Statistical Learning STA4042

INTRODUCTION

SCHOOL OF DATA SCIENCE

What is statistical learning ?

Historically,

Statistical Learning STA4042

INTRODUCTION

SCHOOL OF DATA SCIENCE

What is statistical learning ?

Historically,

Begining of XIX$^{\text{th}}$ century: least-square method (linear regression)

What is statistical learning ?

Historically,

Begining of XIXth century: least-square method (linear regression)

1930's: linear discriminent analysis

Statistical Learning STA4042

INTRODUCTION

SCHOOL OF DATA SCIENCE

What is statistical learning ?

Historically,

Begining of XIXth century: least-square method (linear regression)

1930's: linear discriminent analysis

1990's: Support vector machines, beginning of non linear methods.

Statistical Learning STA4042

INTRODUCTION

SCHOOL OF DATA SCIENCE

What is statistical learning ?

Historically,

Begining of XIXth century: least-square method (linear regression)

1930's: linear discriminent analysis

1990's: Support vector machines, beginning of non linear methods.

2012: Success of Neural networks on Mnist data base

Statistical Learning STA4042

INTRODUCTION

SCHOOL OF DATA SCIENCE

What is statistical learning ?

Historically,

Begining of XIXth century: least-square method (linear regression)

1930's: linear discriminent analysis

exists since the 1980's!

1990's: Support vector machines, beginning of non linear methods.

2012: Success of Neural networks on Mnist data base

Statistical Learning STA4042

INTRODUCTION

SCHOOL OF DATA SCIENCE

What is statistical learning ?

Historically,

Begining of XIXth century: least-square method (linear regression)

1930's: linear discriminent analysis

1990's: Support vector machines, beginning of non linear methods.

2012: Success of Neural networks on Mnist data base

But what is it?

Statistical Learning STA4042

INTRODUCTION

SCHOOL OF DATA SCIENCE

What is statistical learning ?

Historically,

Begining of XIX$^{\text{th}}$ century: least-square method (linear regression)

1930's: linear discriminent analysis

1990's: Support vector machines, beginning of non linear methods.

2012: Success of Neural networks on Mnist data base

But what is it?

Popular definition of machine learning due to Tom Mitchell:

"A computer program is said to learn from experience E with respect to some class of tasks T, and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

What is statistical learning ?

Historically,

Begining of XIX$^{\text{th}}$ century: least-square method (linear regression)

1930's: linear discriminent analysis

1990's: Support vector machines, beginning of non linear methods.

2012: Success of Neural networks on Mnist data base

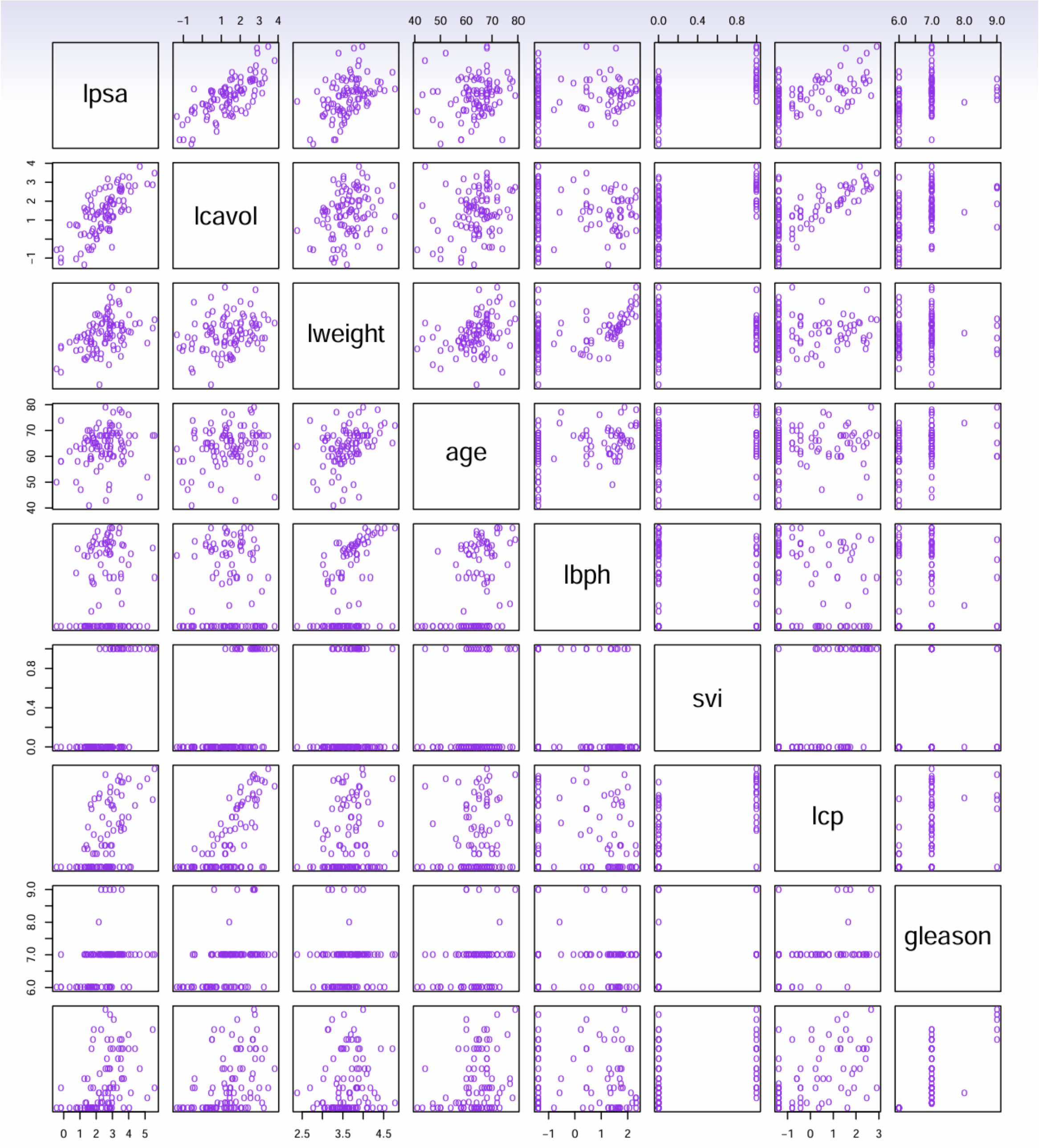
But what is it?

Popular definition of machine learning due to Tom Mitchell:

"A computer program is said to learn from experience E with respect to some class of tasks T, and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

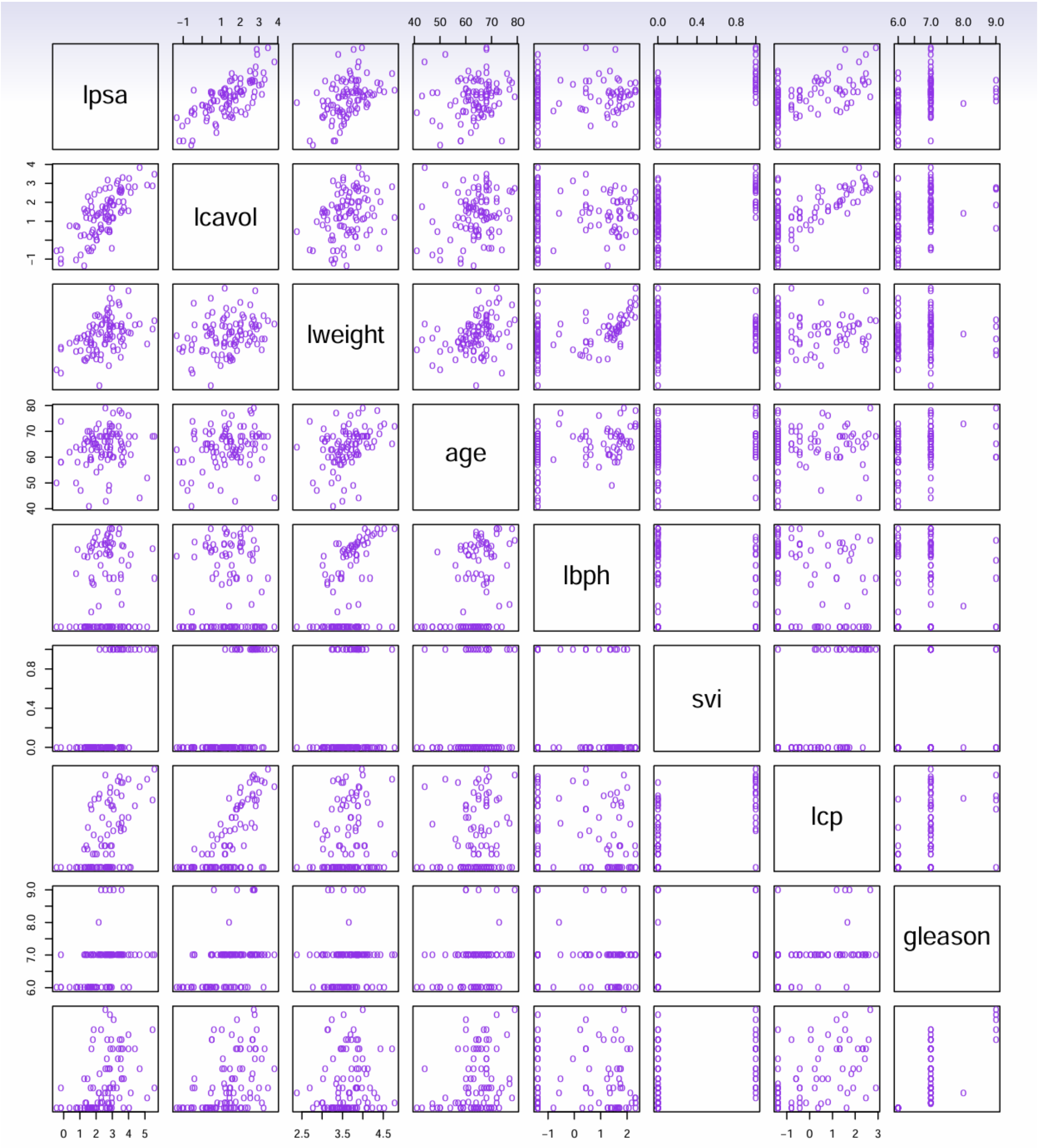
$\longrightarrow$ Let us give some examples

Prostate cancer

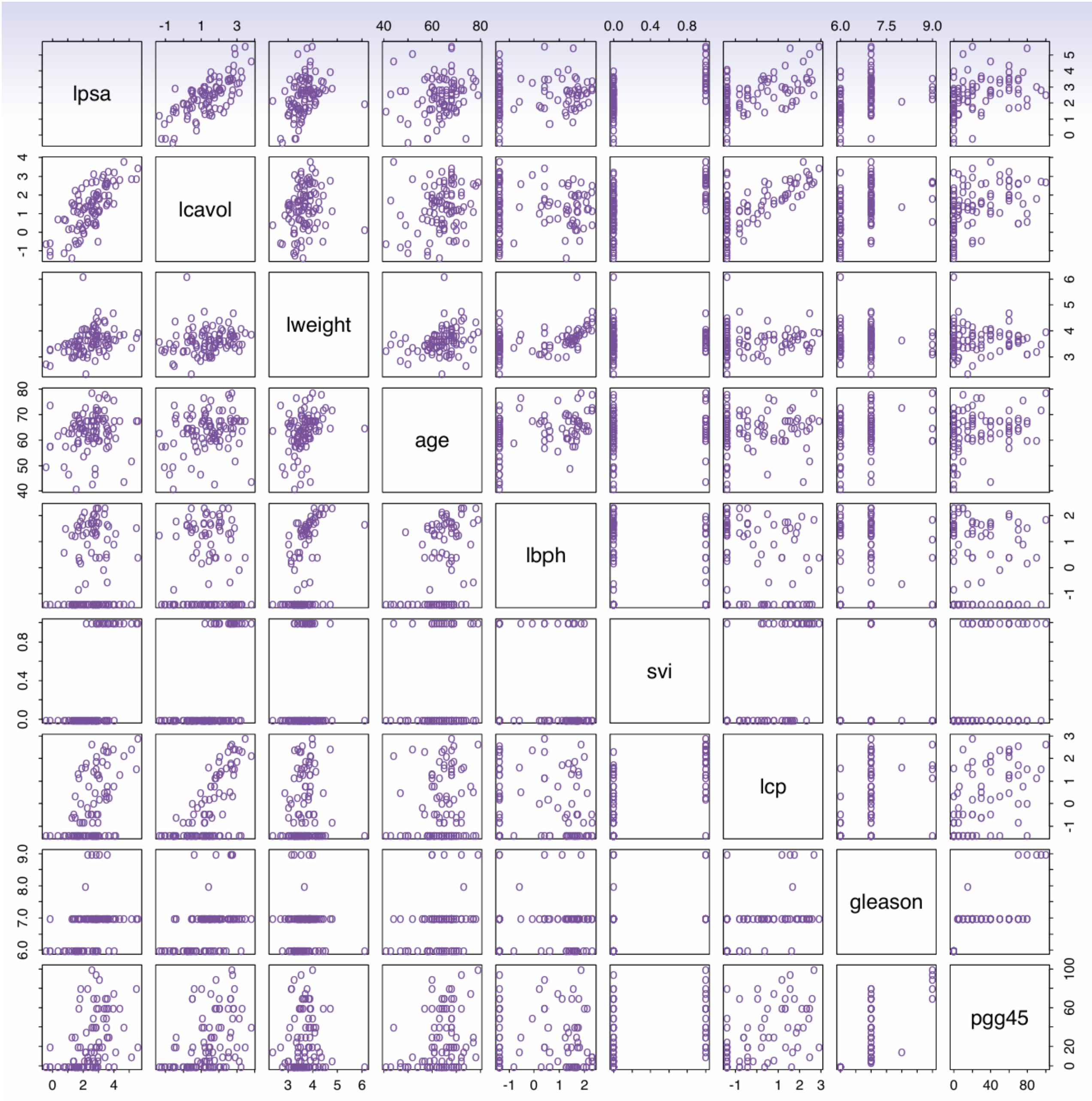


Goal: Predict PSA (enzyma produced by cancer cells) from other indexes.

Prostate cancer

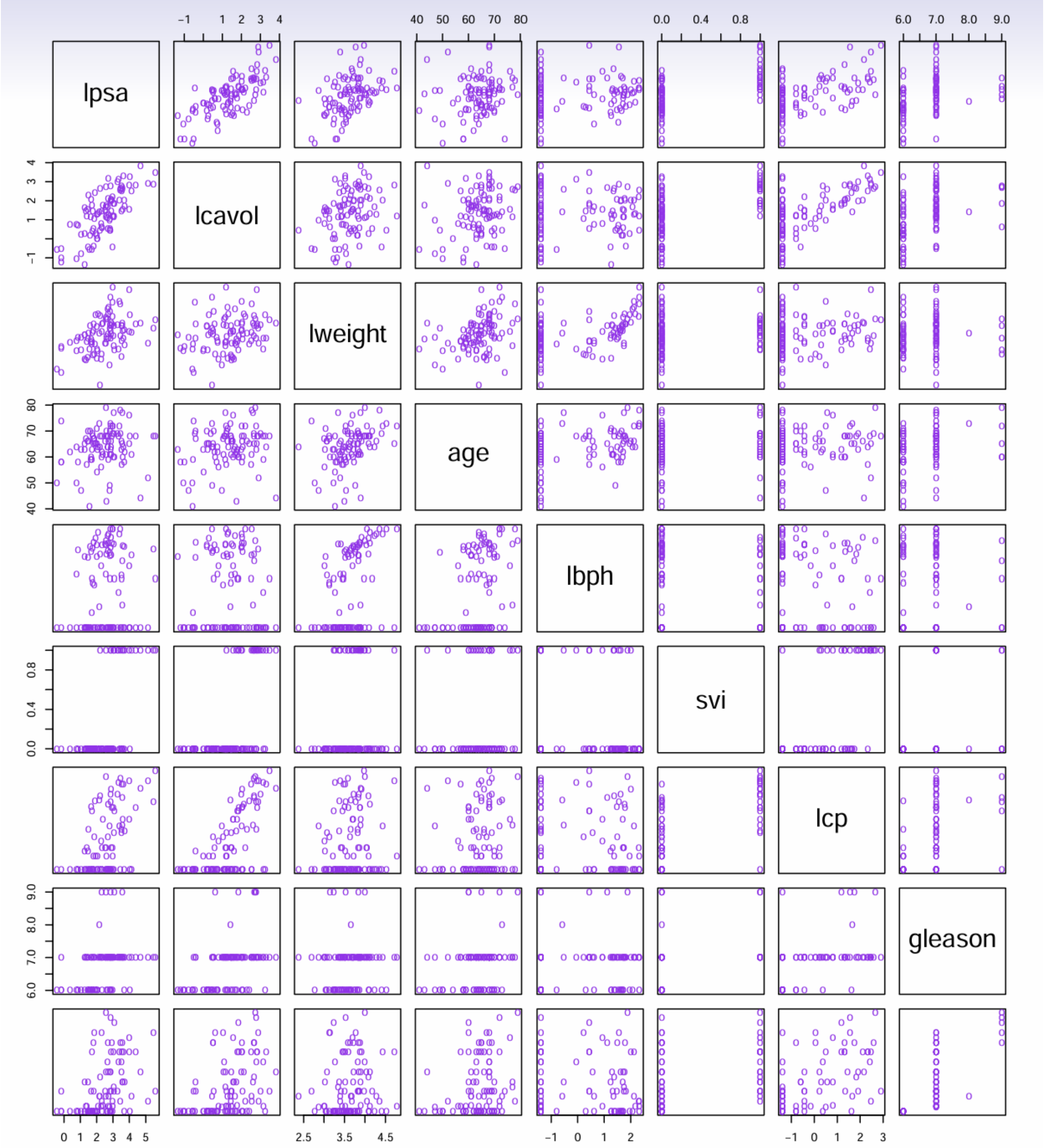


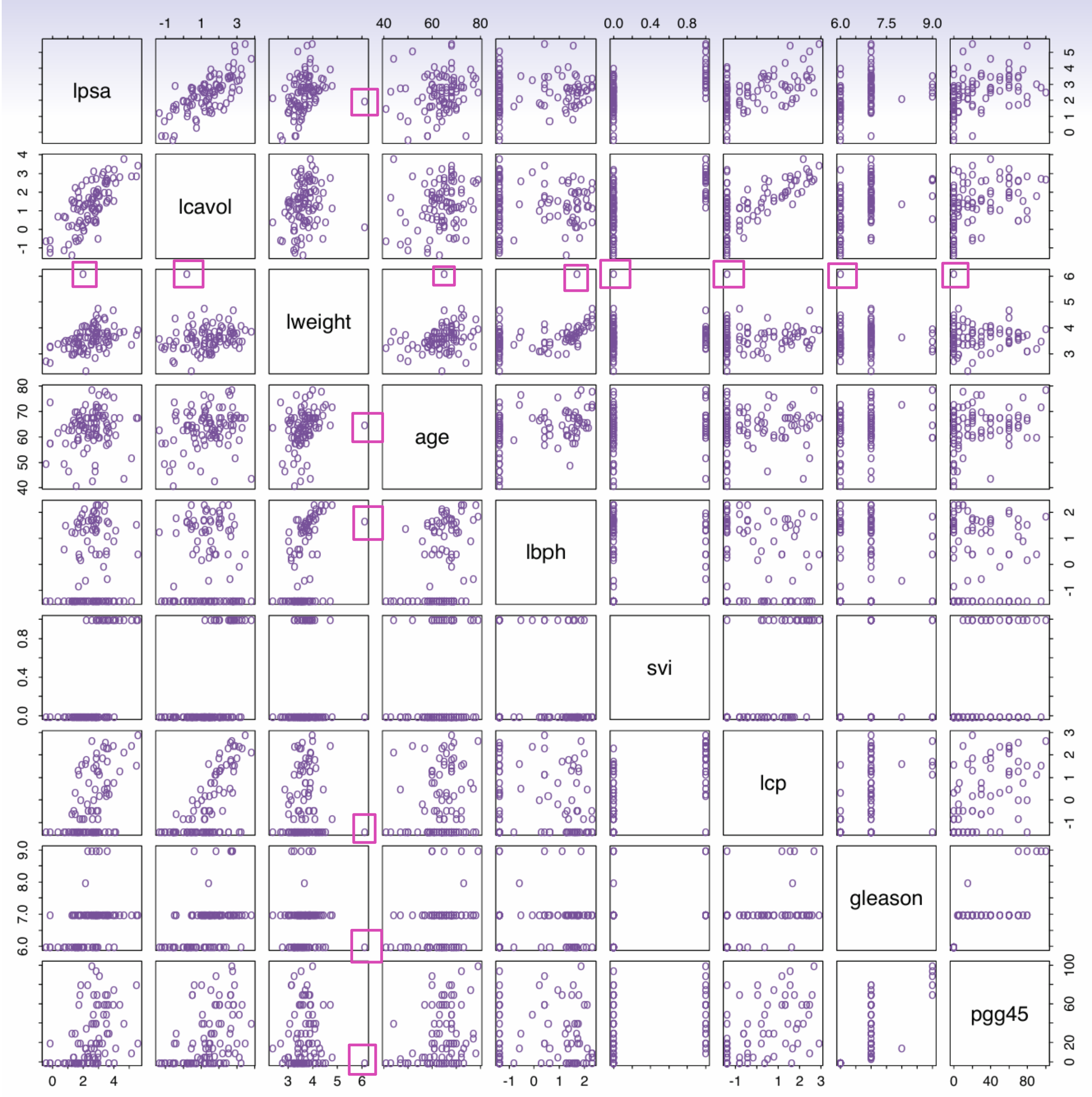
Initial data



Prostate cancer

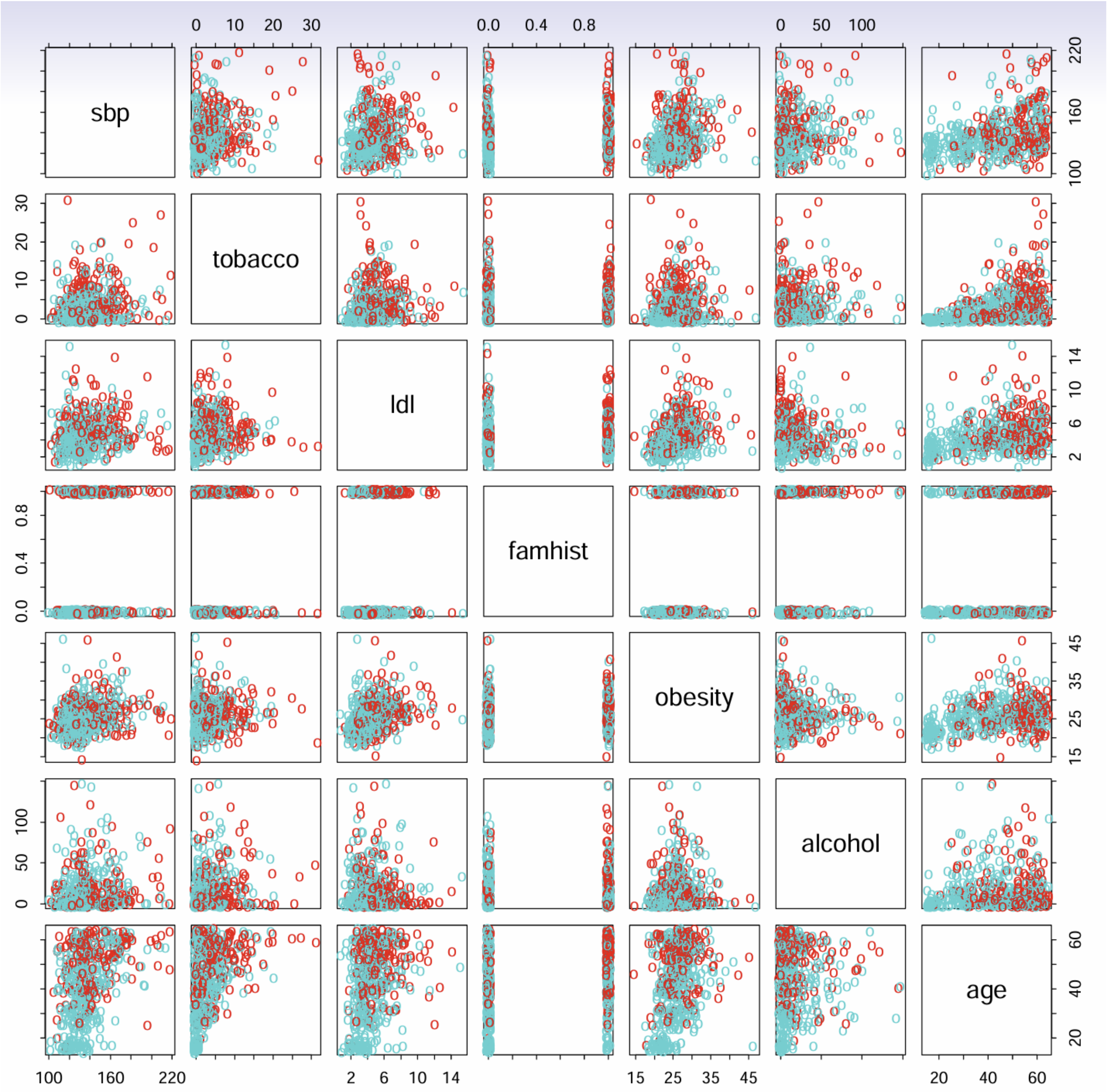
Initial data





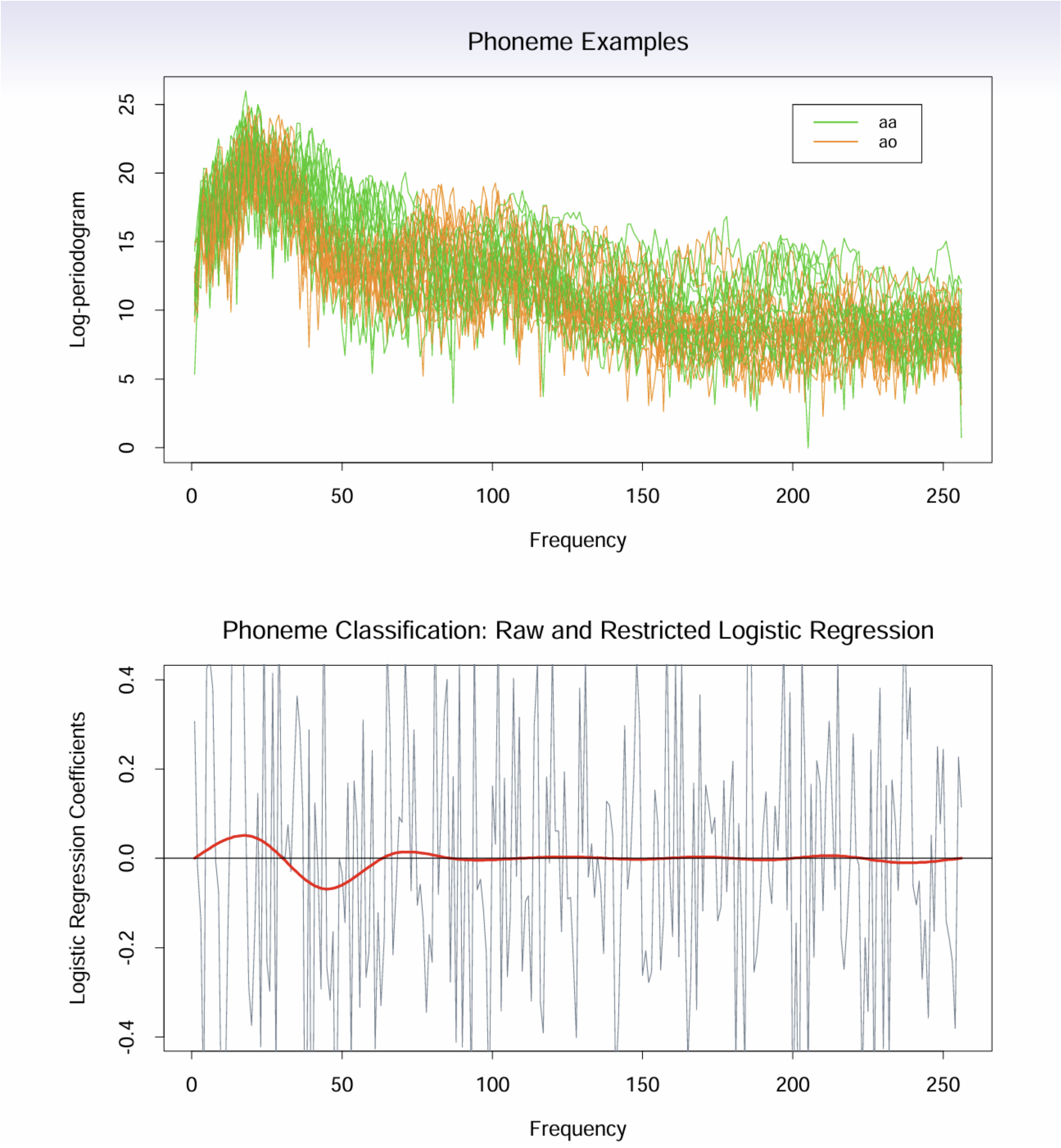
OUTLIERS

Heart strike



Goal: Predict risk for heart attack

Vowel pronunciation frequency



Goal: Recognize vowel from oral pronunciation

Spam & ham

	george	you	hp	free	!	edu	remove
spam	0.00	2.26	0.02	0.52	0.51	0.01	0.28
email	1.27	1.27	0.90	0.07	0.11	0.29	0.01

- Data from 4601 emails sent to an individual (named George, at HP labs, before 2000). Each is labeled as spam or email.
- Input features: relative frequencies of 57 of the most commonly occurring words and punctuation marks in these email messages.

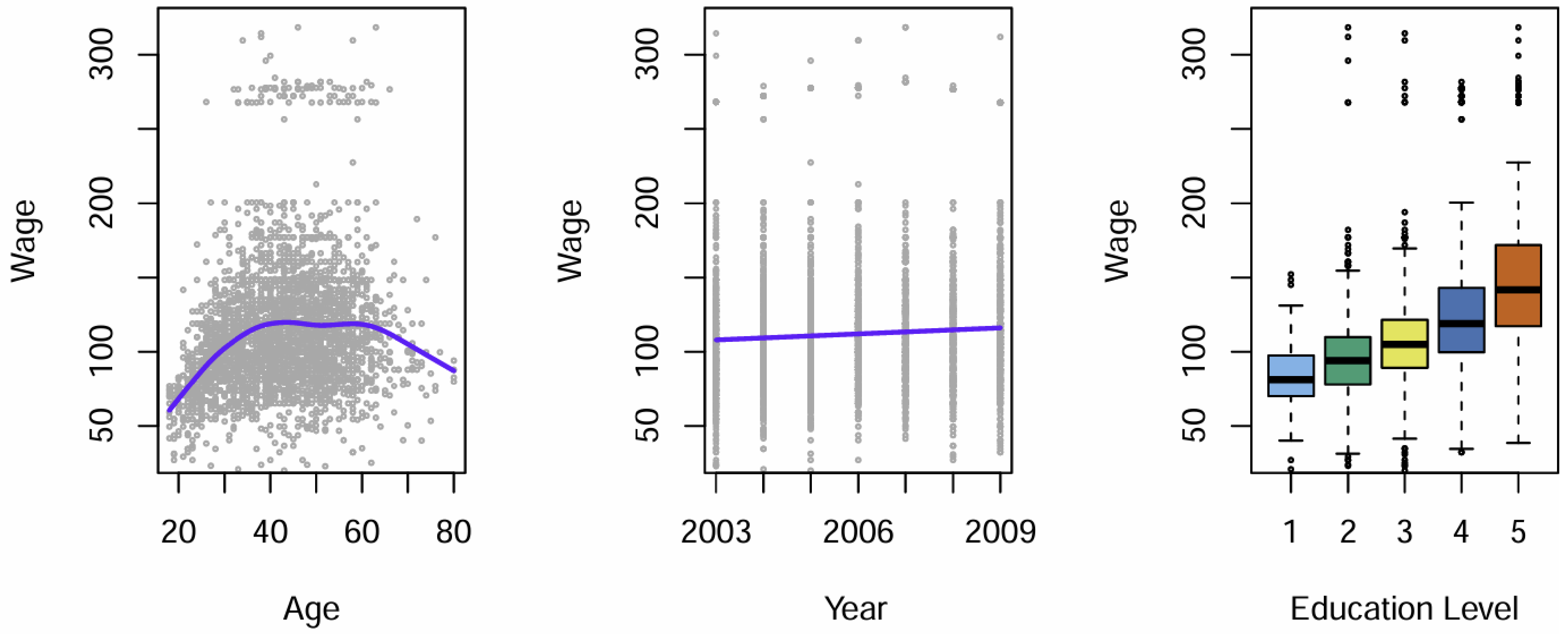
Spam & ham

	george	you	hp	free	!	edu	remove
spam	0.00	2.26	0.02	0.52	0.51	0.01	0.28
email	1.27	1.27	0.90	0.07	0.11	0.29	0.01

- Data from 4601 emails sent to an individual (named George, at HP labs, before 2000). Each is labeled as spam or email.
- Input features: relative frequencies of 57 of the most commonly occurring words and punctuation marks in these email messages.

Goal: build a customized spam lter.

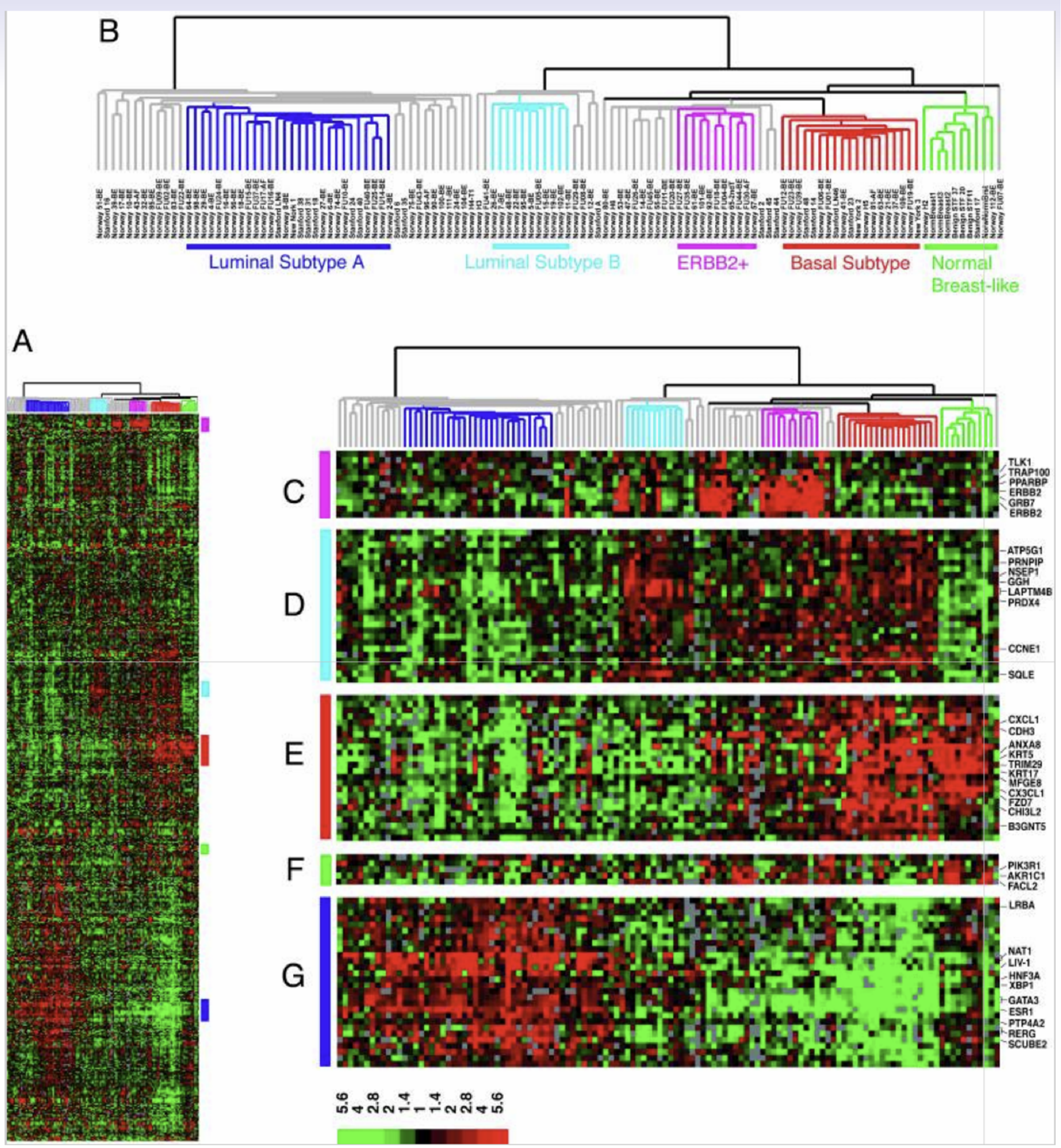
Wages



Income survey data for males from the central Atlantic region of the USA in 2009.

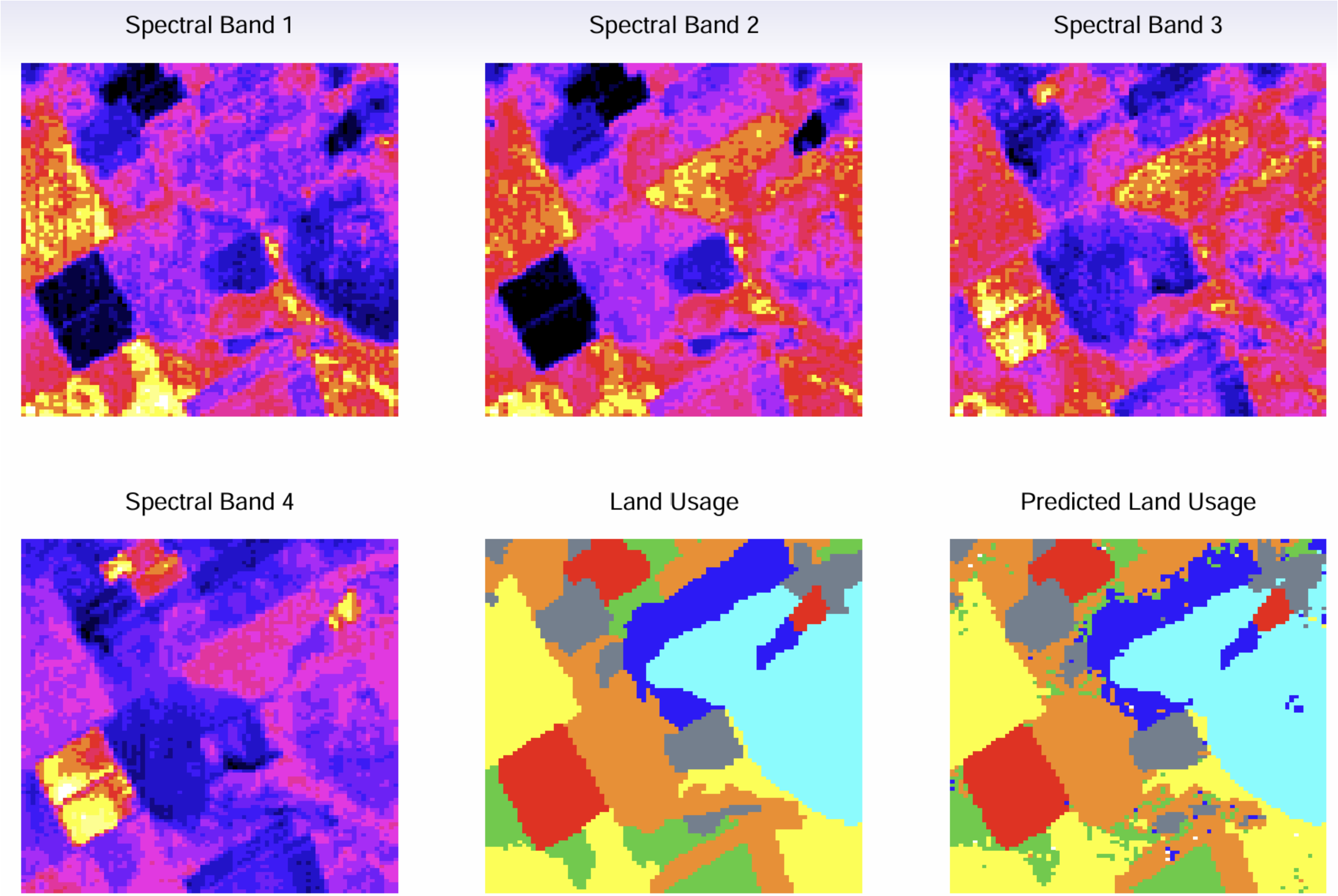
Goal: Predict wages from other quatities.

Genes identification



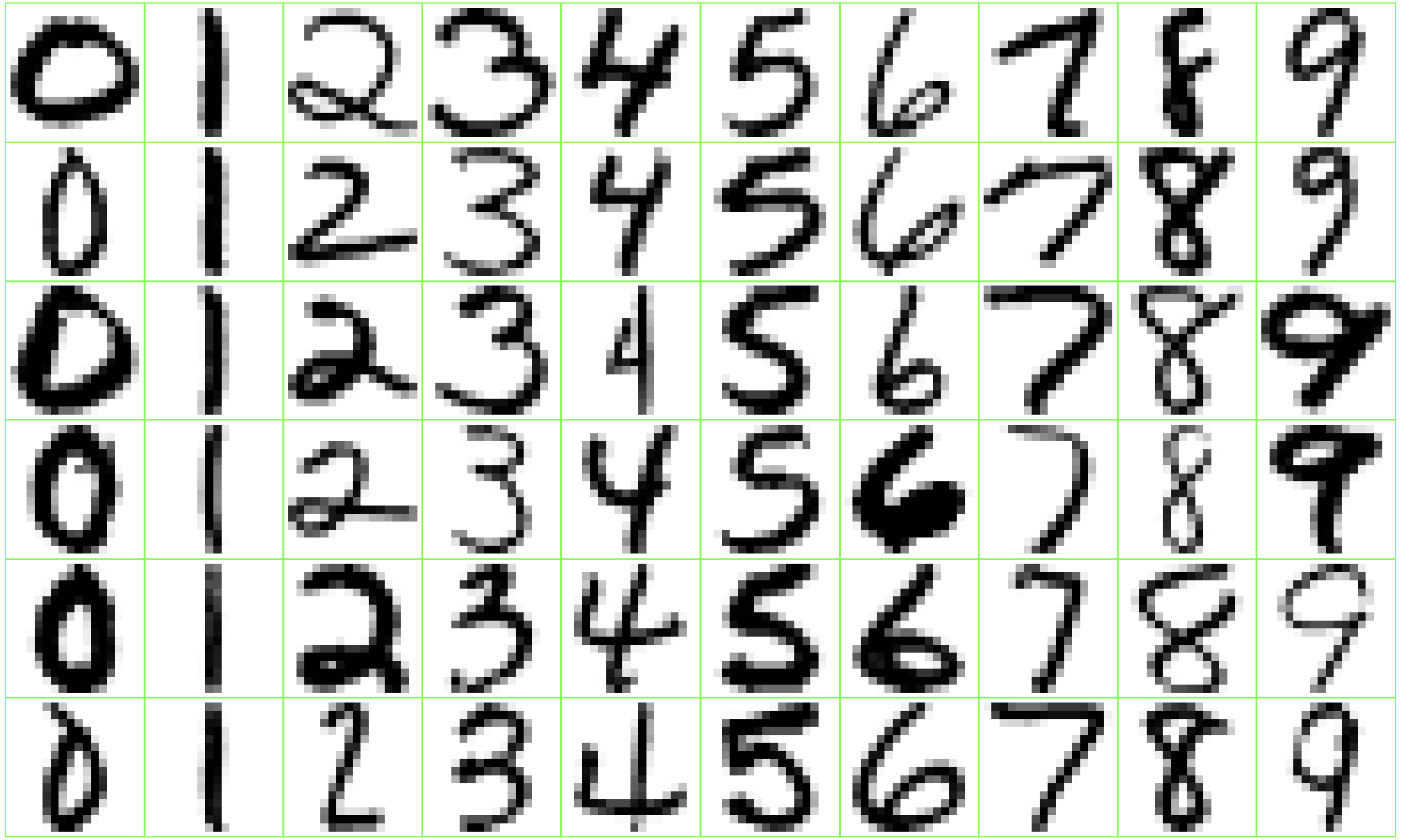
Goal: Identify guilty genes.

Fields identification



Goal: Identify fields usage.

MNIST data base

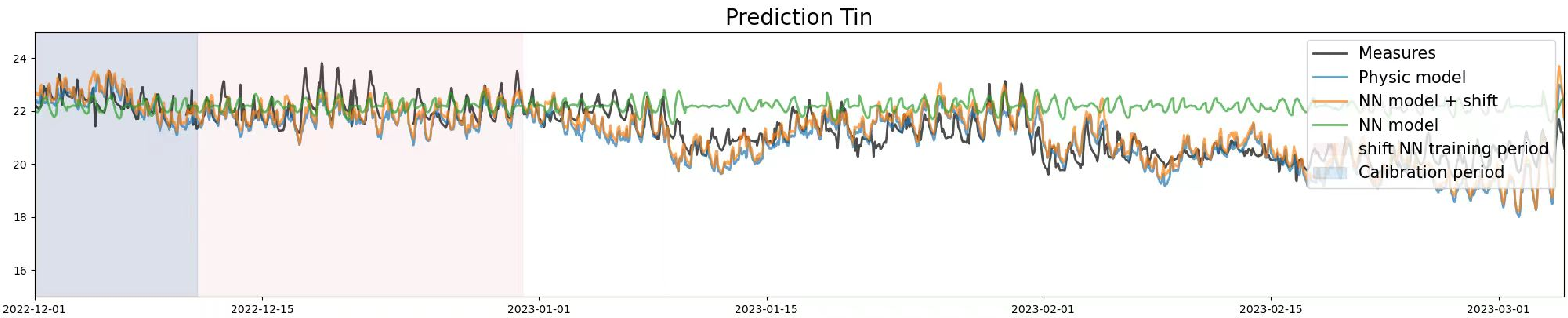


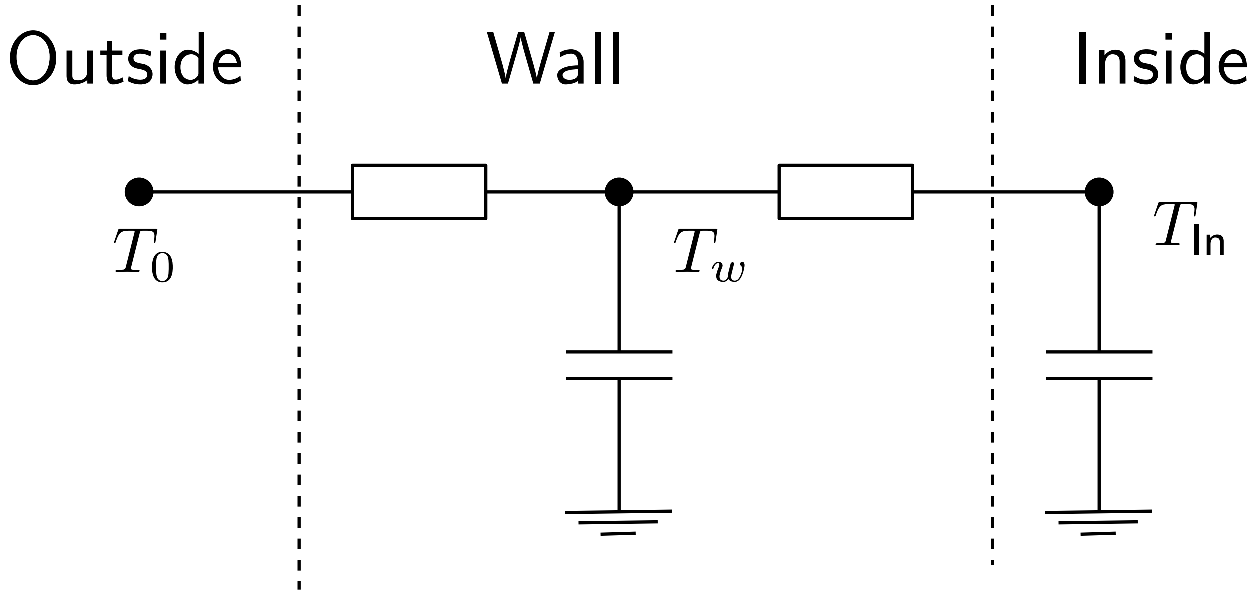
Goal: Identify figures.

Netflix challenge

Goal: predict user's score.

Predict temperature

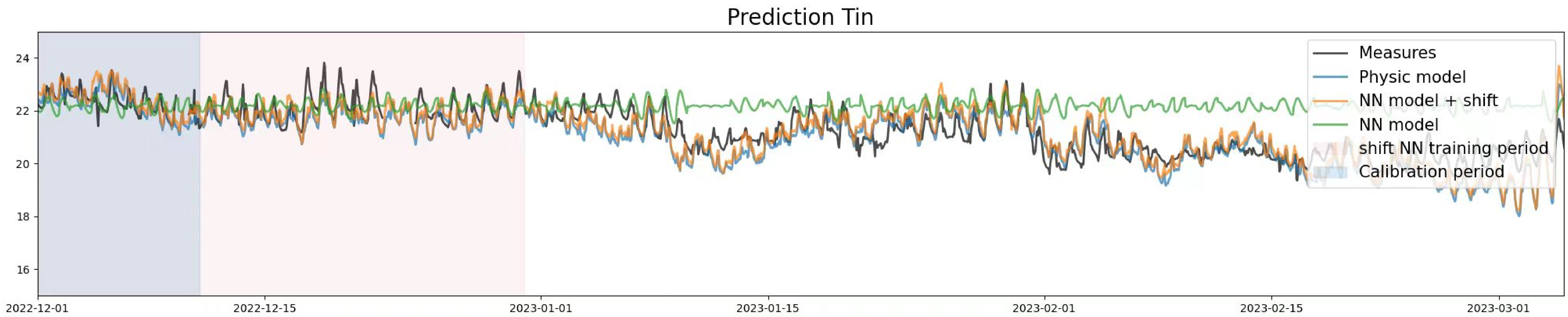


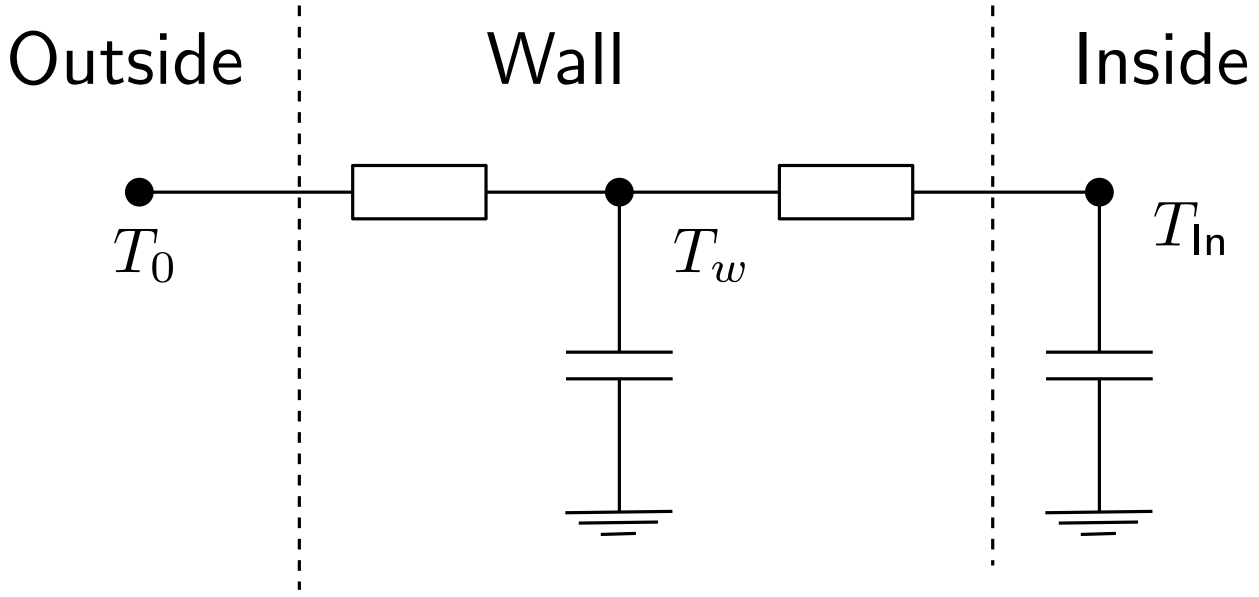


Physical model of Appartment

Goal: predict appartment indoor temperature.

Predict temperature





Physical model of Appartment

VS.

Goal: predict appartment indoor temperature.

Main concepts to classify a given task

Classification/Regression

Main concepts to classify a given task

Classification/Regression

Provide a continuous output.

Main concepts to classify a given task

Classification/Regression

Provide a continuous output.

Provide a discrete output (typically a class).

Main concepts to classify a given task

Classification/Regression

Provide a continuous output.

Provide a discrete output (typically a class).

Supervised/Unsupervised

Main concepts to classify a given task

Classification/Regression

Provide a continuous output.

Provide a discrete output (typically a class).

Supervised/Unsupervised

Tackle the problem from scratch like a baby coming to the world.

Main concepts to classify a given task

Classification/Regression

Provide a continuous output.

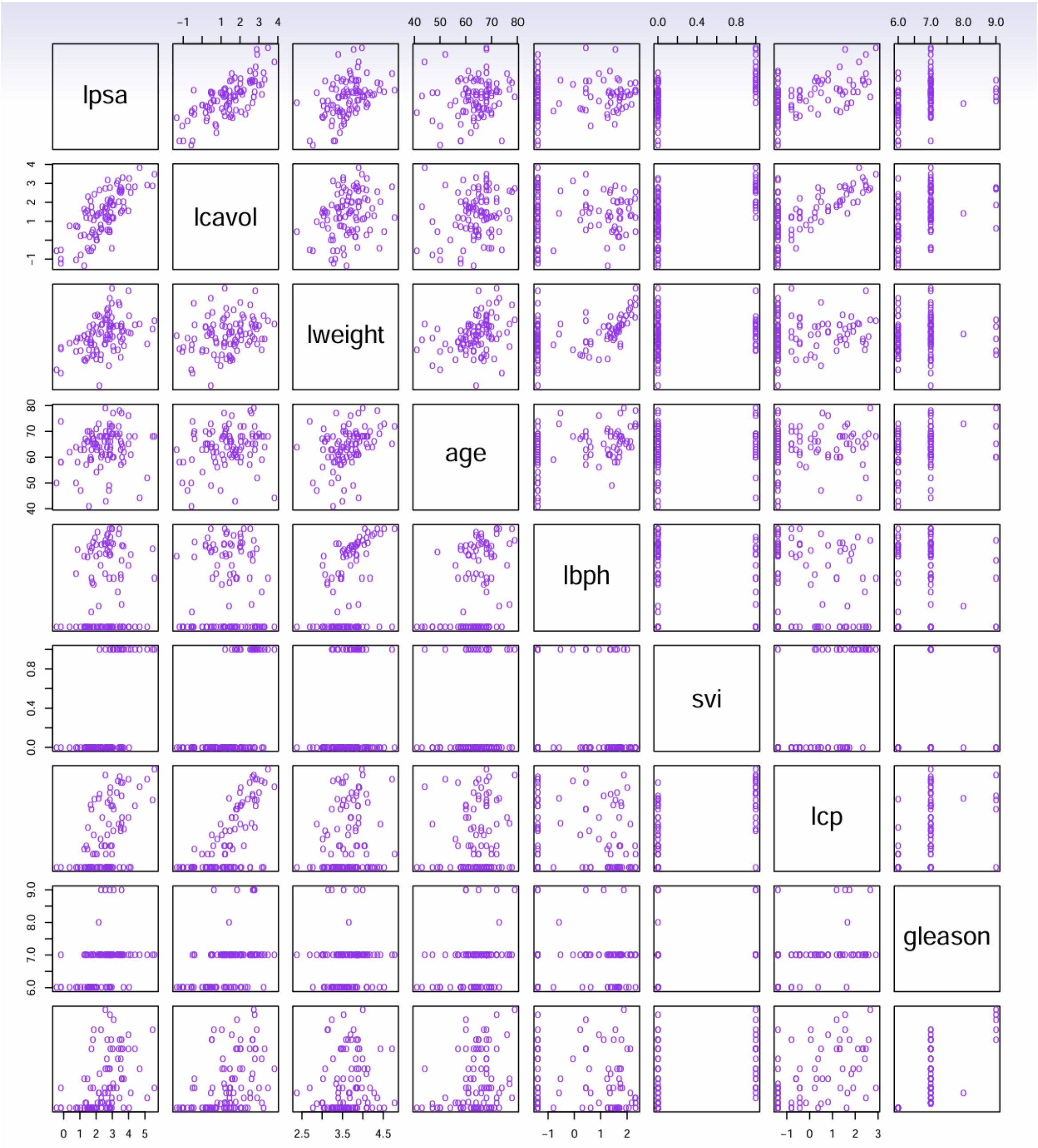
Provide a discrete output (typically a class).

Supervised/Unsupervised

Tackle the problem from scratch like a baby coming to the world.

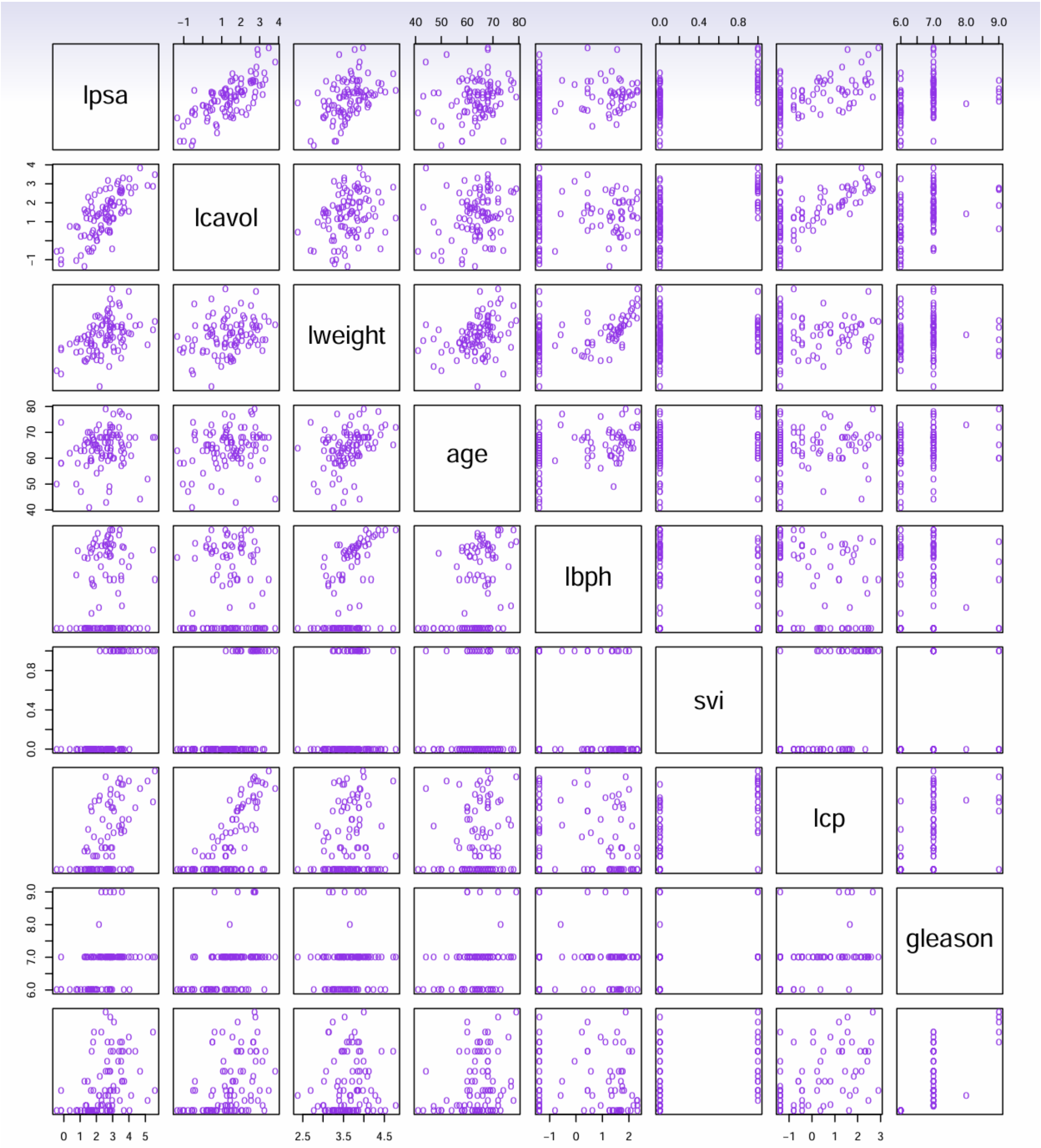
Benefits from a “training dataset” with labeled data to regress or classify “test data”.

Prostate cancer



Goal: Predict PSA (enzyma produced by cancer cells) from other indexes.

Prostate cancer

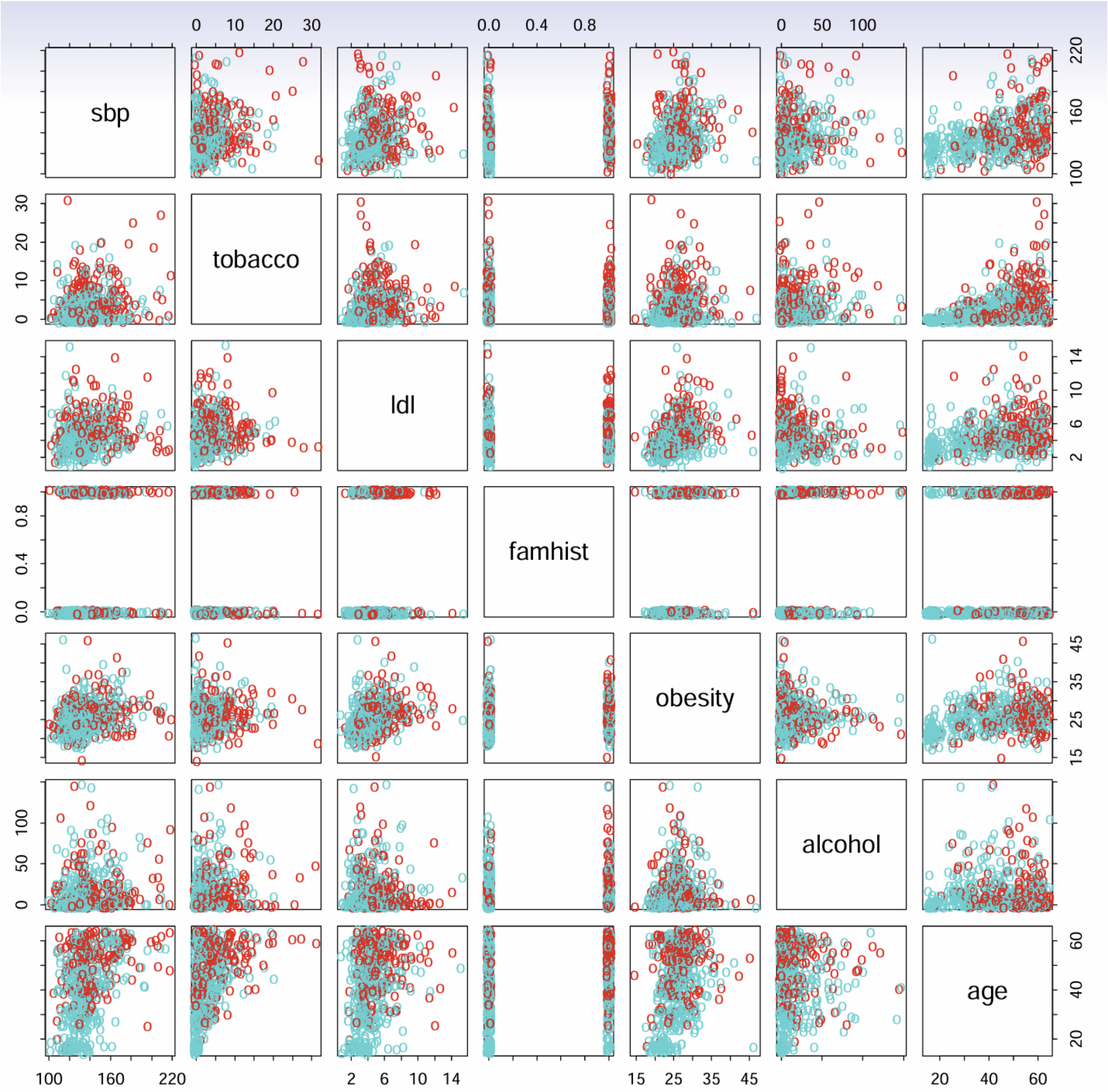


Goal: Predict PSA (enzyma produced by cancer cells) from other indexes.

REGRESSION

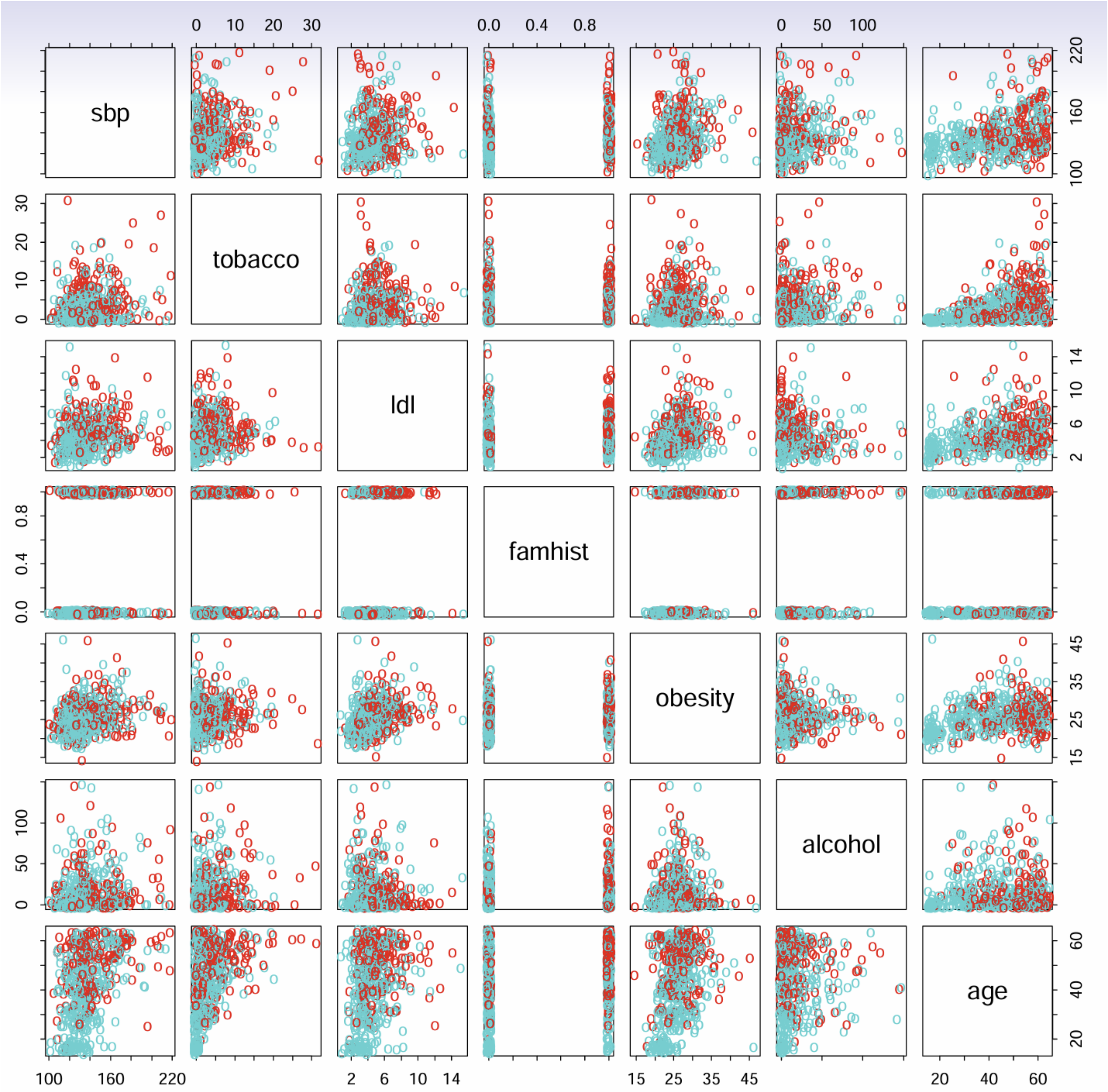
SUPERVISED

Heart strike



Goal: Predict risk for heart attack

Heart strike



Goal: Predict risk for heart attack

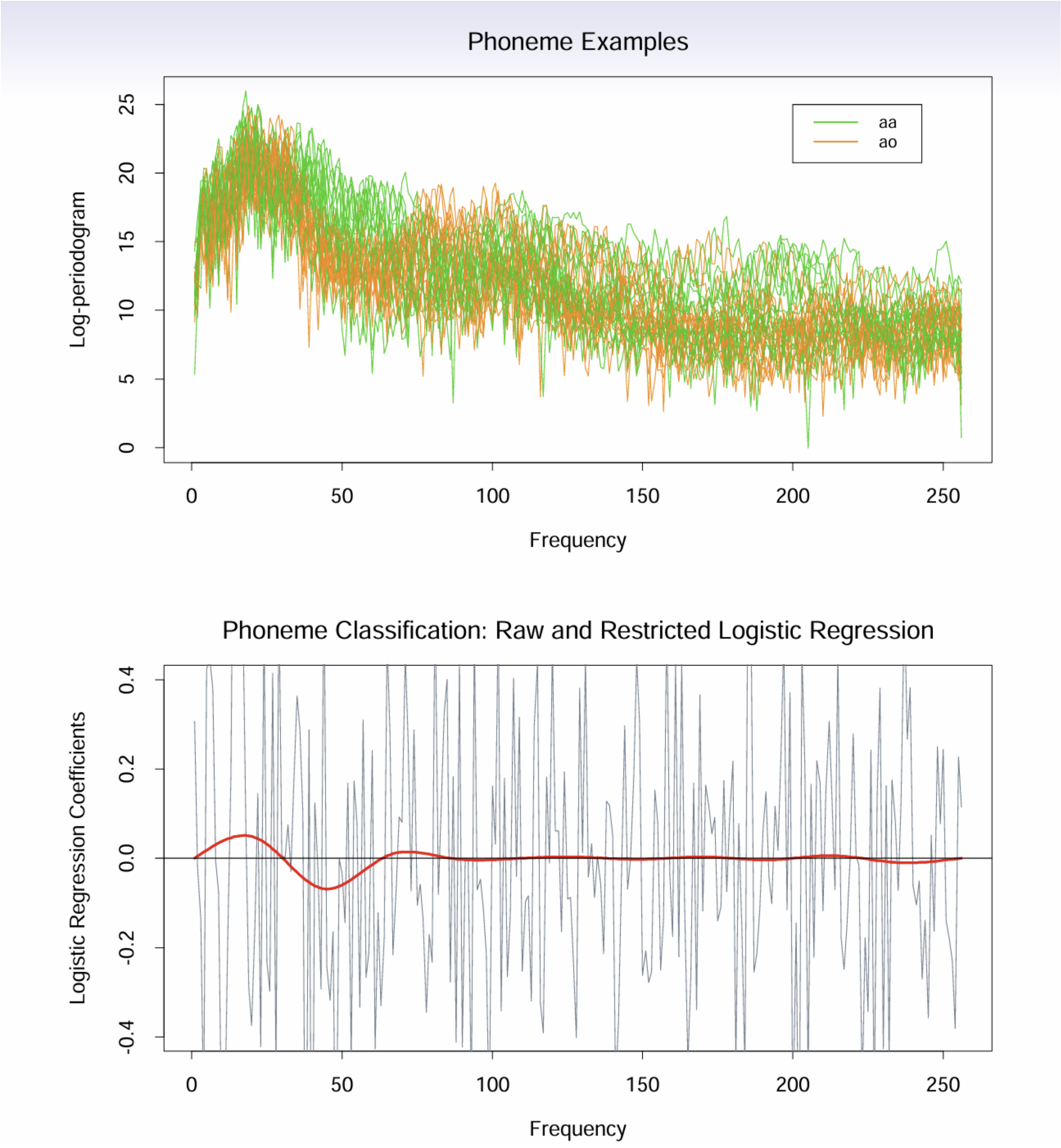
CLASSIFICATION

SUPERVISED

Vowel pronunciation frequency

Goal: Recognize vowel from oral pronunciation

Vowel pronunciation frequency

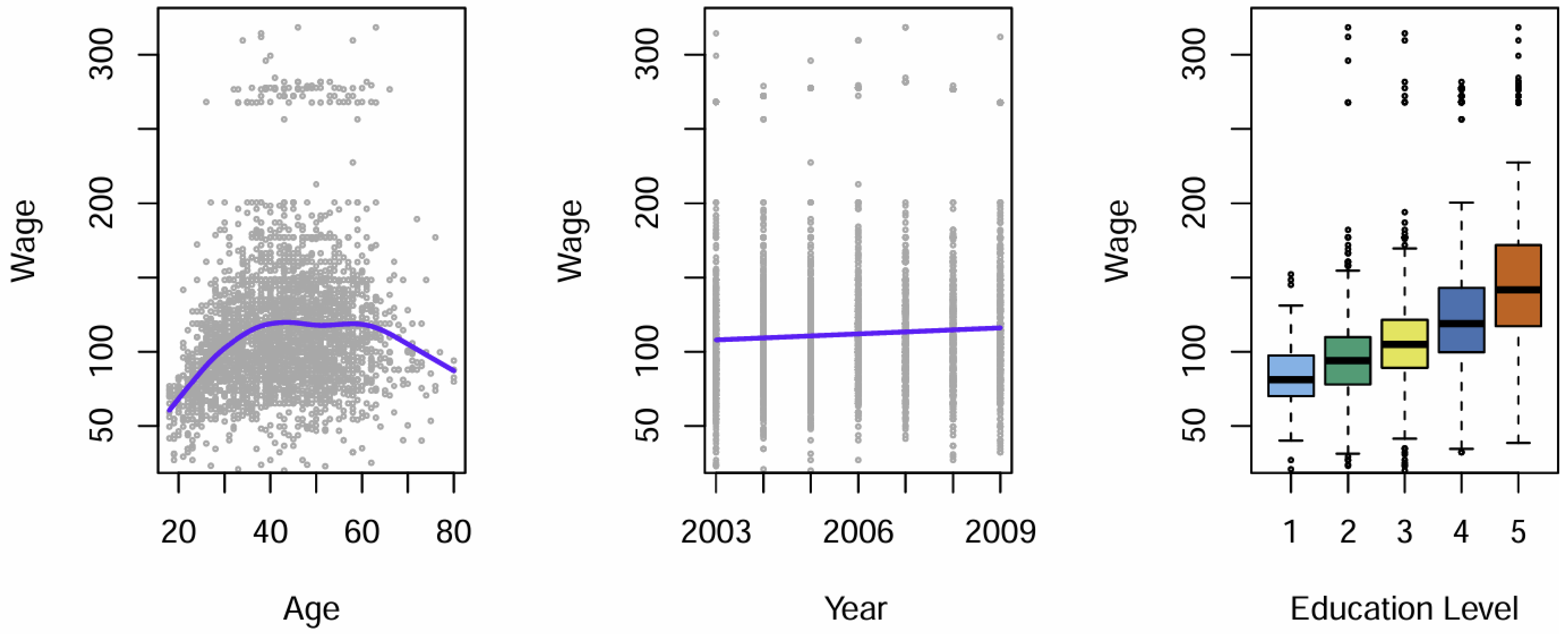


Goal: Recognize vowel from oral pronunciation

CLASSIFICATION

SUPERVISED

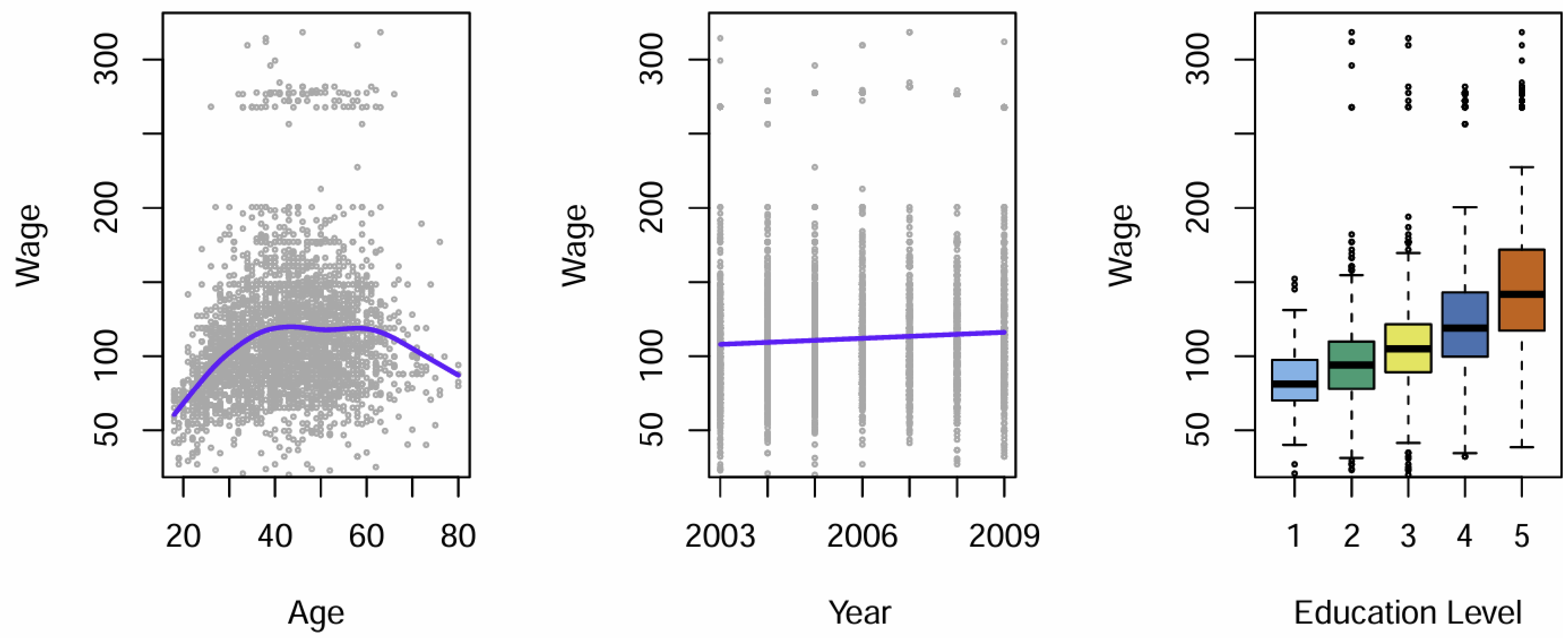
Wages



Income survey data for males from the central Atlantic region of the USA in 2009.

Goal: Predict wages from other quatities.

Wages



REGRESSION
SUPERVISED

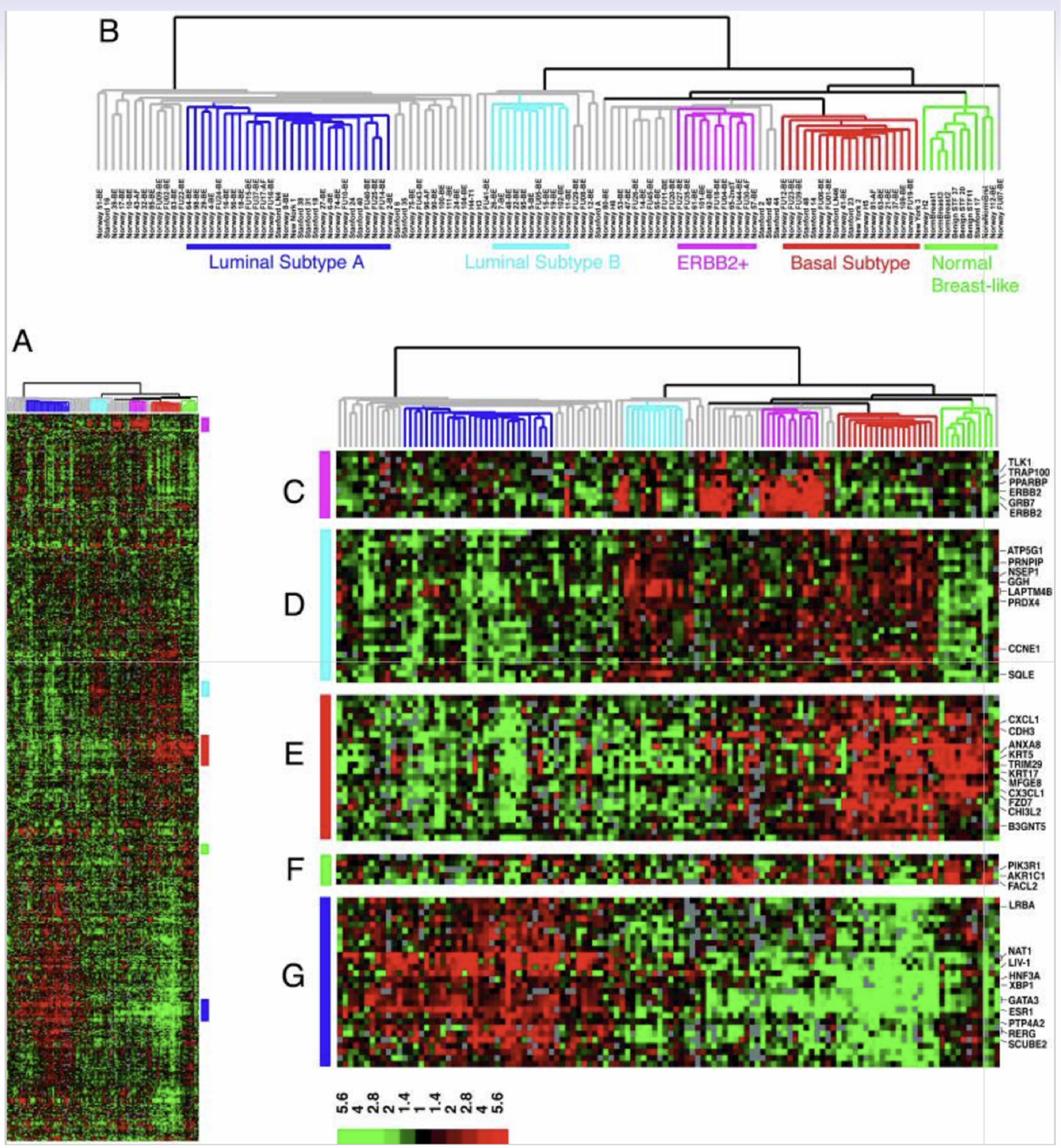
Income survey data for males from the central Atlantic region of the USA in 2009.

Goal: Predict wages from other quatities.

Genes identification

Goal: Identify guilty genes.

Genes identification

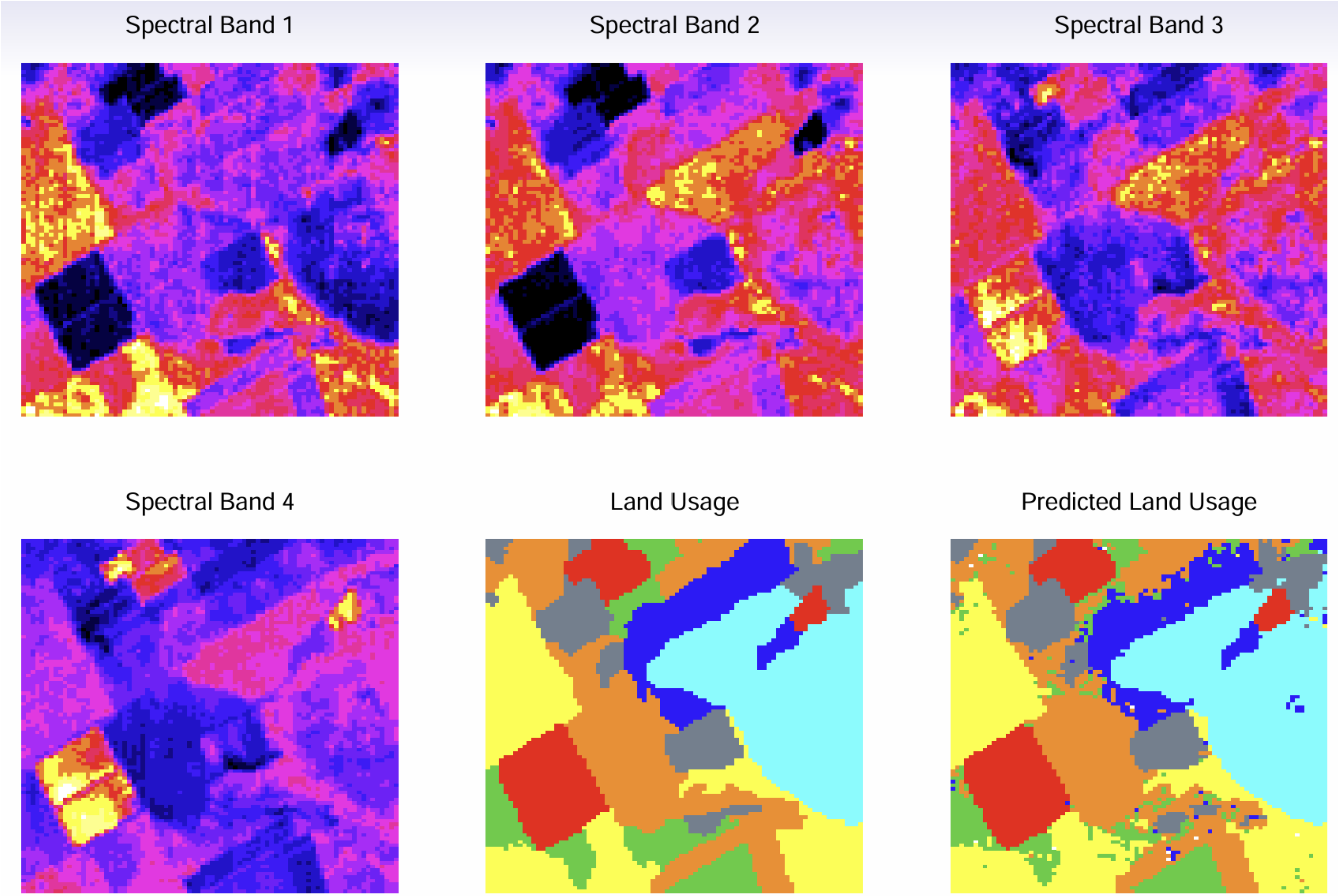


Goal: Identify guilty genes.

CLASSIFICATION

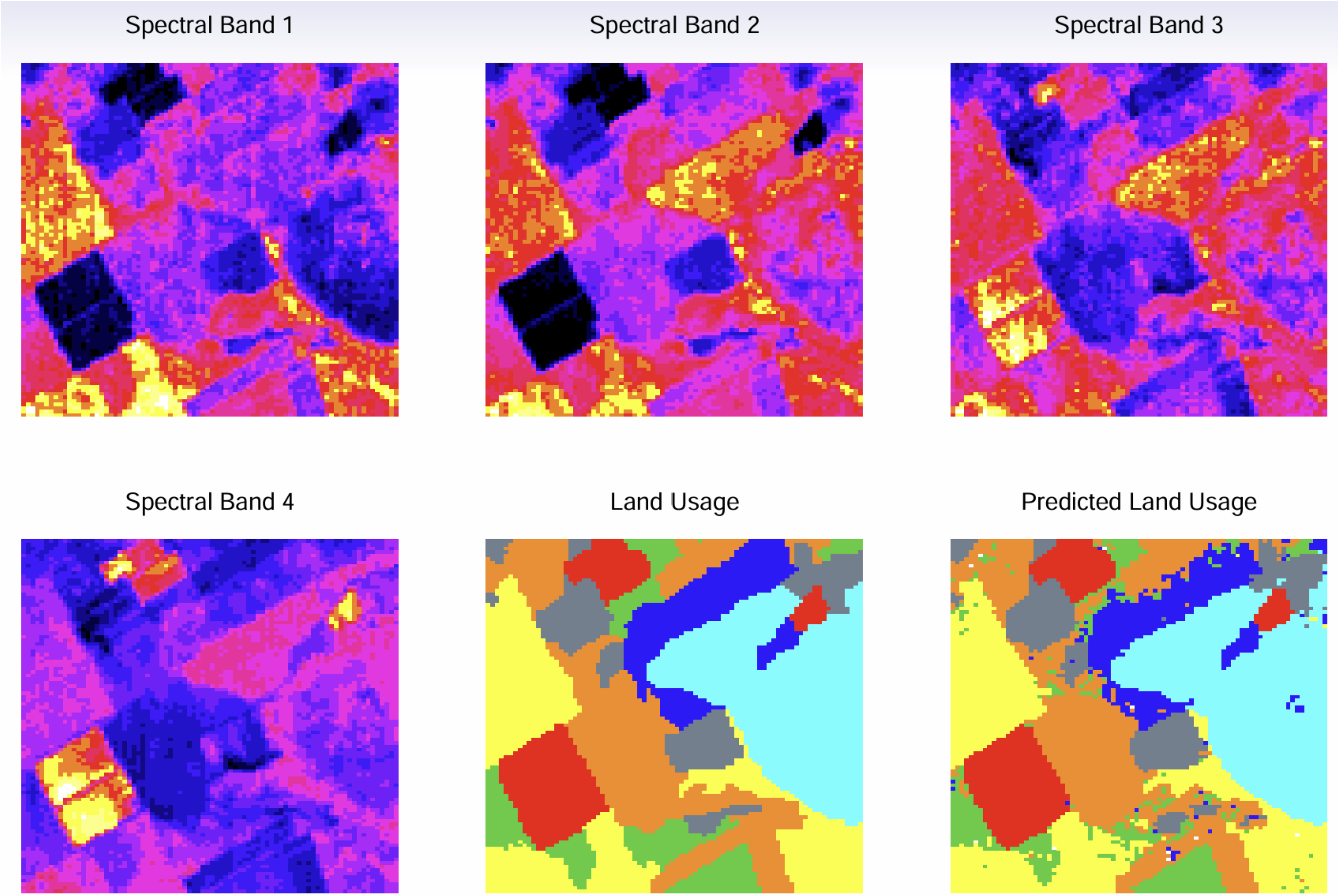
UNSUPERVISED

Fields identification



Goal: Identify fields usage.

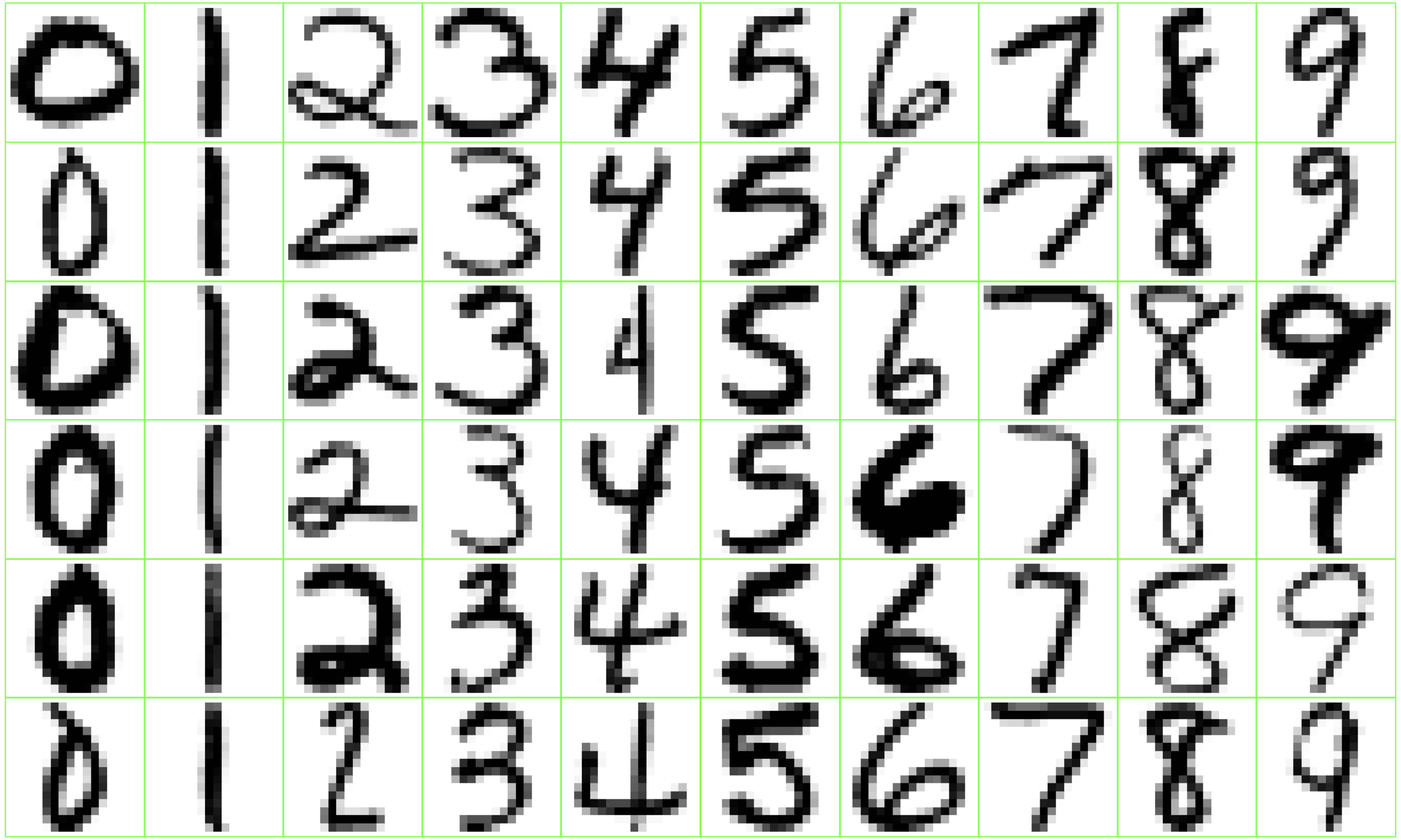
Fields identification



Goal: Identify fields usage.

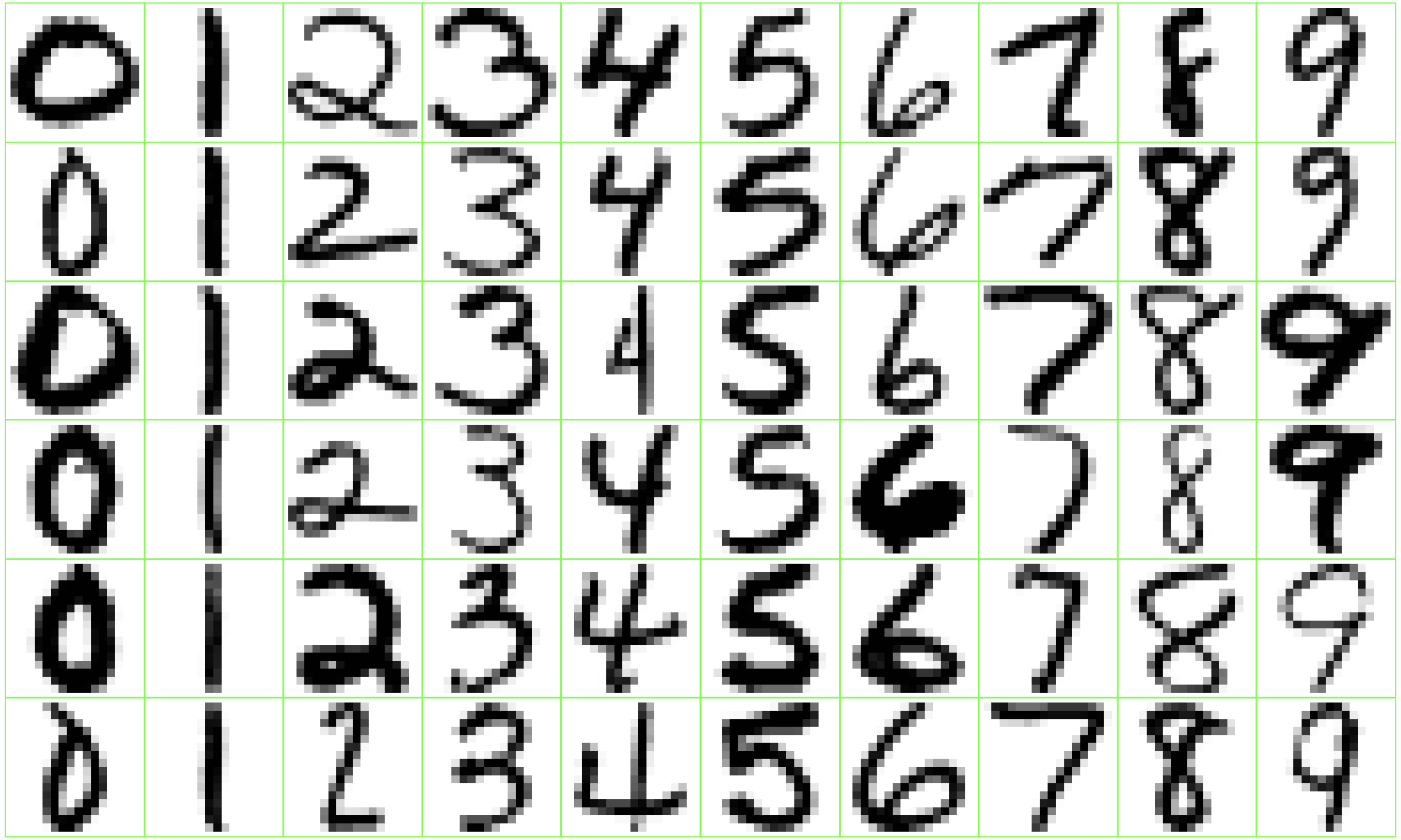
CLASSIFICATION
SUPERVISED

MNIST data base



Goal: Identify figures.

MNIST data base



Goal: Identify figures.

CLASSIFICATION

SUPERVISED

Netflix challenge

Goal: predict user's score.

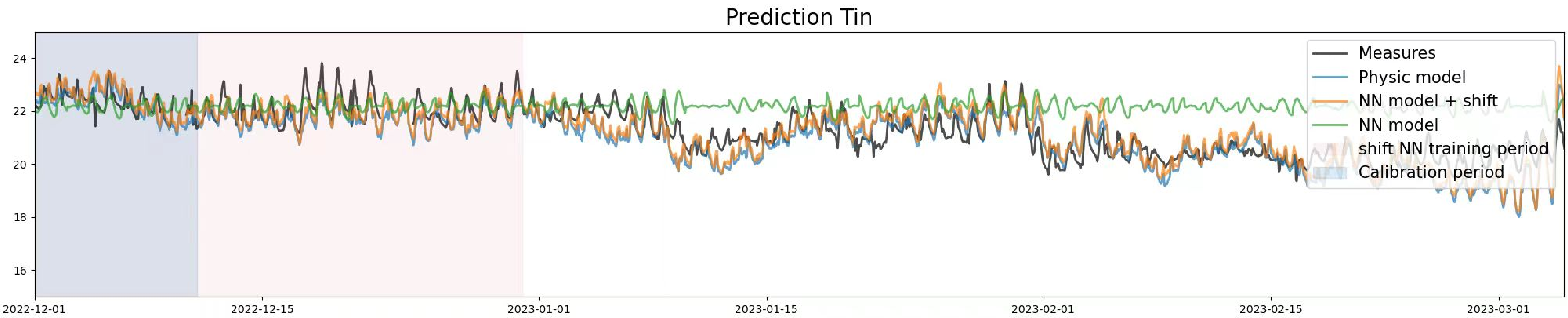
Netflix challenge

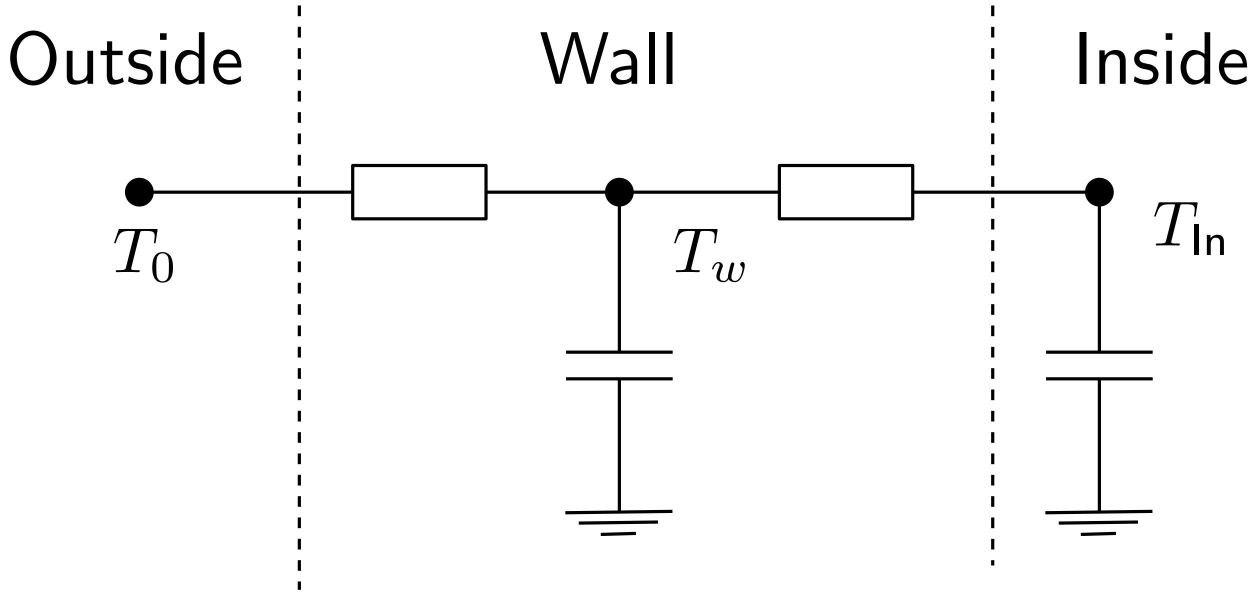
Goal: predict user's score.

REGRESSION

UNSUPERVISED

Predict temperature



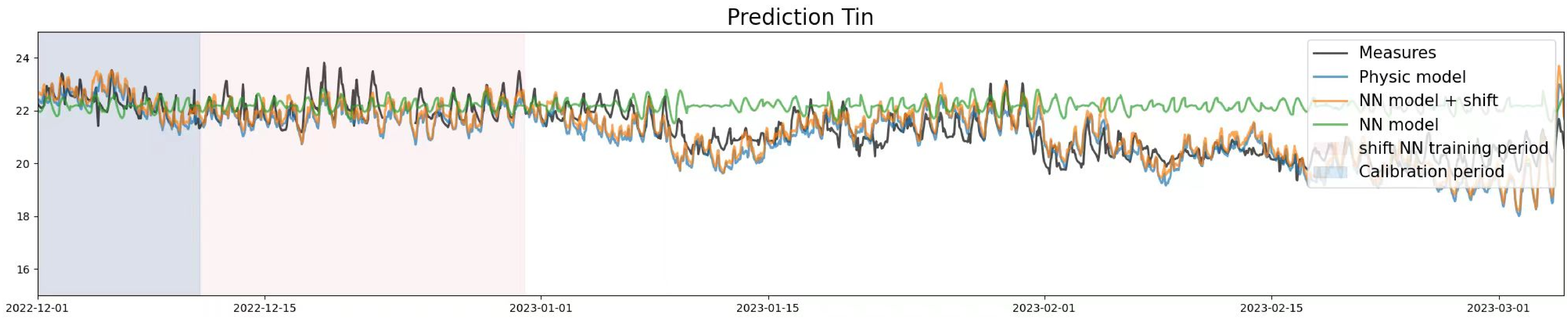


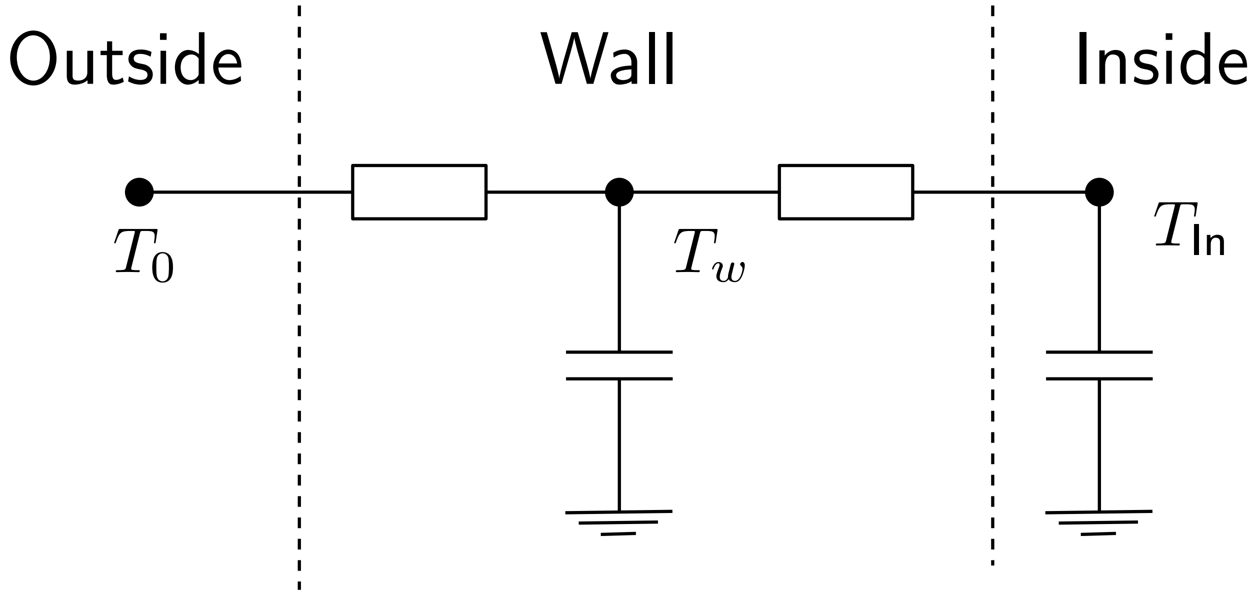
Physical model of Appartment

VS.

Goal: predict appartment indoor temperature.

Predict temperature





Physical model of Appartment

VS.

Goal: predict appartment indoor temperature.

REGRESSION

UNSUPERVISED

Statistical learning or Machine learning ?

Statistical learning or Machine learning ?

Statistical learning	Machine learning

Statistical learning or Machine learning ?

Statistical learning

Machine learning

Arose as a sub field of Artificial Intelligence.

Statistical learning or Machine learning ?

Statistical learning

Arose as a subfield of Statistics.

Machine learning

Arose as a sub field of Artificial Intelligence.

Statistical learning or Machine learning ?

Statistical learning

Arose as a subfield of Statistics.

Machine learning

Arose as a sub field of Artificial Intelligence.

Greater emphasis on large scale applications and prediction accuracy.

Statistical learning or Machine learning ?

Statistical learning	Machine learning
Arose as a subfield of Statistics.	Arose as a sub field of Artificial Intelligence.
Emphasizes models and their interpretability, precision and uncertainty.	Greater emphasis on large scale applications and prediction accuracy.

Statistical learning or Machine learning ?

Statistical learning

Arose as a subfield of Statistics.

Emphasizes models and their interpretability, precision and uncertainty.

Machine learning

Arose as a sub field of Artificial Intelligence.

Greater emphasis on large scale applications and prediction accuracy.

Considers state of the art algorithms, only limited by the technniques and equipment

Statistical learning or Machine learning ?

Statistical learning

Arose as a subfield of Statistics.

Emphasizes models and their interpretability, precision and uncertainty.

Concerns simple methods that allow theoretical guarantees

Machine learning

Arose as a sub field of Artificial Intelligence.

Greater emphasis on large scale applications and prediction accuracy.

Considers state of the art algorithms, only limited by the technniques and equipment

Statistical learning or Machine learning ?

Statistical learning

Arose as a subfield of Statistics.

Emphasizes models and their interpretability, precision and uncertainty.

Concerns simple methods that allow theoretical guarantees

Machine learning

Arose as a sub field of Artificial Intelligence.

Greater emphasis on large scale applications and prediction accuracy.

Considers state of the art algorithms, only limited by the technniques and equipment

In our course, we will provide:
- coding basis to allow you to do your own scripts
- theoretical statistical basis to let you understand why those algorithms work.

Statistical learning or Machine learning ?

Statistical learning

Arose as a subfield of Statistics.

Emphasizes models and their interpretability, precision and uncertainty.

Concerns simple methods that allow theoretical guarantees

Machine learning

Arose as a sub field of Artificial Intelligence.

Greater emphasis on large scale applications and prediction accuracy.

Considers state of the art algorithms, only limited by the technniques and equipment

In our course, we will provide:

- coding basis to allow you to do your own scripts
- theoretical statistical basis to let you understand why those algorithms work.

Data science engineer

Statistical learning or Machine learning ?

Statistical learning

Arose as a subfield of Statistics.

Emphasizes models and their interpretability, precision and uncertainty.

Concerns simple methods that allow theoretical guarantees

Machine learning

Arose as a sub field of Artificial Intelligence.

Greater emphasis on large scale applications and prediction accuracy.

Considers state of the art algorithms, only limited by the technniques and equipment

In our course, we will provide:

- coding basis to allow you to do your own scripts → Data science engineer
- theoretical statistical basis to let you understand why those algorithms work. → Data science researcher

Why “statistical” ?

Why “statistical” ?

> *“Almost all of machine learning can be viewed in probabilistic terms, making probabilistic thinking fundamental. It is, of course, not the only view. But it is through this view that we can connect what we do in machine learning to every other computational science, whether that be in stochastic optimisation, control theory, operations research, econometrics, information theory, statistical physics or bio-statistics. For this reason alone, mastery of probabilistic thinking is essential.”*

Shakir Mohamed, DeepMind

Why “statistical” ?

“Almost all of machine learning can be viewed in probabilistic terms, making probabilistic thinking fundamental. It is, of course, not the only view. But it is through this view that we can connect what we do in machine learning to every other computational science, whether that be in stochastic optimisation, control theory, operations research, econometrics, information theory, statistical physics or bio-statistics. For this reason alone, mastery of probabilistic thinking is essential.”

Shakir Mohamed, DeepMind

...Because there exists some randomness in the data!

Geometric vs. Probabilistic view on data

Geometric vs. Probabilistic view on data

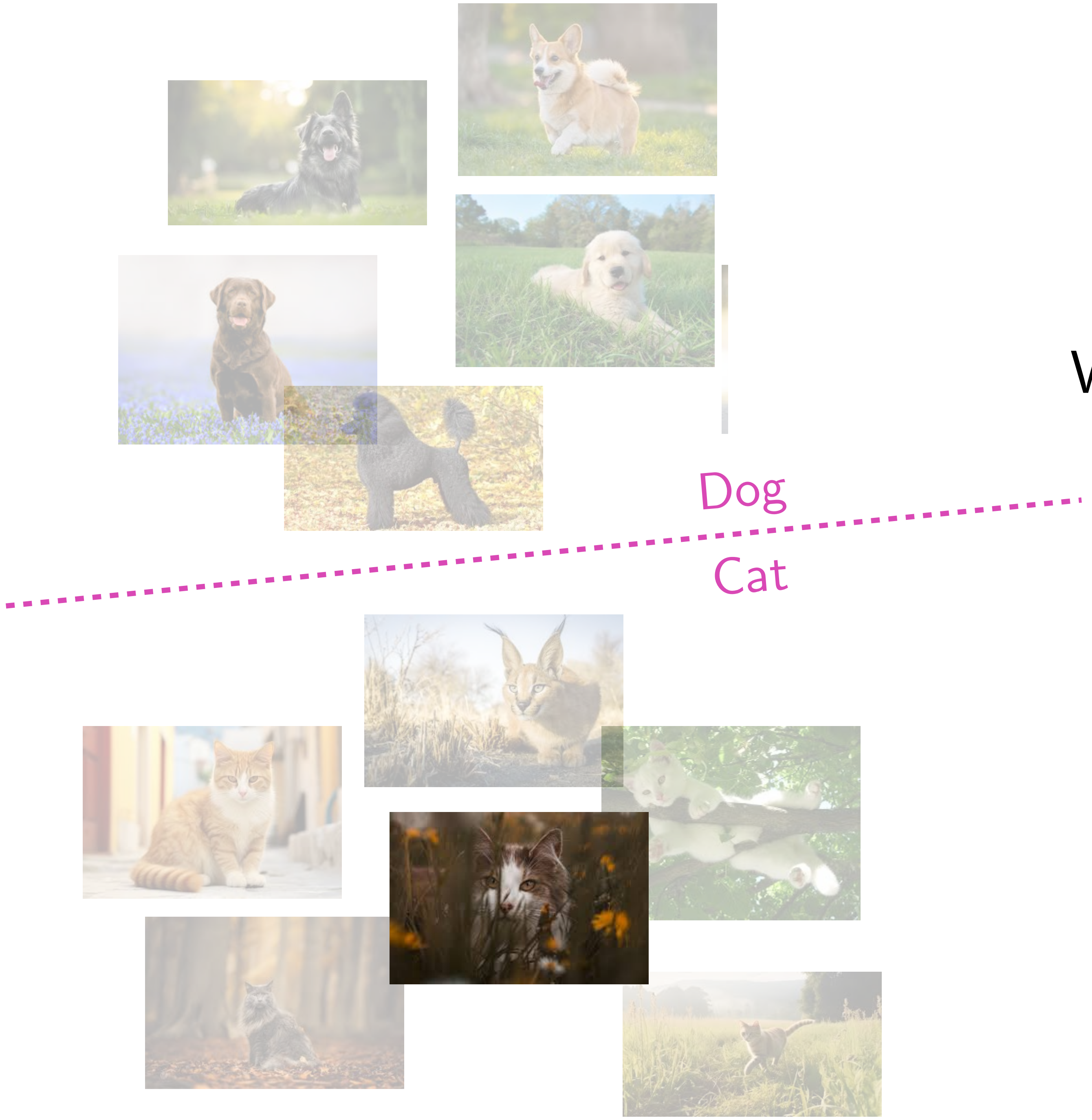
Geometric vs. Probabilistic view on data

Geometric vs. Probabilistic view on data

Geometric vs. Probabilistic view on data

Who is dog/cat ?

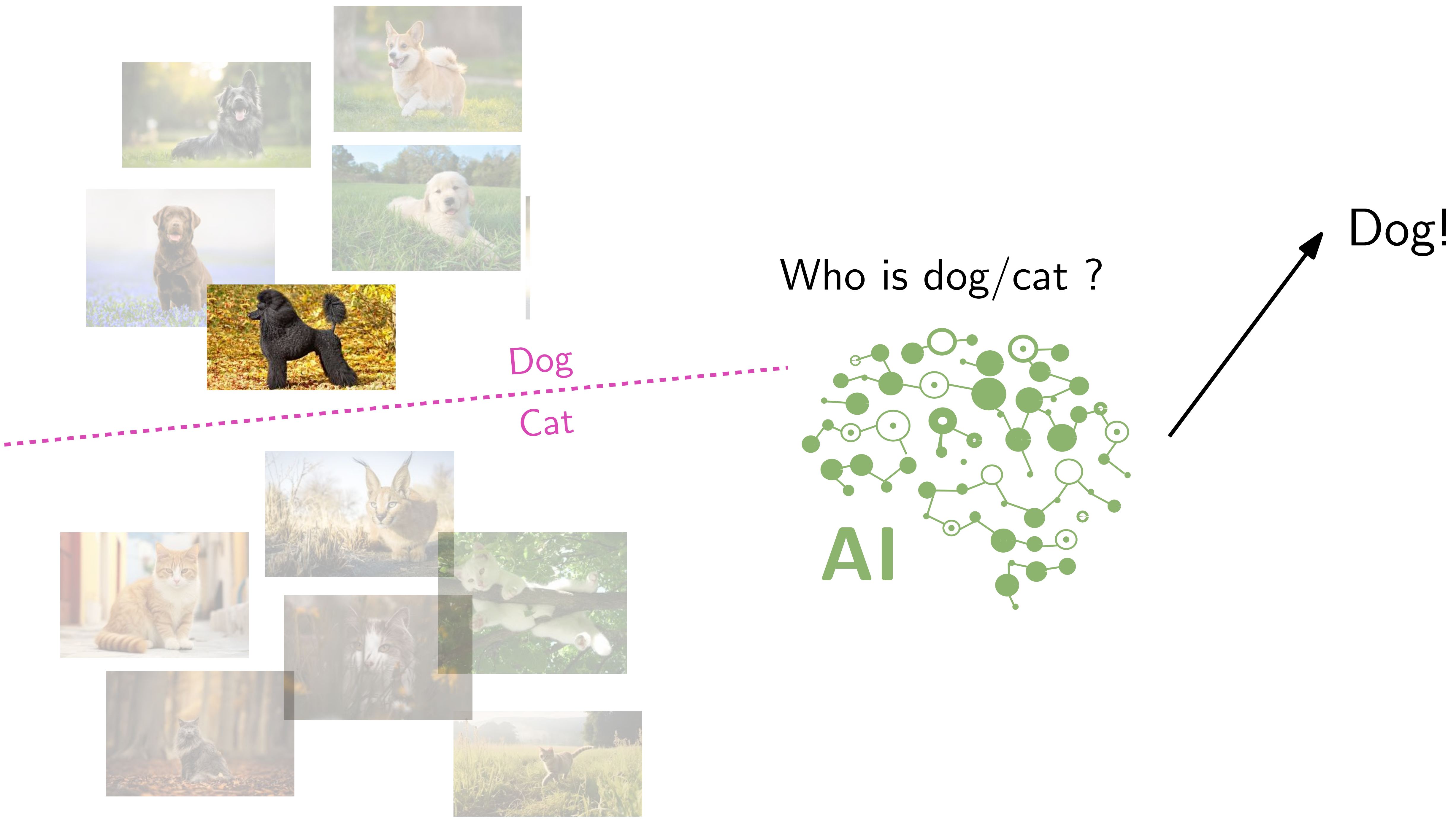
Geometric vs. Probabilistic view on data



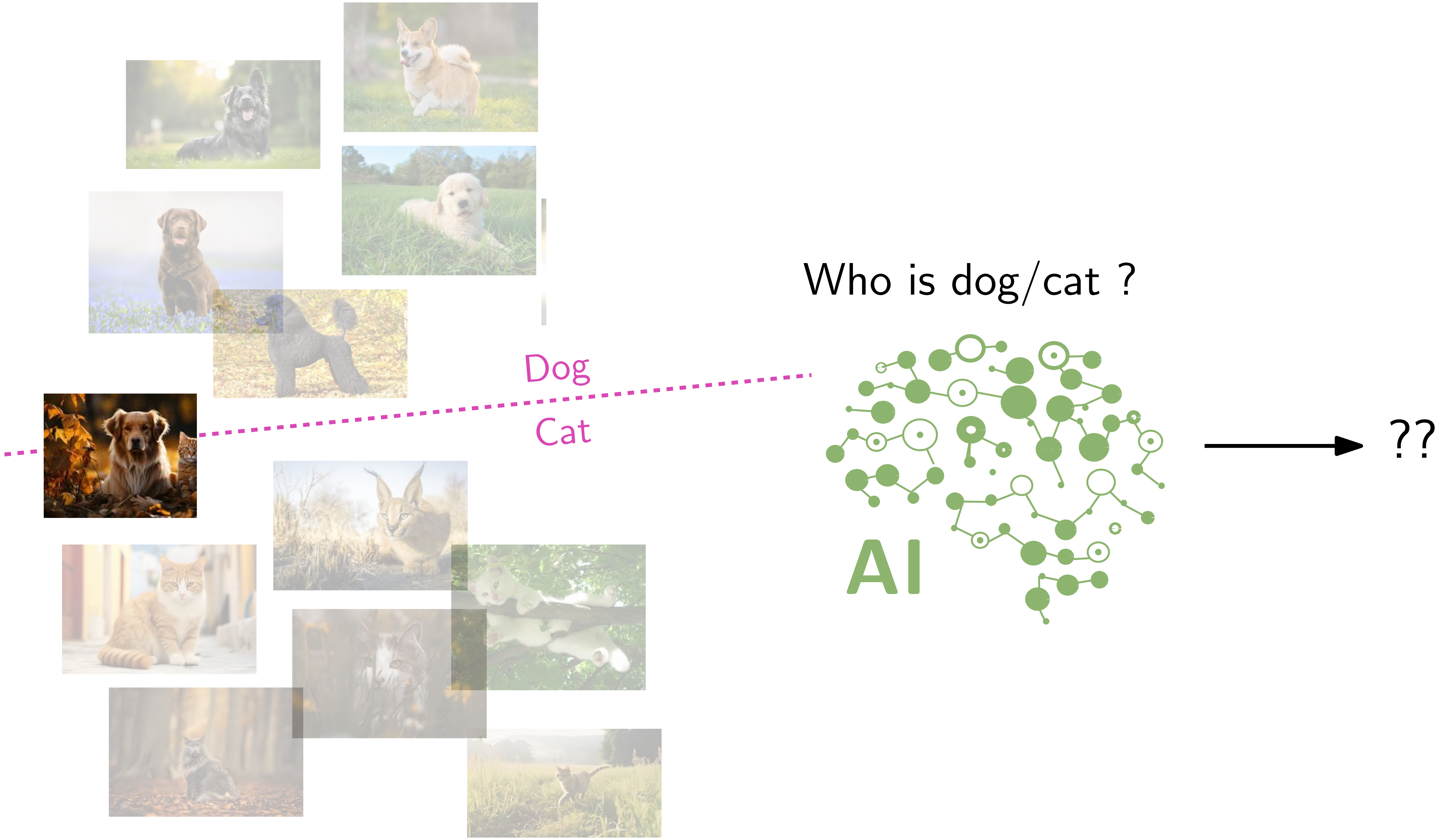
Who is dog/cat ?

Cat!

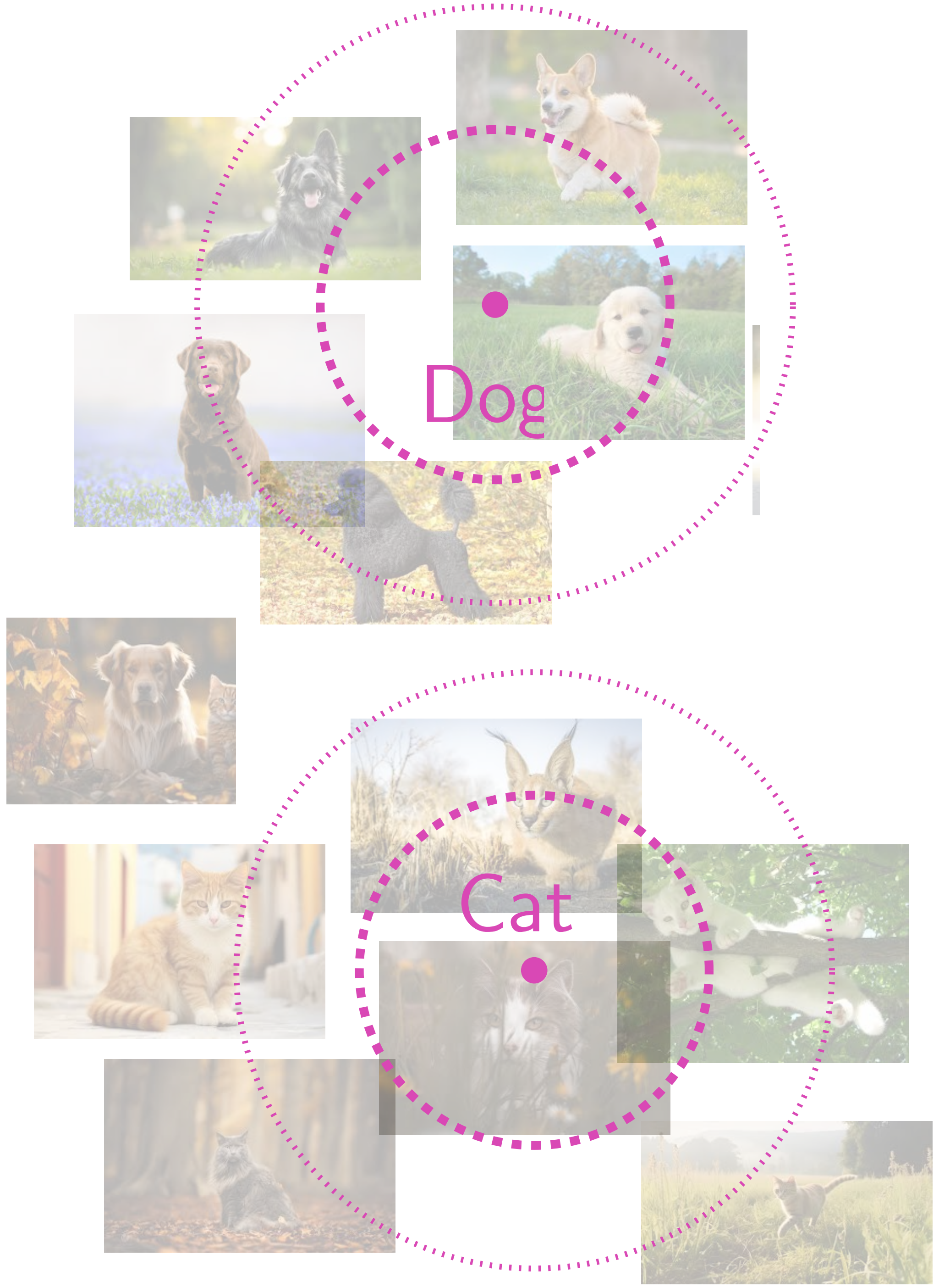
Geometric vs. Probabilistic view on data



Geometric vs. Probabilistic view on data

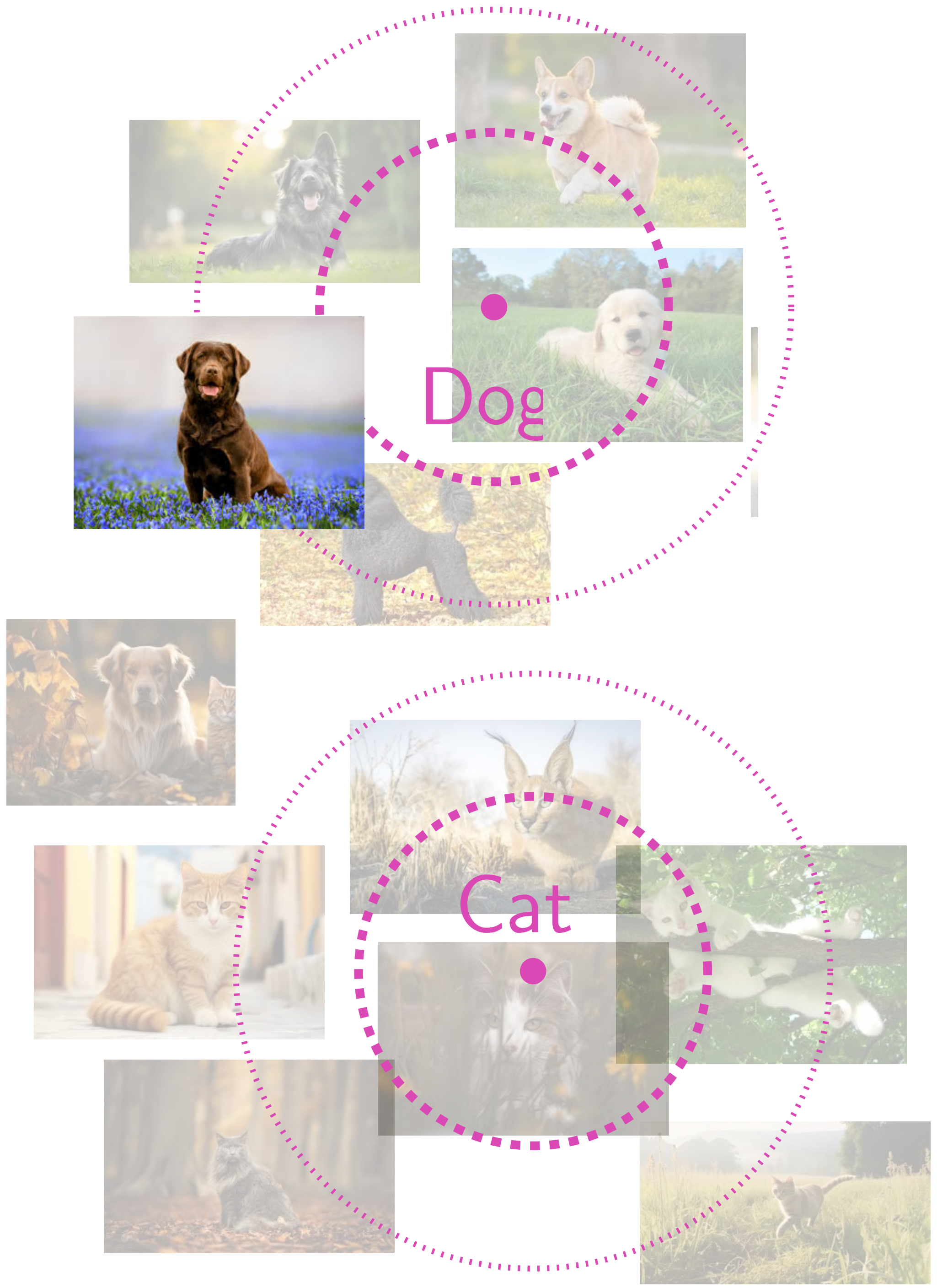


Geometric vs. Probabilistic view on data

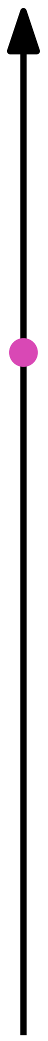


Who is dog/cat ?

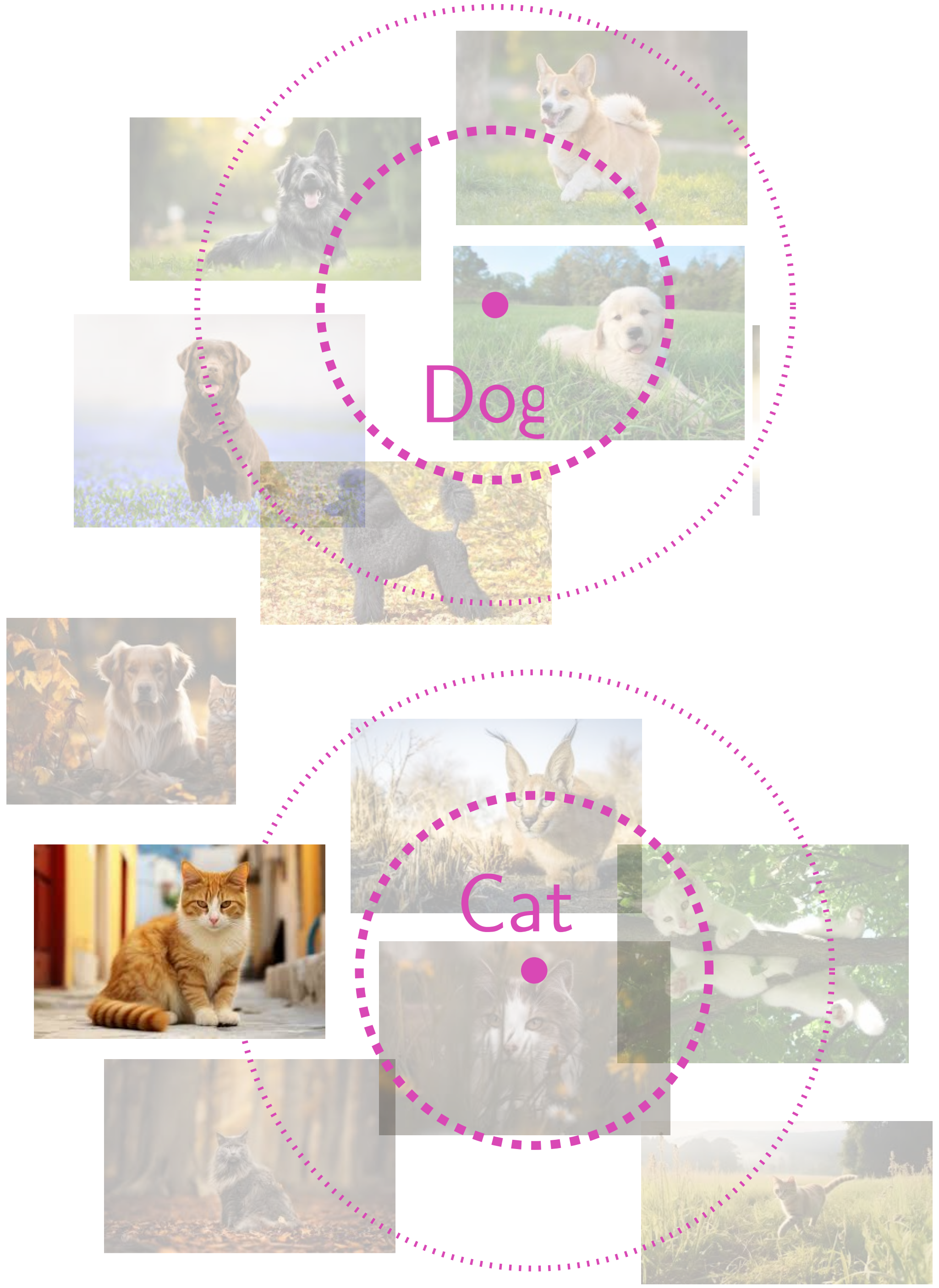
Geometric vs. Probabilistic view on data



Who is dog/cat ?

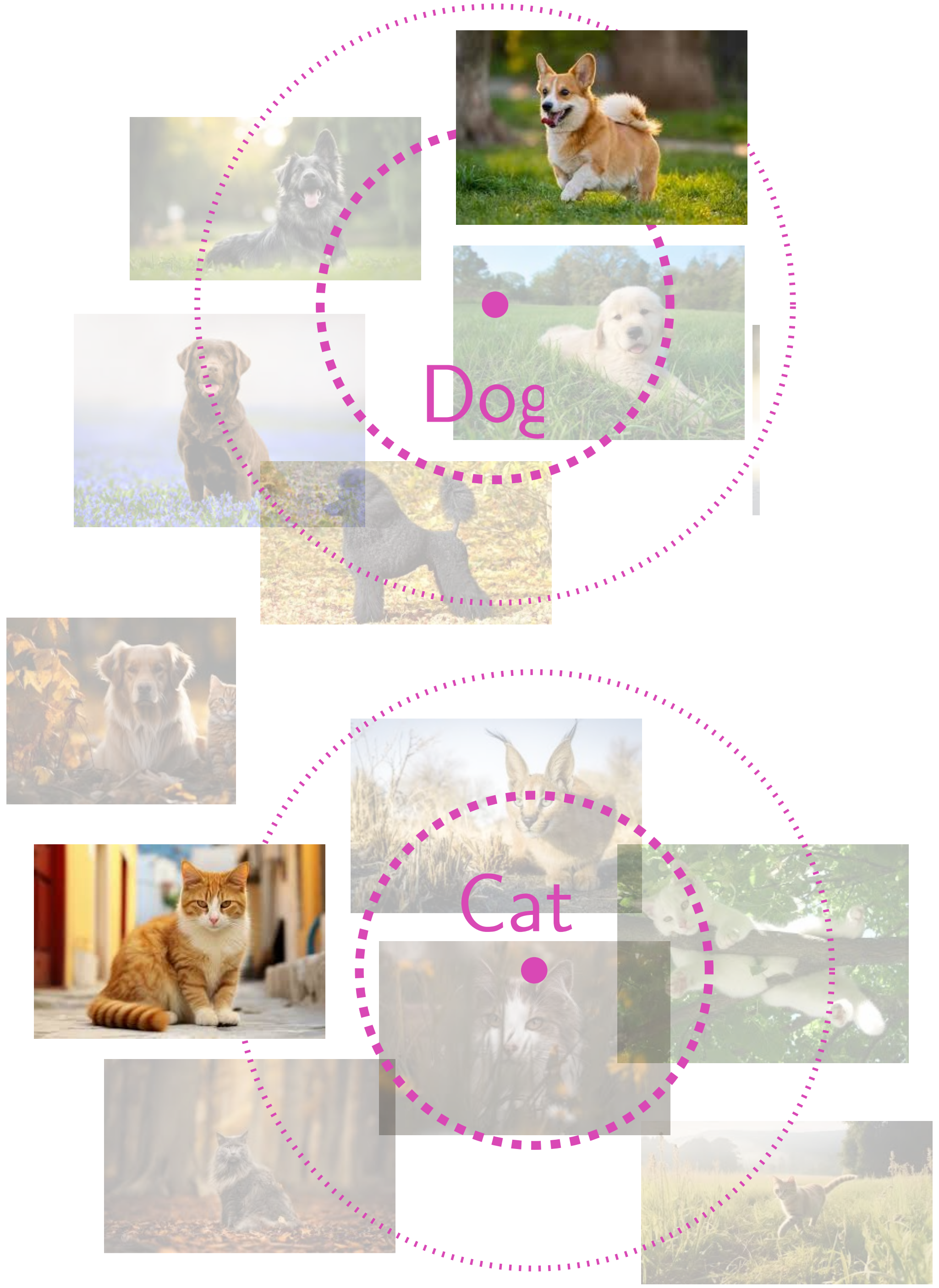


Geometric vs. Probabilistic view on data



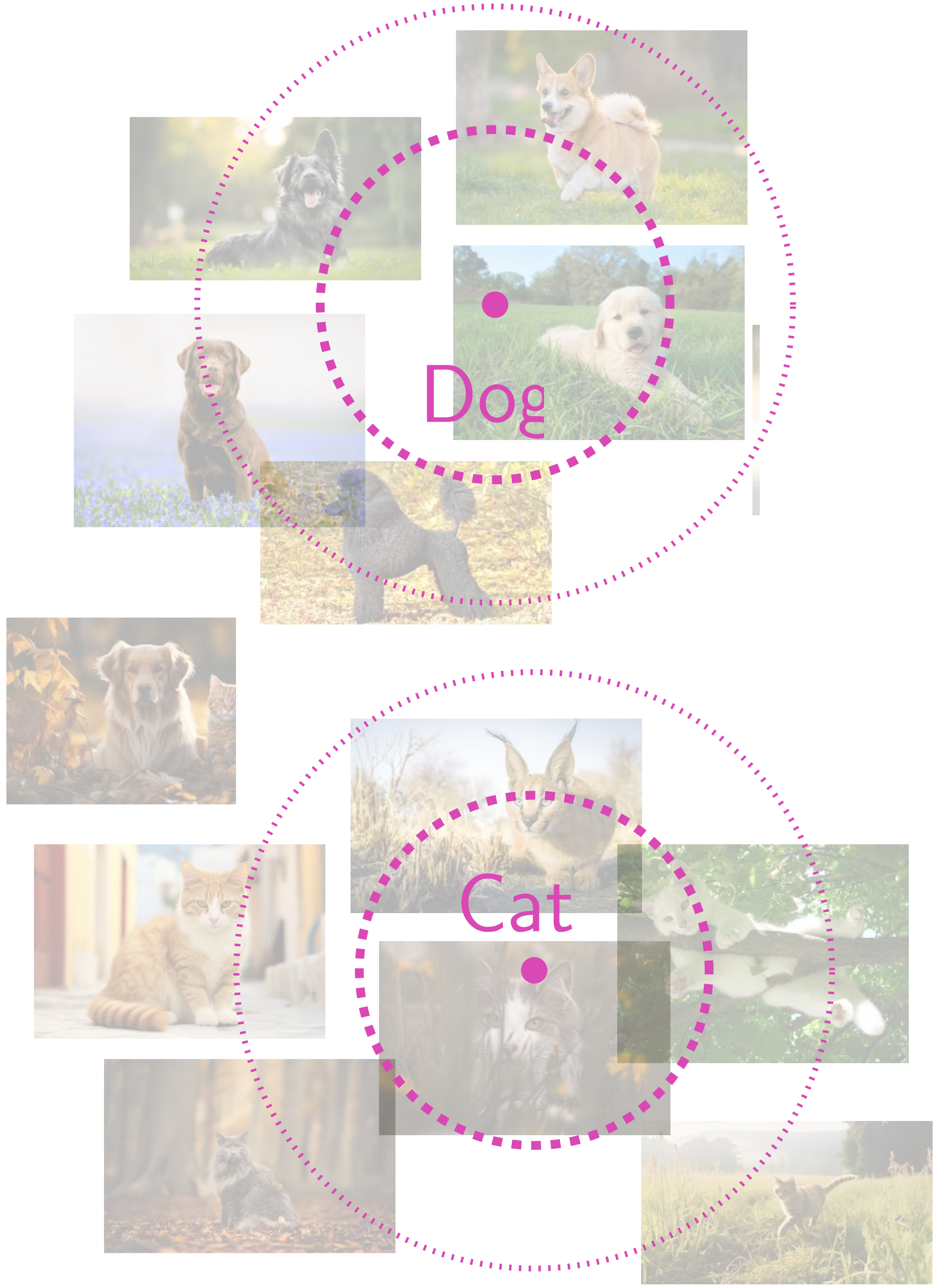
Who is dog/cat ?

Geometric vs. Probabilistic view on data

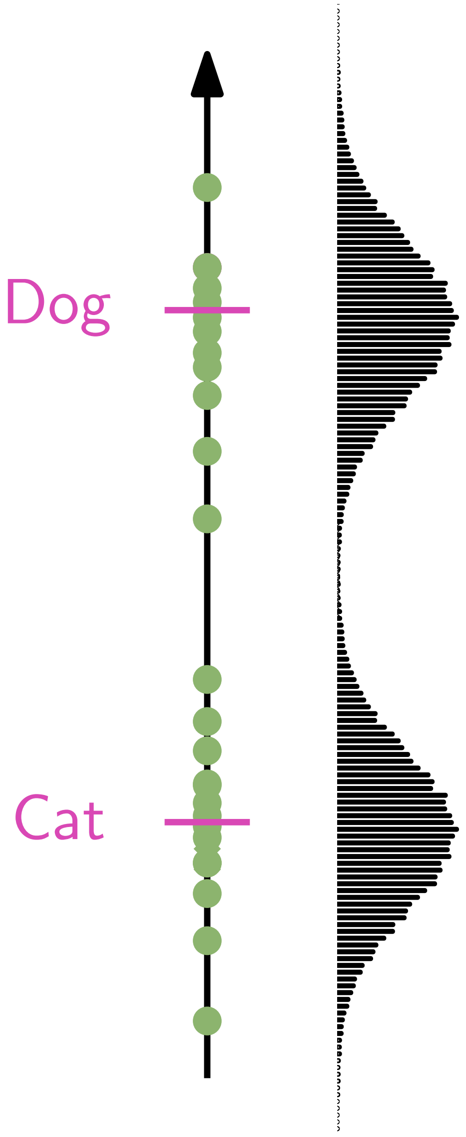


Who is dog/cat ?

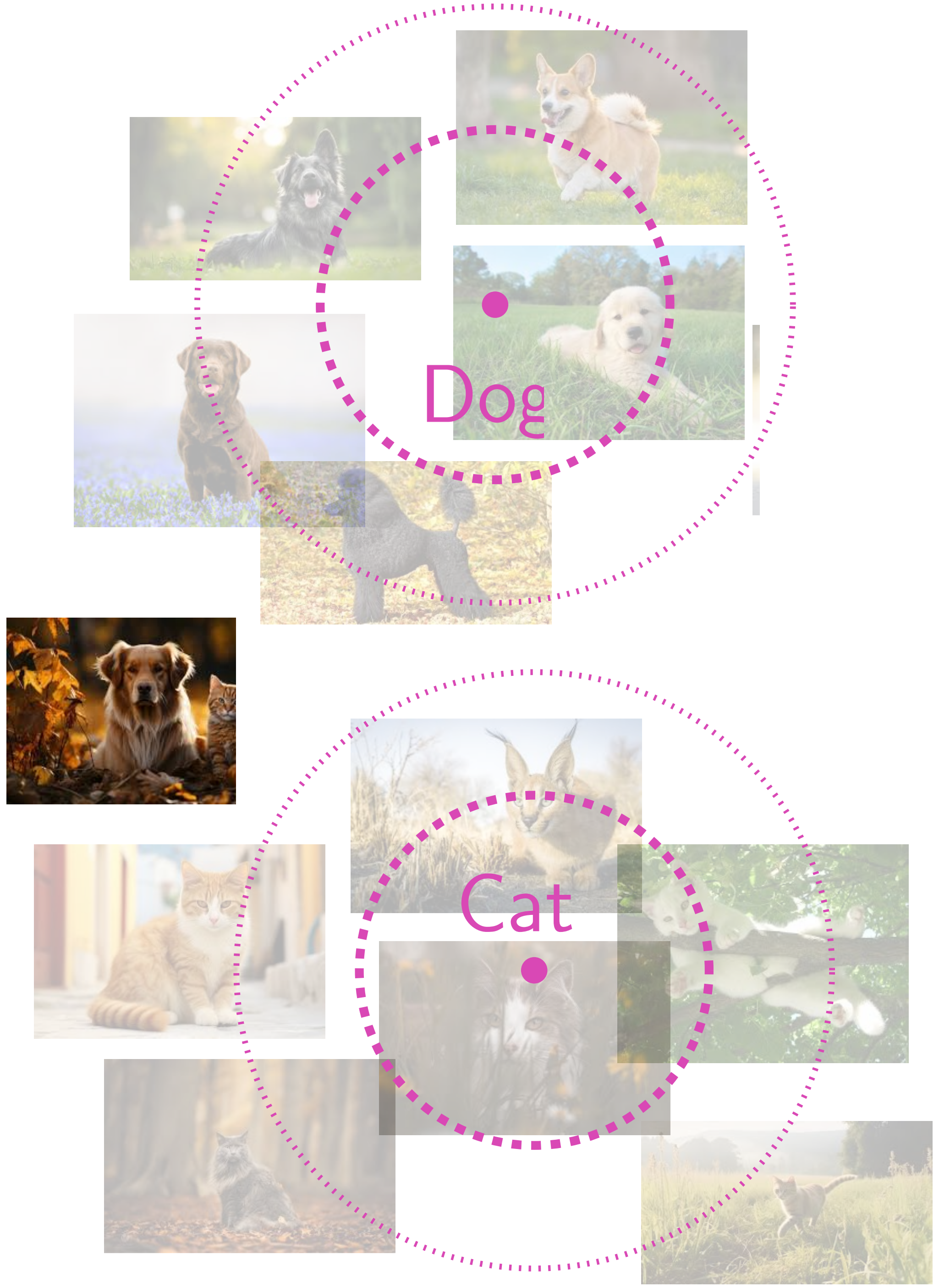
Geometric vs. Probabilistic view on data



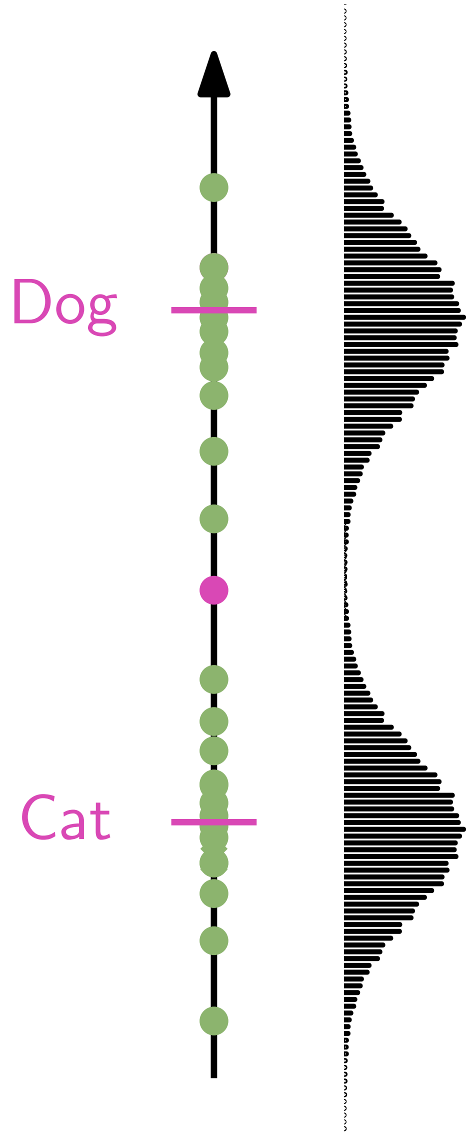
Who is dog/cat ?



Geometric vs. **Probabilistic** view on data

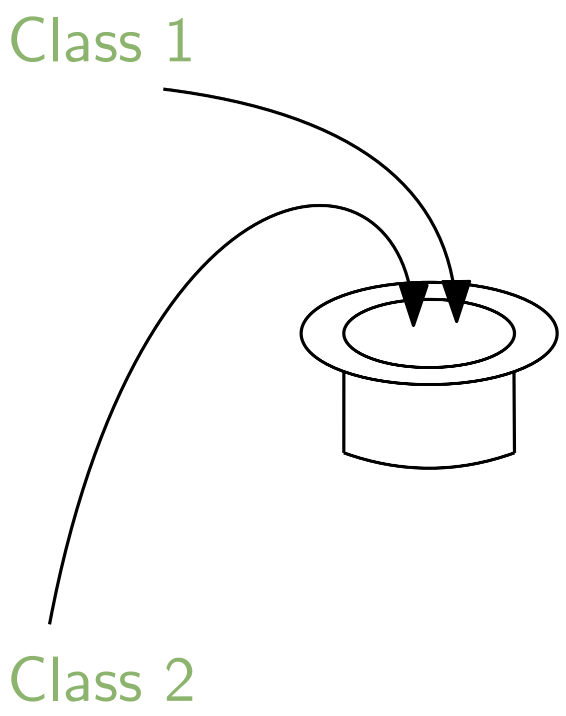


Who is dog/cat ?



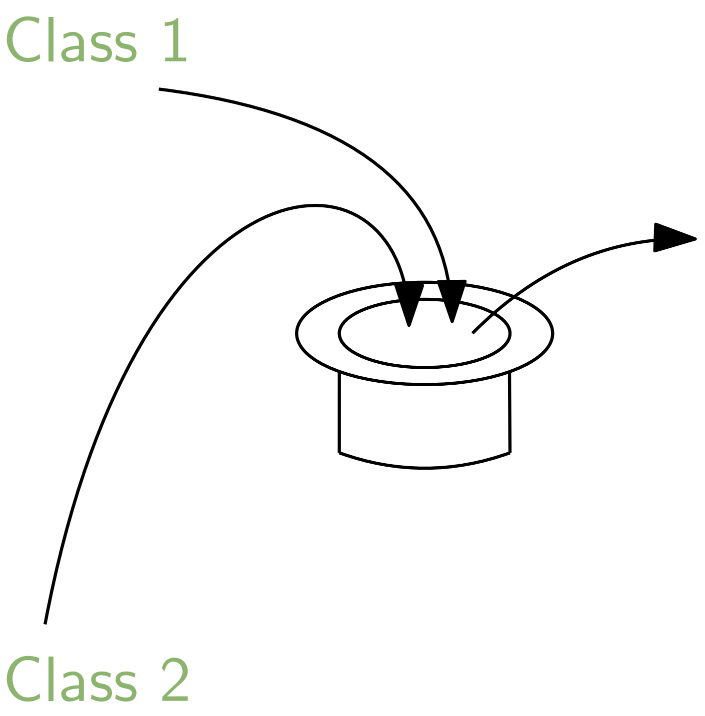
What is random in statistical learning ?

Ex 1: Unsupervised learning



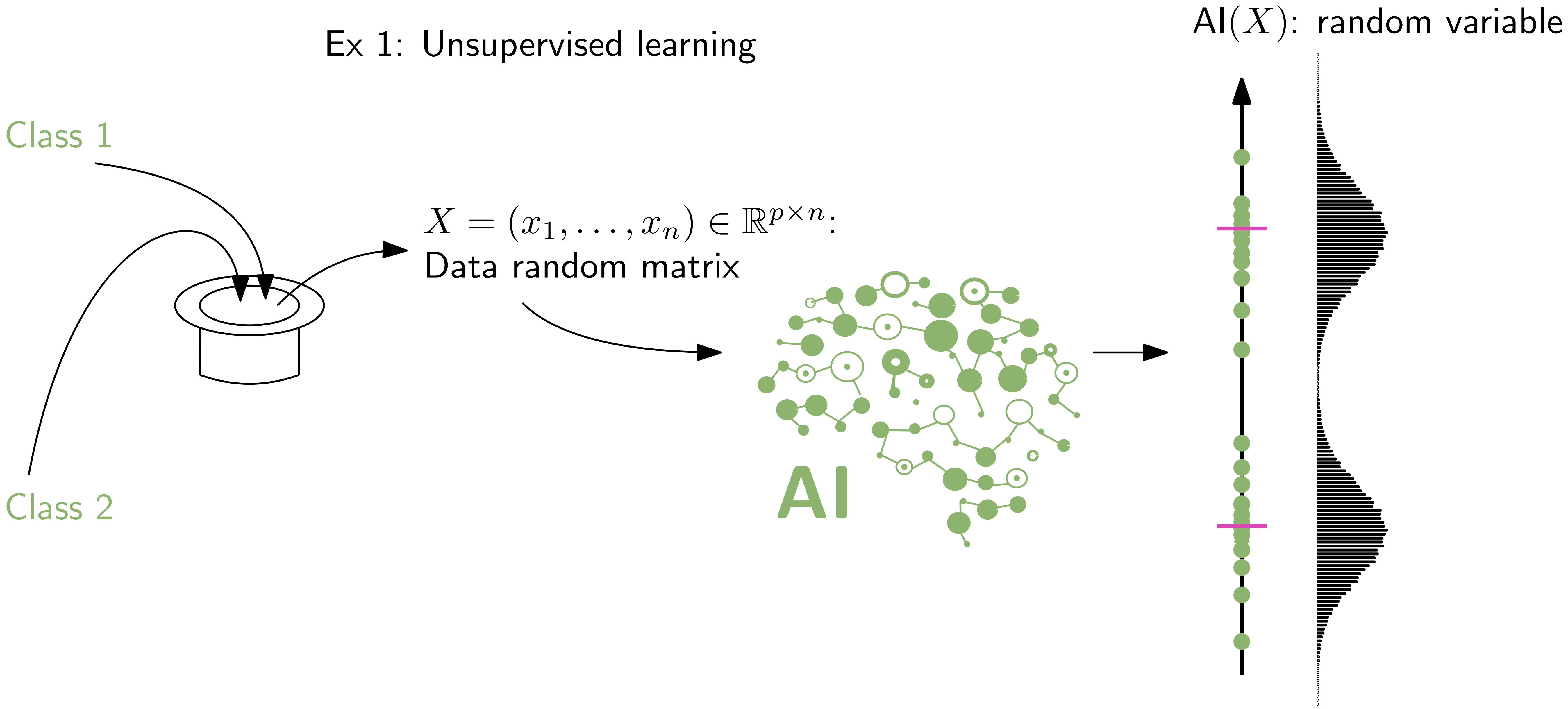
What is random in statistical learning ?

Ex 1: Unsupervised learning



$X = (x_1, \ldots, x_n) \in \mathbb{R}^{p \times n}$:
Data random matrix

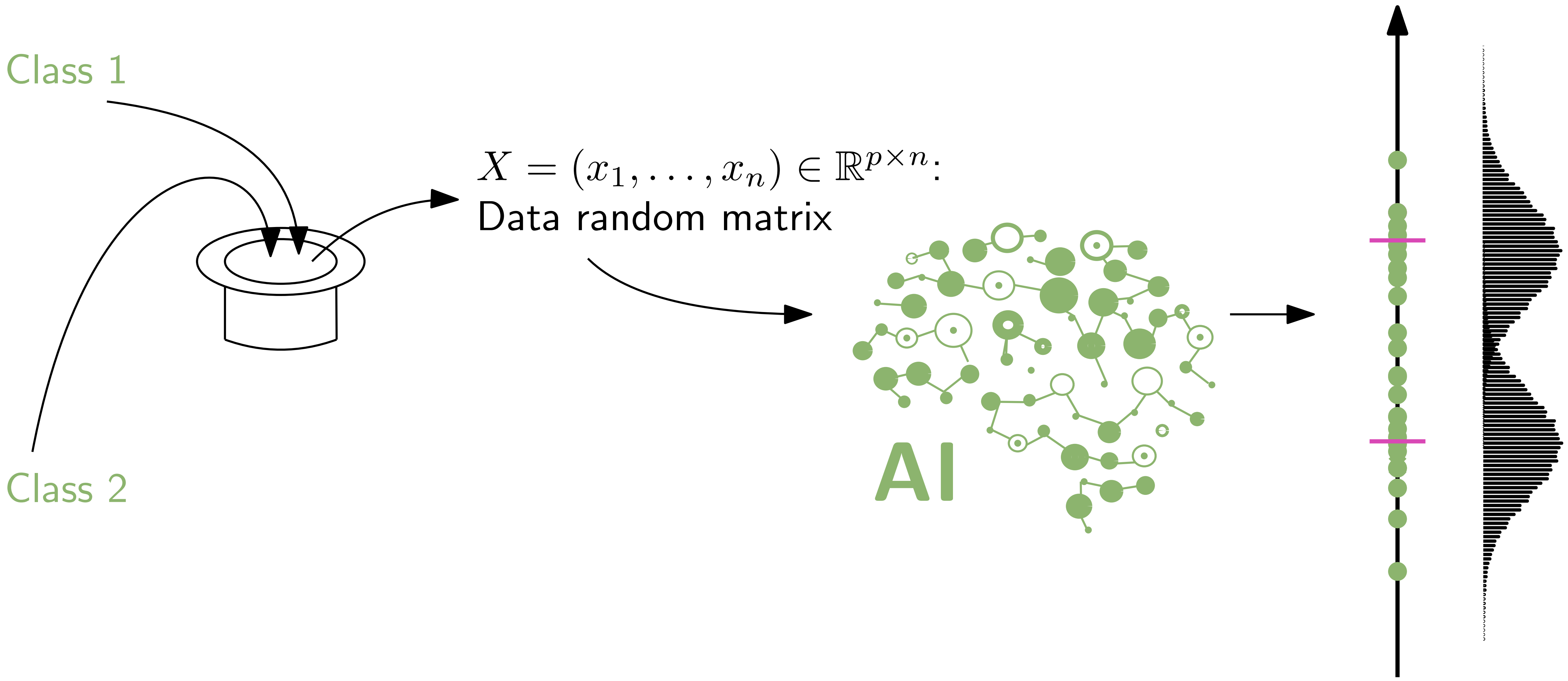
What is random in statistical learning ?



What is random in statistical learning ?

Ex 1: Unsupervised learning

AI(X): random variable



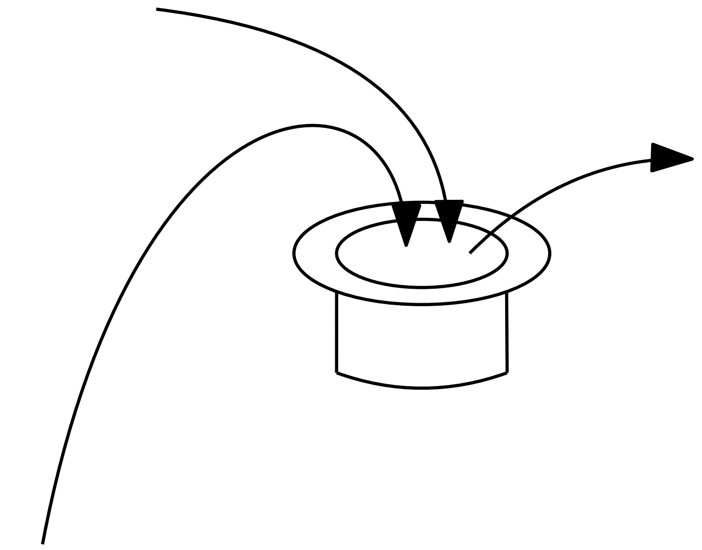
$Y = (y_1, \ldots, y_n) \in \{-1, 1\}^n$

Estimate:

$\mathbb{E}[\|\mathrm{AI}(X) - Y\|]$

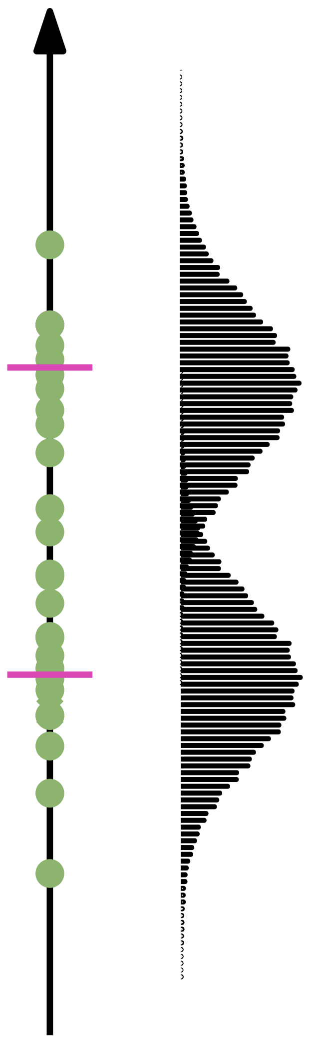
What is random in statistical learning ?

Ex 1: Unsupervised learning



$X = (x_1, \dots, x_n) \in \mathbb{R}^{p \times n}$:
Data random matrix

AI(X): random variable



$Y = (y_1, \dots, y_n) \in \{-1, 1\}^n$

Estimate:

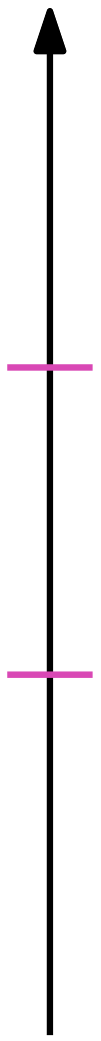
$\mathbb{E}[\|\mathrm{AI}(X) - Y\|]$

$\mathbb{P}(|\mathrm{AI}(x_i) - y_i| \geq 1)$

What is random in statistical learning ?

Ex 2: Supervised learning

Class 1



Class 2

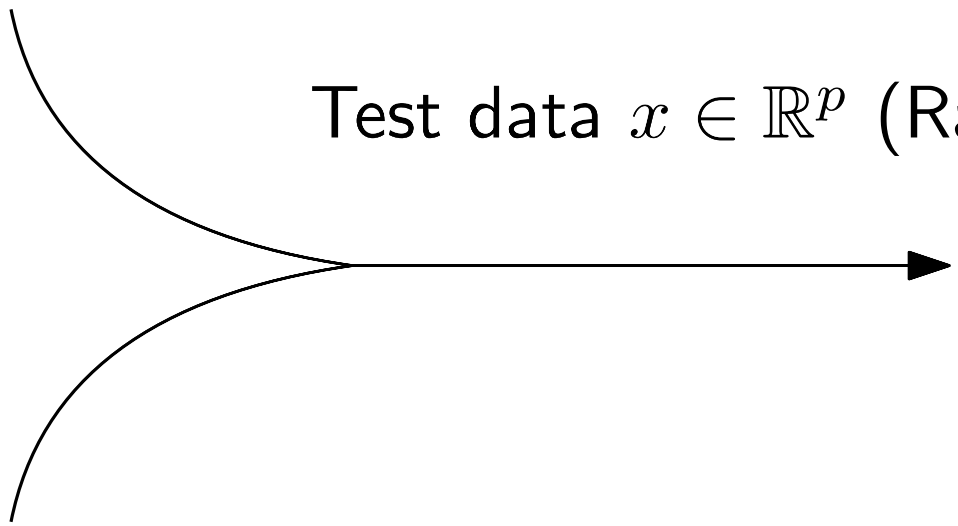
$X = (x_1, \dots, x_n) \in \mathbb{R}^{p \times n}$:
Training data (**random**)

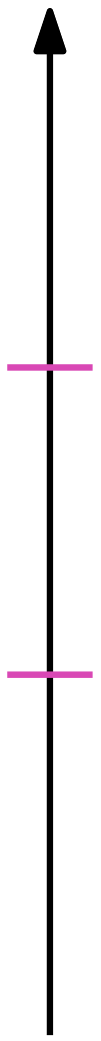
What is random in statistical learning ?

Ex 2: Supervised learning

Class 1

Test data $x \in \mathbb{R}^p$ (Random)





Class 2

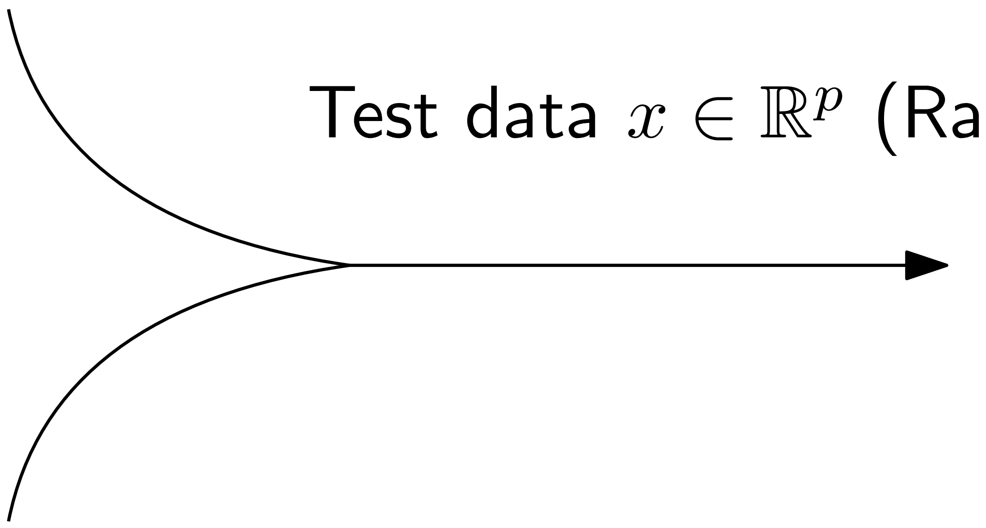
$X = (x_1, \ldots, x_n) \in \mathbb{R}^{p \times n}$:
Training data (**random**)

What is random in statistical learning ?

Ex 2: Supervised learning

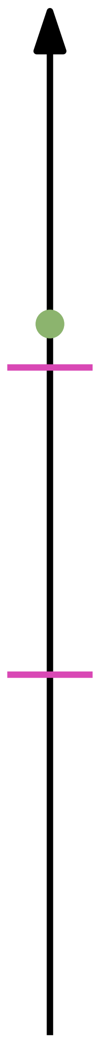
Class 1

Test data $x \in \mathbb{R}^p$ (Random)



Class 2

AI$_X(x)$ (random)



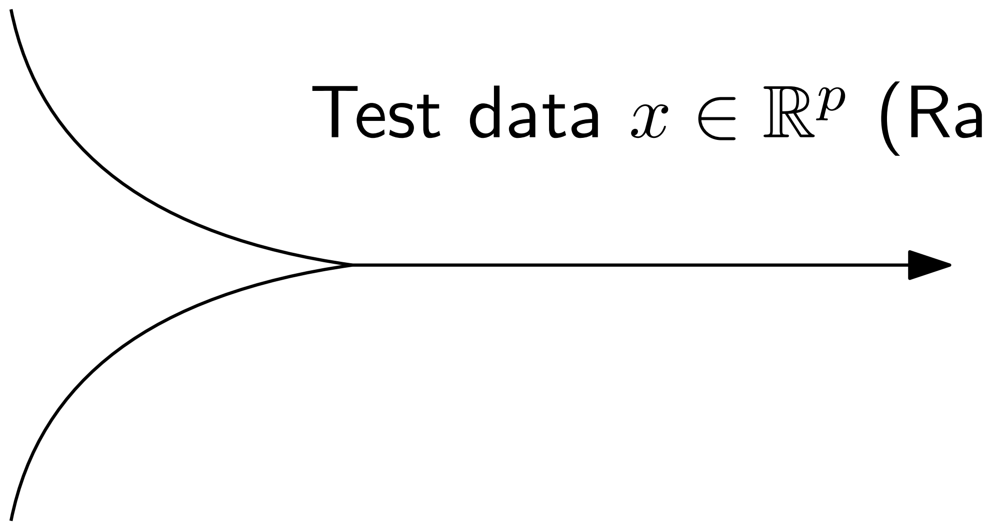
$X = (x_1, \ldots, x_n) \in \mathbb{R}^{p \times n}$:
Training data (**random**)

What is random in statistical learning ?

Ex 2: Supervised learning

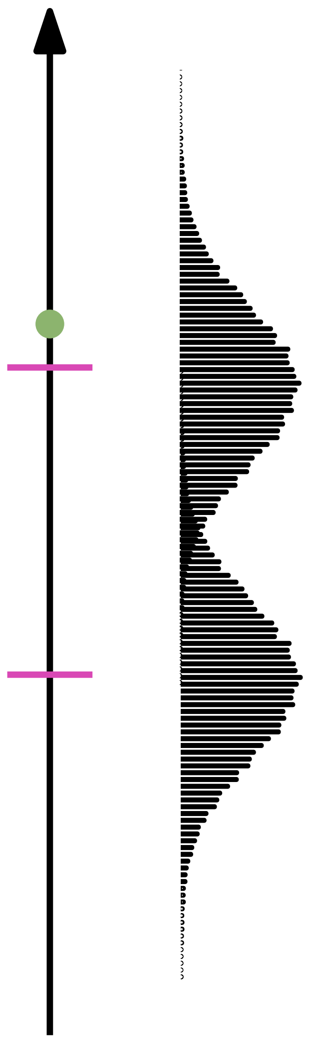
Class 1

Test data $x \in \mathbb{R}^p$ (Random)



Class 2

$\mathsf{AI}_X(x)$ (random)



$y \in \{-1, 1\}$: Label of x

Estimate:
$\mathbb{E}[\|\mathsf{AI}_X(x) - y\|]$
$\mathbb{P}(|\mathsf{AI}_X(x) - y| \geq 1)$

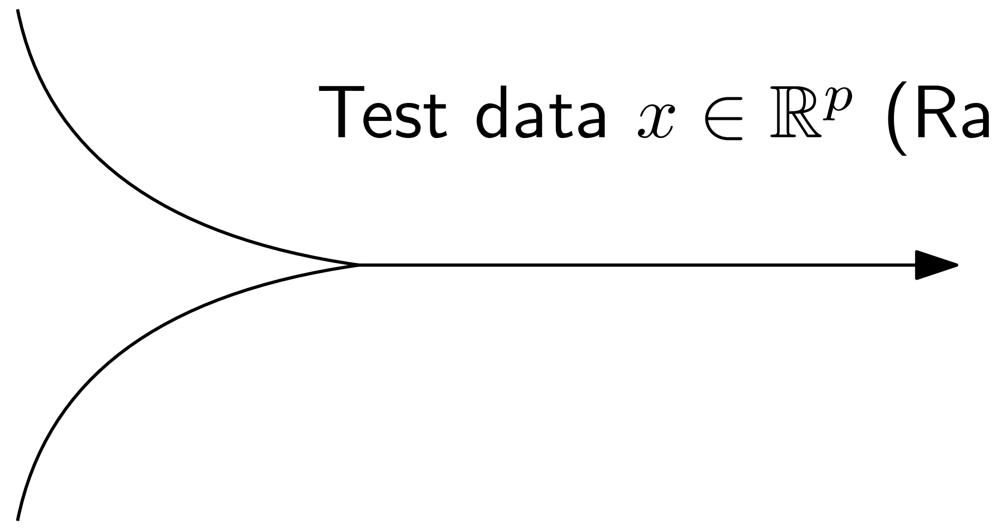
$X = (x_1, \ldots, x_n) \in \mathbb{R}^{p \times n}$:
Training data (**random**)

What is random in statistical learning ?

Ex 2: Supervised learning

Class 1

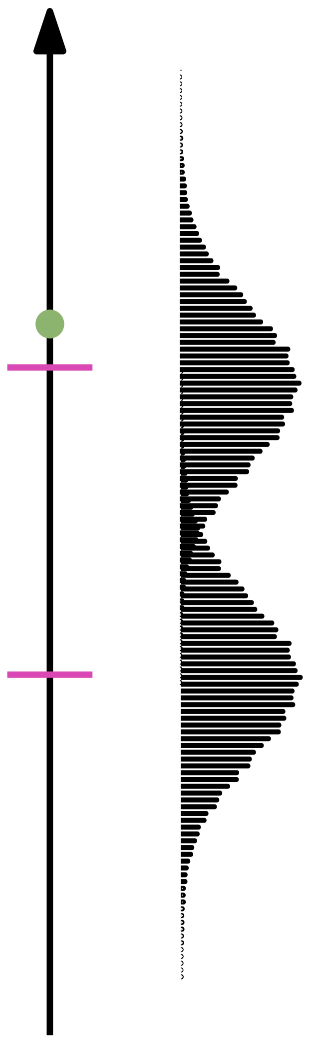
Test data $x \in \mathbb{R}^p$ (Random)



Class 2

$X = (x_1, \ldots, x_n) \in \mathbb{R}^{p \times n}$:
Training data (**random**)

$\text{AI}_X(x)$ (random)



$y \in \{-1, 1\}$: Label of x

Estimate:
$\mathbb{E}[\|\text{AI}_X(x) - y\|]$
$\mathbb{P}(|\text{AI}_X(x) - y| \geq 1)$

Content of the course

I - Major Concepts with examples

Paradigm	Interpretable/ Flexible, accurate	Regression task/ Classification task	Supervised/ Unsupervised	Interpolates/ Extrapolates
Tools & Concepts involved	Parameters Bias/variance trade-off	discriminent analysis k-means loss (mse, logistic, ℓ_1)	Resampling (Cross-validation, Bootstrap, Random forest)	Regularization, overfitting Perameter's degree of liberty
Methods				
• Minimization Problem				
Likelyhood minimization	✓/ ×	× ✓	×/ ✓	✓/ ✓
Regression	✓/ ×	✓/ ×	×/ ✓	✓/ ✓
Logistic Regression	✓/ ×	×/ ✓	✓/ ✓	
Ridge regression	×/ ✓	✓/ (✓)	(✓) / ✓	(✓) / ✓
Multi-task learning	×/ ✓	×/ ✓	✓/ ×	
Empirical risk minimization Support vector machine	×/ ✓	×/ ✓	✓/ ✓	
Neural networks	×/ ✓	✓/ ✓	✓/ ×	✓/ ×
• Spectral methoods				
Principal component analysis (PCA)	×/ ✓	×/ ✓	×/ ✓	
Spectral clustering	×/ ✓	×/ ✓	×/ ✓	
• Monte Carlo		✓/ ×	✓/ (✓)	×/ ✓

Paradigm	Interpretable/ Flexible, accurate	Regression task/ Classification task	Supervised/ Unsupervised	Interpolates/ Extrapolates
Tools & Concepts involved	Parameters Bias/variance trade-off	discriminent analysis k-means loss (mse, logistic, ℓ_1)	Resampling (Cross-validation, Bootstrap, Random forest)	Regularization, overfitting Perameter's degree of liberty
Methods	I			
• Minimization Problem				
Likelyhood minimization	✓/ ×	× ✓	×/ ✓	✓/ ✓
Regression	✓/ ×	✓/ ×	×/ ✓	✓/ ✓
Logistic Regression	✓/ ×	×/ ✓	✓/ ✓	
Ridge regression	×/ ✓	✓/ (✓)	(✓) / ✓	(✓) / ✓
Multi-task learning	×/ ✓	×/ ✓	✓/ ×	
Empirical risk minimization Support vector machine	×/ ✓	×/ ✓	✓/ ✓	
Neural networks	×/ ✓	✓/ ✓	✓/ ×	✓/ ×
• Spectral methods				
Principal component analysis (PCA)	×/ ✓	×/ ✓	×/ ✓	
Spectral clustering	×/ ✓	×/ ✓	×/ ✓	
• Monte Carlo		✓/ ×	✓/ (✓)	×/ ✓

Paradigm	Interpretable/ Flexible, accurate	Regression task/ Classification task	Supervised/ Unsupervised	Interpolates/ Extrapolates
Tools & Concepts involved	Parameters Bias/variance trade-off	discriminent analysis k-means loss (mse, logistic, ℓ_1)	Resampling (Cross-validation, Bootstrap, Random forest)	Regularization, overfitting Perameter's degree of liberty
Methods	I	II		
• Minimization Problem				
Likelyhood minimization	✓/ ×	× ✓	×/ ✓	✓/ ✓
Regression	✓/ ×	✓/ ×	×/ ✓	✓/ ✓
Logistic Regression	✓/ ×	×/ ✓	✓/ ✓	
Ridge regression	×/ ✓	✓/ (✓)	(✓) / ✓	(✓) / ✓
Multi-task learning	×/ ✓	×/ ✓	✓/ ×	
Empirical risk minimization Support vector machine	×/ ✓	×/ ✓	✓/ ✓	
Neural networks	×/ ✓	✓/ ✓	✓/ ×	✓/ ×
• Spectral methods				
Principal component analysis (PCA)	×/ ✓	×/ ✓	×/ ✓	
Spectral clustering	×/ ✓	×/ ✓	×/ ✓	
• Monte Carlo		✓/ ×	✓/ (✓)	×/ ✓

Paradigm	Interpretable/ Flexible, accurate	Regression task/ Classification task	Supervised/ Unsupervised	Interpolates/ Extrapolates
Tools & Concepts involved	Parameters Bias/variance trade-off	discriminent analysis k-means loss (mse, logistic, ℓ_1)	Resampling (Cross-validation, Bootstrap, Random forest)	Regularization, overfitting Perameter's degree of liberty
Methods	I	II	III	
• Minimization Problem				
Likelyhood minimization	✓/ ×	× ✓	×/ ✓	✓/ ✓
Regression	✓/ ×	✓/ ×	×/ ✓	✓/ ✓
Logistic Regression	✓/ ×	×/ ✓	✓/ ✓	
Ridge regression	×/ ✓	✓/ (✓)	(✓) / ✓	(✓) / ✓
Multi-task learning	×/ ✓	×/ ✓	✓/ ×	
Empirical risk minimization Support vector machine	×/ ✓	×/ ✓	✓/ ✓	
Neural networks	×/ ✓	✓/ ✓	✓/ ×	✓/ ×
• Spectral methods				
Principal component analysis (PCA)	×/ ✓	×/ ✓	×/ ✓	
Spectral clustering	×/ ✓	×/ ✓	×/ ✓	
• Monte Carlo		✓/ ×	✓/ (✓)	×/ ✓

Paradigm	Interpretable/ Flexible, accurate	Regression task/ Classification task	Supervised/ Unsupervised	Interpolates/ Extrapolates
Tools & Concepts involved	Parameters Bias/variance trade-off	discriminent analysis k-means loss (mse, logistic, ℓ_1)	Resampling (Cross-validation, Bootstrap, Random forest)	Regularization, overfitting Perameter's degree of liberty
Methods	I	II	III	IV
• Minimization Problem				
Likelyhood minimization	$\checkmark$ / $\times$	$\times$ $\checkmark$	$\times$/ $\checkmark$	$\checkmark$/ $\checkmark$
Regression	$\checkmark$ / $\times$	$\checkmark$ / $\times$	$\times$/ $\checkmark$	$\checkmark$/ $\checkmark$
Logistic Regression	$\checkmark$ / $\times$	$\times$/ $\checkmark$	$\checkmark$/ $\checkmark$	
Ridge regression	$\times$/ $\checkmark$	$\checkmark$/ ($\checkmark$)	($\checkmark$) / $\checkmark$	($\checkmark$) / $\checkmark$
Multi-task learning	$\times$/ $\checkmark$	$\times$/ $\checkmark$	$\checkmark$/ $\times$	
Empirical risk minimization Support vector machine	$\times$/ $\checkmark$	$\times$/ $\checkmark$	$\checkmark$/ $\checkmark$	
Neural networks	$\times$/ $\checkmark$	$\checkmark$/ $\checkmark$	$\checkmark$/ $\times$	$\checkmark$/ $\times$
• Spectral methods				
Principal component analysis (PCA)	$\times$/ $\checkmark$	$\times$/ $\checkmark$	$\times$/ $\checkmark$	
Spectral clustering	$\times$/ $\checkmark$	$\times$/ $\checkmark$	$\times$/ $\checkmark$	
• Monte Carlo		$\checkmark$/ $\times$	$\checkmark$/ ($\checkmark$)	$\times$/ $\checkmark$

Content of the course

I - Major Concepts with examples

a) Interpretable vs. flexible methods.

b) Classification and regression algorithms.

c) Supervised and Unsupervised algorithms.

d) Interpolation vs. Extrapolation.

II - Rigorous probability inferences

Content of the course

I - Major Concepts with examples

a) Interpretable vs. flexible methods.

b) Classification and regression algorithms.

c) Supervised and Unsupervised algorithms.

d) Interpolation vs. Extrapolation.

II - Rigorous probability inferences

a) Bias/variance tradeoff, estimator evaluation, MLE

Content of the course

I - Major Concepts with examples

a) Interpretable vs. flexible methods.

b) Classification and regression algorithms.

c) Supervised and Unsupervised algorithms.

d) Interpolation vs. Extrapolation.

II - Rigorous probability inferences

a) Bias/variance tradeoff, estimator evaluation, MLE

b) Conditioning and Bayes computation, MAP

Content of the course

I - Major Concepts with examples

a) Interpretable vs. flexible methods.

b) Classification and regression algorithms.

c) Supervised and Unsupervised algorithms.

d) Interpolation vs. Extrapolation.

II - Rigorous probability inferences

a) Bias/variance tradeoff, estimator evaluation, MLE

b) Conditioning and Bayes computation, MAP

c) Concentration of the measure phenomenon (basics), the curse of dimension.

Content of the course

I - Major Concepts with examples

a) Interpretable vs. flexible methods.

b) Classification and regression algorithms.

c) Supervised and Unsupervised algorithms.

d) Interpolation vs. Extrapolation.

II - Rigorous probability inferences

a) Bias/variance tradeoff, estimator evaluation, MLE

b) Conditioning and Bayes computation, MAP

c) Concentration of the measure phenomenon (basics), the curse of dimension.

d) Random matrix basics

Work with Python

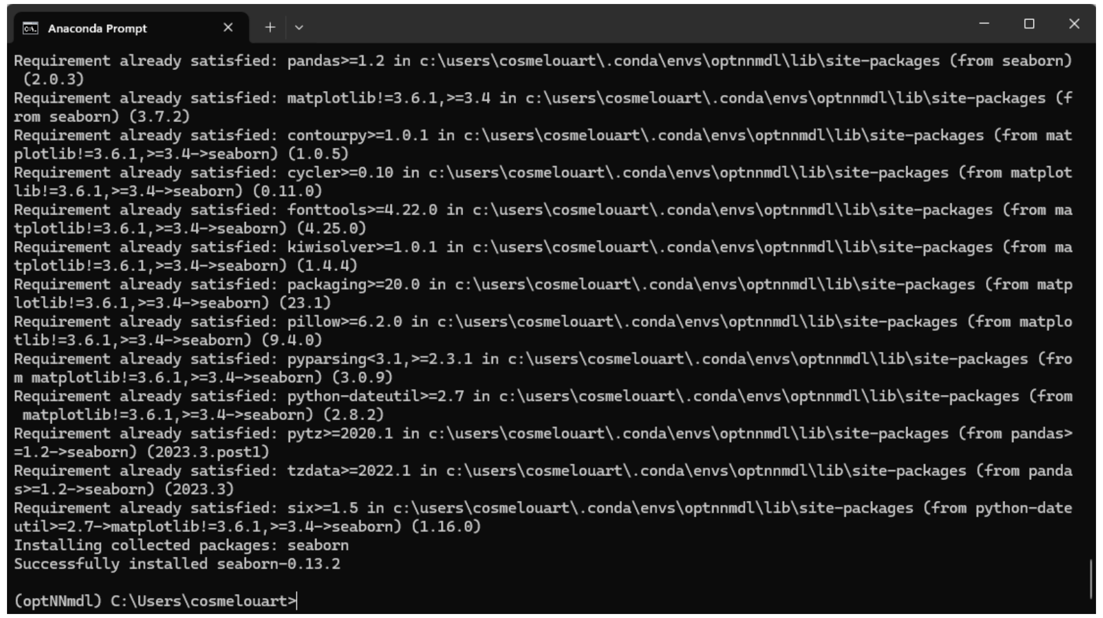
I - Download miniconda from
`https://docs.anaconda.com/miniconda/`

Work with Python

I - Download miniconda from
`https://docs.anaconda.com/miniconda/`

install packages with:

pip install -i https://pypi.tuna.tsinghua.edu.cn/simple some-package

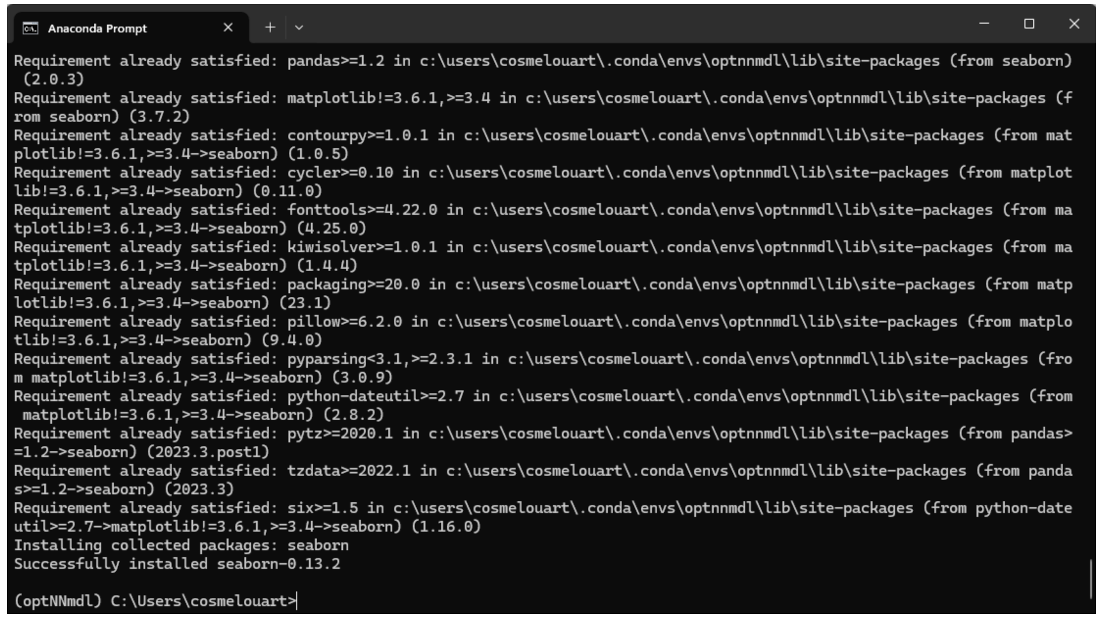


Work with Python

1 - Download miniconda from
`https://docs.anaconda.com/miniconda/`

install packages with:

pip install -i https://pypi.tuna.tsinghua.edu.cn/simple some-package



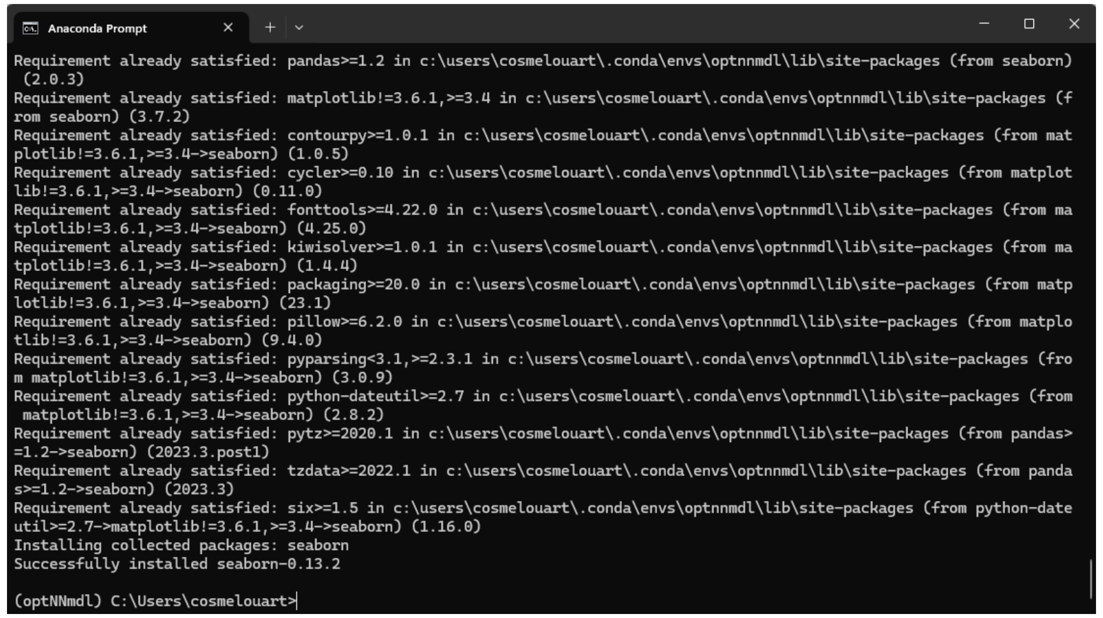
Packages required:
`ipykernel,numpy,pandas,mathplotlib`

Work with Python

I - Download miniconda from
`https://docs.anaconda.com/miniconda/`

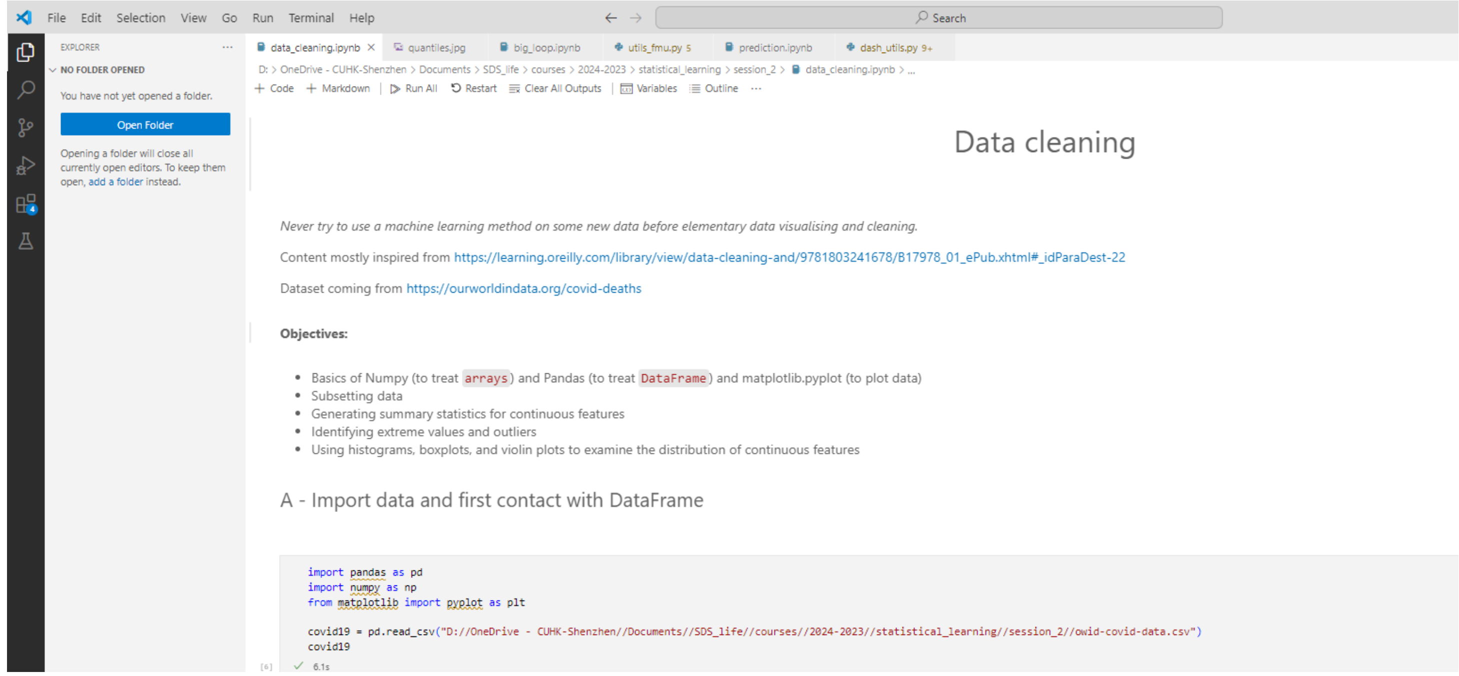
install packages with:

pip install -i https://pypi.tuna.tsinghua.edu.cn/simple some-package



Packages required:
`ipykernel,numpy,pandas,mathplotlib`

II - Download Visual studio code from:
`https://code.visualstudio.com/download`

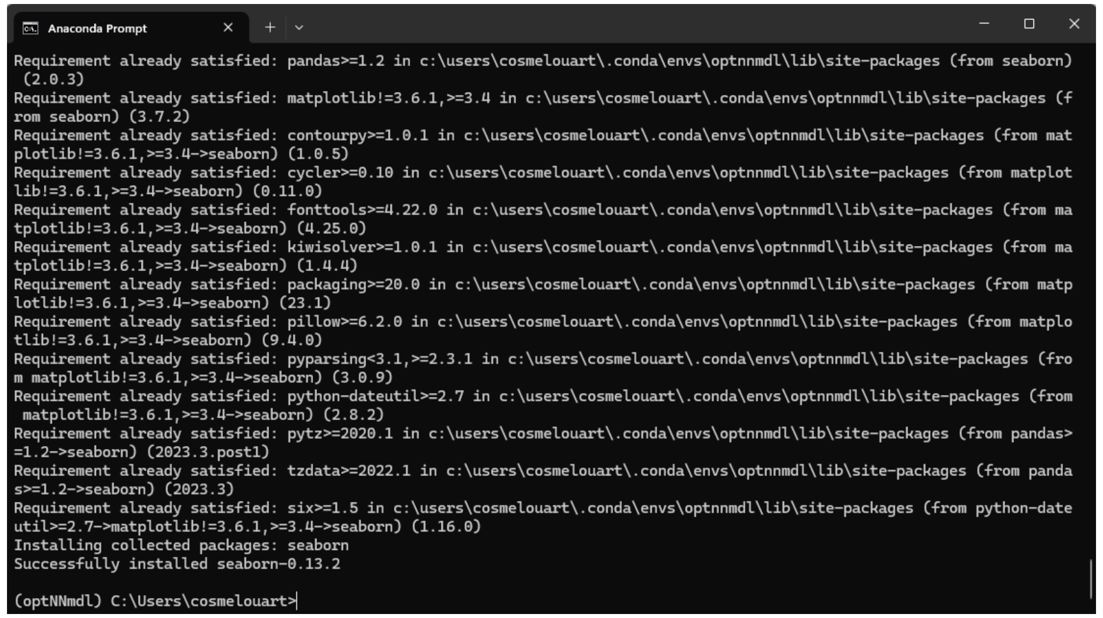


Work with Python

I - Download miniconda from
`https://docs.anaconda.com/miniconda/`

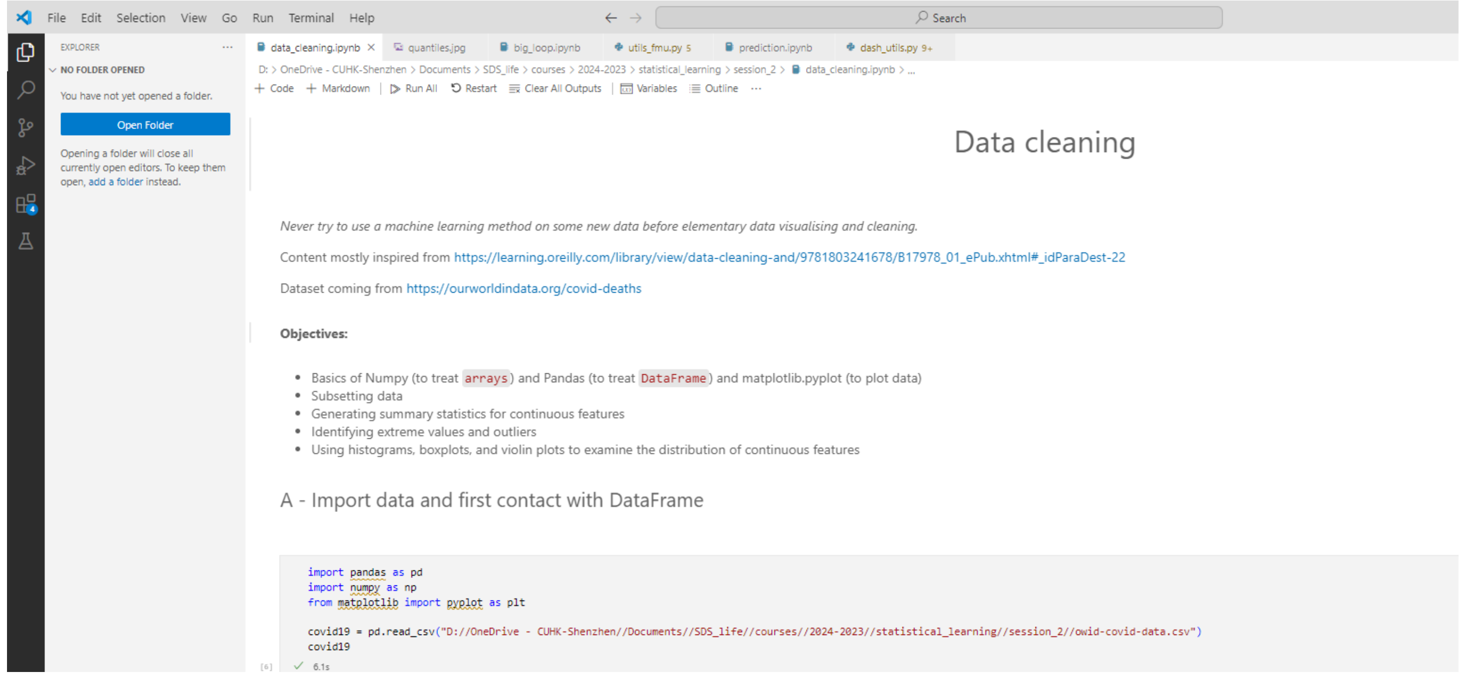
install packages with:

pip install -i https://pypi.tuna.tsinghua.edu.cn/simple some-package



Packages required:
`ipykernel,numpy,pandas,mathplotlib`

II - Download Visual studio code from:
`https://code.visualstudio.com/download`



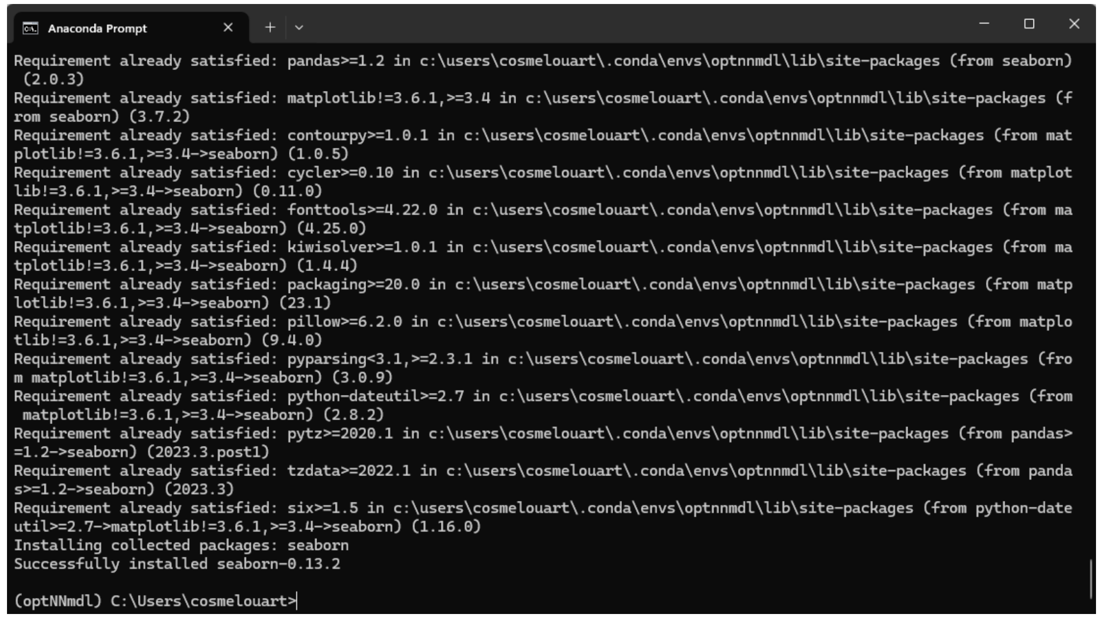
Install extensions: `python,jupyter`

Work with Python

I - Download miniconda from
`https://docs.anaconda.com/miniconda/`

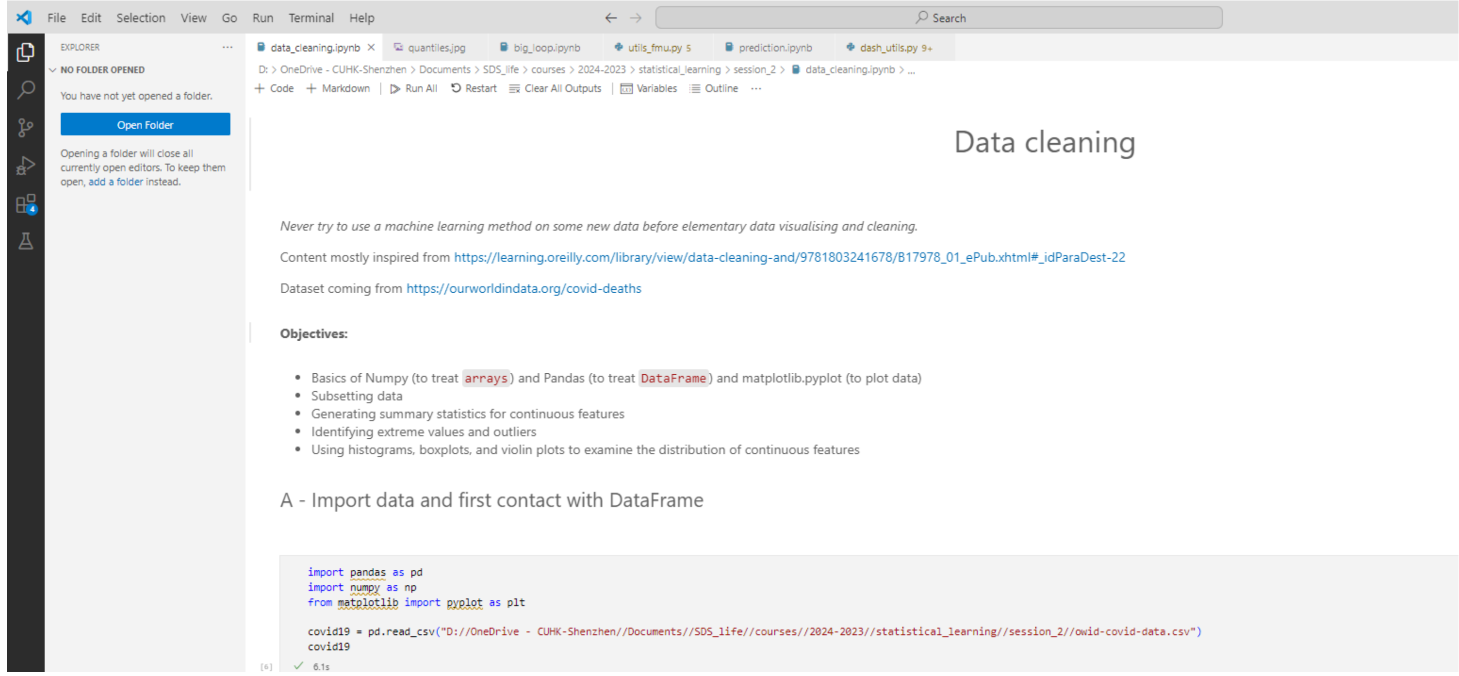
install packages with:

pip install -i https://pypi.tuna.tsinghua.edu.cn/simple some-package



Packages required:
`ipykernel, numpy, pandas, mathplotlib`

II - Download Visual studio code from:
`https://code.visualstudio.com/download`



Install extensions: `python, jupyter`

→ Strongly encouraged to use jupyter (ipynb). (can provide explainations, easier to interact with code and find typos/mistakes)