Internship Experience and Machine Learning Applications in Biomedical Engineering

Jeremiah Simmons

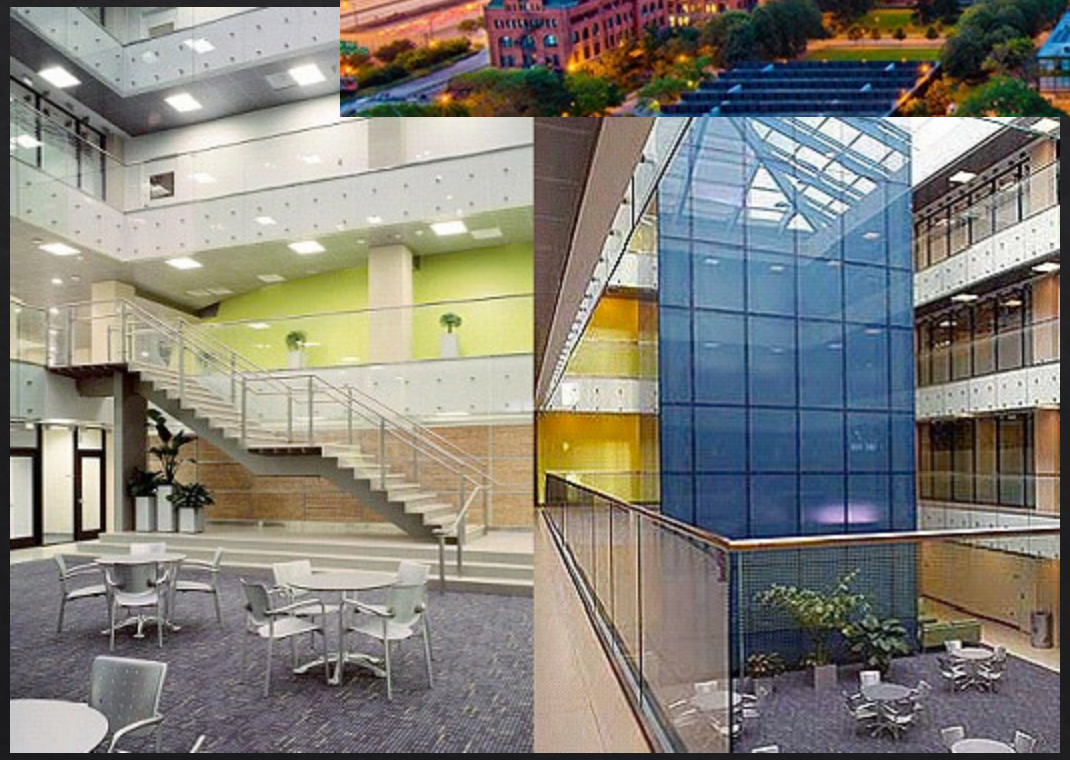
Outline

1. About me
2. Experiences
3. Machine Learning
4. Pattern Recognition
5. Application

A little about who I am

- Junior here at APSU
- Extracurricular: ROTC and Fiji
- Research: Recently joined the FSW team

Where I Went: Illinois Institute of Technology



MEDICAL IMAGING
RESEARCH CENTER

Experiences in Chicago

Other things I experienced

- White Sox Games
- Field Museum
- The Art Institute of Chicago
- Museum of Science and Industry
- Met a lot of different people

What is Machine Learning

- It is a type of artificial intelligence that provides computers with the ability to learn without being explicitly programmed. Focuses on the development of computer programs that can teach themselves to grow and/or change when exposed to new data.

What is Pattern Recognition

- Similar and almost synonymous with machine learning; it is, in actuality, a branch of machine learning that focuses on patterns and regularities of data.

Methods – Pattern Recognition

1. Vocabulary
2. Principal Component Analysis
3. Linear Discriminant Analysis

Vocabulary

- A feature vector is an n-dimensional vector of numerical features that represent some object
- Variance, informally, measures how far a set of numbers are spread out from their mean.
- A class is, for example, an instance, object type, or specific event. Usually labeled as a numerical integer: 0, 1, etc. But, sometimes a string name is used.
- Classifier refers to a program we will “train” to define a decision boundary using pattern analysis with a training set of test data that holds statistical significance. This classifier should be able to accept new data and apply the previously found decision boundary to assign a class label to that new data.

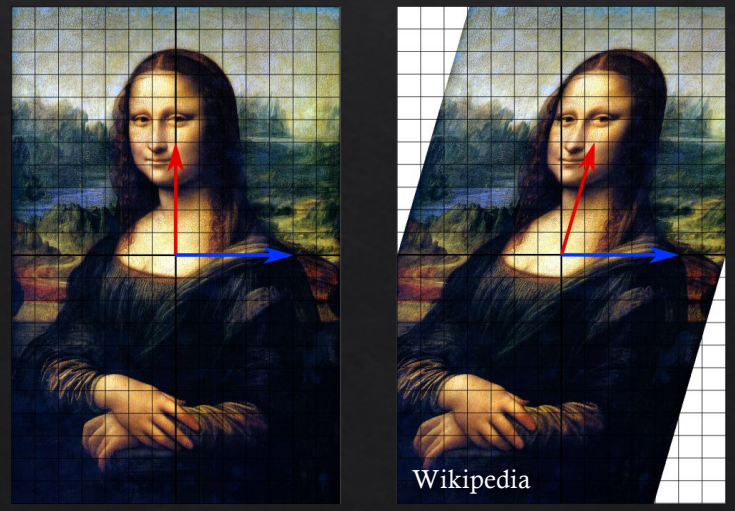
Example

Class Label	Height (inches)	Weight (lbs)
Male (Ricky)	71	240
Male (Jake)	76	190
Male (Anthony)	69	210
Male (Jacob)	69	130

Vocab Cont.

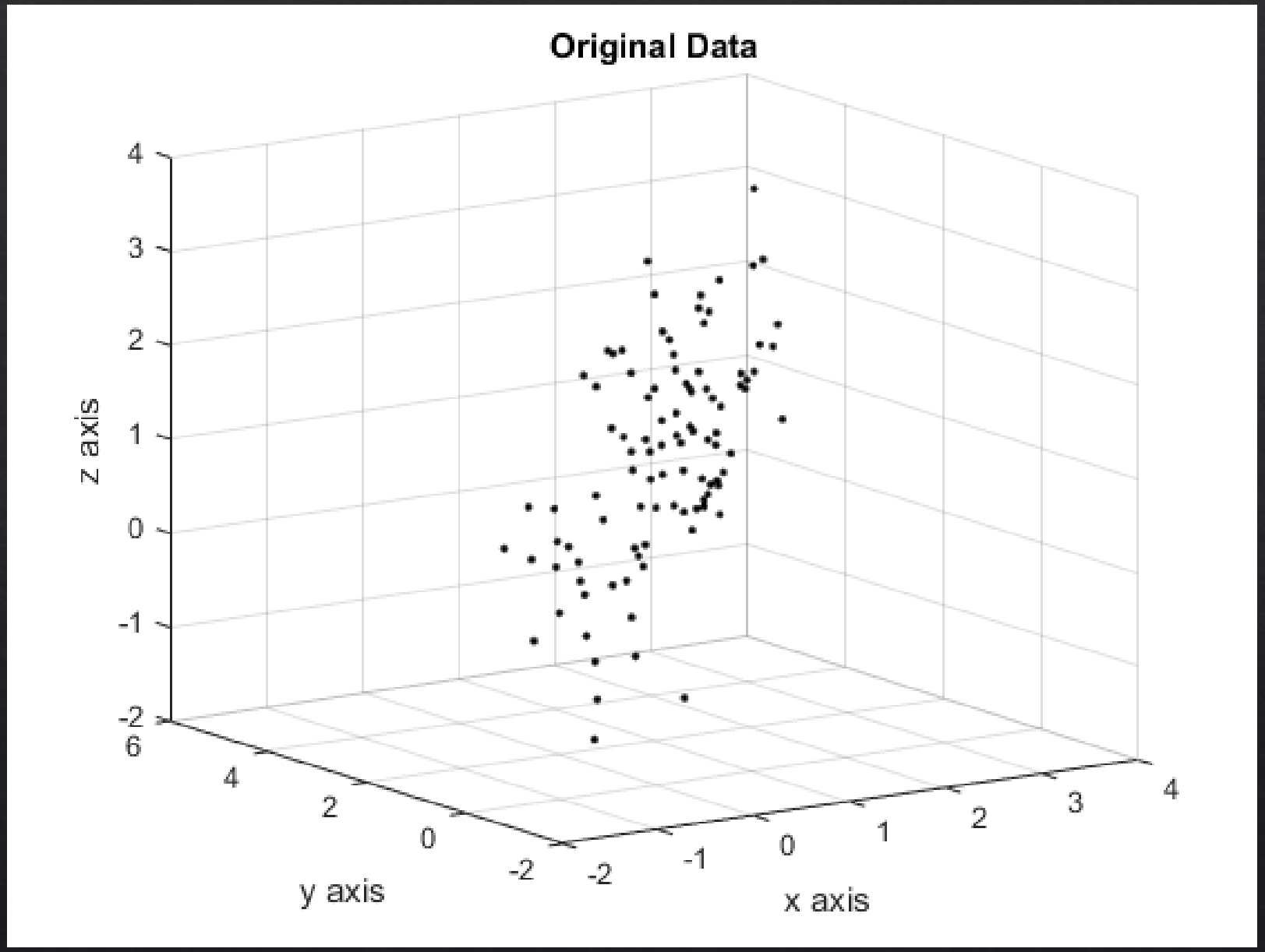
- Eigenvector or characteristic vector does not change directions when a linear transformation is applied. Geometrically an eigenvector, corresponding to a real nonzero eigenvalue, points in a direction that is stretched by the transformation.

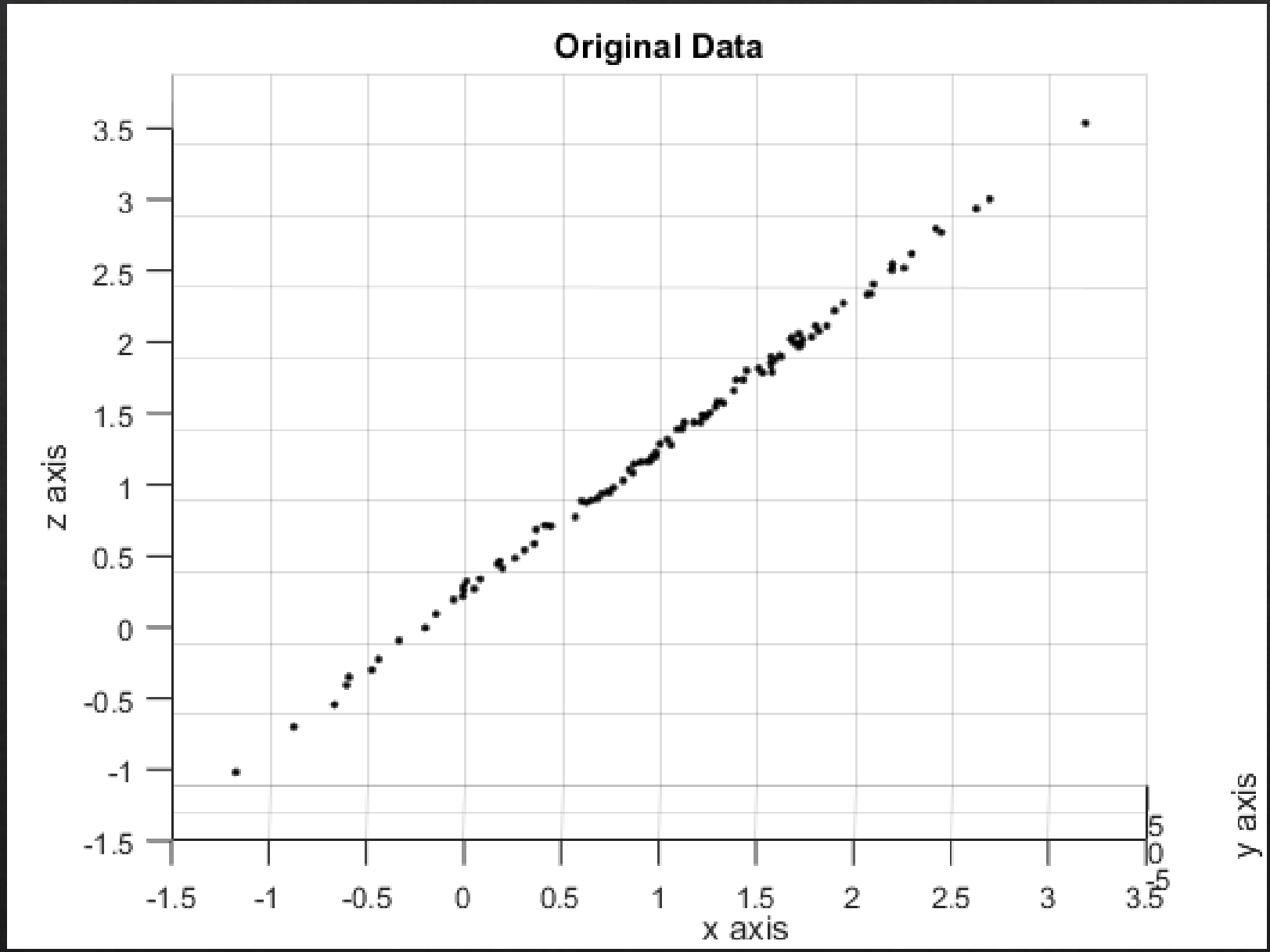
- Eigenvalue: A scalar number associated with an specific eigenvector, and determines how much a transformation stretches or scales an eigenvector. Can be zero.

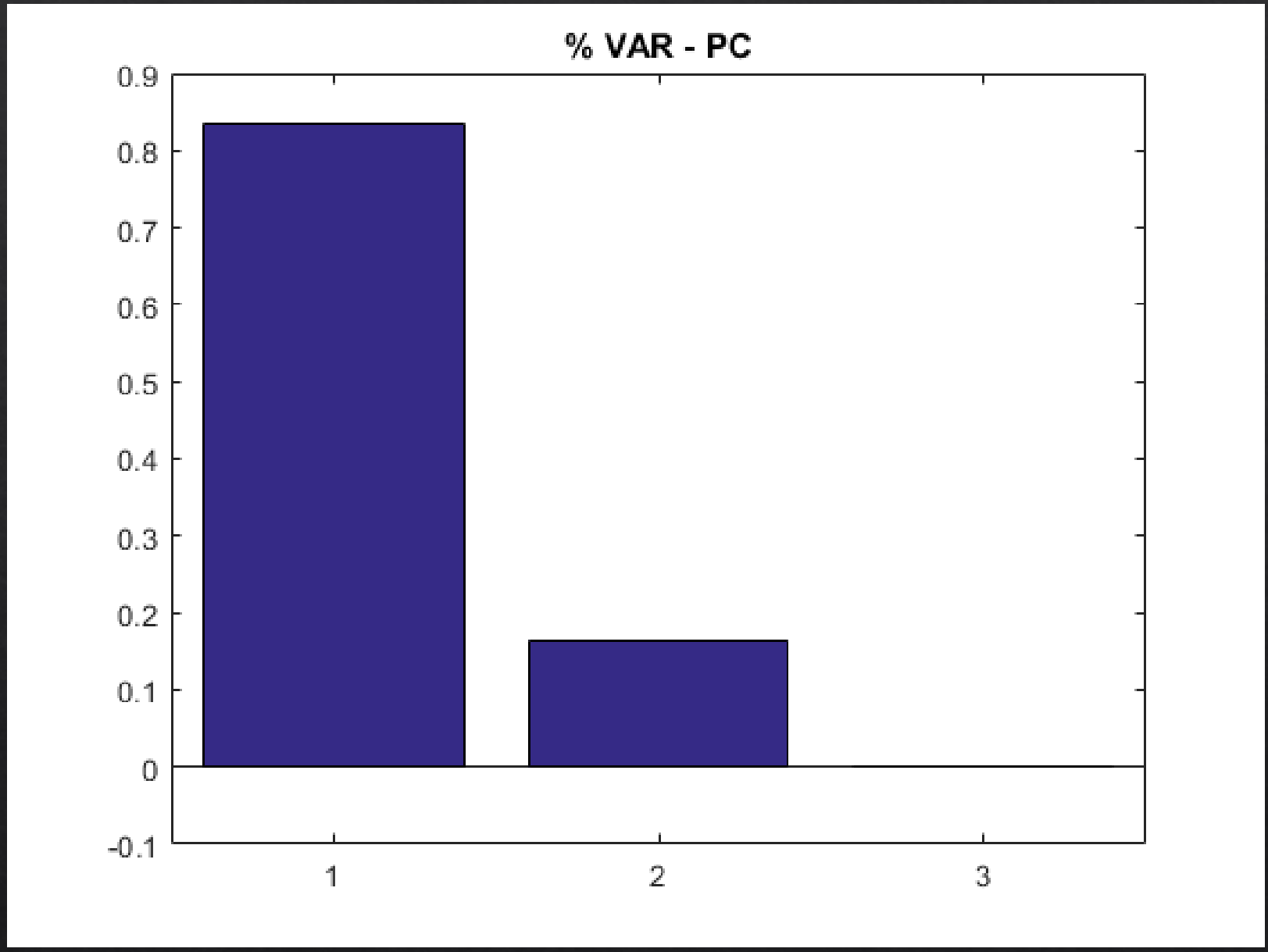


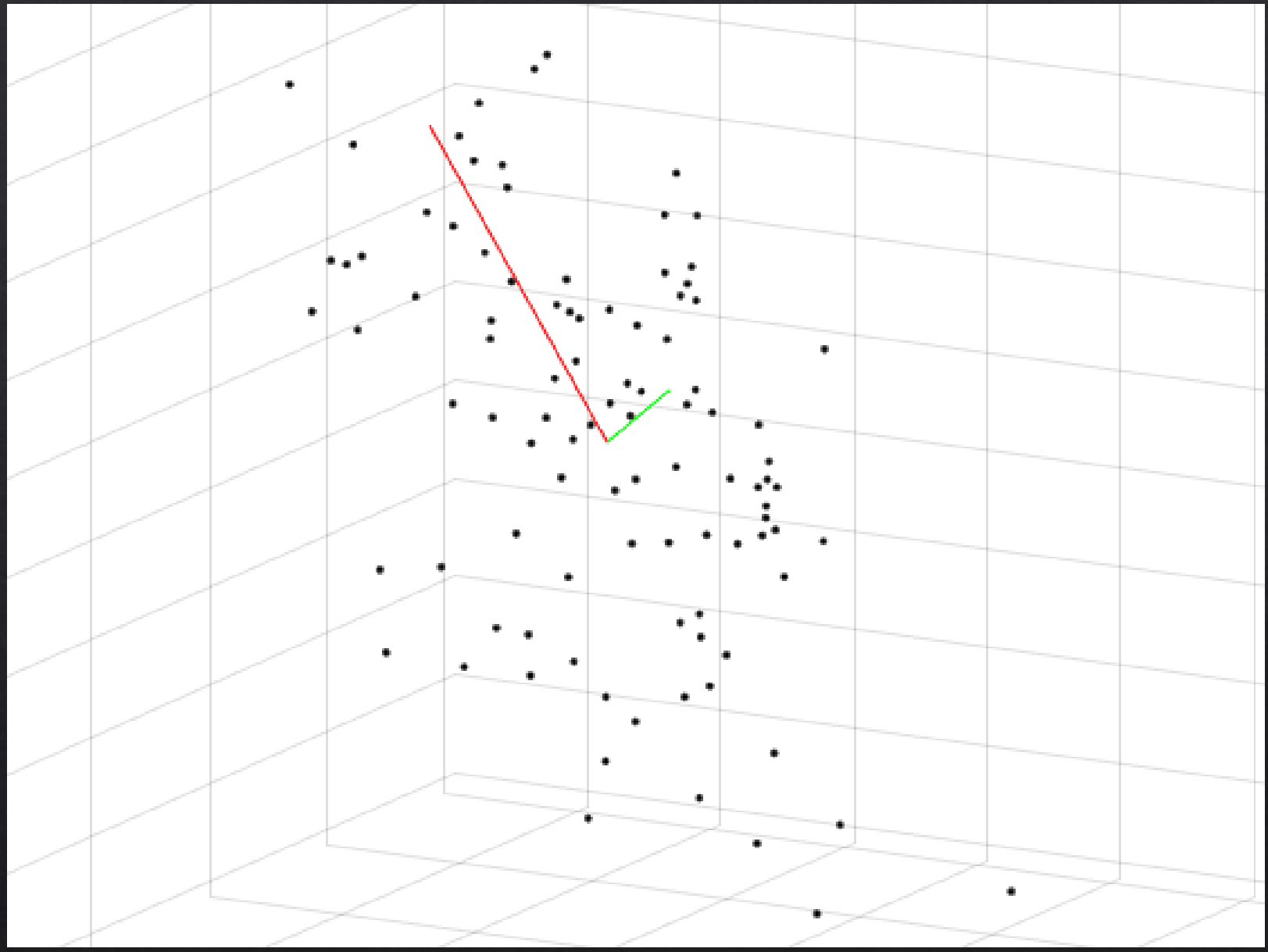
Principle Component Analysis

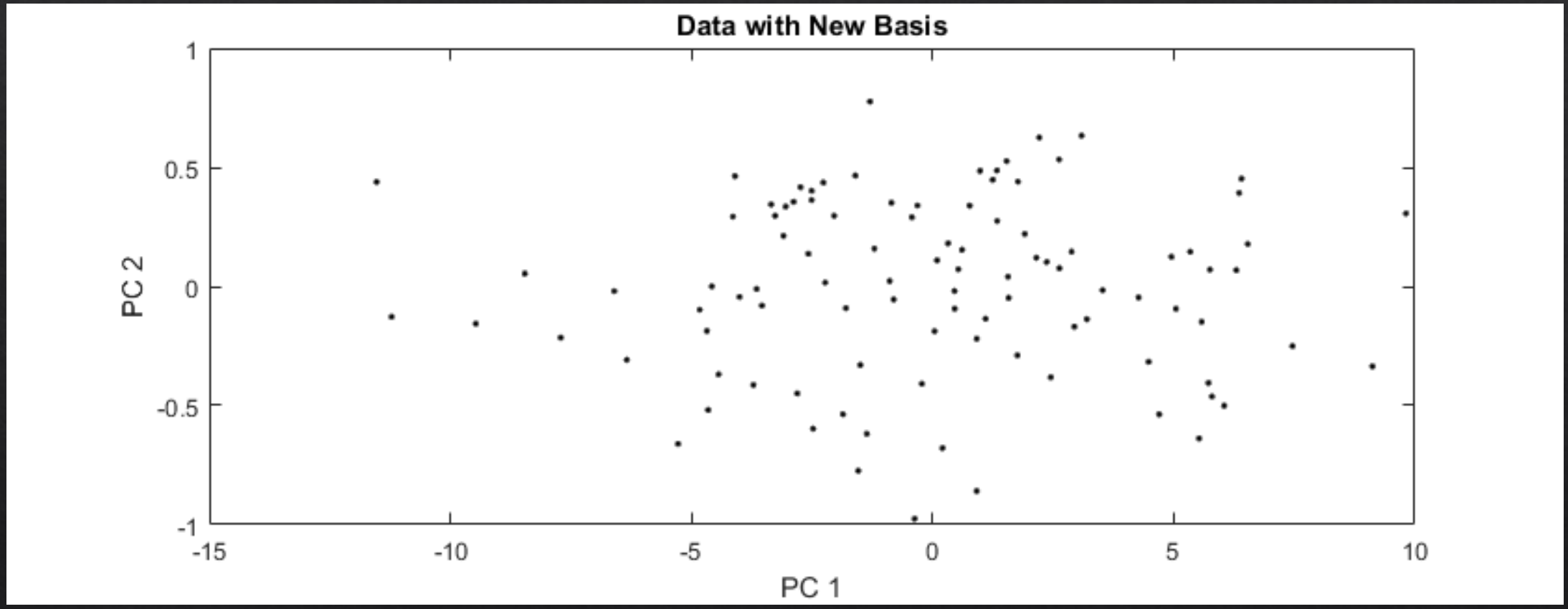
- The goal is to find dimensions that are shared across all features vectors that contain good, representative data for the whole training set. Essentially PCA will reduce the data down to a smaller amount of dimensions while losing little to no statistical significance within the set.
  - Redundant data
  - Invariances


Original Data
z axis
4
3
2
1
0
-1
-2
6
4
2
0
-2
y axis
-2
-1
0
1
2
3
4
x axis


Original Data
z axis
x axis
y axis
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
3
3.5


% VAR - PC
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
1
2
3

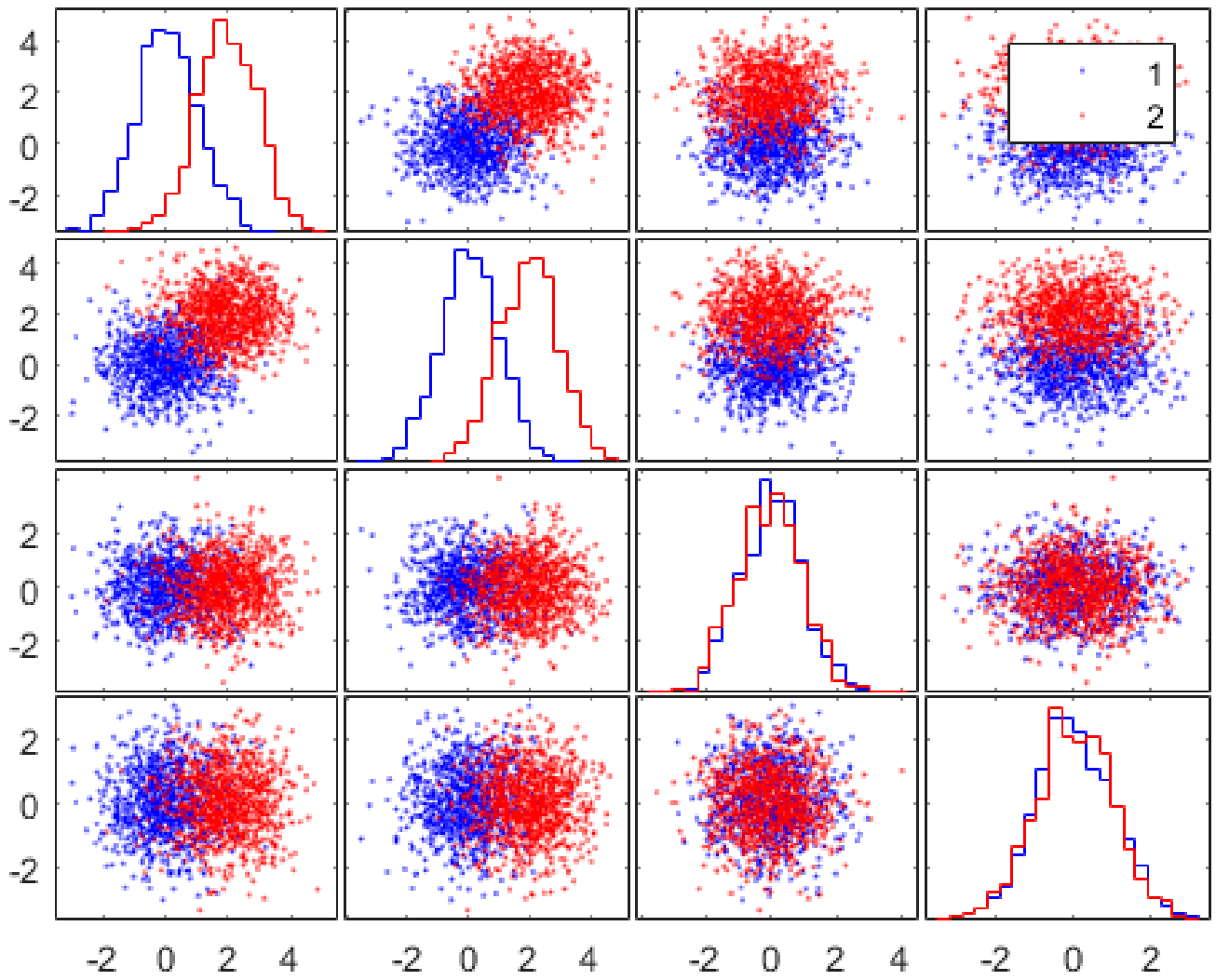


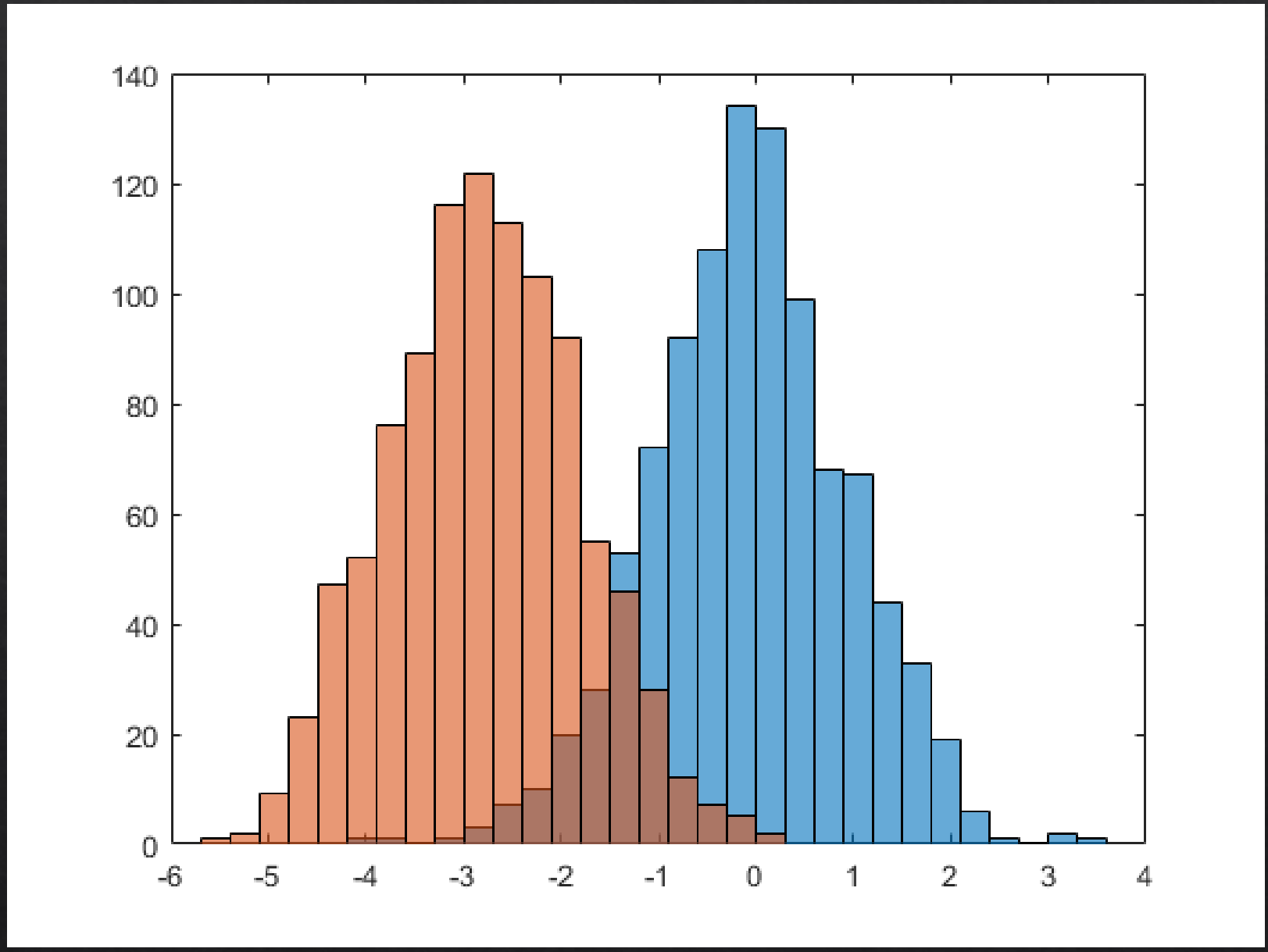

Data with New Basis
1
0.5
0
-0.5
-1
PC 2
-15
-10
-5
0
5
10
PC 1

Linear Discriminant Analysis

- A method used to find a linear combination of features that characterizes or separates two or more multivariate classes
- LDA is a generalized version of Fisher's linear discriminate analysis.
- LDA or FDA work best with classes that are multivariate and have Gaussian distribution (normal distribution).

FLD video


4
2
0
-2
1
2
-2 0 2 4


140
120
100
80
60
40
20
0
-6
-5
-4
-3
-2
-1
0
1
2
3
4

Applications of PCA and FDA/LDA

1. Scenario
2. What is a PET Scan
3. Training a Classifier
4. Performance (Confusion) Matrices
5. Training Methods

Scenario

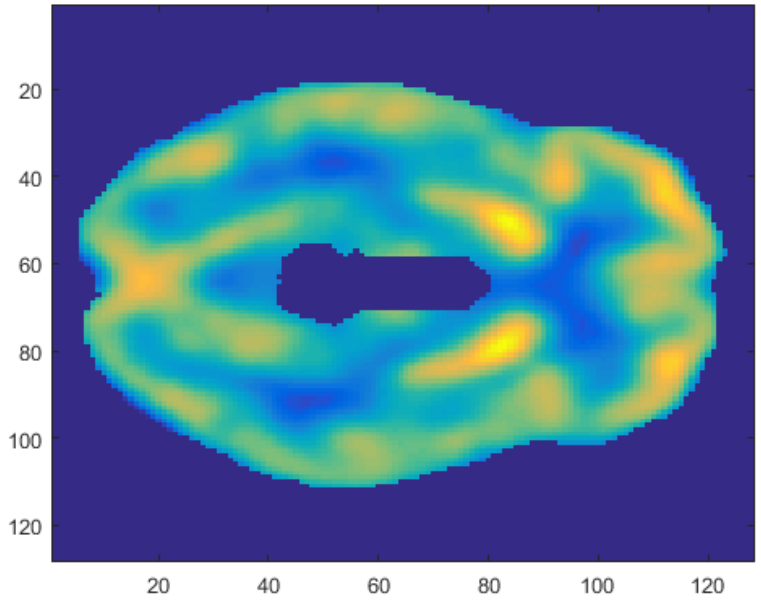
- We want to build a classifier that will classify between normal (NL) and Alzheimer's Disease (AD) via PET (Positron Emission Tomography) scans.
- AD is a serious disease that results in loss of memory, and major changes in speaking, personality, and ability to function generally.
- Imagine we have either been given scans from a database with correct diagnosis or have conducted our own studies in conjunction with a local hospital beforehand.

What is a PET Scan

- A PET scan measures glucose metabolism (utilization of sugar) within the brain. In AD patients, key brain regions stop function normally, and stop consuming resources at the normal rate.

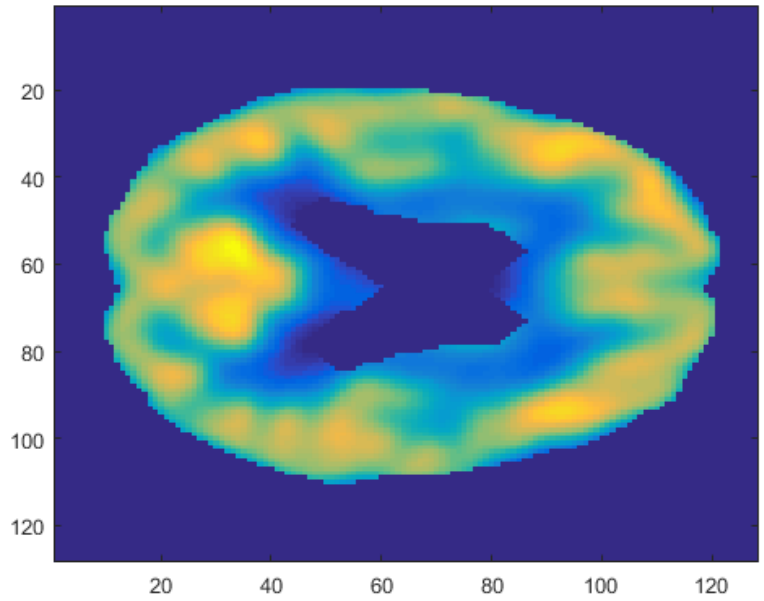
- Key brain areas are usually within the following three lobes: parietal, temporal, and frontal.

- A full scan creates images of individual slices along the brain. These slices combine to create a false 3d image of the patient's brain.

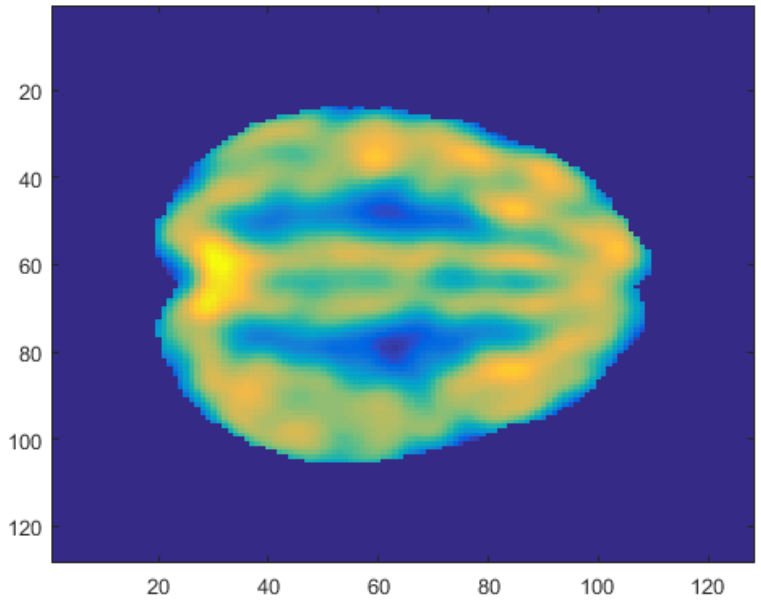


31

48

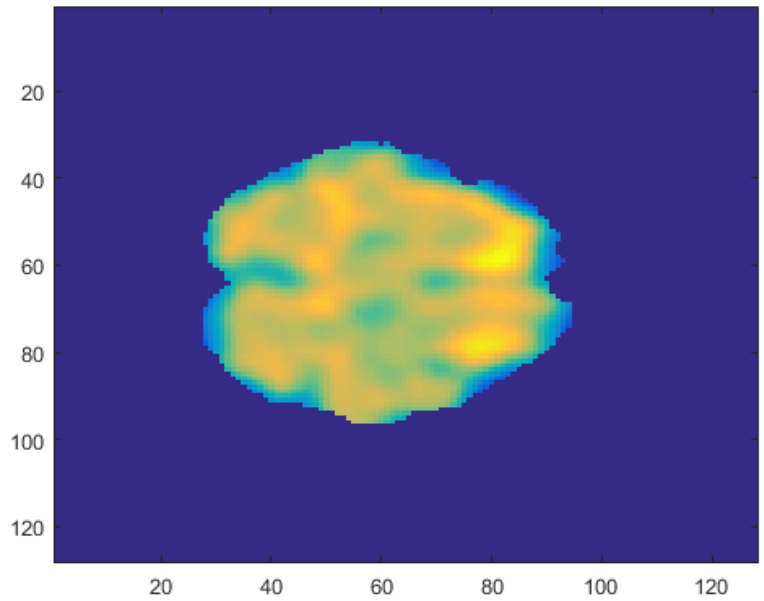


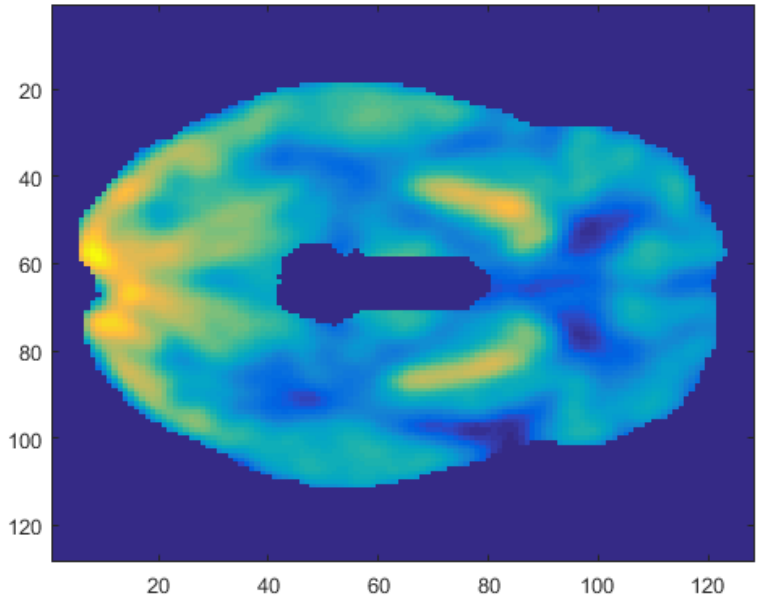
No Alzheimer's



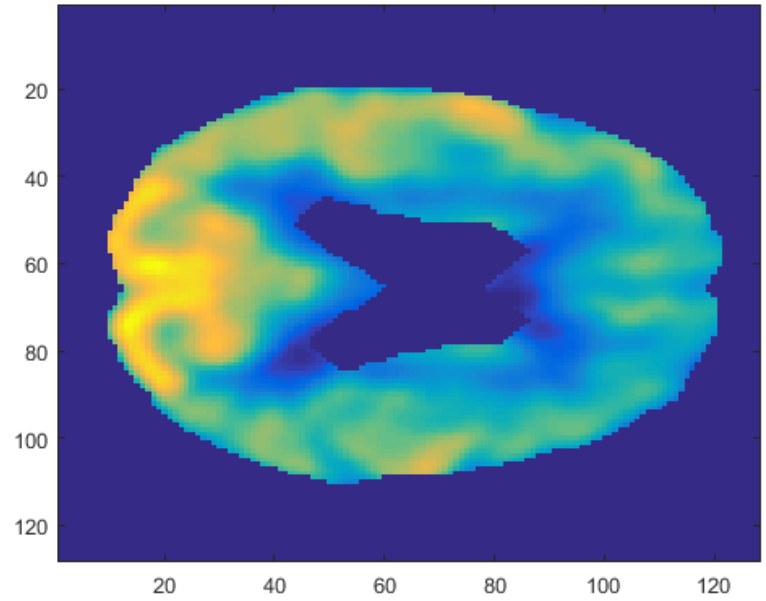
67

78



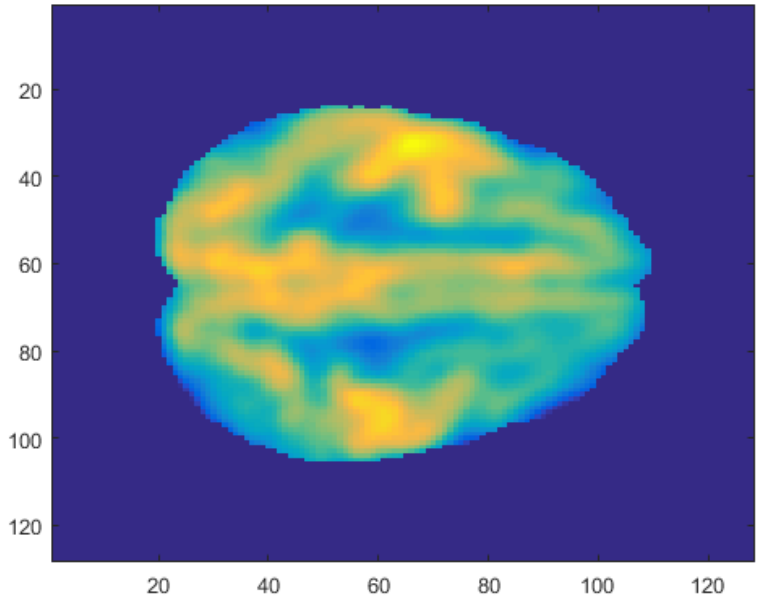


31

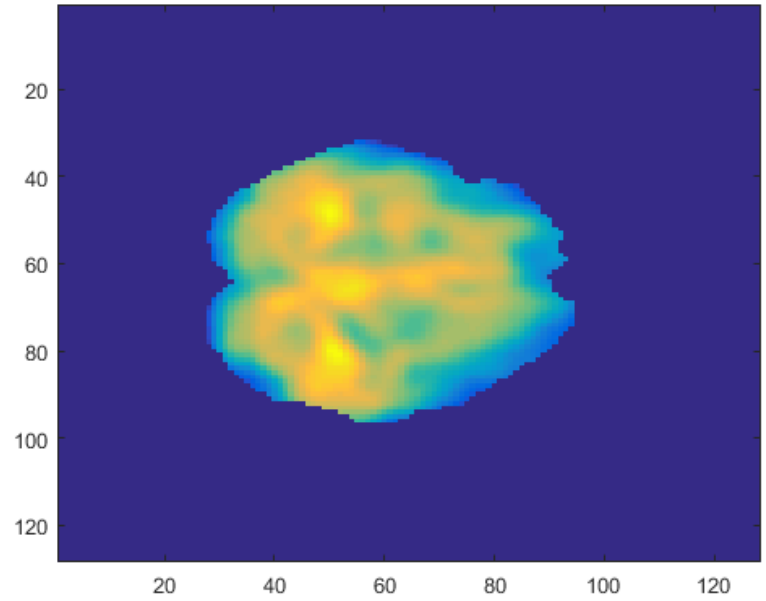


48

Alzheimer's



67



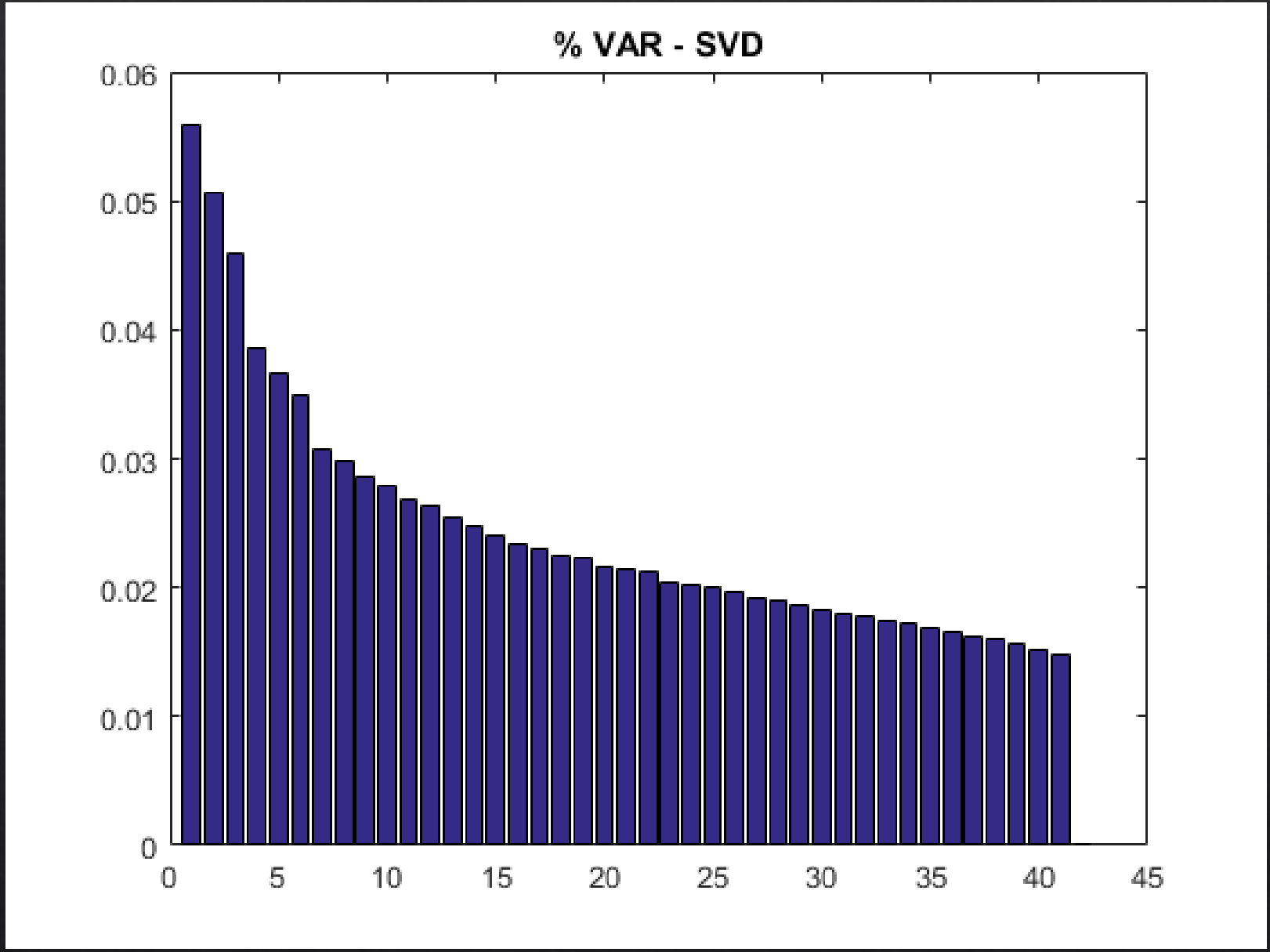
78

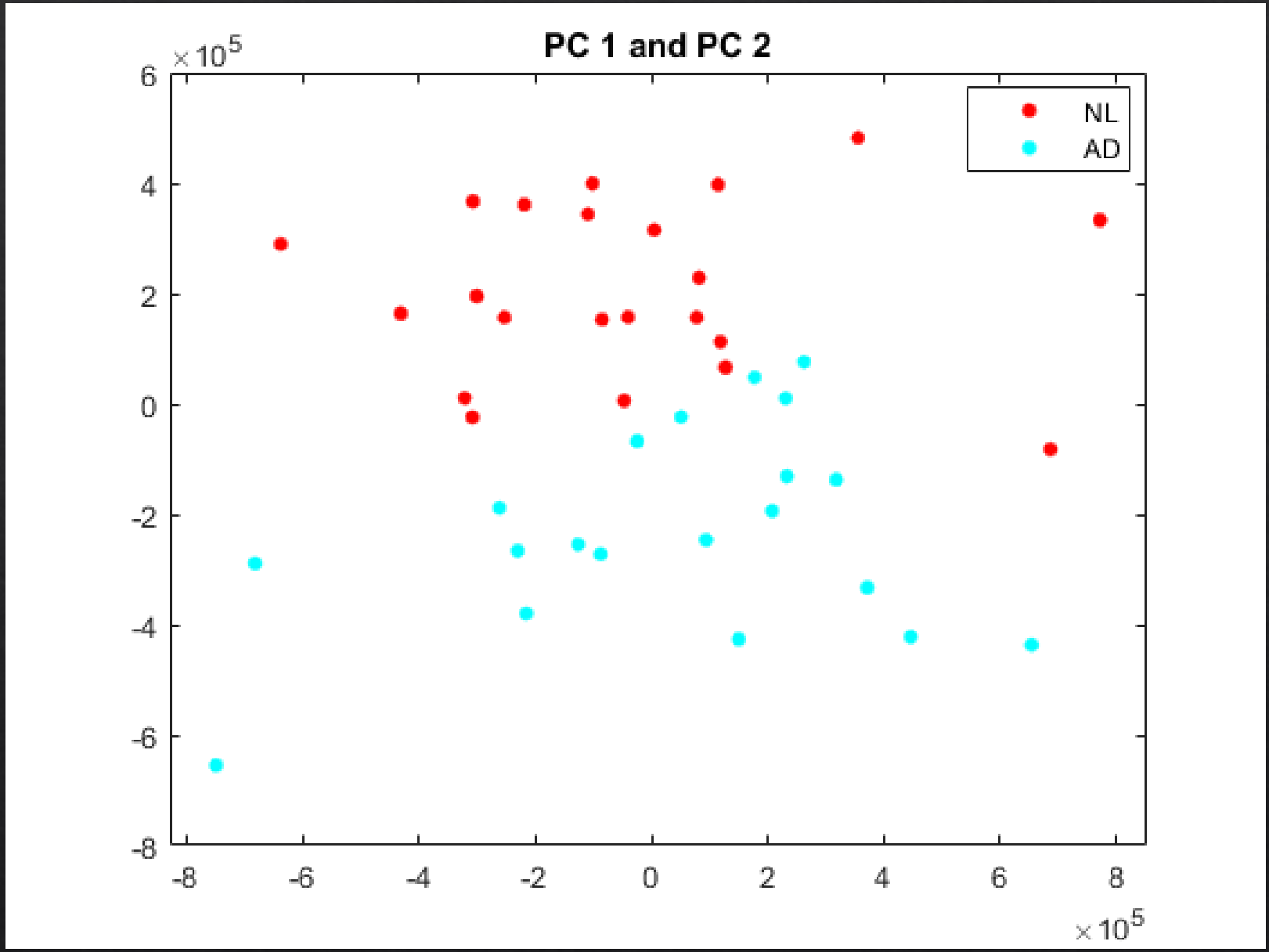
Training a Classifier

1. PCA
2. LDA

PCA

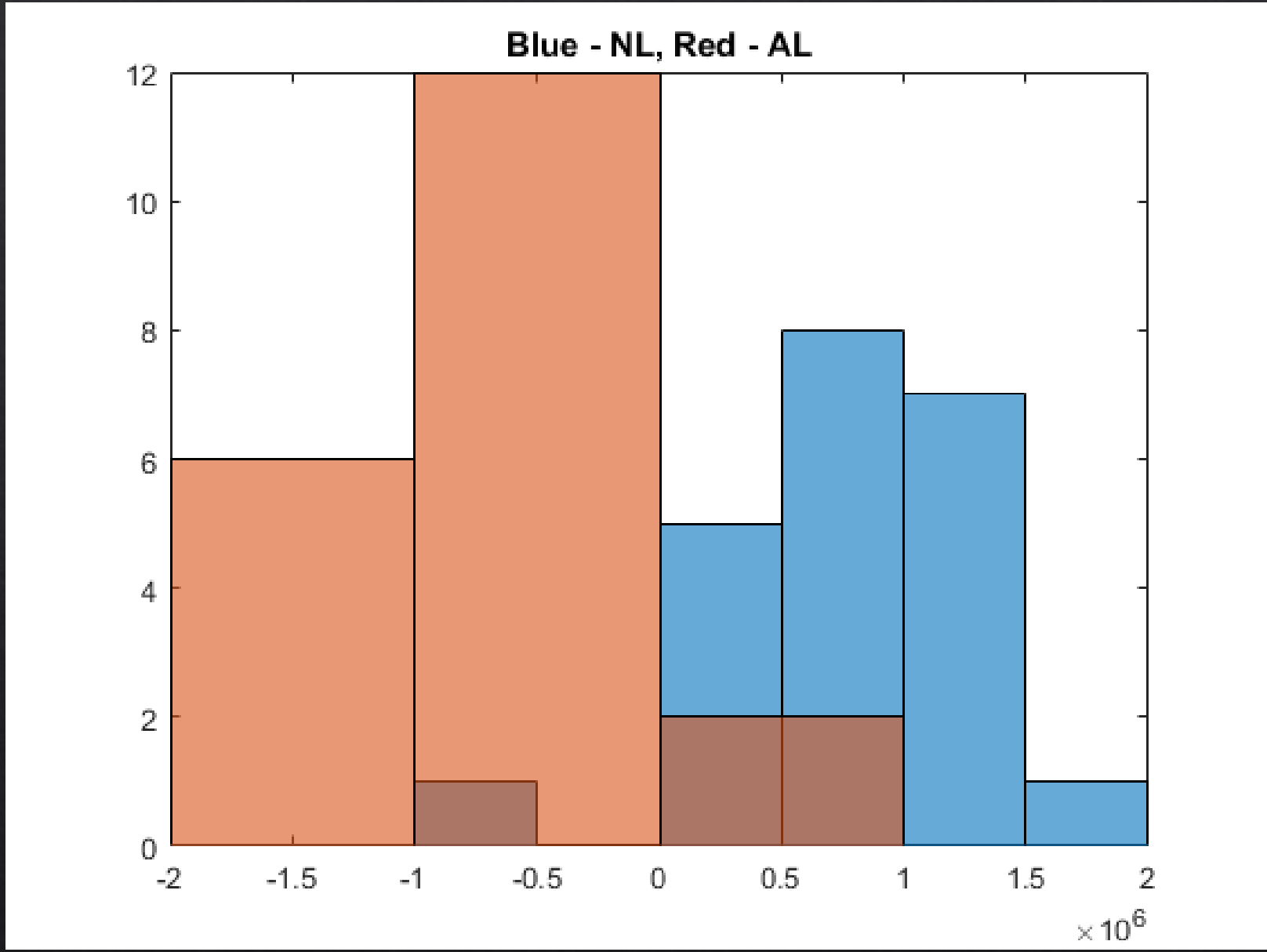
- Image a single scan, what are the images made of?
  - Pixels.
- How is a pixel color determined?
  - By a numerical value, whether it be hexadecimal, decimal, or some other base-n system.
- A single slice image in our training set is 128 by 128 pixels. And there are 91 total slices.
- We use PCA to eliminate all redundant and invariant pixel values.


% VAR - SVD
0.06
0.05
0.04
0.03
0.02
0.01
0
0
5
10
15
20
25
30
35
40
45


PC 1 and PC 2
×10^5
6
4
2
0
-2
-4
-6
-8
-8
-6
-4
-2
0
2
4
6
8
×10^5
NL
AD

LDA

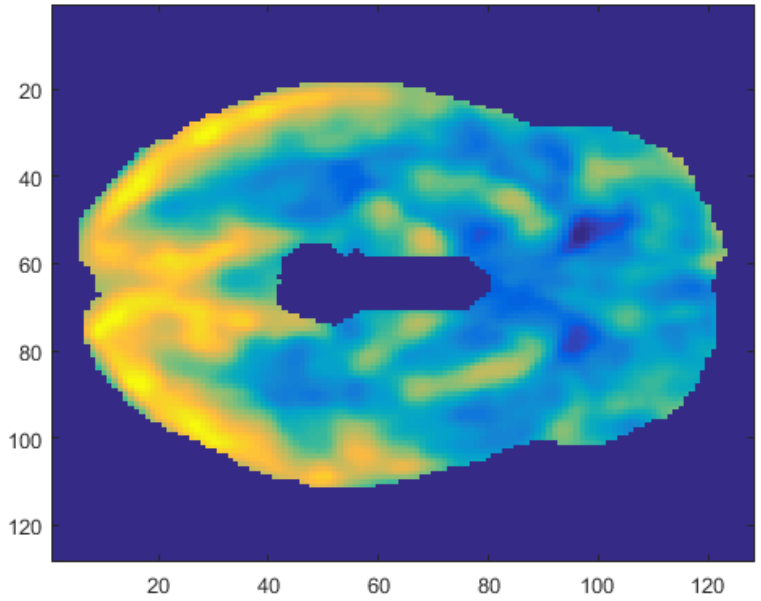
- From looking at two of the dimensions in the PC modified-data it is clear that there may be some linear separation within the training set.


Blue - NL, Red - AL
12
10
8
6
4
2
0
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
×10^6

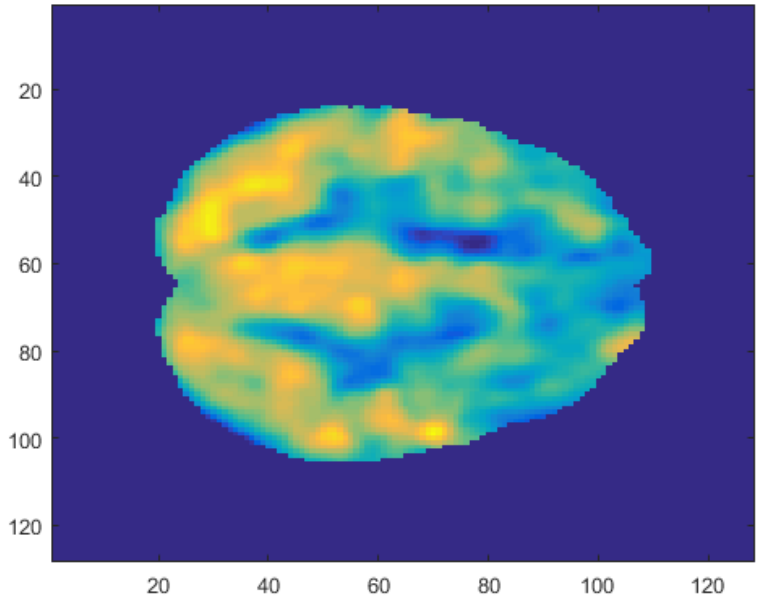
Discriminant Vector

- After applying discriminant analysis, we can define a decision boundary or discriminant vector.

- We then work backwards from the simplified vector and convert it back into a usable PET-like scan.

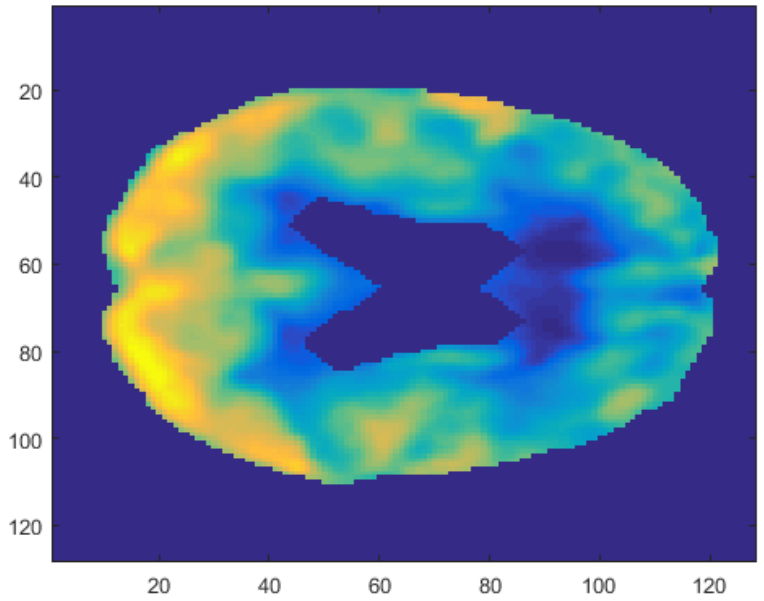

20
40
60
80
100
120
20
40
60
80
100
120

31

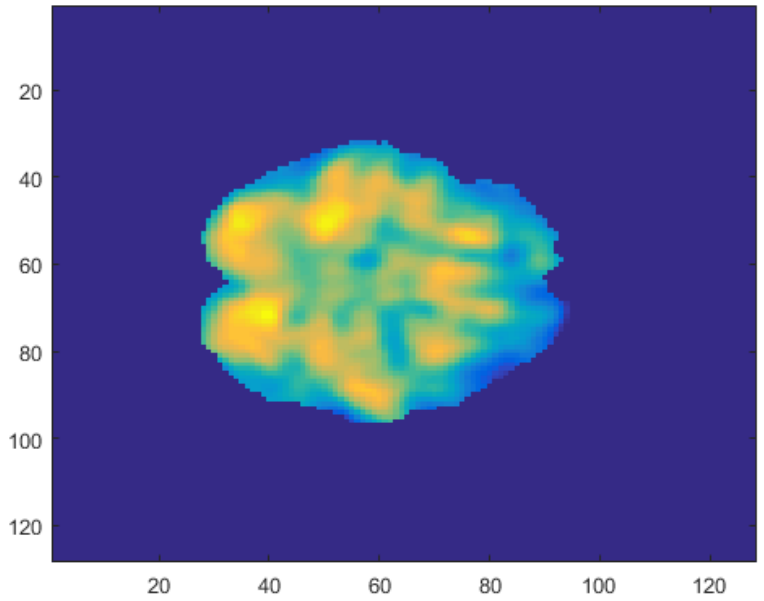

20
40
60
80
100
120
20
40
60
80
100
120

67

48


20
40
60
80
100
120
20
40
60
80
100
120

78


20
40
60
80
100
120
20
40
60
80
100
120

Performance (Confusion) Matrices

- This allows you to determine the performance of your training set and classifier.
- TP = True Positive
- TN = True Negative
- FP = False Positive
- FN = False Negative

TP	**FN**
FP	TN

Other Methods

- Cluster Analysis
- Convex Analysis (also known as a Support Vector Machine (SVM))
- Bayesian Decision Models
- Neural Networks

Questions?