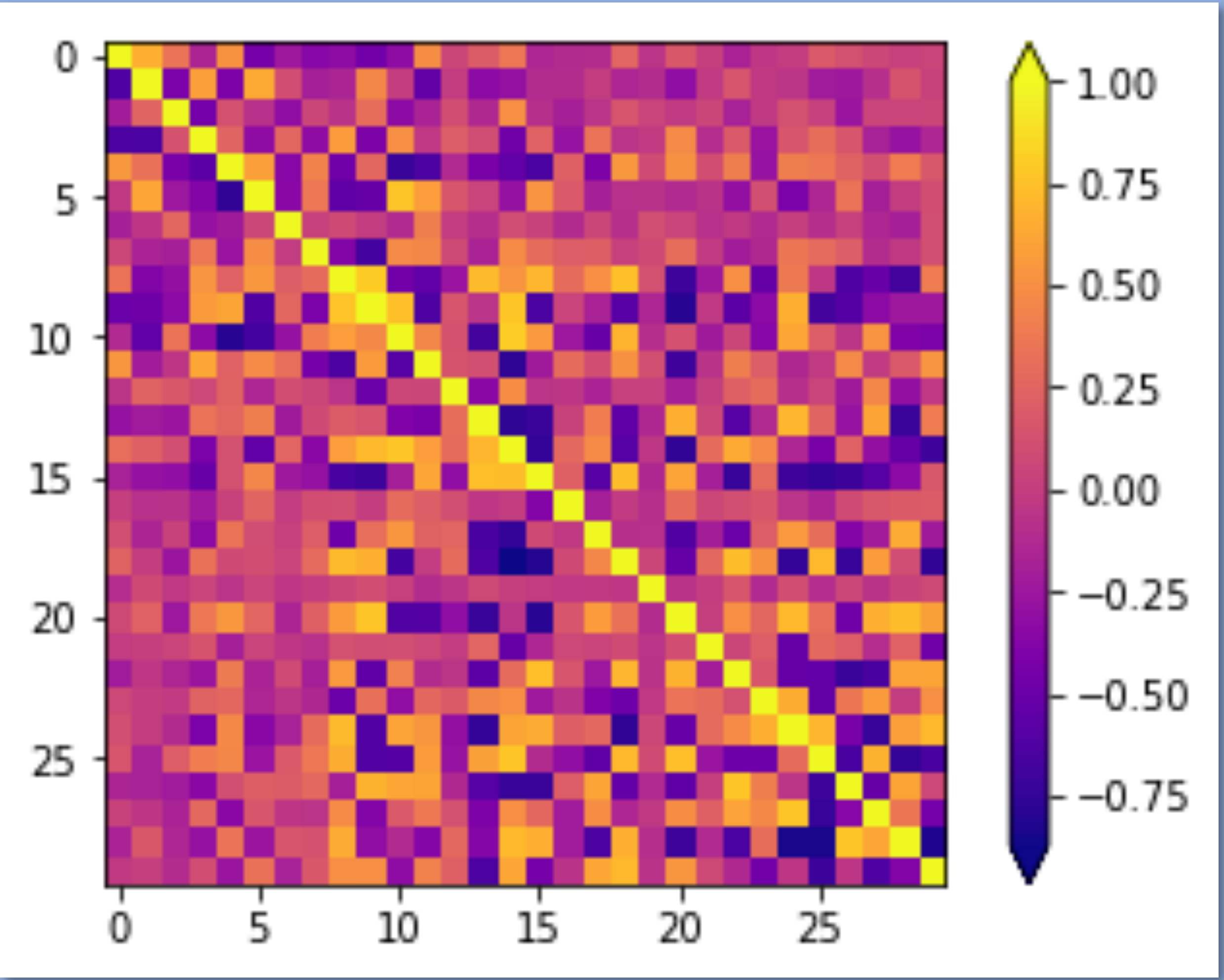
		Predicted classes			
		NC	MCI	AD	
True classes	NC	14	1	0	93.3%
	MCI	0	15	0	100%
	AD	0	1	15	93.8%
		100%	88.2%	100%	Overall Acc: 95.8%

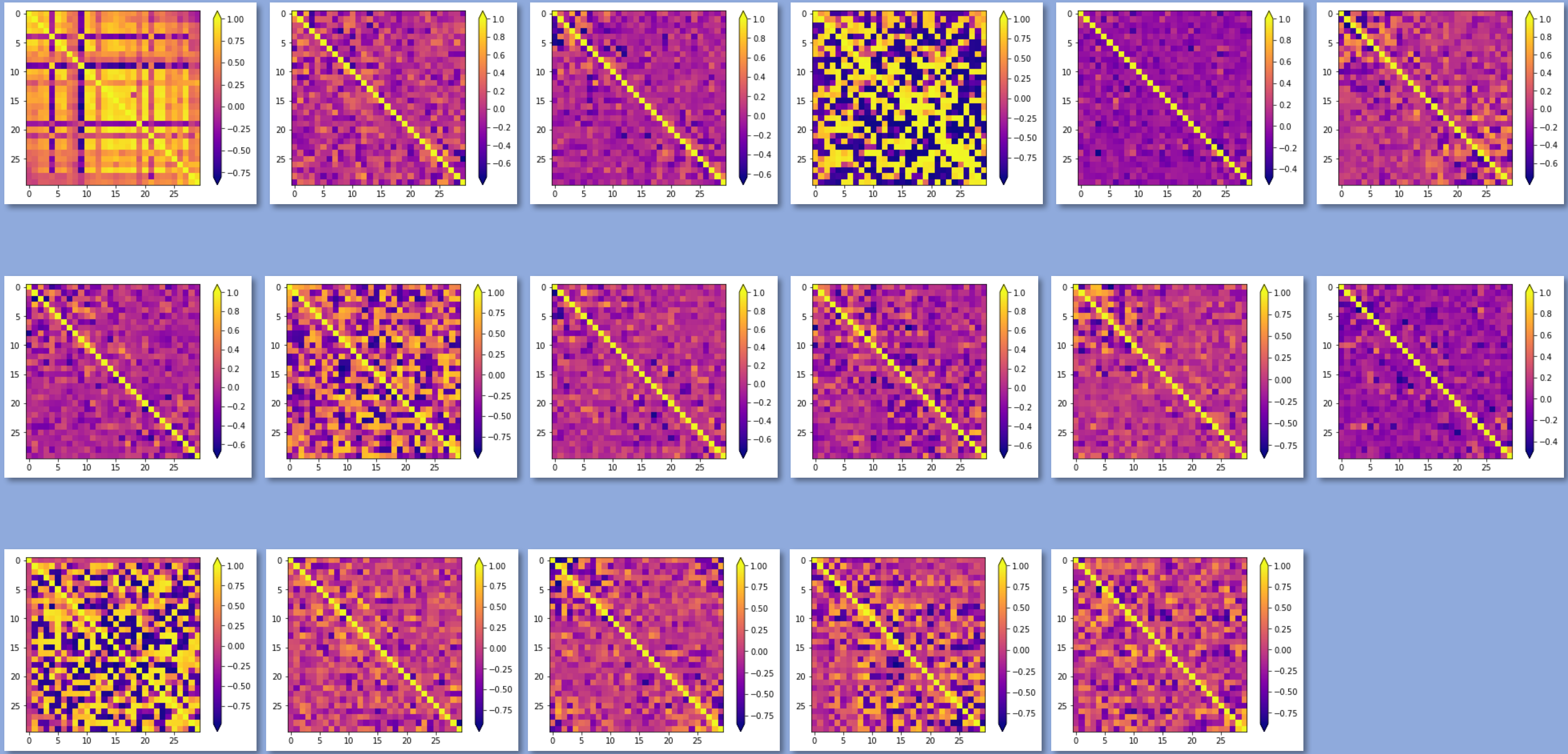


Causality matrix:

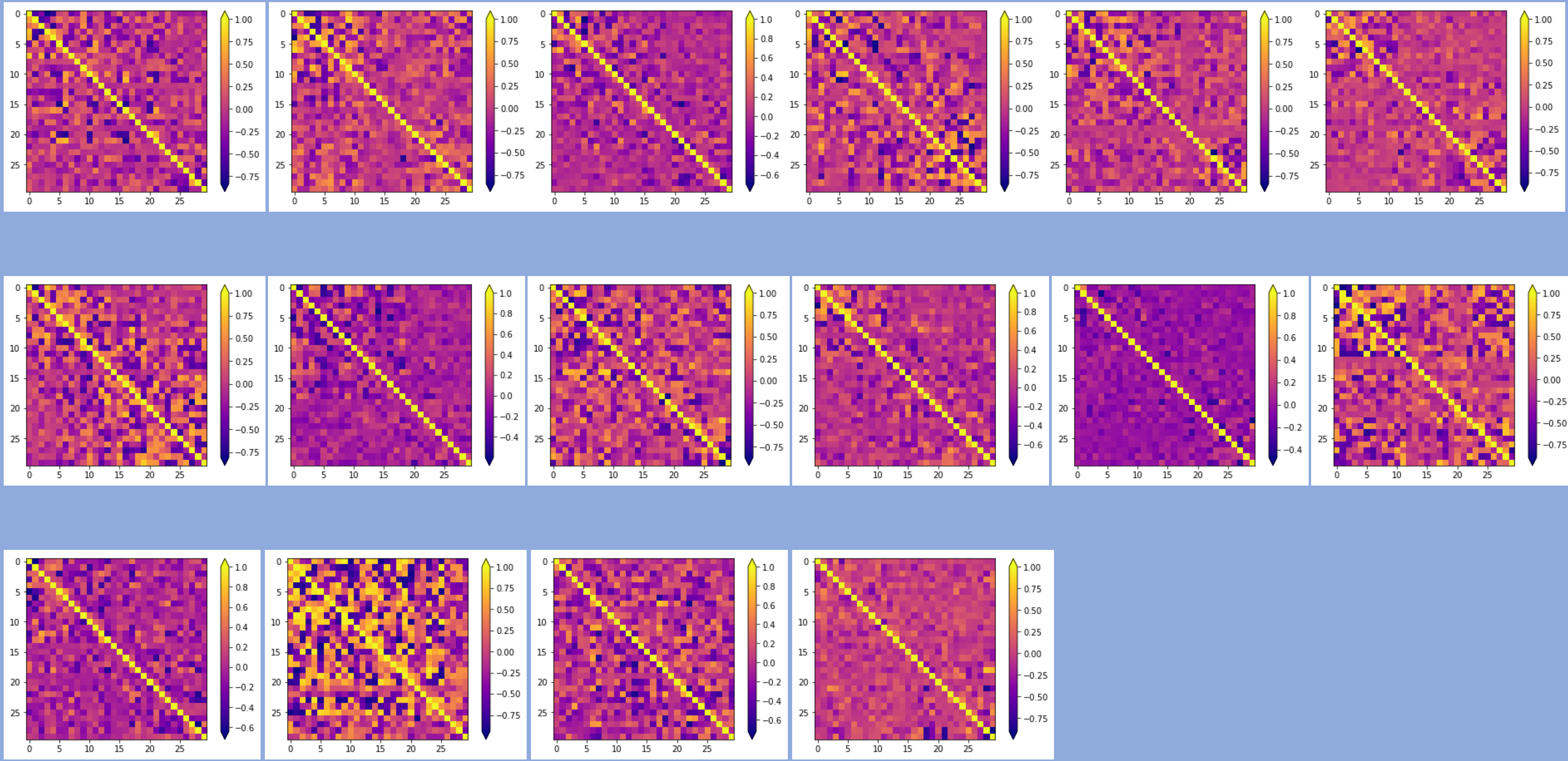
Heat map:

- Visualization of correlation matrices
- Brighter color suggests higher correlation coefficients

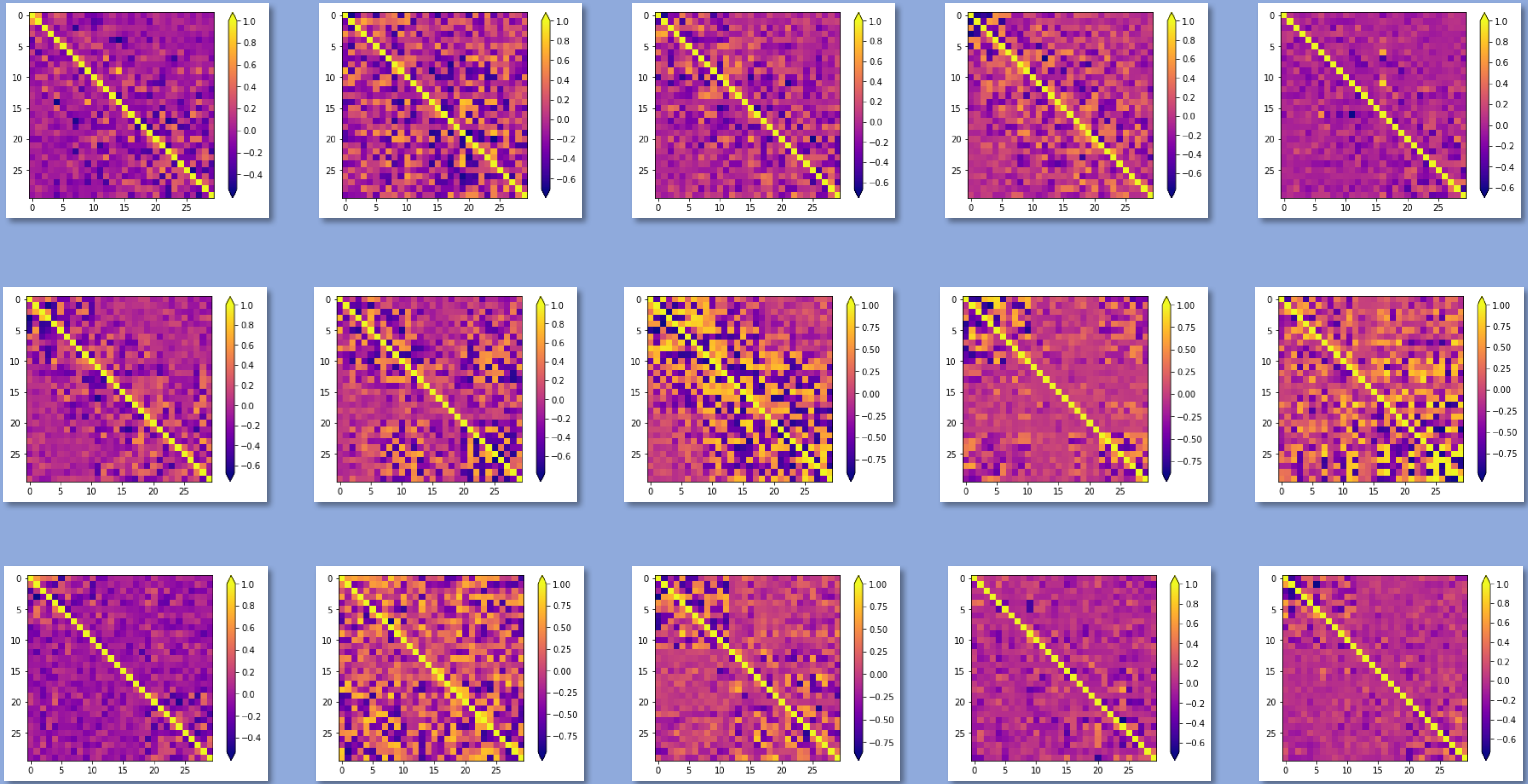




AD



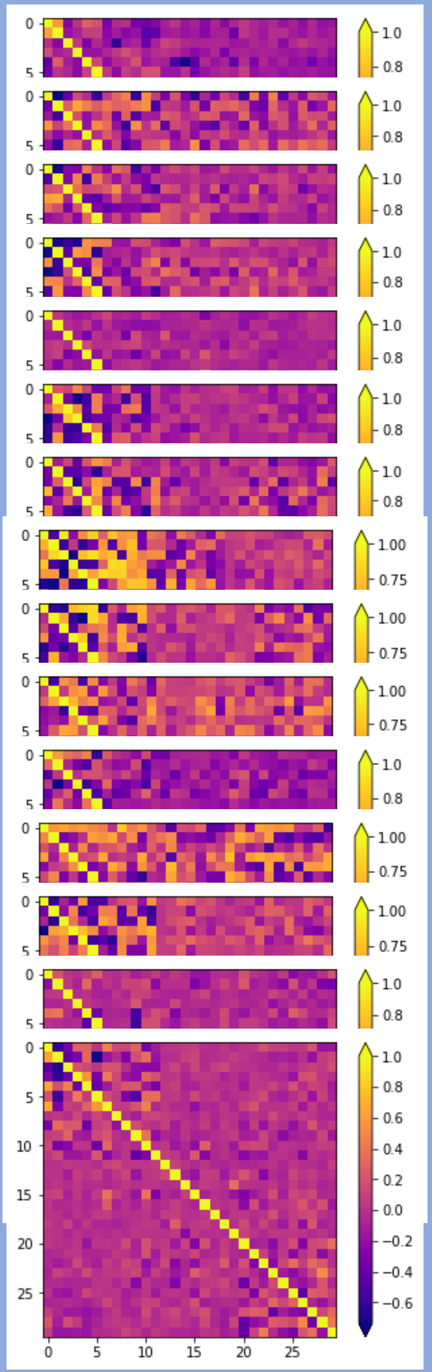
MCI



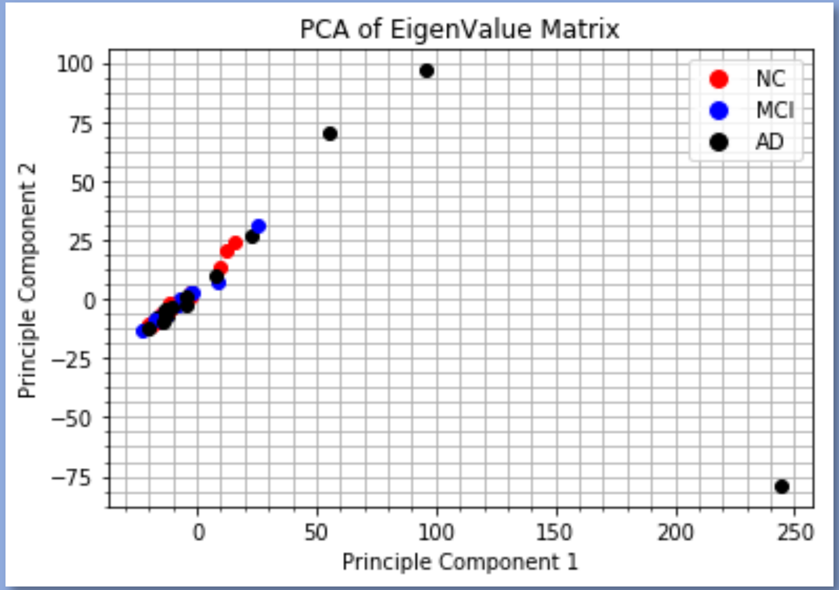
NC

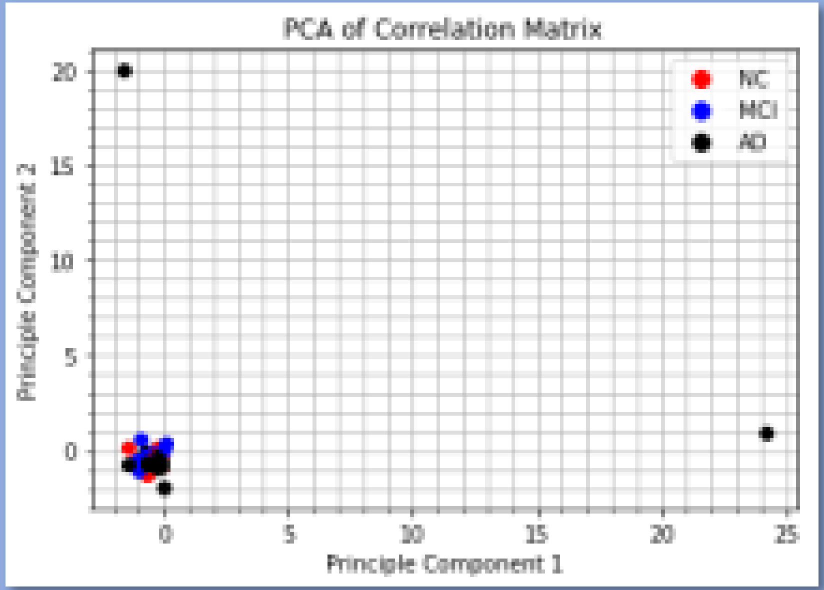
Image classification results:

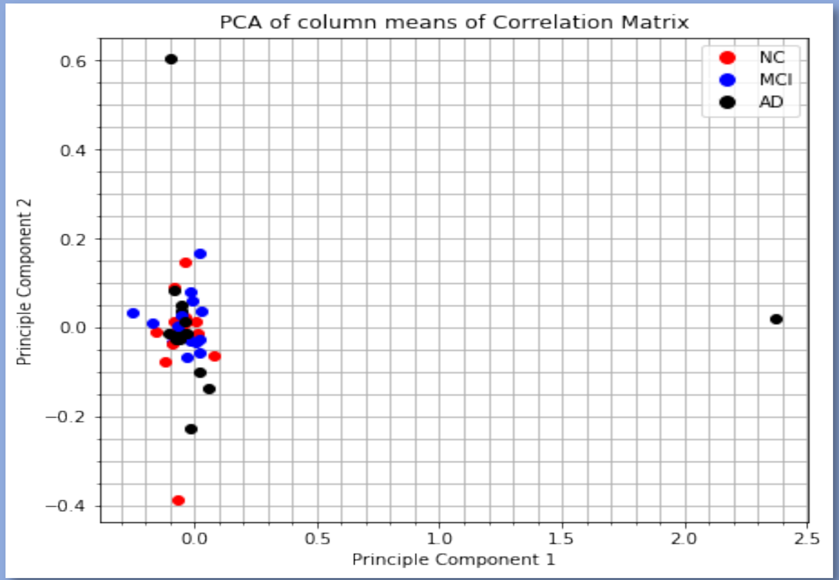
Predicted classes					
		NC	MCI	AD	
True classes	NC	5	4	6	33.3%
	MCI	3	10	3	62.5%
	AD	2	8	7	41.2%
		50%	45.5%	43.8%	Overall Acc: 45.8%

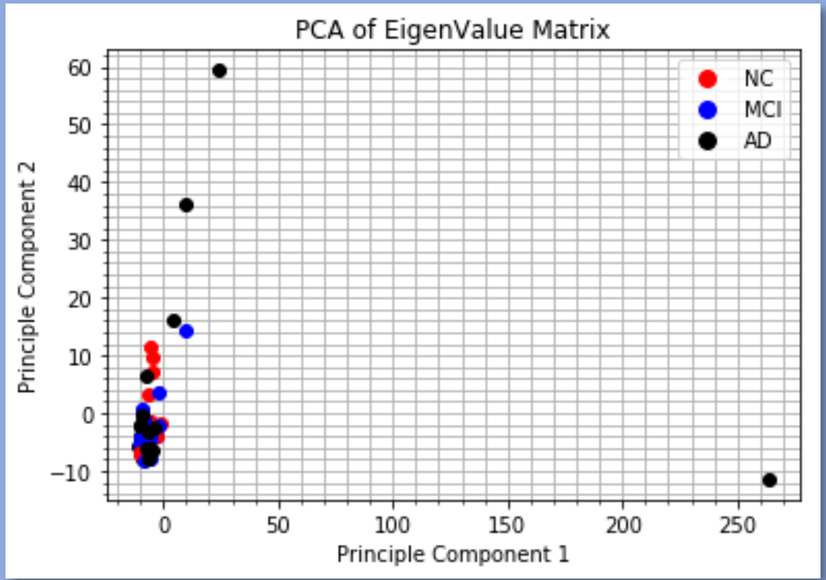


PCA results:

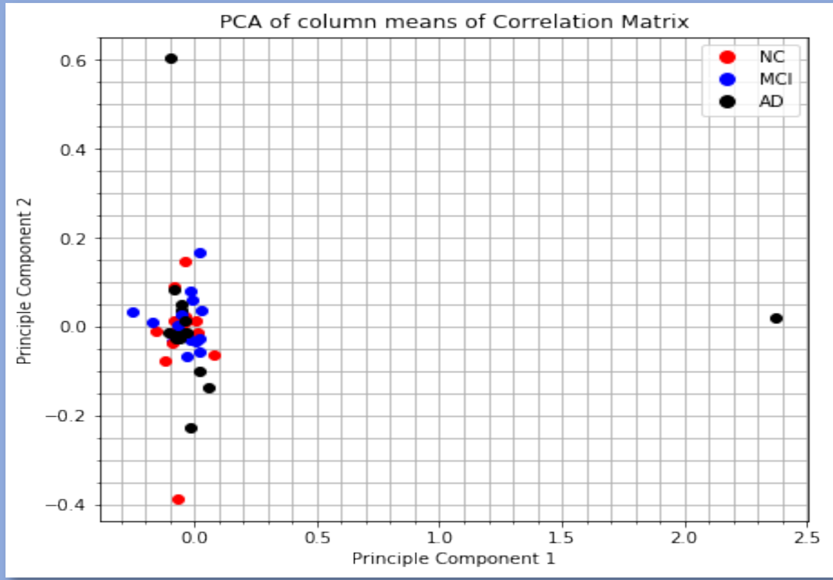

PCA of EigenValue Matrix
NC
MCI
AD
Principle Component 2
Principle Component 1


PCA of Correlation Matrix
NC
MCI
AD
Principle Component 2
Principle Component 1


PCA of column means of Correlation Matrix
NC
MCI
AD
Principle Component 2
Principle Component 1


PCA of EigenValue Matrix
NC
MCI
AD
Principle Component 2
Principle Component 1

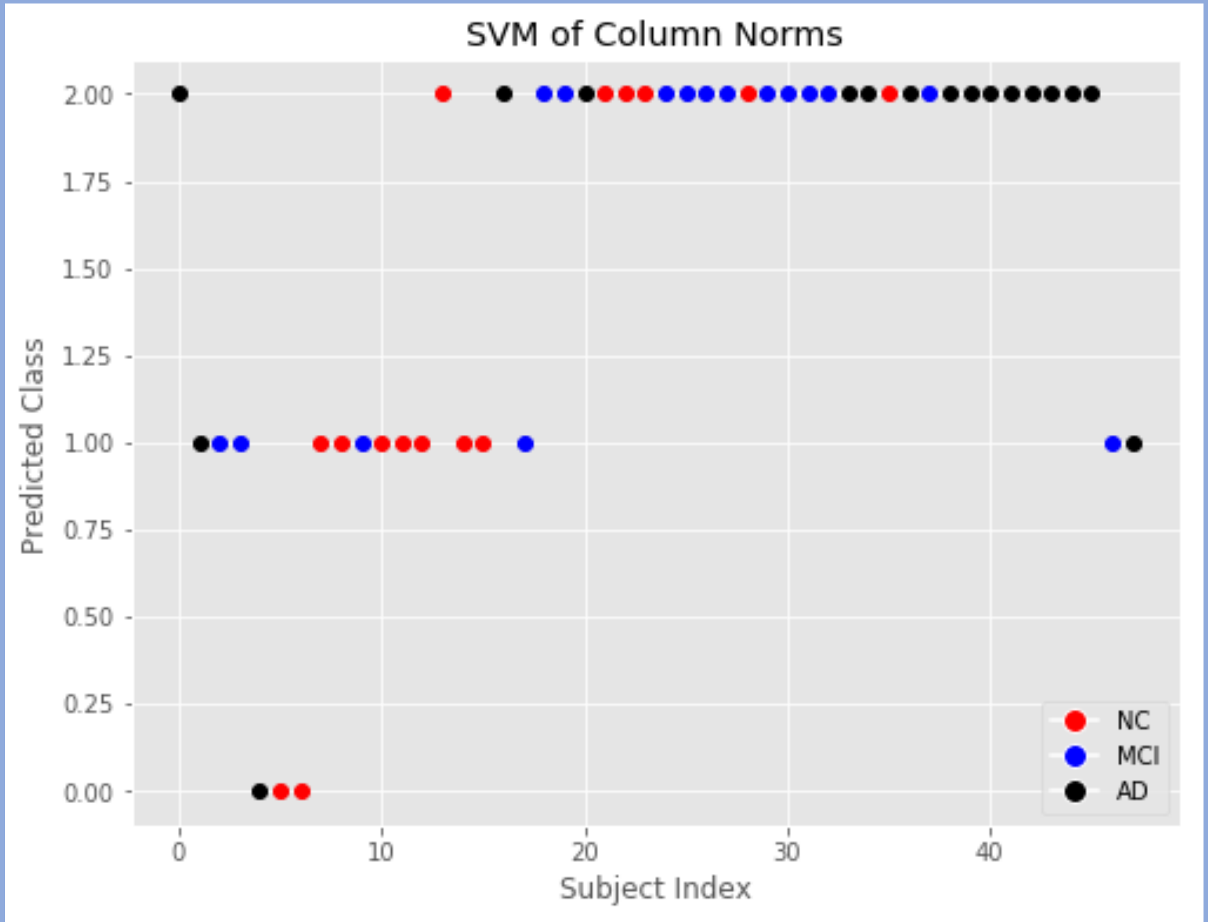
PCA results (Column means):

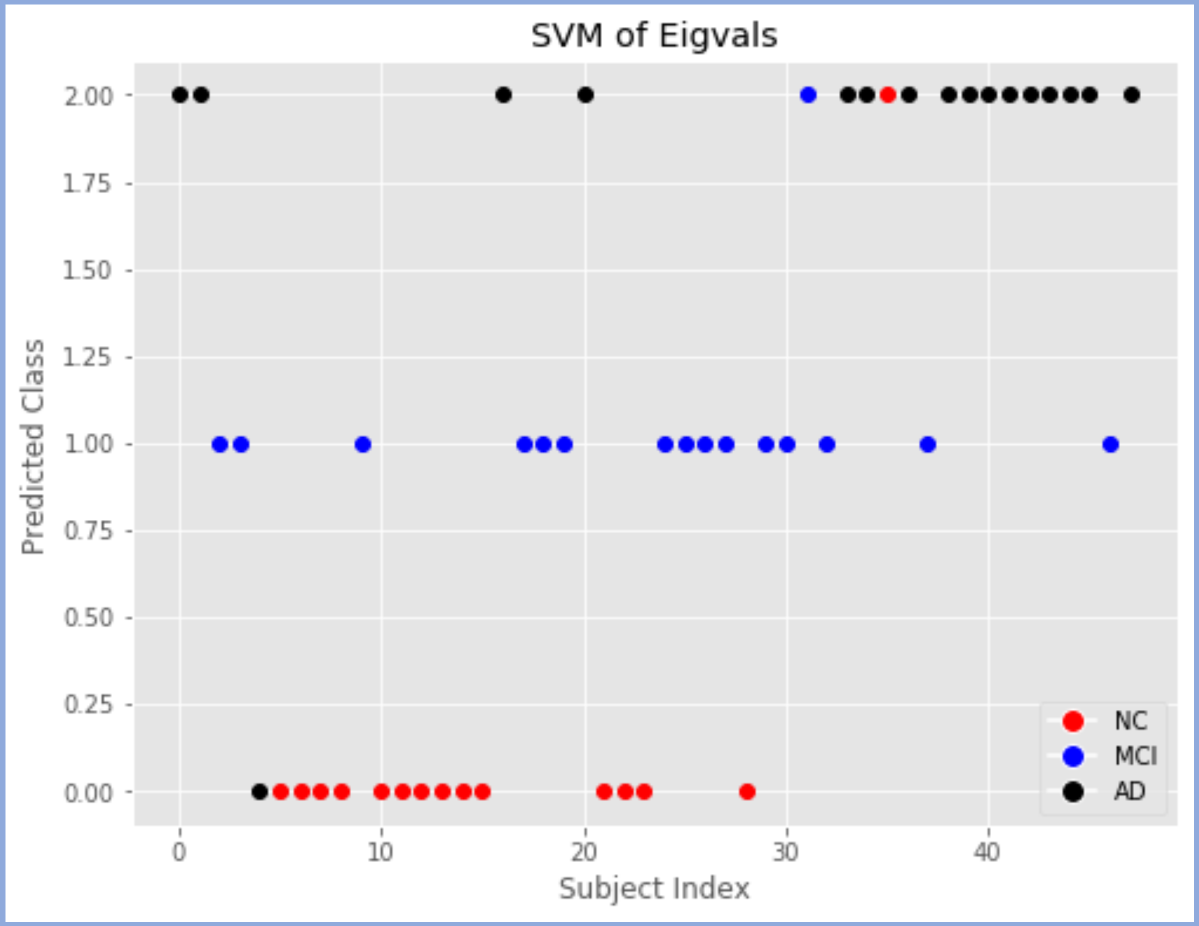


Column-means

		Predicted classes			
		NC	MCI	AD	
True classes	NC	2	1	12	13.33%
	MCI	0	5	11	31.25%
	AD	1	2	14	82.3%
		66.67%	62.5%	37.84%	Overall Acc: 43.75%

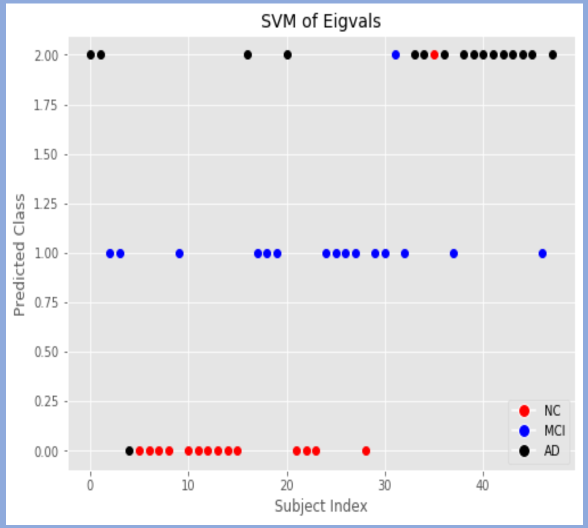
SVM models:


SVM of Column Norms
Predicted Class
Subject Index
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
NC
MCI
AD


SVM of Eigvals
Predicted Class
Subject Index
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
NC
MCI
AD

SVM models:

		Predicted classes			
		NC	MCI	AD	
True classes	NC	14	0	1	93.33%
	MCI	0	15	1	93.75%
	AD	1	0	16	94.12%
		93.33%	100%	88.89%	Overall Acc: 93.75%



Thanks for listening

Q & A